Mill's harm principle, rationality, and Pareto optimality in games

Shaun Hargreaves Heap* Mehmet S. Ismail†

Wednesday 23rd October, 2024

Abstract

Mill's classic argument for liberty requires that people's exercise of freedom should be governed by the harm principle (MHP): that is, an action should not harm another. In this paper, we develop the concept of a Millian harm equilibrium (MHE) in n-person games where players maximize utility subject to the constraint of an MHP. Our main result is in the spirit of the fundamental theorems of welfare economics. We show that for every initial 'reference point' in a game the associated MHE is Pareto efficient and, conversely, every Pareto efficient point can be supported as an MHE for some initial reference point. This is an important result for an old question in political philosophy over whether the exercise of liberty is consistent with order in society and for how we think about policy in a non-ideal world.

Keywords: Pareto optimality, rationality, classical liberalism, harm principle, non-cooperative games

*Department of Political Economy, King's College London, London, United Kingdom. E-mail: s.hargreavesheap@kcl.ac.uk.

†Department of Political Economy, King's College London, London, United Kingdom. E-mail: mehmet.ismail@kcl.ac.uk.

1 Introduction

Is 'order' possible in a society where people are free to do as they please or does anarchy in this sense always entail some kind of 'disorder'? This is a central question in political philosophy. Hobbes (1651) famously answered that the disorder of 'warre' would result; and this provides a reason for people agreeing to constrain their individual liberties in some respects by a Leviathan. In this paper, we embed Mill's (1859) harm principle in non-cooperative game theory to provide a different and more positive conclusion regarding the prospects for liberty.

Two separate strands of formal analysis by economists have hitherto tended to side with Hobbes. In the social choice literature, for example, Sen (1970) famously shows that with some types of individual preference orderings, it is impossible to have a social choice rule that respects the liberal principle of a person's actions being guided by their own preferences at the same time as generating 'order' in the sense of a Pareto efficient allocation of resources.

In the other strand, economists spent much of the first twenty years after World War II establishing the conditions under which a market economy, where agents are free to act in whatever manner they think is best for them, will generate a Pareto efficient allocation of resources. The conditions are restrictive in ways that play into Hobbes's argument. The economy must be 'competitive' and this requires, among other conditions, that property rights are sufficiently well defined so that there are none of the externalities that cause the trouble of 'warre' in Hobbes's analysis. Agents in a competitive economy also have to be price takers, begging a question of where such prices come from. The latter difficulty prompted, in part, the development of non-cooperative game theory where the dismal conclusion for liberty is a key result. A Nash equilibrium is the standard solution concept used in game theory when rational choice agents act freely and the Nash equilibrium in many games is not Pareto optimal: e.g. in the famous Prisoners' Dilemma.

In this paper, we use the game theoretic approach to the analysis of social interaction to propose a different and more positive answer to the question regarding liberty and 'order'. In particular, we argue the exercise of individual liberty will generate Pareto efficiency when individual liberty is understood in the sense of JS Mill: i.e. that a person should be free to do as they please so long as they do not cause harm to others. The difficulty with Mill's proposal has always turned on how to define 'harm' for this purpose (e.g. see Lichtenberg, 2010). The game theoretic framework, however, provides, at least in principle, a natural way to operationalise 'harm'. The pay-offs to each agent in such games are given for each possible combination of agents' actions and so we say that an agent who is guided by Mill's harm principle (MHP) cannot take an action when that action reduces the pay-off, relative to some baseline reference point, of one or more other agents in the game. In short, harm occurs when someone causes a reduction to another person's pay-off. This seems uncontentious as a definition of harm, but its actual operationalisation will obviously depend on the choice of baseline reference point when making this judgement. For our purpose, however, this does not matter because we show that whichever baseline reference is chosen, the Millian harm equilibrium (MHE) is always Pareto efficient.

To put our result in perspective, we note that there have been developments with respect to both strands in the formal analysis of social interaction among free agents. The closest to our result comes initially from Coase (1960). Like us, his analysis is in the spirit of non-cooperative game theory, he takes the pay-offs that agents enjoy in each possible outcome as given and he allows for the troubling kinds of externalities between action that arise when property rights are imperfectly defined and which create the kind of social dilemmas where freedom and efficiency are opposed. He argues in a move that is later also taken by another Nobel laureate Ostrom (2010) that the agents do not have to accept the action sets that are specified by the game. They can conceive of taking actions which are outside this set. For example, with Coase, provided transaction costs are small, the agents have the scope, in effect, to internalise the externalities through bargaining. The point is that when freely chosen actions result in a Pareto inferior outcome, there is a prize to be gained in moving to a Pareto efficient outcome. If the cost of bargaining is not too high, then the people will rationally bargain with each other so as to agree on joint actions that secure this gain. If inefficiency results from the exercise of freedom under this approach, then it is only because the transactions costs of bargaining are too high. Ostrom makes, effectively, the same point, but it is stated in terms of whether the community has the social resources that enable them to establish and police Pareto improving joint actions.

In comparison, Pareto inferior outcomes only arise in our analysis when people are not constrained by the harm principle in the exercise of their liberty. When people are guided by our version of the harm principle then a Pareto efficient outcome will result. When there is more than one Pareto efficient outcome we cannot explain which will obtain as this will depend on what baseline is used for judging harms. But we show in a manner that is analogous to the second fundamental theorem of welfare economics that every Pareto efficient outcome can be reached with a suitable choice of the baseline.

The advances in social choice theory are less relevant to us because they focus on the conditions under which it is possible to have a well-behaved social choice function (i.e. a way of deciding between social outcomes); whereas our concern is with the properties of the outcomes that arise from the interaction of people who exercise their individual freedom. For this, the appropriate framework is non-cooperative game theory. Nevertheless, there are some interesting related results in the social choice literature. Mariotti and Veneziani (2009; 2013; 2020) introduce a notion called "Non-Interference" principle which roughly says that society's preferences should not change following a change in circumstances that affect only one individual and for which everyone else is indifferent. Recently, Mariotti and Veneziani (2020) show that there is inconsistency between their "Non-Interference" principle and the Pareto principle (i.e., if everyone in a society prefers an alternative x to y, then society should prefer x to y) in a non-dictatorship. Our formalization of the harm principle differs from Mariotti and Veneziani's in that ours applies to actions within a non-cooperative game theoretical framework whereas theirs applies to the preferences within a social choice context.

Chung (2019) and Chung and Kogelmann (2020) are concerned with the possibility of a social choice function when people have perspectival disagreements over how to characterise the possible outcomes in society. Although the social choice approach is different to ours, there are several interesting points of comparison in their approach. We focus now on the perspectival disagreements that are possible with their approach and return to the connections between their and our approach more generally in the Conclusion. Their exploration takes a rights approach whereby rights establish a private sphere over which a person can do as they like and the perspectival disagreements can relate either to the evaluation of outcomes in the private sphere and/or to whether particular outcomes belong in the private sphere. For example, with respect to the latter, does having an abortion belong in the private sphere of the individual considering an abortion, where it is a matter of individual choice, or the public sphere where constraints on taking such an action may be in place. As they acknowledge, this perspectival difference can be alternatively cast as a disagreement over what counts as a harm. For example, does 'A' having an abortion harm 'B'? 'B' thinks so and 'A' does not; and as a result there is a dispute over whether having an abortion belongs in the private sphere. We eschew the rights approach and instead work with Mill's harm principle. Our definition of a harm as a pay-off reduction avoids the disputes over what is in the private sphere (or counts as a harm) but we do not, as such, remove perspectival disputes. There is no requirement in non-cooperative game theory that person 'A' should assess 'B' having an abortion in the same way that person 'B' assesses 'A' having an abortion. All that is assumed is that when 'A' acts it is A's assessment that informs their calculation of what to do for the best and likewise when 'B' acts it is B's assessment that informs their actions.

To see this point in ways that makes our contribution clear, consider Gibbard's (1974) representation of the Sen's social choice problem: 'I want my walls to be white but care even more that Parker's be white; suppose Parker wants his walls yellow, but cares even more that mine be yellow' (p. 394). These preferences, as set out by Gibbard, are represented in Figure 1. They represent a particular kind of 'nosiness' (and hence the name of Parker), whereby each, in effect, dislikes to see the other happy with their preferred colour. One could say that there is a dispute about what counts as a harm here. Parker thinks that when Gibbard paints his wall white, it causes him (Parker) a 'harm': the harm of seeing Gibbard happy. Gibbard, though does not see this harm to Parker in painting his wall white. It is not reflected in his preference ordering and he does not take it into account in his own decision making. Equally, this could be cast as a perspectival dispute over each person's private sphere in the sense of Chung (2019) and Chung and Kogelmann (2020). For instance, the same preferences could arise over Gibbard's choice between white and yellow not because Parker has meddlesome preferences and hates seeing Gibbard happy but because Parker believes white paint is carcinogenic and yellow paint is not, while Gibbard holds no such belief about the outcome of painting his house white.

	White	Yellow
Yellow	$2, 2$	$0, 3$
White	$3, 0$	$1, 1$

Figure 1: A game between Gibbard (row player) and Parker (column player)

When Gibbard and Parker act freely so as to best satisfy their preferences, the resulting Nash equilibrium is (White, Yellow), where each causes each other a harm; and this is Pareto inferior to (Yellow, White). This is, of course, a Prisoners' Dilemma interaction and so makes the connection to Hobbes's state of nature when it is cast in the same way.[1] What we show is that in general when people are constrained in their actions by our version of the harm principle, outcomes that are Pareto inferior like (White, Yellow) in Gibbard's Prisoners' Dilemma game, will not result from the exercise of individual liberty. We cannot guarantee that the familiar co-operative solution to this Prisoners' Dilemma, (Yellow, White), will obtain when people are guided by our harm principle. There are three outcomes in this game that are on the Pareto frontier (White, White), (Yellow, Yellow) as well as (Yellow, White) and which obtains

[1] See Chung (2015) for an argument that Hobbes's state of nature should be understood differently, although from which he draws the same conclusion of 'disorder'.

under our harm principle will depend on what is the baseline reference outcome for judging whether an action causes a harm by reducing someone's payoffs.[2]

Our paper proceeds as follows. In the next section, we provide an informal introduction to the challenges posed by introducing the harm principle and sketch how we respond to them. We then set out our approach formally in section 3; and section 4 gives our main result. We reflect on this result in section 5 with some further illustrations. Section 6 concludes the paper.

2 An informal sketch of the challenges of the harm principle and our approach

The MHP does not permit a person to take an action that causes 'harm' to others. Several challenges arise when deciding how to represent this principle as a constraint on actions in non-cooperative games. In this section we offer an informal sketch of how we respond to these challenges by making four key assumptions.

