When Are Local Incentive Constraints Sufficient?

Gabriel Carroll, Department of Economics, MIT

`gdc@mit.edu`

June 24, 2011

Abstract

We study the question of whether local incentive constraints are sufficient to imply full incentive-compatibility, in a variety of mechanism design settings, allowing for probabilistic mechanisms. We give a unified approach that covers both continuous and discrete type spaces. On many common preference domains — including any convex domain of cardinal or ordinal preferences, single-peaked ordinal preferences, and successive single-crossing ordinal preferences — local incentive-compatibility (suitably defined) implies full incentive-compatibility. On domains of cardinal preferences that satisfy a strong nonconvexity condition, local incentive-compatibility is not sufficient. Our sufficiency results hold for dominant-strategy and Bayesian Nash solution concepts and allow for some interdependence in preferences.

Keywords: incentive-compatibility, local incentive constraints, mechanism design, sufficiency

JEL Classifications: C79, D02, D71

1

1   Introduction

In the analysis of mechanism design problems, taking account of all the possible constraints imposed by incentive-compatibility at once can be unwieldy. It can be easier to focus attention on local incentive constraints, ensuring that agents have no incentive to make "small" misreports of their type, and then check at the end of the analysis whether or not the mechanisms obtained are fully incentive-compatible.

In the present paper, we ask the general question of whether local incentive constraints are sufficient on their own to guarantee full incentive-compatibility, and obtain an affirmative answer in a wide array of settings. We allow for arbitrary probabilistic mechanisms, which specify a distribution over some (exogenously specified) outcome space as a function of an agent's type. (Our analysis is mostly worded in terms of a single agent, but we show how it readily extends to multi-agent mechanisms, including allowing a limited degree of interdependence in preferences.)

To clarify the significance of these results, it is useful to distinguish two major branches of mechanism design literature. We give a simple and unified approach that applies to both branches but has slightly different implications for the two. One branch, with roots in axiomatic social choice theory, studies problems without monetary transfers. These include voting (e.g. Gibbard (1973), Satterthwaite (1975), Moulin (1980), Barberà, Sonnenschein, and Zhou (1991), Saporiti (2009)), matching (Roth (1982), Alcalde and Barberà (1994), Bogomolnaia and Moulin (2001)), queueing (Bogomolnaia and Moulin (2002)), and rationing (Sprumont (1991), Ehlers and Klaus (2003)), among others. This literature has recently been influential in applied market design as well; see for example Roth (2008) and references therein. It is commonly taken as given here that each agent submits a ranking over outcomes (such as candidates, in a voting context, or schools or jobs, in a matching context) to the mechanism. Thus, agents report ordinal preferences. Incentive-compatibility typically means that reporting one's true preferences should be a dominant strategy. We will say that such a mechanism is locally incentive-compatible if no agent type can benefit from misreporting by switching some two consecutive outcomes in his preference ranking. We show below that for many of the most common preference domains considered in this literature, local incentive-compatibility implies full incentive-compatibility. Specifically, we show this for domains of ordinal preferences having convex closure (Proposition 2, which actually gives a generalization to *poly-*

hedral type spaces); single-peaked ordinal preferences (Proposition 3); and successive single-crossing ordinal preferences (Proposition 4).

The second large branch of literature concerns settings in which monetary transfers are possible and agents have quasilinear preferences, with applications such as monopoly pricing, auctions, and public projects. Seminal works include Mussa and Rosen (1978), Green and Laffont (1979), Myerson (1981), Myerson and Satterthwaite (1983), Maskin and Riley (1984), Jehiel and Moldovanu (2001), and Jehiel, Meyer-ter-Vehn, Moldovanu, and Zame (2006). It is generally assumed that agents can report a cardinal valuation for each outcome. We show that local incentive-compatibility, suitably formulated, implies full incentive-compatibility whenever the space of agents' cardinal types is convex (Proposition 5). (We should note that this result also has some relevance to the no-transfers literature, as some authors there also allow agents to report cardinal preferences, e.g. Barberà, Bogomolnaia, and van der Stel (1998), Hylland (1980), Zhou (1990).)

Our results on the sufficiency of local incentive constraints are relevant for several reasons. One is that they provide a technical tool to facilitate the researcher's task of analyzing mechanism design problems. This is particularly relevant to the transfers branch of the literature, where analysis typically begins by using local incentive constraints and the envelope theorem to obtain an integral formula for the utility attained by each type of agent (as in e.g. Myerson (1981, Lemma 2)); our sufficiency results provide a general tool to help assure researchers that this reduction of the problem has not neglected important nonlocal constraints. They also can streamline proofs of incentive-compatibility for newly designed mechanisms, since it is enough to show local incentive-compatibility and then invoke sufficiency.

Moreover, our analysis casts light on the *form* of local incentive constraints needed. It is *not* sufficient to specify only that each type of agent should be unable to profitably misreport as any nearby type; one must also specify that each type cannot serve as a profitable misreport for any nearby type. (See the discussion in Subsection 3.1.)

A separate reason our results are relevant is that one may have more literal reasons to impose only a subset of incentive constraints. For example, there may be a monitoring technology that makes it possible to detect and punish reports far away from an agent's true type, in which case the mechanism designer does not need to worry about such misreports (as in Green and Laffont (1986)). One might hope that this would provide an operational way to circumvent impossibility results such as the

3

Gibbard (1973)-Satterthwaite (1975) theorem, which are pervasive in the no-transfers literature; or, in a setting with transfers, one might hope, say, to obtain higher revenue than would be possible with fully incentive-compatible mechanisms. If agents are to report truthfully, then our sufficiency results show that in many settings, having access to such a monitoring technology does not enlarge the space of effective mechanisms.

Relatedly, when designing a mechanism for boundedly rational agents, one might consider that agents are not capable of contemplating every possible misreport of their preferences, and again ask whether this provides an operational way to improve on fully incentive-compatible mechanisms. If the designer believes that agents are at least rational enough to be capable of imitating any nearby type, then in the settings covered by our sufficiency results, imposing only the relevant subset of incentive constraints actually does not enlarge the space of usable mechanisms at all.

In particular, for ordinal type spaces, this idea leads to "computational" versions of many existing impossibility or characterization results. This gives a very general reply to a literature that seeks ways around the Gibbard-Satterthwaite theorem by devising voting mechanisms that are computationally difficult, but not impossible, to strategically manipulate (e.g. Bartholdi, Tovey, and Trick (1989), Bartholdi and Orlin (1991)). On the type spaces where our sufficiency results apply, they immediately imply that any such mechanism is easy to manipulate in some instances, as long as the outcome of the mechanism itself is easy to compute. (Here, as in the preceding literature, we take "easy" to mean computable in polynomial time.) Namely, a manipulator can exhaustively consider each local manipulation — switching some two candidates who are consecutive in the ranking — and compute the outcome of the mechanism; this trial-and-error search is easy and will find an advantageous manipulation in some instances. So a computational-complexity constraint, at least of the naive form, cannot prevent agents from manipulating.[1]

More broadly, there is a tradition in social choice theory of looking for the weakest assumptions necessary to obtain a characterization or impossibility result. Our results can be immediately applied to many axiomatic characterizations (such as those cited in the third paragraph), showing that, say, an axiom requiring dominant-

---

[1]In the Gibbard-Satterthwaite context, stronger results extending this idea are already known (e.g. Isaksson, Kindler, and Mossel (2010)). But our results lead more generally to computational versions of many other existing characterization results by the same argument.

strategy incentive-compatibility can be replaced by local incentive-compatibility without changing the conclusion.

The aforementioned results show that, for many important type spaces, local incentive-compatibility implies full incentive-compatibility. On the other hand, there are type spaces where the implication does not hold. In particular, we show this for domains of cardinal preferences that fail to be convex in a sufficiently strong way (Proposition 6).

Our work connects with several previous papers on mechanism design under a subset of incentive constraints. Green and Laffont (1986), mentioned above, consider a general setup in which the space of messages that agents can send equals the space of types, with exogenous restrictions as to which messages each type is capable of sending, and study when the revelation principle applies. Celik (2006) and Sher and Vohra (2010) consider specific mechanism design problems under subsets of incentive constraints, though their subsets are not local in our sense.

There does not appear to be previously published work asking the broad question of when local and full incentive-compatibility coincide. However, a contemporaneous paper by Sato (2010), independent of ours, does address this question. Sato considers only deterministic mechanisms over ordinal type spaces. For such mechanisms, Sato shows that local incentive constraints are sufficient on all of the ordinal type spaces that we consider (type spaces with convex closure, single-peaked, and successive single-crossing preferences), as well as some others.

This paper also bears some formal resemblance to recent work on general settings with cardinal preferences and transfers. In such a setting, a rule mapping types to outcomes is *implementable* if there exists some accompanying payment function (mapping types to transfers) that makes truthful revelation incentive-compatible. There has recently been much interest in simple conditions ensuring that a rule is implementable, e.g. Saks and Yu (2005), Bikhchandani, Chatterji, Lavi, Mu'alem, Nisan, and Sen (2006), Ashlagi, Braverman, Hassidim, and Monderer (2010). In particular, our work is somewhat reminiscent of a paper by Archer and Kleinberg (2008). They show that local implementability (suitably defined) implies implementability, on any convex space of cardinal types. However, we show that local implies full incentive-compatibility for a given mechanism, consisting of an outcome rule and a payment function together, whereas they show that local implementability by *some* payment functions (possibly using different payment functions on different local neigh-

borhoods) implies full implementability. Thus, both their hypothesis and their conclusion are weaker than ours. Moreover, their sets of local incentive constraints are larger than ours, and their theorem would not hold using our constraint sets; this is discussed in detail in Subsection 3.1. Accordingly, our results on cardinal type spaces do not follow from the result of Archer and Kleinberg, nor vice versa.

