Quantifying Inefficiency*

Yannai A. Gonczarowski Ella Segev

December 16, 2024

Abstract

We axiomatically define a *cardinal* social inefficiency function, which, given a set of alternatives and individuals' vNM preferences over the alternatives, assigns a unique number—the *social inefficiency*—to each alternative. These numbers—and not only their order—are uniquely defined by our axioms despite no exogenously given interpersonal comparison, outside option, or disagreement point. We interpret these numbers as per capita losses in endogenously normalized utility. We apply our social inefficiency function to a setting in which interpersonal comparison is notoriously hard to justify—object allocation without money—leveraging techniques from computer science to prove an approximate-efficiency result for the Random Serial Dictatorship mechanism.

*Gonczarowski: Department of Economics and Department of Computer Science, Harvard University (*e-mail*: yannai@gonch.name); Segev: The Hebrew University Business School and the Federmann Center for the Study of Rationality, The Hebrew University of Jerusalem (*e-mail*: ella.segev@mail.huji.ac.il). The authors thank Cassidy Shubatt for excellent research assistance, and thank Jerry Green, Sergiu Hart, Eric Maskin, Ran Shorrer, Assaf Romm, Tomasz Strzalecki, and Rakesh Vohra for insightful comments and discussions. The authors gratefully acknowledge research support by the following sources. Gonczarowski: National Science Foundation (NSF-BSF grant No. 2343922), Harvard FAS Inequality in America Initiative, and Harvard FAS Dean's Competitive Fund for Promising Scholarship. Segev: Israel Science Foundation (ISF grant No. 301/23).

Contents

1 Introduction **1**

1.1 Further Related Literature 10

2 Model **14**

3 Axioms **15**

4 Characterization **18**

4.1 Explicit Construction 18

4.2 Logical Independence of Our Axioms 21

4.3 Proofs 23

4.3.1 General Properties 23

4.3.2 Characterization of Alternatives with Infinite Inefficiency . . . 25

4.3.3 Structural Results 26

5 Application to Object Allocation **28**

5.1 Model 29

5.2 Theoretical Social Inefficiency Guarantees 30

5.2.1 Preliminaries 31

5.2.2 A Lower Bound on the Social Inefficiency of Truthful Ordinal Mechanisms 32

5.2.3 An Upper Bound on the Social Inefficiency of Random Serial Dictatorship 32

5.3 Computation of Social Inefficiency 35

6 Discussion **37**

References **39**

A Proofs Omitted from Section 4 **A.1**

B Proofs Omitted from Section 5 **A.9**

1 Introduction

The question of how to aggregate the ordinal preferences of individual members of a society over a set of possible alternatives into a societal preference order is at the heart of social choice theory, dating back to Arrow (1951) and Harsanyi (1955). Such ordinal aggregation allows us to compare any two alternatives to determine which of the two is more societally efficient (societally preferred), yet does not generally allow us to quantify the *magnitude* to which one alternative is more societally efficient than another. For example, such ordinal aggregation allows us to identify whether a given alternative is societally efficient (societally most preferred); however, in case a given alternative is not societally efficient, such ordinal aggregation does not generally allow us to quantify *how societally inefficient* such an alternative is. In this paper, we fill this lacuna by axiomatically defining not an ordinal but a cardinal measure of the (in)efficiency of any given alternative, which despite being cardinal and comparable across different settings, is based solely on the (ordinal) von Neumann–Morgenstern (vNM) preferences of the individuals.

One of the early major contributions of theoretical computer science to economic theory is in its abundant use of (and techniques for deriving) *inefficiency-bounding theorems*.[1] Such theorems, rather than ascertaining the efficiency of some equilibrium, mechanism, or strategy, ascertain that it is guaranteed to not be far from efficient, in a precise, quantifiable sense. An early celebrated[2] example of an inefficiency-bounding theorem is that the cumulative driving time in every equilibrium of any of a large class of congestion games is no greater than 4/3 times the optimal cumulative driving time (Roughgarden and Tardos, 2002).[3] This inefficiency-bounding theorem allows for further insights beyond the traditional economic approach of proving that equilibria in some congestion games are inefficient (Braess, 1968) and characterizing in which games they are (in)efficient. Another, recent notable example of an inefficiency-bounding theorem is that in a Bayes–Nash equilibrium of a first-price auction with

[1] These are many times referred to as *approximation theorems* within computer science.

[2] See https://www.acm.org/media-center/2012/may/acm-sigact-presents-godel-prize-for-research-that-illuminated-effects-of-selfish-internet-use.

[3] In computer-science nomenclature, this result proves that the "Price of Anarchy" (Koutsoupias and Papadimitriou, 1999) in any such congestion game is at most 4/3. "Price of Anarchy" is the term used by computer scientists for the multiplicative efficiency loss in equilibrium compared to the first-best alternative (in the case of congestion games, due to unpriced externalities). This terminology originates from computer scientists thinking of this fraction as quantifying the efficiency loss "due to selfish behavior," where "selfish behavior" refers to strategic behavior.

independent bidders' values, the expected welfare (value of the bidder who wins the item) is no less than 86% of the expected optimal welfare (highest value) (Jin and Lu, 2023).[4] This recent bound follows a long line of papers that gradually tighten it (Hartline et al., 2014; Hoy et al., 2018), starting with a bound of approximately 63% (Syrgkanis and Tardos, 2013).[5]

Three key properties of the two above cardinal measures of social inefficiency—the ratio between the cumulative driving time and the optimal cumulative driving time, and the ratio between the expected welfare and the expected optimal welfare—contribute to the economic interpretability of these two inefficiency-bounding theorems. We refer to these three properties as *ordinal aggregation* (over the individuals), *intra-context cardinal comparability*, and *inter-context cardinal comparability*.

Ordinal aggregation (over the individuals). Within a congestion game, if one alternative has higher cumulative driving time than another, then the former is less efficient (taking all individuals into account) than the latter; similarly, within an auction setting, if one alternative has lower expected welfare than another, then the former is less efficient (taking all individuals into account) than the latter. In both cases, this is driven by the utility (or disutility) of every individual being exogenously cast in the same interpersonally comparable unit of measurement (minutes of travel, or dollars). Such ordinal aggregation might be more challenging to achieve when the (dis)utility function of each individual merely represents vNM preferences (and is thus defined only up to an affine transformation). This has been the subject of research and debate within social choice theory for decades.[6]

Intra-context cardinal comparability. Within a specific context (specific congestion game or specific auction setting), if two alternatives have cumulative driving times or (for auctions) welfares that differ by a specific amount (whether or not viewed

[4]In computer-science nomenclature, this result proves that the "Price of Anarchy" in this setting is at least 0.86. (The precise fraction is $1 - 1/e^2$, and in the same paper it is also proved that this bound is tight: It cannot be improved upon.)

[5]Inefficiency-bounding theorems indeed allow for fast-paced progress via different sets of authors building upon each other's insights through publication of gradually tighter bounds.

[6]It is not uncommon for papers in the computer-science literature to assume that the individuals' vNM utility functions are exogenously normalized in a way that allows for meaningful interpersonal comparison (e.g., that utility functions are exogenously given such that the sum of each individual's utilities over all possible alternatives is 1, and that numeric utility values can be compared and summed over across individuals). We do not make such an assumption.

as an absolute difference or as a percentage of the optimal cumulative driving time or welfare for the context), then the efficiency loss between these two alternatives is the same as between every other pair of alternatives in the same context whose cumulative driving times or welfares differ by the same amount (e.g., the loss from a welfare of ten dollars to five dollars is the same as from fifteen dollars to ten dollars). Such cardinal comparability might be more challenging to achieve if the (dis)utility functions of the individuals are not cast in the same unit of measurement. Consider, for example, the Nash social-welfare function (Kaneko and Nakamura, 1979), defined as the product of vNM utilities after normalizing the utility functions so that each individual has zero utility for some exogenously given disagrement point. The economic rationale behind maximizing this function has been repeatedly demonstrated across many settings in both economics and computer science (Nash, 1953; Caragiannis et al., 2019), as has been the economic interpretation of the *order* induced by this function (Kaneko and Nakamura, 1979). And yet, we are not aware of any work that economically motivates this function from a *cardinal* point of view. That is, the Nash social welfare function, its logarithm, its tangent, or any other monotone transformation thereof seem to all be equally motivated from an economic perspective (including by all of the above papers). Indeed, given two alternatives, it is unclear what further economic insight is gained by knowing not only that the Nash social welfare of the former alternative is higher that that of the latter, but also that it is numerically higher by precisely ten (or by precisely ten percent of the optimal Nash social welfare for that context).[7] It is similarly unclear whether and in what sense two pairs of alternatives such that within each pair, the Nash social welfares numerically differ by ten, have the same efficiency loss within each pair. This lack of intra-context cardinal comparability might make cardinal guarantees such as achieving a certain fraction of the optimal Nash social welfare challenging to interpret from an economic perspective.[8]

[7]Luce and Raiffa (1957, p. 16) write about indices that represent orders: "one must develop an almost inhuman self-control not to read into these numbers those properties which numbers usually enjoy. For example, one must keep in mind that it is meaningless to add two together or to compare magnitude of differences between them. If they are used as indices [that represent an order over certain alternatives], then the only meaningful numerical property is order."

[8]For example, in the computer-science literature, maximizing Nash social welfare is commonly associated with choosing fair alternatives (Caragiannis et al., 2019), and there are papers within computer science that develop beautiful mathematical techniques for proving guarantees such as achieving some fraction of the optimal Nash social welfare, interpreting such guarantees as approximate fairness guarantees. However, from an economic perspective, it is unclear how one might justify interpreting 50% of the optimal Nash social welfare as "twice as fair" as 25%, or interpreting

Inter-context cardinal comparability. Considering two seperate contexts (two congestion games or two auction settings), figures such as 4/3 of the optimum or 86% of the expected optimum are comparablely interpreted in each context. This is due to each of these two measures of social inefficiency not simply being an aggregate utility using a common unit, but is furthermore (1) measured relative to a meaningful reference point, and (2) viewed as a fraction of the optimal aggregate utility. Indeed, if one accepts that a cumulative driving time of zero is a meaningful reference point, then 4/3 of the optimal cumulative driving time is in a precise sense "4/3 worse" (relative to the reference point) than the optimal cumulative driving time, and one might make the point that two equilibria, in two different games, that are each "4/3 worse" than the optimum for its respective game are comparable in their inefficiency.[9] Similarly, taking no trade (zero welfare) as a reference point or disagreement point in an auction setting, 86% of the expected optimal welfare is in a precise sense "86% as efficient" (relative to the reference point) as the optimal expected welfare, and hence, in two auction settings, two equlibria that each attain 86% of the respective optimal welfare might be viewed as comparable in their (in)efficiency. Such inter-context comparability might be more challenging to achieve in settings that lack a natural reference point (or where reference points are not clearly comparable across different settings). In such settings, even if we have a way to cardinally aggregate over the individuals using a common unit of measurement in a way that gives intra-context comparability, it might not be clear how to compare across contexts (e.g., it might be hard to normalize by dividing by the optimum, since shifting any aggregate numeric function by any constant changes the ratio of function values between any point and the optimum). Note that a prerequisite for *absolute* inter-context cardinal comparability—i.e., a prerequisite for absolute social inefficiency quantities (like 4/3 or 86%) to have comparable meanings across contexts—is *relative* inter-context cardinal comparability—i.e., for relative social inefficiency differences to have comparable

90% of the optimal Nash social welfare function as losing "twice as much fairness" as 95%, nor is it completely clear whether these should even be measured as percentages of the Nash social welfare, of its logarithm, or of any other monotone transformation thereof.

[9] Of course, a cumulative driving time of zero as a reference point might not be a suitable reference point in some settings. More generally, while, as Luce and Raiffa (1957, p. 33) write, when restricting to ordinal comparisons, "the non-uniqueness of the zero point is of no real concern in any of the applications of utility theory, but the arbitrary unit of measurement gives trouble," in our setting (and more generally, for approximation theorems in general), the non-uniqueness of a zero point is of real concern as well.

meanings across contexts, e.g., for any specific numerical difference of (in)efficiencies to have comparable meaning not only within a context (as in intra-context comparability) but also across contexts.

It is unclear how to state economically meaningful inefficiency-bounding theorems in settings in which a measure of social inefficiency that satisfies the three properties just discussed is not readily available, such as settings without a common unit of measurement with which to compare utilities across individuals, or without a natural exogenously given reference point. In such settings, if we were to have at our disposal an economically meaningful cardinal way—that satisfies the above three properties—for measuring the (in)efficiency of an alternative, this would open the door to using techniques, fine tuned over several decades in the computer-science literature, to bound this (in)efficiency at alternatives of interest (such as equilibria) and derive new economic insights in settings in which economically interpretable inefficiency-bounding theorems might currently be lacking.

In this paper, we initiate the study of axiomatically founding a *cardinal* welfare function—or more specifically, a *cardinal social inefficiency* function—satisfying the above three properties, with the aim of demonstrating a new direction for building sound economic foundations for inefficiency-bounding theorems. Our social inefficiency function is uniquely defined (up to the unit in which it is globally measured, across all contexts) and depends only on the vNM preferences of each individual over the various possible alternatives. The definition of our social inefficiency function does not assume any exogenously given way to compare utility across individuals, nor does it assume any exogenously given reference point or disagreement point (such as having special meaning for a utility value of zero). We demonstrate the use of this social inefficiency function by applying it to a setting in which interpersonal comparison is notoriously challenging to justify and a reference point might not be available: Object allocation without money. In this setting, we use our social inefficiency function to leverage existing mathematical constructions and techniques from the computer-science literature to prove an approximate-efficiency result for the popular Random Serial Dictatorship mechanism.

We define a social inefficiency function as a function that for every *context*—set of possible alternatives, number of individuals, and vNM preferences for the individuals over (lotteries over) these alternatives—and for every specific alternative (or lottery over alternatives), specifies a nonnegative number, which we call the *social ineffi-*

ciency of this alternative given the context.[10,11] After formalizing these definitions in Section 2, in Section 3 we define seven axioms that such a social inefficiency function should satisfy.

Our first two axioms relate to ordinal aggregation (the first of the three properties discussed above). First, *Pareto monotonicity* gives a necessary condition on the ordinal nature of our social inefficiency function in straightforward cases: This axiom requires that if all individuals prefer one alternative over another, then the social inefficiency of the former should be lower than of the latter. Second, *anonymity* gives a necessary condition on the aggregative nature of our social inefficiency function when it is straightforward that two contexts should be treated in the same way from an aggregate point of view: This axiom requires that permuting the set of players does not change the social inefficiency of any alternative.

Our third axiom relates to intra-context cardinal comparability. Inspired by expected utility theory, *expected inefficiency* gives a necessary condition on the cardinal meaning of our social inefficiency function: This axiom requires that the social inefficiency at a lottery between two alternatives is the expected social inefficiency over the lottery. One of our main use cases for social inefficiency (and our application in this paper) is quantifying the social inefficiency of the outcome of *randomized* mechanisms, and this axiom stipulates that this social inefficiency coincides with the expected social inefficiency of the realized outcome, i.e., that the *social aggregation* of inefficiency is risk-neutral.[12]

[10]Regarding dependence on the context, it is useful to recall the words of Arrow (1963, p. 110): "Guilbaud [(1952)] argues that a social welfare function based on individual utilities which are meaningful only up to monotonic transformations must be what is mathematically termed a functional, not a function in the ordinary sense. That is, each social choice must depend on the entire individual ordering."

[11]A main reason for considering the social inefficiency of randomized and not only deterministic alternatives is that an important use case (indeed, our application in this paper) is quantifying the social inefficiency of the outcome of a *randomized* mechanism. More generally, as Arrow (1951, p. 20) writes: "if conceptually we imagine a choice being made between two alternatives, we cannot exclude any probability distribution over those two choices as a possible alternative."

[12]We note that one interpretation of this axiom (in tandem with Pareto monotonicity) is that the social inefficiency at a lottery between two alternatives is the same as at an imaginary alternative in which every individual receives as much utility as in her certainty equivalent for the lottery. Stipulating that the social aggregation is risk-neutral therefore invariably allows the risk preferences *of individuals* to be taken into account in the social inefficiency function, and is one of the drivers of our social inefficiency function taking a utilitarian form (Harsanyi, 1955). While the dependence of utilitarian forms on risk preferences has drawn criticism in some settings, we refer the readers to the excellent, eloquent discussion in the first five pages of the introduction of Dhillon and Mertens (1999) regarding why such criticism might be less applicable to social choice functions not intended

Moving to inter-context cardinal comparability (the last of the three properties discussed above), three axioms provide prerequisites for *relative* inter-context cardinal comparability. The first of these three axioms gives a necessary condition for relative cardinal comparability across contexts with the same individuals and larger or smaller sets of alternatives. This axiom is the familiar *independence of irrelevant alternatives (IIA)*, with two adjustments. First, to achieve cardinal rather than only ordinal comparability, instead of requiring that if the social inefficiency of one alternative is lower than that of another then this order remains the same regardless of changes to irrelevant alternatives, we require that if the social inefficiency of one alternative is lower *by any specific difference* than that of another then this *difference* remains the same regardless of changes to irrelevant alternatives. The second adjustment to IIA is required since we are not given any exogenous reference point such as a disagreement point. Recall that IIA traditionally requires independence of irrelevant alternatives as long as the disagreement point is unchanged (Nash, 1953). In an elegant paper, Roth (1977) explores variants of IIA with respect to various reference points, and in particular considers two alternative, *endogenous* reference points: the *ideal point*—an imaginary alternative in which every individual receives as much utility as in her ideal alternative—and the *point of minimal expectations*—an imaginary alternative in which every individual receives as much utility as in her least-preferred undominated alternative. Our IIA axiom requires independence of irrelevant alternatives as long as both of these points are unchanged.[13]

The second of these three axioms gives a necessary condition for relative cardinal comparability across contexts with the same number of individuals (however not the exact same individuals) and the same alternatives: *independence of irrelevant preferences* considers a scenario in which two contexts with disjoint sets of individuals are composed into one, composed context that captures both original contexts coexisting

as voting rules but rather for measurement and policy discussions. Our social inefficiency function is indeed intended for the latter, which drives our focus on cardinal comparability, both intra-context and inter-context.

[13]We emphasize that the more points we require to be unchanged, the *weaker* IIA becomes. As Arrow (1963, p. 112) writes, a weakening of IIA is indeed required to define a meaningful social welfare function: "If empirically meaningful interpersonal comparisons have to be based on indifference maps, as we have argued, then the Independence of Irrelevant Alternatives must be violated. The information which enables us to assert that individual A prefers x to y more strongly than B prefers y to x must be based on comparisons by A and B of x and y not only to each other but also to other alternatives." Our social inefficiency function is indeed based on comparisons to two additional alternatives—the two endogenous reference points defined by Roth (1977).

alongside each other with no interaction between them.[14] This axiom requires that in such a scenario, the relative social inefficiency between two composed alternatives that differ only on one of the original contexts is independent of the preferences in the original context on which these two composed alternatives coincide. The third of these three axioms gives a necessary condition for relative cardinal comparability across contexts with different numbers of individuals: *duplication* requires that the relative inefficiency between two alternatives in a context remains the same, and does not arbitrarily change, upon composition of copies of that context.

A final axiom relates to absolute (rather than relative) inter-context cardinal comparability. If we had, in each context, an exogenously given optimal alternative, we would have required that this alternative have a social inefficiency of zero. Since we do not have such an exogenous optimal alternative, *feasibility* only requires the much weaker condition that at each context there is *some* alternative with zero social inefficiency.