Game theory has one advantage over the world that the judiciary addresses when trying to decide whether someone's actions cause a 'harm' to another: game theory deals with interactions where the pay-offs, usually captured by utility numbers, to each player in each outcome are given. With players' interests captured in this way by their pay-offs, it seems natural and uncontentious to say that a person suffers a 'harm' when their pay-offs are reduced. We deliberately use utility numbers to capture "pay-offs" in what follows because this

[2] It is worth noting that this thought process involves the players contemplating self harm and, at first glance this may seem strange in the context of the articulation of a harm principle where harm is to be avoided. This first thought, though, is misleading for two reasons. First, the prospective self harm is part of a thought process in an extensive form game and does not involve a player actually deciding to harm themselves. Second it is clear that Mill wants the harm principle to apply only to whether a person's actions harm someone else and not themselves. That is the point of his argument for liberty: it puts you in charge of yourself when your actions only affect yourself. If you cause yourself self harm, then so be it. This is nobody's business. Indeed, Mill plainly countenances such self harms when noting that his sense of liberty allows individuals to engage in what he famously refers to as 'experiments in living'. Indeed he commends such 'experiments in living'. The point about all experiments is that they sometimes fail.

allows for an encompassing definition of possible harms. They could be psychological or symbolic as much as material whereas to use \$ material pay-offs would be to restrict the concept of a harm to a \$ loss alone.

If a harm arises when a person's action reduces another's utility pay-off, a question naturally arises: a reduction relative to what? What is the reference point pay-off for judging whether the action causes a harm? This is the first modelling challenge. Since games contain all the relevant available actions for players in the setting captured by the game, we assume that the reference pay-offs must be given by one of the outcomes in the game. We make no argument over which outcome should be used. Any outcome might serve as the reference. Instead, we seek to characterize in general terms the equilibria that result when players take any of the outcomes in the game as the shared reference point. The specific attributes of an equilibrium that satisfies the MHP may depend on the actual reference point, but we are interested with any general properties of such equilibria.

The next challenge arises because the MHP requires that an individual's action should not cause a harm to any other person and outcomes in normal form games typically result from the joint actions of several players. We need, therefore, some way of making sense of how an individual's choice of action causes an outcome in a game and so judge whether that individual's action causes a harm.

In the context of non-cooperative games, Brams's (1994) seminal "Theory of Moves" framework offers a way of responding to this challenge. It takes one outcome in the game as the reference point (one might think of this as the status quo) and builds an extensive form game on the basis of the possible sequential player deviations from this reference point. In particular, at the reference point outcome, we assume the first player in this sequence decides between 'passing' their turn, 'staying' at the reference outcome, and 'moving' to an alternative outcome by changing their action so as to produce the alternative; and each player thereafter decides in this extensive form game between passing their turn, staying at the outcome they have inherited from the previous decisions of others in this branch of the extensive form game and moving to an alternative one by changing their action. A standard interpretation of this framework is that it is 'as if' the players in the game consider alternatives to the reference outcome through a quasi-bargaining sequential process whereby the first player can accept the reference outcome by 'staying' or propose an alternative by 'moving' to a different outcome, the next player then faces a similar choice to accept what is on the table or propose another, and so on, while 'passing' at any stage offers the turn to another player. To fix this application of Brams's Theory of Moves, consider the Prisoners' Dilemma game below.

	C	D
C	3, 3	1, 4
D	4, 1	2, 2

Suppose DD is the reference outcome and Row is the first player to consider a deviation. The extensive form game is constructed as follows. Row can either 'pass' their turn to Column, 'stay' at DD, or 'move' to CD by changing their action to C. If Row chooses to 'pass' at DD, then, Column faces the same decision at the second node of this extensive form game as the one we have just considered for Row. If Row chooses to 'stay' at DD, then DD remains the outcome. If Row chooses to 'move' to CD, Column at the second decision node in the extensive form game must decide between passing, staying at CD and moving to CC by changing their action from D to C; if Column chooses to pass at CD then Row decides between passing, staying at CD, and moving to DD; and so on. Figure 2 captures these early decision nodes in the extensive form game based on DD as the reference outcome.

The virtue for our purpose of adopting Brams's Theory of Moves in this way is twofold. First, every outcome in the normal form game can be reached through a sequence of decisions in the extensive form game. Second, this extensive form representation of how the outcomes in the game could be reached enables us to identify how an individual's action can be said to cause a particular outcome. Thus, we can say in our illustration that, if Row begins by 'moving' to CD through a 'move' to C from the reference point of DD, then

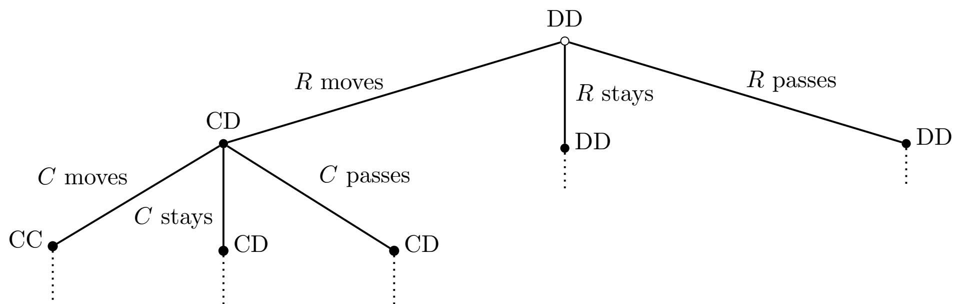

Figure 2: Illustration of early decision nodes in the extensive form game based on DD as the reference outcome

Row has caused CD at this point in the extensive form game. Likewise, if Row initially decides to 'stay' at DD, they have caused DD. However, when Row 'passes' at DD, they do not cause DD because they have exempted themselves from decision making by passing the choice to the Column player. (In effect, giving players an option to 'pass' endogenises the player order for these sequential deviations and so has no influence in a 2×2 symmetric game like the Prisoners' Dilemma. In more complicated n-player games, 'pass' has further technical function of enabling every outcome to be visited through a sequence of player deviations.)

Brams (1994) assumes farsighted or sequential rationality and solves by backward induction the extensive form game created by his procedure of sequential deviations. He calls the outcome of his proposed sequential procedure a nonmyopic equilibrium (see also Brams and Wittman, 1981, and Kilgour, 1984). We adopt the same approach of applying the subgame perfect equilibrium solution concept but we introduce the MHP as a constraint on play in the extensive form game and so call the outcome a Millian harm equilibrium.[3] Thus, to return to the prisoners' dilemma with DD as the initial reference outcome, we ask players whether the subgame perfect equilibrium of the MHP constrained extensive form game based on DD is DD or some other outcome. If it is DD, then DD is the Millian harm equilibrium outcome of the game. If

[3]For a comparison of nonmyopic equilibrium and Millian harm equilibrium, see section 5.

it is not, then the Millian harm equilibrium with DD as the reference outcome is whatever the subgame perfect equilibrium is in this MHP constrained extensive form game. The next challenge is, therefore, to represent the MHP in this extensive form game. This requires three further assumptions.

One is innocuous in the sense that a well-defined finite extensive form game requires a set of terminal nodes. One part of how we do this is by saying that when both players decide to 'stay' at an outcome, then this outcome is implemented. The intuition behind this assumption is that the first decision to 'stay' is like a 'proposal' to implement this outcome and the second decision to 'stay' amounts to an acceptance of this 'proposal'. This naturally produces a terminal node for some branches of the extensive from game. Likewise, if both players decide to pass, then this naturally produces a terminal node. However, we also need to prevent the 'moving' branches of the tree creating what are infinite cycles through the possible outcomes in the game. We do this by preventing a player repeating the same action ('move' or 'stay') twice. Some number of finite repetitions has to be assumed if the extensive form game formed by adopting Brams's Theory of Moves is to be finite. So, some restriction is necessary for the Brams approach to be meaningful (i.e. it is a consequence of adopting that approach). Nevertheless, the choice of the precise number of repetitions that is permitted, once in this instance, is arbitrary. For this reason, we show in section 4.1 that our results hold irrespective of the actual number of finite repetitions that are permitted.

To illustrate the specific once only repetition assumption, suppose, for example, we follow the branch in the extensive form game that begins with Row 'moving' to CD, Column next 'moves' to CC, Row then 'moves' to DC and Column 'moves' to DD. Row cannot 'move' to CD again. Row can either 'stay' or 'pass' but if they choose to 'stay', they cannot choose to 'stay' again at DD. Thus, DD is the terminal outcome either because both players 'stay' or both 'pass' at DD for this truncated move branch of the extensive form game.

The question arises as to what outcome should be implemented if players reach a terminal node through joint 'passing'—either because the truncation rule has been triggered or because players have both chosen to 'pass' at an earlier point in the extensive form game. The terminal outcome cannot be said to have been chosen by the players in these circumstances because neither has decided to stay at this terminal outcome. We assume, therefore, that its reference point is implemented because no other outcome along the path to this terminal outcome, including the terminal outcome itself, has been consciously endorsed by both players through 'stay' decisions.

The next assumption embodies the MHP: a player can only decide to 'stay' at an outcome if that outcome does not harm any other player relative to the reference outcome. We apply the MHP to the 'stay' decision because an individual can only contribute to causing an outcome that is implemented by deciding to 'stay' at that outcome. Of course, it takes more than one decision to 'stay' for an outcome to be implemented. But in so far as one individual could contribute to an outcome being implemented, they would do so by deciding to 'stay' at that outcome. For the same reason, the MHP does not apply to a 'move' decision: the condition for implementation is two 'stay' decisions and so only 'stay' can be said to contribute to causing and potentially implementing an outcome. To illustrate, in the extensive form game that begins with DD as the reference outcome in the prisoners' dilemma, Row 'staying' at DD satisfies the MHP and so does Row 'moving' to CD (because MHP only applies to 'stay' decisions).[4] However, Column's option to 'stay' at CD would not satisfy the MHP (because it harms Row), but moving to CC does (because MHP does not apply to 'move'). Thus, the beginning of the MHP constrained extensive form game looks like Figure 3 (with Column stays at CD faded out as an option as compared with the unconstrained version in Figure 2).

This illustration gives an immediate insight into how sequential rationality and the MHP might combine to make CC the MHE in this extensive form game. The pursuit of the best option for Column at the second CD node by 'staying' is precluded by the MHP. At any later decision node on this branch of the game, Row will not be able to 'stay' at DC for the same reason and so the

[4]NB although Row's move to CD may seem like an act of self harm, it is not because CD is not implemented. To be implemented Column would have to stay at CD and this transgresses the MHP. To discover whether the move to CD is in the interests of Row, we need to solve for what finally happens along this path of the extensive game.