2   Framework

We begin with the general framework. Ensuing sections will give the concrete results.

2.1   Definitions

We will focus on incentives in a mechanism for an individual agent. In Subsection 2.2, we will show how the ideas extend straightforwardly to multi-agent mechanisms with private values. We begin by introducing the definitions for the no-transfers setting; in Subsection 3.5 we will allow for transfers, and also for interdependence.

From the agent's point of view, a mechanism takes the agent's preferences as input and determines an outcome, or a probability distribution over outcomes. We must be explicit about the form of preferences that the agent can announce. In some settings, it is standard practice to assume agents announce their cardinal valuation for each of the possible outcomes. In others (specifically in the no-transfers literature), it is assumed that agents only report an ordinal ranking of outcomes.

This latter assumption entails exogenously restricting the message space of the mechanism to consist of the possible ordinal preferences. This restriction is widely accepted, although it does not yet enjoy solid theoretical foundations. It is often made for analytical tractability, and in practical market design applications it can also be justified by the need to make the mechanism accessible to participants who may have difficulty thinking about their preferences over lotteries. Bogomolnaia and Moulin (2001) give a more detailed discussion on this last point.

Finally, in some settings, one might assume that agents report even coarser information than ordinal preferences (for example, they are required to rank only a limited number of outcomes). We will first give a unified treatment that covers all of the different specifications of preferences, then specialize to define local incentive-compatibility in specific settings.

Let X, the *outcome space*, be any finite set; m will denote its cardinality. Let $\Delta(X)$ denote the space of lotteries over X. The agent is assumed to have expected utility preferences over lotteries. It will be convenient to think of both lotteries over X and utility functions as elements of $\mathbb{R}^m$. If the agent's utility function is u, his payoff from a lottery L is given by the inner product $u \cdot L$.

For subsets of $\mathbb{R}^m$, cl denotes the closure and ∂ the boundary operator. If $u, v \in \mathbb{R}^m$, we write $[u, v]$ for the line segment $\{(1 - \alpha)u + \alpha v \mid \alpha \in [0, 1]\}$.

A *type* is a nonempty subset of $\mathbb{R}^m$. A *type space* is a set of pairwise disjoint types. We henceforth use the term *type space* in preference to *domain*: the latter term suggests only an exogenous restriction of the set of utility functions the agent may have, whereas our notion of a type space conveys both which utility functions are possible and which ones the mechanism is required to treat identically.

Given a type space T, a *mechanism* is a function $f : T \to \Delta(X)$. Thus, the mechanism chooses a distribution over outcomes, based on the agent's (reported) type.

An *incentive constraint* is an ordered pair of types. The interpretation of the constraint (t, t') is that a type t cannot benefit from misreporting as type t'. Accordingly, we say that a mechanism f *satisfies* an incentive constraint (t, t') if, for all $u \in t$, $u \cdot f(t) \geq u \cdot f(t')$; equivalently, $u \cdot (f(t) - f(t')) \geq 0$. A mechanism *satisfies a set* of incentive constraints if it satisfies every constraint in the set.

A mechanism that satisfies the full set of incentive constraints $T \times T$ is *fully incentive-compatible.* This is exactly the usual meaning of incentive-compatibility.

A set S of incentive constraints is *sufficient* if every mechanism that satisfies S is fully incentive-compatible.

We highlight several important kinds of type spaces and define local incentive constraints in each case.

- A type space T is *cardinal* if every type is a singleton. In this case, abusing notation, we will think of types as vectors and T as a subset of $\mathbb{R}^m$. For example, we write $f(u)$ rather than $f(\{u\})$.

  For a cardinal T, a set S of incentive constraints will be called *local incentive constraints* if every $u \in T$ has an open neighborhood N_u in T (with the relative topology) such that $(u, u') \in S$ and $(u', u) \in S$ for every $u' \in N_u$.

- A type space is *ordinal* if every type is of the form $t = \{u \mid u(x_1) > u(x_2) >$

$\cdots > u(x_m)\}$ for some strict ordering $x_1 \succ \cdots \succ x_m$ of the elements of X. We say that t *represents* this ordering. Note that our definition of an ordinal type space does not require that all possible orderings be represented by types.

When types are ordinal, f satisfies a constraint (t, t') if and only if the lottery $f(t)$ first-order stochastically dominates $f(t')$ with respect to the ordering on X represented by t. (This is easy to show.)

Call two ordinal types t, t' *adjacent* if the orderings they represent differ only by a switch of two consecutive outcomes. On any ordinal type space T, the *local incentive constraints* will refer to the set of all constraints (t, t') such that t and t' are adjacent.

- More generally, we can consider *polyhedral* type spaces. In the space of utility functions, $\mathbb{R}^m$, an *open half-space* is a set of the form $\{u \mid u \cdot \lambda > c\}$ for some nonzero $\lambda \in \mathbb{R}^m$ and some constant c. If H is such an open half-space, its closure $\mathrm{cl}(H) = \{u \mid u \cdot \lambda \geq c\}$ is a *closed half-space*, and its boundary $\partial H = \{u \mid u \cdot \lambda = c\}$ is a *hyperplane*. Define an (open) *polyhedron* to be a nonempty set that is the intersection of finitely many open half-spaces. A *polyhedral type space* is a type space consisting of finitely many types that are all polyhedra.

  Say that two disjoint polyhedra t, t' are *adjacent* if $\mathrm{cl}(l) \cap \mathrm{cl}(t')$ contains a nonempty, relatively open subset of a hyperplane. In simpler terms, t and t' are polyhedra that border along a face. We then let the *local incentive constraints* on T be the set of constraints (t, t') such that t and t' are adjacent.

  Any ordinal type space is polyhedral, and one can check that the definitions of adjacency and local incentive constraints for ordinal types agree with those for polyhedral types. (There exists previous literature in mechanism design also using polyhedra to represent ordinal types, e.g. Duggan (1996).) For another example, take the types implied by truncated rankings, i.e. $\{u \mid u(x_1) > \cdots > u(x_p)$ and $u(x_p) > u(y)$ for all $y \neq x_1, \ldots, x_p\}$, for any distinct outcomes $x_1, \ldots, x_p$ with $p < m$ — these are again polyhedral types. Such a type space is natural for studying matching mechanisms in applications such as school choice, where students may be asked to rank, say, 12 favorite schools out of more than 500 available (Abdulkadiroğlu, Pathak, and Roth (2005)).

8

More generally, any mechanism with a *finite* message space gives rise to a polyhedral type space: for each message, the set of utility functions for which it is optimal forms a polyhedron (ignoring boundary issues). Studying local incentives in these type spaces can be helpful for analyzing such mechanisms. Gibbard (1978) gives a fairly complete analysis of dominant-strategy voting mechanisms with arbitrary finite message spaces; much of the analysis focuses on incentives to misreport locally.

We say that a mechanism is *locally incentive-compatible* if it satisfies some set of local incentive constraints (in the cardinal case; or the canonical such set in the polyhedral case).

We are interested in determining whether or not local incentive contraints are sufficient, on various type spaces.

2.2 Mechanisms with multiple agents

As already mentioned, while we focus on single-agent mechanisms, our results apply also with multiple agents, under private values. The extension is similar to arguments in previous literature (Archer and Kleinberg (2008), Heydenreich, Müller, Uetz, and Vohra (2008)), but we spell it out in detail here, as we will further build on it in Subsection 3.5.

Define a mechanism with n agents, type space $T = T_1 \times \cdots \times T_n$, and outcome space X to be a map $f : T \to \Delta(X)$, specifying a (probabilistic) outcome as a function of all the agents' types. Suppose that, for some i, a set S_i of incentive constraints is sufficient for T_i.

One possible notion of incentive-compatibility is to say that f satisfies the incentive constraint $(t_i, t_i') \in T_i \times T_i$ for agent i if, for all t_{-i} and all $u_i \in t_i$, we have $u_i \cdot (f(t_i, t_{-i}) - f(t_i', t_{-i})) \geq 0$. If f satisfies every incentive constraint in S_i for agent i, then holding fixed any profile t_{-i}, the single-agent mechanism $t_i \mapsto f(t_i, t_{-i})$ satisfies S_i and so (by sufficiency) is fully incentive-compatible. Thus, f is fully incentive-compatible in dominant strategies (for agent i).

One can also consider Bayesian incentive-compatibility. Suppose we are given a probability distribution ψ_j over T_j for each agent j, and assume $f(t_i, t_{-i})$ is measurable in t_{-i} for all t_i. Then we can say that f satisfies incentive constraint (t_i, t_i') for agent i if, for all $u_i \in t_i$, we have $u_i \cdot (E_i[f(t_i, t_{-i})] - E_i[f(t_i', t_{-i})]) \geq 0$, where the expectation

is over t_{-i} with respect to the product distribution $\times_{j \neq i} \psi_j$. Again, if f satisfies each incentive constraint in S_i, then the single-agent mechanism $t_i \mapsto E_i[f(t_i, t_{-i})]$ satisfies S_i and so is fully incentive-compatible for agent i. This is the standard notion of Bayesian incentive-compatibility for f. Notice that this argument depends on the agents' types being independently distributed: the expectation E_i needs to be defined in a way that does not depend on t_i.