In Section 4, we state the main result of the first part of this paper: That these seven, logically independent axioms characterize a unique social inefficiency function up to a global choice of unit of measure, that is, up to a positive multiplicative factor applied *simultaneously across all contexts and alternatives*. As we show, the characterized social inefficiency function takes a particularly natural form: It measures the per capita additive loss in welfare relative to the maximum-welfare point, where welfare is defined as the sum of *endogenously* normalized vNM utility functions, each utility function normalized so that its unit is the diameter of its range over the Pareto frontier.[15] This social inefficiency function has additional appealing properties beyond

[14]Formally, possible alternatives of the composed context are pairs of alternatives from the respective original contexts, and each individual's preferences over composed alternatives is induced by her preference over alternatives in her original context.

[15]Arrow (1951, p. 31) writes that "As it stands, [the sum of individual utilities] seems to be excluded by the entire nature of the present approach, since [...] we agreed to reject the idea of cardinal utility, and especially of interpersonally comparable utility. However, presumably the sum of utilities could be reformulated in a way which depends only on the individual orderings and not on the utility indicators." He then proceeds to write that "The only way that I can see of making the sum of utilities depend only on the indifference loci is the following: Since to each individual ordering there corresponds an infinite number of utility indicators, set up an arbitrary rule which assigns to each indifference map one of its utility indicators; then the sum of the particular utility indicators chosen by the rule is a function of the individual orderings and can be used to establish a social ordering." In other words, this involves an arbitrary endogenous normalization. Indeed, this is what our social inefficiency function does, however our rule that "assigns to each indifference map one of its utility indicators" (normalizing so that the unit is the diameter of the range of the utility function over the Pareto frontier) is far from arbitrary, and uniquely follows from our axioms.

those explicitly required by our axioms, such as that the social inefficiency of an alternative is not changed if weakly (or strongly) dominated alternatives are added to, or removed from, the context in question.[16]

In Section 5, we apply the social inefficiency function that emerges from our axiomatic characterization to the study of a setting in which interpersonal comparison is notoriously challenging to justify due to the absence of monetary transfers and the repugnance of measuring using money, and any normalization is challenging due to the absence of a reference point such as an outside option: That of object allocation (without money). We consider the robust social inefficiency guarantee of truthful ordinal mechanisms at their truthtelling equilibrium. We utilize tools that were originally developed within computer science for the study of the "Price of Anarchy" (Filos-Ratsikas et al., 2014), to prove that no such mechanism has a robust social inefficiency (as defined by our axioms) guarantee that improves upon that of the popular Random Serial Dictatorship mechanism by more than 28%.[17] We emphasize two crucial features of our social inefficiency function that give meaning to the statement of this result: First, we can define the robust social inefficiency guarantee of a mechanism (i.e., worst social inefficiency over all object allocation problems) because our social inefficiency function is cardinally comparable across contexts/object allocation problems. Second, no other mechanism being able to improve by more than 28% is interpretable because our axioms pin down a unique *cardinal* social inefficiency function up to a *global* unit of measure (and hence the proportion 28% is well defined independently of this unit of measure). Finally, we consider the task of computing our social inefficiency function over concrete given alternatives and preferences, and

[16]Arrow (1951) criticizes a social choice function for *not* satisfying this property: Regarding the social choice function that normalizes each utility function to the range of $[0, 1]$ (over its entire range, not only over the Pareto frontier) and sums up the normalized utilities, he writes (p. 32) that "It is not hard to see that the suggested assignment of utilities is extremely unsatisfactory. Suppose there are altogether three alternatives and three individuals. Let two of the individuals have the utility 1 for alternative x, .9 for y, and 0 for z; and let the third individual have the utility 1 for y, .5 for x and 0 for z. According to the above criterion, y is preferred to x. Clearly, z is a very undesirable alternative since each individual regards it as worst. If z were blotted out of existence, it should not make any difference to the final outcome; yet, under the proposed rule for assigning utilities to alternatives, doing so would cause the first two individuals to have utility 1 for x and 0 for y, while the third individual has utility 0 for x and 1 for y, so that the ordering by sum of utilities would cause x to be preferred to y."

[17]Our social inefficiency function therefore not only conceptually provides an economically meaningful alternative for the "Price of Anarchy," but also technically allows to rather easily adapt Price-of-Anarchy bounds in various game-theoretic and economic settings into social inefficiency bounds that do not assume that utilities are more than representations of ordinal (vNM) preferences.

construct a computationally feasible algorithm for doing so.

There is a vast body of mathematical techniques within computer science for proving what computer scientists call *approximation theorems*: Theorems that ascertain that some quantity is at least a specific fraction of some reference quantity. This paper aims to make this vast body of work more applicable to economic settings by providing a microfoundation for a cardinal aggregation of individuals' (vNM) preferences. While we focus on cardinally measuring (in)efficiency, the program presented in this paper has the potential to be applicable far more broadly. For example, one might equally imagine axiomatically defining a cardinal social (un)fairness function, microfounding a cardinal notion of approximate fairness and using techniques from computer science (and beyond) to provide approximate fairness guarantees. We indeed hope that *cardinal* social choice theory could facilitate bringing economic theory and theoretical computer science even closer together.

1.1 Further Related Literature

To our knowledge, our paper is the first that, for arbitrary sets of alternatives, axiomatically microfounds a social welfare function that depends *only* on the individuals' ordinal (vNM) preferences over the alternatives, whose *numeric value* (rather than only its induced order) is unique and cannot be arbitrarily additively shifted, and which can be meaningfully compared *across contexts*. We are furthermore not aware of any paper that uses or characterizes the *order* induced by our social welfare function.

The question of how utilities of different individuals should be compared in order to assess the welfare of society is long-standing in welfare economics and social choice theory, two momentous fields whose comprehensive summary is beyond the scope of this literature review. For many years, economists have been split on whether and to what extent interpersonal comparison of utility is justified. In his 1951 book, "Social Choice and Individual Values," Arrow takes the stand that "interpersonal comparison of utilities has no meaning." Later, Harsanyi (1955) argues to the contrary, for utilitarianism (taking the sum of the individuals' utilities) based on the idea that the individuals' utility functions can be meaningfully compared and aggregated. He expands social welfare functions to randomized alternatives and adopts the vNM axioms (von Neumann and Morgenstern, 1953) for societal preferences. He then shows

that the social welfare functions that are a weighted sum of the individuals' cardinal utilities (with weights that may be arbitrarily chosen) are the only ones that are Pareto monotone and satisfy an additional axiom called independence. Following Harsanyi (1955), many other papers construct social welfare functions that take as input a profile of ordinal preferences of the individuals over the given set of alternatives and yield a societal ranking of the alternatives. Among these papers, several influential ones assume given utility functions that are exogenously chosen to convey some additional information beyond only the ordinal preferences (e.g., DeMeyer and Plott, 1971; Maskin, 1978; Sen, 1970; Roberts, 1980; Hammond, 2005). Our paper, in contrast to all of the above, does not assume any exogenous information except for the ordinal (vNM) preferences, and yet pinpoints *unique* weights on the individuals' utility functions to be used for aggregation.

A different approach is taken by Kaneko and Nakamura (1979), who define the "Nash social welfare function." They assume "the existence of a distinguished alternative [...] which represents one of the worst states for all individuals that we may imagine, e.g., we may imagine all the members of the society die," and evaluate social welfare by considering (the product of) the relative increases of the individuals' welfare from this distinguished alternative. The order induced by this function (which is what they axiomatize) is robust to affine transformations of the utility functions that represent the individuals' preferences, while the numeric value depends on the precise choice of utility functions (and can be furthermore arbitrarily monotonically transformed without violating their axioms). Piacquadio (2017) assumes that alternatives are allocations of commodities and introduces an "equal-preference transfer" axiom that considers the relative intensity of two individuals' welfare over two alternatives between which all other members of society are indifferent. He then shows that if the set of alternatives consists of a bundle that is considered worst by all individuals, the social welfare functions in a class that he terms "opportunity-equivalent utilitarianism" are the only ones that meet this axiom in conjunction with several other axioms. In contrast to both of these papers, our construction *does not* depend on a modeling choice of a distinguished worst-case alternative or bundle for all members of society.

The closest paper to ours is that of Dhillon and Mertens (1999), who axiomatize a social welfare function of relative utilitarianism. This paper axiomatically justifies normalizing each individual's (vNM) utility function between 0 and 1 (over all alternatives) and then summing the individuals' normalized utilities to achieve the

social welfare of a given alternative. To do so, this paper assumes that "the set of alternatives is sufficiently encompassing as to include, besides the actual alternatives of interest, each person's best and worst alternative within the 'universal' set." Our social inefficiency function in contrast normalizes the utilities of each individual to be 0 and 1 on the worst and best (from her point of view) alternatives *on the Pareto frontier*, which is endogenously determined for each context. Therefore, we do not need to assume such "richness" (or any other kind of richness) of the set of alternatives. Like all papers preceding ours that we are aware of, Dhillon and Mertens (1999) axiomatize only the order induced by their social welfare function, and like most previous papers, their social welfare function does not allow for comparison across contexts. Another paper that axiomatically motivates a specific normalization of utility functions is that of Fleurbaey and Zuber (2021). In this paper, an alternative is a specification of a positive (single dimensional) consumption level for each individual "living" in that particular alternative; the set of alternatives is then the set of all such consumption profiles. Individuals' preferences are over their own consumption. Assuming differentiable vNM utility functions, this paper introduces axioms that justify the normalization of the utilities of all individuals so that their marginal utilities at a poverty line are the same. Our social inefficiency function in contrast does not assume any structure on the set of alternatives (or on the vNM utility functions).

Fleurbaey and Tadenuma (2014) develop a social-choice theory that allows for comparison across societies, resembling our inter-context comparisons. In their paper, a *social situation* consists of some population, preferences for each individual in that population over her own consumption, and a consumption profile that specifies a consumption bundle for each individual in that population. Fleurbaey and Tadenuma (2014) seek to ordinally compare such social situations. (They are not interested in comparison of lotteries over social situations.) On a conceptual level, while their social situations technically resemble context–alternative pairs in our paper, they seek to answer the question of which of two social situations is preferred (e.g., which country is better off), while we seek to quantify, in every context, the inefficiency of an alternative *as contextualized by the context in question*. On a technical level, this translates in their setting to each individual always having preferences over the same, structured set of consumption bundles for that individual; in contrast, our sets of alternatives are completely structureless (e.g., not even a metric or topological

space) and arbitrarily change between contexts (indeed, in our setting in contrast to theirs, the inefficiency of an alternative in some context might change if new possible alternatives are added to the context). In both Fleurbaey and Tadenuma (2014) and Fleurbaey and Zuber (2021), and in contrast to our paper, the preferences of individuals are monotone, and therefore for every preference profile considered, neither an optimal or ideal point, nor a Pareto frontier of consumption profiles, exist. Furthermore, like all preceding papers of which we are aware, Fleurbaey and Tadenuma (2014) characterize *orderings* that satisfy various axioms rather than characterize a unique real-valued index.

A different stream of literature follows the Nash bargaining solution (Nash, 1953) and, in a setting where society consists of two individuals, searches for a socially preferred *alternative* rather than a social preference order. In this literature, it is usually assumed that such an alternative should be unique. Roth (1977) considers different "Independence of Irrelevant Alternatives (IIA)" axioms and the bargaining solutions they obtain. While Nash's bargaining solution is obtained for the axiom of "IIA other than the disagreement point,", Roth (1977) shows that the bargaining solution of Kalai and Smorodinsky (1975) (see also Raiffa, 1953) is also obtained for the axiom of "IIA other than both the disagreement point and the ideal point" (where the ideal point is as in our paper: an imaginary alternative in which every individual receives as much utility as in her ideal alternative). Roth (1977) also introduces the point of minimal expectations, and characterizes the bargaining solutions that meet either the axiom of "IIA other than the point of minimal expectations" or the axiom of "IIA other than the ideal point." However, he does not consider, as we do, the axiom of "IIA other than both the point of minimal expectations and the ideal point." Nalebuff (2021) looks for "perspective invariant" bargaining solutions in the sense of obeying a modified version of Nash's axioms with respect to the (exogenous) disagreement point as well as a flipped version of the same axioms with respect to the (endogenous) ideal point. He normalizes so that the (exogenous) disagreement point is (0,0) and the (endogenous) ideal point is (1,1) and exhibits such a modification of Nash's axioms that gives a bargaining solution that maximizes the sum of the two individuals' utilities.

Computer scientists have been interested in measuring the inefficiency of a given alternative compared to the optimal efficiency. Within games, Koutsoupias and Papadimitriou (1999) define the Price of Anarchy as the ratio between the worst-case

performance at equilibrium and the optimal performance at any strategy profile—both measured using some welfare (or cost) function (such as the cumulative driving time of all individuals, as studied by Roughgarden and Tardos, 2002). Following that paper, the Price of Anarchy (many times with the sum of exogenously normalized utilities as the welfare function) has been adopted not only in computer science but also in economics, operations research, industrial engineering, and other areas as a measure of the inefficiency of an interactive environment. For example, the Price of Anarchy has been applied to classical operations research areas such as supply chains (e.g., Perakis and Roels, 2007), queuing (see Ghosh and Hassin, 2021 for a survey), load balancing (e.g., Ayesta et al., 2010), and scheduling (e.g., Hoeksma and Uetz, 2019). In contrast to the Price of Anarchy, our social inefficiency function does not require an exogenously given welfare or cost function (or exogenously normalized utility functions) that are cardinally comparable within and between contexts.

2 Model

A *context* is a pair $\left(X, (\succeq_i)_{i=1}^n\right)$ where X is a nonempty finite set called the set of *alternatives*, $n \in \mathbb{N}$ is called the number of *individuals*, and for each $i = 1, \ldots, n$, $\succeq_i$ is a preference (for individual i) over $\Delta(X)$ that satisfies the von Neumann–Morgenstern (vNM) rationality axioms (i.e., that has a vNM utility representation).[18] Let $x, y \in \Delta(X)$. We interpret $x \succeq_i y$ as saying that individual i weakly prefers x over y. A main example of a context is derived from an n-player game, where X is the set of possible strategy profiles of the game, and $\succeq_i$ is the preference relation of player i (over $\Delta(X)$) for every i.

For a preference $\succeq$, we write $x \succ y$ to mean that $x \succeq y$ but not $y \succeq x$, and write $x \sim y$ to mean that both $x \succeq y$ and $y \succeq x$. For a context $C = \left(X, (\succeq_i)_{i=1}^n\right)$, we use the notation $F_C \subseteq \Delta(X)$ to denote the *Pareto frontier* of C: the set of all $x \in \Delta(X)$ such that there does not exist $x' \in \Delta(X)$ such that $x' \succeq_i x$ for every i and $x' \succ_i x$ for some i.

Definition 1 (Social Inefficiency Function)**.** A *social inefficiency function* specifies for each context $C = \left(X, (\succeq_i)_{i=1}^n\right)$ and $x \in \Delta(X)$ a number $I(C, x) \in \mathbb{R}_{\geq 0} \cup \{\infty\}$,

[18] As is standard, $\Delta(X)$ denotes the set of lotteries over elements of X. For $x, y \in \Delta(X)$ and $0 \leq \alpha \leq 1$, we write $\alpha \cdot x + (1 - \alpha) \cdot y$ to denote the lottery whose outcome is with probability α drawn from x and otherwise drawn from y.

called the *social inefficiency* of x with respect to C.[19,20]

While we allow our social inefficiency function to attain the value of infinity, our characterized social inefficiency function does so quite sparingly. After providing an explicit construction for this function in Section 4.1, we discuss the benefits of (sparingly) allowing for infinite social inefficiency. Specifically, a desirable property that we term *invariance to dominated alternatives*[21] cannot hold without (sparingly) allowing for infinite social inefficiency in all of the points in which our characterized social inefficiency function attains this value; see a discussion before Lemma 1.

3 Axioms

For every context $C = \left(X, (\succeq_i)_{i=1}^n\right)$ and $x \in \Delta(X)$, we wish to axiomatically define the *social inefficiency* $I(C, x) \geq 0$ (possibly infinite) of the alternative x within the context C. In this section, we state our axioms for I.

The first axiom is a standard (ordinal) Pareto monotonicity axiom—i.e., that if y is weakly (strictly) Pareto dominated by y, then the social inefficiency at y is weakly (strictly) greater than at x—with one technical tweak that *weakens* this axiom since we allow for a social inefficiency of infinity. The first part of this technical tweak is that if the social inefficiency at y is required to be strictly greater than at x, and the social inefficiency at x is already infinity, then the social inefficiency at y is allowed to be infinity as well, even though technically this makes the social inefficiency at y be the same as at x. The second part of this technical tweak slightly qualifies its first part. One role of strict Pareto monotonicity, across many models and proofs, is to ensure that the characterized function is not constant along every chain of alternatives that is ordered by Pareto dominance. If the first part of the tweak were left unqualified, then the social inefficiency function would in fact have been allowed to be constant

[19] A main example is analylzing the context C derived from a game and taking x to be an equilibrium of the game. One might even take the supremum (or infimum) of $I(C, x)$ over all equilibria x of the game as a measure of the inefficiency of the game. This would serve a similar purpose to the Price of Anarchy from computer science, however will be invariant to strategic equivalence (i.e., not depend on the choice of vNM utility function for each player).

[20] We use the term social inefficiency *function* even though, strictly speaking, I is not a function because the class of all contexts is not a set. A stricter mathematical definition would therefore be that a social inefficiency function is a family of functions $I_C(x)$ that are indexed by contexts. For clarity of exposition, we gloss over this distinction, which is immaterial for this paper.

[21] See Footnote 16 (on Page 9) for Arrow's (1951) criticism of a social choice function for not satisfying this property.

along such a chain, taking the value of infinity along the entire chain even if this were a maximal chain (i.e., a chain that is not a subset of a longer chain). The second part of the technical tweak slightly qualifies its first part so that this cannot happen.

Axiom 1 (Pareto Monotonicity)**.** Let $C = \big(X, (\succeq_i)_{i=1}^n\big)$ be a context. For every $x, y \in \Delta(X)$, if $x \succeq_i y$ for every $i = 1, \ldots, n$, then $I(C, x) \leq I(C, y)$, and if furthermore $x \succ_i y$ for some such i, then either $I(C, x) < I(C, y)$ or both (i) $I(C, x) = I(C, y) = \infty$ and (ii) there exists $w \in \Delta(X)$ with $w \succeq_i x$ for every $i = 1, \ldots, n$ such that $I(C, w) < \infty$.

The second axiom stipulates that the social inefficiency is unchanged if the order of individuals is permuted.

Axiom 2 (Anonymity[22])**.** Let $C = \big(X, (\succeq_i)_{i=1}^n\big)$ be a context, let π be a permutation on $\{1, \ldots, n\}$ and set $C_\pi \triangleq (X, (\succeq_{\pi(i)})_{i=1}^n)$. For every $x \in \Delta(X)$, it holds that $I\big(C_\pi, x\big) = I(C, x)$.

The third axiom stipulates that the social inefficiency of a lottery over alternatives is the expected social inefficiency of the drawn alternative. That is, the social aggregation of inefficiency is risk-neutral. This axiom is hence named *expected inefficiency*, analogously to *expected utility*.