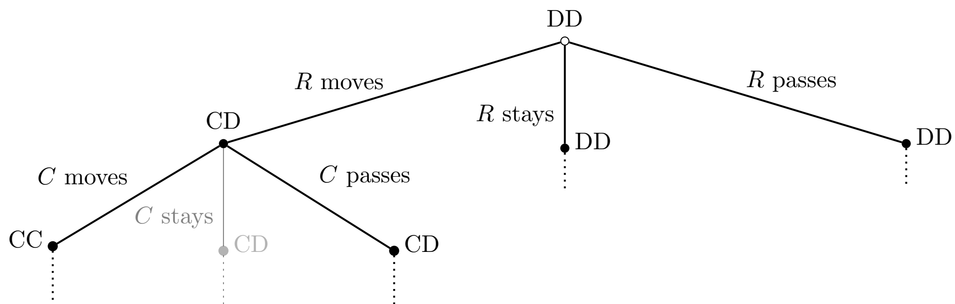

Figure 3: The beginning of the MHP constrained extensive form game where DD is the reference outcome.

only possible terminal options along this branch will be DD (either through Row's initial 'stay' decision or through sequential 'moves' to this terminal node followed by mutual 'passes' leading to its reference point, DD, being implemented) and CC (through mutual 'stay' decisions at this point). Thus either DD or CC will be implemented and sequential rationality will secure CC. The details are, of course, a bit more complicated. One comment is worth making, nevertheless.

Our MHP principle deliberately does not embody sequential rationality. We introduce sequential rationality as a separate assumption. The application of sequential rationality requires an assumption of common knowledge of rationality and there is no reason to bind the MHP to such an assumption. A 'harm' is a 'harm' whether the other player is rational or not. The most that can be said in such circumstances with respect to whether a player's action causes a harm, is: does that action by itself do as much as any individual can do to causing an outcome to be implemented that harms someone else? That is, do they choose to 'stay'.

Two final details in our approach are worth noting at this stage. First, we require only two stay decisions at a particular outcome for it to be implemented in an 'n-person' game. This is because we wish to avoid building in Pareto improvements through some version of unanimity rule that requires everyone to agree on some outcome before it is implemented. Our condition for implementation is, therefore, in general, much weaker than unanimity. However, in a two-person game it does amount to unanimity and this supplies another part of the intuition behind why CC emerges as the Millian harm equilibrium in the prisoners' dilemma illustration. In section 4.1.4, we extend our model to the case in which any player can unilaterally implement the outcome on their turn. These and other natural possible modifications do not affect our results.

Second, a final key assumption for more complicated games than the two person prisoners' dilemma is that if say player i decides to stay at an outcome but the next player j decides to reject this proposal by moving to another outcome, then the reference point is updated to the outcome proposed by i through their decision to 'stay'.[5] In effect, this outcome has been endorsed by i, it satisfies the MHP for i and it could have been implemented by j; so, it is natural to use this as the (new) reference point for judging future deviations. Thus, in general, the decision to 'stay' is also a decision to change the reference point and this can only be done by a player when to do so satisfies the MHP.

3 The harm principle in non-cooperative games

3.1 The setup

Let $G = (A_i, u_i)_{i \in N}$ denote a normal form game in which $N = \{1, 2, \ldots, n\}$ denotes a society whose members are called players, A_i finite pure action set of player i, $u_i : A \to \mathbb{R}$ player i's Bernoulli utility function representing strict preferences over the set of action profiles $A = \times_{i \in N} A_i$. Let $a = (a_1, a_2, ..., a_n) \in A$ denote a pure action profile in game G.[6] As is standard in normal form games, every action profile is associated with an outcome (i.e., a pay-off profile) and vice versa. We use the terms "action profile" and "outcome" interchangeably.

A profile a Pareto dominates a' if for all i, $u_i(a) \geqslant u_i(a')$ with at least

[5] This is reminiscent of 'evolving' status quos in dynamic bargaining and democratic deliberation models; see, e.g., Baron (1996), Kalandrakis (2004), Bowen and Zahran (2012), and Chung and Duggan (2020). We refer the interested reader to section 4.1.7

[6] Our definitions can be extended to the games with mixed strategies in a straightforward way. We keep the current framework for its simplicity.

one strict inequality. A profile is called Pareto optimal or efficient if there is no other profile that Pareto dominates it. A profile is called weakly Pareto optimal if there is no other profile in which everyone is strictly better off.

Fix a game $G = (A, u)$ and action profile $a_0 \in A$. Given a player function I, we define an associated extensive form game with perfect information denoted by $\Gamma(a_0, I) = (N, X, I, u, \Sigma, H)$. We refer to this extensive form game as simply "Γ". The interpretation of Γ is that starting from a_0, each player i sequentially decides to 'stay', 'move' or 'pass' in game G until the play terminates. Table 1 summarises our notation.[7]

Let X denote a game tree, $x \in X$ a node in the tree, x_0 the root of the game tree, and $z \in Z$ a terminal node, which is a node that is not a predecessor of any other node.

Name	Notation	Element
Players	$N = \{1, 2, \ldots, n\}$	i
Information set	H	h
Nodes	X	x
Terminal nodes	Z	z
Player function	$I : X \to N$	
State function	$S : X \to A$	
Actions at node x	$A_i(x)$	$a_i(x)$
Strategy profiles	Σ	σ
Path of play of σ	$[\sigma]$	x_j
Terminal node of $[\sigma]$	$\overline{[\sigma]}$	
Reference points of σ	$R(\sigma)$	
Utility of i at σ	$u_i(\sigma)$	
Extensive form game	$\Gamma(a_0, I) = (N, X, I, u, \Sigma, H)$	

Table 1: Extensive form game notation

[7] For a standard textbook on extensive form games, see, e.g., Osborne and Rubinstein (1994).

3.1.1 Player function

Let $[x_m] = \{x_0, x_1, x_2, ..., x_m\}$ denote the *path of play* between node x_0 and node $x_m \in X$ where for every $j = 0, 1, ..., m-1$, x_{j+1} is an immediate successor of x_j. Let $I : X \to N$ be the player function, where $I(x)$ gives the "active" player who moves at node x. The only restriction we impose on the player function is the following. For player i, let $[x_m]_i = |\{x' \in [x_m] | I(x') = i\}|$, i.e., the number of times player i is active during the path of play $[x_m]$. We assume that for every player i, every player $j \neq i$, and every path of play $[x_m]$, $|[x_m]_i - [x_m]_j| \leqslant 1$. In other words, if a player has been active $\bar{m}$ times in some path of play, then every other player should have been active at least $\bar{m} - 1$ times. This assumption ensures that every player has more or less the same number of moves to play on every path. Beyond this assumption, notice that the order in which players take turn is not fixed and at every non-terminal node x the next active player may depend on the particular action chosen by player $I(x)$.

3.1.2 Actions, strategies, and information sets

Let $h \in H$ denote an information set, which is a singleton. With a slight abuse of notation, an information set h at node x is denoted by x, i.e., $h = x$. A *subgame* $\Gamma|x$ of a game Γ is the game Γ restricted to an information set $h = x$ and all of its successors in Γ.

Next, we introduce $S : X \to A$, called the *state function*, that maps each node $x \in X$ to an action profile $a \in A$. We define S by induction. The state at the root x_0 of the game is defined as a_0, i.e. $S(x_0) = a_0$, which is the action profile in G where the extensive form game Γ starts. Let $x \in X$, $x \neq x_0$, be a node, $x' \in X$ the immediate predecessor of x, $i = I(x')$, and a'_i player i's action that leads to node x. Assume that $S(x') = a$. Then, define $S(x) = (a'_i, a_{-i})$. In other words, at every node x, the state $S(x)$ is given by the action profile $(a'_i, a_{-i}) \in A$ such that player $i = I(x')$ changes only the i'th component of the state at x'.

Let $A_i(x)$ denote the set of pure actions of player i at x. For each $x \in X$,

$A_i(x)$ is defined as follows. First, define $X'(a_i, x) = \{x' \in X | a_i$ is chosen at x', $S(x') = S(x)$, and x' is a predecessor of $x\}$. Then, $A_i(x) = \{a_i \in A_i |\ X'(a_i, x) = \varnothing\} \cup \{p\}$, where p stands for pass. For example, suppose that $i = I(x_0) = I(x'')$, $x_0 \neq x''$, and $S(x_0) = S(x'')$. If i chooses a_i at x_0, then $a_i \notin A_i(x'')$ because a_i has already been chosen at state $S(x_0)$. This means once a player returns to a state, they cannot make the same decision as they made last time they were at this state. This is the assumption that prevents infinite cycling through always choosing to 'move' at each decision node and/or through one player staying and the other(s) passing, repeatedly at the same state. When this truncation rule is binding at a decision node x, players can only choose to pass at x.

Let $A'_i = \bigcup_{x \in X_i} A_i(x)$ denote player i's set of all pure actions where X_i is player i's set of all information sets. Let $\Sigma_i = \bigtimes_{x \in X_i} A_i(x)$ denote the set of all pure strategies of i where a pure strategy $\sigma_i \in \Sigma_i$ is a function $\sigma_i : X_i \to A'_i$ satisfying $\sigma_i(x) \in A_i(x)$ for all $x \in X_i$. Let $\sigma \in \Sigma$ denote a pure strategy profile and $u_i(\sigma)$ its (Bernoulli) utility for player i.

Let $a'_i \to x'$ denote player i's action $a'_i \in A_i(x)$ that leads to node $x' \in X$. Let $[\sigma] = \{x \in X | \sigma_i(x') \to x$ for some $i \in N, x' \in X\} \cup \{x_0\}$ be the path of play of σ and $\overline{[\sigma]}$ be the terminal node in $[\sigma]$.

3.1.3 Reference points, terminal nodes, and utility functions

For a given strategy profile σ, let

$$R(\sigma) = \{a \in A | x \in [\sigma], \sigma_i(x) = a_i \in A_i(x), a = S(x)\} \cup \{a_0\}$$

be the set of all *reference points* of σ. In other words, a state is called a reference point if the player who acts at the associated node "stays" at it: that is chooses not to change it. The initial reference point a_0 is included in $R(\sigma)$.

Note that given a profile σ, for every decision node $y \neq x_0$ there is a unique reference point $a_y \in R(\sigma)$ where $S(x) = a_y$ for some predecessor x of y. The unique reference point at x_0 is a_0 by definition. Thus, we can refer to *the* reference point at every node $y \in X$.

We next define 'off-path' reference points. Let $y \neq x_0$ be a non-terminal node. The set of reference points of $(\sigma|y)$, denoted by $R_{|y}(\sigma)$, is defined as $R(\sigma|y)$ except that a_0 is replaced with a_y, which is the reference point at y, $a_y \in R(\sigma)$. The intuition is that if we restrict a strategy profile σ to a node y, then the initial reference point of $(\sigma|y)$ should be a_y and not necessarily a_0.

Game Γ comes to an end under two situations. First, let x' be a node and x be a (not necessarily immediate) successor of x' such that $I(x') = i$, $I(x) = j$, $j \neq i$, and the reference point at x is $S(x')$ where $S(x') = S(x) = b$. If player i stays at b by choosing b_i, making b the reference point, and player $j \neq i$ also stays at b by choosing b_j, then node x is called a *terminal node*. Second, let $\{x^1, x^2, ..., x^n, x^{n+1}\}$ be a path of play such that for every player $i \in N$ there exists x^m such that $i = I(x^m)$ where $n \geqslant m \geqslant 1$, and for every m, x^{m+1} is an immediate successor of x^m. Node $x^{n+1} \in X$ is called a *terminal node* if every i chooses p (i.e., pass) at x^m. In plain words, the game terminates if either (i) two distinct players choose to stay at a state (the first one is like a 'proposal' to implement this state and the second one amounts to an acceptance of this 'proposal'), or (ii) every player consecutively passes their turn.