3   Sufficiency

In this section we show that local incentive constraints are sufficient on a variety of common type spaces. All proofs absent from the main text are in Appendix A.

3.1   Cardinal type spaces

Recalling that a cardinal type space is identified with a subset of $\mathbb{R}^m$, we can state our first sufficiency result:

Proposition 1 *On a convex cardinal type space T, any set of local incentive constraints is sufficient.*

We present the proof in detail, since the proofs of most of our other sufficiency results (Propositions 2, 3, 5) follow the same model. To prove that an agent of type u never wants to misreport as type v, we restrict attention to types along the line segment $[u, v]$, effectively reducing to one dimension; we then break the segment into short pieces for which local incentive constraints apply, and combine these local incentive constraints into the incentive constraint (u, v) using the kind of supermodularity or "revealed-preference" argument that is familiar elsewhere in the mechanism design literature (see e.g. Myerson (1981, Lemma 2), Rochet (1987, Theorem 1)).

Proof: Let S be a set of local incentive constraints and f a mechanism satisfying S. For types u, v, write $u \leftrightarrow v$ if (u, v) and (v, u) are both in S. By definition, every $u \in T$ has some neighborhood N_u in T such that $u \leftrightarrow v$ for all $v \in N_u$.

Fix arbitrary $u, v \in T$. We want to show that $u \cdot (f(u) - f(v)) \geq 0$.

For any $\alpha \in [0, 1]$, define $u_\alpha = (1 - \alpha)u + \alpha v$. Convexity implies $u_\alpha \in T$. Let

$$A = \{\alpha \mid \text{there exist } 0 = \alpha_0 < \alpha_1 < \cdots < \alpha_r \leq 1 \text{ with}$$

$$u_{\alpha_0} \leftrightarrow u_{\alpha_1} \leftrightarrow \cdots \leftrightarrow u_{\alpha_r} \text{ and } \alpha_r = \alpha\}.$$

Clearly, if $\alpha \in A, \alpha < \alpha' \le 1$, and $u_\alpha \leftrightarrow u_{\alpha'}$, then $\alpha' \in A$. Now let $\overline{\alpha} = \sup A \ge 0$. If $\overline{\alpha} = 0$ then $\overline{\alpha} \in A$. If $\overline{\alpha} > 0$, then for α sufficiently close to $\overline{\alpha}$ we have $u_\alpha \leftrightarrow u_{\overline{\alpha}}$; since we can choose $\alpha \in A$ arbitrarily close, we again get $\overline{\alpha} \in A$. Moreover, if $\overline{\alpha} < 1$, then $u_{\overline{\alpha}} \leftrightarrow u_\alpha$ for α just slightly larger than $\overline{\alpha}$; this implies $\alpha \in A$, contradicting $\overline{\alpha} = \sup A$. Therefore, we get $\overline{\alpha} = 1$ and $1 \in A$.

So we have $0 = \alpha_0 < \alpha_1 < \cdots < \alpha_r = 1$ with $u_{\alpha_k} \leftrightarrow u_{\alpha_{k+1}}$ for each k. Now write out the local incentive constraints:

$$
\begin{aligned}
u_{\alpha_k} \cdot (f(u_{\alpha_k}) - f(u_{\alpha_{k+1}})) &\ge 0, \\
u_{\alpha_{k+1}} \cdot (f(u_{\alpha_{k+1}}) - f(u_{\alpha_k})) &\ge 0.
\end{aligned}
$$

Multiplying by α_{k+1} and α_k, respectively, and adding gives

$$
[\alpha_{k+1} u_{\alpha_k} - \alpha_k u_{\alpha_{k+1}}] \cdot (f(u_{\alpha_k}) - f(u_{\alpha_{k+1}})) \ge 0.
$$

But one directly calculates that $\alpha_{k+1} u_{\alpha_k} - \alpha_k u_{\alpha_{k+1}} = (\alpha_{k+1} - \alpha_k)u$. Since $\alpha_{k+1} - \alpha_k > 0$, we can divide through to obtain

$$
u \cdot (f(u_{\alpha_k}) - f(u_{\alpha_{k+1}})) \ge 0.
$$

Now we can sum over $k = 0, 1, \ldots, r - 1$, and telescoping gives

$$
u \cdot (f(u) - f(v)) = u \cdot (f(u_{\alpha_0}) - f(u_{\alpha_r})) \ge 0.
$$

$\square$

Proposition 1 applies to any convex cardinal type space. This includes, for example, the full space of utility functions on X; or the space of utility functions that are increasing with respect to some partial order on X; or the space of supermodular or submodular utility functions, given a lattice structure on X; or the space of utility functions satisfying some concavity conditions.

The proof of Proposition 1 clearly uses both parts of the definition of local incentive constraints — that each u should have a neighborhood N_u with both $(u, u') \in S$ and $(u', u) \in S$ for $u' \in N_u$. A seemingly more natural way to define local incentive constraints would only require $(u, u') \in S$. Under this definition, Proposition 1 would no longer hold. For example, suppose $X = \{x, y\}$ and T is the full space of all cardinal

11

types. Consider the mechanism f given by $f(u) = x$ if $u(x) < u(y)$ and $f(u) = y$ otherwise. This f meets the weaker definition of local incentive-compatibility, but is not fully incentive-compatible. (Requiring only $(u', u) \in S$ would also not be enough: with the same X and T, consider the mechanism $f(u) = x$ if $u(x) = u(y) - 1$ and $f(u) = y$ otherwise.)[2]

By contrast, the local-to-global result of Archer and Kleinberg (2008, Corollary 3.7), on implementability in a quasilinear setting, effectively requires stronger local incentive constraints. They assume implementability throughout each N_u — that is, for each u, there should be some payment function p_u (specifying a payment for each agent type) so that the mechanism-with-transfers (f, p_u) satisfies incentive constraints (u', u'') for all $u', u'' \in N_u$. The analogue of our constraints in their setting would be to merely require that (f, p_u) should satisfy constraints (u, u') and (u', u) for all $u' \in N_u$. This requirement is a local form of *weak monotonicity*, which is not enough to imply their implementability conclusion without further restrictions; see e.g. Bikhchandani, Chatterji, Lavi, Mu'alem, Nisan, and Sen (2006, Example S1) or Saks and Yu (2005, Section 7).

Unlike the local constraints of Archer and Kleinberg (2008), ours can be expressed succinctly in terms of local maxima: f is locally incentive-compatible if every $u \in T$ is a local maximum of both the functions $v \mapsto u \cdot f(v)$ and $v \mapsto v \cdot (f(u) - f(v))$. With this interpretation, local incentive-compatibility can potentially be checked by first- and second-order conditions at points where f is differentiable. This convenience is relevant in making the reduction from global to local incentive-compatibility a useful one: if one wishes to check incentive constraints directly, then even local incentive-compatibility can require checking many constraints when T is high-dimensional, since it is necessary to check constraints in every direction at each u.

3.2 Polyhedral type spaces

Next we consider polyhedral type spaces. Our main result here is:

---

[2]A referee points out that a mechanism on a cardinal type space is fully incentive-compatible if and only if the indirect utility function $u \mapsto u \cdot f(u)$ is convex, with $f(u)$ belonging to the subdifferential at each point u. Our two local conditions can be viewed loosely as local forms of these requirements: the subdifferential condition at u is equivalent to satisfying (u', u) for all $u' \in T$, so our requiring this for all $u' \in N_u$ gives a local form of the subdifferential condition, and then imposing the additional constraints (u, u') ensures convexity.

Proposition 2 *Let T be a polyhedral type space such that $\cup_{t \in T} \operatorname{cl}(t)$ is convex. Then the set of local incentive constraints is sufficient.*

The argument is essentially the same as for Proposition 1. For utility functions u and v, we consider the line segment $[u, v]$; this segment passes through various types in succession. By jiggling v a bit if necessary, we can ensure that any two successive types along this line segment are adjacent polyhedra, and then we can just add up the corresponding local incentive constraints as before.

A particular case of Proposition 2 is that on the full space of all ordinal types over a given X, the local incentive constraints are sufficient.[3] (The union of the closures of all types is simply all of $\mathbb{R}^m$.) Proposition 2 also applies when T consists of all ordinal types that respect a given partial ordering on X. For example, Bogomolnaia and Moulin (2002) consider an allocation problem with real objects and a null object; all types have the same preference ordering on the real objects, but rank the null object differently relative to the real objects.

3.3 Single-peaked preferences

The preceding results have focused on essentially convex type spaces. One important nonconvex type space is that of single-peaked preferences.

Fix an ordering $x_1, \ldots, x_m$ of the outcomes in X. A strict preference ordering $\succ$ over X is *single-peaked* if there exists some outcome x_{p^*} such that, whenever $q < p \leq p^*$ or $q > p \geq p^*$, we have $x_p \succ x_q$. An ordinal type is *single-peaked* if it represents a single-peaked ordering.