Axiom 3 (Expected Inefficiency)**.** Let $C = \big(X, (\succeq_i)_{i=1}^n\big)$ be a context. For every $x, y \in \Delta(X)$ and $0 \leq \alpha \leq 1$, it holds that $I\big(C, \alpha \cdot x + (1 - \alpha) \cdot y\big) = \alpha \cdot I(C, x) + (1 - \alpha) \cdot I(C, y)$.

The fourth axiom is a cardinal version of IIA, stipulating that the relative social inefficiency between two alternatives is independent of the addition or removal of other alternatives that do not change the *ideal point*—an imaginary alternative in which every individual receives as much utility as in her ideal alternative—and the *point of minimal expectations*—an imaginary alternative in which every individual receives as much utility as in her least-preferred undominated alternative. For a set $A \subseteq \Delta(X)$, we use $\max_{\succeq_i} A$ (resp. $\min_{\succeq_i} A$) to denote some $x \in A$ such that $x \succeq_i y$ (resp. $x \preceq_i y$) for every $y \in A$.

[22]The following, weaker yet bulkier variant of anonymity suffices just as well for our characterization. Let $C = \big(X, (\succeq_i)_{i=1}^n\big)$ be a context, let π be a permutation on $\{1, \ldots, n\}$, and set $C_\pi \triangleq \big(X, (\succeq_{\pi(i)})_{i=1}^n\big)$. For every $x, y \in \Delta(X)$, it holds that $I(C_\pi, x) - I(C_\pi, y) = I(C, x) - I(C, y)$ whenever neither of the two sides of this equality is "$\infty - \infty$." That is, the *relative* social inefficiency between two alternatives is unchanged if the order of individuals is permuted.

Definition 2 (Ideal Point and Point of Minimal Expectations (Roth, 1977))**.** Let $C = \left(X, (\succeq_i)_{i=1}^n\right)$ be a context, let $\emptyset \neq X' \subset X$, and set $C' \triangleq \left(X', (\succeq_i|_{\Delta(X')})_{i=1}^n\right)$.

- C and C' are said to have *the same ideal point* if $\max_{\succeq_i} X' \sim_i \max_{\succeq_i} X$ for every $i = 1, \ldots, n$.
- C and C' are said to have *the same point of minimal expectations* if $\min_{\succeq_i} F_{C'} \sim_i \min_{\succeq_i} F_C$ for every $i = 1, \ldots, n$.

Axiom 4 (Independence of Irrelevant Alternatives (IIA))**.** Let $C = \left(X, (\succeq_i)_{i=1}^n\right)$ be a context, let $\emptyset \neq X' \subset X$, and set $C' \triangleq \left(X', (\succeq_i|_{\Delta(X')})_{i=1}^n\right)$. If C and C' have the same ideal point and the same point of minimal expectations, then for every $x, y \in \Delta(X')$ it holds that $I(C', x) - I(C', y) = I(C, x) - I(C, y)$ whenever neither of the two sides of this equality is "$\infty - \infty$."

Before introducing the fifth and sixth axioms, following the definition by von Neumann and Morgenstern (1953) of the composition of two games, we define an analogous notion of composition of several contexts into one, composed context that models the original contexts as coexisting alongside each other, with disjoint sets of individuals, without any interaction between these contexts. Let $C^1 = \left(X^1, (\succeq_i^1)_{i=1}^{n^1}\right)$ and $C^2 = \left(X^2, (\succeq_i^2)_{i=1}^{n^2}\right)$ be two contexts (with possibly different numbers of individuals). We define the *composed context* $C^1 \oplus C^2$ as the context that captures these two contexts existing concurrently and without affecting each other. Formally, $C^1 \oplus C^2 \triangleq \left(X^1 \times X^2, (\succeq_i)_{i=1}^{n_1+n_2}\right)$ where for every $x, y \in \Delta(X^1 \times X^2)$, we have that:

1. For every $i = 1, \ldots, n^1$, it holds that $x \succeq_i y$ if and only if $x^1 \succeq_i^1 y^1$, where $x^1, y^1 \in \Delta(X^1)$ are the respective marginal distributions of x, y over X^1.
2. For every $i = n^1+1, \ldots, n^1+n^2$, it holds that $x \succeq_i y$ if and only if $x^2 \succeq_{i-n^1}^2 y^2$, where $x^2, y^2 \in \Delta(X^2)$ are the respective marginal distributions of x, y over X^2.

The fifth axiom stipulates that in a composed context, the relative social inefficiency between two composed alternatives that differ only on one of the original contexts is independent of the preferences in the original context on which these two composed alternatives coincide.

Axiom 5 (Independence of Irrelevant Preferences (IIP))**.** Let $C = \left(X, (\hat{\succeq}_i)_{i=1}^n\right)$ be a context, and let $D = \left(Y, (\succeq_i)_{i=1}^m\right)$ and $D' = \left(Y, (\succeq_i')_{i=1}^m\right)$ be two contexts with the same

number of individuals and same alternatives. For every $x, x' \in \Delta(X)$ and $y \in \Delta(Y)$, it holds that $I\big(C \oplus D, (x, y)\big) - I\big(C \oplus D, (x', y)\big) = I\big(C \oplus D', (x, y)\big) - I\big(C \oplus D', (x', y)\big)$ whenever neither of the two sides of this equality is "$\infty - \infty$."

The sixth axiom facilitates comparability of relative inefficiencies across contexts with different numbers of individuals, by ensuring that the relative inefficiency between two alternatives in a context remains the same, and does not arbitrarily change, upon self-composition of the context. (We define the composition of more than two contexts analogously to the composition of two contexts.)

Axiom 6 (Duplication)**.** Let $C = \big(X, (\succeq_i)_{i=1}^n\big)$ be a context. For every $x, x' \in \Delta(X)$ and $k \in \mathbb{N}$, it holds that $I\big(\oplus_{j=1}^k C, (x, \ldots, x)\big) - I\big(\oplus_{j=1}^k C, (x', \ldots, x')\big) = I(C, x) - I(C, x')$ whenever neither of the two sides of this equality is "$\infty - \infty$."

The seventh and final axiom ensures that in each context, at least one alternative has social inefficiency zero—i.e., is considered most efficient.

Axiom 7 (Feasibility)**.** Let $C = \big(X, (\succeq_i)_{i=1}^n\big)$ be a context. There exists $x \in \Delta(X)$ such that $I(C, x) = 0$.

4 Characterization

In this section, we constructively define a social inefficiency function and prove that this social inefficiency function is uniquely characterized by our seven axioms from Section 3, up to a global choice of unit of measure.

4.1 Explicit Construction

Let $C = \big(X, (\succeq_i)_{i=1}^n\big)$ be a context, and for every $i = 1, \ldots, n$, let u_i be some vNM utility representation of $\succeq_i$. Let $u_i^{max} \triangleq \max_{x \in F_C} u_i(x)$ and $u_i^{min} \triangleq \min_{x \in F_C} u_i(x)$. Let $V(C, x) \triangleq \frac{1}{n} \sum_{i=1}^n \frac{u_i(x) - u_i^{min}}{u_i^{max} - u_i^{min}}$ for every $x \in \Delta(X)$; if the denominator in one of these fractions is zero, then we follow the convention that a negative number divided by zero equals minus infinity and that zero divided by zero equals zero. We note that $V(C, x)$ does not depend on the choice of vNM utility representations of $(\succeq_i)_{i=1}^n$. One can view $V(C, \cdot)$ as a per capita utilitarian social-welfare function that normalizes the utility function of each individual so that its maximum and minimum over the

Pareto frontier F_C are 1 and 0, respectively. We define the social inefficiency at a point $x \in \Delta(X)$ as the loss in the value of $V(C, \cdot)$ compared to its maximum possible value:

$$\hat{I}(C, x) \triangleq \max_{x' \in X} V(C, x') - V(C, x).$$

We note that $\hat{I}(C, x) \geq 0$ for every context C and $x \in \Delta(X)$. Hence (and since $\hat{I}(C, x)$ does not depend on the choice of vNM utility representations of $(\succeq_i)_{i=1}^n$), we have that $\hat{I}$ is a social inefficiency function.

The social inefficiency function $\hat{I}$ has a natural interpretation: It is the per capita additive utility loss (compared to the first-best) when utility functions are endogenously normalized so that the unit of each utility function is the diameter of its range over the Pareto frontier.[23]

It is straightforward to verify that the social inefficiency function $\hat{I}$ satisfies our seven axioms. An additional appealing property of this social inefficiency function is that for every $x \in \Delta(X)$, the individuals' preferences over the Pareto frontier F_C and over x (but not over the rest of $\Delta(X)$) uniquely determine the social inefficiency of x—a property that we term *invariance to dominated alternatives*.

Definition 3 (Invariance to Dominated Alternatives)**.** A social inefficiency function I satisfies *invariance to dominated alternatives* if for every context $C = \left(X, (\succeq_i)_{i=1}^n\right)$ and every $X' \subset X$ such that $\Delta(X')$ weakly dominates $X \setminus X'$,[24] it holds that $I(C', x) = I(C, x)$ for every $x \in \Delta(X')$, where $C' \triangleq \left(X', (\succeq_i|_{\Delta(X')})_{i=1}^n\right)$.

Invariance to dominated alternatives—which we show in Lemma 3 below to be implied by Pareto monotonicity, IIA, and feasibility—is a desirable property, which relieves models that use our social inefficiency function from the burden of encompassing "universally worst" alternatives such as "all the members of the society die" (Kaneko and Nakamura (1979) explicitly state that their formalism requires such a point, and that such a point may take starkly different meanings across contexts; Dhillon and Mertens (1999) explicitly require such a worst point to be included for every individual, though do not require that it is the same point for all individuals), and allows such models to focus only on modeling undominated alternatives and other

[23] If the diameter of the range of the utility function of some individual over the Pareto frontier is zero (i.e., if this utility function is constant over the Pareto frontier), then the normalized utility function is the limit of normalizations of the utility function whose unit tends to zero.

[24] That is, for every $x \in X \setminus X'$ there exists $x' \in \Delta(X')$ such that $x' \succeq_i x$ for every $i = 1, \ldots, n$.

alternatives of interest.[25]

If we had at our disposal vNM utility functions that are exogenously normalized in a way that allows for interpersonal comparison, then a social inefficiency function that measures the loss of efficiency compared to the welfare-maximizing alternative might naturally depend only on the (interpersonally comparable) utility values at the welfare-maximizing alternative and at the alternative whose efficiency is evaluated.[26] Invariance to dominated alternatives is a natural generalization of this property for settings in which an exogenous way to compare utilities across individuals is not available. Indeed, invariance to dominated alternatives precisely requires that the social inefficiency of x is uniquely determined by the individuals' preferences over x as well as over all alternatives that are "potential welfare maximizers," since the Pareto frontier is precisely the set of all points that are each a welfare-maximizer for *some* choice of vNM utility representations of the individuals' respective preferences.

We say that an individual $i \in \{1, \ldots, n\}$ is *frontier-indifferent* (in C) if $x \sim_i y$ for every $x, y \in F_C$; otherwise, we say that i is *frontier-concerned*. Note that the denominator in the definition of $\hat{I}$ is zero (only) for frontier-indifferent i. Therefore, $\hat{I}(C, x)$ is infinite if (and only if) x is less preferred by a frontier-indifferent i than all (equivalent, any) Pareto-frontier points, which we denote by $F_C \succ_i x$. While allowing for infinite inefficiency might seem unappealing at first glance, the following lemma shows that these alternatives having infinite inefficiency is necessitated already by Pareto monotonicity and invariance to dominated alternatives (and hence similarly necessitated by Pareto monotonicity, IIA, and feasibility).

Lemma 1. *Let I be a social inefficiency function that satisfies Pareto monotonicity and invariance to dominated alternatives, let $C = \left(X, (\succeq_i)_{i=1}^n\right)$ be a context, and let $x \in \Delta(X)$. If there exists a frontier-indifferent $i \in \{1, \ldots, n\}$ such that $F_C \succ_i x$, then $I(C, x) = \infty$.*

Figure 1 schematically illustrates where the social inefficiency function $\hat{I}$ attains different values: All (possibly randomized) alternatives with social inefficiency zero (a nonempty set) are contained in the Pareto frontier, which is a subset of the set of all alternatives with finite social inefficiency (all the alternatives over which the Pareto frontier is not strictly preferred by any frontier-indifferent individual), which

[25] See also Footnote 16 (on Page 9) for Arrow's (1951) criticism of a social choice function for not satisfying invariance to dominated alternatives.

[26] This is the case for the Price of Anarchy in computer science.

is in turn a subset of the full set of all alternatives. Each of these containments is strict in some contexts and weak in others.

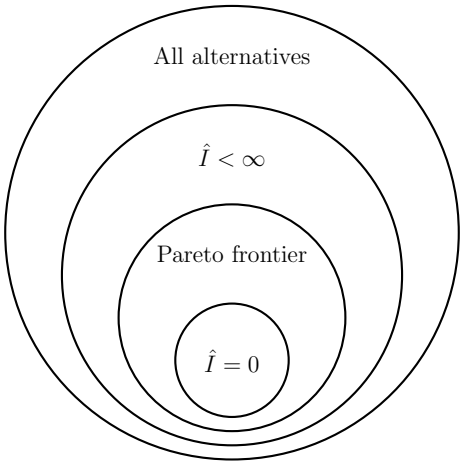


Figure 1: Schematic illustration of where different social inefficiency values are attained.

Our main result is that our seven axioms uniquely define the social inefficiency function $\hat{I}$, up to a single positive (finite) multiplicative constant that applies to all contexts and alternatives.

Theorem 1. *A social inefficiency function I satisfies Pareto monotonicity, anonymity, expected inefficiency, IIA, IIP, duplication, and feasibility if and only if there exists a constant $0 < c < \infty$ such that for every context $C = \left(X, (\succeq_i)_{i=1}^n\right)$ and $x \in \Delta(X)$, it holds that $I(C, x) = c \cdot \hat{I}(C, x)$. Furthermore, dropping any of these seven axioms invalidates this statement.*

4.2 Logical Independence of Our Axioms

In this section we show that our seven axioms are logically independent. That is, if we drop any one of them, Theorem 1 ceases to hold (as claimed in the second part of that theorem). We do so by demonstrating, for each of the seven axioms, a social inefficiency function that satisfies all of our axioms except that axiom.

Example 1 (Pareto Monotonicity is Logically Independent of the Other Axioms)**.** The constant zero social inefficiency function, $I(C, x) = 0$, satisfies all of our axioms except Pareto monotonicity.

Example 2 (Anonymity is Logically Independent of the Other Axioms)**.** Using the notation from the definition of $\hat{I}$, for every context C, first set $V'(C,x) \triangleq \frac{1}{1-2^n}\sum_{i=1}^{n} 2^{-i} \cdot \frac{u_i(x)-u_i^{min}}{u_i^{max}-u_i^{min}}$ for all $x \in \Delta(X)$ and then $I(C,x) = \max_{x' \in X} V'(C,x') - V'(C,x)$. This social inefficiency function satisfies all of our axioms except anonymity.

Example 3 (Expected Inefficiency is Logically Independent of the Other Axioms)**.** Using the notation from the definition of $\hat{I}$, for every context C, first set $V'(C,x) \triangleq \frac{1}{n}\sum_{i=1}^{n}\left(\frac{u_i^{max}-u_i(x)}{u_i^{max}-u_i^{min}}\right)^2$ for all $x \in \Delta(X)$ and then $I(C,x) = V'(C,x) - \min_{x' \in X} V'(C,x')$. This social inefficiency function satisfies all of our axioms except expected inefficiency.

The remaining four (counter)examples of social inefficiency functions all induce, for each context, the same order over its alternatives as induced by our social inefficiency function $\hat{I}$, and yet these examples differ cardinally from $\hat{I}$ (and from each other) in various nuanced ways. This showcases that axiomatizing a cardinal social welfare function is a far stricter requirement than axiomatizing the order that it induces.

Example 4 (IIA is Logically Independent of the Other Axioms)**.** The social inefficiency function $I(C,x) = \mathrm{rad}\big(|X|\big) \cdot \hat{I}(C,x)$ (where X is the set of alternatives in C and $\mathrm{rad}\big(|X|\big)$, the radical of $|X|$, is the product of the distinct prime factors of $|X|$) satisfies all of our axioms except IIA.

Definition 4 (Frontier Dimension)**.** Let C be a context. We define the *frontier dimension* of C, which we denote by d_C, as the number of frontier-concerned individuals in C.

Example 5 (IIP is Logically Independent of the Other Axioms)**.** Using the notation from the definition of $\hat{I}$, for every context C, first set $V'(C,x) \triangleq \frac{1}{d_C}\sum_{i=1}^{n} \frac{u_i(x)-u_i^{min}}{u_i^{max}-u_i^{min}}$ for all $x \in \Delta(X)$ and then $I(C,x) = \max_{x' \in X} V'(C,x') - V'(C,x)$. This social inefficiency function satisfies all of our axioms except IIP.

Example 6 (Duplication is Logically Independent of the Other Axioms)**.** The social inefficiency function $I(C,x) = 2^n \cdot \hat{I}(C,x)$ (where n is the number of individuals in C) satisfies all of our axioms except duplication.[27]

Example 7 (Feasibility is Logically Independent of the Other Axioms)**.** The shifted social inefficiency function $I(C,x) = 1 + \hat{I}(C,x)$ satisfies all of our axioms except feasibility.

[27] As part of our proof of Theorem 1, in Lemma 10 we characterize all social inefficiency functions that satisfy all of our axioms except duplication.

4.3 Proofs

In this section, we prove Lemma 1 and Theorem 1, as well as a number of intermediate results, some of which might be of independent interest. Several proofs are relegated to Appendix A.

4.3.1 General Properties

In this section, we state and prove several notable properties of all social inefficiency functions that satisfy various subsets of our axioms. In addition to using these properties throughout the proof of Lemma 1 and Theorem 1 below, these properties are also of independent interest and might be viewed as additional desirable properties of social inefficiency functions that are implied by our axioms.

The first property, which immediately follows from Pareto monotonicity, is *symmetry*: Two alternatives between which all individuals are indifferent must have the same social inefficiency.

Definition 5 (Symmetry)**.** A social inefficiency function I satisfies *symmetry* if for every context $C = \left(X, (\succeq_i)_{i=1}^n\right)$ and every $x, y \in \Delta(x)$ such that $x \sim_i y$ for every $i = 1, \ldots, n$, it holds that $I(C, x) = I(C, y)$.

Lemma 2. *If a social inefficiency function satisfies Pareto monotonicity, then it satisfies symmetry.*

Recall from Section 4.1 the definition of invariance to dominated alternatives: for every $x \in \Delta(X)$, the individuals' preferences over the Pareto frontier F_C and x (but not over the rest of $\Delta(X)$) uniquely determine the social inefficiency of x. As noted in that section, this property is implied by Pareto monotonicity, IIA, and feasibility.

Lemma 3. *If a social inefficiency function satisfies Pareto monotonicity, IIA, and feasibility, then it satisfies invariance to dominated alternatives.*

Proof. Let $C = \left(X, (\succeq_i)_{i=1}^n\right)$ be a context and let $X' \subset X$ be a set of alternatives such that $\Delta(X')$ weakly dominates $X \setminus X'$. Let $C' \triangleq \left(X', (\succeq_i|_{\Delta(X')})_{i=1}^n\right)$.