Let $\sigma \in \Sigma$ be a strategy profile, $\overline{[\sigma]} = z$ its terminal node, and $a \in R(\sigma)$ the reference point at z. We define the *outcome* of σ as a. With slight abuse of notation we use the same utility function for $u_i(\sigma)$ and $u_i(a)$ because their outcomes, and hence their utilities are the same. In summary, for every player i, $u_i(\sigma) = u_i(a)$, where a is the reference point at $z = \overline{[\sigma]}$. Put simply, the reference point at the terminal node is implemented as the outcome of the relevant strategy profile under both (i) and (ii) above. The reason the reference point at the terminal node is implemented as the outcome in condition (ii) is that no other state from the reference point to the terminal node has been endorsed by any player through a 'stay' decision and that players have consciously chosen not to stay at the terminal node. As mentioned earlier, (ii) is in part a technical condition that prevents infinite cycling. In section 4 (proof of the main theorem), we show that there is always an MHE whose outcome is attained under condition (i).

An illustrative example

To illustrate our notation, we return to the prisoners' dilemma (PD). Let $\Gamma(a_0, I)$ be the extensive form game that begins this time with the reference point $a_0 = \text{CC}$. Assume that Row (player 1) moves first, Column (player 2) moves second, and this sequential order strictly alternates irrespective of players' choices. Starting from CC players might end up at DD if they play as follows (see Figure 4). Row unilaterally switches their action to D, hence 'moving' to DC. Column then moves to DD, where Row chooses D to 'stay' which makes DD the new reference point. Column also stays at DD, where both players receive pay-offs of (2, 2). If, instead of staying, both Row and Column choose to pass at DD, then the implemented outcome would be the reference point at this node, which is CC. The difference between passing and staying is that passing changes the order of play but does not change the reference point.

Note that a 'cycle' cannot be repeated. Suppose, for example, that Row moves to DC from CC, Column moves to DD, Row moves to CD, and Column moves back to CC. Then, Row cannot choose D again at CC. Row can now only either stay or pass at CC.

We have not so far imposed any restrictions on the players' choices such as 'rationality' or 'harm principle'. We next introduce the MHP.

3.1.4 The harm principle

Our specification of the harm principle applies when a player stays. We make this assumption for two reasons. First, in a dynamic strategic setting, the classical liberal has no reason to be concerned with the properties of any transitional (i.e., non-reference point) states in the extensive form game, particularly if they are purely mental constructs. Second, in contrast when a player chooses to stay, this matters for everyone because either another player follows this by choosing to stay and this becomes the implemented outcome; or, in so far as the play moves to another outcome, the reference point changes through the stay decision and this conditions future play and the eventual out-

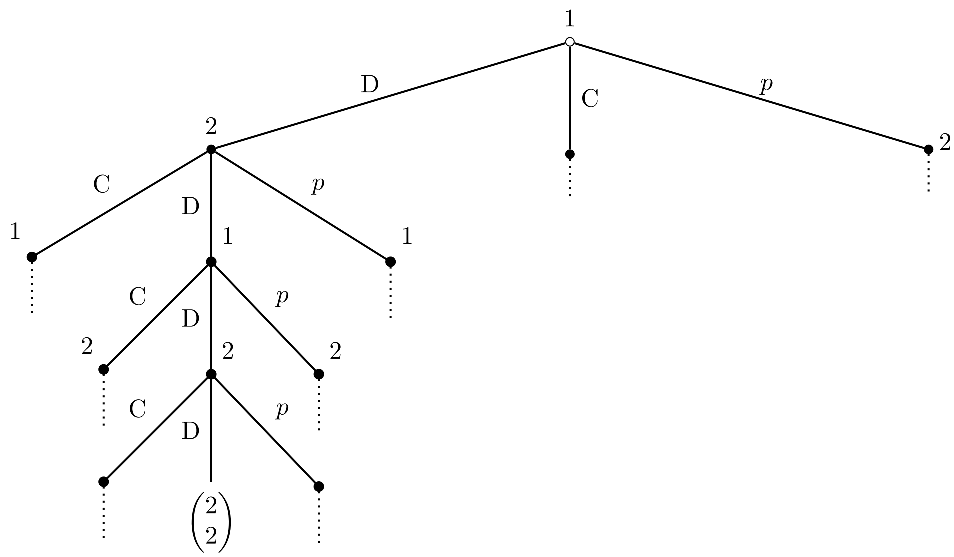

Figure 4: An illustrative example where a_0 = CC in the PD. Row (player 1) moves first and Column (player 2) moves second.

come. Indeed, an individual can only influence the character of the eventual outcome either directly or indirectly by choosing to stay because this changes the reference point. The point is that the only other way that a terminal node is reached is by mutual decisions to pass, in which case the original or prior reference point is implemented and the decision to pass has not affected the implemented outcome.

Definition 1 (Harm principle)**.** Let Γ be a game, σ a strategy profile, $x \in X$ a non-terminal node, $b \in R(\sigma)$ the reference point at x, and $S(x) = a$. Action $\sigma_i(x) = a_i$ is said to satisfy the harm principle (MHP) at x if for every $j \neq i$, $u_j(a) \geqslant u_j(b)$. Strategy profile σ satisfies the MHP at x_0 if for every i and every x as defined above, $\sigma_i(x) = a_i$ satisfies the MHP. Finally, strategy profile σ satisfies the *harm principle* if for every non-terminal x', $(\sigma|x')$ satisfies the MHP at x'.

In plain words, a player's stay action satisfies the MHP if their decision does not harm others with respect to the current reference point (e.g., see Figure 3).

Accordingly, a strategy profile satisfies the MHP if every player's every stay action (both on-path and off-path) under that strategy profile satisfies the MHP.

Of note, the harm principle implies neither Pareto optimality nor even Pareto improvement from a reference point. In section 4.1.1, we illustrate that assuming the harm principle may lead a society to a Pareto inferior outcome compared to the initial reference point. Even in situations in which the MHP leads to a Pareto improvement, the outcome of the game may be Pareto dominated as we illustrate in section 4.1.3.

3.2 The Millian harm equilibrium

We assume that players are sequentially rational—i.e., players maximize their utility at every node given the others' strategies—and are additionally constrained by Mill's harm principle (MHP) in their action choices in Γ. Moreover, we assume that G, Γ, and the previous sentence are common knowledge (Lewis, 1969; Aumann, 1976). First, we define subgame perfect equilibrium (Selten, 1965; Nash, 1951).

A pure strategy profile $\sigma \in \Sigma$ in game Γ is called a *subgame perfect equilibrium* (SPE) if for every player i and for every non-terminal $x \in X$ where $i = I(x)$, $u_i(\sigma|x) \geqslant u_i(\sigma_i', \sigma_{-i}|x)$ for every $\sigma_i'|x \in \Sigma_i|x$. Put differently, σ is a subgame perfect equilibrium if it constitutes a Nash equilibrium in every subgame of Γ.

Definition 2 (Millian harm equilibrium)**.** Let $G = (A, u)$ be a game. A pure strategy profile $\sigma^* \in \Sigma$ that satisfies the harm principle is called a *Millian harm equilibrium* (MHE) in G if for every player i and for every non-terminal $x \in X$ where $i = I(x)$

$$u_i(\sigma^*|x) \geqslant u_i(\sigma_i', \sigma_{-i}^*|x)$$

for every $\sigma_i'|x \in \Sigma_i|x$ such that $(\sigma_i', \sigma_{-i}^*) \in \Sigma$ satisfies the harm principle.

In plain words, a strategy profile is an MHE if at every node the active player plays a best response under the constraint of the harm principle. Like subgame perfect equilibrium, in finite games Millian harm equilibria can be computed using backward induction under the constraint of the MHP. Note that an MHE is *not* equivalent to a strategy profile that is both a subgame perfect equilibrium and satisfies the harm principle, in part because in general there may be no SPE that satisfies the MHP, but as we show in section 4 an MHE always exists.

An MHE in G depends, of course, on $\Gamma(a_0, I)$, i.e., the initial reference point, a_0, and the player function I. But for now it is important to note that the MHP *per se* does not require Pareto optimality of the outcome. Players simply act independently and maximize their individual utility; they do not act to maximize the pay-offs of others. They can stay wherever they want as long as the outcome does not harm others with respect to the reference point and there could always be other outcomes that are as good for the individual who decides to 'stay' and which would be better for the other players. We illustrate this point in section 4.1.1 with an example where the MHP by itself does not produce a Pareto efficient outcome (see also section 4.1.3). We next show under what conditions the MHE outcomes are Pareto optimal in n-person games.

4 The main theorem

In this section, we first show that the MHE exists under general conditions in normal form games.

Lemma 1 (Existence)**.** *Let $G = (A, u)$ be a game with strict preferences. For every initial reference point $a_0 \in A$, for every player function I, there exists an MHE associated to a_0 in pure strategies.*

Proof. We fix an initial reference point a_0 and a player function I.

Notice that for every a_0, the game Γ always possesses a pure subgame perfect equilibrium. This is true because Γ is a well-defined finite extensive form game with perfect information. To see this, notice that the root of the game is x_0 where $S(x_0) = a_0$ and that every player function I gives a unique player at every non-terminal node by construction of Γ. Because there are finitely many players and nodes, the game Γ ends after finitely many steps. This implies that there is always a subgame perfect equilibrium in pure strategies.

Next, we assume that players act according to the MHP, which essentially puts a constraint on their choices in Γ. This implies that they have fewer (finitely many) choices under the MHP than they have under Γ. Because the MHP is common knowledge, the constrained game—i.e., the game in which all strategy profiles satisfy the MHP—is still of perfect information. Let σ^* be a subgame perfect equilibrium in the constrained game, which exists by the same arguments as above. We note that σ^* is an MHE in Γ because σ^* satisfies the MHP and at every node every active player plays a best response among the profiles that satisfy the MHP, since by construction all those profiles satisfy the MHP. This concludes the proof that σ^* is an MHE. □

We next show under what conditions the uniqueness of the MHE outcome is guaranteed from an initial reference point.

Lemma 2 (Uniqueness). *Let $G = (A, u)$ be a game with strict preferences. For every initial reference point $a_0 \in A$ and every player function I, the MHE outcome associated to a_0 is unique.*

Proof. Given an initial reference point a_0, and a player function I, the associated Γ possesses a pure subgame perfect equilibrium as shown in the proof of Lemma 1. We next show that this subgame perfect equilibrium outcome is unique. The reason is that no matter which player moves on a non-terminal node either (i) the player has a unique pure best response or (ii) the pure best responses all lead to the same outcome because the preferences of the players are strict in G. Thus, the subgame perfect equilibrium outcome in Γ must be unique. Analogously, the subgame perfect equilibrium outcome in Γ which is constrained by the MHP must also have a unique outcome. Together with Lemma 1, this implies that the MHE outcome must be unique. □

Finally, we illustrate the relationship between the harm principle, rationality, and efficiency in n-person normal form games.