Single-peaked preferences have been popular in voting theory ever since Black's (1948) observation that the rule choosing the median of the voters' favorite outcomes is dominant-strategy incentive-compatible. Single-peaked preferences are also important in economic applications because single-peakedness is the same as quasiconcavity of the utility function (aside from issues of indifference). Moulin (1980) characterizes dominant-strategy incentive-compatible deterministic voting systems under single-peaked preferences. (Moulin assumes the outcome space is the whole real line, but his proofs carry through almost unchanged for a finite outcome space.) Ehlers, Peters, and Storcken (2002) extend this work to probabilistic mechanisms. Sprumont (1991),

---

[3]An analogous result also holds if we allow indifferences — so that for each weak order on X, the set of utility functions representing it constitutes a type — with an appropriate definition of local incentive constraints. We omit the details here.

Barberà, Jackson, and Neme (1997), and Ehlers and Klaus (2003) study rationing problems when consumers have single-peaked preferences over quantities.

The space of single-peaked ordinal types does not meet the convexity condition of Proposition 2. However, we still have the result:

Proposition 3 *Fix an ordering $x_1, \ldots, x_m$ of the elements of X. On the space of single-peaked ordinal types, the set of local incentive constraints is sufficient.*

The argument is a slight extension of that used for Proposition 2. In general, in an ordinal type space, say that a utility function v is *accessible* from another utility function u if the segment $[u, v]$ is contained in the union of the closures of all types. In this case we can apply the argument of adding up local incentive constraints from Propositions 1 and 2. Now, in the single-peaked ordinal type space, it is no longer true (as it was for Proposition 2) that all v in a given type t' are accessible from $u \in t$, but we actually only need to be able to find some such v for each u. Lemma 5 in the appendix shows that this can be done.

One can also consider the space of single-dipped ordinal types (Klaus, Peters, and Storcken (1997)), or of single-peaked ordinal preferences on a tree (Demange (1982), Danilov (1994)). It is straightforward to extend the proof to cover each of these cases, showing that the local incentive constraints are again sufficient.

3.4   Single-crossing preferences

Besides single-peaked preferences, another economically important class of ordinal type spaces is given by single-crossing preferences. These are defined as follows: Fix an ordering $x_1, \ldots, x_m$ of the elements of X. A sequence $\succ_1, \ldots, \succ_r$ of distinct strict preference orderings is a *single-crossing preference domain* if the following holds: whenever $p < q$ and $x_q \succ_k x_p$ for some k, we also have $x_q \succ_l x_p$ for all $l > k$. Single-crossing ordinal preferences arise in economic models such as the redistributive taxation models of Roberts (1977) and Meltzer and Richard (1981) (see Saporiti (2009) for references to other applications). Just as with single-peaked preferences, preferences coming from any single-crossing domain satisfy a median voter property — the voting scheme that chooses the outcome most preferred by the voter with the median preference is dominant-strategy incentive-compatible. More generally, Saporiti (2009) characterizes dominant-strategy incentive-compatible voting schemes on any maximal single-crossing preference domain.

For any strict preference ordering $\succ$ on X, let $V(\succ) = \{(p, q) \mid p < q, x_p \prec x_q\}$. By definition, a sequence of preference orderings $\succ_1, \ldots, \succ_r$ is a single-crossing preference domain if and only if $V(\succ_1) \subseteq \cdots \subseteq V(\succ_r)$. In fact, these inclusions must all be strict, since any ordering $\succ$ can be uniquely reconstructed from $V(\succ)$. Therefore $|V(\succ_1)| < \cdots < |V(\succ_r)|$. Call the domain a *successive single-crossing preference domain* if $|V(\succ_{k+1})| = |V(\succ_k)| + 1$ for each $k = 1, \ldots, r - 1$.

This covers the domains considered in Saporiti (2009) — any maximal single-crossing preference domain $\succ_1, \ldots, \succ_r$ is successive. For suppose that $|V(\succ_{k+1})| - |V(\succ_k)| > 1$ for some k. There must be some two alternatives x_p, x_q that are ranked consecutively by $\succ_k$, with $x_p \succ_k x_q$ but $x_q \succ_{k+1} x_p$. Single-crossing ensures $p < q$. By switching the positions of x_p and x_q in $\succ_k$, we get a new ordering $\succ'$ with $V(\succ') = V(\succ_k) \cup \{(p, q)\}$, and hence $V(\succ_k) \subset V(\succ') \subset V(\succ_{k+1})$. This means that $\succ_1, \ldots, \succ_k, \succ', \succ_{k+1}, \ldots, \succ_r$ is again a single-crossing preference domain, contradicting maximality.

For any successive single-crossing preference domain $\succ_1, \ldots, \succ_r$, call the corresponding space of ordinal types $T = \{t_1, \ldots, t_r\}$ a *successive single-crossing ordinal type space*. In this case, the local incentive constraints are precisely those of the form (t_k, t_{k+1}) or (t_{k+1}, t_k). We shall show that on such a type space, the local incentive constraints are sufficient. This result may be surprising, since these incentive constraints are especially parsimonious — each type is adjacent to just two other types.

Proposition 4 *On any successive single-crossing ordinal type space, the local incentive constraints are sufficient.*

The strategy of proof is a little different from that used for the previous propositions. Instead of breaking a single line segment into short pieces, we find a sequence of parallel line segments, each connecting two consecutive types t_k, t_{k+1}, but such that each segment need not begin where the previous one ended. (As pointed out by a referee, this method has some precedent in Gibbard (1977, Lemma 2), where a similar argument is applied to the full ordinal type space; and the argument can be applied to the space of single-peaked ordinal types as well.)

Proof: Suppose the mechanism f satisfies the local incentive constraints. Fix any two types $t_l, t_{l'}$, and let $u \in t_l$. We wish to show that $u \cdot (f(t_l) - f(t_{l'})) \geq 0$. We will show this for $l' > l$; the proof for $l' < l$ is similar.

15

In fact it suffices to show that

$$u \cdot (f(t_k) - f(t_{k+1})) \geq 0 \qquad \text{for } k \geq l, \tag{1}$$

since then we can sum up (1) for $k = l, l+1, \ldots, l'-1$ to obtain $u \cdot (f(t_l) - f(t_{l'})) \geq 0$.

So fix $k \geq l$, and also define $M = \max_x u(x) - \min_x u(x)$. Write $V(\succ_{k+1}) \setminus V(\succ_k) = \{(p, q)\}$ by successiveness; then $p < q$, and $\succ_k$ ranks x_p just above x_q. Because $u \in t_l$ with $l \leq k$, single-crossing implies that $u(x_p) > u(x_q)$ also. Let v be any utility function representing $\succ_k$ such that $v(x_p) - v(x_q) < u(x_p) - u(x_q)$, and $|v(x) - v(y)| > M$ for all distinct outcomes $x, y \in X$ other than x_p and x_q. Because $\succ_k$ ranks x_p and x_q consecutively, we can do this. Then the utility function $v - u$ ranks every pair of outcomes in the same way as v does, except $\{x_p, x_q\}$. Since $V(\succ_{k+1}) = V(\succ_k) \cup \{(p, q)\}$, this means that $v - u$ represents $\succ_{k+1}$.

So, $v \in t_k$ and $v - u \in t_{k+1}$. The local incentive constraints give

$$\begin{aligned} v \cdot (f(t_k) - f(t_{k+1})) &\geq 0, \\ [v - u] \cdot (f(t_{k+1}) - f(t_k)) &\geq 0. \end{aligned}$$

Adding these two gives exactly (1), and this completes the proof. $\qquad\square$

The hypothesis of successiveness in Proposition 4 cannot be dropped, even if the set of local incentive constraints is modified in the natural way. That is, it is not the case that, for any single-crossing ordinal type space $\{t_1, \ldots, t_r\}$, the set consisting of the incentive constraints (t_k, t_{k+1}) and (t_{k+1}, t_k), for $1 \leq k < r$, is sufficient. For a counterexample, consider the three orderings

$$\begin{aligned} \succ_1: \quad & x_1 \ \succ_1 \ x_2 \ \succ_1 \ x_3 \ \succ_1 \ x_4 \\ \succ_2: \quad & x_2 \ \succ_2 \ x_1 \ \succ_2 \ x_3 \ \succ_2 \ x_4 \\ \succ_3: \quad & x_4 \ \succ_3 \ x_2 \ \succ_3 \ x_1 \ \succ_3 \ x_3 \end{aligned}$$

and the corresponding ordinal types t_1, t_2, t_3. Let f map the types to lotteries over (x_1, x_2, x_3, x_4) as follows:

$$f(t_1) = (1/4, 1/4, 1/2, 0); \qquad f(t_2) = (0, 1/2, 1/2, 0); \qquad f(t_3) = (1/2, 0, 0, 1/2).$$

Then f satisfies the incentive constraints $(t_1, t_2), (t_2, t_1), (t_2, t_3), (t_3, t_2)$, but not (t_1, t_3),

so it is not fully incentive-compatible. (The line of the proof of Proposition 4 that fails is the statement $V(\succ_{k+1}) \setminus V(\succ_k) = \{(p,q)\}$ in the third paragraph. More broadly, the approach of the proof fails because if we take, say, the utility function u representing $\succ_1$ with $u(x_1) = 4$, $u(x_2) = 3$, $u(x_3) = 2$, $u(x_4) = 1$, then we cannot find any v such that v represents $\succ_2$ and $v - u$ represents $\succ_3$.)