We first claim that there exists $\hat{x} \in \Delta(X')$ such that $I(C, \hat{x}) = 0$. To see this, first note that by feasibility there exists $x \in \Delta(X)$ (not necessarily in $\Delta(X')$) such that $I(C, x) = 0$. Write $x = \sum_{\ell=1}^L \alpha_\ell \cdot x_\ell$ where $\alpha_1, \ldots \alpha_L \in [0, 1]$ and $x_1, \ldots, x_L \in X$. By assumption, for every $\ell = 1, \ldots, L$ there exist $\hat{x}_\ell \in \Delta(X')$ such that $\hat{x}_\ell \succeq_i x_\ell$ for

every $i = 1, \ldots, n$. Therefore, $\hat{x} \triangleq \sum_{\ell=1}^{L} \alpha_\ell \cdot \hat{x}_\ell \in \Delta(X')$ satisfies $\hat{x} \succeq_i x$ for every $i = 1, \ldots, n$. By Pareto monotonicity, $I(C, \hat{x}) = 0$ as claimed.

By feasibility, there exist $\hat{x}' \in \Delta(X')$ such that $I(C', \hat{x}') = 0$. Since by assumption C and C' have the same ideal point and the same point of minimal expectations, by IIA we have that

$$0 \leq I(C, \hat{x}') = I(C, \hat{x}') - I(C, \hat{x}) = I(C', \hat{x}') - I(C', \hat{x}) = -I(C', \hat{x}) \leq 0,$$

and therefore $I\big(C', \hat{x}\big) = 0$. Now, for every $x \in \Delta(X')$, by IIA we have that

$$I(C', x) = I(C', x) - I(C', \hat{x}) = I(C, x) - I(C, \hat{x}) = I(C, x). \qquad \square$$

We next establish that any social inefficiency function that satisfies symmetry and invariance to dominated alternatives (or alternatively, by to Lemmas 2 and 3, any social inefficiency function that satisfies Pareto monotonicity, IIA, and feasibility) also satisfies *neutrality*: it must be invariant to renaming of the alternatives in any context.

Definition 6 (Isomorphic Contexts)**.** Let $C^1 = \big(X^1, (\succeq_i^1)_{i=1}^n\big)$ and $C^2 = \big(X^2, (\succeq_i^2)_{i=1}^n\big)$ be two contexts with the same number of individuals. Let $\phi : X^1 \to X^2$ be a bijection between X^1 and X^2. We extend ϕ in the natural way to be defined over $\Delta(X^1)$, making it a bijection between $\Delta(X^1)$ and $\Delta(X^2)$. We say that ϕ is an *isomorphism* between C^1 and C^2 if for every $x, y \in \Delta(X^1)$ and $i = 1, \ldots, n$, it holds that $x \succeq_i^1 y$ if and only if $\phi(x) \succeq_i^2 \phi(y)$. If there exists an isomorphism between C^1 and C^2, then we say that C^1 and C^2 are *isomorphic*.

Definition 7 (Neutrality)**.** A social inefficiency function I satisfies *neutrality* if for every pair $C^1 = \big(X^1, (\succeq_i^1)_{i=1}^n\big)$ and $C^2 = \big(X^2, (\succeq_i^2)_{i=1}^n\big)$ of isomorphic contexts with isomorphism ϕ between C^1 and C^2 and for every $x \in \Delta(X^1)$, it holds that $I(C^1, x) = I\big(C^2, \phi(x)\big)$.

Lemma 4. *If a social inefficiency function satisfies symmetry and invariance to dominated alternatives, then it satisfies neutrality.*

Finally, the following lemma, beyond its use in our proofs below, paints together with Pareto monotonicity and feasibility the schematic, qualitative picture of where our social inefficiency function attains different values that is illustrated in Figure 1

in Section 4.1: All (possibly randomized) alternatives with social inefficiency zero (a nonempty set by feasibility) are contained (by Pareto monotonicity) in the Pareto frontier, which is a subset (by the below lemma) of the set of all alternatives with finite social inefficiency,[28] which is in turn a subset of the full set of all alternatives. As already noted, each of these containments is strict in some contexts and weak in others.

Lemma 5. *If a social inefficiency function I satisfies Pareto monotonicity, feasibility, and invariance to dominated alternatives, then $I(C,x) < \infty$ for every context $C = \big(X, (\succeq_i)_{i=1}^n\big)$ and every $x \in F_C$.*

4.3.2 Characterization of Alternatives with Infinite Inefficiency

In this section, we characterize the alternatives that have infinite social inefficiency according to any social inefficiency function that satisfies Pareto monotonicity, expected inefficiency, IIA, and feasibility as the alternatives that are suboptimal for one or more frontier-indifferent individuals. One direction of this characterization is given by Lemma 1 from Section 4.1, which we now prove. The other direction is given by a converse to Lemma 1, which we state next.

Proof of Lemma 1. Let $C = \big(X, (\succeq_i)_{i=1}^n\big)$ be a context and let $\hat{i} \in \{1,\dots,n\}$ be frontier-indifferent.

For every $i = 1,\dots,n$, let u_i be a vNM utility representation of $\succeq_i$; choose $u_{\hat{i}}$ so that $u_{\hat{i}}(F_C) = 0$. Let $X' \triangleq X \times \{1,2\}$, and define a context $C' = \big(X', (\succeq'_i)_{i=1}^n\big)$ by defining $\succeq'_i$ for every individual i through the vNM utility representation u'_i that for every $(x,j) \in X'$ satisfies

$$u'_i\big((x,j)\big) \triangleq \begin{cases} j \cdot u_i(x) & i = \hat{i}, \\ u_i(x) & i \neq \hat{i}. \end{cases}$$

For every $j \in \{1,2\}$, let $C^j \triangleq \big(X \times \{j\}, (\succeq'_i|_{\Delta(X\times\{j\})})_{i=1}^n\big)$ and note that C^j is obtained from C' by removing weakly dominated alternatives. Observe that for every $j \in \{1,2\}$, the mapping $x \mapsto (x,j)$ is an isomorphism between C and C^j.

Let $x \in \Delta(X)$ such that $F_C \succ_{\hat{i}} x$. By invariance to dominated alternatives and by neutrality (Lemma 4, where symmetry is by Lemma 2), each applied twice, we have

[28] This set characterized by Lemma 1 and its converse Lemma 6 in Section 4.3.2.

that

$$I\big(C',(x,1)\big) = I\big(C^1,(x,1)\big) = I(C,x) = I\big(C^2,(x,2)\big) = I\big(C',(x,2)\big).$$

Notice that $(x,1) \sim'_i (x,2)$ for every $i \neq \hat{i}$, and furthermore $(x,1) \succ'_{\hat{i}} (x,2)$ (since $u_{\hat{i}}(x) < u_{\hat{i}}(F_C) = 0$). By Pareto monotonicity, therefore, since $I\big(C',(x,1)\big) = I\big(C',(x,2)\big)$ it must be that $I\big(C',(x,1)\big) = I\big(C',(x,2)\big) = \infty$. Therefore, by the above chain of equalities, $I(C,x) = \infty$. □

The following lemma provides a converse for Lemma 1 for social inefficiency functions that also satisfy expected inefficiency. The proof of this lemma, which is relegated to the appendix, is somewhat technically involved.

Lemma 6. *Let I be a social inefficiency function that satisfies Pareto monotonicity, expected inefficiency, IIA, and feasibility, let $C = \big(X,(\succeq_i)_{i=1}^n\big)$ be a context, and let $x \in \Delta(X)$. If for every frontier-indifferent $i \in \{1,\ldots,n\}$ it holds that $F_C \sim_i x$, then $I(C,x) < \infty$.*

4.3.3 Structural Results

In this section, we derive several structural results, culminating in the proof of our main characterization result, Theorem 1.

We have so far used expected inefficiency in a rather weak way: Solely to ascertain the finiteness of the inefficiency of various alternatives. The next lemma builds on top of the characterization from Section 4.3.2, using the full power of expected efficiency.

Lemma 7. *If a social inefficiency function I satisfies Pareto monotonicity, expected inefficiency, IIA, and feasibility, then for every context $C = \big(X,(\succeq_i)_{i=1}^n\big)$ and for every vNM utility representations $u_1,\ldots,u_n$ of $\succeq_1,\ldots,\succeq_n$, there exist constants $c^1,\ldots,c^n > 0$ such that*

$$I(C,x) = \max_{x' \in X} \left\{ \sum_{i=1}^n c^i \cdot \big(u_i(x') - \max_{x'' \in X} u_i(x'')\big) \right\} - \sum_{i=1}^n c^i \cdot \big(u_i(x) - \max_{x'' \in X} u_i(x'')\big)$$

for every $x \in \Delta(X)$. Furthermore, for every $i \in \{1,\ldots,n\}$, it is the case that $c^i = \infty$ if and only if i is frontier-indifferent.[29]

[29] Here and henceforth, we use the convention that $\infty \cdot 0 = 0$.

The next lemma builds on top of Lemma 7, using the full power of IIA, to derive our first characterization—of all social inefficiency functions that satisfy the four axioms used so far: Pareto monotonicity, expected inefficiency, IIA, and feasibility. For better insight into the statement of this lemma, observe that for every context $C = \left(X, (\succeq_i)_{i=1}^n\right)$, the number of frontier-concerned individuals in C can be any integer between 0 and n, *except for* 1.

Lemma 8. *A social inefficiency function I satisfies Pareto monotonicity, expected inefficiency, IIA, and feasibility if and only if for every $n \in \mathbb{N}$ and for every $G \subseteq \{1, \ldots, n\}$ with $|G| \neq 1$ there exist finite positive constants $c^1_{n,G}, \ldots, c^n_{n,G}$ such that*

$$I(C, x) = \max_{x' \in X} \left\{ \sum_{i=1}^{n} c^i_{n,G} \frac{u_i(x') - u_i^{min}}{u_i^{max} - u_i^{min}} \right\} - \sum_{i=1}^{n} c^i_{n,G} \frac{u_i(x) - u_i^{min}}{u_i^{max} - u_i^{min}}$$

for every context with n individuals $C = \left(X, (\succeq_i)_{i=1}^n\right)$ whose set of frontier-concerned individuals is G, for every vNM utility representations $u_1, \ldots, u_n$ of $\succeq_1, \ldots, \succeq_n$, and for every $x \in \Delta(X)$. Furthermore, for every $i \in \{1, \ldots, n\} \setminus G$, the choice of $c^i_{n,G}$ is arbitrary.

The following lemma, which characterize all social inefficiency functions that satisfy all of our axioms except IIP and duplication, follows quite readily from applying the anonymity axiom (which we have not used so far) to the guarantee of Lemma 8. Recall from Section 4.2 that the frontier dimension d_C of a context C is defined as the number of frontier-concerned individuals in C, and recall that this number can be any integer between 0 and the number of individuals in the context, except for 1.

Lemma 9. *A social inefficiency function I satisfies Pareto monotonicity, anonymity, expected inefficiency, IIA, and feasibility if and only if for every $n \in \mathbb{N}$ and for every $d \in \{0, 2, 3, \ldots, n\}$ there exists a constant $0 < c_{n,d} < \infty$ such that $I(C, x) = c_{n,d} \cdot \hat{I}(C, x)$ for every context $C = \left(X, (\succeq_i)_{i=1}^n\right)$ with n individuals and $d_C = d$ and for every $x \in \Delta(X)$. Furthermore, whenever $d = 0$, the choice of $c_{n,d}$ is arbitrary.*

The following lemma leverages the IIP axiom (which we have not used so far) to build on top of Lemma 9 and characterize all social inefficiency functions that satisfy all of our axioms except duplication. In particular, the first part of this lemma characterizes $\hat{I}$ for any fixed $n \notin \{3, 4\}$ as the social inefficiency function for contexts with n individuals that uniquely (up to the unit in which it is globally measured, across

all such contexts) satisfies Pareto monotonicity, anonymity, expected inefficiency, IIA, IIP, and feasibility.

Lemma 10. *A social inefficiency function I satisfies Pareto monotonicity, anonymity, expected inefficiency, IIA, IIP, and feasibility if and only if both of the following hold.*

- *For every $n \in \mathbb{N} \setminus \{3, 4\}$, there exists a constant $0 < c_n < \infty$ such that $I(C, x) = c_n \cdot \hat{I}(C, x)$ for every context $C = \left(X, (\succeq_i)_{i=1}^n\right)$ with n individuals and for every $x \in \Delta(X)$. Furthermore, for $n = 1$, the choice of c_n is arbitrary.*
- *For every $n \in \{3, 4\}$, there exist constants $0 < c_n < \infty$ and $0 < c'_n < \infty$ such that both of the following hold.*
  - *$I(C, x) = c_n \cdot \hat{I}(C, x)$ for every context $C = \left(X, (\succeq_i)_{i=1}^n\right)$ with n individuals and $d_C \neq n - 1$ and for every $x \in \Delta(X)$.*
  - *$I(C, x) = c'_n \cdot \hat{I}(C, x)$ for every context $C = \left(X, (\succeq_i)_{i=1}^n\right)$ with n individuals and $d_C = n - 1$ and for every $x \in \Delta(X)$.*

Finally, in the appendix we leverage duplication (the final axiom that we have not used so far) to build on top of Lemma 10 and prove Theorem 1, characterizing $\hat{I}(C, x)$ as the unique social inefficiency function (up to the unit in which it is globally measured, across all contexts) that satisfies all seven of our axioms.

5 Application to Object Allocation

In this section, we apply our social inefficiency function to a setting in which due to the absence of monetary transfers and outside options, and to the repugnance of measuring using money,[30] interpersonal comparison is challenging, and meaningful precise cardinal values (that is, not up to an affine transformation) for utilities are challenging to justify: the object allocation setting. We explore this setting both from a theoretical perspective and from a computational one. Several proofs are relegated to Appendix B.

[30] See Roth (2007) for a survey of repugnance as a constraint in market design.

5.1 Model

In an object allocation context—to which, in line with the literature on matching theory, we henceforth refer as an *object allocation problem*—there are n individuals and n objects for some $n \in \mathbb{N}$. Each individual has vNM preferences over (lotteries over) the objects, that are strict (no ties) between any two objects. $X = X(n)$ is the set of all the perfect matchings of the n objects to the n individuals, and the individuals' vNM preferences $\succeq = (\succeq_i)_{i=1}^n$ over matchings are induced by their vNM preferences over the objects.

When n is clear from context, we denote by $\mathcal{P}$ the set of all vNM preferences over $\Delta(X) = \Delta\big(X(n)\big)$ that are induced by vNM preferences over (lotteries over) the objects that are strict between any two objects. We call an element of $\mathcal{P}$ a *(vNM) preference*, and an element of $\mathcal{P}^n$ a *profile of (vNM) preferences*, denoting a general element of $\mathcal{P}^n$ by $\succeq = (\succeq_i)_{i=1}^n$. We denote by $\mathcal{R}$ the set of all (strict ordinal) rankings of the n objects. We call an element of $\mathcal{R}$ a *ranking* and an element of $\mathcal{R}^n$ a *profile of rankings*, denoting a general element of $\mathcal{R}^n$ by $R = (R_i)_{i=1}^n$. For a profile of preferences $\succeq = (\succeq_i)_{i=1}^n \in \mathcal{P}^n$, for every $i = 1, \ldots, n$ we denote by $R(\succeq_i) \in \mathcal{R}$ the ranking of the objects that is induced by $\succeq_i$, and denote $R(\succeq) = \big(R(\succeq_i)\big)_{i=1}^n \in \mathcal{R}^n$.

Let $\mathcal{M} = \big\{(X, (\succeq_i)_{i=1}^n) \;\big|\; n \in \mathbb{N} \;\&\; X = X(n) \;\&\; (\succeq_i)_{i=1}^n \in \mathcal{P}^n\big\}$ be the set of all object allocation problems. An *ordinal (matching) mechanism* is a mapping from profiles of rankings $R = (R_1, \ldots, R_n)$ over n objects to distributions in $\Delta(X) = \Delta(X(n))$. Given an object allocation problem $\big(X, (\succeq_i)_{i=1}^n\big) \in \mathcal{M}$, an ordinal mechanism μ induces a game between the n individuals, in which (i) each individual i reports a ranking $R_i \in \mathcal{R}$, (ii) the mechanism outcome $\mu\big((R_i)_{i=1}^n\big) \in \Delta(X)$ is chosen. An ordinal mechanism is *truthful* if for every object allocation problem $\big(X, (\succeq_i)_{i=1}^n\big) \in \mathcal{M}$, for each individual $i = 1, \ldots, n$, *truthtelling* (i.e., playing $R(\succeq_i)$) is a dominant strategy in the induced game. To reduce clutter, we abuse notation by writing $\mu(\succeq)$ instead of $\mu\big(R(\succeq)\big)$.

Random Serial Dictatorship (RSD) (Abdulkadiroğlu and Sönmez, 1998) is a truthful ordinal mechanism that has found wide use in allocating "objects" as varied as housing units (by both governments and colleges), internships, and school slots, just to name a few examples. This mechanism, in uniformly random order over the individuals, matches each individual to the not-yet-matched object that is ranked highest on her reported ranking.

5.2 Theoretical Social Inefficiency Guarantees

In this section, we apply our social inefficiency function to prove an approximate-efficiency result for the widely used Random Serial Dictatorship (RSD) mechanism.[31] We take a robust approach, considering the social inefficiency guarantee of a truthful ordinal mechanism to be the worst social inefficiency (over all object allocation problems) of the truthtelling equilibrium of that mechanism. We prove that it is impossible for any truthful ordinal mechanism to provide a robust social inefficiency guarantee that is less than $\sim 72\%$ of that of RSD.

Theorem 2. *There does not exist any ordinal mechanism μ such that*

$$\sup_{(X,\succeq)\in\mathcal{M}} \hat{I}\big((X,\succeq),\mu(\succeq)\big) < \frac{1}{2\ln 2} \sup_{(X,\succeq)\in\mathcal{M}} \hat{I}\big((X,\succeq),RSD(\succeq)\big).$$

We note that while the statement of Theorem 2 does not require for μ to be truthful, this theorem is mostly interesting when μ is truthful, as only then this theorem considers μ in equilibrium.

We emphasize that we would not have even been able to phrase Theorem 2 had we not had a social inefficiency function that both allows for comparison across contexts (giving meaning to the supremum operation) and is cardinal (giving meaning to the notion of "$\frac{1}{2\ln 2}$ of the social inefficiency"). Indeed, since our axiomatic definition uniquely defines the social inefficiency function $\hat{I}$ up to a single positive multiplicative constant that applies to all contexts, by Theorem 1 we therefore have the following corollary of Theorem 2:

Corollary 1. *For every social inefficiency function I that satisfies Pareto monotonicity, anonymity, expected inefficiency, IIA, IIP, duplication, and feasibility, there does not exist any ordinal mechanism μ such that*

$$\sup_{(X,\succeq)\in\mathcal{M}} I\big((X,\succeq),\mu(\succeq)\big) < \frac{1}{2\ln 2} \sup_{(X,\succeq)\in\mathcal{M}} I\big((X,\succeq),RSD(\succeq)\big).$$

In the remainder of this section, we prove two bounds that together imply Theorem 2: A lower bound on the social inefficiency guarantee of truthful ordinal mechanisms in Theorem 3, and an upper bound on the social inefficiency guarantee of RSD

[31] An approximate-efficiency result is especially of interest for this widely used mechanism in light of a well-known inefficiency result for it (Bogomolnaia and Moulin, 2001).

in Theorem 4. In both bounds, we essentially adapt proofs from Filos-Ratsikas et al. (2014), who bound the "Price of Anarchy" of RSD and other ordinal mechanisms, that is, they bound the asymptotic order of magnitude of the *ratio* (rather than difference) between the sum of utilities at the mechanism output and the optimal sum of utilities. These bounds of Filos-Ratsikas et al. (2014) are for the case of utility functions that are exogenously normalized to have range $[0, 1]$ for each individual; that is, they depend on exogenously supplied cardinal utilities both for specifying an *exogenously given interpersonal comparison* to define welfare, and an *exogenously given "zero"* to give meaning to fractional comparison of welfare. These two features of their analysis—taking the ratio and having exogenous normalization for the utility functions—are common to many papers in computer science; by adapting the proofs of Filos-Ratsikas et al. (2014) with almost no conceptual change, we also demonstrate the potential of a large body of approximation results in computer science to be leveraged to give approximate-optimality results regarding our social inefficiency function, which in contrast to standard ones in computer science, depends only on the ordinal preferences and therefore assumes nothing about interpersonal comparison or an exogenously given, special zero utility value such as that of an outside option or a disagreement point.