Main Theorem (Efficiency)**.** *Let* $G = (A, u)$ *be a game with strict preferences. For every Pareto optimal outcome* $a \in A$ *there exists an initial reference point* $a_0 \in A$ *such that for every player function* I*, the associated MHE outcome is* a*. Conversely, for every initial reference point* $a_0 \in A$ *and every player function* I*, the associated MHE outcome is Pareto optimal.*

The proof of this theorem is in the Appendix 6. Here we give an informal sketch of the proof.

Given a game G with strict preferences and a player function I, we first show that if an initial reference point a_0 is Pareto optimal then it is the MHE outcome from a_0. By way of contradiction, suppose that $a' \neq a_0$ is the MHE outcome. It implies that there exists at least one player who chose to stay (i.e., changed the reference point) in the path of play of an MHE, σ^*. Every player who did *not* stay receives a strictly greater pay-off at a' than a_0 because σ^* satisfies the MHP. In addition, every player who did stay must, due to sequential rationality and the MHP, receive a strictly greater pay-off at a' than a_0. As a result, a' Pareto dominates a_0, which contradicts the supposition that a_0 is Pareto optimal.

Second, we show that for an initial reference point a_0 that is not Pareto optimal, the MHE from a_0 must be Pareto optimal. By way of contradiction, suppose that b is the MHE outcome from a_0 and b is Pareto dominated by some action profile $a \neq b$. Let σ^* be an MHE from a_0 such that the first time a player stays at b on the path of play of σ^*, the next player (say, i) also stays at b by choosing b_i, hence terminating the game. Notice that if (i) there is a path from b to a along which the MHP is satisfied, then b_i cannot be a best response of player i because (ii) every player (including i) receives a strictly greater pay-off at a than b by our supposition that b is Pareto dominated by a, and (iii) no other player can stay at an action profile which harms player i along the path to a since the MHP applies and b is the reference point. We next show that (i) is true. First, notice that players can reach from b to a in at most n moves by the following path of play. At every node, the active player i plays move a_i except when $a_i = b_i$, in which case player i plays p (i.e., pass). The MHP is not violated along this path of play because no player stays. Second, this path of play does not overlap with the path of play of σ^* because if it did, then the active player at the overlapping node would have a profitable deviation to the path towards a. The reason is that if the state of the overlapping node x is a, then the active player would have a profitable deviation from σ^* to stay at a and make a the reference point because they are strictly better off at a and staying at a satisfies the MHP. By backward induction, the two paths of play cannot include an immediate predecessor x' of node x because the active player at x' would have a profitable deviation to x, where the next player would stay. By analogous backward induction reasoning, one can conclude that the path of play from b to a and the path of play of σ^* have an empty intersection. Thus, statement (i) holds as well. As desired, we reach a contradiction: b cannot be the MHE outcome from a_0.

4.1 Discussion of the assumptions

We next discuss how different assumptions in the definition of Γ and the MHE affect the results.

4.1.1 The harm principle

To see why the MHP is essential for the main theorem, first notice that the MHE definition would reduce to subgame perfect equilibrium if the harm principle were not assumed. Consider the following simple example and suppose that the MHP is *not* assumed.

	L	R
L	$4,3$	$1,4$
R	$2,1$	$3,2$

Let the initial reference point be (1,4). Suppose that Row moves first, Column moves second, and this order strictly alternates. Row would not choose to pass at (1,4) because Column would then choose to pass too, making (1,4) as the outcome. On grounds of sequential rationality, Row's best response is to move from (1,4) to (3,2), where Column as well as Row would stay, making it the outcome. To see this, first notice that Column would not gain by moving to (2,1) from (3,2) because Row would not move to (4,3) as Row anticipates that Column would then go back to (1,4) where Row would have to either stay or pass. If Row stays at (1,4), then Column would simply make (1,4) the outcome by staying too. If Row passes at (1,4), then Column would also pass, making (1,4) the outcome. Second, notice that Row would not move back to (1,4) from (3,2), because Column would then stay there. Thus, without the harm principle and starting at (1,4), players would end up at (3,2), and this is Pareto dominated by (4,3).[8]

4.1.2 Sequential rationality

Sequential rationality is also a necessary assumption for the main theorem because a strategy profile might satisfy the harm principle alone and yield a Pareto inferior outcome with respect to the initial reference point. The following 2×2 game provides a simple example.

	L	R
L	$2, 2$	$0, 3$
R	$1, 0$	$4, 4$

Suppose that the reference point is (2,2), Column moves first, Row moves second, and this order strictly alternates. Consider the strategy profile in which Column moves from (2,2) to (0,3) where Row stays, making (0,3) the updated reference point. Row's choice of L satisfies the harm principle since it does not harm Column player. Next, Column moves back to (2,2) and Row moves to (1,0) where first Column stays and then Row stays, making (1,0) the outcome. Column's decision to stay at (1,0) satisfies the harm principle since it does not harm Row player with respect to the updated reference point (0,3). Anticipating this and if the players were sequentially rational, Column

[8] We will discuss later in section 5 how one could obtain the outcome (4,3) as a nonmyopic equilibrium outcome starting from the initial reference point (1,4), using Brams's (1994) "two-sidedness convention".

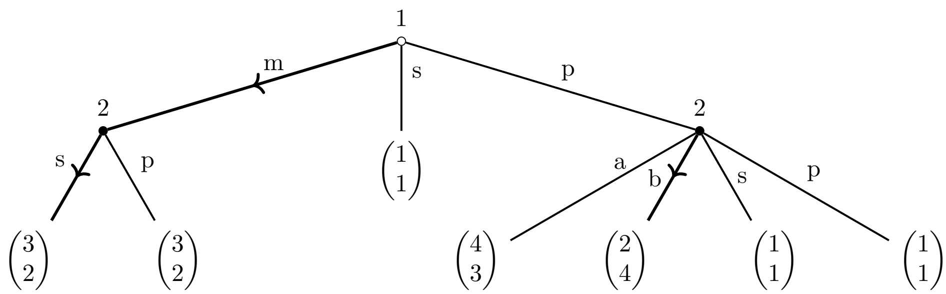

Figure 5: An extensive form game in which the MHE outcome from reference point (1,1) is (3,2), which is Pareto dominated.

would *not* stay at (1,0). But in the absence of the assumption of rationality, the aforementioned moves cannot be ruled out and it results in an outcome, (1,0), that is strictly Pareto dominated by (2,2).

4.1.3 Normal form structure

The normal form structure of game G is also necessary for the main theorem. We now assume both the MHP and sequential rationality and illustrate this with the extensive form game given in Figure 5.

Suppose that the initial reference point pay-off profile is (1,1) and player 1 moves first. There is a unique MHE in this game and it is Pareto dominated. To see this, notice that the best response of player 1 is to choose m, moving to (3,2) because if player 1 chooses to pass (i.e., p), then the best response of player 2 would be to move to (2,4), which would satisfy the MHP with respect to (1,1). Thus, player 1 moves to (3,2), where player 2 stays (or passes). The MHE outcome (3,2) coincides with the SPE outcome in this game. Although (3,2) is a Pareto improvement over the reference point (1,1), it is Pareto dominated by (4,3).

One reason why the MHP and sequential rationality of players do not immediately imply Pareto optimality is that the MHP puts a mild constraint on the behaviour of players. It restricts players from causing harm to others relative to the reference point, but beyond that the MHP does not require players to maximize the pay-off of others.

4.1.4 Unilateral termination

In this section, we consider the modification of our model where each player has the opportunity to unilaterally terminate the game. In that case, our results would remain valid as long as the MHP applies to 'termination' decisions as well. Consider the model presented in section 3 with the modifications outlined below, holding everything else fixed. For every i and every non-terminal x, let $A_i(x) \cup \{t\}$ be player i's set of available actions at node x. If a player plays action t at x, then the terminal node is reached and the outcome is defined as a where $a = S(x)$. In section 3.1.3, drop the terminal node condition (i) where the game terminates if two players choose to stay at a state. Add the following line to Definition 1. Action $\sigma_i(x) = t$ is said to satisfy Mill's harm principle (MHP) at x if for every $j \neq i$, $u_j(a) \geqslant u_j(b)$. Under this modification of our model, the MHE associated with an initial node a_0 may differ from the MHE under the original setup. However, all three of our theorems would remain valid for the analogous reasons to the ones used in the proofs of respective theorems.

One could also consider the following modification to our model in section 3. Suppose that the game terminates if m players ($n \geqslant m > 2$) choose to stay at a state instead of two players as is assumed in condition (i) in section 3.1.3. This modification would not affect the application of the main arguments in the proofs of the three theorems. Thus, the theorems would remain valid in under this modification too.

4.1.5 Preferences

Strict vs weak preferences

One might wonder what happens to the Millian harm equilibria when there are indifferences between the outcomes in G. In that case, Lemma 1 would remain valid, though Lemma 2 would no longer hold. This is because subgame perfect equilibrium outcomes in Γ need not be unique, which implies that MHE outcomes need not be unique either. For analogous reasons as in the proof of the main theorem we can conclude that every Pareto optimal profile must be an MHE outcome, and an MHE outcome cannot be strictly Pareto dominated. Moreover, for every initial reference point and every I there would always be an MHE that is Pareto optimal.

Preference changes

As is standard in game theory, preferences are taken as given in our model. Here, we briefly discuss how our results might change if players self-report their preferences. Under this assumption, there would certainly be games where it is beneficial to misrepresent one's own preferences. However, all of our results would remain valid with respect to the reported preferences, though not necessarily with respect to the private preferences, as long as these reported preferences are strict.

4.1.6 Repetitive moves

We next consider a modification of the basic model where a player can repeat the same move at the same state more than once.

Recall the definition of the set of available actions at a node x: $A_i(x) = \{a_i \in A_i | \ X'(a_i, x) = \varnothing\} \cup \{p\}$. Now suppose that the cardinality of $X'(a_i, x)$ is a finite number. Put differently, if player i chooses action a_i at some node $x \in X$, then player i can choose a_i at state $S(x)$ finitely many times.

Then, the three theorems would still hold because the associated extensive form game $\Gamma(a_0, I)$ would still have a finite horizon, and this is one of the key assumptions to guarantee the existence of an MHE. None of these theorems immediately extend to the case in which a move can be repeated infinitely many times. In section 6, we pose an open problem in that case.

4.1.7 The player function

While the three theorems hold for any player function I, the associated MHE would potentially be different for different I (see, e.g., the example in subsection 5.2). However, this does not change the conclusion of, e.g., the main theorem that any such MHE is Pareto efficient.