3.5 Transfers and interdependent preferences

We now return to the setting of cardinal preferences. However, we generalize in two new directions. First, we consider the transfers setting, in which agents have quasi-linear utility in outcomes and money, and a mechanism specifies both a lottery over outcomes and a transfer for each agent. Second, we allow for the possibility of inter-dependent preferences, where each agent's utility for each outcome depends on the other agents' types. Numerous recent works prove possibility and impossibility results with transfers and interdependence (Jehiel and Moldovanu (2001), Jehiel, Meyer-ter-Vehn, and Moldovanu (2010), Jehiel, Meyer-ter-Vehn, Moldovanu, and Zame (2006), Bikhchandani (2006)), and it is natural to ask to what extent our methods apply here.

We adopt new notations and terminology, for this subsection only, in order to describe these extensions. For clarity, it will help to explicitly write out the dependence of the mechanism on all n agents' types, as in Subsection 2.2. (Of course, a single agent is a special case.) Each agent i's type space T_i is now assumed to be a subset of an arbitrary finite-dimensional Euclidean space, not necessarily $\mathbb{R}^m$. Write $T = T_1 \times \cdots \times T_n$. Agent i's utility is now represented by a function $u_i : T \to \mathbb{R}^m$ specifying his utility for each outcome as a function of the entire type profile. (The private-values case discussed previously is the special case where $T_i \subseteq \mathbb{R}^m$ and $u_i(t_1, \ldots, t_n) = t_i$.)

To allow for transfers, a *mechanism* is now a *pair* (f,p), where $f : T \to \Delta(X)$ specifies a lottery over outcomes for each type profile, and $p : T \to \mathbb{R}^n$ is a *payment function* specifying the net transfer each agent receives. We write $p_i(t)$ for the ith component of $p(t)$, representing agent i's transfer.

If the true type profile is t and the agents report profile t', then agent i's realized utility is $u_i(t) \cdot f(t') + p_i(t')$. An *incentive constraint* for agent i is again a pair $(t_i, t_i') \in T_i \times T_i$. We will emphasize here the Bayesian notion of incentive-compatibility, so assume a distribution ψ_j on each agent's type space T_j is given. The mechanism (f,p)

satisfies the incentive constraint (t_i, t'_i) if

$$E_i[u_i(t_i, t_{-i}) \cdot f(t_i, t_{-i}) + p_i(t_i, t_{-i})] \geq E_i[u_i(t_i, t_{-i}) \cdot f(t'_i, t_{-i}) + p_i(t'_i, t_{-i})].$$

Here the expectations are with respect to the product distribution $\times_{j \neq i} \psi_j$ on other agents' types; it is presupposed that the expressions inside the expectations are measurable in t_{-i}, and both expectations are finite. (As in Subsection 2.2, the assumption of independently distributed types is crucial.)

A set S_i of incentive constraints will again be called *local incentive constraints* if every $t_i \in T_i$ has an open neighborhood N_{t_i} in T_i such that $(t_i, t'_i) \in S_i$ and $(t'_i, t_i) \in S_i$ for all $t'_i \in N_{t_i}$. S_i is *sufficient* for agent i if every mechanism that satisfies it must satisfy the full set of incentive constraints $T_i \times T_i$.

Dominant-strategy incentive-compatibility has an analogue in the interdependent setting, namely *ex post* incentive-compatibility (Chung and Ely (2006), Jehiel, Meyer-ter-Vehn, Moldovanu, and Zame (2006)), which demands Bayesian incentive-compatibility for all probability distributions simultaneously. Our result (Proposition 5 below) is expressed in terms of Bayesian incentive-compatibility, but an immediate corollary is that the same result holds using ex post incentive-compatibility instead.

To obtain a sufficiency result, we need to restrict interdependence by assuming that, for each fixed t_{-i}, the utility function $u_i(\cdot, t_{-i}) : T_i \to \mathbb{R}^m$ is linear in t_i. Under this restriction, we have:

Proposition 5 *In the setting with transfers and interdependent utility linear in own type, if agent i has a convex type space T_i, then every set of local incentive constraints is sufficient for agent i.*

The proof is a straightforward extension of that for Proposition 1.

The linearity assumption warrants some comments. It is satisfied trivially in the private-values case (hence, Proposition 1 is a special case of Proposition 5). It is also satisfied by many concrete models appearing in the interdependent preferences literature, including Dasgupta and Maskin (2000, Examples 2, 3, 4, 5); Jehiel and Moldovanu (2001); Fieseler, Kittsteiner, and Moldovanu (2003) (under their assumption A2); Bikhchandani (2006, Example 1); and Bergemann and Morris (2009, Section 3). On the other hand, it is quite restrictive, relative to the space of all well-behaved utility functions $u_i : T \to \mathbb{R}^m$ an agent might have.

The linearity assumption is crucial in our analysis, ensuring that the convexity condition in Proposition 5 extends that of Proposition 1. To understand this, notice (as observed also in Archer and Kleinberg (2008)) that we can think of each type of agent i as specifying a utility function $X \times T_{-i} \to \mathbb{R}$, and given the priors ψ_j, a mechanism induces a distribution over $X \times T_{-i}$ for each t_i. In order to apply the argument from the proof of Proposition 1, we essentially need agent i's type space to be a convex subset of the linear space of all functions $X \times T_{-i} \to \mathbb{R}$. This is exactly the combination of linearity and convexity assumptions we have made above.

The preceding paragraph does not show that sufficiency fails when the linearity assumption is violated, only that the method of proof used here (adding up incentive constraints along a line) cannot be used. The question of how much further sufficiency can be generalized is taken up in more extensively in the online appendix, which suggests that sufficiency results do not exist much beyond what can be proven with the present method; as well as Section 4 below, which shows how sufficiency fails under a condition somewhat stronger than nonconvexity.

We have here extended Proposition 1 to allow for transfers and/or interdependence with utility linear in own type. The same extension can be applied to Propositions 2 and 3. Such results may potentially be useful, for example, in analyzing mechanism design problems in such settings when the message spaces are constrained to be finite.

4   Insufficiency

The previous section gave numerous classes of type spaces on which local incentive constraints are sufficient. The discussion is not complete without giving some cases where local incentive constraints are not sufficient. We restrict ourselves here to cardinal type spaces. Proposition 6 below identifies a large class of such type spaces — roughly, those which violate convexity in a strong enough way — for which we can construct mechanisms that are locally, but not fully, incentive-compatible. (Sato (2010, Proposition 4.2) gives an analogous result for ordinal type spaces.)

It is unclear just how far Proposition 6 can be sharpened. Proposition 1 showed that if the type space is convex, any local incentive constraints are sufficient, but the converse is not true. The question of exactly characterizing those type spaces T for which all local incentive constraints are sufficient appears to be subtle. This topic is explored further in the online appendix. Proposition OA-2 of that document

gives a nontrivial example of a nonconvex type space for which all local incentive constraints are sufficient; on the other hand, Proposition OA-1 gives a kind of converse to Proposition 1 for *finite* cardinal type spaces. The details are somewhat technical, so we refer the reader to the online appendix, and for now proceed to give our simpler result.

In the space $\mathbb{R}^m$, let Π be the subspace of vectors whose sum of components is zero. Let a *fair open half-space* be a set of the form $H = \{u \mid u \cdot \lambda > 0\}$ for some nonzero $\lambda \in \Pi$. Say that a cardinal type space T is *fairly separated* if there is some fair open half-space H such that the set $T \cap H$ is not connected.

Proposition 6 *Let T be a cardinal type space that is fairly separated. Then there exists a set of local incentive constraints that is not sufficient.*

Fair separatedness certainly implies nonconvexity. To further indicate the relationship between the two concepts, a little graphical intuition is in order.

For concreteness, suppose X has four elements. By additive and multiplicative renormalization, we can map every utility function either to a point on the unit sphere in the three-dimensional space Π, or to the origin. This sphere is illustrated in Figure 1. The upper hemisphere, whose boundary is shown dashed in the figure, corresponds to a fair open half-space. If T contains the types labeled u and v, but does not contain any type along the thick curve (or any type cardinally equivalent to it), then T is fairly separated.

If T were to consist of all possible utility functions except 0 and w (and anything cardinally equivalent to them), then T would be nonconvex. Nonetheless, on this T, any local incentive constraints are sufficient; this is just Proposition OA-2 of the online appendix. Excluding the whole curve from T, rather than just the one point w, is enough for insufficiency, by Proposition 6.

