Discussion Paper Series

IZA DP No. 18986

October 2026

Opportunity-Sensitive Social Welfare

Tim Wienand ⓡ
Erasmus University Rotterdam
and Tinbergen Institute

Brice Magdalou ⓡ
CEE-M, University of Montpellier,
CNRS, INRAE and Institut Agro

Richard Nock ⓡ
Google Research

Paul Hufe
University of Bristol, CESifo,
HCEO, IFS and IZA@LISER

The IZA Discussion Paper Series (ISSN: 2365-9793) ("Series") is the primary platform for disseminating research produced within the framework of the IZA@LISER Network, an unincorporated international network of labour economists coordinated by the Luxembourg Institute of Socio-Economic Research (LISER). The Series is operated by LISER, a Luxembourg public establishment (établissement public) registered with the Luxembourg Business Registers under number J57, with its registered office at 11, Porte des Sciences, 4366 Esch-sur-Alzette, Grand Duchy of Luxembourg.

Any opinions expressed in this Series are solely those of the author(s). LISER accepts no responsibility or liability for the content of the contributions published herein. LISER adheres to the European Code of Conduct for Research Integrity. Contributions published in this Series present preliminary work intended to foster academic debate. They may be revised, are not definitive, and should be cited accordingly. Copyright remains with the author(s) unless otherwise indicated.

IZA DP No. 18986 October 2026

Opportunity-Sensitive Social Welfare*

Abstract

We develop a framework to evaluate income distributions from the perspective of an opportunity-egalitarian social planner. Through an original axiomatic route that bridges normative economics and decision theory under ambiguity, we characterize a class of social welfare functions governed by a single parameter capturing the planner's aversion to inequality of opportunity. We obtain a 'mean-divergence' representation that separates economic efficiency from a penalty for inequality of opportunity, as well as a stochastic dominance criterion for robust welfare comparisons. We apply these theoretical results to analyze opportunity-sensitive social welfare in the United States across birth cohorts and regions. Mapping standard estimates of intergenerational mobility into our framework, we show that welfare rankings vary substantially with the assumed degree of inequality-of-opportunity aversion. We transparently gauge the extent of this normative ambiguity by providing bounds on the ranking of birth cohorts and regions, respectively.

JEL classification

D31, D63, I31

Keywords

inequality of opportunity, intergenerational mobility, welfarism, utilitarianism, social welfare measurement

Corresponding author

Paul Hufe
paul.hufe@bristol.ac.uk

* Random author order. We acknowledge financial support from ANR (ANR-24-CE26-3823-02) and UKRI (MR/X033333/1). We are grateful to seminar and workshop audiences at the University of Bristol, the University of Montpellier, SOFI Stockholm, Tinbergen Institute, the French-Japanese Webinar in Economics (FJWE), the Normative Economics & Economic Policy Webinar (NE&EP), IIPF 2026, and the FUR Conference 2026 for their constructive feedback and suggestions. This paper has benefited from discussions with Mohammed Abdellaoui, Gilles Dufrénot, Francisco Ferreira, Nicolas Gravel, Markus Jäntti, Paola Manzini, Juan Moreno-Ternero, Mauricio Ribeiro, and Peter Wakker. All remaining errors are our own.

1. Introduction

Equality of opportunity is widely regarded as a fundamental ideal of distributive justice (Roemer and Trannoy, 2016). While inequalities resulting from individual effort or well-informed choices may be morally acceptable, they become unfair when driven by circumstances beyond the individual's control, such as parental income, race, or biological sex. Assessing the extent of these unfair inequalities is therefore essential to evaluating a society's distributive performance and to guiding the design of fair public policies.

In this paper, we develop a new normative framework to evaluate income distributions from the perspective of an opportunity-egalitarian social planner. We characterize a novel class of opportunity-sensitive social welfare functions governed by a single, transparent parameter that captures the planner's degree of aversion to inequality of opportunity. The resulting framework is both normatively transparent and easy to implement in empirical applications. The inequality-of-opportunity aversion parameter governs how rapidly the planner's concern shifts toward disadvantaged circumstance groups and directly translates into endogenous welfare weights for each of them.

We take these measures to the data and evaluate income distributions in the United States across birth cohorts and regions. In line with the existing literature, we estimate intergenerational mobility and interpret the resulting estimates as proxy indicators of equality of opportunity. These estimates replicate well-established findings on mobility patterns in the US, both across birth cohorts and commuting zones. However, by interpreting these estimates through our framework, we also show that their welfare implications are ambiguous as they crucially depend on assumptions about the appropriate degree of inequality-of-opportunity aversion.

To fix ideas, consider a stylized example of two societies with identical marginal income distributions. In the first society, an individual's income prospects are entirely determined by circumstances beyond the individual's control, whereas in the second, they are independent of such factors. A utilitarian planner, who evaluates welfare solely on the basis of the marginal distribution of anonymous outcomes, would be strictly indifferent between the two. By contrast, an opportunity-egalitarian planner would regard the first society as substantially less desirable. We capture this intuition by partitioning the population into distinct 'types' based on observable circumstances and by penalizing disparities in type-specific welfare derived from these conditional income prospects, according to the planner's degree of inequality-of-opportunity aversion.

Theoretical foundations. In the first part of the paper, we provide a complete axiomatic characteri-

zation of this class of social welfare functions. To cleanly isolate our core normative requirements, we begin by establishing a baseline representation under the assumption of a uniform demographic distribution across types. The resulting representation takes a double-expectation form, in which the welfare of each type, evaluated by the expected utility of its income distribution, is aggregated via a concave transformation that captures the social planner's inequality-of-opportunity aversion. This aggregation is achieved by taking the expectation over the demographic distribution of types. By strengthening our independence axioms, we further show that this transformation reduces to a one-parameter exponential family. Moreover, under the usual monotonicity and scale-invariance requirements and the assumption that the social planner is averse to inequality of opportunity, the utility function is necessarily logarithmic in income. We then extend the analysis to societies with arbitrary, heterogeneous demographic weights. This realistic setting yields a highly tractable criterion. Formally, letting $q(s)$ denote the demographic weight of type s and $U(\pi_s)$ the expected utility of its income distribution, social welfare takes the following explicit form when the planner is averse to unequal opportunities ($\theta > 0$):

$$V_{\theta,u}(f,q) = -\frac{1}{\theta} \ln \left(\sum_{s \in \mathcal{S}} q(s) e^{-\theta U(\pi_s)} \right) .$$

This structure admits multiple insightful representations. Beyond this baseline functional form, we show that the planner's evaluation can also be expressed as a minimization problem involving the Kullback-Leibler divergence, which assigns endogenous normative weights to socially disadvantaged types. Furthermore, we establish an exact, closed-form 'mean-divergence' representation. This specific formulation uses a Bregman divergence to structurally separate economic efficiency from the inequality-of-opportunity penalty, a trade-off that also naturally reduces to a mean-variance approximation. Finally, we exploit this rich analytical structure to establish a stochastic dominance criterion that allows for an unambiguous ranking of societies across a broad class of completely alternating transformations.

Empirical analysis. In the second part of the paper, we use this framework to assess the welfare implications of intergenerational mobility estimates in the United States. In line with the empirical literature on intergenerational income mobility, we quantify the extent to which children's incomes are predicted by their parents' incomes and interpret these estimates as proxy indicators of equality of opportunity. Mapping these empirical estimates into our framework, we assess income distributions from the perspective of an opportunity-egalitarian social planner. Concretely, we use the Panel Study of Income Dynamics (PSID) to estimate mobility at the national level across the 1940s–1980s birth cohorts, and data from Opportunity Insights to estimate mobility differences across Commuting Zones (CZs) for children born around 1980. We then show that the associated welfare comparisons are subject to normative ambiguity, i.e., they crucially depend on assumptions about the appropriate

degree of inequality-of-opportunity aversion. We quantify the extent of normative ambiguity in welfare comparisons by constructing dominance sets for the rankings of societies over plausible ranges of inequality-of-opportunity aversion. The resulting bounds provide a transparent tool for welfare comparisons across societies. For example, we show that the substantial drop in intergenerational mobility between the 1950s and the 1960s does not lead to a corresponding drop in social welfare when assuming a low but plausible extent of inequality-of-opportunity aversion. More generally, our analysis reveals substantial normative ambiguity both across birth cohorts and regions in the United States: Unambiguous welfare rankings are rare, even when restricting the analysis to a small subset of the full range of potential parameter values.

Related literature. Our paper contributes to two main strands of the economics literature: the theoretical literature on normative economics and the empirical literature on intergenerational mobility and equality of opportunity.

First, within normative economics, we contribute to the theoretical literature on equality of opportunity and the assessment of risky social situations. Firmly embedded within the ex-ante perspective, our approach evaluates societies by treating currently observed type-specific distributions as objective income prospects. A defining feature of our framework is that we do not impose expected utility representations or von Neumann-Morgenstern rationality *a priori* on individuals, a standard premise in foundational works like Fleurbaey (2010) and Grant *et al.* (2010). Instead, this expected utility structure emerges endogenously from the social planner's evaluation of aggregate risks. Furthermore, while our social welfare function shares the 'generalized utilitarian' structure of Grant *et al.* (2010), we reach it via a less demanding axiomatic route that isolates the core normative evaluation from population variations. This deliberate separation allows our framework to explicitly tackle the aggregation of unequally sized types, distinguishing our approach from recent theoretical contributions that assume symmetric groups or uniform demographic weights (Fleurbaey *et al.*, 2024; Andreoli *et al.*, 2026).[1]

Second, we contribute to the empirical literature on intergenerational mobility and equality of opportunity (Chetty *et al.*, 2014a, 2026; Davis and Mazumder, 2026; Jácome *et al.*, 2025; Nybom and Stuhler, 2024; Mello *et al.*, forthcoming; Hufe *et al.*, 2022; Roemer and Trannoy, 2016; Ramos and Van de gaer, 2016). The extent of intergenerational mobility is commonly interpreted as a proxy for equal opportunities in a society (Deutscher and Mazumder, 2023; Chetty *et al.*, 2014a; Alesina *et al.*, 2018). Despite this normative motivation, much of the existing literature on intergenerational mobility sidesteps discussions about its

[1] Although grounded in an ex-ante framework, our methodology establishes formal connections to ex-post egalitarianism (Grant *et al.*, 2012) and ex-post prioritarianism (Adler *et al.*, 2025). Finally, for a recent and insightful exploration of the boundaries between welfarism and equality of opportunity, see Carroll (2025), who characterizes the conditions under which these two criteria coincide in resource allocation problems.

normative underpinnings. We address this shortcoming by embedding standard intergenerational mobility estimates into our axiomatically characterized social welfare function. This mapping makes the normative foundations explicit and allows us to evaluate the implications of intergenerational mobility estimates for social welfare assessments. Relatedly, several recent papers have explored the connection between the literature on equality of opportunity and intergenerational mobility (Deutscher and Mazumder, 2023; Björklund and Jäntti, 2020; Adermon *et al.*, forthcoming). These papers mostly analyze whether estimates from both branches of the literature yield similar empirical conclusions. Our approach differs by incorporating standard intergenerational mobility estimates into an axiomatically grounded analytical framework that represents the preferences of an opportunity-egalitarian social planner.

Our contribution falls within the field of normative economics, but the formal framework remains relevant to other areas of economic theory. In particular, it exhibits a structural connection to decision theory under ambiguity. A detailed mapping of our axioms and results to the ambiguity literature is provided in Appendix B.

Outline. The remainder of the paper is organized as follows. Section 2 establishes the formal framework. Section 3 introduces the baseline axioms governing the planner's preferences over uniform type partitions. Section 4 derives the main representation theorems for the social welfare function in this uniform setting. Section 5 extends this characterization to the domain of arbitrary demographic distributions across types. Section 6 explores equivalent representations of the welfare criterion and establishes a stochastic dominance criterion. Section 7 examines the basic properties of the welfare function. In Section 8, we use our framework to evaluate opportunity-sensitive social welfare in the US across birth cohorts and commuting zones. Finally, Section 9 concludes the paper.

2. Setup

For a topological space $\mathcal{A}$, we indicate by $|\mathcal{A}|$ its cardinality whenever $\mathcal{A}$ is finite. We let $\Sigma_{\mathcal{A}}$ be the Borel σ-algebra on $\mathcal{A}$, and we denote by $\Delta(\mathcal{A})$ the set of finitely supported probability measures on $(\mathcal{A}, \Sigma_{\mathcal{A}})$, endowed with the weak topology. For $p \in \Delta(\mathcal{A})$ we denote by $p(a)$, instead of $p(\{a\})$, the probability assigned to $a \in \mathcal{A}$.

Consider a finite population $\mathcal{N}$ consisting of at least two individuals. A *type* $s \in \mathcal{S}$, with $\mathcal{S}$ a partition of $\mathcal{N}$, is a population subgroup sharing characteristics that give rise to unequal opportunities. We focus on subgroups defined by parental income; however, the framework is general enough to incorporate other sources of unequal opportunities, such as biological sex, ethnicity, or neighborhood quality during

childhood. We do not impose an order relation over types. We denote by $q \in \Delta(\mathcal{S})$ the demographic type weights, defined by the population shares $q(s) = |s|/|\mathcal{N}|$ for all $s \in \mathcal{S}$, which represent the distribution of individuals over $\mathcal{S}$ and which is known by the social planner.

An *income distribution* is denoted $\pi \in \Delta(\mathcal{Y})$, with $\mathcal{Y}$ a compact interval in $(0, \infty)$. The mean income is $\mu(\pi) = \sum_{y \in \mathcal{Y}} \pi(y) y$.[2] Type-specific income distributions are denoted by $\pi_s \in \Delta(\mathcal{Y})$, for all $s \in \mathcal{S}$.

An *opportunity profile* $f \in \mathcal{F}_{\mathcal{S}}$ assigns a type-specific income distribution $f(s) = \pi_s \in \Delta(\mathcal{Y})$ to each $s \in \mathcal{S}$. The set of all opportunity profiles $\mathcal{F}_{\mathcal{S}} = \Delta(\mathcal{Y})^{|\mathcal{S}|}$, for a given partition $\mathcal{S}$, is endowed with the corresponding product topology. Denoting by $\mathbb{E}_q$ the expectation operator according to $q \in \Delta(\mathcal{S})$, the income distribution across the entire population is $\mathbb{E}_q[f] = \sum_{s \in \mathcal{S}} q(s) \pi_s$, and the mean income is $\mu\left(\mathbb{E}_q[f]\right) = \sum_{s \in \mathcal{S}} q(s) \mu(\pi_s)$. An equal-opportunity profile that assigns the same income distribution $\pi \in \Delta(\mathcal{Y})$ to all types $s \in \mathcal{S}$, is denoted by $f_\pi \in \mathcal{F}_{\mathcal{S}}$.

Finally, a *society* is a pair $(f, q) \in \mathcal{F}_{\mathcal{S}} \times \Delta(\mathcal{S})$.

Example. Suppose there are two types in society: 20% of the population have rich parents (R), and 80% have poor parents (P). The set of types is therefore $\mathcal{S}$ = {R, P}. The set of observed incomes is {\$1, \$2}. The type-specific income distributions are π_{R} = (0.1, 0.9) and π_{P} = (0.6, 0.4). This society is characterized by the following opportunity profile f and demographic type weights q:

$$f = \begin{array}{c} \\ \mathrm{R} \\ \mathrm{P} \end{array} \begin{array}{c} \begin{array}{cc} \$1 & \$2 \end{array} \\ \begin{pmatrix} 0.1 & 0.9 \\ 0.6 & 0.4 \end{pmatrix} \end{array}, \quad \text{and} \quad q = \begin{pmatrix} 0.2 \\ 0.8 \end{pmatrix}.$$

The overall income distribution is $\mathbb{E}_q[f]$ = (0.5, 0.5) = 0.2 × (0.1, 0.9) + 0.8 × (0.6, 0.4), with mean income $\mu(\mathbb{E}_q[f])$ = \$1.5. Moreover, the mean incomes of individuals with rich and poor parents are $\mu(\pi_{\mathrm{R}})$ = \$1.9 and $\mu(\pi_{\mathrm{P}})$ = \$1.4, respectively.

3. Social planner's preferences

Initially, we assume that the demographic type weights $q \in \Delta(\mathcal{S})$ are fixed and uniform across all societies, with $q(s) = 1/|\mathcal{S}|$ for all $s \in \mathcal{S}$. This assumption allows us to separate considerations relating to equality of opportunity from those relating to variation in population composition. It will be relaxed

[2] Note that, since any $\pi \in \Delta(\mathcal{Y})$ is finitely supported, the sum contains only a finite number of non-zero terms $\pi(y)$. It is therefore well-defined and is equivalent to a summation over the finite support of π.

in Section 5.

The social planner's preferences are represented by a binary relation $\succcurlyeq$ on $\mathcal{F}_{\mathcal{S}}$: $f \succcurlyeq f'$ means that opportunity profile $f \in \mathcal{F}_{\mathcal{S}}$ is weakly preferred to profile $f' \in \mathcal{F}_{\mathcal{S}}$, given the uniform distribution of types in $\mathcal{S}$. We denote by $\sim$ and $\succ$ the symmetric and asymmetric parts of $\succcurlyeq$, respectively. We assume that $\succcurlyeq$ is a weak order (transitive and complete), non-degenerate (the asymmetric part is non-empty), and continuous (the upper and lower contour sets are closed). Under these assumptions, there exists a continuous social welfare function $W : \mathcal{F}_{\mathcal{S}} \to \mathbb{R}$ which represents $\succcurlyeq$ so that, for all $f, f' \in \mathcal{F}_{\mathcal{S}}$, we have $f \succcurlyeq f' \iff W(f) \geq W(f')$.

The first axiom we impose on $\succcurlyeq$ reflects unanimity in social improvement: If the social planner prefers all type-specific income distributions for one opportunity profile over those for another, then they must prefer the former profile to the latter.

Axiom 1 (Strict Pareto). *For all $f = (\pi_s)_{s\in\mathcal{S}} \in \mathcal{F}_{\mathcal{S}}$ and all $f' = (\pi'_s)_{s\in\mathcal{S}} \in \mathcal{F}_{\mathcal{S}}$, we have: $\forall s \in \mathcal{S}$, $f_{\pi_s} \succcurlyeq f_{\pi'_s} \Rightarrow f \succcurlyeq f'$. If, in addition, there exists at least one $s \in \mathcal{S}$ such that $f_{\pi_s} \succ f_{\pi'_s}$, then $f \succ f'$.*

The second axiom is a 'row separability condition'. It states that the ranking of opportunity profiles does not depend on types that have the same income distribution in both profiles. For any $f \in \mathcal{F}_{\mathcal{S}}$ and any $\pi \in \Delta(\mathcal{Y})$, let $(f_{-s}, \pi) \in \mathcal{F}_{\mathcal{S}}$ be the opportunity profile such that $(f_{-s}, \pi)(t) = \pi_t$ for all $t \neq s$ and $(f_{-s}, \pi)(s) = \pi$.

Axiom 2 (Independence to common type-specific distributions). *For all $f, f' \in \mathcal{F}_{\mathcal{S}}$, all $\pi, \pi' \in \Delta(\mathcal{Y})$ and all $s \in \mathcal{S}$, $(f_{-s}, \pi) \succcurlyeq (f'_{-s}, \pi) \Rightarrow (f_{-s}, \pi') \succcurlyeq (f'_{-s}, \pi')$.*

Consider the following example:

$$\begin{array}{c} \\ \mathrm{R} \\ \mathrm{P} \end{array}\!\!\begin{array}{c} \begin{array}{cc} \$1 & \$2 \end{array} \\ \begin{pmatrix} 0.2 & 0.8 \\ 0.4 & 0.6 \end{pmatrix} \end{array} \succcurlyeq \begin{array}{c} \\ \mathrm{R} \\ \mathrm{P} \end{array}\!\!\begin{array}{c} \begin{array}{cc} \$1 & \$2 \end{array} \\ \begin{pmatrix} 0.5 & 0.5 \\ 0.4 & 0.6 \end{pmatrix} \end{array} \implies \begin{array}{c} \\ \mathrm{R} \\ \mathrm{P} \end{array}\!\!\begin{array}{c} \begin{array}{cc} \$1 & \$2 \end{array} \\ \begin{pmatrix} 0.2 & 0.8 \\ 0.3 & 0.7 \end{pmatrix} \end{array} \succcurlyeq \begin{array}{c} \\ \mathrm{R} \\ \mathrm{P} \end{array}\!\!\begin{array}{c} \begin{array}{cc} \$1 & \$2 \end{array} \\ \begin{pmatrix} 0.5 & 0.5 \\ 0.3 & 0.7 \end{pmatrix} \end{array}.$$

The third axiom specifies that type identities are irrelevant for the social planner's evaluations. Specifically, it states that permuting the rows of an opportunity profile does not affect social welfare. Let $\mathcal{P}$ be the set of all permutations $\sigma : \mathcal{S} \to \mathcal{S}$.

Axiom 3 (Anonymity). *For all $f, f' \in \mathcal{F}_{\mathcal{S}}$, if there exists $\sigma \in \mathcal{P}$ such that $\pi_{\sigma(s)} = \pi'_s$ for all $s \in \mathcal{S}$, then $f \sim f'$.*

Consider the following example:

$$f = \begin{array}{c} \\ \text{R} \\ \text{P} \end{array}\!\!\begin{array}{c} \begin{array}{cc} \$1 & \$2 \end{array} \\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{array} \sim \begin{array}{c} \\ \text{R} \\ \text{P} \end{array}\!\!\begin{array}{c} \begin{array}{cc} \$1 & \$2 \end{array} \\ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{array} = f'.$$

This example illustrates that the anonymity axiom rules out some plausible egalitarian preferences. For example, one may argue that f should be preferred to f' because it reduces inequalities between dynasties by re-ranking the rich and the poor across generations. However, in this work, we evaluate individual opportunity sets rather than the fortunes of multigenerational dynasties. From this vantage point, type labels carry no moral weight beyond the opportunity set they are attached to.

The fourth axiom captures inequality-of-opportunity aversion. It states that any convergence of the income distributions of two types reduces inequality of opportunity.