5.2.1 Preliminaries

Before proving our bounds, we start with a few useful definitions and lemma. We say that a matching $x \in X = X(n)$ is *ex-post Pareto efficient* with respect to a profile of rankings $R \in \mathcal{R}^n$ if there does not exist $x' \in X$ that, according to R, is ranked weakly higher than x by all individuals and strictly higher than x by some individual. When the object allocation problem $(X, \succeq)$ discussed is clear from context, we say that a matching is *ex-post Pareto efficient* as a shorthand for saying that it is ex-post Pareto efficient with respect to $R(\succeq)$. An ordinal mechanism μ is *ex-post Pareto efficient* if for every profile of rankings R, it is the case that $\mu(R)$ is supported only on matchings that are ex-post Pareto efficient with respect to R. It is well known that RSD is ex-post Pareto efficient.

We emphasize that ex-post Pareto efficiency is defined with respect to Pareto dominance by pure matchings, whereas the Pareto frontier $F_{(X,\succeq)}$ is defined with respect to Pareto dominance by randomized matchings. Nonetheless, for object allocation these two concepts are in fact equivalent.

Lemma 11. *For every $(X, \succeq) \in \mathcal{M}$, every matching $x \in X$ that is ex-post Pareto efficient is on the Pareto frontier $F_{(X,\succeq)}$.*

5.2.2 A Lower Bound on the Social Inefficiency of Truthful Ordinal Mechanisms

We start by proving the following lower bound on the social inefficiency guarantee of every ordinal mechanism.

Theorem 3. *For every $n \in \mathbb{N}$ and every ordinal mechanism μ,*

$$\sup_{\succeq \in \mathcal{P}^n} \hat{I}\big((X, \succeq), \mu(\succeq)\big) \geq \frac{1}{2} - \frac{1}{2n}.$$

The proof of Theorem 3, which is relegated to the appendix, is an adaptation of the proof of Lemma 5 in Filos-Ratsikas et al. (2014), which lower-bounds the order of magnitude of the *ratio* between the sum of utilities at the mechanism output and the optimal sum of utilities in ordinal mechanisms. The proof of Theorem 3 in fact turns out to be slightly simpler than that of their Lemma 5.

5.2.3 An Upper Bound on the Social Inefficiency of Random Serial Dictatorship

To complement Theorem 3, we prove the following *upper* bound on the social inefficiency of the Random Serial Dictatorship (RSD) mechanism.

Theorem 4. *For every $n \in \mathbb{N}$,*

$$\sup_{\succeq \in \mathcal{P}^n} \hat{I}\big((X, \succeq), RSD(\succeq)\big) \leq \ln 2 \approx 0.693.$$

Our analysis is again heavily inspired by that of Filos-Ratsikas et al. (2014). As already noted, there are two main differences between our social inefficiency function and their measure of social inefficiency. First (as is common in computer science), the normalization of their utility functions is exogenous rather than endogenous: Theirs are normalized so that their minimum (over all matchings) is always 0 and their maximum is always 1, which exogenizes also interpersonal comparison and is affected by preferences over objects that would never be awarded by an ex-post Pareto efficient mechanism like RSD. Second (as is also common in computer science), they analyze

the ratio between the optimal and attained sum of utilities rather than the difference between the two, which is again affected by the choice of normalization. As we now show in Lemma 12, the first of these differences is relatively easy to bridge due to our analysis seeking a robust social inefficiency guarantee (this should be contrasted with our analysis in Section 5.3 below). The second of these differences requires local adaptations to their proofs (mostly in the calculations in our proof of Lemma 14 below).

Let $\mathcal{UR}$ be the set of *unit range* utility functions over X, that is, functions from X to $[0,1]$ whose minimum is 0 and whose maximum is 1. For every profile $u = (u_i)_{i=1}^n$ of utility functions over X (which we extend over $\Delta(X)$ by taking expectation) and for every $x \in \Delta(X)$, define $U_u(x) \triangleq \sum_{i=1}^n u_i(x)$ and $I_u(x) \triangleq \max_{x' \in X} U_u(x') - U_u(x)$. For every profile $u = (u_i)_{i=1}^n$ of utility functions over X, for every $i = 1, \ldots, n$ we denote by $R(u_i) \in \mathcal{R}$ the ranking of the objects that is induced by u_i, and denote $R(u) = \big(R(u_i)\big)_{i=1}^n \in \mathcal{R}^n$; furthermore, for every ordinal mechanism μ, to reduce clutter we abuse notation by writing $\mu(u)$ instead of $\mu\big(R(u)\big)$.

Lemma 12. $\sup_{\succeq \in \mathcal{P}^n} \hat{I}\big((X, \succeq), \mu(\succeq)\big) \leq \frac{1}{n} \sup_{u \in \mathcal{UR}^n} I_u\big(\mu(u)\big)$ *for every ex-post Pareto efficient ordinal mechanism* μ *and* $n \in \mathbb{N}$.

Proof. Given a profile $u = (u_i)_{i=1}^n$ of utility functions over X, we say that u_i is *normalized* (with respect to the profile u) if (i) the maximum of u_i is 1, (ii) if u_i is not constant over the entire Pareto frontier (with respect to u), then the minimum of u_i over the Pareto frontier is 0, and (iii) if u_i is constant over the entire Pareto frontier, then the (overall) minimum of u_i is negative. Denote by $\mathcal{N}$ the profiles $u = (u_i)_{i=1}^n$ of utility functions over X such that u_i is normalized for every $i = 1, \ldots, n$. By definition, $\sup_{\succeq \in \mathcal{P}^n} \hat{I}\big((X, \succeq), \mu(\succeq)\big) = \frac{1}{n} \sup_{u \in \mathcal{N}} I_u\big(\mu(u)\big)$. It therefore suffices to show that $\sup_{u \in \mathcal{N}} I_u\big(\mu(u)\big) \leq \sup_{u \in \mathcal{UR}^n} I_u\big(\mu(u)\big)$. Since μ is ex-post Pareto efficient, this follows by induction from the following claim, which we now prove: For every profile $u = (u_i)_{i=1}^n$ such that each u_i is either normalized with respect to u or in $\mathcal{UR}$, for every $i = 1, \ldots, n$, and for every $\varepsilon > 0$, there exists $u_i' \in \mathcal{UR}$ satisfying $R(u_i') = R(u_i)$, such that $I_u(x) \leq I_{(u_i', u_{-i})}(x) + \varepsilon$ for every matching $x \in X$ that is ex-post Pareto efficient with respect to $R(u)$.

If $u_i \in \mathcal{UR}$, then we can set $u_i' = u_i$ and the proof is complete; assume therefore that $u_i \notin \mathcal{UR}$; since u_i is normalized, we therefore have that there exists $x \in X$ such

that $u_i(x) < 0$. Let $L \triangleq \min_{x \in X} u_i(x) < 0$. Define u_i' as follows:

$$u_i'(x) = \begin{cases} \varepsilon + (1-\varepsilon)u_i(x) & u_i(x) \geq 0, \\ \varepsilon\big(1 - u_i(x)/L\big) & u_i(x) < 0. \end{cases}$$

Note that by definition, $R(u_i') = R(u_i)$. Furthermore, since u_i is normalized and there exists $x \in X$ such that $u_i(x) < 0$, we have that $u_i' \in \mathcal{UR}$. Note furthermore that for every matching $x \in X$ that is ex-post Pareto efficient with respect to $R(u)$, by Lemma 11 this matching is on the Pareto frontier (with respect to u), and hence $0 \leq u_i'(x) - u_i(x) \leq \varepsilon$ and therefore $0 \leq U_{(u_i', u_{-i})}(x) - U_u(x) \leq \varepsilon$. Fixing some $x_{max} \in \arg\max_{x \in X} U_u(x)$, for every matching $x \in X$ that is ex-post Pareto efficient with respect to $R(u)$ we therefore have that

$$I_u(x) = U_u(x_{max}) - U_u(x) \leq U_{(u_i', u_{-i})}(x_{max}) - U_{(u_i', u_{-i})}(x) + \varepsilon \leq I_{(u_i', u_{-i})}(x) + \varepsilon. \quad \square$$

The following lemma, along with its proof which is relegated to the appendix, is adapted with only minor adjustments from the analogous Lemma 3 in Filos-Ratsikas et al. (2014) (which proves an analogous result for the ratio, instead of difference, between the realized and optimal welfare). For every $\varepsilon > 0$, let $\mathcal{UR}_\varepsilon$ be the set of utility functions $u \in \mathcal{UR}$ whose range is contained in $[0, \varepsilon) \cup (1-\varepsilon, 1]$.

Lemma 13. $\sup_{u \in \mathcal{UR}^n} I_u\big(\mu(u)\big) = \sup_{u \in \mathcal{UR}_\varepsilon^n} I_u\big(\mu(u)\big)$ *for every ordinal mechanism* μ, $n \in \mathbb{N}$, *and* $\varepsilon > 0$.

The proof of the following lemma builds upon a result from the proof of Lemma 4 in Filos-Ratsikas et al. (2014).

Lemma 14. $\lim_{\varepsilon \to 0} \sup_{u \in \mathcal{UR}_\varepsilon^n} I_u\big(RSD(u)\big) \leq (\ln 2) \cdot n$ *for every* $n \in \mathbb{N}$.

Proof. Let $0 < \varepsilon < 1/n^3$ and let $u = (u_i)_{i=1}^n \in \mathcal{UR}_\varepsilon$. It is enough to prove that $I_u\big(RSD(u)\big) - 2n\varepsilon < (\ln 2) \cdot n$. By definition of $\mathcal{UR}_\varepsilon$, there exists $k \in \{1, \ldots, n\}$ such that $\big|k - \max_{x \in X} U_u(x)\big| < \varepsilon n < 1/n^2$. Filos-Ratsikas et al. (2014) prove that under these conditions, we have (in our notation) that

$$U_u\big(RSD(u)\big) \geq \sum_{l=0}^{\lfloor k/2 \rfloor} \frac{k-2l}{n-l} - n\varepsilon.$$

Since $\max_{x\in X} U_u(x) - n\varepsilon < k$ we have that[32]

$$I_u\big(RSD(u)\big) - 2n\varepsilon < k - \sum_{\ell=0}^{\lfloor k/2\rfloor} \frac{k-2\ell}{n-\ell}.$$

Note that $d\big(\frac{k-2x}{n-x}\big)/dx = \frac{k-2n}{(n-x)^2} < 0$. Therefore,

$$\sum_{\ell=0}^{\lfloor k/2\rfloor} \frac{k-2\ell}{n-\ell} > \int_0^{k/2} \frac{k-2x}{n-x}dx.$$

Combining these two, we have that

$$I_u\big(RSD(u)\big) - 2n\varepsilon < k - \int_0^{k/2} \frac{k-2x}{n-x}dx = \int_0^{k/2} \left(2 - \frac{k-2x}{n-x}\right)dx =$$
$$\int_0^{k/2} \frac{2n-2x-k+2x}{n-x}dx = \int_0^{k/2} \frac{2n-k}{n-x}dx = (2n-k)\int_0^{k/2} \frac{1}{n-x}dx =$$
$$(2n-k)\Big(\ln n - \ln\big(n-\tfrac{k}{2}\big)\Big) = (2n-k)\left(\ln \frac{1}{1-\frac{k}{2n}}\right).$$

Writing $k' = \frac{k}{n} \leq 1$, we then have that

$$I_u\big(RSD(u)\big) - 2n\varepsilon < n\cdot(2-k')\left(\ln \frac{1}{1-\frac{k'}{2}}\right) \leq n\cdot \ln 2. \qquad \square$$

Finally, Theorem 4 follows from the above three lemmas.

Proof of Theorem 4. Combining Lemmas 12 through 14, we have that

$$\sup_{\succeq\in\mathcal{P}^n} \hat{I}\big((X,\succeq), RSD(\succeq)\big) \leq \frac{1}{n}\sup_{u\in\mathcal{UR}^n} I_u\big(RSD(u)\big) = \frac{1}{n}\lim_{\varepsilon\to 0}\sup_{u\in\mathcal{UR}^n_\varepsilon} I_u\big(RSD(u)\big) \leq \ln 2.$$

$\square$

5.3 Computation of Social Inefficiency

Complementing the robust guarantee of the previous section, one might wish to compute the social inefficiency of some matching—such as a realized mechanism output—

[32] At this point, our analysis diverges from that of Filos-Ratsikas et al. (2014), who go on to bound the order of magnitude of the ratio $\frac{U_u(RSD(u))}{\max_{x\in X} U_u(x)}$.

on concrete given preferences. This raises computational concerns, which we discuss and resolve in this section.

To compute our social inefficiency function $\hat{I}$ for an object allocation problem, letting $u_1, \ldots, u_n$ be vNM utility representations of the individuals' preferences, we need to compute (recalling the notation of the definition of $\hat{I}$ from Section 4.1): (1) the per-individual maximum and minimum Pareto-frontier utilities u_i^{max} and u_i^{min}, and (2) the linear translation $\max_{x' \in X} V(C, x')$. Once the per-individual maximum and minimum Pareto-frontier utilities are known, the linear translation can be readily computed using any of the well-known computationally feasible algorithms for the optimal assignment problem (e.g., Kuhn, 1955; Munkres, 1957). For the per-individual maximum and minimum Pareto-frontier utilities, by Lemma 11 we need to compute, for each individual, her most preferred and least preferred objects among those matched to her in ex-post Pareto efficient matchings. While the former is trivially her favorite object, it is at first glance unclear how one might feasibly compute the latter. Indeed, Saban and Sethuraman (2015) show that it is NP-complete (i.e., computationally infeasible) to answer, for a given individual i and object o, the question of whether there exists an ex-post Pareto efficient matching that matches o to i. We nonetheless prove that it is computationally feasible to compute each individual i's least preferred object among those matched to her in ex-post Pareto efficient matchings, which in turn allows us to feasibly compute our social inefficiency function. Our algorithm is presented in Algorithm 1.

Proposition 1. *For every object allocation problem $\left(X, (\succeq_i)_{i=1}^n\right) \in \mathcal{M}$ and individual $\hat{i} \in \{1, \ldots, n\}$, the function call* FindMinParetoMatch$(n, (\succeq_i)_{i=1}^n, \hat{i})$ *returns the object least preferred by $\hat{i}$ among the objects matched to $\hat{i}$ in ex-post Pareto efficient matchings.*

Proposition 1 follows from the following two lemmas.

Lemma 15. *For every object allocation problem $\left(X, (\succeq_i)_{i=1}^n\right) \in \mathcal{M}$, individual $\hat{i}$, and object o, if o is not matched to $\hat{i}$ in any ex-post Pareto efficient matching, then the function call* Test$(n, (\succeq_i)_{i=1}^n, \hat{i}, o)$ *returns False.*

Lemma 16. *For every object allocation problem $\left(X, (\succeq_i)_{i=1}^n\right) \in \mathcal{M}$, individual $\hat{i}$, and object o, if o is the object least preferred by $\hat{i}$ among the objects matched to $\hat{i}$ in ex-post Pareto efficient matchings, then the function call* Test$(n, (\succeq_i)_{i=1}^n, \hat{i}, o)$ *returns True.*

Algorithm 1 Finds the object least preferred by individual $\hat{i}$ among the objects matched to $\hat{i}$ in ex-post Pareto efficient matchings. Runs in time polynomial in n (i.e., computationally feasible).

function FINDMINPARETOMATCH$(n, (\succeq_i)_{i=1}^n, \hat{i})$
  for all object o, from $\hat{i}$'s least preferred to $\hat{i}$'s most preferred, **do**
    if TEST$(n, (\succeq_i)_{i=1}^n, \hat{i}, o)$ **then**
      return o
    end if
  end for
end function
function TEST$(n, (\succeq_i)_{i=1}^n, \hat{i}, o)$
  $V \leftarrow$ all individuals except $\hat{i}$ and all objects except o
  $E \leftarrow$ all individual–object pairs $(i, o') \in V^2$ s.t. $o' \succ_i o$
  $G \leftarrow (V, E)$
  // The next line is calculable in polynomial time (Hopcroft and Karp, 1973).
  $\sigma \leftarrow$ max-cardinality matching in the bipartite graph G
  if $|\sigma| = n - 1$ **then**
    return True
  else
    return False
  end if
end function

6 Discussion

In this paper, we axiomatically define a *cardinal* social welfare function, or more specifically a cardinal social inefficiency function, whose numeric value (rather than induced order) is uniquely defined and comparable across contexts. For achieving this goal, we introduce seven axioms that fully characterize our social inefficiency function. Like any axiomatization, these axioms can obviously be debated upon. Changing one of them might lead to a very different social inefficiency function. We do not claim that our set of axioms is necessarily the only possible one or even the best possible one. Our aim is to demonstrate the applicability of the axiomatic approach for microfounding an economically meaningful quantitative comparison of the inefficiency of different alternatives across contexts.

As shown by Lemma 1, two desirable properties of our social inefficiency function—Pareto monotonicity (even in its weakened form that we use) and invariance to (addition or removal of) dominated alternatives—necessitate allowing our social inefficiency

function to attain the value of infinity on some specific Pareto dominated alternatives when there are frontier-indifferent individuals. In our explicit construction of the social inefficiency function—which normalizes each utility function by scaling it to be in units of the diameter of its range over the Pareto frontier—this has a natural interpretation: Normalizing the utility functions of frontier-indifferent individuals by dividing these functions by zero.[33] The resulting generalized interpretation of utilitarianism uniquely satisfies all of our axioms. That is, we succeed in satisfying the axioms even in the presence of frontier-indifferent individuals by appealing to a degree of freedom not traditionally afforded by the social choice literature (but indeed naturally afforded by our cardinal approach)—that of leaving the societal preference order as unspecified between the specific (Pareto dominated) alternatives that are considered inferior to the Pareto frontier by frontier-indifferent individuals.

Beyond demonstrating the usefulness of our social inefficiency function and the types of theorems that it can fascilitate, an additional purpose of our application section, Section 5, is to demonstrate the relative ease at which existing computer-science analyses can be adapted to yield results about our social inefficiency function. For this reason, we make no attempt to develop further techniques to get a better fraction than $\frac{1}{2\ln 2}$ in Theorem 2. As is common in computer science, approximation theorems allow for progress to be made gradually, and we hence leave the question of tighter bounds in the object allocation application as an interesting open problem.[34]

While we demonstrate the usefulness of our social inefficiency function within noncooperative game theory, we emphasize that its applicability is potentially much broader. Indeed, it can be applied to any setting that features a set of alternatives and preferences of individuals over them. For example, within a bargaining setting, our social inefficiency function essentially frames negotiations as, instead of bargaining with respect to an exogenous disagreement point *à la* Nash, bargaining with respect to an *endogenous* agreement *region* (the Pareto frontier). This might be seen as in line with many negotiation texts for executives that discuss the "Zone of Possible

[33]For a more accurate interpretation as a limit of normalizations, see Footnote 23.