Fair separatedness might not seem like a natural condition on a type space, so we give one economically important example. Fix an ordering $x_1, \ldots, x_m$ of the outcomes, and let T be the cardinal type space consisting of all quasiconcave utility functions — a cardinal analogue of the single-peaked ordinal type space considered in Section 3.3. Then T is fairly separated. For example, take any $1 \le p < p' < p'' \le m$, and let $H = \{u \mid u(x_p) - 2u(x_{p'}) + u(x_{p''}) > 0\}$. If $u \in T \cap H$, then either $u(x_p) > u(x_{p'})$ or $u(x_{p''}) > u(x_{p'})$. So $\{u \in T \cap H \mid u(x_p) > u(x_{p'})\}$ and $\{u \in T \cap H \mid u(x_{p''}) > u(x_{p'})\}$ are two open, nonempty subsets of $T \cap H$, whose union

Figure 1: Illustration of fair separatedness

is all of $T \cap H$, and whose intersection is empty (any u satisfying both inequalities would violate quasiconcavity). Hence, $T \cap H$ is not connected. By Proposition 6, there are local incentive constraints on T that are insufficient. Note that this result for single-peaked cardinal types contrasts with our sufficiency result for the single-peaked ordinal type space (Proposition 3). The accessibility argument underlying that proposition (Lemma 5 in the appendix) fails with single-peaked cardinal types.

Proof of Proposition 6: Let $H = \{u \mid u \cdot \lambda > 0\}$ with $\lambda \in \Pi$. There exist lotteries L, L' on X such that $H = \{u \mid u \cdot L > u \cdot L'\}$. Indeed, let L be any lottery with full support, and just let $L' = L - \delta \lambda$, where $\delta > 0$ is chosen small enough so that all components of L' are still positive.

Now write $T \cap H = T_a \cup T_b$, where T_a, T_b are open, disjoint, and nonempty. Consider the mechanism f defined as follows: if $u \in T_a$, then $f(u) = L$; otherwise, $f(u) = L'$.

Let $S = (T \times T) \setminus (T_b \times T_a)$. This is a set of local incentive constraints: If $u \in T_a$, let $N_u = T_a$; if $u \in T_b$, let $N_u = T_b$; and if $u \in T \setminus H$, let $N_u = T$. In each case we have $(u, u'), (u', u) \in S$ for all $u' \in N_u$.

One readily checks that f satisfies the incentive constraints S, but does not satisfy any incentive constraint in $T_b \times T_a$ and so is not fully incentive-compatible. Thus, S is not sufficient. $\square$

A similar construction can be applied in the context of Subsection 3.5, to generate many examples with interdependent preferences, nonlinear in own type, for which

local incentive constraints are not sufficient.

5   Conclusion

This paper has examined the question of whether or not a small set of local incentive constraints is sufficient to ensure that all other incentive constraints are automatically satisfied, allowing for probabilistic mechanisms. We have obtained affirmative answers in many of the most common mechanism design settings. With convex spaces of cardinal types, local incentive constraints are sufficient to imply full incentive-compatibility. This result allows for monetary transfers under quasilinear utility, and for interdependence so long as each agent's utility is linear in his own type. Local incentive constraints are also sufficient on polyhedral (including ordinal) type spaces with convex closure, as well as single-peaked or successive single-crossing ordinal type spaces. Our proofs follow a unified analytical approach based on a simple supermodularity argument that applies across different settings. For cardinal type spaces that are not convex, the argument does not apply, and with a strengthening of nonconvexity we have insufficiency — there are mechanisms that are locally but not fully incentive-compatible.

The sufficiency results provide an immediate strengthening of many existing impossibility and characterization theorems, and a negative answer to a possible line of inquiry as to whether one could obtain new mechanisms by ignoring nonlocal incentive constraints on grounds of bounded rationality or monitoring technology. Most importantly, they facilitate the technical analysis of mechanism design problems in these settings by ensuring that one can focus on local incentive constraints without any loss, avoiding the need for separate verifications of full incentive-compatibility.

Our analysis on cardinal type spaces in particular also sheds some light on the form of local incentive constraints that should be considered in order to ensure full incentive-compatibility. A naive formulation is not sufficient. On the other hand, our local incentive constraints are still substantially weaker than requiring incentive-compatibility throughout a neighborhood of each type (as required for the formally similar result of Archer and Kleinberg (2008)), and arguably easier to verify in applications.

A Omitted proofs

Here we present the proofs that were omitted from the main text. We begin with some technical lemmas.

Lemma 1 *Let the polyhedron t be written as an intersection of open half-spaces, $t = \cap_{k=1}^r H_k$. Then the boundary of t is*

$$\partial t = \text{cl}(t) \setminus t = \bigcup_{\emptyset \neq K \subseteq \{1,\dots,r\}} \left[\left(\bigcap_{k \in K} \partial H_k \right) \cap \left(\bigcap_{k \notin K} H_k \right) \right].$$

Proof: Write $\text{cl}(t) = \cap_k \text{cl}(H_k) = \cap_k (\partial H_k \cup H_k)$; distribute on the right side; then remove $t = \cap_k H_k$ from both sides. $\quad\square$

In particular, $\partial t \subseteq \cup_k \partial H_k$.

Lemma 2 *Suppose T is a polyhedral type space. Fix $u \in \mathbb{R}^m$, and let $t, t' \in T$ be distinct, nonadjacent types. Then there exist finitely many hyperplanes, each one passing through u, whose union contains $\text{cl}(t) \cap \text{cl}(t')$.*

Proof: Since t, t' are open and disjoint, neither one can intersect the closure of the other, so $\text{cl}(t) \cap \text{cl}(t') = \partial t \cap \partial t'$. Now suppose $t = \cap_{k=1}^r H_k$ and $t' = \cap_{k'=1}^{r'} H'_{k'}$. Applying Lemma 1 to both t and t', and then distributing the intersection operator, we get

$$\partial t \cap \partial t' = \bigcup_{\substack{\emptyset \neq K \subseteq \{1,\dots,r\} \\ \emptyset \neq K' \subseteq \{1,\dots,r'\}}} B(K, K'),$$

where

$$B(K, K') = \left(\bigcap_{k \in K} \partial H_k \right) \cap \left(\bigcap_{k \notin K} H_k \right) \cap \left(\bigcap_{k' \in K'} \partial H'_{k'} \right) \cap \left(\bigcap_{k' \notin K'} H'_{k'} \right).$$

It therefore suffices to show that each set $B(K, K')$ is contained in a hyperplane that passes through u. We may assume $B(K, K')$ is nonempty.

$B(K, K')$ is a relatively open subset of $P(K, K') = (\cap_{k \in K} \partial H_k) \cap (\cap_{k' \in K'} \partial H'_{k'})$. The set $P(K, K')$ is an affine set, that is, an intersection of hyperplanes. If $P(K, K')$ is itself a hyperplane, then since $B(K, K') \subseteq \text{cl}(t) \cap \text{cl}(t')$, it follows that t and t' are adjacent. This contradicts the hypothesis. Therefore $P(K, K')$ is an affine set

23

of dimension at most $m - 2$, and we can then find a hyperplane containing both $P(K, K')$ and u. □

Lemma 3 *Let T be any polyhedral type space, and $u \in \mathbb{R}^m$ any utility function. Then there exists a finite collection Z of hyperplanes such that*

- *for any $t \in T$, ∂t is contained in the union of the hyperplanes in Z;*

- *if t, t' are distinct, nonadjacent types, and $w \in \mathrm{cl}(t) \cap \mathrm{cl}(t')$, then some hyperplane in Z passes through both u and w.*

Proof: Immediate from Lemmas 1 and 2. □

As in the text, if T is a polyhedral type space and u, v two utility functions, we say that v is *accessible* from u if $[u, v] \subseteq \cup_{t \in T} \mathrm{cl}(t)$.

Lemma 4 *Let T be any polyhedral type space, let $u \in \mathbb{R}^m$, and let Z be the set identified by Lemma 3. Let $v \in \mathbb{R}^m$ be any utility function not lying on any hyperplane in Z. Suppose v is accessible from u. Let $t_0, t_1, \ldots, t_r$ be the types intersecting the segment $[u, v]$ in order. Then t_k and t_{k+1} are adjacent, for each $k = 0, \ldots, r - 1$.*

Note in the statement that the phrase "in order" makes sense, since each type t_k intersects the segment $[u, v]$ in a subsegment, and these subsegments must be disjoint.

Proof: As in the proof of Proposition 1, define $u_\alpha = (1 - \alpha)u + \alpha v$ for $\alpha \in [0, 1]$. Any hyperplane in Z can contain at most one point of $[u, v]$: otherwise it would contain the entire segment and in particular would contain $u_1 = v$, contradicting the choice of v.

As noted above, each type t_k intersects $[u, v]$ in a subsegment $\{u_\alpha \mid \alpha \in J_k\}$, where J_k is an open subinterval of $[0, 1]$ — that is, J_k is of the form (γ, δ), $[0, \delta)$, or $(\gamma, 1]$. Write $\gamma_k = \inf J_k$, $\delta_k = \sup J_k$. By the assumption that the t_k are in order, we have $\delta_k \leq \gamma_{k+1}$ for each $k = 0, \ldots, r - 1$.

Next we show that in fact $\delta_k = \gamma_{k+1}$ for each k. Suppose instead that $\delta_k < \gamma_{k+1}$. By assumption, the union of the closures of all types in T contains all of $[u, v]$. So for any α with $\delta_k < \alpha < \gamma_{k+1}$, then, u_α is in the closure of some type $\widehat{t}$. Such a point u_α cannot belong to $\widehat{t}$ proper, so it belongs to $\partial\widehat{t}$. Then u_α belongs to one of the hyperplanes in Z. But there are infinitely many choices of u_α. Since Z is finite and each hyperplane in Z meets $[u, v]$ at most once, we have a contradiction.