Axiom 4 (Inequality of opportunity aversion). *For all $f, f' \in \mathcal{F}_{\mathcal{S}}$, if there exist $s, s' \in \mathcal{S}$ such that $\pi_t = \pi'_t$ for all $t \neq s, s'$ and, furthermore, $\pi_s = \pi_{s'} = (1/2)\,(\pi'_s + \pi'_{s'})$, then $f \succcurlyeq f'$.*

Consider the following example:

$$f = \begin{array}{c} \\ \text{R} \\ \text{P} \end{array}\!\!\begin{array}{c} \begin{array}{cc} \$1 & \$2 \end{array} \\ \begin{pmatrix} 0.3 & 0.7 \\ 0.3 & 0.7 \end{pmatrix} \end{array} \succcurlyeq \begin{array}{c} \\ \text{R} \\ \text{P} \end{array}\!\!\begin{array}{c} \begin{array}{cc} \$1 & \$2 \end{array} \\ \begin{pmatrix} 0.2 & 0.8 \\ 0.4 & 0.6 \end{pmatrix} \end{array} = f'.$$

This example also illustrates that our axiom on inequality-of-opportunity aversion has no immediate implications for overall inequality. To see this note that $\mathbb{E}_q[f] = \mathbb{E}_q[f'] = (0.3, 0.7)$, illustrating that inequality of opportunity is lower in f than f' while total inequality is the same.

The fifth axiom is a 'column separability condition'. It is analogous to the independence axiom in the expected utility model of von Neumann and Morgenstern (1944). Specifically, it states that mixing two equal-opportunity profiles with a common equal-opportunity profile does not alter the ranking of the first two profiles.

Axiom 5 (Independence to common equal opportunities). *For all $\pi, \pi', \pi'' \in \Delta(\mathcal{Y})$ and all $\alpha \in (0,1)$, $\quad f_\pi \succcurlyeq f_{\pi'} \;\Rightarrow\; \alpha f_\pi + (1-\alpha) f_{\pi''} \succcurlyeq \alpha f_{\pi'} + (1-\alpha) f_{\pi''}$.*

Consider the following example:

$$\begin{pmatrix} 0.2 & 0.8 \\ 0.2 & 0.8 \end{pmatrix} \succcurlyeq \begin{pmatrix} 0.4 & 0.6 \\ 0.4 & 0.6 \end{pmatrix} \implies \alpha \begin{pmatrix} 0.2 & 0.8 \\ 0.2 & 0.8 \end{pmatrix} + (1-\alpha) \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \succcurlyeq \alpha \begin{pmatrix} 0.4 & 0.6 \\ 0.4 & 0.6 \end{pmatrix} + (1-\alpha) \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}.$$

The next axiom strengthens A5 by relaxing the requirement that the initial profiles must be equal-opportunity profiles. Because $\succcurlyeq$ is assumed to be a continuous weak order, it implies A5, but the converse does not hold.

Axiom 5' (Strong independence to common equal opportunities). *For all $f, f' \in \mathcal{F}_{\mathcal{S}}$, all $\pi, \pi' \in \Delta(\mathcal{Y})$ and all $\alpha \in (0,1)$,* $\alpha f + (1-\alpha) f_{\pi} \succcurlyeq \alpha f' + (1-\alpha) f_{\pi} \Rightarrow \alpha f + (1-\alpha) f_{\pi'} \succcurlyeq \alpha f' + (1-\alpha) f_{\pi'}$.

Consider the following example:

$$\begin{aligned} & \alpha \begin{pmatrix} 0.2 & 0.8 \\ 0.3 & 0.7 \end{pmatrix} + (1-\alpha) \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \succcurlyeq \alpha \begin{pmatrix} 0.4 & 0.6 \\ 0.1 & 0.9 \end{pmatrix} + (1-\alpha) \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \\ \Longrightarrow \quad & \alpha \begin{pmatrix} 0.2 & 0.8 \\ 0.3 & 0.7 \end{pmatrix} + (1-\alpha) \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \succcurlyeq \alpha \begin{pmatrix} 0.4 & 0.6 \\ 0.1 & 0.9 \end{pmatrix} + (1-\alpha) \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

4. Representation with a uniform type distribution

We derive the main representation results in the baseline setting with fixed demographic type weights and relate them to the existing normative economics literature. All proofs of theorems and properties presented in the main text are provided in Appendix A. The connections to comparable results in decision theory under ambiguity are detailed in Appendix B.

Theorem 1. *Let $q(s) = 1/|\mathcal{S}|$ for all $s \in \mathcal{S}$, and $|\mathcal{S}| \geq 3$.*[3] *The social planner's preference relation $\succcurlyeq$ satisfies A1-A5 iff it can be represented by a function $W_{\phi,u,q} : \mathcal{F}_{\mathcal{S}} \to \mathbb{R}$, defined by:*

$$W_{\phi,u,q}(f) = \phi^{-1}\left(\sum_{s \in \mathcal{S}} q(s)\phi\left(U(\pi_s)\right)\right), \tag{1}$$

where $U(\pi_s) = \sum_{y \in \mathcal{Y}} \pi_s(y)u(y)$ is the expected utility of type s with a utility function $u : \mathcal{Y} \to \mathbb{R}$, continuous and non-constant, and ϕ is a strictly increasing and weakly concave function. Moreover, two functions $W_{\phi,u,q}$ and $W_{\phi',u',q}$ represent the same preferences iff there exist $a, A > 0$ and $b, B \in \mathbb{R}$ such that $u' = au + b$ and $\phi'(av + b) = A\phi(v) + B$, for all v in the range of U.

Theorem 1 provides a social welfare function based on a two-stage evaluation: Expected utility is first computed within each type, and these type-specific welfare levels are then aggregated through an in-

[3] Theorem 1 and all other theorems in this section also work for $|\mathcal{S}| = 2$ when invoking an additional axiom called the 'Thomsen condition' (see Wakker, 1988; Debreu, 1960).

creasing and concave transformation capturing inequality-of-opportunity aversion. The first-stage utility function is left entirely unrestricted. Moreover, by treating the income distribution of their own type as the relevant income prospect of individuals, our framework naturally aligns with the ex-ante equality of opportunity literature (Fleurbaey and Peragine, 2013).

The social welfare function presented in Theorem 1 shares the 'generalized utilitarian' structure found in Grant *et al.* (2010), while placing it on different normative foundations. Grant *et al.* (2010) operate within the Harsanyi tradition, focusing on an impartial observer's preferences over hypothetical identity lotteries ('accidents of birth') versus objective risks ('life chances'), while explicitly assuming that individuals are von Neumann-Morgenstern expected utility maximizers. In their framework, the observer's preferences are defined over the entire product space of outcome lotteries and identity lotteries, and their identity-lottery independence axiom (Axiom 3, page 1947) requires the observer to evaluate probability mixtures of different identity distributions. By contrast, our framework abstracts from subjective individual preferences, and takes the perspective of an opportunity-egalitarian social planner who evaluates the objective opportunity structure of society.

Moreover, we fix the partition of types and their demographic weights $q \in \Delta(\mathcal{S})$ at the outset of the axiomatic derivation. Because we do not require the planner to evaluate probability mixtures of different demographic distributions, our row separability condition (A2) is structurally less demanding than the identity-lottery independence axiom used by Grant *et al.* (2010). Our approach ensures that the initial axioms strictly capture the normative evaluation of conditional income distributions, unconfounded by variations in population structure. We address the distinct normative issue of comparing societies with different type partitions and demographic type weights separately in Section 5.

Along similar lines, our framework departs from the approach proposed by Fleurbaey (2010) for assessing risky social situations. While Fleurbaey (2010) introduces the elegant 'expected equally distributed equivalent' (EEDE) class of criteria, he assumes that individual ex-ante preferences are correctly measured by expected utility (as in Grant *et al.* (2010), where the authors impose von Neumann-Morgenstern rationality on individuals from the outset). In contrast, our column separability condition (A5) imposes independence directly on the social planner's preferences, but restricts it to the subset of equal-opportunity profiles. This ensures that the expected utility representation of a type emerges endogenously from the planner's evaluation of aggregate social risk in the absence of inequality of opportunity, rather than being assumed as a behavioral premise.

Comparable structures for assessing inequality of opportunity have recently been explored by Fleurbaey *et al.* (2024) and Andreoli *et al.* (2026). To cleanly isolate their respective normative criteria, these frameworks proceed by treating types symmetrically. For instance, Fleurbaey *et al.* (2024) propose a

nested 'between-group prioritarian' social welfare function while assuming that types share the same size. Similarly, Andreoli *et al.* (2026) evaluate opportunity allocations using a uniform expected utility valuation, effectively weighting all types equally. By leaving the issue of arbitrary demographic type weights unaddressed in their core axiomatic characterizations, these approaches are limited in their direct applicability to societies with heterogeneous type distributions.

Theorem 1 provides a fairly general expression for the social welfare function in the double-expectation form. From a practical perspective, however, this expression is too general, as it assumes a rather arbitrary choice of functions u and ϕ. Replacing A5 by the stronger version A5' restricts ϕ to a single-parameter function.

Theorem 2. *Let $q(s) = 1/|\mathcal{S}|$ for all $s \in \mathcal{S}$, and $|\mathcal{S}| \geq 3$. The social planner's preferences relation $\succcurlyeq$ satisfies A1-A4 and A5' iff it can be represented by a function $V_{\theta,u,q} \equiv W_{\phi_\theta,u,q}$, with $W_{\phi_\theta,u,q}$ as defined in Theorem 1, and with the following function ϕ_θ:*

$$\phi_\theta(t) = \begin{cases} t\,, & \text{for } \theta = 0\,, \\ -\exp\left(-\theta t\right)\,, & \text{for } \theta > 0\,. \end{cases} \tag{2}$$

Moreover, two functions $V_{\theta,u,q}$ and $V_{\theta',u',q}$ represent the same preferences iff there exist $a > 0$ and $b \in \mathbb{R}$ such that $u' = au + b$ and $\theta' = \theta/a$. [4]

In Theorem 2, the function ϕ becomes translation-scale invariant and takes a single-parameter exponential form. The resulting structure is comparable to that identified by Blackorby and Donaldson (1982). However, their function ϕ_θ depends on individual utilities, whereas in our case, it depends on the expected utilities of types. We obtain the following representation:

$$V_{\theta,u,q}(f) = \begin{cases} \sum_{s\in\mathcal{S}} q(s)U(\pi_s)\,, & \text{for } \theta = 0\,, \\ (-1/\theta)\ln\left(\sum_{s\in\mathcal{S}} q(s)\left(e^{-\theta U(\pi_s)}\right)\right)\,, & \text{for } \theta > 0\,. \end{cases} \tag{3}$$

In this work, we abstract from individual preferences and focus exclusively on the social planner's normative judgments. As a result, the functions u and ϕ are not completely independent. Theorem 8 in Appendix C shows that the usual additional requirements of monotonicity (A9) and scale invariance (A10) imply that the utility function u must be logarithmic in the presence of inequality-of-opportunity aversion ($\theta > 0$). These additional requirements are not particularly restrictive for applications that consider monetary outcomes. Therefore, we also use the representation of Theorem 8 when applying

[4] $\phi_\theta^{-1}(t) = -(1/\theta)\ln(-t)$ if $\theta \neq 0$ and $\phi_\theta^{-1}(t) = t$ otherwise.

our framework to the intergenerational mobility estimates in Section 8.

5. Representation on the extended domain

We extend the framework to compare societies $(f, q) \in \mathcal{F}_{\mathcal{S}} \times \Delta(\mathcal{S})$ with potentially different populations, type structures, and non-uniform demographic type weights q. Hence, we lift the restriction that types are equally sized subgroups of the population, which may be restrictive for empirical applications.

Let $\mathcal{I}$ be the (countably infinite) set of potential individuals. Any population is therefore a finite subset $\mathcal{N} \subset \mathcal{I}$. All possible sets of types can be written as $\mathcal{T} = \{\mathcal{S} \mid \exists \mathcal{N} \subset \mathcal{I} \,:\, \mathcal{S} \text{ is a partition of } \mathcal{N}\}$. Note that this setup encompasses both the case where types correspond to groups of individuals and the limiting case in which $\mathcal{S}$ uniquely identifies each individual. The set of all possible societies is denoted by:

$$\mathcal{D} = \bigcup_{\mathcal{S} \in \mathcal{T}} (\mathcal{F}_{\mathcal{S}} \times \Delta(\mathcal{S})) \ . \tag{4}$$

The social planner's preference $\succcurlyeq^{\star}$ is now defined on the extended domain $\mathcal{D}$, endowed with the disjoint union topology (where each component $\mathcal{F}_{\mathcal{S}} \times \Delta(\mathcal{S})$ carries the product of their respective topologies), and $(f, q) \succcurlyeq^{\star} (f', q')$ means that society (f, q) is weakly preferred to society (f', q'). We denote by $\sim^{*}$ and $\succ^{*}$ its symmetric and asymmetric parts. As with the local preference $\succeq$ defined in Section 3, we assume that $\succcurlyeq^{\star}$ is a continuous weak order on $\mathcal{D}$, which directly implies continuity with respect to both opportunity profiles f and demographic type weights q within any fixed type partition $\mathcal{S}$. We introduce three additional axioms to characterize the social planner's preferences on the extended domain. To do so, we now formally treat the local preference introduced in Section 3 as a family of relations $\{\succcurlyeq_{\mathcal{S}}\}_{\mathcal{S} \in \mathcal{T}}$, where each $\succcurlyeq_{\mathcal{S}}$ evaluates profiles on a specific partition $\mathcal{S}$.

The first of these axioms requires that, when two societies are based on the same type partition and have a uniform distribution across types, the preference on the extended domain coincides with the local preference. This axiom allows the locally defined axioms of Section 3 to be imported into the general framework.

Axiom 6 (Consistency)**.** *For all $f, f' \in \mathcal{F}_{\mathcal{S}}$ and all $\mathcal{S} \in \mathcal{T}$, if $q(s) = 1/|\mathcal{S}|$ for all $s \in \mathcal{S}$, then $(f, q) \succcurlyeq^{\star} (f', q) \;\Leftrightarrow\; f \succcurlyeq_{\mathcal{S}} f'$.*

The second of these axioms is Dalton's principle of populations, which states that an identical replication of the population does not alter social welfare. Let $(f, q) \in \mathcal{F}_{\mathcal{S}} \times \Delta(\mathcal{S})$ be a society defined on a finite

population $\mathcal{N}$, where $\mathcal{S}$ is a partition of $\mathcal{N}$. We say that $(f', q') \in \mathcal{F}_{\mathcal{S}'} \times \Delta(\mathcal{S}')$ defined on a population $\mathcal{N}'$ is a 'k-fold replication' of (f, q), with $k \geq 1$, if there exists a bijection $b : \mathcal{S} \to \mathcal{S}'$ such that for all $s \in \mathcal{S}$ we have $|b(s)| = k \cdot |s|$, $q'(b(s)) = q(s)$ and $f'(b(s)) = f(s)$. In other words, both societies have the same number of types, and each type in $\mathcal{S}$ can be uniquely mapped to a corresponding type in $\mathcal{S}'$ that contains exactly k times as many individuals, while maintaining the exact same demographic weight and type-specific income distribution.

Axiom 7 (Principle of Populations). *For all $(f, q) \in \mathcal{F}_{\mathcal{S}} \times \Delta(\mathcal{S})$ and all $(f', q') \in \mathcal{F}_{\mathcal{S}'} \times \Delta(\mathcal{S}')$, if (f', q') is a k-fold replication of (f, q), then $(f, q) \sim^{\star} (f', q')$.*

The third of these axioms states that splitting a type into two subgroups with the same income distribution as the original type leaves social welfare unchanged.

Axiom 8 (Invariance to Type-Split). *For all $(f, q) \in \mathcal{F}_{\mathcal{S}} \times \Delta(\mathcal{S})$ and all $(f', q') \in \mathcal{F}_{\mathcal{S}'} \times \Delta(\mathcal{S}')$, if the following three conditions hold, then $(f, q) \sim^{\star} (f', q')$:*

(i) $\mathcal{S}$ is obtained from $\mathcal{S}'$ by replacing one type $s_{ab} \in \mathcal{S}'$ with two distinct types $s_a, s_b \notin \mathcal{S}'$,

(ii) For all types $s \in \mathcal{S}' \setminus \{s_{ab}\}$ we have $q(s) = q'(s)$, and $\pi_s = \pi'_s$,

(iii) $q(s_a) + q(s_b) = q'(s_{ab}) > 0$, and $\pi_{s_a} = \pi_{s_b} = \pi'_{s_{ab}}$.

Consider the following example:

$$(f, q) = \begin{array}{c} \\ s_a \\ s_b \end{array} \begin{array}{c} \begin{array}{cc} \$1 & \$2 \end{array} \\ \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 0 \end{pmatrix} \end{array}, \begin{pmatrix} 0.2 \\ 0.5 \\ 0.3 \end{pmatrix} \quad \sim^{\star} \quad \begin{array}{c} \\ s_{ab} \end{array} \begin{array}{c} \begin{array}{cc} \$1 & \$2 \end{array} \\ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{array}, \begin{pmatrix} 0.2 \\ 0.8 \end{pmatrix} = (f', q') .$$

These three axioms suffice to establish that the representation of local preferences over opportunity profiles $f \in \mathcal{F}_{\mathcal{S}}$ obtained in Section 4 can also represent preferences in the extended domain over societies $(f, q) \in \mathcal{D}$. We add these axioms to those used in Theorem 1 to obtain our next result.[5]

Theorem 3. *Assume that for each $\mathcal{S} \in \mathcal{T}$, the local preference $\succcurlyeq_{\mathcal{S}}$ over uniform partitions satisfies A1-A5. The social planner's preferences relation $\succcurlyeq^{\star}$ satisfies A6-A8 iff, for all $(f, q) \in \mathcal{D}$, it can be represented by the function $W_{\phi,u}(f, q) = \phi^{-1}\left(\sum_{s \in \mathcal{S}} q(s)\phi\left(U(\pi_s)\right)\right)$, with ϕ and u as characterized in Theorem 1.*

[5] Note that while the local representation in Theorem 1 strictly requires at least three essential types to ensure additive separability, Theorem 3 applies to all societies in $\mathcal{D}$, including those with only one or two types. This is a direct consequence of A8: Any society with fewer than three types can be viewed as ordinally equivalent to a society with three or more identical sub-types, thereby inheriting the additive structure without additional axioms.

Note that this result is based on Theorem 1, which is the most general. However, it also applies to the social welfare functions obtained in Theorem 2 and Theorem 8 (Appendix C). We denote the social welfare functions characterized in Theorem 2, but applied to the extended domain, by $V_{\theta,u}(f,q)$.

6. Equivalent representations and stochastic dominance

6.1. Equivalent representations

We introduce two equivalent representations and an approximation of the social welfare function $V_{\theta,u}(f,q)$, to illustrate different perspectives on how aversion to inequality of opportunity is captured in our framework. To establish the first equivalent representation, we define Kullback-Leibler divergences.

Definition 1. *Consider two distributions $p, q \in \Delta(\mathcal{S})$, with p absolutely continuous to q. The set of such distributions p is denoted $\Delta^q(\mathcal{S})$. A Kullback-Leibler divergence D_{KL} is defined by:* [6]

$$D_{KL}(p||q) = \sum_{s\in\mathcal{S}} p(s) \ln \frac{p(s)}{q(s)} \,.$$

Kullback-Leibler divergences are a special case of a more general class called Bregman divergences (see Magdalou and Nock, 2011, and Definition 2 below).[7] An important property of Kullback-Leibler divergences is that $D_{KL}(p||q) \geq 0$, with equality if and only if $p = q$. By letting $U = (U(\pi_s))_{s\in\mathcal{S}}$ be the list of expected utilities of all types in $f \in \mathcal{F}_{\mathcal{S}}$ distributed according to $q = (q(s))_{s\in\mathcal{S}}$, and recalling that $\mathbb{E}_q[U] = \sum_{s\in\mathcal{S}} q(s)U(\pi_s)$, we obtain the following result.

Theorem 4. *An equivalent representation of the social welfare function $V_{\theta,u}(f,q)$ is:*

$$V_{\theta,u}(f,q) = \begin{cases} \mathbb{E}_q[U] \,, & \text{for } \theta = 0 \,, \\ \min_{p\in\Delta^q(\mathcal{S})} \left[\mathbb{E}_p[U] + \frac{1}{\theta} D_{KL}(p||q)\right] \,, & \text{for } \theta > 0 \,, \end{cases} \tag{5}$$

where $p \in \Delta^q(\mathcal{S})$. The optimal weights $p^\star_{\theta,u}(f,q;\cdot)$ that solve this minimization problem are such that, for all $s \in \mathcal{S}$:

$$p^\star_{\theta,u}(f,q;s) = \frac{q(s)e^{-\theta U(\pi_s)}}{\sum_{s\in\mathcal{S}} q(s)\left(e^{-\theta U(\pi_s)}\right)} \,. \tag{6}$$

[6] By absolute continuity, $q(s) = 0$ implies $p(s) = 0$. Moreover, in empirical analyses, it is generally the case that $q(s) > 0$ for all $s \in \mathcal{S}$. Because $\lim_{t\to 0^+} t \ln t = 0$, the contribution of state $s \in \mathcal{S}$ to the divergence $D_{KL}(p||q)$ is zero if $p(s) = 0$.

[7] Hufe *et al.* (2022) apply them as an index of inequality of opportunity.

While $q(s)$ is the demographic weight of type s in the population, $p(s)$ can be interpreted as a normative weight applied to type s by the social planner. Equation (6) furthermore shows that optimal type weights $p^{\star}_{\theta,u}(f,q;s)$ are negatively related to types' expected utilities $U(\pi_s)$ and that the strength of this negative relationship is governed by the inequality-of-opportunity aversion parameter θ. Hence, while the representation in Theorem 2 captures inequality-of-opportunity aversion by a concave transformation of types' expected utilities $U(\pi_s)$, the dual representation of Theorem 4 captures inequality-of-opportunity aversion through a multiplicative weighting of $U(\pi_s)$ with normative weights $p(s)$.