In section 3, we put a restriction on player function I that in every path of play each player has more or less equal number of nodes at which they are active. We next show that the main theorem would not hold in general if we let the player function be arbitrary. Let I' be a player function such that for every non-terminal node $x \in X$, $I'(x) = 1$. Clearly, the main theorem would not hold if the player function were I'. To see this, consider the PD with the initial reference point DD. Then, player 1 cannot by themself move to CC. Thus, player 1 would stay at DD, which is Pareto dominated.

A different way to interpret the player function I is that it may be chosen by Nature in the beginning of the game according to the stochastic process described below. Fix a game $G = (A, u)$ and action profile $a_0 \in A$. Let $q \in \Delta N$ be a probability distribution over the set of players N such that for every player i, $q(i) > 0$ and $\sum_i q(i) = 1$. For a given probability distribution $q \in \Delta N$, we define an associated extensive form game with perfect information and Nature move denoted by $\Gamma'(a_0, q) = (N, X', I', u', S', H')$. At the root, x'_0, of $\Gamma'(a_0, q)$, Nature randomly chooses a player function I, and then players play the game $\Gamma(a_0, I) = (N, X, I, u, S, H)$.

Let $I' : X' \to N$ denote the player function in $\Gamma'(a_0, q)$, where $I'(x)$ is the active player at node $x \in X'$. At x'_0, Nature chooses player function $I : X \to N$, where $X \subset X'$, according to the following process. The probability player j is the active player at a non-terminal node x_m is given by the conditional probability $P(j|x_{m-1})$, where x_{m-1} is the immediate predecessor of x_m, which is defined as follows. Let $\underline{m} = floor(\frac{m}{n})$. If $x_{\underline{m}n+1} = m$, then $P(j|x_{m-1}) = q(j)$. If $x_{\underline{m}n+1} < m$, then

$$P(j|x_{m-1}) = \begin{cases} 0, \text{ if } j = I'(x_{\underline{m}n+1}), \text{ or } j = I'(x_{\underline{m}n+2}), ..., \text{ or } j = I'(x_{m-1}) \\ \frac{q(j)}{\sum_i q(i) - \sum_{i=\underline{m}n+1}^{m-1} q(I(x_i))}, \text{ else.} \end{cases}$$

Notice that the player function I defined as above satisfies the restriction we put in section 3. Thus, irrespective of the realisation of Nature's randomisation, the three theorems would remain valid in $\Gamma(a_0, I)$.

5 Comparison with Theory of Moves and illustrations

In this section, we give several illustrations of MHEs in games. Since we have based our development of the harm principle on Brams's (1994) game theoretical framework, we focus on how our MHE differs from his alternative solution concept of nonmyopic equilibrium—as well, of course, as the differences between MHE and the Nash equilibrium concept.

5.1 Relation to Nash equilibrium and nonmyopic equilibrium

It is well-known that Pareto optimality and the Nash equilibrium are logically distinct concepts in the sense that neither concept is a refinement of the other. As we show in the main theorem the MHEs coincide with Pareto optimal profiles. Thus, there is no logical relationship between the set of MHEs and the set of Nash equilibria.

The nonmyopic equilibrium and MHE concepts in general give different predictions. This can be seen in the relation that each solution concept has with the generation of Pareto optimal outcomes. An MHE is always Pareto optimal and every Pareto optimal outcome can be supported as an MHE with a suitable choice of reference point. In contrast, Brams and Ismail (2022) show that while there is always a nonmyopic equilibrium that is Pareto optimal, a) not all nonmyopic equilibria are Pareto optimal and b) not all Pareto optimal profiles are nonmyopic equilibria.

In Brams's (1994, p. 217) categorization of 57 2×2 games, a) is unusual. It only occurs in the PD where the nonmyopic equilibria are (3,3) and (2,2). In comparison, while (3,3) is an MHE, (2,2) is not. The main reason why (2,2) is not supported as an MHE when (2,2) is the reference point is the harm principle. This is because the harm principle prevents a player from making the outcome (1,4) or (4,1). It is also instructive to sketch why (3,3) is a nonmyopic equilibrium because it depends on a *two-sidedness convention*.

If the initial reference point is (3,3), this is not necessary as it is clear that an initial deviation will trigger a further deviation, resulting in no one benefiting. If, however, the initial reference point is (4,1) or (1,4), the corresponding nonmyopic equilibrium is (3,3) only because of a *two-sidedness convention* (Brams, 1994, p. 28). The convention means that with a given reference point, if one player, by moving, can induce a better outcome than by staying, but the other player by moving can induce a Pareto-superior outcome, then the other player's move takes precedence. Thus, at the initial reference point (1,4), if Row moves first to (2,2), then Column would stay at (2,2), instead of moving to (4,1). Alternatively, if Column moves first from (1,4) to (3,3), then Row would stay at (3,3) rather than move to (4,1) because Column would not stay at (4,1). Since (3,3) Pareto dominates (2,2), the *two-sidedness convention* applies; hence the nonmyopic equilibrium from (1,4) is (3,3). For similar reasons, the nonmyopic equilibrium from (4,1) is also (3,3).

Returning to the example in subsection 4.1.1, we can again see the part played by the *two-sidedness convention*. The game in subsection 4.1.1 has a unique nonmyopic equilibrium, (4,3).[9] While it is clear why the nonmyopic equilibrium from the initial reference point of (4,3) is (4,3), it is not immediately obvious why, from the initial reference point of (1,4), the nonmyopic equilibrium is also (4,3). This is because, as discussed in subsection 4.1.1, if Row moves first from (1,4), the outcome would be (3,2). If Column moves first from (1,4), then the outcome would be (4,3). Thus, by the two-sidedness convention, Column moves first, and hence the nonmyopic equilibrium is (4,3).

b) is not so unusual and seems likely to occur more frequently in larger games. For example, consider the 2×3 game below (Brams and Ismail, 2022, p. 356).

	C	D	E
A	$6,1$	$4,4$	$1,6$
B	$5,2$	$3,3$	$2,5$

[9]The game in subsection 4.1.2 also has a unique nonmyopic equilibrium, (4,4), but it does not require the two-sidedness convention.

The unique nonmyopic equilibrium in this game is (2,5), but this is only one of five Pareto optimal outcomes.

In addition to the nonmyopic equilibrium, Brams (1994) explores the concept of power and its variations in games, including threat power, order power, and moving power. Here, we discuss threat power, as it allows for implementing the outcome (3,3) in the PD. Brams (1994) formalizes two types of threats—compellent and deterrent—following a distinction made by Schelling (1966). Essentially, a player's compellent threat involves committing to a strategy to induce the opponent to play a best response to that strategy. In contrast, a deterrent threat occurs when a player threatens to deviate from an agreed-upon strategy if the opponent fails to play as agreed. For example, if a player, say Row, has a deterrent threat in the PD, it would be rational for each player to choose C because if Column deviates to D, then Row would implement the threat by choosing D. Here, Row essentially deters Column from choosing D.

5.2 A three-person illustrative example

We next illustrate the Millian harm equilibria in a three-person game presented in Figure 6. Throughout this example, we assume that the initial reference point is (A,D,E).

Assume that Row moves first, Column second, Matrix third, and this order strictly alternates. Figure 7 illustrates part of the game tree where the arrows show the on-path moves of the MHE, which can be described as follows. Row moves to (8, 8, 4), and Column stays at (8, 8, 4), which makes it the reference point. A best response of Matrix is to stay at (8, 8, 4), making it the outcome of the MHE from (3, 1, 2). Notice that Matrix can move to (4, 4, 5), but cannot stay in matrix F because this would violate the MHP with respect to the reference point (8, 8, 4).

Now, assume that Matrix moves first, Column second, Row third, and this order strictly alternates. The initial reference point is (3, 1, 2) as before. We explain the on-path actions of the MHE as follows. Matrix moves to (7, 5,

E	C	D
A	$1, 6, 1$	$3, 1, 2$
B	$2, 7, 3$	$8, 8, 4$

F	C	D
A	$5, 2, 6$	$7, 5, 8$
B	$6, 3, 7$	$4, 4, 5$

Figure 6: Millian harm equilibria in a three-person illustrative game

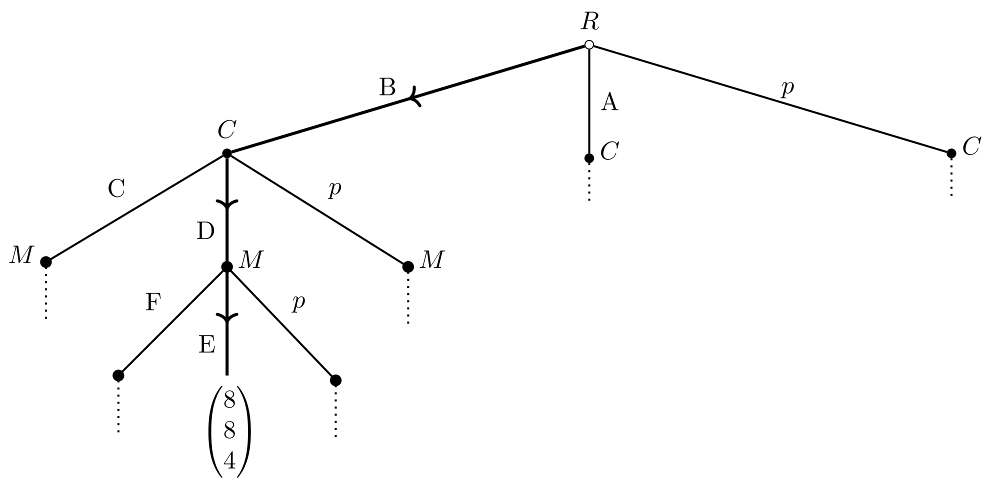

Figure 7: Part of the game tree of Γ presented in Figure 6 where the arrows illustrate the MHE path. Row moves first, Column second, and Matrix third. (The full game tree is not shown due to space constraints.)

8) by playing F, where both Column and then Row stay. The reason why it is a best response for Column to stay at (7, 5, 8) is that (i) Column receives their highest pay-off in matrix F, (ii) Matrix player would prefer to stay at any outcome in matrix F rather than moving to matrix E, and (iii) every outcome in matrix F satisfies the MHP with respect to the reference point (3, 1, 2). For analogous reasons, it is also a best response for Row to stay at (7, 5, 8). Thus, (7, 5, 8) is the outcome of the MHE from the initial reference point (3, 1, 2).

At the outset, it looks like Row and Column should be able to implement their most preferred outcome (8, 8, 4) in the game. However, as shown above this is not possible if Matrix is the first-mover at the initial reference point. This three-person example illustrates that the player function I can affect the

MHE associated with any reference point, but I does not affect the conclusion that the MHEs are Pareto optimal.

6 Conclusion

In this paper we supply a new answer to an old question in political philosophy. We show that in n-person games, when people are free to select the action that best satisfies their preferences subject to a version of Mill's harm principle, the result is a Pareto efficient outcome. This is a surprising sense of 'order' that the exercise of freedom produces—surprising, that is, relative to the formal insights provided by economics on this question. For example, game theory does not expect, in general, this exercise of freedom to produce Pareto efficient outcomes, as is well illustrated by the standard analysis of the famous prisoners' dilemma game.