24

This establishes $\delta_k = \gamma_{k+1}$. On the other hand, $u_{\delta_k} \in \mathrm{cl}(t_k)$, and $u_{\gamma_{k+1}} \in \mathrm{cl}(t_{k+1})$. So, $u_{\delta_k} \in \mathrm{cl}(t_k) \cap \mathrm{cl}(t_{k+1})$. If t_k, t_{k+1} are not adjacent types, then some hyperplane of Z passes through u_{δ_k} and u. This again contradicts the fact that each hyperplane of Z can intersect $[u, v]$ only once. So t_k, t_{k+1} are adjacent. $\qquad\square$

Proof of Proposition 2: Suppose $u \in t$, for some type $t \in T$, and let t' be any other type. Suppose the mechanism f satisfies the local incentive constraints. We wish to show that $u \cdot (f(t) - f(t')) \geq 0$.

Let Z be given by Lemma 3. Because t' is an open set, we can choose $v \in t'$ not lying on any of the hyperplanes in Z. By convexity, v is accessible from u, so Lemma 4 applies, and the successive types t_k, t_{k+1} identified in that lemma are adjacent for each k.

Again define $u_\alpha = (1 - \alpha)u + \alpha v$. For each $k = 0, \ldots, r$, pick any α_k with $u_{\alpha_k} \in t_k \cap [u, v]$. The local incentive constraints (t_k, t_{k+1}) and (t_{k+1}, t_k) for $k = 0, \ldots, r-1$ ensure that

$$
\begin{aligned}
u_{\alpha_k} \cdot (f(t_k) - f(t_{k+1})) &\geq 0, \\
u_{\alpha_{k+1}} \cdot (f(t_{k+1}) - f(t_k)) &\geq 0.
\end{aligned}
$$

From here we proceed exactly as in the proof of Proposition 1 to reach the conclusion

$$
u \cdot (f(t) - f(t')) = u \cdot (f(t_0) - f(t_r)) \geq 0.
$$

$\qquad\square$

Lemma 5 *Let T be the space of single-peaked ordinal types. Fix $u \in t$ for some $t \in T$. For any $t' \in T$, there exists a nonempty open set contained in t' such that every v in the open set is accessible from u.*

Proof: Let u, t' be as in the lemma. We know that u is strict (i.e. gives different values to different outcomes) since it belongs to an ordinal type. Then a sufficient condition for v to be accessible from u is that $(1-\alpha)u+\alpha v$ be single-peaked whenever it is strict: the set $\{\alpha \in [0, 1] \mid (1-\alpha)u+\alpha v \text{ is not strict}\}$ is finite, hence $(1-\alpha)u+\alpha v$ will be in the closure of some single-peaked ordinal type for each α.

We first construct some v that is accessible from u. Let x_p be the outcome ranked highest by u, and let $x_{p'}$ be the outcome ranked highest by t'. If $p' = p$, then any

25

$v \in t'$ is accessible from u: since $u(x_q)$ and $v(x_q)$ are both increasing in q for $q \le p$ and decreasing for $q \ge p$, the same is true of any weighted average $(1-\alpha)u + \alpha v$, so each such weighted average is single-peaked (as long as it is strict).

Now suppose $p' > p$ (the case $p' < p$ is similar). So $u(x_q)$ is decreasing and $v(x_q)$ must be increasing in q for $p \le q \le p'$. Choose $v(x_{p'})$ and $v(x_{p'-1})$ arbitrarily, with $v(x_{p'-1}) < v(x_{p'})$. If $p < p' - 1$ then successively choose $v(x_q)$ for $q = p' - 2, p' - 3, \ldots, p'$, such that

$$\frac{v(x_{q+2}) - v(x_{q+1})}{u(x_{q+1}) - u(x_{q+2})} < \frac{v(x_{q+1}) - v(x_q)}{u(x_q) - u(x_{q+1})}. \tag{2}$$

This can be done by choosing $v(x_q)$ low enough at each step. Finally, we can choose $v(x_q)$ for $q > p'$ or $q < p$ so that v represents the ordering given by t'.

Now we will show that, for $\alpha \in [0, 1]$, $(1-\alpha)u + \alpha v$ is single-peaked whenever it is strict. That is, we claim that $(1-\alpha)u(x_q) + \alpha v(x_q)$ is increasing in q for q up to some peak, and decreasing after that. Both $u(x_q)$ and $v(x_q)$ are increasing in q for $q \le p$, and decreasing in q for $q \ge p'$, so we focus on the range $p \le q \le p'$. We will show that there cannot exist any $q \in \{p, \ldots, p'-2\}$ such that

$$(1-\alpha)u(x_q) + \alpha v(x_q) > (1-\alpha)u(x_{q+1}) + \alpha v(x_{q+1}) \tag{3}$$

and

$$(1-\alpha)u(x_{q+1}) + \alpha v(x_{q+1}) < (1-\alpha)u(x_{q+2}) + \alpha v(x_{q+2}) \tag{4}$$

simultaneously hold; this will prove the claim.

Suppose (3) and (4) do both hold, for some q. (3) implies

$$\frac{1-\alpha}{\alpha} > \frac{v(x_{q+1}) - v(x_q)}{u(x_q) - u(x_{q+1})}$$

while (4) implies

$$\frac{1-\alpha}{\alpha} < \frac{v(x_{q+2}) - v(x_{q+1})}{u(x_{q+1}) - u(x_{q+2})}.$$

(Note we used the fact that $u(x_q) > u(x_{q+1}) > u(x_{q+2})$ to make sure the signs don't switch when we divide. We know $\alpha > 0$ since (4) is violated at $\alpha = 0$.) Combining these two inequalities gives a contradiction of (2), completing the proof of the claim.

At this point we have shown that any $v \in t'$ satisfying the inequalities (2) is accessible from u. Since these inequalities carve out a nonempty open subset of t',

the lemma is proven. □

Proof of Proposition 3: Suppose that T is the space of single-peaked ordinal types. Let $u \in t$, for some $t \in T$, and let t' be any other type. We wish to show that $u \cdot (f(t) - f(t')) \geq 0$, for any f satisfying the local incentive constraints.

Let Z again be the set of hyperplanes promised to us by Lemma 3 (with respect to T and u). By Lemma 5, we can choose a $v \in t'$ that is accessible from u and does not lie on any of the hyperplanes in Z. Accessibility ensures that Lemma 4 applies. From here onward we just repeat the argument used to prove Proposition 2. □

Proof of Proposition 5: Suppose that the mechanism (f, p) satisfies the set S_i of local incentive constraints for agent i. Consider any two types $t_i, t_i' \in T_i$. Write $t_\alpha = (1 - \alpha)t_i + \alpha t_i'$. As in the proof of Proposition 1, we can find $0 = \alpha_0 < \alpha_1 < \cdots < \alpha_r = 1$ with $(t_{\alpha_k}, t_{\alpha_{k+1}}), (t_{\alpha_{k+1}}, t_{\alpha_k}) \in S_i$ for each $k = 0, \ldots, r - 1$. These local incentive constraints give

$$E_i[u_i(t_{\alpha_k}, t_{-i}) \cdot (f(t_{\alpha_k}, t_{-i}) - f(t_{\alpha_{k+1}}, t_{-i})) + (p_i(t_{\alpha_k}, t_{-i}) - p_i(t_{\alpha_{k+1}}, t_{-i}))] \geq 0,$$
$$E_i[u_i(t_{\alpha_{k+1}}, t_{-i}) \cdot (f(t_{\alpha_{k+1}}, t_{-i}) - f(t_{\alpha_k}, t_{-i})) + (p_i(t_{\alpha_{k+1}}, t_{-i}) - p_i(t_{\alpha_k}, t_{-i}))] \geq 0.$$

Multiply by α_{k+1} and α_k, respectively, and add:

$$E_i \left[[\alpha_{k+1}u_i(t_{\alpha_k}, t_{-i}) - \alpha_k u_i(t_{\alpha_{k+1}}, t_{-i})] \cdot (f(t_{\alpha_k}, t_{-i}) - f(t_{\alpha_{k+1}}, t_{-i})) + \right.$$
$$\left. [\alpha_{k+1} - \alpha_k] \cdot (p_i(t_{\alpha_k}, t_{-i}) - p_i(t_{\alpha_{k+1}}, t_{-i}))] \geq 0. \quad (5)$$

Because utility is linear in own type, and t_{α_k} is equal to the weighted average $(\alpha_k/\alpha_{k+1})t_{\alpha_{k+1}} + (1 - \alpha_k/\alpha_{k+1})t_i$, we know that for each realization of t_{-i},

$$u_i(t_{\alpha_k}, t_{-i}) = \frac{\alpha_k}{\alpha_{k+1}}u_i(t_{\alpha_{k+1}}, t_{-i}) + \left(1 - \frac{\alpha_k}{\alpha_{k+1}}\right)u_i(t_i, t_{-i}).$$

Rearranging gives

$$\alpha_{k+1}u_i(t_{\alpha_k}, t_{-i}) - \alpha_k u_i(t_{\alpha_{k+1}}, t_{-i}) = (\alpha_{k+1} - \alpha_k)u_i(t_i, t_{-i}).$$

Applying this identity and dividing through (5) by the constant $\alpha_{k+1} - \alpha_k > 0$ gives

$$E_i[u_i(t_i, t_{-i}) \cdot (f(t_{\alpha_k}, t_{-i}) - f(t_{\alpha_{k+1}}, t_{-i})) + (p_i(t_{\alpha_k}, t_{-i}) - p_i(t_{\alpha_{k+1}}, t_{-i}))] \geq 0.$$

Summing over $k = 0, \ldots, r - 1$ gives

$$E_i[u_i(t_i, t_{-i}) \cdot (f(t_i, t_{-i}) - f(t'_i, t_{-i})) + (p_i(t_i, t_{-i}) - p_i(t'_i, t_{-i}))] \geq 0$$

which shows that the incentive constraint (t_i, t'_i) is satisfied.