To establish the following equivalent representation of $V_{\theta,u}(f,q)$, we define Bregman divergences.

Definition 2. *Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a differentiable and strictly convex function, with derivative $\varphi^{(1)}$. The Bregman divergence with generator φ, denoted D_φ, between two real numbers x and y, is defined by $D_\varphi(x||y) = \varphi(x) - \varphi(y) - (x-y)\varphi^{(1)}(y)$.*

A Bregman divergence is a measure of 'distance' between two quantities, called a divergence because it does not satisfy the triangular inequality. Like the Kullback-Leibler divergence, a Bregman divergence satisfies $D_\varphi(x||y) \geq 0$ for all $x, y \in \mathbb{R}$, and $D_\varphi(x||y) = 0$ if and only if $x = y$. We have:[8]

Theorem 5. *An equivalent representation of the social welfare function $V_{\theta,u}(f,q)$ is:*

$$V_{\theta,u}(f,q) = \begin{cases} \mathbb{E}_q[U]\,, & \text{for } \theta = 0\,, \\ \underbrace{\mathbb{E}_q[U]}_{\text{Efficiency}} - \underbrace{{}^{1}\!/_{\theta} D_{K_U}(-\theta||0)}_{\text{Inequality of opportunity}}\,, & \text{for } \theta > 0\,, \end{cases} \tag{7}$$

with $K_U(\tau) = \ln\left(\mathbb{E}_q[e^{\tau U}]\right)$ for any $\tau \in \mathbb{R}$. Moreover, when $\theta > 0$ is close to zero, or when the dispersion of expected utilities across types is small, we have $V_{\theta,u}(f,q) \approx \mathbb{E}_q[U] - \frac{\theta}{2}\,\mathrm{var}_q(U)$.

Theorem 5 provides a closed-form representation, which we call the 'mean-divergence' criterion. In this representation, disparities in type-specific expected utilities are exactly captured by the Bregman divergence D_{K_U}. Given the well-established statistical properties of Bregman divergences (Magdalou and Nock, 2011), this representation makes the planner's trade-off between efficiency and inequality of opportunity explicit.

Theorem 5 also shows that $V_{\theta,u}(f,q)$ can be approximated by a 'mean-variance' criterion. In this representation, the weight assigned to disparities across types, as measured by the variance $\mathrm{var}_q(U)$, is

[8] Note that if type-specific expected utilities are constant in Theorem 5, the function K_U becomes affine and the divergence trivially collapses to zero ($D_{K_U} = 0$). This reflects the economic intuition: The penalty vanishes whenever types face equivalent expected welfare. Otherwise, when U is non-degenerate, K_U is strictly convex and generates a proper Bregman divergence.

directly proportional to the planner's degree of inequality-of-opportunity aversion. In this representation, the inequality-of-opportunity aversion parameter θ plays a role analogous to the Arrow-Pratt coefficient of absolute risk aversion in the decision theory literature. The closed form of the mean-divergence criterion based on the Bregman divergence goes beyond the local mean-variance approximation by accounting for all moments of the distribution of types' expected utilities.[9]

6.2. Stochastic dominance

We establish a stochastic dominance criterion that allows for unambiguous ranking of societies for a wide class of social planners' preferences over inequality of opportunity. For any given value of the inequality-of-opportunity aversion parameter θ, two societies can always be ranked; however, the ranking may change as this parameter varies. Therefore, we are interested in the preorder relation which corresponds to the dominance of a society over another, for all admissible values of θ. To establish our stochastic dominance result, the following definition is needed.

Definition 3. *A function $\varphi : (0, \infty) \to \mathbb{R}$ is said completely monotone if all the derivatives $\varphi^{(n)}$ of φ exist, and they satisfy $(-1)^{(n)}\varphi^{(n)}(x) \geq 0$ for all integers $n \geq 0$, with $\varphi^{(0)} = \varphi$. A function $\varphi : (0, \infty) \to \mathbb{R}$ such that $\varphi^{(1)}$ is completely monotone is said completely alternating.*[10]

The notion of completely alternating functions has been investigated in the context of stochastic dominance for risky decisions (see Whitmore, 1989). In this literature, the utility function of an expected utility maximizer is typically assumed to be increasing and concave (positive first derivative and negative second derivative, respectively). The commonly used CARA and CRRA utility functions have these properties and, more generally, are completely alternating (positive third derivative, negative fourth derivative, and so on).

For the social welfare function $V_{\theta,u}(f, q)$, the social planner transforms the expected utility $U(\pi_s)$ of each type $s \in \mathcal{S}$ using the function ϕ_θ defined in Equation (2), which is a particular completely alternating function. This raises the question of whether ϕ_θ can be replaced with alternative completely alternating functions. In Definition 3, we restrict the domain of these transformations to $(0, \infty)$, which implicitly assumes that all expected utilities are strictly positive. This restriction is made without loss of generality, because the underlying utility function u is defined up to an increasing affine transfor-

[9] The following expression highlights the analogy between the Bregman divergence and the variance. For any $\theta > 0$, we have: $D_{K_U}(-\theta||0) = \ln \mathbb{E}_q\left[e^{-\theta U}\right] - \mathbb{E}_q[\ln e^{-\theta U}]$. Hence, the Bregman divergence uses the logarithm applied to an exponential transformation of utility, whereas the variance uses the square function.

[10] Adding the restriction $\varphi \geq 0$, it is called a Bernstein function.

mation: As the income domain $\mathcal{Y}$ is compact, any continuous utility function u attains a finite global minimum on $\mathcal{Y}$, allowing a shift of u to ensure strict positivity.

Theorem 6. *Let $(f,q),(f',q') \in \mathcal{D}$ be two societies defined respectively on type partitions $\mathcal{S}$ and $\mathcal{S}'$. Assuming, without loss of generality, that the utility function u takes strictly positive values, the following two statements are equivalent:*

(i) $\sum_{s\in\mathcal{S}} q(s)\varphi\left(U(\pi_s)\right) \geq \sum_{s\in\mathcal{S}'} q'(s)\varphi\left(U(\pi'_s)\right)$, *for all* $\varphi : (0,\infty) \to \mathbb{R}$ *completely alternating*,

(ii) $V_{\theta,u}(f,q) \geq V_{\theta,u}(f',q')$, *for all* $\theta \geq 0$.

Theorem 6 establishes that the stochastic dominance of society (f,q) over society (f',q') according to $V_{\theta,u}$ with a fixed first-stage utility function u, extends to a broader class of social welfare functions than the one-parameter exponential family defined in Equation (2).[11] Specifically, it implies dominance for any evaluation where expected utilities are transformed by a completely alternating function. This includes standard transformations such as CARA and CRRA (restricted to strictly positive aversion parameters), as well as the logarithmic function. While completely alternating functions form a proper subclass of the general increasing and weakly concave transformations characterized in Theorem 1, they still cover a wide range of normative judgments on how to incorporate inequality-of-opportunity aversion when comparing societies. Moreover, when adopting logarithmic utility (Theorem 8), Statement (ii) of Theorem 6 depends on a single parameter, allowing for easy implementation in empirical applications.

7. Properties

We present the main properties satisfied by the social welfare function $V_{\theta,u}(f,q)$.

Property 1. *If $u(y) \geq 0$ for all $y \in \mathcal{Y}$, then $V_{\theta,u}(f,q) \geq 0$.*

Property 1 shows that the non-negativity of the social welfare function follows from the simple requirement that u be positive. Because u is defined up to an increasing affine transformation, this requirement entails no loss of generality, as discussed in the previous section.

Property 2. $\frac{\partial V_{\theta,u}}{\partial U(\pi_s)}(f,q) = p^{\star}_{\theta,u}(f,q;s)$.

[11] In stochastic dominance theory, preorder (i) is called a 'completion' of preorder (ii).

Property 2 establishes that the marginal effect of the expected utility of type $s \in \mathcal{S}$ equals $p^\star_{\theta,u}(f,q;s)$, i.e., the normative weight that the social planner assigns to that type. If the social planner is inequality-of-opportunity neutral ($\theta = 0$), then $p^\star_{\theta,u}(f,q;s) = q(s)$, meaning that types' expected utilities are weighted according to their demographic weights.

Property 3. $\frac{\partial V_{\theta,u}}{\partial u(y)}(f,q) = \pi^\star_{\theta,u}(f,q;y)$, *where* $\pi^\star_{\theta,u}(f,q;y) = \sum_{s\in\mathcal{S}} p^\star_{\theta,u}(f,q;s)\pi_s(y)$.

Property 3 identifies the impact of the utility associated with income $y \in \mathcal{Y}$ on social welfare. In the overall population, $\pi(y) = \sum_{s\in\mathcal{S}} q(s)\pi_s(y)$ represents the proportion of individuals earning income y. If the social planner is inequality-of-opportunity neutral ($\theta = 0$), this is exactly the marginal effect of $u(y)$ on social welfare, since $p^\star_{\theta,u}(f,q;s) = q(s)$. Otherwise, if the planner is inequality-of-opportunity averse ($\theta > 0$), the demographic weights $q(s)$ are replaced by the normative weights $p^\star_{\theta,u}(f,q;s)$, which transforms $\pi(y)$ into the weighted average $\pi^\star_{\theta,u}(f,q;y)$.

Property 4. $V_{0,u}(f,q) = \sum_{s\in\mathcal{S}} q(s)U(\pi_s) = \sum_{y\in\mathcal{Y}} \pi(y)u(y)$, *and*
$V_{\infty,u}(f,q) = \min_{s\in\mathcal{S}:q(s)>0} U(\pi_s)$.

Property 4 illustrates the role of the inequality-of-opportunity aversion parameter through its polar cases. At $\theta = 0$, the criterion collapses to standard utilitarianism: Social welfare reduces to the expected utility of the aggregate income distribution. Conversely, as $\theta \to \infty$, the planner adopts a strict Maximin perspective, focusing solely on the worst-off type. Assuming an affine u, the function $V_{\infty,u}(f,q)$ recovers the opportunity-sensitive measure of van de Gaer (1993), namely $\min_{s\in\mathcal{S}} \mu(\pi_s)$.

The following result further clarifies the role of θ in our framework beyond the polar cases discussed above. For a given society $(f,q) \in \mathcal{D}$, we simplify notation by writing $p^\star_{\theta,u}(f,q;s) = p^\star_{\theta,u}(s)$ in what follows.

Theorem 7. *Let* $\theta, \theta' \geq 0$. *The following three statements are equivalent:*

(i) $\theta < \theta'$,

(ii) $V_{\theta,u}(f,q) > V_{\theta',u}(f,q)$, *for all societies* $(f,q) \in \mathcal{D}$ *exhibiting inequality of opportunity, i.e., where there exist* $s, s' \in \mathcal{S}$ *with* $q(s), q(s') > 0$ *and* $U(\pi_s) \neq U(\pi_{s'})$,

(iii) For any two-type society with $q = (1/2, 1/2)$ *and an opportunity profile* $f = (\pi_s, \pi_{s'}) \in \mathcal{F}_\mathcal{S}$ *such that* $U(\pi_s) < U(\pi_{s'})$, *we have* $p^\star_{\theta',u}(s) > p^\star_{\theta,u}(s) \geq p^\star_{\theta,u}(s') > p^\star_{\theta',u}(s')$.

Theorem 7 establishes that the higher the inequality-of-opportunity aversion parameter θ, the lower the social welfare. This reflects a growing concern about disparities in expected utilities between types.

Statement (iii) of the theorem concerns a simple society profile consisting of two types, s and s', represented with equal population shares. Recalling that $p^\star_{\theta,u}(s)$ measures the marginal impact of the expected utility of type s (Property 2), if $U(\pi_s) < U(\pi_{s'})$, then $p^\star_{\theta,u}(s) \geq p^\star_{\theta,u}(s')$, with equality only if $\theta = 0$. If inequality-of-opportunity aversion increases from θ to θ', then the gap between normative weights increases. This result complements Property 4, which states that, if θ tends to infinity in such a society with two types, then $p^\star_{\theta,u}(s)$ tends to 1 and $p^\star_{\theta,u}(s')$ tends to 0.

Figure 1 summarizes the properties of our framework in graphical form using the mean-divergence representation from Theorem 5.[12] Panels (a)–(b) illustrate changes in $V_{\theta,u}(f,q)$ resulting from shifts in the normative parameter θ. Panel (a) shows a baseline case without inequality-of-opportunity aversion: if $\theta = 0$, normative weights $p^\star_{\theta,u}(s)$ correspond to demographic weights $q(s)$, and social welfare $V_{\theta,u}(f,q)$ is equal to the efficiency component $\mathbb{E}_q[U]$, i.e., the average of expected type utilities. Panel (b) introduces inequality of opportunity aversion: if $\theta > 0$, normative weights $p^\star_{\theta,u}(s)$ diverge from demographic weights $q(s)$, with higher weights placed on types with lower expected type utilities. As a result, $V_{\theta,u}(f,q)$ is lower than $\mathbb{E}_q[U]$ with the difference equal to the inequality of opportunity component ${}^1\!/_\theta D_{K_U}$, i.e., the Bregman divergence of expected type utilities.

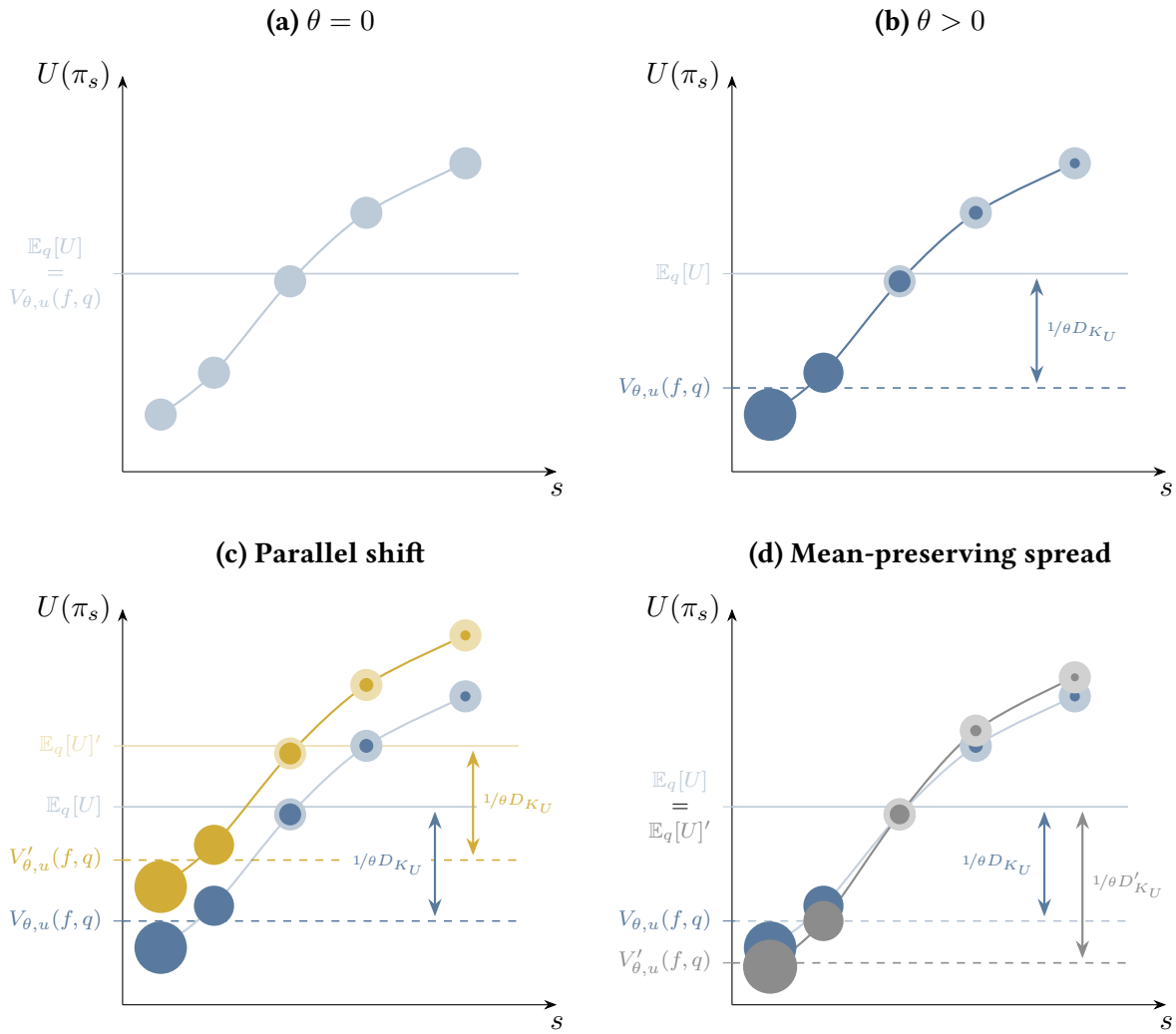
Panels (c)–(d) illustrate changes in $V_{\theta,u}(f,q)$ resulting from shifts in empirical distributions. Panel (c) shows the effect of a parallel shift in type outcomes: for a given $\theta > 0$, social welfare $V_{\theta,u}(f,q)$ increases, whereas inequality of opportunity ${}^1\!/_\theta D_{K_U}$ remains unaffected. Panel (d) illustrates a mean-preserving spread in type outcomes: for a given $\theta > 0$, social welfare $V_{\theta,u}(f,q)$ decreases due to an increase in inequality of opportunity ${}^1\!/_\theta D_{K_U}$, whereas efficiency $\mathbb{E}_q[U]$ remains unaffected.

8. Opportunity-sensitive social welfare in the United States

In this section, we take our axiomatic framework to the data and evaluate opportunity-sensitive social welfare in the United States. First, we assess national differences across time by comparing birth cohorts from the 1940s to the 1980s. Second, we examine regional differences for cohorts born around 1980 by comparing commuting zones (CZ). In both applications, we define types $s \in \mathcal{S}$ based on measures of parental income, thereby linking our analysis to the extensive literature on intergenerational income mobility in the US. Our analysis proceeds based on the additive decomposition of social welfare into efficiency $\mathbb{E}_q[U]$ and inequality of opportunity ${}^1\!/_\theta D_{K_U}$ (Theorem 5). As illustrated above, the quantitative impact of inequality of opportunity on social welfare depends on the degree of inequality-

[12] For the sake of illustration, we assume uniform demographic type weights $q(s)$, impose logarithmic utility, and order types s by expected type utilities $U(\pi_s)$.

Figure 1: Comparative statics of $V_{\theta,u}(f,q)$



Note: This figure illustrates the properties of our measures based on the mean-divergence representation from Theorem 5: Social welfare $V_{\theta,u}(f,q)$ decomposes into efficiency, i.e., expected type utilities $\mathbb{E}_q[U]$ and inequality of opportunity, i.e., the θ-weighted Bregman divergence in expected type utilities $1/\theta D_{K_U}$.

of-opportunity aversion θ. We therefore also develop a method for computing dominance sets to rank societies. This allows us to transparently illustrate which welfare rankings are robust to disagreement over the appropriate level of inequality-of-opportunity aversion, and which are subject to normative ambiguity.

8.1. Trend across birth cohorts (1940s–1980s)

Data. The Panel Study of Income Dynamics (PSID) is the longest-running household survey in the world, collecting rich income and family background information for a nationally representative sample

of US households since 1968. Children of sampled PSID households are invited to become separate sampling units once they move out of their parental household, enabling income data to be linked across generations. This feature makes the PSID one of the key data sources for analyses of intergenerational mobility in the US (among others, Solon, 1992; Mello *et al.*, forthcoming; Hufe *et al.*, 2025; Hartley *et al.*, 2022; Vosters and Nybom, 2017; Mazumder, 2018; Nybom and Stuhler, 2024).

We use all available PSID waves (1968–2023) and process the data in accordance with standard practices in the recent intergenerational mobility literature. First, we calculate family incomes for each PSID household in each available year. The focus on family rather than individual income allows for the inclusion of population groups with historically lower labor force attachment rates in the analysis, such as ethnic minorities and women (Jácome *et al.*, 2025). We measure family income as the sum of the household head's and their co-resident spouse's total labor income, deflated to 2019 USD. We exclude family members other than the household head and their spouse from the calculation of family income to avoid mechanical correlations in family incomes between co-resident parents and children. Second, we use the time series on family incomes to construct a measure of parental income y^p during the child's adolescence, consistent with recent evidence emphasizing the importance of parental income during this period on later-life outcomes (Eshaghnia *et al.*, 2024; Carneiro *et al.*, 2021). In particular, we calculate a weighted average of family income in the parental household, with weights w decreasing with the temporal distance to the child's 15th birthday ($w = \max[0; 1 - |\text{Age} - 15|/15]$).[13] Third, we construct a measure of children's income y^c during their prime working years. We focus on an age interval rather than an annual snapshot to smooth out transitory fluctuations in children's income and to account for differences in lifecycle earnings profiles (Mello *et al.*, forthcoming). Specifically, we calculate the average family income in the child's household over ages 30-40, assigning equal weight to each annual observation.[14] We retain only observations with at least three family income observations in the considered age range. Fourth, we drop observations with missing values or incomes of zero for either y^p or y^c. Since we measure family income averaged over several years, the share of individuals affected by this restriction is small. Finally, we reweight the resulting analysis sample to address selective attrition in the PSID (Ward, 2023). In particular, for each decadal birth cohort in our analysis sample, we rescale the PSID cross-sectional weights to match the relative frequencies of sex $\times$ race-cells in the corresponding pooled Current Population Survey (CPS, 1969-2023).

[13] Since the PSID records parental incomes only starting in 1968, parental income in the 1940s cohort is not centered on the child's 15th birthday, and we observe parental incomes at later points in their lifecycle than in the other cohorts. In robustness analyses, we show that our results are largely unchanged when replacing log parental income with parental income quintiles (Table S.2). Assuming that parental income quintiles are relatively stable during this life period, this suggests that differences in the timing of parental income measurement do not drive our results.

[14] Ideally, we would like to calculate this average over an even longer age interval. However, extending the age interval would compromise the comparability of birth cohorts in terms of their age structures, as not all considered cohorts are equally represented at each age.