The difference in our analysis of prisoners' dilemma types of situations is the inclusion of a version of Mill's harm principle as a constraint on the choice of actions in games. This can be contrasted with the other explanation from economics (e.g. from Coase and Ostrom) over how the 'disorder' of Pareto inefficiency might be avoided when the players step outside the game and agree, through bargaining or the reliance on social capital resources, on joint actions which secure a Pareto improvement. These derivations of 'order' rely on institutional props or supplements, whereas Mill's harm principle has, since Mill's time at least, been regarded as constitutive of what liberty means in a liberal society. Liberty on this account has never been a licence for a free-for-all, it has always been constrained by the harm principle. Of course, there are always disputes over what counts as harm in liberal societies. But, what we show is that when preferences are given, harm is understood as a pay-off reduction and people are constrained by our version of MHP, there can be disputes over what is a harm and Pareto efficiency will still result. In this sense, our demonstration of Pareto efficiency is intrinsic to what freedom means in the liberal tradition of political theory and is therefore more powerful than Coase's and Ostrom's.

It may be useful to put this result into perspective by returning to Gibbard's example of him and Parker painting their walls. Both can cause each other a harm, in our sense of a pay-off reduction, if they paint their house the colour they like most. Our version of Mill's harm principle constrains individual actions so as to avoid this happening and it produces instead a Pareto efficient outcome. This is, of course, a contrived example, not least because the 'harms' are not exactly publicly defensible as they arise from a kind of malevolence. Hobbes's harms in the state of nature that each experiences from the threat of others who arm themselves, and which create an equivalent of the prisoners' dilemma, are much less controversial. Nevertheless, this is what makes the Gibbard example instructive because it too generates a Prisoners' Dilemma and our result holds more generally for all such interactions.

Gibbard's example is also useful in locating our contribution in the literature. In the game theoretic framework, we take a person's preferences as given, if they are 'weird' in a way that seeing a happy neighbour causes discomfort (that registers with a lower 'utility' number) is 'weird', then so be it. De gustibus est non disputandum in game theory. We do not require, as a result, an agreement over what substantively causes harm to a person. There is no need for Parker to agree with Gibbard that Parker's happiness actually causes Gibbard harm. In other words, there can be perspectival disagreement in the sense of Chung (2019) and Chung and Kogelmann (2020). Their social choice approach is different to ours. We ask what will happen when people act rationally but are constrained by a harm principle, whereas they ask whether there is a social choice function when there are perspectival disagreements that satisfy a liberal conception of rights. For this purpose, they assume that liberal rights define an individual's private space where individual decisions hold sway. Even when there is an agreement over what is in this private domain, Chung (2019, Theorem 4) shows that perspectival disagreements over the outcomes in that private domain are enough to produce an impossibility result for a social choice function in these circumstances that respects the liberal's private sphere and satisfies (strong) Pareto efficiency. Although this is a contrasting result to the one we produce, from our non-cooperative perspective, it is not so surprising. For example, if house painting is in the private domain when Parker and Gibbard enjoy rights, they will decide to paint their house the Pareto inferior (yellow, white). This makes our result all the more arresting and the difference between it and theirs would appear to be not so much due to perspectival differences as the introduction of our version of Mill's harm principle as compared with their conception of liberal rights that establish an individual's private space.

For this reason, our version of Mill's harm principle (and the comparison with the liberal rights in the social choice approach) is worth further comment. To begin with our harm principle, our definition of a harm as a pay-off reduction seems attractive because it is quite general. However, our version of the harm principle assumes that the assessment of any person's pay-off reduction is made according to that person's preferences; and this could be controversial when there are perspectival disagreements. Why should Parker's judgement of harm to Gibbard, for instance, be guided by Gibbard's perspective on the matter if Parker does not agree with that perspective?

We offer two defences to this aspect of our argument. First, Parker does not have to accept Gibbard's view as the correct view, but, pragmatically, in a game theoretic context, Parker does have to anticipate how Gibbard will act; and Gibbard will plausibly be guided in this by his own view of the situation, including regarding how harms to him arise when he is applying the harm principle. This is a bit like the common knowledge of rationality assumption in standard game theory, whereby each recognises that each will act to satisfy their preferences as they see them, and this is so, even though a particular person's preferences may make no sense to another person. Second, we require that Gibbard does the same when assessing harms to Parker: that is, he accepts Parker's perspective as the relevant one for how harms arise for Parker. This is for the same reason that Parker does but it reveals an additional desirable feature of our harm principle. When there are perspectival differences, in effect, our principle affords equal standing to each perspective. Each person's perspective guides their own assessment of harm to themselves. We do not treat any perspective differently. No perspective is elevated above another in the definition of a harm. This is an attractive feature in any genuinely plural (and liberal) societies where perspectival differences are to be expected. Some will be ineluctable in liberal societies and such differences should therefore command equal respect.

Even if these arguments for our harm principle are accepted, in practice, our harm principle is likely to converge with a rights approach in some respects. This is because in a non-ideal world, we cannot take people's preferences as given in the way that game theory does. Indeed, if people know that people are guided by our version of the harm principle, then they will have reason, on occasion, strategically to pretend harms/discomforts where none exist (i.e. the principle is not strategy proof). For this reason, it is not practical always to allow for such personal definitions of what causes discomfort outside the ideal world of game theory and expect that Pareto efficiency will still result from the harm principle with respect to whatever are people's real as opposed to their strategic preferences. Instead, in a non-ideal world, we find that in so far as the prospect of causing harm constrains action, it is often because harms are given a public definition and legal force through the rights that people can expect to enjoy.

Our reliance on rights in a non-ideal world appears to move us closer to the social choice approach, with its similar reliance, and this may seem to favour a Chung-like impossibility result over our contrary Pareto efficiency result. However, matters are not so simple because rights arise for us in a non-ideal world and rights do not function exactly in the manner of the social choice approach in such worlds. In this respect, the social choice approach, whereby one asks whether a social choice function is possible in a world where people have liberal rights in their sense of private spaces, is an ideal-type question. If the comparison between their approach and ours is to be made instead on the non-ideal terrain, then the concept of rights needs adjustment from that found in the social choice approach. There are two respects in which this appears potentially important.

First, rights in practice do not just establish private spheres. They may do this but they also put constraints on actions more generally and they thereby also create expectations regarding other people's behaviours. This means that rights also affect the character of the feasible outcomes for a social choice problem. Second, in so far as rights encode some publicly determined definition of what a harm is, then the existence of such rights casts a questioning shadow over whether perspectival disagreements of the Chung (2019) type should really be a datum in the formation of public policy (and hence the non-ideal version of the social choice problem). The point is this. The public articulation of what is a harm through the creation of rights has become necessary in our view because private definitions of a harm collide and they cannot be treated as reliable indicators of people's real preferences in some circumstances. The purpose of the rights intervention is, in other words, to resolve those differences in a particular way. Thus, one might argue that public policy has already dealt with these differences and so why should they feature in a public policy discussion of what to do in a world where such rights exist? Or to express this slightly differently, if rights establish that painting your house is in a person's private domain, why should Gibbard think that his possible irritation with Parker's painting decision should enter into the post rights social choice problem? Of course, there can be a dispute over whether this is the correct assignment of rights. But, if the social choice approach takes the rights as given, then shouldn't Gibbard and Parker also? In short, model consistency in a non-ideal world would seem to demand that they should, so to speak, move on from this particular dispute.

We leave these as open questions for the social choice approach in a non-ideal world. For our harm approach, the challenge of a non-ideal world is that articulated or revealed preferences cannot plausibly be relied upon to define what is a harm in all circumstances. There will be cases where 'de gustibus est disputandum', not least because aspects of people's preferences are ill-formed and uncertain, especially in a fast changing world, and not just because they may be subject to strategic manipulation. It is in these cases that public policy intervention is required through Parliament and legal adjudication. In short, the gap between the ideal world and the non-ideal that seems likely to arise sets a particular public policy agenda in our framework.

To put this last point slightly differently and more sharply, it is often argued in economics textbooks, for example, that the policy agenda in our non-ideal world is set by the existence of prisoners' dilemma type interactions. They are a brute fact about our social world and governments constrain our actions in the market by devising policies, like taxes, that internalise the externalities that create the prisoners' dilemmas. From the perspective of this paper, though, it is not the occurrence of prisoners' dilemmas in social and economic life that should occasion this retreat from liberty. Rather, it is a retreat from the liberal conception of liberty that is responsible for making prisoners' dilemma interactions problematic. Such retreats may occur either because we have lost sight of the harm principle in the constitution of liberty or because we have failed to define adequately through legislation and legal judgement what constitutes harm. The simple point from a policy perspective is that it is important to get the source of the problem right.

In short, our result is important both for a fundamental question in political theory regarding how liberty needs to be restrained for social order to result and for guiding policy discussion in a non-ideal world.

Appendix

Proof of the main theorem

Proof of the first part. We first show that if an initial reference point a_0 is Pareto optimal, then for every player function I, the associated MHE outcome is a_0 in game $\Gamma(a_0, I)$, proving the first part of the theorem.

To reach a contradiction, suppose that a_0 is not the MHE outcome, and the outcome of an MHE σ^* from a_0 is given by some $a' \neq a_0$. We know that the MHE outcome from each initial reference point is unique by Lemma 2. Because $a' \neq a_0$ and a' is the outcome of σ^*, it must be that $R(\sigma^*)\backslash\{a_0\}$ is non-empty. Too see this, suppose that $R(\sigma^*) = \{a_0\}$. It implies that $a' = a_0$, which is a contradiction. Thus, there exists at least one player who stayed along the path of play of σ^*, that is, changed a reference point in $R(\sigma^*)$.

For every player j who did *not* stay along the path of play of σ^*, it must be that $u_j(a') > u_j(a_0)$. This is because for every reference point $a \in R(\sigma^*)$, $u_j(a) \geqslant u_j(a_0)$ due to the fact that σ^* satisfies the MHP. For every player i who did stay along the path of play of σ^*, it must be that $u_i(a') > u_i(a_0)$ due to two reasons.

First, player i's pay-off cannot be diminished before i stays because every preceding stay decision must satisfy the MHP. To see this, let $x \in X$ be the node such that player i stays for the first time at $a \in R(\sigma^*)$, and b be the reference point at x. Then, it must be that $u_i(b) > u_i(a_0)$ due to the MHP, that is, other players could not have stayed and harmed player i.

Second, it would not be optimal for player i to stay by playing a_i unless $u_i(a') > u_i(a_0)$. That is, if i stays at a where $S(x) = a$, then i eventually must benefit from this action due to sequential rationality and the MHP. If a_i is optimal at x, i.e., $\sigma_i^*(x) = a_i$, then $u_i(a') \geqslant u_i(b)$ because otherwise player i would *not* stay at a, changing the reference point b. Notice that i can move to another state or play p, in which case the minimum pay-off i would receive is $u_i(b)$. This is because (i) if someone else stays at a state different than b, then i cannot be harmed, and (ii) if everyone passes, then the outcome would be b. But we also have that $u_i(b) > u_i(a_0)$. Therefore, $u_i(a') > u_i(a_0)$.