[34]Of course, ours is not the first economics paper that derives an approximation result and leaves potentially tightening that result as an open problem. For example (for approximately optimizing a monopolist's revenue rather than an aggregate, social welfare function), Hart and Nisan (2017) prove that for a two-item monopolist, optimally pricing each item separately guarantees at least 50% of the optimal revenue, a guarantee that was later improved to $\sim 62\%$ by Hart and Reny (2019), with a tight bound on this guarantee still remaining an open problem.

Agreement (ZOPA)" as a key object that frames negotiations.[35]

As noted in the introduction, our analysis in this paper is a demonstration of a potentially larger program of providing a microfoundation for what computer scientists call *approximation theorems*. These theorems, which ascertain that some quantity is at least a specific fraction of some reference quantity, are emblematic of theoretical computer science, and as such, tools and techniques for deriving them have been developed over many decades. While we focus on cardinally measuring (in)efficiency, one might imagine axiomatically defining cardinal measures of other objectives of interest such as (un)fairness, (in)equity, and so forth. Such cardinal measures, of (in)efficiency or otherwise, can also be used beyond bounding the loss in a certain objective in specific alternatives, e.g., for deriving *cardinal* comparative statics: Quantifying not only the direction, but also the magnitude, of change in a certain objective as certain model parameters are changed. Overall, our paper demonstrates a new direction for providing a sound economic foundation for approximation theorems and other quantitative theorems as we enter the second quarter-century of intense EconCS cross-polination.

References

A. Abdulkadiroğlu and T. Sönmez. Random serial dictatorship and the core from random endowments in house allocation problems. *Econometrica*, 66(3):689–701, 1998.

K. J. Arrow. *Social Choice and Individual Values*. Yale University Press, 1951.

K. J. Arrow. *Social Choice and Individual Values*. John Wiley & Sons, second edition, 1963.

U. Ayesta, O. Brun, and B. J. Prabhu. Price of anarchy in non-cooperative load balancing. In *Proceedings of the 29th IEEE Conference on Computer Communications (INFOCOM)*, pages 1–5, 2010.

A. Bogomolnaia and H. Moulin. A new solution to the random assignment problem. *Journal of Economic Theory*, 100(2):295–328, 2001.

[35] See, e.g., https://online.hbs.edu/blog/post/understanding-zopa. We thank Yuval Gonczarowski for introducing us to this notion.

D. Braess. Über ein paradoxon aus der verkehrsplanung. *Unternehmensforschung*, 12:258–268, 1968.

I. Caragiannis, D. Kurokawa, H. Moulin, A. D. Procaccia, N. Shah, and J. Wang. The unreasonable fairness of maximum Nash welfare. *ACM Transactions on Economics and Computation*, 7(3), 2019.

F. DeMeyer and C. R. Plott. A welfare function using "relative intensity" of preference. *The Quarterly Journal of Economics*, 85(1):179–186, 1971.

A. Dhillon and J.-F. Mertens. Relative utilitarianism. *Econometrica*, 67(3):471–498, 1999.

A. Filos-Ratsikas, S. K. S. Frederiksen, and J. Zhang. Social welfare in one-sided matchings: Random priority and beyond. In *Proceedings of the 7th International Symposium on Algorithmic Game Theory (SAGT)*, pages 1–12, 2014.

M. Fleurbaey and K. Tadenuma. Universal social orderings: An integrated theory of policy evaluation, inter-society comparisons, and interpersonal comparisons. *Review of Economic Studies*, 81(3):1071–1101, 2014.

M. Fleurbaey and S. Zuber. Fair utilitarianism. *American Economic Journal: Microeconomics*, 13(2):370–401, 2021.

S. Ghosh and R. Hassin. Inefficiency in stochastic queueing systems with strategic customers. *European Journal of Operational Research*, 295(1):1–11, 2021.

G. T. Guilbaud. Les théories de l'intérêt général et le problème logique de l'agrégation. *Economie appliquée*, 5(4):501–584, 1952.

P. J. Hammond. Utility invariance in non-cooperative games. In *Advances in Public Economics: Utility, Choice and Welfare: A Festschrift for Christian Seidl*, pages 31–50. Springer, 2005.

J. C. Harsanyi. Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *Journal of Political Economy*, 63(4):309–321, 1955.

S. Hart and N. Nisan. Approximate revenue maximization with multiple items. *Journal of Economic Theory*, 172:313–347, 2017.

S. Hart and P. J. Reny. The better half of selling separately. *ACM Transactions on Econonomics and Computation*, 2019.

J. Hartline, D. Hoy, and S. Taggart. Price of anarchy for auction revenue. In *Proceedings of the 15th ACM Conference on Economics and Computation (EC)*, page 693–710, 2014.

R. Hoeksma and M. Uetz. The price of anarchy for utilitarian scheduling games on related machines. *Discrete Optimization*, 31:29–39, 2019.

J. E. Hopcroft and R. M. Karp. An $n^{5/2}$ algorithm for maximum matchings in bipartite graphs. *SIAM Journal on Computing*, 2(4):225–231, 1973.

D. Hoy, S. Taggart, and Z. Wang. A tighter welfare guarantee for first-price auctions. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, pages 132–137, 2018.

Y. Jin and P. Lu. First price auction is $1-1/e^2$ efficient. *Journal of the ACM*, 70(5), 2023.

E. Kalai and M. Smorodinsky. Other solutions to Nash's bargaining problem. *Econometrica*, pages 513–518, 1975.

M. Kaneko and K. Nakamura. The Nash social welfare function. *Econometrica*, 47 (2):423–435, 1979.

E. Koutsoupias and C. Papadimitriou. Worst-case equilibria. In *Proceedings of the 16th Annual Symposium on Theoretical Aspects of Computer Science (STACS)*, pages 404–413, 1999.

H. W. Kuhn. The Hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 2(1–2):83–97, 1955.

R. D. Luce and H. Raiffa. *Games and Decisions: Introduction and Critical Survey*. John Wiley & Sons, New York, 1957.

E. Maskin. A theorem on utilitarianism. *The Review of Economic Studies*, 45(1): 93–96, 1978.

J. Munkres. Algorithms for the assignment and transportation problems. *Journal of the Society for Industrial and Applied Mathematics*, 5(1):32–38, 1957.

B. Nalebuff. A perspective-invariant approach to Nash bargaining. *Management Science*, 67(1):577–593, 2021.

J. Nash. Two-person cooperative games. *Econometrica*, 21(1):128–140, 1953.

G. Perakis and G. Roels. The price of anarchy in supply chains: Quantifying the efficiency of price-only contracts. *Management Science*, 53(8):1249–1268, 2007.

P. G. Piacquadio. A fairness justification of utilitarianism. *Econometrica*, 85(4): 1261–1276, 2017.

H. Raiffa. Arbitration schemes for generalized two-person games. In H. W. Kuhn and A. W. Tucker, editors, *Contributions to the Theory of Games, Volume II*, pages 361–388. Princeton University Press, 1953.

K. W. Roberts. Interpersonal comparability and social choice theory. *The Review of Economic Studies*, 47(2):421–439, 1980.

A. E. Roth. Independence of irrelevant alternatives, and solutions to Nash's bargaining problem. *Journal of Economic Theory*, 16(2):247–251, 1977.

A. E. Roth. Repugnance as a constraint on markets. *Journal of Economic Perspectives*, 21(3):37–58, 2007.

T. Roughgarden and E. Tardos. How bad is selfish routing? *Journal of the ACM*, 49 (2):236–259, 2002.

D. Saban and J. Sethuraman. The complexity of computing the random priority allocation matrix. *Mathematics of Operations Research*, 40(4):1005–1014, 2015.

A. Sen. Interpersonal aggregation and partial comparability. *Econometrica*, pages 393–409, 1970.

V. Syrgkanis and E. Tardos. Composable and efficient mechanisms. In *Proceedings of the 45th Annual ACM Symposium on Theory of Computing (STOC)*, pages 211–220, 2013.

J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, 1953.

A Proofs Omitted from Section 4

Proof of Lemma 2. Let $C = \big(X, (\succeq_i)_{i=1}^n\big)$ be a context and let $x, y \in \Delta(x)$ such that $x \sim_i y$ for every $i = 1, \ldots, n$. By Pareto monotonicity (applied twice) we have both that $I(C, x) \leq I(C, y)$ and that $I(C, x) \geq I(C, y)$, and hence $I(C, x) = I(C, y)$, and so I satisfies symmetry. □

Proof of Lemma 4. Let $C^1 = \big(X^1, (\succeq_i^1)_{i=1}^n\big)$ and $C^2 = \big(X^2, (\succeq_i^2)_{i=1}^n\big)$ be two isomorphic contexts and let ϕ be an isomorphism between C^1 and C^2.

We first prove that for the special case in which $X^1 \cap X^2 = \emptyset$, it indeed holds that $I(C^1, x) = I\big(C^2, \phi(x)\big)$ for every $x \in \Delta(X^1)$. Let $X \triangleq X^1 \cup X^2$ and define a context $C = \big(X, (\succeq_i)_{i=1}^n\big)$ by defining $\succeq_i$ for every i using the isomorphism ϕ as follows: Let $x, y \in \Delta(X)$. By definition of X, there exist $\alpha \in [0, 1]$, $x^1 \in \Delta(X^1)$, and $x^2 \in \Delta(X^2)$ such that $x = \alpha \cdot x^1 + (1 - \alpha) \cdot x^2$. Similarly, there exist $\beta \in [0, 1]$, $y^1 \in \Delta(X^1)$, and $y^2 \in \Delta(X^2)$ such that $y = \beta \cdot y^1 + (1 - \beta) \cdot y^2$. We define $\succeq_i$ so that $x \succeq_i y$ if and only if $\alpha \cdot \phi(x^1) + (1 - \alpha) \cdot x^2 \succeq_i^2 \beta \cdot \phi(y^1) + (1 - \beta) \cdot y^2$. (Equivalently, if and only if $\alpha \cdot x^1 + (1-\alpha) \cdot \phi^{-1}(x^2) \succeq_i^1 \beta \cdot y^1 + (1-\beta) \cdot \phi^{-1}(y^2)$.) Note that $C^1 = \big(X^1, (\succeq_i|_{\Delta(X^1)})_{i=1}^n\big)$ and that $C^2 = \big(X^2, (\succeq_i|_{\Delta(X^2)})_{i=1}^n\big)$.

Note that for every $x \in \Delta(X^1)$, we have that $x \sim_i \phi(x)$ for every $i = 1, \ldots, n$. By symmetry, we therefore have that $I(C, x) = I\big(C, \phi(x)\big)$. Therefore, by invariance to dominated alternatives (applied twice), we have that

$$I(C^1, x) = I(C, x) = I\big(C, \phi(x)\big) = I\big(C^2, \phi(x)\big),$$

completing the proof for the special case in which $X^1 \cap X^2 = \emptyset$.

For the general case in which possibly $X^1 \cap X^2 \neq \emptyset$, let X^3 be a set such that $|X^3| = |X^1|$ (and hence $|X^3| = |X^2|$) and $(X^1 \cup X^2) \cap X^3 = \emptyset$. Since $|X^3| = |X^1|$, there exists a bijection ψ between X^3 and X^1. We extend ψ in the natural way to be defined over $\Delta(X^3)$, making it a bijection between $\Delta(X^3)$ and $\Delta(X^1)$. For every $i = 1, \ldots, n$, we define $\succeq_i^3$ over $\Delta(X^3)$ using the bijection ψ as follows: Let $x, y \in \Delta(X^3)$. We define $\succeq_i^3$ so that $x \succeq_i^3 y$ if and only if $\psi(x) \succeq_i^1 \psi(y)$. Note that ψ is an isomorphism between $C^3 \triangleq \big(X^3, (\succeq_i^3)_{i=1}^n\big)$ and C^1, and therefore $\phi \circ \psi$ is an isomorphism between C^3 and C^2. For every $x \in \Delta(X^1)$, we therefore have by the above special case (applied twice) that $I(C^1, x) = I\big(C^3, \psi^{-1}(x)\big) = I\big(C^2, \phi(x)\big)$, and so I satisfies neutrality. □

Proof of Lemma 5. We first prove the lemma for contexts in which every individual $i \in \{1, \ldots, n\}$ is frontier-indifferent. By feasibility, there exists $\hat{x} \in \Delta(X)$ such that $I(C, \hat{x}) = 0$. For every $x \in F_C$, since by assumption $x \sim_i \hat{x}$ for all $i = 1, \ldots, n$, by symmetry (Lemma 2) we have that $I(C, x) = 0 < \infty$, concluding the proof of the lemma for contexts in which every $i \in \{1, \ldots, n\}$ is frontier-indifferent.

We move on to prove the lemma for all other contexts. By invariance to dominated alternatives, it suffices to prove the lemma for such contexts under the assumption that there exists $x_0 \in X$ such that $x_0 \sim_i \min_{\succ_i} F_C$ for all $i = 1, \ldots, n$. Let $x \in F_C$. Note that $x \succeq_i x_0$ for all $i = 1, \ldots, n$, with $x \succ_i x_0$ for at least one (frontier-concerned) i. Therefore, by Pareto monotonicity either $I(C, x) < I(C, x_0)$, in which case $I(C, x) < \infty$ and we're done, or there exists $w \in \Delta(X)$ with $w \succeq_i x$ for all $i = 1, \ldots, n$ such that $I(C, w) < \infty$. In the latter case, since $x \in F_C$, we have that $w \sim_i x$ for all $i = 1, \ldots, n$, and hence by symmetry (Lemma 2), $I(C, x) < \infty$ as well. □

Proof of Lemma 6. Let $C = \big(X, (\succeq_i)_{i=1}^n\big)$ be a context. If all $i = 1, \ldots, n$ are frontier-indifferent, then every $x \in \Delta(X)$ such that $F_C \sim_i x$ for every frontier-indifferent $i \in \{1, \ldots, n\}$ satisfies $x \in F_C$, and by Lemma 5 (where invariance to dominated alternatives is by Lemma 3) we have that $I(C, x) < \infty$. Assume henceforth, therefore, that not all i are frontier-indifferent.

For every i, let u_i be a vNM utility representation of $\succeq_i$. For every $x \in \Delta(X)$, we write $u(x) \triangleq \big(u_1(x), \ldots, u_n(x)\big) \in \mathbb{R}^n$, and for a set $X' \subseteq \Delta(X)$, we write $u(X') \triangleq \big\{u(x) \;\big|\; x \in X'\big\}$. For a set $A \subset \mathbb{R}^n$, we write $\mathrm{Conv}(A)$ to denote the convex hull of the points in A. Let $u^1, \ldots, u^L \in \mathbb{R}^n$ be the extreme points of $u(F_C)$, that is, $u(F_C) = \mathrm{Conv}\big(\{u^1, \ldots, u^L\}\big)$ and $u^\ell \notin \mathrm{Conv}\big(\{u^1, \ldots, u^L\} \setminus \{u^\ell\}\big)$ for every $\ell = 1, \ldots, L$. Since not all i are frontier-indifferent, we have that $L \geq 2$. For every $\ell = 1, \ldots, L$, let $x_\ell \in X$ such that $u(x_\ell) = u^\ell$.

By neutrality (Lemma 4, where symmetry is by Lemma 2 and invariance to dominated alternatives is by Lemma 3), assume without loss of generality that $X \cap [0, 1/2] = \emptyset$. For every $i = 1, \ldots, n$, let $u_i' : X \cup [0, 1/2] \to \mathbb{R}$ be the vNM utility function such that $u_i'(x) = u_i(x)$ for every $x \in X$, and for every $\varepsilon \in [0, 1/2]$ it holds that $u_i'(\varepsilon) = \frac{1}{L}\sum_{\ell=1}^L u_i^\ell$ if i is frontier-indifferent, and $u_i'(\varepsilon) = \frac{1}{L}\sum_{\ell=1}^L u_i^\ell + \varepsilon$ otherwise; let $\succeq_i'$ be the vNM preferences represented by u_i'. For every $\varepsilon \in [0, 1/2]$, let $u^\varepsilon \triangleq u'(\varepsilon) = \big(u_1'(\varepsilon), \ldots, u_n'(\varepsilon)\big) \in \mathbb{R}^n$. For every $\varepsilon \in [0, 1/2]$, let $X^\varepsilon \triangleq X \cup \{\varepsilon\}$, and denote $C^\varepsilon \triangleq \big(X^\varepsilon, (\succeq_i' |_{\Delta(X^\varepsilon)})_{i=1}^n\big)$.

Let $u^{max} \triangleq \big(\max_{x\in X} u_1(x), \ldots, \max_{x\in X} u_n(x)\big) \in \mathbb{R}^n$. We claim that there exists $\varepsilon > 0$ such that $x_1, \ldots, x_L \in F_{C^\varepsilon}$ and $u^\varepsilon \leq u^{max}$.

Let $\hat{\ell} \in \{1, \ldots, L\}$. We observe that $u^{\hat{\ell}} \notin \mathrm{Conv}\big(\{u^0, u^1, \ldots, u^L\} \setminus \{u^{\hat{\ell}}\}\big)$.[36] Indeed, assume the converse for contradiction, then there exist $\alpha_0, \ldots, \alpha_{\hat{l}-1}, \alpha_{\hat{l}+1}, \ldots, \alpha_L \in [0,1]$ with $\sum_{\substack{\ell=0 \\ \ell\neq\hat{\ell}}}^{L} \alpha_\ell = 1$ such that $u^{\hat{\ell}} = \sum_{\substack{\ell=0 \\ \ell\neq\hat{\ell}}}^{L} \alpha_\ell u^\ell = \frac{\alpha_0}{L} u^{\hat{\ell}} + \sum_{\substack{\ell=1 \\ \ell\neq\hat{\ell}}}^{L} (\alpha_\ell + \frac{\alpha_0}{L}) u^\ell$. Therefore, since $L \geq 2$, we have that $u^{\hat{\ell}} = \sum_{\substack{\ell=1 \\ \ell\neq\hat{\ell}}}^{L} \frac{1}{1-\frac{\alpha_0}{L}} (\alpha_\ell + \frac{\alpha_0}{L}) u^\ell$, and note that $\sum_{\substack{\ell=1 \\ \ell\neq\hat{\ell}}}^{L} \frac{1}{1-\frac{\alpha_0}{L}} (\alpha_\ell + \frac{\alpha_0}{L}) = 1$, contradicting the fact that $u^{\hat{\ell}} \notin \mathrm{Conv}\big(\{u^1, \ldots, u^L\} \setminus \{u^{\hat{\ell}}\}\big)$. So indeed $u^{\hat{\ell}} \notin \mathrm{Conv}\big(\{u^0, u^1, \ldots, u^L\} \setminus \{u^{\hat{\ell}}\}\big)$.