Table A.1 displays descriptive statistics for our analysis sample separately for each decadal birth cohort from the 1940s (1940-1949) to the 1980s (1980-1989). The sample is gender-balanced in all birth cohorts. The share of non-Hispanic Whites declines substantially, reflecting changes in the ethnic composition of the US. Across all birth cohorts, we measure children's income, on average, close to their 35th birthday. Our variables of interest, parental income y^p and child income y^c, generally increase over time. The increase in child income y^c, however, is moderate relative to GDP per capita growth rates during this period. This discrepancy can be attributed to our unit of analysis: As we consider family income rather than individual incomes, the increasing number of single-person households depresses income growth over time (Lundberg *et al.*, 2016).[15] Furthermore, due to our focus on labor income, we exclude capital income, which has grown particularly strongly over the observation period (Piketty *et al.*, 2018). Generally, the documented cross-cohort trend of family incomes in the child generation is consistent with evidence from other data sources, including the Current Population Survey (CPS) and the National Longitudinal Surveys (NLS) (Davis and Mazumder, 2026; Jácome *et al.*, 2025). Furthermore, we observe pronounced increases in the standard deviation of both y^p and y^c, reflecting the well-documented increase in income inequality in the US (Piketty *et al.*, 2018).

Intergenerational mobility. We use these data to estimate the intergenerational elasticity of income separately for each birth cohort as follows:

$$\ln y_i^c = \alpha + \beta^{IGE} \ln y_i^p + X_i'\delta + \epsilon_i, \tag{8}$$

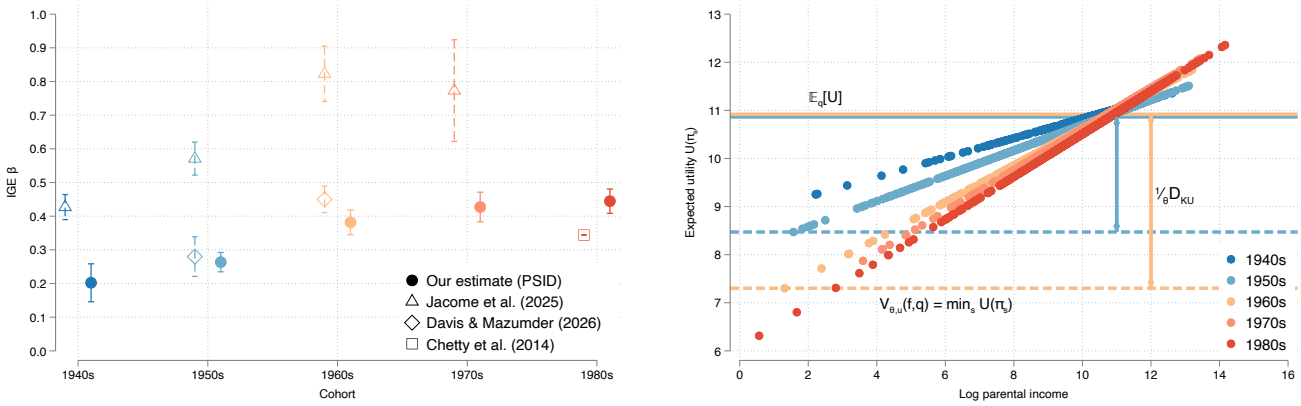
where $\ln y_i^c$ and $\ln y_i^p$ are the logarithms of children's and parents' income, respectively. The vector X_i includes the first- and second-order polynomials of child age to account for lifecycle effects within birth cohorts. ϵ_i is the error term. The coefficient of interest is β^{IGE}, which captures the intergenerational elasticity of income (IGE). This coefficient is often interpreted as a measure of inequality of opportunity (Deutscher and Mazumder, 2023): If it is high, children's incomes are strongly associated with their parents' incomes, limiting the influence of their own efforts and choices on economic outcomes.

Figure 2a shows the evolution of the IGE in the US. Solid markers represent our estimates, whereas hollow markers show estimates from other studies for the birth cohorts we cover. We observe a marked increase in the IGE from 0.20 in the 1940s to 0.44 in the 1980s. This increase is particularly pronounced from the 1950s (0.26) to the 1960s (0.38). The IGE remains relatively stable at a high level across subsequent birth cohorts.

This pattern is consistent with previous literature. Davis and Mazumder (2026) use data from two cohorts of the NLS to document a pronounced increase from the 1950s to the 1960s, which they attribute

[15] We account for the impact of changing household compositions in sensitivity analyses (Table S.3).

Figure 2: Intergenerational mobility across birth cohorts (1940s–1980s)



(a) Intergenerational elasticity $\hat{\beta}^{IGE}$

(b) Expected type utilities $\hat{U}(\pi_s)$

Data: PSID.

Note: This figure shows trends in intergenerational mobility in the US across the birth cohorts 1940s–1980s. Panel (a) shows point estimates and 95% confidence intervals for the intergenerational elasticity β^{IGE} based on Equation (8). Solid markers indicate our estimates; hollow markers indicate comparable estimates from studies cited in the legend. Lower values of $\hat{\beta}^{IGE}$ indicate higher intergenerational mobility. Panel (b) shows the distribution of predicted type utilities $\hat{U}(\pi_s)$ based on Equation (9). The additional horizontal and vertical lines illustrate the decomposition of social welfare $V_{\theta,u}(f,q)$ into efficiency $\mathbb{E}_q[U]$ and inequality of opportunity $1/\theta D_{K_U}$ for the 1950s and 1960s cohorts, based on the mean-divergence representation from Theorem 5, evaluated at $\theta \to \infty$.

to rising inequality in the 1980s, particularly affecting the 1960s birth cohort. This increase and a subsequent stabilization of the IGE are also visible in the estimates of Jácome *et al.* (2025). Their estimates are based on imputing parental income from observable characteristics, leading to an underestimation of the variance of y_i^p and a corresponding upward bias in the IGE estimates (Cortes Orihuela *et al.*, 2025). While the *levels* of their IGE estimates cannot be compared with our estimates based on observed y_i^p, it is reassuring that the *trends* across birth cohorts align. Lastly, we can compare our estimates to Chetty *et al.* (2014a), who base their estimates of the IGE for the birth cohort around 1980 on a large tax data sample. In this data, they observe children at comparatively young ages (29-32) and construct y_i^p and y_i^c as averages over relatively short time intervals of five and two years, respectively. This data structure likely depresses their IGE estimates relative to those based on more comprehensive measures of children's and parents' lifecycle incomes (Mazumder, 2016).

Overall, we conclude that our estimates of IGE trends in the US are consistent with the existing literature, suggesting that intergenerational mobility has decreased for American children born in the post-war period.

Opportunity-sensitive social welfare. The IGE directly maps onto our theoretical framework. To see this, recall that an opportunity-sensitive planner evaluates social welfare by assessing the distribution of

expected type utilities $U(\pi_s)$. These expected type utilities can be calculated using data inputs from IGE estimations: Types $s \in \mathcal{S}$ are defined by levels of log parental income $\ln y^p$, demographic weights $q(s)$ are given by the empirical distribution across $\ln y^p$, and children's utility is captured by log-transformed incomes, consistent with Theorem 8. Thus, using the linearity assumption inherent in the IGE, we can estimate expected type utilities $U(\pi_s)$ as the conditional expectation of $\ln y^c$ given $\ln y^p_s$:[16]

$$\hat{U}(\pi_s) = \mathbb{E}[\ln y^c | \ln y^p_s] = \hat{\alpha} + \hat{\beta}^{IGE} \times \ln y^p_s. \tag{9}$$

Figure 2b translates this mapping into graphical form and provides an empirical analog to the representation of our theoretical framework in Figure 1. Each data point represents the estimated expected utility $\hat{U}(\pi_s)$ for a given parental income level. Cohort-specific slopes correspond to the point estimates $\hat{\beta}^{IGE}$ in Figure 2a, indicating a decrease in intergenerational mobility over time. Focusing on the cohorts of the 1950s and 1960s as an illustrative example, the solid horizontal lines indicate that their average expected type utilities are almost identical. Thus, social welfare differences between these birth cohorts are not driven by efficiency $\mathbb{E}_q[U]$. At the same time, the increase in $\hat{\beta}^{IGE}$ from the 1950s to the 1960s, together with a rise in the dispersion of parental income, induces a spread in $\hat{U}(\pi_s)$, leading to an increase in inequality of opportunity $1/\theta D_{K_U}$. The strength of this change is governed by the extent of inequality-of-opportunity aversion θ. In the figure, we highlight the limiting case of $\theta \to \infty$, i.e., the case where social welfare is determined solely by the outcomes of the most disadvantaged type. In this case, social welfare is substantially higher in the 1950s than in the 1960s.

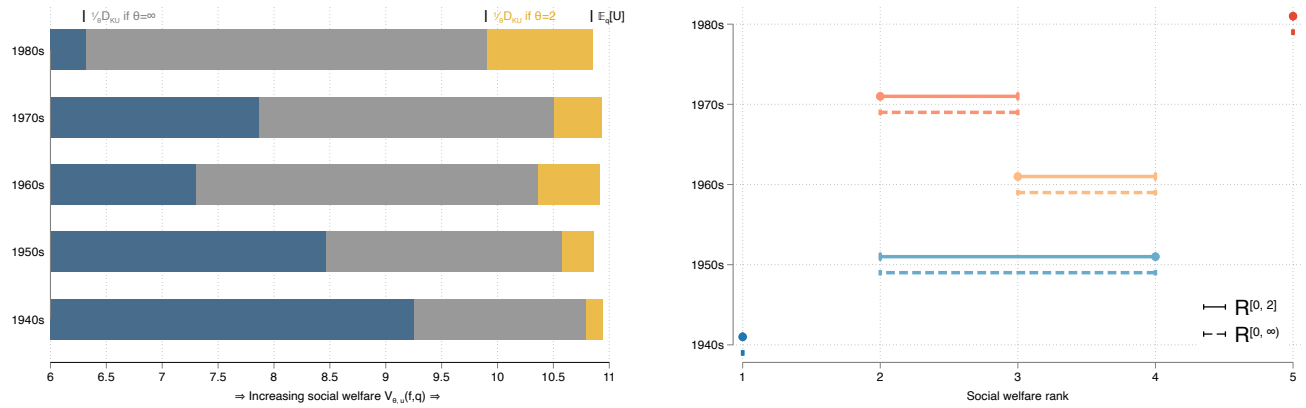
Figure 3a extends this analysis to all birth cohorts in our sample. In particular, we trace social welfare $V_{\theta,u}(f, q)$ for different levels of inequality-of-opportunity aversion θ. The full length of the bars indicates the efficiency component $\mathbb{E}_q[U]$. The stacked bar segments capture the decrease in social welfare due to inequality of opportunity, $1/\theta D_{K_U}$.

The results suggest that the efficiency component $\mathbb{E}_q[U]$ does not vary substantially across birth cohorts in the US: The cross-cohort standard deviation of $\mathbb{E}_q[U]$ is merely 0.039. The attenuated variation in $\mathbb{E}_q[U]$ is unsurprising, given the relative stability of family incomes over time (Table A.1). Instead, cross-cohort differences in social welfare are mostly driven by the inequality-of-opportunity component $1/\theta D_{K_U}$. The corresponding cross-cohort standard deviation of $1/\theta D_{K_U}$ is 0.306 at $\theta = 2$, i.e., roughly eight times larger than for $\mathbb{E}_q[U]$.[17]

[16] We demean age controls within each cohort before estimating Equation (8). Therefore, the level effects of age are absorbed by $\hat{\alpha}$, ensuring that the prediction shown in Equation (9) is evaluated at the cohort's mean age. Generally, age controls have a negligible effect as there is little variation around the average age at which income is measured (Table A.1).

[17] It is even larger at $\theta \to \infty$, when welfare is evaluated at the lowest observed parental income. Welfare rankings are unchanged when types are identified by parental income quintiles, suggesting that the dominance results presented below

Figure 3: Social welfare across birth cohorts (1940s–1980s)



(a) Decomposition

(b) Dominance sets for ranking

Data: PSID.

Note: This figure shows trends in social welfare across the birth cohorts 1940s–1980s. Panel (a) shows the decomposition of social welfare $V_{\theta,u}(f,q)$ into efficiency $\mathbb{E}_q[U]$ and inequality of opportunity $^1/_\theta D_{K_U}$, based on the mean-divergence representation from Theorem 5, evaluated at $\theta = 2$ and $\theta \to \infty$. Panel (b) shows dominance sets for the social welfare ranking based on $\theta \in [0, 2]$ (solid lines) and $\theta \in [0, \infty)$ (dashed lines). Rank 1 (5) corresponds to the cohort with the highest (lowest) social welfare.

Figure 3a also illustrates that cross-cohort rankings are sensitive to the choice of θ. For example, the 1950s are ranked lower than the 1960s for low values of inequality-of-opportunity aversion, whereas this ranking reverses as θ increases. To assess the extent of normative ambiguity in our cross-cohort comparisons, we next propose a method to compute dominance sets for the ranking of societies.

Welfare ranking. In particular, we modify the rank inference procedure of Mogstad *et al.* (2024), who convert confidence intervals on intergenerational mobility coefficients into confidence sets for mobility rankings. In their application, Mogstad *et al.* (2024) focus on the expected income rank of children who grew up with parents at the 25th percentile of the national income distribution. However, their procedure is equally applicable to other mobility statistics such as the IGE. It can be summarized as follows. Denote by $\hat{\psi}_j$ any intergenerational mobility estimate for society $j \in \mathcal{J}$, where a higher $\hat{\psi}$ indicates higher mobility. First, they fix a desired level of statistical significance α. Second, they construct corresponding confidence intervals for all pairwise differences $\hat{\psi}_j - \hat{\psi}_k$. These confidence intervals are simultaneous, i.e., they jointly cover all pairwise differences $\hat{\psi}_j - \hat{\psi}_k$ across societies in $\mathcal{J}$ at level $1 - \alpha$. Third, they count the number of societies $k \neq j$ for which j can be unambiguously classified as having higher (lower) mobility at the chosen level of statistical significance. Denoting by n_j^+ (n_j^-) the number of societies $k \neq j$ for which the simultaneous confidence interval on $\hat{\psi}_j - \hat{\psi}_k$ is entirely above (below) zero, they obtain the confidence set

are not driven by sensitivity to the tail (Table S.2).

$$R_j^{1-\alpha} \;=\; \{\, n_j^- + 1,\; \ldots,\; |\mathcal{J}| - n_j^+ \,\}. \tag{10}$$

The width of this confidence set depends on the degree of statistical uncertainty in intergenerational mobility estimates: If estimates are noisy, simultaneous confidence intervals are wide, few societies k can be placed unambiguously above or below j, and $R_j^{1-\alpha}$ is large, potentially covering the entire range of ranks.

We adapt the logic of this procedure to normative ambiguity, i.e., we derive dominance sets for welfare rankings when the appropriate normative parameter θ is unknown. First, we fix an interval $[\theta_{\min}, \theta_{\max}]$ of admissible values of the aversion parameter. Second, we compute pairwise welfare differences $V_{\theta,u}^j(f,q) - V_{\theta,u}^k(f,q)$ across this interval. Third, denoting by n_j^+ (n_j^-) the number of societies $k \neq j$, for which $V_{\theta,u}^j(f,q) - V_{\theta,u}^k(f,q)$ is positive (negative) for all $\theta \in [\theta_{\min}, \theta_{\max}]$, we construct the dominance set $R_j^{[\theta_{\min},\theta_{\max}]}$. The width of this dominance set depends on the degree to which pairwise welfare orderings vary across $\theta \in [\theta_{\min}, \theta_{\max}]$: If orderings are sensitive to θ, few societies k can be placed unambiguously above or below j, and $R_j^{[\theta_{\min},\theta_{\max}]}$ grows large.

The resulting object is set-identified rather than probabilistic: It illustrates which ranks for j are robust to normative disagreement across $\theta \in [\theta_{\min}, \theta_{\max}]$. The criterion becomes more stringent as the θ-interval widens. In the limit, requiring the welfare ordering to hold for $\theta \in [0, \infty)$, it is equivalent to a ranking by stochastic dominance (Theorem 6).[18] However, the procedure is flexible enough to accommodate less stringent criteria by restricting $[\theta_{\min}, \theta_{\max}]$ to a subrange that reflects a narrower class of normative positions.[19]

Figure 3b shows dominance sets $R^{[0,2]}$ for US birth cohorts from the 1940s to the 1980s. There is no clear ranking among the birth cohorts 1950s–1970s, i.e., their social welfare rankings vary with assumptions about θ. In contrast, the 1940s and 1980s cohorts can be unambiguously classified as having the highest and lowest social welfare, respectively. One may ask whether the unambiguous classification of these two cohorts is an artifact of a rather restrictive interval for the θ parameter. To answer this question, we also consider a more demanding criterion by calculating $R^{[0,\infty)}$. The results show that $R^{[0,2]}$ and $R^{[0,\infty)}$

[18] Note $R_j^{[\theta_{\min},\theta_{\max}]}$ reflects the full range of ranks society j could hold in any complete ranking consistent with the unambiguous pairwise dominance preorder over $[\theta_{\min}, \theta_{\max}]$. This differs conceptually from a sharp attainable rank set, which would include only the realized ranks of society j at each θ. While our approach provides a theoretical outer bound on the attainable rank set, the empirical difference is small–see Tables S.5 and S.10.

[19] To implement this in practice, we use a θ grid sufficiently fine to rule out pairwise rank reversals between neighboring grid nodes. Specifically, the rate of welfare changes with respect to θ can be bounded analytically. At a given rate, a pairwise welfare difference that has the same sign over a short enough θ-interval cannot reach zero in between. We use this insight to check every pairwise difference over the whole grid, refining the nodes as needed, and close the limit $\theta \to \infty$ using the maximin value from Property 4. Therefore, the computed sets exactly reflect continuous-interval dominance.

are fully overlapping across all cohorts, implying that the 1940s (1980s) cohort ranks best (worst) for any possible level of inequality-of-opportunity aversion. Importantly, we recall that while this dominance result is derived under the assumption of logarithmic income utility within types, the ranking extends to a much wider class of second-stage opportunity-aversion functions, including CRRA and CARA (Theorem 6).

Additional analyses. We conduct a series of additional analyses to probe the sensitivity of our results. First, income data from surveys may be distorted by measurement error, particularly in the tails (e.g., Brewer *et al.*, 2017). Table S.1 shows that dominance sets are largely robust to trimming, suggesting that mismeasured tails are not driving our results. Second, in line with the literature on intergenerational mobility, our baseline specification uses a linear mapping between $\ln y^p$ and $\ln y^c$ to estimate expected type utilities (Equation (8)). Table S.2 shows that dominance sets are largely similar whether we estimate $U(\pi_s)$ using second- or third-order polynomials in $\ln y^p$ or use non-parametric estimates based on parental income quintiles. Third, our baseline specification uses total family labor income as the outcome of interest. However, one may argue that this income concept does not adequately capture economic well-being as it neglects differences in household size and transfer income. Table S.3 shows that dominance sets remain largely unchanged when we replace labor income with disposable income after taxes and transfers. However, dominance sets widen when we correct for differences in household size by using an equivalence scale. This result is driven by the fact that average outcomes improve in later cohorts once we account for the decrease in US household sizes over time (Lundberg *et al.*, 2016). Lastly, Mogstad *et al.* (2024) show that any ranking of societies based on estimated objects is subject to statistical uncertainty. Therefore, we augment our analysis to account for sampling variation in $V_{\theta,u}(f,q)$.[20] Table S.4 shows that confidence sets for statistical uncertainty are wider than dominance sets for normative ambiguity at $\theta = 0$ and $\theta = 2$. This result is driven by the relatively small sample sizes underlying each cohort-specific estimate in the PSID, and the relatively small differences in $\mathbb{E}_q[U]$ between cohorts. As $\theta \to \infty$, statistical confidence sets narrow because differences in $V_{\theta,u}(f,q)$ are now governed by the relatively large cross-cohort utility differences among the worst-off types.

8.2. Differences across regions (Commuting Zones)

Data. We draw on the data produced by Chetty *et al.* (2014a, 2026). These papers are based on the universe of federal income tax returns, linking parents and their children across generations. The sample is large and representative of the entire population, allowing the authors to calculate indicators of

[20] We provide a detailed methodological overview of our approach to estimating dominance sets, confidence sets, and combined sets in Appendix E.

intergenerational mobility at various levels of regional disaggregation. Many summary statistics from these analyses are publicly available and have become important inputs for analyses of intergenerational mobility in the US, among others, in Biasi (2023); Mogstad *et al.* (2024); Derenoncourt (2022); Gallagher *et al.* (2019).

The data of Chetty *et al.* (2014a) cover the birth cohorts 1980–1985.[21] Their data preparation protocol parallels our processing of the PSID in several important dimensions. First, Chetty *et al.* (2014a) also focus on family income. Concretely, they compute family income as the sum of the individual incomes of married partners from labor and capital, plus the non-taxable portion of Social Security and disability benefits. Second, parental income y^p is measured during children's adolescence by averaging across 1996–2000, i.e., when children are between 14 and 20 years old. Third, the child income y^c is the average of annual incomes in 2011 and 2012, i.e., when children are between 26 and 32 years old. The data cover 95% of the targeted population, and are therefore broadly representative of children in the relevant birth cohorts.