As a result, it implies that for every player i', the inequality $u_{i'}(a') > u_{i'}(a_0)$ is satisfied, irrespective of whether i' stays or not along the path of play of σ^*. This contradicts to our supposition that a_0 is Pareto optimal. Therefore, a_0 must be the outcome of σ^*. □

Proof of the second part. Next, we show that for every initial reference point $a_0 \in A$, and every player function I, the associated MHE outcome in game $\Gamma(a_0, I)$ is Pareto optimal. In the first part of the proof we already showed that if a_0 is Pareto optimal, the associated MHE outcome is Pareto optimal. It is left to show that for an initial reference point a_0 that is not Pareto optimal, the MHE associated with a_0 must be Pareto optimal. Let σ be an MHE from a_0, which exists by Lemma 1.

To reach a contradiction, suppose that the outcome of σ is b, and b is

Pareto dominated by some action profile $a \neq b$. We obtain a contradiction in two main steps.

Step 1: Given σ, we construct an MHE σ^* from a_0 such that there exists a player i who chooses b_i at some node y in $[\sigma^*]$ and makes the reference point, $b \in R(\sigma^*)$, at y the outcome. Since b is the outcome of σ, the outcome of any MHE from a_0 must be b by Lemma 2.

Consider path of the play, $[\sigma]$, of σ excluding the terminal node $\overline{[\sigma]}$. Note that whether the active player at the penultimate node in $[\sigma]$ stayed or passed, b must have been the reference point at some point during the path of the play of σ. Let i' be the player who makes b the reference point for the first time at some node $y' \in [\sigma^*]$. Let $y \in [\sigma^*]$ be an immediate successor of y' such that $\sigma_{i'}(y') \to y$ and $i = I(y)$ be the player who moves at y. Since the outcome of σ is b, it must be optimal (i.e., a best response under the constraint of the MHP) for player i to stay at y and make $S(y) = b$ the outcome. Player i's stay action, b_i, is available at node y because b is the reference point for the first time at y'. In addition, choosing b_i clearly satisfies the MHP because b is already the reference point at y. We then construct σ^* such that $\sigma_i^*(y) = b_i$. If $\sigma_i(y) = b_i$, then define $\sigma^* = \sigma$. If $\sigma_i(y) \neq b_i$, then for every player m and every non-terminal node $\hat{x} \in X \backslash \{y\}$ define $\sigma_m^*(\hat{x}) = \sigma_m(\hat{x})$. As desired, we have constructed an MHE σ^* from a_0 such that there is player i who makes the reference point, $b \in R(\sigma^*)$, at y the outcome in the first opportunity.

Step 2: We next show that σ^* and hence σ cannot actually be an MHE because $\sigma_i^*(y)$ cannot be player i's optimal choice at y. In other words, player i has a unilateral profitable deviation from σ^* and this deviations satisfies the MHP. Notice that if (i) there exists a path of play from b to a along which the MHP is satisfied, then $\sigma_i^*(y)$ cannot be optimal because (ii) for every player m (including i) $u_m(a) > u_m(b)$, and (iii) no other player can stay at an action profile which harms player i along the path because b is the reference point.

We first show (i). Let $b' \in A$ and $a' \in A$ be two action profiles. We first show that for any player function I, there is always a path of play between b' and a'. Let $[b', a']$ be the path of play from b' to a' with the following property. For every node y' in this path of play, the active player at y', $i' = I(y')$, chooses

$a_{i'}$ except when $a_{i'} = b_{i'}$, in which case player i' chooses to pass, p. Notice that if the players follow this path, then the play would reach to a' from b' in at most n moves. Note that no player stays along the path of play, so no action in the constructed path of play violates the MHP.

Now, let $[b, a]$ be the path of play from b to a constructed as above. We next show that $[b, a] \cap [\sigma^*] = \varnothing$, i.e., the constructed path of play does not overlap with the path of play of σ^*. In other words, we make sure that σ^* does not prescribe players to choose actions at some nodes in $[\sigma^*]$ such that these actions then make the constructed path, $[b, a]$, infeasible due to the history of play.

Let $x \in [b, a]$ be a node such that $S(x) = a$. Then, it must be that $x \notin [\sigma^*]$ because if $x \in [\sigma^*]$, then it would be a unilateral profitable deviation for the active player at x to stay at a and make a the reference point. This is because $u_m(a) > u_m(b)$ and staying at a satisfies the MHP. To see why staying at a satisfies the MHP, suppose (to reach a contradiction) that there exists a player $\hat{i}$ such that $u_{\hat{i}}(a) < u_{\hat{i}}(\hat{a})$, where $\hat{a}$ is the reference point at x. But we know that the outcome of σ^* is b and that $u_{\hat{i}}(b) < u_{\hat{i}}(a)$, which implies that $u_{\hat{i}}(b) < u_{\hat{i}}(\hat{a})$. Thus, either player $\hat{i}$ harms themself by staying at b, or someone else harms $\hat{i}$. It implies that either the sequential rationality of $\hat{i}$ is violated or the MHP is violated, a contradiction. As a result, if there exists a node $x \in [\sigma^*]$ such that $S(x) = a$, then the active player would stay at a, making it the reference point. But if a is the reference point, then b cannot be the outcome of the MHE σ^* because every player is strictly better off at a. This leads to a contradiction to our supposition that the outcome of σ^* is b. This establishes that $x \notin [\sigma^*]$.

Let $y \in [b, a]$ be an immediate predecessor of x. Then, it must be that $y \notin [\sigma^*]$ because if $y \in [\sigma^*]$, then the player at y would have a unilateral profitable deviation by moving to a, anticipating that the next player would stay at a as shown in the previous paragraph. Note that this deviation would not violate the MHP by construction of $[b, a]$. Next, let $y' \in [b, a]$ be a (not necessarily immediate) predecessor of y. By backward induction, notice that $y' \notin [\sigma^*]$ because if $y' \in [\sigma^*]$, then the player $I(y')$ would have a unilateral profitable deviation from $\sigma^*_{I(y')}(y')$ by playing an action that leads to a node $y'' \in [b, a]$, where y'' is an immediate successor of y'. Thus, the paths of play $[b, a]$ and $[\sigma^*]$ have an empty intersection. As a result, statement (i) holds: there is a path of play from b to a such that no player violates the MHP, given the path of play of σ^*.

Statement (ii) holds by our supposition that b is Pareto dominated. Statement (iii) holds by definition of the MHP: player i has a unilateral profitable deviation at b by playing a_i unless $b_i = a_i$, in which case i has a unilateral profitable deviation by playing p, entering the path from b to a as constructed above. By the MHP, no other player in the path can reduce i's pay-off. In addition, no player stays in the path from b to a. Thus, player i will eventually receive a strictly greater pay-off by deviating to the constructed path because $u_i(a) > u_i(b)$.

As a result, if the outcome b of MHE σ^* is Pareto dominated by some a, then player i who stays at b and make b the outcome would have a unilateral profitable deviation from σ^*. This contradicts to our supposition that the MHE outcome from a_0 is b. As desired, this implies that the MHE outcome from any a_0 must be Pareto optimal. □

Statements and Declarations

Competing Interests The authors have no financial or non-financial interests to disclose.

Funding The authors conducted this study without external funding.

Data availability No data were used in the preparation of this manuscript.

References

Aumann, R. J. (1976). Agreeing to disagree. *The Annals of Statistics 4*(6), 1236–1239.

Baron, D. P. (1996). A dynamic theory of collective goods programs. *American Political Science Review 90*(2), 316–330.

Bowen, T. R. and Z. Zahran (2012). On dynamic compromise. *Games and Economic Behavior 76*(2), 391–419.

Brams, S. (1994). *Theory of Moves*. Cambridge, UK: Cambridge University Press.

Brams, S. J. and M. S. Ismail (2022). Every normal-form game has a Pareto-optimal nonmyopic equilibrium. *Theory and Decision*, 349–362.

Brams, S. J. and D. Wittman (1981). Nonmyopic equilibria in 2 × 2 games. *Conflict Management and Peace Science 6*(1), 39–62.

Chung, H. (2015). Hobbes's State of Nature: A Modern Bayesian Game-Theoretic Analysis. *Journal of the American Philosophical Association 1*(3), 485–508.

Chung, H. (2019). The impossibility of liberal rights in a diverse world. *Economics & Philosophy 35*(1), 1–27.

Chung, H. and J. Duggan (2020). A Formal Theory of Democratic Deliberation. *American Political Science Review 114*(1), 14–35.

Chung, H. and B. Kogelmann (2020). Diversity and rights: a social choice-theoretic analysis of the possibility of public reason. *Synthese 197*, 839–865.

Coase, R. H. (1960). The problem of social cost. *The Journal of Law and Economics 3*, 1–44.

Gibbard, A. (1974). A Pareto-consistent libertarian claim. *Journal of Economic Theory 7*(4), 388–410.

Hobbes, T. (1651). *Leviathan Or The Matter, Forme, Amp Power Of A Common-wealth Ecclesiasticall And Civill*. London.

Kalandrakis, A. (2004). A three-player dynamic majoritarian bargaining game. *Journal of Economic Theory 116*(2), 294–14.

Kilgour, D. M. (1984). Equilibria for far-sighted players. *Theory and Decision 16*(2), 135–157.

Lewis, D. (1969). *Convention: A Philosophical Study.* Harvard University Press.

Lichtenberg, J. (2010). Negative duties, positive duties, and the "new harms". *Ethics 120*(3), 557–578.

Mariotti, M. and R. Veneziani (2009). 'Non-interference' implies equality. *Social Choice and Welfare 32*(1), 123–128.

Mariotti, M. and R. Veneziani (2013). On the impossibility of complete non-interference in Paretian social judgements. *Journal of Economic Theory 148*(4), 1689–1699.

Mariotti, M. and R. Veneziani (2020). The Liberal Ethics of Non-Interference. *British Journal of Political Science 50*(2), 567–584.

Mill, J. S. (1859). *On Liberty.* John W. Parker and Son: London.

Nash, J. (1951). Non-Cooperative Games. *The Annals of Mathematics 54*(2), 286–295.

Osborne, M. J. and A. Rubinstein (1994). *A Course in Game Theory.* MIT Press.

Ostrom, E. (2010). Beyond Markets and States: Polycentric Governance of Complex Economic Systems. *American Economic Review 100*(3), 641–672.

Schelling, T. (1966). *Arms and Influence.* New Haven, CT: Yale University Press.

Selten, R. (1965). Spieltheoretische Behandlung eines Oligopolmodells mit Nachfrageträgheit. *Zeitschrift für die gesamte Staatswissenschaft*.

Sen, A. (1970). The impossibility of a Paretian liberal. *Journal of Political Economy 78*(1), 152–157.