Since $u^0, u^1, \ldots, u^{\hat{\ell}-1}, u^{\hat{\ell}+1}, \ldots, u^L \in u\big(\Delta(X)\big)$ and $u^{\hat{\ell}} \in u(F_C)$, we have that $\mathrm{Conv}\big(\{u^0, u^1, \ldots, u^L\} \setminus \{u^{\hat{\ell}}\}\big) \cap \{a \in \mathbb{R}^n \mid a > u^{\hat{\ell}}\} = \emptyset$, where for $a, b \in \mathbb{R}^n$, we write $a > b$ to denote that $a \geq b$ and $a \neq b$. Therefore, and since $u^{\hat{\ell}} \notin \mathrm{Conv}\big(\{u^0, u^1, \ldots, u^L\} \setminus \{u^{\hat{\ell}}\}\big)$, we have that $\mathrm{Conv}\big(\{u^0, u^1, \ldots, u^L\} \setminus \{u^{\hat{\ell}}\}\big) \cap \{a \in \mathbb{R}^n \mid a \geq u^{\hat{\ell}}\} = \emptyset$. Since each of these two disjoint sets is closed and one of them is bounded, we have that the distance between them is positive: $d\big(\mathrm{Conv}(\{u^0, u^1, \ldots, u^L\} \setminus \{u^{\hat{\ell}}\}), \{a \in \mathbb{R}^n \mid a \geq u^{\hat{\ell}}\}\big) > 0$. Since $d\big(\mathrm{Conv}(\{u^\varepsilon, u^1, \ldots, u^L\} \setminus \{u^{\hat{\ell}}\}), \{a \in \mathbb{R}^n \mid a \geq u^{\hat{\ell}}\}\big)$ is a continuous function of ε, there exists $\varepsilon_{\hat{\ell}} > 0$ such that $d\big(\mathrm{Conv}(\{u^\varepsilon, u^1, \ldots, u^L\} \setminus \{u^{\hat{\ell}}\}), \{a \in \mathbb{R}^n \mid a \geq u^{\hat{\ell}}\}\big) > 0$ for every $0 < \varepsilon \leq \varepsilon_{\hat{\ell}}$. For every such ε, therefore $\big\{u'(y) \;\big|\; y \in \Delta(X^\varepsilon) \;\&\; \forall i : y \succeq_i x^{\hat{\ell}} \;\&\; \exists i : y \succ_i x^{\hat{\ell}}\big\} = \mathrm{Conv}\big(\{u^\varepsilon, u^1, \ldots, u^L\}\big) \cap \{a \in \mathbb{R}^n \mid a > u^{\hat{\ell}}\} = \emptyset$, and hence $x^{\hat{\ell}} \in F_{C^\varepsilon}$. Observe that for every frontier-concerned i, it is the case that $u_i^0 < u_i^{max}$. Therefore, there exists $\varepsilon_0 > 0$ such that $u^\varepsilon \leq u^{max}$ for every $0 < \varepsilon < \varepsilon_0$. Taking $\varepsilon \triangleq \min\{\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_L\} > 0$, we therefore have that $x_1, \ldots, x_L \in F_{C^\varepsilon}$ and $u^\varepsilon \leq u^{max}$, as claimed.

Since $x_1, \ldots, x_L \in F_{C^\varepsilon}$, by Lemma 5 (where invariance to dominated alternatives is by Lemma 3) we have that $I(C^\varepsilon, x_\ell) < \infty$ for all $\ell = 1, \ldots, L$. Therefore, by expected inefficiency, $I\big(C^\varepsilon, \frac{1}{L}\sum_{\ell=1}^{L} x_\ell\big) < \infty$ and hence by symmetry (Lemma 2), $I(C^\varepsilon, 0) < \infty$. Since all of the extreme points of $u(F_C)$ are in $u(F_{C^\varepsilon})$ and since $u^\varepsilon \geq u^0 \geq \big(\min_{x\in F_C} u_1(x), \ldots, \min_{x\in F_C} u_n(x)\big)$, we have that C and C^ε have the same point of minimal expectations. Since all of the extreme points of $u(F_C)$ are in $u(F_{C^\varepsilon})$ and since $u^\varepsilon \leq u^{max}$, these two contexts also have the same ideal point.

We now prove that for every $x \in \Delta(X)$ such that $F_C \sim_i x$ for every frontier-indifferent i, it holds that $I(C^\varepsilon, x) < \infty$. Indeed, since $\varepsilon \sim_i' 0$ and $x \sim_i' 0$ for every

[36] We emphasize that in this expression, one should not confuse the definition of u^0 (which equals u^ε for $\varepsilon = 0$) with the definition of $u^1, \ldots, u^L$ (which equal $u(x_1), \ldots, u(x_L)$). We use similar symbols for these differently defined points for ease of notation in the part of the proof that now commences.

frontier-indifferent i, and $\varepsilon \succ_i' 0$ for every other i, there exists $0 < \alpha < 1$ such that $\alpha \cdot \varepsilon + (1-\alpha) \cdot x \succeq_i' 0$ for every $i = 1, \ldots, n$. By Pareto monotonicity, we have that both $I(C^\varepsilon, \varepsilon) < I(C^\varepsilon, 0) < \infty$ and $I(C^\varepsilon, \alpha \cdot \varepsilon + (1-\alpha) \cdot x) \leq I(C^\varepsilon, 0) < \infty$. By both of these and by expected inefficiency, we have that $I(C^\varepsilon, x) < \infty$.

Let $x \in \Delta(X)$ such that $F_C \sim_i x$ for every frontier-indifferent i; we need to prove that $I(C, x) < \infty$. By feasibility, there exists $\hat{x} \in \Delta(X)$ such that $I(C, \hat{x}) = 0 < \infty$. By Lemma 1 (where invariance to dominated alternatives is by Lemma 3), we have that $F_C \sim_i \hat{x}$ for every frontier-indifferent i. By the previous paragraph, both $I(C^\varepsilon, x) < \infty$ and $I(C^\varepsilon, \hat{x}) < \infty$. By IIA, we have that $I(C, x) - I(C, \hat{x}) = I(C^\varepsilon, x) - I(C^\varepsilon, \hat{x})$. Since we have shown that three of the four social inefficiencies in this equation are finite, we therefore have that the fourth social inefficiency, $I(C, x)$, is finite as well. □

Proof of Lemma 7. Let $C = \left(X, (\succeq_i)_{i=1}^n\right)$ be a context and let $u_1, \ldots, u_n$ be vNM utility representations of $\succeq_1, \ldots, \succeq_n$. Let $X' = \{x \in X \mid \forall \text{frontier-indifferent } i : F_C \sim_i x\}$. By Lemma 6, $I(C, x) < \infty$ for all $x \in \Delta(X')$. By symmetry (Lemma 2), $I(C, x)$ depends on x only through the values $u_1(x), \ldots, u_n(x)$. By expected inefficiency and since all the u_i functions are linear, $I(C, \cdot)$ can therefore be written over $\Delta(X')$ in the form $I(C, x) = a - \sum_i c^i \cdot \left(u_i(x) - \max_{x'' \in X} u_i(x'')\right)$ for some finite constants $a, (c^i)_{\text{frontier-concerned } i}$ where the sum is over all frontier-concerned i. While this representation may not be unique, by Pareto monotonicity it is possible to choose such a representation where $c^i > 0$ for every frontier-concerned i.

For every frontier-indifferent i, let $c^i \triangleq \infty$. By Lemma 1 (where invariance to dominated alternatives is by Lemma 3), we have that $I(C, x) = a - \sum_{i=1}^n c^i \cdot \left(u_i(x) - \max_{x'' \in X} u_i(x'')\right)$ for every $x \in \Delta(X)$.

Since by definition of a social inefficiency function, $I(C, x) \geq 0$ for every $x \in X$, we have that $a \geq \max_{x' \in X} \left\{\sum_{i=1}^n c^i \cdot (u_i(x') - \max_{x'' \in X} u_i(x''))\right\}$. By feasibility, there exists $x \in \Delta(X)$ such that $a - \sum_{i=1}^n c^i \cdot \left(u_i(x) - \max_{x'' \in X} u_i(x'')\right) = 0$, and therefore $a \leq \max_{x' \in X} \left\{\sum_{i=1}^n c^i \cdot (u_i(x') - \max_{x'' \in X} u_i(x''))\right\}$. By both of these, $a = \max_{x' \in X} \left\{\sum_{i=1}^n c^i \cdot (u_i(x') - \max_{x'' \in X} u_i(x''))\right\}$. □

Proof of Lemma 8. The "if" direction is straightforward to check. We prove the "only if" direction.

Let $n \in \mathbb{N}$, and let $G \subseteq \{1, \ldots, n\}$ such that $|G| \neq 1$. Let $\hat{X}_{n,G} \triangleq G \cup \{0\}$, and define a context $\hat{C}_{n,G} \triangleq \left(\hat{X}_{n,G}, (\hat{\succeq}_i)_{i=1}^n\right)$ by defining $\hat{\succeq}_i$ for every individual i through

the vNM utility representation $\hat{u}_i$ that for every $i = 1, \ldots, n$ and $x \in \hat{X}_{n,G}$ satisfies

$$\hat{u}_i(x) \triangleq \begin{cases} 1 & x = i, \\ 0 & x \neq i. \end{cases}$$

Note that the set of frontier-concerned individuals in $\hat{C}_{n,G}$ is G. By Lemma 7, there exist constants $c^1_{\hat{C}_{n,G}}, \ldots, c^n_{\hat{C}_{n,G}} > 0$ such that $I(\hat{C}_{n,G}, x) = \max_{x' \in \hat{X}_{n,G}} \left\{ \sum_{i=1}^n c^i_{\hat{C}_{n,G}} \cdot (\hat{u}_i(x') - \max_{x'' \in \hat{X}_{n,G}} \hat{u}_i(x'')) \right\} - \sum_{i=1}^n c^i_{\hat{C}_{n,G}} \cdot (\hat{u}_i(x) - \max_{x'' \in \hat{X}_{n,G}} \hat{u}_i(x''))$ for every $x \in \Delta(\hat{X}_{n,G})$. For every $i \in G$, denote $c^i_{n,G} \triangleq c^i_{\hat{C}_{n,G}}$, which by Lemma 7 is finite. For every $i \in \{1, \ldots, n\} \setminus G$, let $0 < c^i_{n,G} < \infty$ be arbitrary.

Let $C = \left(X, (\succeq_i)_{i=1}^n\right)$ be a context with n individuals, in which the set of frontier-concerned individuals is G. By neutrality (Lemma 4, where symmetry is by Lemma 2 and invariance to dominated alternatives is by Lemma 3), assume without loss of generality that $X \cap \hat{X}_{n,G} = \emptyset$.

Let $X' \triangleq X \cup \hat{X}_{n,G}$ and define a context $C' \triangleq \left(X', (\succeq'_i)_{i=1}^n\right)$ where for every i, the preference $\succeq'_i$ is obtained from $\succeq_i$ by stipulating for every $x \in \hat{X}_{n,G}$ that $x \sim'_i \max_{\succeq_i} F_C$ if $\hat{u}_i(x) = 1$ and $x \sim'_i \min_{\succeq_i} F_C$ if $\hat{u}_i(x) = 0$. Note that $\hat{C}_{n,G} = (\hat{X}_{n,G}, (\succeq'_i|_{\Delta(\hat{X}_{n,G})})_{i=1}^n)$ as well as that $\hat{C}_{n,G}$ and C' have the same ideal point and the same point of minimal expectations.

For every $i = 1, \ldots, n$, choose a vNM utility representation u'_i of $\succeq'_i$ that coincides with $\hat{u}_i$ over $\hat{X}_{n,G}$. By Lemma 7, there exist constants $c^1_{C'}, \ldots, c^n_{C'} > 0$ such that $I(C', x) = \max_{x' \in X'} \left\{ \sum_{i=1}^n c^i_{C'} \cdot (u'_i(x') - \max_{x'' \in X'} u'_i(x'')) \right\} - \sum_{i=1}^n c^i_{C'} \cdot \left(u'_i(x) - \max_{x'' \in X'} u'_i(x'')\right)$ for every $x \in \Delta(X')$. By Lemma 7, we have that $c^i_{C'}$ and $c^i_{\hat{C}_{n,G}}$ are finite for every $i \in G$. Therefore, $I(C', x)$ and $I(\hat{C}_{n,G}, x)$ are finite for every $x \in \hat{X}_{n,G}$. Therefore, for every $i \in G$ by IIA we have that

$$c^i_{C'} = I(C', 0) - I(C', i) = I(\hat{C}_{n,G}, 0) - I(\hat{C}_{n,G}, i) = c^i_{\hat{C}_{n,G}} = c^i_{n,G}.$$

For every $i \notin G$, by Lemma 7 we have that $c^i_{C'} = \infty$.

For every $i = 1, \ldots, n$, let u_i be some vNM utility representation of $\succeq_i$. By invariance to dominated alternatives (Lemma 3), for every $x \in \Delta(X)$ we have that

$$I(C, x) = I(C', x) =$$

$$\max_{x'\in X}\left\{\sum_{i=1}^{n} c^i_{C'}\cdot\left(u'_i(x')-\max_{x''\in X}u'_i(x'')\right)\right\}-\sum_{i=1}^{n} c^i_{C'}\cdot\left(u'_i(x)-\max_{x''\in X}u'_i(x'')\right)=$$

$$\max_{x'\in X}\left\{\sum_{i\in G} c^i_{n,G}\cdot\left(u'_i(x')-\max_{x''\in X}u'_i(x'')\right)+\sum_{i\notin G}\infty\cdot\left(u'_i(x')-\max_{x''\in X}u'_i(x'')\right)\right\}-$$

$$\left(\sum_{i\in G} c^i_{n,G}\cdot\left(u'_i(x)-\max_{x''\in X}u'_i(x'')\right)+\sum_{i\notin G}\infty\cdot\left(u'_i(x)-\max_{x''\in X}u'_i(x'')\right)\right)=$$

$$\max_{x'\in X}\left\{\sum_{i\in G} c^i_{n,G}\cdot\frac{u'_i(x')-{u'_i}^{min}}{{u'_i}^{max}-{u'_i}^{min}}+\sum_{i\notin G}\frac{u'_i(x')-{u'_i}^{min}}{{u'_i}^{max}-{u'_i}^{min}}\right\}-$$

$$\left(\sum_{i\in G} c^i_{n,G}\cdot\frac{u'_i(x)-{u'_i}^{min}}{{u'_i}^{max}-{u'_i}^{min}}+\sum_{i\notin G}\frac{u'_i(x)-{u'_i}^{min}}{{u'_i}^{max}-{u'_i}^{min}}\right)=$$

$$\max_{x'\in X}\left\{\sum_{i=1}^{n} c^i_{n,G}\cdot\frac{u'_i(x')-{u'_i}^{min}}{{u'_i}^{max}-{u'_i}^{min}}\right\}-\sum_{i=1}^{n} c^i_{n,G}\cdot\frac{u'_i(x)-{u'_i}^{min}}{{u'_i}^{max}-{u'_i}^{min}}=$$

$$\max_{x'\in X}\left\{\sum_{i=1}^{n} c^i_{n,G}\cdot\frac{u_i(x')-u_i^{min}}{u_i^{max}-u_i^{min}}\right\}-\sum_{i=1}^{n} c^i_{n,G}\cdot\frac{u_i(x)-u_i^{min}}{u_i^{max}-u_i^{min}}.\qquad\square$$

Lemma 9. The "if" direction follows by construction of $\hat{I}$. We prove the "only if" direction.

Let $n\in\mathbb{N}$ and $d\in\{0,2,3,\ldots,n\}$. For every $G\subseteq\{1,\ldots,n\}$ with $|G|=d$ and for every $i\in G$, let $0<c^i_{n,G}<\infty$ be the constant guaranteed to exist by Lemma 8. If $d>0$, then let $c_{n,d}\triangleq n\cdot c^1_{n,\{1,\ldots,d\}}$. Otherwise, let $0<c_{n,d}<\infty$ be arbitrary.

Let $C=\left(X,(\succeq_i)_{i=1}^n\right)$ be a context with n individuals and $d_C=d$. Let G be the set of frontier-concerned individuals in C (hence, $|G|=d$), and let $\hat{C}_{n,G}=\left(\hat{X}_{n,G},(\hat{\succeq}_i)_{i=1}^n\right)$ be as in the proof of Lemma 8.

Let $i\in G$. Let π be a permutation such that $\pi(1)=\{i\}$ and $\left\{\pi(2),\ldots,\pi(d)\right\}=G\setminus\{i\}$. Therefore, $\left\{\pi(d{+}1),\ldots,\pi(n)\right\}=\{1,\ldots,n\}\setminus G$. Note that in the context $C'\triangleq\left(\hat{X}_{n,G},(\hat{\succeq}_{\pi(i)})_{i=1}^n\right)$, individuals $\{1,\ldots,d\}$ are frontier-concerned and individuals $\{d{+}1,\ldots,n\}$ are frontier-indifferent. Therefore, by Lemma 8 (applied twice) and by anonymity,

$$c^i_{n,G}=I(\hat{C}_{n,G},0)-I(\hat{C}_{n,G},i)=I(C',0)-I(C',i)=c^1_{n,\{1,\ldots,d\}}=\frac{c_{n,d}}{n}.$$

For every $i=1,\ldots,n$, let u_i be some vNM utility representation of $\succeq_i$. We have shown that $c^i_{n,G}=\frac{c_{n,d}}{n}$ for all $i\in G$. Letting $c^i_{n,G}\triangleq\frac{c_{n,d}}{n}$ for all $i\notin G$ (recall that by

Lemma 8 the choice of these is arbitrary), by Lemma 8 we have for every $x \in \Delta(X)$ that

$$I(C,x) = \max_{x' \in X} \left\{ \sum_{i=1}^{n} \frac{c_{n,d}}{n} \frac{u_i(x') - u_i^{min}}{u_i^{max} - u_i^{min}} \right\} - \sum_{i=1}^{n} \frac{c_{n,d}}{n} \frac{u_i(x) - u_i^{min}}{u_i^{max} - u_i^{min}} = c_{n,d} \cdot \hat{I}(C,x). \quad \square$$

Proof of Lemma 10. The "if" direction follows by construction of $\hat{I}$.[37] We prove the "only if" direction. Throughout the proof, we make use of the fact that the frontier dimension of a composed context equals the sum of frontier dimensions of the original contexts from which it is composed; this is because an individual in a composed context is frontier-concerned if and only if that individual is frontier-concerned in their original context.

For every $n \in \mathbb{N}$ and $d \in \{2,3,\ldots,n\}$, let $0 < c_{n,d} < \infty$ be the constant guaranteed to exist by Lemma 9. For $n > 1$, let $c_n \triangleq c_{n,n}$. For $n = 1$, let $0 < c_n < \infty$ be arbitrary. For $n \in \{3,4\}$, let $c'_n \triangleq c_{n,n-1}$.

By Lemma 9, it suffices to prove for all $n \in \mathbb{N} \setminus \{3,4\}$ and $d \in \{2,3,\ldots,n-1\}$ that $c_{n,d} = c_n$ (recall that the choice of $c_{n,0}$ is arbitrary), and to prove for all $n \in \{3,4\}$ and $d \in \{2,3,\ldots,n-2\}$ that $c_{n,d} = c_n$.