We do not have access to the individual-level data used by Chetty *et al.* (2014a). Therefore, we combine various summary statistics from Opportunity Insights to reconstruct for each CZ the joint distribution of children's family incomes (y^c) and parental income quintiles defined at the national level (r^p). We proceed in two steps. First, for each CZ, we construct a frequency-weight matrix $w^{CZ}(r^c, r^p)$ over children's and parents' national income positions. Concretely, in the publicly available data, parents are grouped into their five national income quintiles $r^p = 1, \ldots, 5$. For each parental income quintile, we know the CZ-specific share of children in that quintile $q^{CZ}(r^p)$, and the CZ-specific parent-child quintile transition $t^{CZ}(k(r^c) \mid r^p)$, i.e., the probability that a child of origin r^p reaches the child income quintile k containing income percentile r^c. Furthermore, we know the national within-quintile placement $h(r^c \mid k(r^c), r^p)$, i.e., the share of children of origin r^p who reach percentile r^c within k. Based on these inputs, we construct CZ-specific frequency-weight matrices as follows:

$$w^{CZ}(r^c, r^p) \;=\; q^{CZ}(r^p) \times t^{CZ}\big(k(r^c) \mid r^p\big) \times h(r^c \mid k(r^c), r^p), \tag{11}$$

where we increase the resolution within children's income quintiles by assuming that the placement probability $h(r^c \mid k(r^c), r^p)$ is common across CZ. Second, we combine $w^{CZ}(r^c, r^p)$ with the national percentile-income schedule $g(y^c \mid r^c)$ provided by Opportunity Insights, allowing us to back out the CZ-level joint distribution of y^c and r^p. We assess the accuracy of this procedure in two steps. First, we marginalize the resulting joint distributions over r^p and compare various moments of y^c with the

[21] Their main analysis sample covers the cohorts 1980–1982. However, various CZ-level statistics required for our analysis, such as quintile transition probabilities, are reported for the 1980–1985 cohort in the public data. This difference matters little in practice given the stability of various rank-mobility statistics in neighboring birth cohorts (Chetty *et al.*, 2014b).

corresponding moments computed by Chetty *et al.* (2014a) on individual-level microdata (Figure A.1). Second, we analogously compare the joint distributions of child ranks and parental ranks r^p (Figure A.2). Reassuringly, both tests show that the distributions align very closely.[22] In line with the PSID application, we restrict resulting data to observations with positive y^c, renormalize frequency weights within each CZ, and express monetary values in 2019 USD.

Table A.2 provides summary statistics for the 50 largest CZs, which we will focus on in our analysis. Population size varies substantially across CZs, from around $85,000$ children in Austin to more than $1,018,000$ in Los Angeles. While the sample is gender-balanced across all regions, CZs vary in their ethnic composition: The share of non-Hispanic Whites is highest in Manchester and lowest in Miami. Average child income ranges from around $\$45,000$ in Jacksonville and Orlando to approximately $\$65,000$ in Minneapolis.

Intergenerational mobility. In this application, we cannot rely on standard measures like the IGE or the rank-rank slope (RRS). On the one hand, we cannot compute the IGE as we do not have access to $\ln y^p$. On the other hand, we do not compute the RRS used by Chetty *et al.* (2014a, 2026), as our normative framework assesses child outcomes in monetary terms and not ranks. Taking account of these restrictions, we estimate intergenerational income mobility separately for each CZ as follows:

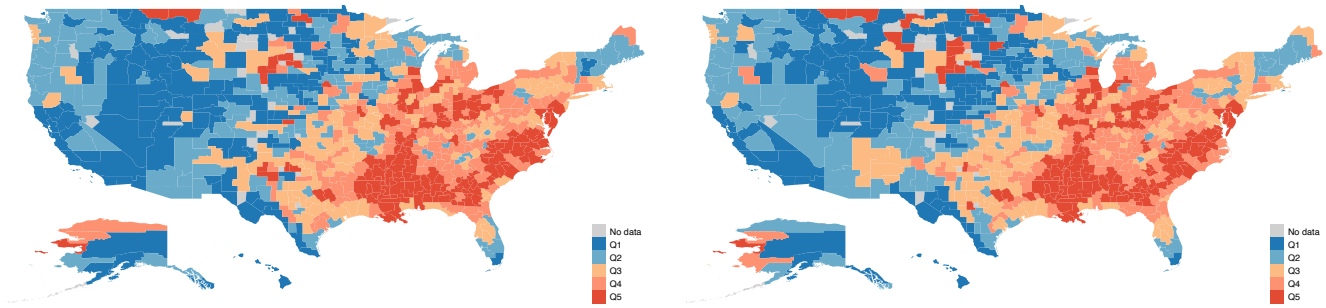
$$\ln y_i^c = \alpha + \beta^{IRS} r_i^p + \epsilon_i, \tag{12}$$

where $\ln y_i^c$ is the log of child income and r_i^p is the parental income quintile, represented by its midpoint rank, $r_i^p\ in\{10, 30, 50, 70, 90\}$. As we do not observe individual children, an observation i is a rank-quintile (r^c, r^p) pair from the reconstructed joint distribution, weighted by its frequency $w^{CZ}(r^c, r^p)$. The coefficient of interest, β^{IRS}, captures the average change in log child income associated with a one-rank increase in parental income. We call this measure income-rank slope (IRS) and interpret it as a summary measure of intergenerational mobility at the CZ level.

Figure 4 shows that the IRS replicates much of the regional variation embedded in standard mobility measures. In particular, we compare CZ-specific rank-rank slopes (RRS) estimated by Chetty *et al.* (2014a) in Panel (a) to our estimates of the IRS in Panel (b). Both estimators of intergenerational mobility yield similar spatial patterns: CZs in the South exhibit the lowest levels, while those in the West and Midwest exhibit the highest. The Spearman rank correlation coefficient between the RRS and the IRS is 0.95, indicating a strong concordance between the two measures. The correlation increases to 0.99

[22] The accuracy of our procedure decreases somewhat in the upper tail around the 99th income percentile of children's incomes—a result of the fact that the national-level placement function $h(r^c \mid k(r^c), r^p)$ cannot capture CZ-specific idiosyncracies in the tails of the distributions.

Figure 4: Intergenerational mobility across commuting zones



(a) Rank-rank slope (RRS) (b) Income-rank slope (IRS)

Data: Opportunity Insights.
Note: This figure shows intergenerational mobility in the US across commuting zones (CZs). Panel (a) shows the variation of rank-rank slopes (RRS) estimated by Chetty *et al.* (2014a). Panel (b) shows the variation of income-rank slopes (IRS) estimated based on Equation (12). CZs are grouped by quintiles of the RRS and IRS, respectively. Q1 (Q5) corresponds to the group of CZs with the highest (lowest) intergenerational mobility.

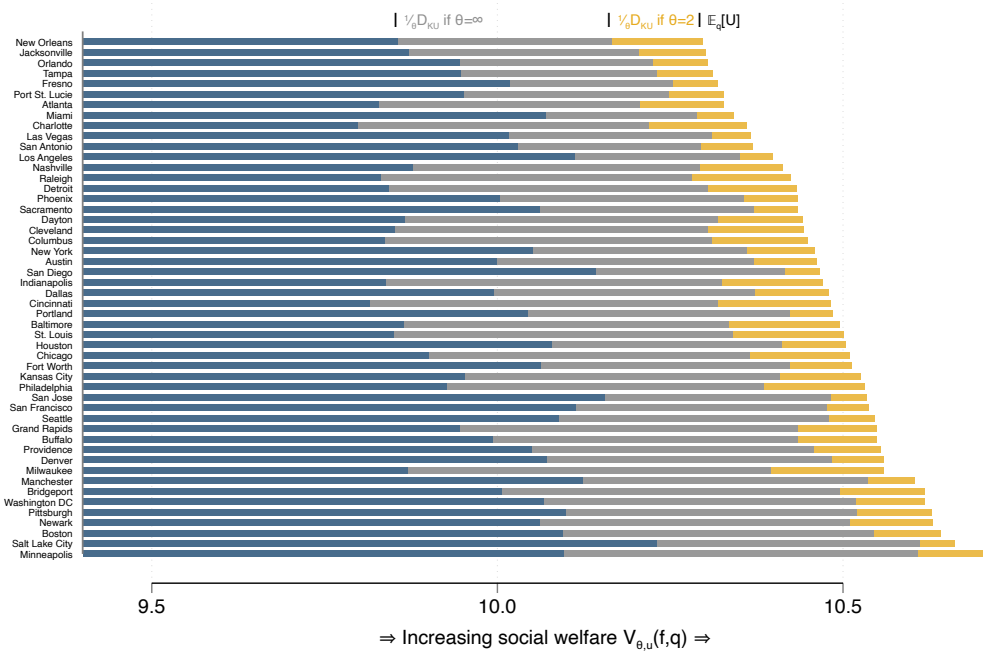
when considering only the 50 largest CZs, which are the focus of our analysis.

Opportunity-sensitive social welfare. Figure 5a illustrates our decomposition of social welfare $V_{\theta,u}(f,q)$ into efficiency $\mathbb{E}_q[U]$ and inequality of opportunity $1/\theta D_{K_U}$ based on the representation from Theorem 5 for the 50 largest CZs in the US. As before, the full length of the bars indicates $\mathbb{E}_q[U]$. The stacked bar segments show decreases in social welfare driven by $1/\theta D_{K_U}$ across different values of θ.

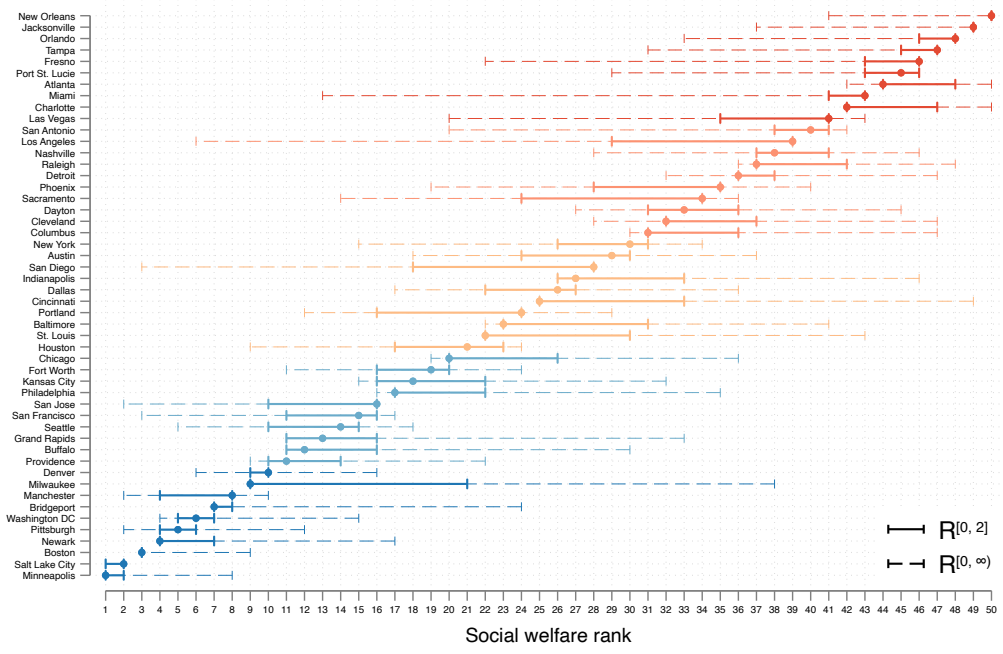
Compared with the cross-cohort comparison, regional differences in the efficiency component $\mathbb{E}_q[U]$ are much more pronounced: The cross-CZ standard deviation in $\mathbb{E}_q[U]$ is 0.104. Furthermore, the variation of $\mathbb{E}_q[U]$ exceeds the variation of inequality of opportunity for low levels of inequality-of-opportunity aversion: When setting $\theta = 2$, the cross-CZ standard deviation in $1/\theta D_{K_U}$ is 0.034, i.e., it is less than half of the corresponding variation in $\mathbb{E}_q[U]$. However, variation in inequality of opportunity grows with the aversion parameter θ: When $\theta \to \infty$, the cross-CZ standard deviation in $1/\theta D_{K_U}$ is 0.110, i.e., it is on par with the corresponding variation in $\mathbb{E}_q[U]$.

Figure 5a also illustrates strong heterogeneity in the importance of inequality of opportunity for social welfare comparisons. For example, Los Angeles ranks 39th in social welfare based on efficiency, $\mathbb{E}_q[U]$. However, compared with other CZs, social welfare declines less with increasing inequality-of-opportunity aversion, suggesting Los Angeles offers better opportunities for children from lower parental-income ranks than many other CZs. Conversely, Bridgeport ranks 7th when considering

Figure 5: Social welfare across commuting zones



(a) Decomposition



(b) Dominance sets for ranking

Data: Opportunity Insights.

Note: This figure shows variation in social welfare across the 50 largest commuting zones (CZs). Panel (a) shows the decomposition of social welfare $V_{\theta,u}(f,q)$ into efficiency $\mathbb{E}_q[U]$ and inequality of opportunity $^1/_\theta D_{K_U}$, based on the mean-divergence representation from Theorem 5, evaluated at $\theta = 2$ and $\theta \to \infty$. Panel (b) shows dominance sets for the social welfare ranking based on $\theta \in [0, 2]$ (solid lines) and $\theta \in [0, \infty)$ (dashed lines). Rank 1 (50) corresponds to the CZ with the highest (lowest) social welfare.

$\mathbb{E}_q[U]$. However, social welfare decreases more sharply as aversion to inequality of opportunity increases. At the highest level of θ, Bridgeport is outperformed by Los Angeles. Hence, the comparative evaluation of Los Angeles and Bridgeport crucially depends on our normative stance: How much do we care about the outcomes of children in the lowest tail of the parental income distribution, as opposed to other children who are better off but still worse off than the average?[23] In light of this normative ambiguity, we again calculate dominance sets to summarize the range of potential rankings for each CZ.

Welfare ranking. Figure 5b shows dominance sets for the 50 largest CZs in the US. Solid markers indicate the ranking of each CZ by efficiency $\mathbb{E}_q[U]$, and solid (dashed) lines indicate the dominance sets $R^{[0,2]}$ ($R^{[0,\infty)}$).

In contrast to the cross-cohort comparison, the results show markedly higher normative ambiguity. Considering $\theta \in [0, 2]$, only Boston, Jacksonville, and New Orleans can be ranked unambiguously at the 3rd, 49th, and 50th ranks, respectively. The average (median) dominance set comprises 5.6 (6) ranks across the 50 considered CZs. Thus, the amount of normative ambiguity is comparable to the amount of statistical uncertainty in mobility rankings documented by Mogstad *et al.* (2024), which ranges between 3 and 7 ranks for the 95% simultaneous confidence sets of the top and bottom five CZs, ranked in terms of absolute upward mobility at the 25th percentile of parental income. The extent of normative ambiguity increases considerably when considering the full range of θ. In this case, no CZ can be ranked unambiguously, and the average (median) dominance set increases to 17.6 (18) ranks across the 50 considered CZs.

Additional analyses. We again conduct additional analyses to probe the sensitivity of these results. First, Figure A.1 shows that our reconstruction of CZ-specific income distributions is less accurate in the tails. To assess the implications for dominance sets, we trim the child income distributions at the bottom and top. Table S.6 shows that trimming the child income distribution narrows dominance sets. This change is driven by changes in the inequality-of-opportunity component $1/\theta D_{K_U}$, where excluding the tails increases the homogeneity of income-rank slopes across CZs. Rankings by the efficiency component $\mathbb{E}_q[U]$ remain largely unaffected. Second, we again assess the importance of the linearity assumption when estimating expected type utilities $U(\pi_s)$ (Equation (12)). Table S.7 shows results are robust to various functional-form assumptions, similar to the application based on birth cohorts. Third,

[23] This intuition also relates to recent studies that evaluate societies using absolute mobility statistics, i.e., the average outcomes of children from a fixed position in the parental income distribution (among others, Boustan *et al.*, 2025; Britto *et al.*, forthcoming). Assuming expected child outcomes rise monotonically with parental rank, each level of θ can be translated into the average parental income rank considered by the social planner. For example, when evaluating societies based on the average outcomes of children at the lowest parental-income percentile, researchers implicitly assume $\theta \to \infty$. We illustrate this link formally in Appendix S.2.

Chetty *et al.* (2014a) observe children relatively early in their lifecycle at ages 26-32. To assess the sensitivity of our results to lifecycle bias, we repeat our analysis by replacing the national percentile-income schedule $g(y^c \mid r^c)$ of Chetty *et al.* (2014a) with the corresponding schedule of Chetty *et al.* (2026), in which children's incomes are observed at ages 31–37.[24] Table S.8 shows that dominance sets are largely unchanged when considering children at later ages. Finally, we also incorporate statistical uncertainty into our welfare comparisons. When considering normative ambiguity alone, dominance sets on average span 5.6 ranks for $\theta \in [0, 2]$ and 17.6 ranks for $\theta \in [0, \infty)$ (Table S.9). These sets increase to 9.6 ranks and 22.7 ranks when additionally accounting for statistical uncertainty via simultaneous confidence sets. This moderate increase is consistent with the results of Mogstad *et al.* (2024), who show that statistical uncertainty in ranks is small when focusing on mobility estimates in large samples, such as the 50 largest CZs in the US. As a consequence, normative ambiguity and statistical uncertainty contribute in comparable measure when inequality-of-opportunity aversion is restricted to $\theta \in [0, 2]$. However, when it is left unrestricted, normative ambiguity becomes the primary barrier to a clear-cut welfare ranking of these regions.

9. Conclusion

This paper develops a framework for evaluating income distributions when inequality of opportunity is a source of social concern. We characterize a tractable class of social welfare functions in which a single parameter governs aversion to inequality of opportunity. The resulting criterion admits a mean-divergence representation that separates average welfare from a structural penalty for disparities in opportunities, as well as a stochastic dominance criterion for welfare comparisons that are robust to the planner's degree of opportunity aversion.

Our theoretical framework lends itself to straightforward empirical implementation and accommodates standard measures of intergenerational mobility, which are the most prevalent proxies for unequal opportunities in the current economics literature. Using data from the United States, we show that robust rankings of opportunity-sensitive social welfare are the exception rather than the rule: Welfare rankings vary substantially with the assumed degree of inequality-of-opportunity aversion.

Importantly, by providing an explicit mapping from observable income distributions to opportunity-

[24] This procedure is valid if the CZ-level mapping between parental income ranks and children's income ranks is the same across both data sources. Chetty *et al.* (2014a) provide initial support for this assumption by showing that the RRS stabilizes when children are in their early 30s, suggesting that any lifecycle bias is likely confined to changes in marginal income distributions. Figure S.1 further supports this assumption by showing that CZ-level estimates of the RRS and absolute upward mobility at the 25th percentile of parental incomes are broadly concordant in both data sources.

sensitive social welfare, our framework makes normative judgments transparent and allows researchers to characterize conclusions that remain valid despite disagreement over the appropriate trade-off between efficiency and equality of opportunity.

References

Aczél, J. (1966). *Lectures on Functional Equations and Their Applications*. New York: Academic Press.

Aczél, J. (2005). Extension of a generalized Pexider equation. *Proceedings of the American Mathematical Society*, **133** (11), 3227–3234.

Adermon, A., Brandén, G. and Nybom, M. (forthcoming). The relationship between intergenerational mobility and equality of opportunity. *Scandinavian Journal of Economics*.

Adler, M., Bossert, W., Cato, S. and Kamaga, K. (2025). Ex post approaches to prioritarianism and sufficientarianism. *Theoretical Economics*, **20** (4), 1367–1410.

Alesina, A., Stantcheva, S. and Teso, E. (2018). Intergenerational mobility and preferences for redistribution. *American Economic Review*, **108** (2), 521–554.

Andreoli, F., Faure, M., Gravel, N. and Kundu, T. (2026). Evaluating allocations of opportunities. *International Economic Review*, **67** (1), 365–397.

Anscombe, F. J. and Aumann, R. J. (1963). A definition of subjective probability. *The Annals of Mathematical Statistics*, **34** (1), 199–205.

Arrow, K. (1965). *Aspects of the Theory of Risk-Bearing*. Helsinki: Yrjö Jahnssonin Säätiö.

Biasi, B. (2023). School finance equalization increases intergenerational mobility. *Journal of Labor Economics*, **41** (1), 1–38.

Björklund, A. and Jäntti, M. (2020). Intergenerational mobility, intergenerational effects, sibling correlations, and equality of opportunity: A comparison of four approaches. *Research in Social Stratification and Mobility*, **70**, 100455.

Blackorby, C. and Donaldson, D. (1982). Ratio-scale and translation-scale full interpersonal comparability without domain restrictions: Admissible social-evaluation functions. *International Economic Review*, **23** (2), 249–268.

Boustan, L., Jensen, M. F., Abramitzky, R., Jácome, E., Manning, A., Pérez, S., Watley, A., Adermon, A., Arellano-Bover, J., Åslund, O., Connolly, M., Deutscher, N., Gielen, A. C., Giesing, Y., Govind, Y., Halla, M., Hangartner, D., Jiang, Y., Karmel, C., Landaud, F., Macmillan, L., Martinez, I., Polo, A., Poutvaara, P., Rapoport, H., Roman, S., Salvanes, K. G., San, S., Siegenthaler, M., Sirugue, L., Soria Espin, J., Stuhler, J., Violante, G. L., Webbink, D., Weber, A., Zhang, J., Zheng, A. and Zohar, T. (2025). Intergenerational mobility of immigrants in 15 destination countries. *National Bureau of Economic Research Working Paper*, **33558**.

Brewer, M., Etheridge, B. and O'Dea, C. (2017). Why are Households that Report the Lowest Incomes So Well-off? *Economic Journal*, **127** (605), F24–F49.

Britto, D., Fonseca, A., Pinotti, P., Sampaio, B. and Warwar, L. (forthcoming). Intergenerational mobility in the land of inequality. *Review of Economics and Statistics*.

Brockett, P. and Golden, L. (1987). A class of utility functions containing all the common utility

functions. *Management Science*, **33** (8), 955–964.

Carneiro, P., López García, I., Salvanes, K. G. and Tominey, E. (2021). Intergenerational mobility and the timing of parental income. *Journal of Political Economy*, **129** (3), 757–788.

Carroll, G. (2025). Is equal opportunity different from welfarism?, Working Paper, Yale Department of Economics.

Cerreia-Vioglio, S., Maccheroni, F., Marinacci, M. and Montrucchio, L. (2011). Uncertainty averse preferences. *Journal of Economic Theory*, **146** (4), 1275–1330.

Chetty, R., Friedman, J. N., Hendren, N., Jones, M. R. and Porter, S. R. (2026). The Opportunity Atlas: Mapping the Childhood Roots of Social Mobility. *American Economic Review*, **116** (1), 1–51.