$\square$

References

Abdulkadiroğlu, A., P. A. Pathak, and A. E. Roth (2005): "The New York City High School Match," *American Economic Review*, 95, 364-367.

Alcalde, J., and S. Barberà (1994): "Top Dominance and the Possibility of Strategy-Proof Stable Solutions to Matching Problems," *Economic Theory*, 4, 417-435.

Archer, A., and R. Kleinberg (2008): "Truthful Germs are Contagious: A Local to Global Characterization of Truthfulness," in *Proceedings of the 9th ACM Conference on Electronic Commerce* (EC-08), 21-30.

Ashlagi, I., M. Braverman, A. Hassidim, and D. Monderer (2010): "Monotonicity and Implementability," *Econometrica*, 78, 1749-1772.

Barberà, S., A. Bogomolnaia, and H. van der Stel (1998): "Strategy-Proof Probabilistic Rules for Expected Utility Maximizers," *Mathematical Social Sciences*, 35, 89-103.

Barberà, S., M. O. Jackson, and A. Neme (1997): "Strategy-Proof Allotment Rules," *Games and Economic Behavior*, 18, 1-21.

Barberà, S., H. Sonnenschein, and L. Zhou (1991): "Voting by Committees," *Econometrica*, 59, 595-609.

Bartholdi, J. J. III, and J. B. Orlin (1991): "Single Transferable Vote Resists Strategic Voting," *Social Choice and Welfare*, 8, 341-354.

Bartholdi, J. J. III, C. A. Tovey, and M. A. Trick (1989): "The Computational Difficulty of Manipulating an Election," *Social Choice and Welfare*, 6, 227-241.

Bergemann, D., and S. Morris (2009): "Robust Implementation in Direct Mechanisms," *Review of Economic Studies*, 76, 1175-1204.

Bikhchandani, S. (2006): "Ex Post Implementation in Environments with Private Goods," *Theoretical Economics*, 1, 369-393.

Bikhchandani, S., S. Chatterji, R. Lavi, A. Mu'alem, N. Nisan, and A. Sen (2006): "Weak Monotonicity Characterizes Deterministic Dominant-Strategy Implementation," *Econometrica*, 74, 1109-1132.

Black, D. (1948): "On the Rationale of Group Decision-Making," *Journal of Political Economy*, 56, 23-34.

Bogomolnaia, A., and H. Moulin (2001): "A New Solution to the Random Assignment Problem," *Journal of Economic Theory*, 100, 295-328.

Bogomolnaia, A., and H. Moulin (2002): "A Simple Random Assignment Problem with a Unique Solution," *Economic Theory*, 19, 623-635.

Celik, G. (2006): "Mechanism Design with Weaker Incentive Compatibility Constraints," *Games and Economic Behavior*, 56, 37-44.

Chung, K.-S., and J. C. Ely (2006): "Ex-Post Incentive Compatible Mechanism Design," unpublished paper, Northwestern University.

Danilov, V. I. (1994): "The Structure of Non-Manipulable Social Choice Rules on a Tree," *Mathematical Social Sciences*, 27, 123-131.

Dasgupta, P., and E. Maskin (2000): "Efficient Auctions," *Quarterly Journal of Economics*, 115, 341-388.

Demange, G. (1982): "Single-Peaked Orders on a Tree," *Mathematical Social Sciences*, 3, 389-396.

Duggan, J. (1996): "A Geometric Proof of Gibbard's Random Dictatorship Theorem," *Economic Theory*, 7, 365-369.

Ehlers, L., and B. Klaus (2003): "Probabilistic Assignments of Identical Indivisible Objects and Uniform Probabilistic Rules," *Review of Economic Design*, 8, 249-268.

Ehlers, L., H. Peters, and T. Storcken (2002): "Strategy-Proof Probabilistic Decision Schemes for One-Dimensional Single-Peaked Preferences," *Journal of Economic Theory*, 105, 408-434.

Fieseler, K., T. Kittsteiner, and B. Moldovanu (2003): "Partnerships, Lemons, and Efficient Trade," *Journal of Economic Theory*, 113, 223-234.

Gibbard, A. (1973): "Manipulation of Voting Schemes: A General Result," *Econometrica*, 41, 587-601.

Gibbard, A. (1977): "Manipulation of Schemes that Mix Voting with Chance," *Econometrica*, 45, 665-681.

Gibbard, A. (1978): "Straightforwardness of Game Forms with Lotteries as Outcomes," *Econometrica*, 46, 595-614.

Green, J. R., and J.-J. Laffont (1979): *Incentives in Public Decision-Making.* Amsterdam: North-Holland.

Green, J. R., and J.-J. Laffont (1986): "Partially Verifiable Information and Mechanism Design," *Review of Economic Studies*, 53, 447-456.

Heydenreich, B., R. Müller, M. Uetz, and R. Vohra (2008): "Characterization of Revenue Equivalence," Meteor Research Memorandum RM/08/01, Maastricht University.

Hylland, A. (1980): "Strategy Proofness of Voting Procedures with Lotteries as Outcomes and Infinite Sets of Strategies," unpublished paper, Harvard University.

Isaksson, M., G. Kindler, and E. Mossel (2010): "The Geometry of Manipulation: A Quantitative Proof of the Gibbard-Satterthwaite Theorem," in *2010 IEEE 51st Annual Symposium on the Foundations of Computer Science* (FOCS 2010), 319-328.

Jehiel, P., and B. Moldovanu (2001): "Efficient Design with Interdependent Valuations," *Econometrica*, 69, 1237-1259.

Jehiel, P., M. Meyer-ter-Vehn, and B. Moldovanu (2010): "Locally Robust Implementation and Its Limits," unpublished paper, Ecole Nationale des Ponts et Chaussées.

Jehiel, P., M. Meyer-ter-Vehn, B. Moldovanu, and W. R. Zame (2006): "The Limits of Ex Post Implementation," *Econometrica*, 74, 585-610.

Klaus, B., H. Peters, and T. Storcken (1997): "Strategy-Proof Division of a Private Good when Preferences are Single-Dipped," *Economics Letters*, 55, 339-346.

Maskin, E., and J. Riley (1984): "Monopoly with Incomplete Information," *RAND Journal of Economics*, 15, 171-196.

Meltzer, A. H., and S. F. Richard (1981): "A Rational Theory of the Size of Government," *Journal of Political Economy*, 89, 914-927.

Moulin, H. (1980): "On Strategy-Proofness and Single Peakedness," *Public Choice*, 35, 437-455.

Mussa, M., and S. Rosen (1978): "Monopoly and Product Quality," *Journal of Economic Theory*, 18, 301-317.

Myerson, R. B. (1981): "Optimal Auction Design," *Mathematics of Operations Research*, 6, 58-73.

Myerson, R. B., and M. A. Satterthwaite (1983): "Efficient Mechanisms for Bilateral Trading," *Journal of Economic Theory*, 29, 265-281.

Roberts, K. W. S. (1977): "Voting over Income Tax Schedules," *Journal of Public Economics*, 8, 329-340.

Rochet, J.-C. (1987): "A Necessary and Sufficient Condition for Rationalizability in a Quasi-Linear Context," *Journal of Mathematical Economics*, 16, 191-200.

Roth, A. E. (1982): "The Economics of Matching: Stability and Incentives," *Mathematics of Operations Research*, 7, 617-628.

Roth, A. E. (2008): "What Have We Learned from Market Design?," *Economic Journal*, 118, 285-310.

Saks, M., and L. Yu (2005): "Weak Monotonicity Suffices for Truthfulness on Convex Domains," in *Proceedings of the 6th ACM Conference on Electronic Commerce* (EC-05), 286-293.

Saporiti, A. (2009): "Strategy-Proofness and Single-Crossing," *Theoretical Economics*, 4, 127-163.

Sato, S. (2010): "A Sufficient Condition for the Equivalence of Strategy-Proofness and Nonmanipulability by Preferences Adjacent to the Sincere One," unpublished paper, Fukuoka University, October 2010 version.

Satterthwaite, M. A. (1975): "Strategy-proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions," *Journal of Economic Theory*, 10, 187-217.

Sher, I., and R. Vohra (2010): "Optimal Selling Mechanisms on Incentive Graphs," unpublished paper, University of Minnesota.

Sprumont, Y. (1991): "The Division Problem with Single-Peaked Preferences: A Characterization of the Uniform Allocation Rule," *Econometrica*, 59, 509-519.

Zhou, L. (1990): "On a Conjecture by Gale about One-Sided Matching Problems," *Journal of Economic Theory*, 52, 123-135.