We first prove that for every $n \in \mathbb{N}$ and $d \in \{2,\ldots,n-2\}$, it holds that $c_{n,d} = c_n$. Note that for $n \leq 3$, this claim is vacuous and trivially holds, so it suffices to prove this claim for $n \geq 4$. Let $\hat{C}_{d,\{1,\ldots,d\}}$ and $\hat{C}_{n-d,\{1,\ldots,n-d\}}$ be as in the proof of Lemma 8 (and note that $d \geq 2$ and $n - d \geq 2$ since $n \geq 4$). Let C' be the context obtained from $\hat{C}_{n-d,\{1,\ldots,n-d\}}$ by making all individuals indifferent between all alternatives. By definition of $\hat{I}$, by Lemma 9 (applied twice), and by IIP,

$$\begin{aligned} c_{n,d} = c_{n,d} \cdot n \left(\hat{I}\big(\hat{C}_{d,\{1,\ldots,d\}} \oplus C', (0,0)\big) - \hat{I}\big(\hat{C}_{d,\{1,\ldots,d\}} \oplus C', (1,0)\big) \right) = \\ n \left(I\big(\hat{C}_{d,\{1,\ldots,d\}} \oplus C', (0,0)\big) - I\big(\hat{C}_{d,\{1,\ldots,d\}} \oplus C', (1,0)\big) \right) = \\ n \left(I\big(\hat{C}_{d,\{1,\ldots,d\}} \oplus \hat{C}_{n-d,\{1,\ldots,n-d\}}, (0,0)\big) - I\big(\hat{C}_{d,\{1,\ldots,d\}} \oplus \hat{C}_{n-d,\{1,\ldots,n-d\}}, (1,0)\big) \right) = \\ c_{n,n} \cdot n \left(\hat{I}\big(\hat{C}_{d,\{1,\ldots,d\}} \oplus \hat{C}_{n-d,\{1,\ldots,n-d\}}, (0,0)\big) - \hat{I}\big(\hat{C}_{d,\{1,\ldots,d\}} \oplus \hat{C}_{n-d,\{1,\ldots,n-d\}}, (1,0)\big) \right) = \end{aligned}$$

[37] To see this for $n = 3$, note that the there is no context with 3 individuals that is a composition of contexts that both have nonzero frontier dimensions, and so IIP has no bite (on top of the result of Lemma 9) for $n = 3$. For $n = 4$, the only contexts with 4 individuals that are a composition of contexts that both have nonzero frontier dimensions are some contexts with frontier dimension 4, which are each a composition of contexts with frontier dimension 2; consequently, IIP is uninformative (on top of the result of Lemma 9) regarding contexts with 4 individuals and frontier dimension 3.

$$c_{n,n} = c_n.$$

It remains to prove that $c_{n,n-1} = c_n$ for every $n \geq 5$. Let $\hat{C}_{n-3,\{1,\ldots,n-3\}}$ and $\hat{C}_{3,\{1,2\}}$ be as in the proof of Lemma 8 (and note that $n-3 \geq 2$ since $n \geq 5$). Let C' be the context obtained from $\hat{C}_{3,\{1,2\}}$ by making all individuals indifferent between all alternatives. By definition of $\hat{I}$, by Lemma 9 (applied twice), by IIP, and by the first part of this proof with $d = n-3$,

$$\begin{gathered}
c_{n,n-1} = \\
c_{n,n-1} \cdot n \left(\hat{I}\big(\hat{C}_{n-3,\{1,\ldots,n-3\}} \oplus \hat{C}_{3,\{1,2\}}, (0,0)\big) - \hat{I}\big(\hat{C}_{n-3,\{1,\ldots,n-3\}} \oplus \hat{C}_{3,\{1,2\}}, (1,0)\big) \right) = \\
n \left(I\big(\hat{C}_{n-3,\{1,\ldots,n-3\}} \oplus \hat{C}_{3,\{1,2\}}, (0,0)\big) - I\big(\hat{C}_{n-3,\{1,\ldots,n-3\}} \oplus \hat{C}_{3,\{1,2\}}, (1,0)\big) \right) = \\
n \left(I\big(\hat{C}_{n-3,\{1,\ldots,n-3\}} \oplus C', (0,0)\big) - I\big(\hat{C}_{n-3,\{1,\ldots,n-3\}} \oplus C', (1,0)\big) \right) = \\
c_{n,n-3} \cdot n \left(\hat{I}\big(\hat{C}_{n-3,\{1,\ldots,n-3\}} \oplus C', (0,0)\big) - \hat{I}\big(\hat{C}_{n-3,\{1,\ldots,n-3\}} \oplus C', (1,0)\big) \right) = \\
c_{n,n-3} = c_n. \qquad \square
\end{gathered}$$

Proof of Theorem 1. The "if" direction follows by construction of $\hat{I}$. We prove the "only if" direction. (For the "furthermore" part, see Section 4.2.) For every $n \in \mathbb{N}$, let $0 < c_n < \infty$ be the constant guaranteed to exist by Lemma 10; for $n = 3, 4$, let $0 < c'_n < \infty$ be the constant guaranteed to exist by that lemma. Let $c \triangleq c_2$. By Lemma 10, it suffices to prove for all integers $n \geq 3$ that $c_n = c$ (recall that the choice of c_1 is arbitrary), and that $c'_3 = c'_4 = c$.

Let $n \in \mathbb{N}$, and let $\hat{C}_{n,\{1,\ldots,n\}}$ and $\hat{C}_{2,\{1,2\}}$ be as in the proof of Lemma 8. By definition of $\hat{I}$, by Lemma 10 (applied four times), and by duplication (applied twice),

$$\begin{gathered}
c_n = c_n \cdot n\big(\hat{I}(\hat{C}_{n,\{1,\ldots,n\}}, 0) - \hat{I}(\hat{C}_{n,\{1,\ldots,n\}}, 1)\big) = n\big(I(\hat{C}_{n,\{1,\ldots,n\}}, 0) - I(\hat{C}_{n,\{1,\ldots,n\}}, 1)\big) = \\
n \left(I\big(\hat{C}_{n,\{1,\ldots,n\}} \oplus \hat{C}_{n,\{1,\ldots,n\}}, (0,0)\big) - I\big(\hat{C}_{n,\{1,\ldots,n\}} \oplus \hat{C}_{n,\{1,\ldots,n\}}, (1,1)\big) \right) = \\
c_{2n} \cdot n \left(\hat{I}\big(\hat{C}_{n,\{1,\ldots,n\}} \oplus \hat{C}_{n,\{1,\ldots,n\}}, (0,0)\big) - \hat{I}\big(\hat{C}_{n,\{1,\ldots,n\}} \oplus \hat{C}_{n,\{1,\ldots,n\}}, (1,1)\big) \right) = c_{2n} = \\
c_{2n} \cdot 2 \left(\hat{I}\big(\oplus_{j=1}^{n} \hat{C}_{2,\{1,2\}}, (0,\ldots,0)\big) - \hat{I}\big(\oplus_{j=1}^{n} \hat{C}_{2,\{1,2\}}, (1,\ldots,1)\big) \right) = \\
2 \left(I\big(\oplus_{j=1}^{n} \hat{C}_{2,\{1,2\}}, (0,\ldots,0)\big) - I\big(\oplus_{j=1}^{n} \hat{C}_{2,\{1,2\}}, (1,\ldots,1)\big) \right) = \\
2\big(I(\hat{C}_{2,\{1,2\}}, 0) - I(\hat{C}_{2,\{1,2\}}, 1)\big) = c_2 \cdot 2\big(\hat{I}(\hat{C}_{2,\{1,2\}}, 0) - \hat{I}(\hat{C}_{2,\{1,2\}}, 1)\big) = c_2 = c.
\end{gathered}$$

Finally, let $n \in \{3,4\}$, and let $\hat{C}_{n,\{1,\ldots,n-1\}}$ be as in the proof of Lemma 8. By the first part of this proof, $c_{2n} = c$. Therefore, by definition of $\hat{I}$, by Lemma 10 (applied twice), and by duplication,

$$\begin{aligned} c'_n = c'_n \cdot n\big(\hat{I}(\hat{C}_{n,\{1,\ldots,n-1\}},0) - \hat{I}(\hat{C}_{n,\{1,\ldots,n-1\}},1)\big) = \\ n\big(I(\hat{C}_{n,\{1,\ldots,n-1\}},0) - I(\hat{C}_{n,\{1,\ldots,n-1\}},1)\big) = \\ n\left(I\big(\hat{C}_{n,\{1,\ldots,n-1\}} \oplus \hat{C}_{n,\{1,\ldots,n-1\}},(0,0)\big) - I\big(\hat{C}_{n,\{1,\ldots,n-1\}} \oplus \hat{C}_{n,\{1,\ldots,n-1\}},(1,1)\big)\right) = \\ c_{2n} \cdot n\left(\hat{I}\big(\hat{C}_{n,\{1,\ldots,n-1\}} \oplus \hat{C}_{n,\{1,\ldots,n-1\}},(0,0)\big) - \hat{I}\big(\hat{C}_{n,\{1,\ldots,n-1\}} \oplus \hat{C}_{n,\{1,\ldots,n-1\}},(1,1)\big)\right) = \\ c_{2n} = c. \qquad \square \end{aligned}$$

B Proofs Omitted from Section 5

Proof of Lemma 11. We need to prove that for every ex-post Pareto efficient matching $x \in X$ and for every $y \in \Delta(X)$ such that $y \succeq_i x$ for all $i = 1,\ldots,n$, it holds that $y = x$. We prove this by induction over the number of individuals n. The case of $n = 1$ is immediate. Assume that $n \geq 2$ and that the claim holds for $n-1$.

Since x is ex-post Pareto efficient, there exists $\hat{i} \in \{1,\ldots,n\}$ that is matched to her favorite object in x (otherwise, there would exist a Pareto-improving trading cycle, contradicting the ex-post Pareto efficiency of x). Since $y \succeq_{\hat{i}} x$ and since the individuals' preferences over the objects are strict, $\hat{i}$ is matched with probability 1 to her favorite object in y as well. Let $N' \triangleq \{1,\ldots,n\} \setminus \{\hat{i}\}$ and let O' be the set of all objects except for the one matched to $\hat{i}$ by x. Let X' be the set of all perfect matchings between N' and O', and denote by $x' \in X'$ and $y' \in \Delta(X')$ the respective restrictions of x and y to N'. Note that x' is ex-post Pareto efficient with respect to $R\big((\succeq_i |_{\Delta(X')})_{i \in N'}\big)$ and that $y' \succeq_i x'$ for every $i \in N'$. By the induction hypothesis, $y' = x'$. Therefore, $y = x$. $\square$

Proof of Theorem 3. We say that μ is *anonymous* if for every profile of rankings $R = (R_i)_{i=1}^n \in \mathcal{R}^n$ and for every permutation π on $\{1,\ldots n\}$, it holds that $\pi^{-1}\big(\mu(\pi(R))\big) = \mu(R)$, where (i) $\pi(R) \triangleq (R_{\pi(1)},\ldots,R_{\pi(n)})$ and (ii) $\pi^{-1}\big(\mu(\pi(R))\big)$ is defined so that if $\mu\big(\pi(R)\big) = \sum_{\ell=1}^L p_\ell \cdot x_\ell$ where $x_\ell \in X$ for all $1 \leq \ell \leq L$, then $\pi^{-1}\big(\mu(\pi(R))\big) \triangleq \sum_{\ell=1}^L p_\ell \cdot \pi^{-1}(x_\ell)$, where $\pi^{-1}(x_\ell)$ matches i to whoemever $\pi^{-1}(i)$ is matched in x_ℓ, for every $i = 1,\ldots,n$.

We start with the case in which μ is anonymous. Let $0 < \varepsilon < 1/n$. For each $i = 1, \ldots, n$, consider the following utility function over objects.

$$u_i(j) \triangleq \begin{cases} 1 - (j-1)\varepsilon & j \leq i, \\ \frac{n-j}{n}\varepsilon & j > i, \end{cases}$$

and let $\succeq_i$ be the preference over $\Delta(X)$ whose vNM utility representation is u_i. Note that $R(\succeq_i) = R(\succeq_j)$ for all $i, j \in \{1, \ldots, n\}$. Therefore, all matchings are ex-post Pareto efficient, and by Lemma 11 are on the Pareto frontier. Furthermore, since μ is ordinal and anonymous, when the reported profile of rankings is $R(\succeq)$, every matching in X is output by μ with the same probability. Therefore, the sum of utilities at $\mu(\succeq)$ is

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} u_i(j) \leq \frac{1}{n}\sum_{i=1}^{n}\big(i + (n-i)\varepsilon\big) \leq \frac{n+1}{2} + n\varepsilon.$$

Note that the matching that maximizes the sum of utilities matches each individual $i = 1, \ldots, n$ to the same-indexed object $j = i$. The sum of utilities of that matching is

$$\sum_{i=1}^{n} u_i(i) = \sum_{i=1}^{n}\big(1 - (i-1)\varepsilon\big) \geq n - n^2\varepsilon.$$

Therefore, since all matchings are on the Pareto frontier,[38]

$$\hat{I}\big((X, \succeq), \mu(\succeq)\big) \geq 1 - n\varepsilon - \left(\frac{n+1}{2n} + \varepsilon\right) = \frac{n-1}{2n} - (n+1)\varepsilon \xrightarrow[\varepsilon \to 0]{} \frac{n-1}{2n} = \frac{1}{2} - \frac{1}{2n}.$$

We move on to the case in which μ is not anonymous. Let $0 < \varepsilon < 1/n$. Let $\hat{\mu}$ be such that for every $R \in \mathcal{R}^n$, it holds that $\hat{\mu}(R) = \sum_\pi \frac{1}{n!} \cdot \pi^{-1}\big(\mu(\pi(R))\big)$, where the sum is over all permutations on $\{1, \ldots, n\}$. By the anonymous case, there exists a preference profile $\succeq \in \mathcal{P}^n$ such that $\hat{I}\big((X, \succeq), \hat{\mu}(\succeq)\big) \geq \frac{n-1}{2n} - (n+1)\varepsilon$. Therefore, there exists a permutation π on $\{1, \ldots, n\}$ such that $\hat{I}\big((X, \succeq), \pi^{-1}(\mu(\pi(\succeq)))\big) \geq \frac{n-1}{2n} - (n+1)\varepsilon$. We conclude the proof by focusing on the preference profile $\pi(\succeq)$ and noting that $\hat{I}\big((X, \pi(\succeq)), \mu(\pi(\succeq))\big) = \hat{I}\big((X, \succeq), \pi^{-1}(\mu(\pi(\succeq)))\big) \geq \frac{n-1}{2n} - (n+1)\varepsilon \xrightarrow[\varepsilon \to 0]{} \frac{1}{2} - \frac{1}{2n}$. □

[38] Recall that our social inefficiency function is a per capita measure, i.e., it averages over the individuals' utility losses.

Proof of Lemma 13. The result follows by induction from the following claim, which we now prove: For every profile $u = (u_i)_{i=1}^n \in \mathcal{UR}$ and $i = 1, \ldots, n$ such that[39] $u_i(X) \cap [\varepsilon, 1-\varepsilon] \neq \emptyset$, there exists $u_i' \in \mathcal{UR}$ satisfying $R(u_i') = R(u_i)$ and $\left|u_i'(X) \cap [\varepsilon, 1-\varepsilon]\right| < \left|u_i(X) \cap [\varepsilon, 1-\varepsilon]\right|$, such that $I_u\big(\mu(u)\big) \leq I_{(u_i', u_{-i})}\big(\mu(u)\big)$.

Let $[0, \bar{l}] \subseteq [0, \varepsilon)$, $[l, r] \subseteq [\varepsilon, 1-\varepsilon]$, and $[\bar{r}, 1] \subseteq (1-\varepsilon, 1]$ be the smallest such segments such that $u_i(X) \subseteq [0, \bar{l}] \cup [l, r] \cup [\bar{r}, 1]$. These are well defined because the three intersections $u_i(X) \cap [0, \varepsilon)$, $u_i(X) \cap [\varepsilon, 1-\varepsilon]$, and $u_i(X) \cap (1-\varepsilon, 1]$ are all nonempty by assumption. Let $\hat{l} = \frac{\bar{l}+\varepsilon}{2}$ and $\hat{r} = \frac{\bar{r}+1-\varepsilon}{2}$. For every $\delta \in \{\hat{l} - l, 0, \hat{r} - r\}$, define the utility function u_i^δ as follows:

$$u_i^\delta(X) = \begin{cases} u_i(X) & u_i(X) \notin [l, r], \\ u_i(X) + \delta & u_i(X) \in [l, r]. \end{cases}$$

Note that $u_i^0 = u_i$ and that $R(u_i^\delta) = R(u_i)$ for every δ. Note also that for every $\delta \in \{\hat{l} - l, \hat{r} - r\}$, we have that $\left|u_i^\delta(X) \cap [\varepsilon, 1-\varepsilon]\right| < \left|u_i(X) \cap [\varepsilon, 1-\varepsilon]\right|$. It remains to show that $I_{(u_i^0, u_{-i})}\big(\mu(u)\big) \leq \max_{\delta \in \{\hat{l}-l, \hat{r}-r\}} \big\{I_{(u_i^\delta, u_{-i})}\big(\mu(u)\big)\big\}$.

Let $x_{max} \in \arg\max_{x \in X} U_u(x)$. Note that the function $f(\delta) \triangleq U_{(u_i^\delta, u_{-i})}(x_{max}) - U_{(u_i^\delta, u_{-i})}\big(\mu(u)\big)$ is a linear function. Therefore, at one of the extreme values $\hat{\delta} \in \{\hat{l} - l, \hat{r} - r\}$, the function f attains a weakly higher value than at the intermediate value $\delta = 0$. Let $u_i' \triangleq u_i^{\hat{\delta}}$. We conclude by observing that $I_u\big(\mu(u)\big) = f(0) \leq f(\hat{\delta}) \leq I_{(u_i', u_{-i})}\big(\mu(u)\big)$. □

Proof of Lemma 15. Assume for contradition that True is returned. Then the max-cardinality matching σ in G is of size $n-1$. Therefore, $x \triangleq \sigma \cup (\hat{i}, o) \in X$ is a perfect matching of the n individuals to the n objects such that each individual except $\hat{i}$ is matched to an object that she strictly prefers to o. Let $x' \in X$ be an ex-post Pareto efficient matching such that $x' \succeq_i x$ for all $i = 1, \ldots, n$. Then, in x', each individual except $\hat{i}$ is matched to an object that she strictly prefers to o. Therefore, o is matched to $\hat{i}$ in x', in contradiction to o not being matched to $\hat{i}$ in any ex-post Pareto efficient matching. □

Proof of Lemma 16. Let $x \in X$ be an ex-post Pareto efficient matching that matches o to $\hat{i}$. Assume for contradiction that in x, some individual $j \neq \hat{i}$ is matched to an object that she strictly prefers less than o. Let $y \in X$ be the matching obtained from

[39] We denote $u_i(X) \triangleq \{u_i(x) \mid x \in X\}$.

x by exchanging the objects of $\hat{i}$ and j. Then, j strictly prefers y to x. Since x is ex-post Pareto efficient, we therefore have that $\hat{i}$ is matched in y to an object she strictly prefers less than o. Let $y' \in X$ be an ex-post Pareto efficient matching such that $y' \succeq_i y$ for all $i = 1, \ldots, n$. We claim that $\hat{i}$ is matched in y' to an object she strictly prefers less than o; indeed, if this were not the case, then it would hold that $y' \succeq_i x$ for all $i = 1, \ldots, n$ and $y' \succ_j x$, in contradiction to x being ex-post Pareto efficient. We conclude that y' is an ex-post Pareto efficient matching that matches to $\hat{i}$ an object that she strictly prefers less than o, contradicting the definition of o. Therefore, in x each individual $j \neq \hat{i}$ is matched to an object that she strictly prefers to o. Hence, during the computation of the function call $\textsc{Test}(n, (\succeq_i)_{i=1}^n, \hat{i}, o)$, the max-cardinality matching σ in G is of size $n-1$. Thus, True is returned. $\square$