—, Hendren, N., Kline, P. and Saez, E. (2014a). Where is the land of opportunity? The geography of intergenerational mobility in the United States. *Quarterly Journal of Economics*, **129** (4), 1553–1623.

—, —, —, — and Turner, N. (2014b). Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility. *American Economic Review*, **104** (5), 141–47.

Cortes Orihuela, J., Díaz, J. D., Gutiérrez Cubillos, P. and Troncoso, P. A. (2025). The two-sample two-stage least squares method to estimate the intergenerational earnings elasticity. *Journal of Economic Inequality*, **23**, 549–570.

Davis, J. M. and Mazumder, B. (2026). The decline in intergenerational mobility after 1980. *Review of Economics and Statistics*, **108** (1), 1–15.

Debreu, G. (1960). Topological methods in cardinal utility theory. In K. J. Arrow, S. Karlin and P. Suppes (eds.), *Mathematical Methods in the Social Sciences, 1959*, Stanford: Stanford University Press, pp. 16–26.

Derenoncourt, E. (2022). Can you move to opportunity? Evidence from the great migration. *American Economic Review*, **112** (2), 369–408.

Deutscher, N. and Mazumder, B. (2023). Measuring intergenerational income mobility: A synthesis of approaches. *Journal of Economic Literature*, **61** (3), 988–1036.

Dupuis, P. and Ellis, R. (1997). *A Weak Convergence Approach to the Theory of Large Deviations*. Wiley Series in Probability and Statistics, New York: John Wiley & Sons.

Ergin, H. and Gul, F. (2009). A theory of subjective compound lotteries. *Journal of Economic Theory*, **144** (3), 899–929.

Eshaghnia, S. S., Heckman, J. J., Landersø, R. and Qureshi, R. (2024). Intergenerational transmission of family influence. *National Bureau of Economic Research Working Paper*, **33023**.

Fleurbaey, M. (2010). Assessing risky social situations. *Journal of Political Economy*, **118** (4), 649–680.

—, Moramarco, D. and Peragine, V. (2024). Measuring inequality and welfare when some inequalities matter more than others. *SERIES Working Paper*, **03/2024**.

— and Peragine, V. (2013). Ex ante versus ex post equality of opportunity. *Economica*, **80** (317), 118–130.

Gallagher, R., Kaestner, R. and Persky, J. (2019). The geography of family differences and intergen-

erational mobility. *Journal of Economic Geography*, **19** (3), 589–618.

Ghirardato, P. and Marinacci, M. (2002). Ambiguity made precise: A comparable foundation. *Journal of Economic Theory*, **102** (2), 251–289.

Gilboa, I. and Schmeidler, D. (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, **18** (2), 141–153.

Grant, S., Kajii, A., Polak, B. and Safra, Z. (2010). Generalized Utilitarianism and Harsanyi Impartial Observer Theorem. *Econometrica*, **78** (6), 1939–1971.

—, —, — and — (2012). Equally-distributed equivalent utility, ex post egalitarianism and utilitarianism. *Journal of Economic Theory*, **147** (4), 1545–1571.

—, Polak, B. and Strzalecki, T. (2009). Second-order expected utility, Working Paper, SSRN.

Hansen, L. and Sargent, T. (2001). Robust control and model uncertainty. *American Economic Review*, **91** (2), 60–66.

Hartley, R. P., Lamarche, C. and Ziliak, J. P. (2022). Welfare reform and the intergenerational transmission of dependence. *Journal of Political Economy*, **130** (3), 523–565.

Hufe, P., Kanbur, R. and Peichl, A. (2022). Measuring unfair inequality: Reconciling equality of opportunity and freedom from poverty. *Review of Economic Studies*, **89** (6), 3345–3380.

—, Kobus, M., Peichl, A. and Schüle, P. (2025). Multidimensional Equality of Opportunity in the United States. *European Economic Review*, **180**, 105162.

Jácome, E., Kuziemko, I. and Naidu, S. (2025). Mobility for all: Representative intergenerational mobility estimates over the twentieth century. *Journal of Political Economy*, **133** (1), 306–354.

Klibanoff, P., Marinacci, M. and Mukerji, S. (2005). A smooth model of decision making under ambiguity. *Econometrica*, **73** (6), 1849–1892.

Kreps, D. and Porteus, E. (1978). Temporal resolution of uncertainty and dynamic choice theory. *Econometrica*, **46** (1), 185–200.

Lundberg, S., Pollak, R. A. and Stearns, J. (2016). Family inequality: Diverging patterns in marriage, cohabitation, and childbearing. *Journal of Economic Perspectives*, **30** (2), 79–102.

Maccheroni, F., Marinacci, M. and Ruffino, D. (2013). Alpha as ambiguity: Robust mean-variance portfolio analysis. *Econometrica*, **81** (3), 1075–113.

—, — and Rustichini, A. (2006). Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica*, **74** (6), 1447–1498.

Magdalou, B. and Nock, R. (2011). Income distributions and decomposable divergence measures. *Journal of Economic Theory*, **146** (6), 2440–2454.

Mazumder, B. (2016). Estimating the intergenerational elasticity and rank association in the united states: Overcoming the current limitations of tax data. In *Inequality: Causes and Consequences, Research in Labor Economics*, vol. 43, Emerald Group Publishing, pp. 83–129.

— (2018). Intergenerational mobility in the United States: What we have learned from the PSID. *The*

ANNALS of the American Academy of Political and Social Science, **680** (1), 213–234.

Mello, U., Nybom, M. and Stuhler, J. (forthcoming). A life-cycle estimator of intergenerational income mobility. *Review of Economics and Statistics*.

Mogstad, M., Romano, J. P., Shaikh, A. M. and Wilhelm, D. (2024). Inference for ranks with applications to mobility across neighbourhoods and academic achievement across countries. *Review of Economic Studies*, **91** (1), 476–518.

Nock, R., Magdalou, B., Briys, E. and Nielsen, F. (2011). On tracking portfolios with certainty equivalents on a generalization of Markowitz model: The fool, the wise and the adaptive. In *Proceedings of the 28th International Conference on Machine Learning (ICML)*, Bellevue, WA: Omnipress, pp. 73–80.

Nybom, M. and Stuhler, J. (2024). Interpreting trends in intergenerational mobility. *Journal of Political Economy*, **132** (8), 2531–2570.

Piketty, T., Saez, E. and Zucman, G. (2018). Distributional national accounts: Methods and estimates for the United States. *Quarterly Journal of Economics*, **133** (2), 553–609.

Pratt, J. (1964). Risk aversion in the small and in the large. *Econometrica*, **32** (1/2), 122–136.

Ramos, X. and Van de gaer, D. (2016). Approaches to inequality of opportunity: Principles, measures and evidence. *Journal of Economic Surveys*, **30** (5), 855–883.

Roemer, J. and Trannoy, A. (2016). Equality of opportunity: Theory and measurement. *Journal of Economic Literature*, **54** (4), 1288–1332.

Savage, L. J. (1954). *The Foundations of Statistics*. New York: John Wiley & Sons.

Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica*, **57** (3), 571–587.

Shaked, M. and Shanthikumar, J. (2007). *Stochastic Orders*. Springer Series in Statistics, New York: Springer.

Solon, G. (1992). Intergenerational Income Mobility in the United States. *American Economic Review*, **82** (3), 393–408.

Strzalecki, T. (2011). Axiomatic foundations of multiplier preferences. *Econometrica*, **79** (1), 47–73.

van de Gaer, D. (1993). *Equality of Opportunity and Investment in Human Capital*. Ph.D. thesis, KU Leuven, Leuven.

von Neumann, J. and Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton: Princeton University Press.

Vosters, K. and Nybom, M. (2017). Intergenerational Persistence in Latent Socioeconomic Status: Evidence from Sweden and the United States. *Journal of Labor Economics*, **35** (3), 869–901.

Wakker, P. (1988). The algebraic versus the topological approach to additive representations. *Journal of Mathematical Psychology*, **32** (4), 421–435.

Ward, Z. (2023). Intergenerational mobility in american history: Accounting for race and measurement error. *American Economic Review*, **113** (2), 3213–3248.

Whitmore, G. (1989). Stochastic dominance for the class of completely monotonic utility functions. In T. B. Fomby and T. K. Seo (eds.), *Studies in the Economics of Uncertainty*, New York, NY: Springer, pp. 77–88.

A. Proofs

Proof of Theorem 1. Assume that the axioms stated in the theorem are satisfied. The set $\Delta(\mathcal{Y})$, endowed with the weak topology, forms a convex mixture space, which ensures it is a connected topological space. Consequently, the space of opportunity profiles $\mathcal{F}_{\mathcal{S}} = \Delta(\mathcal{Y})^{|\mathcal{S}|}$, endowed with the product topology, is also a connected space.

Then, we say that a type $s \in \mathcal{S}$ is essential for the relation $\succeq$ if there exists $f \in \mathcal{F}_{\mathcal{S}}$ and two conditional distributions $\pi_s, \pi'_s \in \Delta(\mathcal{Y})$ such that $(f_{-s}, \pi_s) \succ (f_{-s}, \pi'_s)$. By assumption, $\succeq$ is non-degenerate, meaning that $f \succ f'$ for some $f, f' \in \mathcal{F}_{\mathcal{S}}$. This implies there exist at least two income distributions $\pi, \pi' \in \Delta(\mathcal{Y})$ such that $f_\pi \succ f_{\pi'}$. Indeed, if $f_\pi \sim f_{\pi'}$ held for all $\pi, \pi' \in \Delta(\mathcal{Y})$, the weak part of A1 would force $f \sim f'$ for all $f, f' \in \mathcal{F}_{\mathcal{S}}$, contradicting non-degeneracy. Now consider these two distributions $\pi, \pi' \in \Delta(\mathcal{Y})$, a state $s \in \mathcal{S}$ and two profiles $f, f' \in \mathcal{F}_{\mathcal{S}}$ such that, respectively, $f'(t) = \pi'_t = \pi'$ for all $t \in \mathcal{S}$, and $f(t) = \pi_t = \pi'$ for all $t \neq s$ and $f(s) = \pi_s = \pi$. Because $f_\pi \succ f_{\pi'}$ and $\pi_t = \pi'_t$ for all $t \neq s$, A1 implies that $f = (f_{-s}, \pi_s) \succ (f_{-s}, \pi'_s) = f'$. Since this holds for any $s \in \mathcal{S}$, all types are essential. Because it is assumed that $|\mathcal{S}| \geq 3$, at least three types are essential. The relation $\succcurlyeq$ is a continuous weak order. By adding A2, the prerequisites of Theorem 4.1 in Wakker (1988) are satisfied, hence there exist a list of continuous functions $H_s : \Delta(\mathcal{Y}) \to \mathbb{R}$ for all $s \in \mathcal{S}$, defined up to an increasing affine transformation,[25] such that $\succcurlyeq$ can be represented by the function $W(f) = \sum_{s \in \mathcal{S}} H_s(\pi_s)$.

Now, we prove that $H_s(\pi_s) = q(s)H(\pi_s)$, under A3. Consider any $s, s' \in \mathcal{S}$ and any $f \in \mathcal{F}_{\mathcal{S}}$. Let f' be the profile obtained by permuting the conditional lotteries of types s and s', while leaving all other types unchanged. From A3, we have $f \sim f'$. Applying the additive representation, we deduce that $H_s(\pi_s) + H_{s'}(\pi_{s'}) = H_s(\pi_{s'}) + H_{s'}(\pi_s)$. Rearranging yields $H_s(\pi_s) - H_{s'}(\pi_s) = H_s(\pi_{s'}) - H_{s'}(\pi_{s'})$. Since this must hold for any arbitrary conditional distributions, the difference $H_s - H_{s'}$ must be a constant on $\Delta(\mathcal{Y})$. Because the uniqueness theorem for additive representations allows each function H_s to be shifted by an independent constant without altering the overall preference ranking, we can calibrate these type-specific constants to absorb this difference. This normalization establishes $H_s = H_{s'} = \tilde{H}$ for all $s, s' \in \mathcal{S}$, so that $W(f) = \sum_{s \in \mathcal{S}} \tilde{H}(\pi_s)$. By assumption, $q(s) = 1/|\mathcal{S}|$ for all $s \in \mathcal{S}$. Since W is unique up to a positive affine transformation, we can multiply the entire evaluation by the positive constant $1/|\mathcal{S}|$ without altering the represented preferences. By defining $H(\pi) = |\mathcal{S}| \cdot \tilde{H}(\pi)$, we deduce that $H_s(\pi_s) = q(s)H(\pi_s)$.

Then, we establish that $H(\pi_s) = \phi(U(\pi_s))$, with $U(\pi_s)$ the expected utility of π_s, and ϕ an increasing function. For two distributions $\pi, \pi' \in \Delta(\mathcal{Y})$, we define $\pi \succcurlyeq^\star \pi' \Leftrightarrow f_\pi \succcurlyeq f_{\pi'}$. By construction, the

[25] The uniqueness result for additive representations on connected spaces applies here (see Debreu, 1960, Theorem 3).

subspace of equal-opportunity profiles is isomorphic to the space of distributions $\Delta(\mathcal{Y})$, with respect to the probability mixture operation. Since the relation $\succcurlyeq$, when restricted to equal-opportunity profiles, satisfies the axioms of weak order, continuity, and independence (A5), it follows that the induced relation $\succcurlyeq^\star$ on $\Delta(\mathcal{Y})$ satisfies the von Neumann-Morgenstern (vNM) axioms. By the vNM expected utility theorem, there exists an affine function $U : \Delta(\mathcal{Y}) \to \mathbb{R}$ that represents $\succcurlyeq^\star$. Therefore, there exists a (continuous) utility function over incomes $u : \mathcal{Y} \to \mathbb{R}$ such that the evaluation of a distribution $\pi \in \Delta(\mathcal{Y})$ is given by $U(\pi) = \sum_{y\in\mathcal{Y}} \pi(y)u(y)$. Thus, $U(\pi) \geq U(\pi') \Leftrightarrow \pi \succcurlyeq^\star \pi' \Leftrightarrow f_\pi \succcurlyeq f_{\pi'}$. Moreover, we know from the previous representation that, for any $f \in \mathcal{F}_\mathcal{S}$, we have $W(f) = \sum_{s\in\mathcal{S}} q(s)H(\pi_s)$. We deduce that $W(f_\pi) = H(\pi)$. It follows that $U(\pi)$ and $H(\pi)$ represent $\succcurlyeq^\star$, hence there exists an increasing function ϕ such that $H(\pi) = \phi(U(\pi))$. From A1, we know that u is non-constant and ϕ is strictly increasing. For the sake of consistency with the rest of the paper, we prefer the ordinally equivalent representation in (1), rather than $W(f) = \sum_{s\in\mathcal{S}} q(s)\phi(U(\pi_s))$.

The next step is to show that ϕ is weakly concave. Consider profiles $f, f' \in \mathcal{F}_\mathcal{S}$ as defined in A4. By assuming A4 we have, after simplification (by multiplying by $|\mathcal{S}|$ and dividing by 2):

$$f \succcurlyeq f' \iff (1/2)\,\phi(U(\pi_s)) + (1/2)\,\phi(U(\pi_{s'})) \geq (1/2)\,\phi(U(\pi'_s)) + (1/2)\,\phi(U(\pi'_{s'}))\ .$$

By definition, $\pi_s = \pi_{s'} = (1/2)(\pi'_s + \pi'_{s'})$. Because U is affine, $U(\pi_s) = U(\pi_{s'}) = (1/2)\,U(\pi'_s) + (1/2)\,U(\pi'_{s'})$. After simplification, one obtains:

$$f \succcurlyeq f' \iff \phi((1/2)(U(\pi'_s) + U(\pi'_{s'}))) \geq (1/2)\,\phi(U(\pi'_s)) + (1/2)\,\phi(U(\pi'_{s'}))\ .$$

Since ϕ is continuous, this midpoint concavity implies global weak concavity of ϕ.

Finally, we deal with the uniqueness of the functions u and ϕ. First, u is obtained in a vNM framework, which implies that u and u' represent the same preferences if there exist $a > 0$ and $b \in \mathbb{R}$ such that $u' = au + b$. We have previously established that H, hence also ϕ, is defined up to an increasing affine transformation. Hence ϕ and ϕ' represent the same preferences if there exist $A > 0$ and $B \in \mathbb{R}$ such that $\phi'(u') = A\phi(u) + B$. The proof of the consistency of $W_{\phi,u,q}$ with the considered axioms is left to the reader. □

Proof of Theorem 2. Assume that the axioms stated in the theorem are satisfied. Because $\succcurlyeq$ is a continuous weak order, A5' is stronger than A5, thus A5 also hold. We know from Theorem 1 that $f \succeq f'$ implies $\phi^{-1}\left(\sum_{s\in\mathcal{S}} q(s)\phi(U(\pi_s))\right) \geq \phi^{-1}\left(\sum_{s\in\mathcal{S}} q(s)\phi(U(\pi'_s))\right)$, with ϕ a strictly increasing function. Let $\mathcal{U} = U(\Delta(\mathcal{Y})) \subseteq \mathbb{R}$ be the range of the expected utility function, which is an interval. Let $u = (u_s)_{s\in\mathcal{S}}$ and $u' = (u'_s)_{s\in\mathcal{S}}$ with, respectively, $u_s = U(\pi_s)$ and $u'_s = U(\pi'_s)$, for all $s \in \mathcal{S}$. We first

define a relation $\succeq^*$ over $\mathcal{U}^{|\mathcal{S}|}$ induced by $\succeq$ such that, for all $f, f' \in \mathcal{F}_\mathcal{S}$, we have $u \succeq^* u' \iff f \succeq f'$. By construction, $\succeq^*$ is represented by the continuous function $H_{\phi,u,q}(u) = \phi^{-1}(\sum_{s\in\mathcal{S}} q(s)\phi(u_s))$ on $\mathcal{U}^{|\mathcal{S}|}$. Then, define $c = U(\pi)$ and $c' = U(\pi')$. Because U is affine, A5' implies that, for all $f, f' \in \mathcal{F}_\mathcal{S}$, all $\pi, \pi' \in \Delta(\mathcal{Y})$ and all $\alpha \in (0,1)$:

$$\begin{aligned} & \phi^{-1}\left(\sum_{s\in\mathcal{S}} q(s)\phi\left(\alpha u_s + (1-\alpha)c\right)\right) \geq \phi^{-1}\left(\sum_{s\in\mathcal{S}} q(s)\phi\left(\alpha u'_s + (1-\alpha)c\right)\right) \\ \iff \quad & \phi^{-1}\left(\sum_{s\in\mathcal{S}} q(s)\phi\left(\alpha u_s + (1-\alpha)c'\right)\right) \geq \phi^{-1}\left(\sum_{s\in\mathcal{S}} q(s)\phi\left(\alpha u'_s + (1-\alpha)c'\right)\right). \end{aligned}$$

By letting $1_{|\mathcal{S}|}$ be a vector of one repeated $|\mathcal{S}|$ times, it is equivalent to write:

$$\alpha u + (1-\alpha)c \cdot 1_{|\mathcal{S}|} \succcurlyeq^\star \alpha u' + (1-\alpha)c \cdot 1_{|\mathcal{S}|} \iff \alpha u + (1-\alpha)c' \cdot 1_{|\mathcal{S}|} \succcurlyeq^\star \alpha u' + (1-\alpha)c' \cdot 1_{|\mathcal{S}|}.$$

Let $v = \alpha u + (1-\alpha)c \cdot 1_{|\mathcal{S}|}$, $v' = \alpha u' + (1-\alpha)c \cdot 1_{|\mathcal{S}|}$, and $k = (1-\alpha)(c'-c)$. After simplification:

$$v \succcurlyeq^\star v' \iff v + k \cdot 1_{|\mathcal{S}|} \succcurlyeq^\star v' + k \cdot 1_{|\mathcal{S}|}.$$

By continuity, this translation invariance extends globally to all $u, u' \in \mathcal{U}^{|\mathcal{S}|}$ and any admissible translation k. From Theorem 3, page 257 in Blackorby and Donaldson (1982), we know that $\succcurlyeq^\star$ can be represented by the function $H_{\phi_\theta,u,q}(u) = \phi_\theta^{-1}\left(\sum_{s\in\mathcal{S}} q(s)\phi_\theta(u_s)\right)$, with ϕ_θ as defined in (2). Note that, in Blackorby and Donaldson (1982), $\succcurlyeq^\star$ compares elements of $\mathbb{R}^{|\mathcal{S}|}$, and not elements of $\mathcal{U}^{|\mathcal{S}|} \subset \mathbb{R}^{|\mathcal{S}|}$ as we do. Drawing on the results established in Aczél (1966), Section 2.1.4, this representation remains valid on our restricted domain, given that $\mathcal{U}$ is an interval. Then, the restriction on the domain of θ results from the concavity of ϕ_θ (and A4). We know that the function u can be replaced by $u' = au + b$, with $a > 0$ and $b \in \mathbb{R}$. If $H_{\phi_\theta,u,q}$ and $H_{\phi_{\theta'},u',q}$ represent the same preferences, we have $H_{\phi_{\theta'},u',q} = \phi_{\theta'}^{-1}\left(\sum_{s\in\mathcal{S}} q(s)\left(-e^{-\theta'(au_s+b)}\right)\right) = \phi_{\theta'}^{-1}\left(e^{-\theta' b}\sum_{s\in\mathcal{S}} q(s)\left(-e^{-\theta' au_s}\right)\right)$. Therefore, we must also have $\theta' = \theta/a$. To sum up, $\succcurlyeq$ can be represented by the function $W_{\phi_\theta,u,q}(f) = \phi_\theta^{-1}\left(\sum_{s\in\mathcal{S}} q(s)\phi_\theta\left(U(\pi_s)\right)\right)$. All the axioms considered here are satisfied by this representation. □

Proof of Theorem 3. Assume that the axioms stated in the theorem are satisfied. Because the set of potential individuals $\mathcal{I}$ is countably infinite and populations $\mathcal{N}$ are finite, the set of all possible partitions $\mathcal{T}$ is countable. The extended domain $\mathcal{D}$ is thus a countable union of separable spaces. Relation $\succeq^*$ is assumed to be a continuous weak order, hence representable by a continuous function $H : \mathcal{D} \to \mathbb{R}$, so that $(f,q) \succeq^* (f',q') \iff H(f,q) \geq H(f',q')$. In Theorem 1, we established that the local preference $\succcurlyeq_\mathcal{S}$ on a fixed partition $\mathcal{S}$ of types with $|\mathcal{S}| \geq 3$, is represented by $W_{|\mathcal{S}|}(f) = \phi_{|\mathcal{S}|}^{-1}\left(\sum_{s\in\mathcal{S}} \overline{q}(s)\phi_{|\mathcal{S}|}(U_{|\mathcal{S}|}(\pi_s))\right)$, with $\overline{q}(s) = 1/|\mathcal{S}|$ for all $s \in \mathcal{S}$. Note that this representation depends only on the partition size $|\mathcal{S}|$

and not on the specific partition identity. Indeed, within any given partition, A3 ensures all types are treated symmetrically. Furthermore, across different partitions, A8 allows any two uniform societies of the same size and with identical conditional income distributions to be transformed into one another via type-splitting and merging, guaranteeing their ordinal equivalence in $\mathcal{D}$.

To establish that the cardinal representation (u_N, ϕ_N) is independent of partition size and population scale, consider two uniformly distributed societies (f, q) and (f', q') defined on partitions of sizes N and N', respectively. Because the distributions are uniform, each type in the first society contains an identical number $c \in \mathbb{N}$ of individuals, and each type in the second society contains $c' \in \mathbb{N}$ individuals. Because the set of potential individuals $\mathcal{I}$ is countably infinite, we can use A7 to replicate the first society by a factor of $c'N'$, and the second society by a factor of cN. Both replicated societies now share the exact same total population size, $cc'NN'$. Then, we apply A8 to split each of the N types of the first replicated society into N' identical subtypes, and each of the N' types of the second replicated society into N identical subtypes. Both transformed societies now consist of exactly NN' uniform types, with each type containing exactly cc' individuals. Because these two resulting societies share the exact same demographic structure, a type-preserving bijection can be drawn between their partitions. By A7 (applied as a 1-fold replication), evaluating a profile on one partition is strictly equivalent to evaluating it on the other. These profiles can be treated as operating on a common domain $\mathcal{F}_{\mathcal{S}^\star}$ with $|\mathcal{S}^\star| = NN'$. By the uniqueness result of Theorem 1 on this common domain, (u_N, ϕ_N) and $(u_{N'}, \phi_{N'})$ must both be positive affine transformations of this refined representation $(u_{NN'}, \phi_{NN'})$, yielding a unique pair (u, ϕ). Thus, a local preference $\succcurlyeq_{\mathcal{S}}$ on a fixed partition $\mathcal{S}$ is represented by $W(f) = \phi^{-1}\left(\sum_{s\in\mathcal{S}} \overline{q}(s)\phi(U(\pi_s))\right)$. A6 implies, for all $f, f' \in \mathcal{F}_{\mathcal{S}}$, that $(f, \overline{q}) \succcurlyeq^\star (f', \overline{q}) \iff f \succcurlyeq f'$. We can therefore calibrate H globally such that, for any uniformly distributed society $(f, \overline{q}) \in \mathcal{D}$ of arbitrary size, we have:

$$H(f, \overline{q}) = \phi^{-1}\left(\sum_{s\in\mathcal{S}} \overline{q}(s)\phi\left(U(\pi_s)\right)\right). \tag{13}$$

Consider now an arbitrary society $(f, q) \in \mathcal{D}$ defined on a partition $\mathcal{S}$, where the distribution $q \in \Delta(\mathcal{S})$ consists of strictly positive rational probabilities. Let $K \in \mathbb{N}$ be a common denominator for all $q(s)$ such that for each $s \in \mathcal{S}$, we can write $q(s) = k_s/K$ with $k_s \in \mathbb{N}$. To evaluate this society, we first apply A7 to replicate the entire population by a factor K, ensuring that each type's absolute size becomes a multiple of K. Then, by iterative applications of A8, we can partition each type $s \in \mathcal{S}$ into k_s distinct subtypes. Each new subtype is assigned a demographic weight of $1/K$ and receives the exact same conditional income distribution π_s as the original type s. This procedure maps the original society (f, q) to a new society $(f', \overline{q}') \in \mathcal{D}$ defined on a finer partition $\mathcal{S}'$ of size $|\mathcal{S}'| = K$. The distribution $\overline{q}'$ is uniform over $\mathcal{S}'$, so that $\overline{q}'(s) = 1/K$ for all $s \in \mathcal{S}'$. A8 guarantees that this procedure leaves the

extended evaluation unchanged. Hence $(f, q) \sim^\star (f', \overline{q}')$, which implies $H(f, q) = H(f', \overline{q}')$. Since $\overline{q}'$ is a uniform society, we evaluate it using the representation in (13):

$$H(f', \overline{q}') = \phi^{-1}\left(\sum_{s \in \mathcal{S}'} \frac{1}{K}\phi\left(U(\pi_s)\right)\right). \tag{14}$$

Grouping the terms of this sum back according to the original types $s \in \mathcal{S}$ (since each original type s was split into k_s identical sub-types), the sum simplifies to:

$$H(f, q) = \phi^{-1}\left(\sum_{s \in \mathcal{S}} \frac{k_s}{K}\phi\left(U(\pi_s)\right)\right) = \phi^{-1}\left(\sum_{s \in \mathcal{S}} q(s)\phi\left(U(\pi_s)\right)\right). \tag{15}$$

Finally, the set of strictly positive rational distributions is dense in the probability simplex $\Delta(\mathcal{S})$. Since the extended preference relation $\succcurlyeq^\star$ is a continuous weak order, the representation $H(f, q)$ must be continuous with respect to q. Therefore, the functional form derived in (15) extends to all real-valued $q \in \Delta(\mathcal{S})$, including boundary points where terms with $q(s) = 0$ trivially vanish. In conclusion, we immediately see that $W_{\phi,u}(f, q)$ satisfies A6–A8. □

Proof of Theorem 4. When $\theta = 0$, the expression of $V_{\theta,u}(f, q)$ in (5) is immediately deduced from (2). Now assume that $\theta > 0$. Let $H_{\theta,u}(f, q, p) = \sum_{s \in \mathcal{S}} p(s)U(\pi_s) + \frac{1}{\theta}D_{KL}(p||q)$, with $p \in \Delta^q(\mathcal{S})$. We first prove that $p^\star_{\theta,u}(f, q; \cdot) = \arg\min_{p \in \Delta^q(\mathcal{S})} H_{\theta,u}(f, q, p)$ can be written as in (6). We then establish that $H_{\theta,u}(f, q, p^\star_{\theta,u}(f, q; \cdot)) = V_{\theta,u}(f, q)$.

Because $p \in \Delta^q(\mathcal{S})$, we have $p^\star_{\theta,u}(f, q; s) = 0$ as soon as $q(s) = 0$. Now, denote by $\overline{\mathcal{S}} \subseteq \mathcal{S}$ the set of s such that $q(s) > 0$. Let us assume that the minimum is an interior solution on $\overline{\mathcal{S}}$, meaning $p(s) > 0$ for all $s \in \overline{\mathcal{S}}$. The Lagrangian of this minimization problem, with $\nu = (\nu_s)_{s \in \overline{\mathcal{S}}}$, can be written as:

$$L(p, \lambda, \nu) = \sum_{s \in \overline{\mathcal{S}}} p(s)\left[U(\pi_s) + \frac{1}{\theta}\ln\frac{p(s)}{q(s)}\right] + \lambda\left[\sum_{s \in \overline{\mathcal{S}}} p(s) - 1\right] - \sum_{s \in \overline{\mathcal{S}}} \nu_s p(s).$$

For all $s \in \overline{\mathcal{S}}$, the first Karush-Kuhn-Tucker (KKT) optimality condition needs $\nu_s p(s) = 0$, hence $\nu_s = 0$. Then, the stationarity condition $\frac{\partial L(p,\lambda,\nu)}{\partial p(s)} = 0$ leads to:

$$p^\star_{\theta,u}(f, q; s) = q(s)\exp\left(\theta\left(-U(\pi_s) - \lambda\right) - 1\right). \tag{16}$$

Because $q(s) > 0$, the right hand-side is positive, hence $p^\star_{\theta,u}(f, q; s) > 0$ for all $s \in \overline{\mathcal{S}}$. Moreover,

$\sum_{t\in\overline{\mathcal{S}}} p^\star_{\theta,u}(f,q;t) = 1$, which is equivalent, by using (16), to:

$$\exp\left(-\theta\lambda - 1\right) = \frac{1}{\sum_{t\in\overline{\mathcal{S}}} q(t)\exp\left(-\theta U(\pi_t)\right)}\,. \tag{17}$$

One also knows from (16) that $p^\star_{\theta,u}(f,q;s) = q(s)\exp\left(-\theta U(\pi_s)\right)\exp\left(-\theta\lambda - 1\right)$. Hence, $p^\star_{\theta,u}(f,q;s)$ as defined in (6) is obtained by substituting (17) into (16). The other KKT conditions are satisfied. We verify that $p^\star_{\theta,u}(f,q;s) > 0$, which is consistent with our initial interiority assumption. Finally, because the objective function is strictly convex on the probability simplex, the global minimum is unique.

Let's now prove that $H_{\theta,u}(f,q,p^\star_{\theta,u}(f,q;\cdot)) = V_{\theta,u}(f,q)$. By using $p^\star_{\theta,u}(f,q;s)$, we have:

$$\begin{aligned} H_{\theta,u}(f,q,p^\star_{\theta,u}(f,q;\cdot)) &= \sum_{s\in\mathcal{S}} p^\star_{\theta,u}(f,q;s)\left[U(\pi_s) + \frac{1}{\theta}\ln\left(\frac{\exp\left(-\theta U(\pi_s)\right)}{\sum_{t\in\mathcal{S}} q(t)\exp\left(-\theta U(\pi_t)\right)}\right)\right] \\ &= -\frac{1}{\theta}\ln\left(-\sum_{t\in\mathcal{S}} q(t)\phi_\theta\left(U(\pi_t)\right)\right) \times \sum_{s\in\mathcal{S}} p^\star_{\theta,u}(f,q;s)\,. \end{aligned}$$

The result is deduced from $\sum_{s\in\mathcal{S}} p^\star_{\theta,u}(f,q;s) = 1$, and $\phi_\theta^{-1}(t) = -(1/\theta)\ln(-t)$ when $\theta > 0$. □

Proof of Theorem 5. When $\theta = 0$, the expression of $V_{\theta,u}(f,q) = \mathbb{E}_q[U]$ is immediately deduced from (2). For $\theta > 0$, we first need a definition. The 'cumulant-generating function' of a real-valued random variable X can be written, for any $\tau \in \mathbb{R}$, by $K_X(\tau) = \ln\left(\mathbb{E}[e^{\tau X}]\right)$. By definition, such a function is convex. If we denote by $K_X^{(n)}$ the n-th derivative of K_X, a Taylor expansion of order m around zero yields, as τ tends to zero, $K_X(\tau) = \sum_{n=1}^m K_X^{(n)}(0)\frac{\tau^n}{n!} + \mathcal{O}(\tau^{m+1})$. The cumulants are sufficient statistics to characterize the full distribution of the random variable X, with the first two cumulants $K_X^{(1)}(0) = \mathbb{E}[X]$ and $K_X^{(2)}(0) = \mathrm{var}(X)$.

If we let $U = (U(\pi_s))_{s\in\mathcal{S}}$ be the list of expected utilities of all types in $f \in \mathcal{F}_\mathcal{S}$, distributed according to $q = (q(s))_{s\in\mathcal{S}}$, we have $V_{\theta,u}(f,q) = (-1/\theta)\ln\left(\mathbb{E}_q\left[e^{-\theta U}\right]\right)$ (see Theorem 2). Therefore, we can write $V_{\theta,u}(f,q) = (-1/\theta)K_U(-\theta)$. By limiting the expansion of K_U to order 2, one obtains the local approximation, as θ tends to zero:

$$V_{\theta,u}(f,q) = \mathbb{E}_q[U] - (\theta/2)\,\mathrm{var}_q(U) + \mathcal{O}(\theta^2)\,.$$

Following the standard Arrow-Pratt logic, the mean-variance approximation is accurate either when the planner's aversion θ is close to zero, or when the disparities in type-specific expected utilities are small (so that higher-order cumulants are negligible relative to the variance). Then, from Definition 2, $D_{K_U}(-\theta||0) = K_U(-\theta) - K_U(0) + \theta K_U^{(1)}(0)$. Moreover, by definition, $K_U(0) = 0$. Because $V_{\theta,u}(f,q) =$

$(-1/\theta)K_U(-\theta)$, we also have $V_{\theta,u}(f,q) = \mathbb{E}_q[U] - (1/\theta)\, D_{K_U}(-\theta||0)$. □

Proof of Theorem 6. We know that $V_{\theta,u}(f,q) = \phi_\theta^{-1}\left(\mathbb{E}_q\left[\phi_\theta(U)\right]\right)$, with $U = (U(\pi_s))_{s\in\mathcal{S}}$. Because ϕ_θ^{-1} is strictly increasing, $V_{\theta,u}(f,q) \geq V_{\theta,u}(f',q')$ if and only if $\mathbb{E}_q[\phi_\theta(U)] \geq \mathbb{E}_{q'}[\phi_\theta(U')]$. Moreover, because u can be assumed strictly positive, $U(\pi_s) > 0$ for all $\pi_s \in \Delta(\mathcal{Y})$. Thus, U and U' are strictly positive random variables. It follows that (i) $\Leftrightarrow$ (ii) can be deduced from Theorem 5.A.4 in Shaked and Shanthikumar (2007) and Theorem 1 in Brockett and Golden (1987) (see also Whitmore, 1989). (i) $\Rightarrow$ (ii) is immediate as ϕ_θ is completely alternating for all $\theta \geq 0$.

Let's prove (ii) $\Rightarrow$ (i). If φ is completely alternating, then $\varphi^{(1)}$ is completely monotone and, from Hausdorff-Bernstein-Widder theorem with $t > 0$, we know that $\varphi^{(1)}(t) = \int_{[0,\infty)} \exp(-rt)m(\mathrm{d}r)$ for some positive measure m on $[0,\infty)$. Isolating the case $r = 0$, this expression can be written as $\varphi^{(1)}(t) = c + \int_{(0,\infty)} \exp(-rt)m(\mathrm{d}r)$ for some constant $c \geq 0$ and measure m on $(0,\infty)$. Assume that $\varphi(0^+) = 0$.[26] By integration over $(0,z]$ and reversing the order of integration, we deduce that $\varphi(z) = cz + \int_{(0,\infty)} \left[(-1/r)\exp(-rt)\right]_0^z m(\mathrm{d}r) = cz + \int_{(0,\infty)} (1/r)\left[1 - \exp(-rz)\right] m(\mathrm{d}r)$. Statement (ii) is equivalent to $\mathbb{E}_q[-\exp(-\theta U)] \geq \mathbb{E}_{q'}[-\exp(-\theta U')]$ for all $\theta \in (0,\infty)$ and $\mathbb{E}_q[U] \geq \mathbb{E}_{q'}[U']$ if $\theta = 0$. Assume first that $\theta = 0$. If (ii) is true, it follows that:

$$\mathbb{E}_q[cU] \geq \mathbb{E}_{q'}[cU'], \tag{18}$$

for any $c \geq 0$, which establishes statement (i) for any purely linear function $\varphi(z) = cz$. Then, for all $\theta > 0$, statement (ii) implies that:

$$\mathbb{E}_q\left[(1/\theta)\left[1 - \exp(-\theta U)\right]\right] \geq \mathbb{E}_{q'}\left[(1/\theta)\left[1 - \exp(-\theta U')\right]\right]. \tag{19}$$

[26] This normalization entails no loss of generality. First, if $\varphi(0^+)$ is finite and non-zero, the reasoning applies to the shifted function $\varphi^\star(t) = \varphi(t) - \varphi(0^+)$, which preserves the dominance between distributions (see Brockett and Golden, 1987). Second, if $\varphi(0^+) = -\infty$, because u is continuous and assumed to be strictly positive on the compact domain $\mathcal{Y}$, all expected utilities U and U' are bounded below by $t_0 = \min_{y\in\mathcal{Y}} u(y) > 0$. Integrating $\varphi^{(1)}$ over $[t_0, z]$ instead of $(0,z]$ ensures that all terms and integrals remain finite and yields the same conclusion.

For any completely alternating function φ, one deduces that:

$$\begin{aligned}
\mathbb{E}_q[\varphi(U)] &\overset{\text{def}}{=} \mathbb{E}_q\left[cU + \int_{(0,\infty)} (1/r)\left[1-\exp\left(-rU\right)\right]\, m(\mathrm{d}r)\right] \\
&= c\,\mathbb{E}_q[U] + \int_{(0,\infty)} \mathbb{E}_q\left[(1/r)\left[1-\exp\left(-rU\right)\right]\right] m(\mathrm{d}r) \\
&\overset{(18 \text{ and } 19)}{\geq} c\,\mathbb{E}_{q'}[U'] + \int_{(0,\infty)} \mathbb{E}_{q'}\left[(1/r)\left[1-\exp\left(-rU'\right)\right]\right] m(\mathrm{d}r) \\
&= \mathbb{E}_{q'}\left[cU' + \int_{(0,\infty)} (1/r)\left[1-\exp\left(-rU'\right)\right]\, m(\mathrm{d}r)\right] \overset{\text{def}}{=} \mathbb{E}_{q'}[\varphi(U')]\,, \qquad (20)
\end{aligned}$$

Therefore, (ii) implies that statement (i) is true. □

Proof of Property 1. Let $y \in \mathcal{Y}$. By assumption $u(y) \geq 0$, thus $U(\pi_s) \geq 0$ for all $s \in \mathcal{S}$. If $\theta = 0$, then $V_{0,u}(f,q) = \sum_{s\in\mathcal{S}} q(s)U(\pi_s) \geq 0$. Now assume that $\theta > 0$. We have $\exp(-\theta U(\pi_s)) \leq \exp(0) = 1$. It follows that $\sum_{s\in\mathcal{S}} q(s)\exp(-\theta U(\pi_s)) \in (0,1]$, and that $\ln\left(\sum_{s\in\mathcal{S}} q(s)\exp(-\theta U(\pi_s))\right) \in (-\infty, 0]$. Finally, $V_{\theta,u}(f,q)$ is equal to this last term, times $-(1/\theta)$, thus $V_{\theta,u}(f,q) \geq 0$. □

Proof of Property 2. If $\theta = 0$, we have $V_{0,u}(f,q) = \sum_{s\in\mathcal{S}} q(s)U(\pi_s)$, which yields $\frac{\partial V_{0,u}}{\partial U(\pi_s)}(f,q) = q(s)$. Since $p^\star_{0,u}(f,q;s) = q(s)$, the property trivially holds. Assume now that $\theta > 0$. Let $\tilde{V}_{\theta,u}(f,q) = \sum_{t\in\mathcal{S}} q(t)\exp(-\theta U(\pi_t))$, so that $V_{\theta,u}(f,q) = -(1/\theta)\ln\tilde{V}_{\theta,u}(f,q)$. One observes that $\partial\tilde{V}_{\theta,u}/\partial U(\pi_s) = -\theta q(s)\exp(-\theta U(\pi_s))$. Thus:

$$\frac{\partial V_{\theta,u}}{\partial U(\pi_s)}(f,q) = -\frac{1}{\theta}\frac{-\theta q(s)\exp(-\theta U(\pi_s))}{\tilde{V}_{\theta,u}(f,q)} = p^\star_{\theta,u}(f,q;s)\,.$$

□

Proof of Property 3. We have $U(\pi_s) = \sum_{y\in\mathcal{Y}} \pi_s(y)u(y)$. If $\theta = 0$, then $V_{0,u}(f,q) = \sum_{y\in\mathcal{Y}} \pi(y)u(y)$, which yields $\frac{\partial V_{0,u}}{\partial u(y)}(f,q) = \pi(y)$. Since $p^\star_{0,u}(f,q;s) = q(s)$, $\pi^\star_{0,u}(f,q;y) = \sum_{s\in\mathcal{S}} q(s)\pi_s(y) = \pi(y)$, meaning that the property holds. Assume now that $\theta > 0$. Let $\tilde{V}_{\theta,u}(f,q) = \sum_{s\in\mathcal{S}} q(s)\exp(-\theta U(\pi_s))$, such that $V_{\theta,u}(f,q) = -(1/\theta)\ln\tilde{V}_{\theta,u}(f,q)$. We have $\frac{\partial\tilde{V}_{\theta,u}}{\partial u(y)} = -\theta\sum_{s\in\mathcal{S}} \pi_s(y)q(s)\exp\left(-\theta U(\pi_s)\right)$, recalling that $U(\pi_s) = \sum_{y\in\mathcal{Y}} \pi_s(y)u(y)$. It follows that:

$$\frac{\partial V_{\theta,u}}{\partial u(y)}(f,q) = -\frac{1}{\theta}\frac{\partial\tilde{V}_{\theta,u}/\partial u(y)}{\tilde{V}_{\theta,u}(f,q)} = \sum_{s\in\mathcal{S}}\left[\frac{q(s)\exp\left(-\theta U(\pi_s)\right)}{\tilde{V}_{\theta,u}(f,q)}\right]\pi_s(y) = \sum_{s\in\mathcal{S}} p^\star_{\theta,u}(f,q;s)\pi_s(y)\,.$$

□

Proof of Property 4. The case $V_{0,u}(f,q)$ is stated in equation (5). With Theorem 4, it suffices to calculate

$\lim_{\theta\to\infty} p^\star_{\theta,u}(f,q;s)$. Let $\overline{\mathcal{S}} = \{s \in \mathcal{S} \mid q(s) > 0\}$ be the support of q. Let $U_{\min} = \min_{s\in\overline{\mathcal{S}}} U(\pi_s)$ and $\mathcal{S}_{\min} = \{s \in \overline{\mathcal{S}} \,|\, U(\pi_s) = U_{\min}\}$. For any $s \notin \overline{\mathcal{S}}$, we trivially have $p^\star_{\theta,u}(f,q;s) = 0$ for all $\theta \geq 0$. For $s \in \overline{\mathcal{S}}$, by multiplying and dividing $p^\star_{\theta,u}(f,q;s)$ by $\exp(\theta U_{\min})$, one has:

$$p^\star_{\theta,u}(f,q;s) = \frac{q(s)\exp\left(\theta(U_{\min} - U(\pi_s))\right)}{\sum_{t\in\mathcal{S}} q(t)\exp\left(\theta(U_{\min} - U(\pi_t))\right)} = \frac{q(s)\exp\left(\theta(U_{\min} - U(\pi_s))\right)}{\sum_{t\in\mathcal{S}_{\min}} q(t) + \sum_{t\notin\mathcal{S}_{\min}} q(t)\exp\left(\theta(U_{\min} - U(\pi_t))\right)}\,.$$

Moreover, for all $t \in \overline{\mathcal{S}} \setminus \mathcal{S}_{\min}$, we have $\lim_{\theta\to\infty}\exp\left(\theta(U_{\min} - U(\pi_t))\right) = \exp(-\infty) = 0$. It follows that:

$$\lim_{\theta\to\infty} p^\star_{\theta,u}(f,q;s) = \begin{cases} \frac{q(s)}{\sum_{t\in\mathcal{S}_{\min}} q(t)}\,, & \text{for } s \in \mathcal{S}_{\min}\,, \\ 0\,, & \text{for } s \notin \mathcal{S}_{\min}\,. \end{cases} \tag{21}$$

By Theorem 4, we have $V_{\theta,u}(f,q) = \mathbb{E}_{p^\star_{\theta,u}}[U] + D_{KL}(p^\star_{\theta,u}||q)/\theta$. Observe from (21) that $D_{KL}(p^\star_{\theta,u}||q)$ converges to a finite constant when $\theta \to \infty$, so that $\lim_{\theta\to\infty} D_{KL}(p^\star_{\theta,u}||q)/\theta = 0$. One deduces that $\lim_{\theta\to\infty} V_{\theta,u}(f,q) = \lim_{\theta\to\infty}\sum_{s\in\mathcal{S}} p^\star_{\theta,u}(f,q;s)U(\pi_s) = U_{\min}$. □

Proof of Theorem 7. We first prove (i) ⇔ (ii). If $\theta = 0$, then $V_{0,u}(f,q) = \sum_{s\in\mathcal{S}} q(s)U(\pi_s)$ and the equivalence is a result of Jensen's inequality. Now assume that $\theta, \theta' > 0$. We know from the proof of Theorem 5 that $V_{\theta,u}(f,q) = (-1/\theta)\,K_U(-\theta)$, hence:

$$\frac{\partial V_{\theta,u}(f,q)}{\partial\theta} = \frac{K_U(-\theta) + \theta K_U^{(1)}(-\theta)}{\theta^2}\,. \tag{22}$$

By definition, and because we restrict our evaluation to societies where there exist $s, s' \in \mathcal{S}$ such that $U_s \neq U_{s'}$, the CGF $K_U(-\theta)$ is strictly convex. Thus $K_U(x) - K_U(y) - (x-y)K_U^{(1)}(y) > 0$ for all $x, y \in \mathbb{R}$, with $x \neq y$. Let $x = 0$ and $y = -\theta$. It follows that $K_U(0) > K_U(-\theta) + \theta K_U^{(1)}(-\theta)$. Moreover, $K_U(0) = \ln\left(\sum_{s\in\mathcal{S}} q(s)e^{-0\times U(\pi_s)}\right) = 0$, thus the numerator of (22) is strictly negative, and $\partial V_{\theta,u}(f,q)/\partial\theta < 0$.

Next, we prove (i) ⇒ (iii). Assume $\theta < \theta'$. Consider an opportunity profile $f = (\pi_s, \pi_{s'}) \in \mathcal{F}_{\mathcal{S}}$, with $U(\pi_s) < U(\pi_{s'})$ and $q = (1/2, 1/2)$. Let $p^\star_{\theta,u}(f,q;s) = p^\star_{\theta,u}(s)$ for all $s \in \mathcal{S}$. From Theorem 4 we have, for $t = s, s'$:

$$p^\star_{\theta,u}(t) = \frac{\exp\left(-\theta U(\pi_t)\right)}{\exp\left(-\theta U(\pi_s)\right) + \exp\left(-\theta U(\pi_{s'})\right)}\,.$$

If $\theta = 0$, then $p^\star_{\theta,u}(s) = p^\star_{\theta,u}(s') = 1/2$. If $\theta > 0$, then $U(\pi_s) < U(\pi_{s'})$ implies $p^\star_{\theta,u}(s) > p^\star_{\theta,u}(s')$. After simplification, we also have $p^\star_{\theta,u}(s) = \left[1 + \exp\left(-\theta(U(\pi_{s'}) - U(\pi_s)\right)\right]^{-1}$. One deduces that:

$$\frac{\partial p^\star_{\theta,u}(s)}{\partial\theta} = (U(\pi_{s'}) - U(\pi_s)) \times \exp\left(-\theta(U(\pi_{s'}) - U(\pi_s)\right) \times p^\star_{\theta,u}(s)^2 > 0\,. \tag{23}$$

Thus, $\theta < \theta'$ implies that $p^\star_{\theta,u}(s) < p^\star_{\theta',u}(s)$. Because $p^\star_{\theta,u}(s') = 1 - p^\star_{\theta,u}(s)$, we have $\frac{\partial p^\star_{\theta,u}(s')}{\partial \theta} < 0$ and $p^\star_{\theta,u}(s') > p^\star_{\theta',u}(s')$.

We now prove (iii) $\Rightarrow$ (i). Assume statement (iii) holds. Suppose, by contradiction, $\theta \geq \theta'$. If $\theta = \theta'$, then $p^\star_{\theta,u}(s) = p^\star_{\theta',u}(s)$, which contradicts the strict inequality in (iii). If $\theta > \theta'$, the strict monotonicity established by the positive derivative in (23) implies $p^\star_{\theta,u}(s) > p^\star_{\theta',u}(s)$, which again contradicts (iii). Therefore, we must have $\theta < \theta'$. $\square$

B. Formal connections with decision theory under ambiguity

Our axiomatic characterization bridges the literature on inequality of opportunity with the literature on decision theory under ambiguity. Formally, our setup is equivalent to the Anscombe and Aumann (1963) framework: The 'states of the world' correspond to the social types $s \in \mathcal{S}$, while the lotteries assigned to each state are type-specific income distributions $\pi_s \in \Delta(\mathcal{Y})$. An 'act', which assigns a lottery to each state of the world, corresponds to an opportunity profile $f \in \mathcal{F}_{\mathcal{S}}$. Because our framework is entirely finite and discrete in states (to directly accommodate empirical data), we have $\mathcal{F}_{\mathcal{S}} = \Delta(\mathcal{Y})^{|\mathcal{S}|}$.

Discussion related to Theorem 1. Our baseline representation in Theorem 1 is directly connected to the 'second-order utility' (SOEU) and 'smooth ambiguity' models characterized by Klibanoff *et al.* (2005), Ergin and Gul (2009), and Grant *et al.* (2009).[27] However, our framework departs from these models in two fundamental ways: the nature and derivation of the outer expectation, and the characterization of (ambiguity / inequality-of-opportunity) aversion.

First, in the ambiguity literature, the second-order measure operates as a representation of a decision-maker's *subjective uncertainty* regarding the 'true' probability model (Klibanoff *et al.*, 2005; Ergin and Gul, 2009). To derive this subjective prior, these models rely on heavy axiomatic machinery, such as complex behavioral assumptions over hypothetical second-order acts (Assumptions 2 and 3 in Klibanoff *et al.*, 2005), rich Savage-type product state spaces designed to capture unobservable second-order risks (Axioms 4 and 5b in Ergin and Gul, 2009), or the necessity to either impose a highly demanding ambiguity aversion axiom or couple standard uncertainty aversion with a demanding translation invariance condition (A.5 for Theorem 2, or A.6–A.7 for Theorem 3 in Grant *et al.*, 2009).[28] By contrast, our probability distribution $q \in \Delta(\mathcal{S})$ is discrete, uniform, and represents perfectly known, *objective* demographic weights across types. That allows us to bypass the need to axiomatize unobservable subjective beliefs over a state space. By substituting the heavy machinery of subjective priors with the anonymity requirement (A3) and a weak independence axiom applied to equal opportunity profiles (A5), we recover the exact SOEU structure using a much leaner framework.

Second, our inequality-of-opportunity aversion axiom (A4) differs from the usual axioms of uncertainty aversion. Standard approaches in decision theory rely on global mixing conditions across all states

[27] The dynamic choice model of Kreps and Porteus (1978) also shares this recursive structure. However, it relies on temporal consistency to evaluate the timing of uncertainty resolution, whereas our static representation builds on an independence axiom to compare societies.

[28] It is worth noting that while Grant *et al.* (2009)'s translation invariance (A.7) is not required to establish our baseline SOEU representation (Theorem 1), we explicitly impose a strong form of translation invariance along the equal-opportunity line via A5', but only in Theorem 2 to restrict the planner's preferences to the single-parameter exponential family.

simultaneously (Gilboa and Schmeidler, 1989; Schmeidler, 1989) or comparisons to ambiguity-neutral benchmarks (Ghirardato and Marinacci, 2002). Instead, our axiom operates *locally* between two types, providing a direct formal equivalent to the Pigou-Dalton transfer principle from the standard literature on inequality measurement.

Discussion related to Theorem 2. In the literature on decision under ambiguity, Strzalecki (2011) provides an axiomatic foundation for the single-parameter exponential form, which we obtain in Theorem 2. However, his axiomatic route is different. Because he defines this family of preferences as a subclass of the 'variational preferences' introduced by Maccheroni *et al.* (2006), he is required to first assume their full suite of axioms to establish this baseline. To then restrict this class down to the exponential form, he additionally imposes the 'P2 Sure-Thing Principle' of Savage (1954), over the entire Anscombe-Aumann domain of acts. By contrast, our derivation takes the SOEU general class as a starting point (Theorem 1). Rather than imposing Savage's global P2 axiom over a general space of acts, we rely on the local row-separability condition A2, already used to derive our Theorem 1. Indeed, on our finite product domain, repeated applications of A2 naturally yield the sure-thing implications necessary for the representation. To restrict the general SOEU representation to the exponential family, we do not need a novel axiom, but only a strengthening of the existing one: The independence condition A5 is replaced by A5'. Note that A5 is comparable, but adapted to our framework, to the 'risk independence' axiom in Cerreia-Vioglio *et al.* (2011), whereas A5' corresponds to the 'weak certainty independence' axiom in Maccheroni *et al.* (2006) for variational preferences.

Discussion related to Theorem 4. The dual representation of the exponential family in Theorem 4, using the Kullback-Leibler (KL) divergence, has already been identified by Strzalecki (2011). In fact, he demonstrates that this specific exponential form constitutes the unique class of functions lying at the exact intersection between variational preferences (Maccheroni *et al.*, 2006) and SOEU preferences. This unique intersection corresponds to the 'multiplier preferences' introduced by Hansen and Sargent (2001). However, the mathematical derivation differs substantially. The duality established by Strzalecki (2011) builds on results from the 'theory of large deviations' discussed in Dupuis and Ellis (1997). By contrast, because our objective framework already yields the exponential form as a baseline (Theorem 2), we obtain the dual KL representation much more transparently through a simple optimization argument. Furthermore, rather than penalizing model misspecification as in the decision theory literature, the KL divergence measures the endogenous displacement of normative weights (induced by the profile of type-specific expected utilities and the social planner's aversion to inequality of opportunity) from demographic weights.

Discussion related to Theorem 5. Our Theorem 5 establishes an additive separation within the representation function: It disentangles the efficiency component from the inequality component, with

the latter explicitly integrating the planner's aversion parameter. In standard risk theory, such an additive separation is achieved via the mean-variance local approximation of the Expected Utility (EU) model (typically assuming CARA or CRRA preferences), as pioneered by Arrow (1965) and Pratt (1964). Broadening this classic framework, Nock *et al.* (2011) identify a closed-form 'mean-divergence' representation for the EU model that is comparable to what we obtain in Theorem 5. They demonstrate that for exponential families of distributions, the risk premium translates exactly into a Bregman divergence.

In the specific context of SOEU preferences, Maccheroni *et al.* (2013) identified a generalization of the Arrow-Pratt approximation. They derive a mean-variance evaluation (based on a second-order Taylor expansion) where the ambiguity premium is additively separated and determined by the variance of expected returns across subjective models, multiplied by the agent's ambiguity aversion index. As the authors note, their local approximation becomes exact only under restrictive conditions, namely CARA utility combined with a normal distribution of risks. Our mean-variance approximation in Theorem 5 relies on a different variance penalty term. Instead of capturing an ambiguity premium through the variance of unobservable subjective beliefs, our variance term measures a purely objective reality: the variance of expected outcomes across observable demographic types.

Furthermore, our mathematical derivation takes an original route by explicitly relying on the cumulant generating function. By restricting this expansion to the second order, we recover the mean-variance approximation, mapping the planner's inequality aversion to the second cumulant. The true analytical advantage of our framework lies beyond this quadratic truncation. The Bregman divergence approach allows us to isolate the first-order cumulant (the expectation of type-specific expected utilities) from all distributional features beyond the mean, by operating through an algebraic identity rather than relying on an infinite series expansion. It encapsulates these remaining dimensions into an exact, closed-form expression. Consequently, rather than being forced to rely on a local approximation, our model yields a globally tractable structure in which this exact divergence serves as a comprehensive penalty for inequality of opportunity.

C. Representation with a logarithmic utility function

This section follows on from Sections 3 and 4, where we have assumed a fixed distribution $q \in \Delta(\mathcal{S})$ over types, uniform across all societies, such that $q(s) = 1/|\mathcal{S}|$ for all $s \in \mathcal{S}$. Under Axioms A1-A5', Theorem 2 in Section 4 characterizes the following social welfare function, for any opportunity profile $f \in \mathcal{F}_{\mathcal{S}}$:

$$V_{\theta,u,q}(f) = \begin{cases} \sum_{s\in\mathcal{S}} q(s)U(\pi_s)\,, & \text{for } \theta = 0\,, \\ (-1/\theta)\ln\left(\sum_{s\in\mathcal{S}} q(s)\left(e^{-\theta U(\pi_s)}\right)\right)\,, & \text{for } \theta > 0\,. \end{cases} \tag{24}$$

We establish that two additional (yet standard) axioms imply that, within this framework, strictly positive inequality-of-opportunity aversion necessarily leads to a logarithmic utility function in income.

The following axiom states that, in the absence of inequality of opportunity, social welfare increases with additional income, regardless of the individual who receives it. Specifically, among equal-opportunity profiles, one profile is preferred to another if it can be obtained from the latter through a 'favorable probability shift'.

Axiom 9 (Monotonicity). *For all $\pi, \pi' \in \Delta(\mathcal{Y})$, if there exist $\epsilon > 0$ and two incomes $x < x'$ such that $\pi(y) = \pi'(y)$ for all $y \neq x, x'$ and, furthermore, $\pi(x) = \pi'(x) - \epsilon$ and $\pi(x') = \pi'(x') + \epsilon$, then $f_\pi \succ f_{\pi'}$.*

Consider the following example:

$$f_\pi = \begin{array}{c} \\ \text{R} \\ \text{P} \end{array}\!\!\begin{array}{c} \begin{array}{cc} \$1 & \$2 \end{array} \\ \begin{pmatrix} 0.2 & 0.8 \\ 0.2 & 0.8 \end{pmatrix} \end{array} \succ \begin{array}{c} \\ \text{R} \\ \text{P} \end{array}\!\!\begin{array}{c} \begin{array}{cc} \$1 & \$2 \end{array} \\ \begin{pmatrix} 0.4 & 0.6 \\ 0.4 & 0.6 \end{pmatrix} \end{array} = f_{\pi'}\,.$$

The next axiom requires that multiplying all incomes by a positive constant (e.g., converting between currencies) should not affect the ranking of two opportunity profiles. For any profile $f \in \mathcal{F}_{\mathcal{S}}$ and any $\lambda > 0$, we define a new profile λf by $(\lambda f)(s) = \pi_s^\lambda$ where, for any $s \in \mathcal{S}$ and any $y \in \mathcal{Y}$, $\pi_s^\lambda(y) = \pi_s(y/\lambda)$. That is, profile λf corresponds to f where the income of all individuals is multiplied by λ.

Axiom 10 (Scale invariance). *For all $f, f' \in \mathcal{F}_{\mathcal{S}}$ and all $\lambda > 0$, if $\lambda f, \lambda f' \in \mathcal{F}_{\mathcal{S}}$, then $f \succcurlyeq f' \;\Rightarrow\; \lambda f \succcurlyeq \lambda f'$.*

Consider the following example, where we set $\lambda = 3$:

$$\begin{array}{c} \\ \mathrm{R} \\ \mathrm{P} \end{array}\begin{array}{c} \begin{array}{cc} \$1 & \$2 \end{array} \\ \begin{pmatrix} 0.2 & 0.8 \\ 0.2 & 0.8 \end{pmatrix} \end{array} \succcurlyeq \begin{array}{c} \\ \mathrm{R} \\ \mathrm{P} \end{array}\begin{array}{c} \begin{array}{cc} \$1 & \$2 \end{array} \\ \begin{pmatrix} 0.4 & 0.6 \\ 0.4 & 0.6 \end{pmatrix} \end{array} \implies \begin{array}{c} \\ \mathrm{R} \\ \mathrm{P} \end{array}\begin{array}{c} \begin{array}{cc} \$3 & \$6 \end{array} \\ \begin{pmatrix} 0.2 & 0.8 \\ 0.2 & 0.8 \end{pmatrix} \end{array} \succcurlyeq \begin{array}{c} \\ \mathrm{R} \\ \mathrm{P} \end{array}\begin{array}{c} \begin{array}{cc} \$3 & \$6 \end{array} \\ \begin{pmatrix} 0.4 & 0.6 \\ 0.4 & 0.6 \end{pmatrix} \end{array}.$$

The following result demonstrates the implications of these two new axioms, in addition to those of Theorem 2. The proof is provided at the end of this Appendix.

Theorem 8. *Let $q(s) = 1/|\mathcal{S}|$ for all $s \in \mathcal{S}$, and $|\mathcal{S}| \geq 3$. The social planner's preferences relation $\succcurlyeq$ satisfies A1-A5' (as in Theorem 2) and A9-A10 iff it can be represented by the function $V_{\theta,u_\sigma,q}$, with $V_{\theta,u_\sigma,q}$ as defined in Theorem 2, and with the following utility function u_σ. If $\theta = 0$ then, for all $y \in \mathcal{Y}$:*

$$u_\sigma(y) = \begin{cases} -y^\sigma\,, & \text{for } \sigma < 0\,, \\ \ln y\,, & \text{for } \sigma = 0\,, \\ y^\sigma\,, & \text{for } \sigma > 0\,. \end{cases} \tag{25}$$

If $\theta > 0$, then $u_\sigma(y) = \ln y$, for all $y \in \mathcal{Y}$ and all $\sigma \in \mathbb{R}$. To simplify the notation, we let $V_{\theta,q} = V_{\theta,\ln,q}$.

The resulting restriction is significant when inequality-of-opportunity aversion is strictly positive: The social welfare function now has only one parameter, namely θ, which captures the planner's aversion to inequality of opportunity. Whilst this restriction is relatively strong from a theoretical perspective, it is not particularly restrictive for empirical studies, where the use of a logarithmic utility function is fairly standard. A real advantage is that the normative choices are limited to selecting the value of θ.

Note that this representation remains valid on the extended domain (as presented in Section 5), where we compare societies $(f, q) \in \mathcal{D}$, with potentially different populations, type structures, and type distributions. In this case, we denote the social welfare function $V_\theta(f, q)$ instead of $V_{\theta,q}(f)$.

Proof of Theorem 8. Assume that the axioms stated in the theorem are satisfied. From Theorem 2 we know that, under A1-A5', preference $\succcurlyeq$ can be represented by the function $V_{\theta,u,q}$. Because u is defined up to a positive affine transformation, we assume, without loss of generality, that $u \geq 0$. Consider $\pi, \pi' \in \Delta(\mathcal{Y})$ as in A9, so that $f_\pi \succ f_{\pi'}$. This is equivalent to $\sum_{y\in\mathcal{Y}} \pi(y)u(y) > \sum_{y\in\mathcal{Y}} \pi'(y)u(y)$ or, by definition, to $\epsilon(u(x') - u(x)) > 0$ with $x < x'$ and $\epsilon > 0$. Therefore, A9 implies that u is strictly increasing.

Now, assume A10, let $\lambda > 0$ and consider two profiles $f, f' \in \mathcal{F}_\mathcal{S}$ such that $f \succcurlyeq f'$. By definition of λf,

we have $V_{\theta,u,q}(\lambda f) = \phi_\theta^{-1}\left(\sum_{s\in\mathcal{S}} q(s)\phi_\theta\left(\sum_{y\in\mathcal{Y}} \pi_s^\lambda(\lambda y)u(\lambda y)\right)\right)$. Recalling that $\pi_s^\lambda(\lambda y) = \pi_s(y)$ and denoting $u^\lambda(y) = u(\lambda y)$, we have $V_{\theta,u,q}(\lambda f) = V_{\theta,u^\lambda,q}(f)$. Using A10, $f \succcurlyeq f' \Leftrightarrow \lambda f \succcurlyeq \lambda f'$. For any $\lambda > 0$ such that the overlap $\mathcal{Y} \cap \lambda^{-1}\mathcal{Y}$ forms an interval, this implies that $V_{\theta,u,q}$ and $V_{\theta,u^\lambda,q}$ represent the same preference, restricted to profiles supported on this interval. Applying the uniqueness result of Theorem 2 to this interval, u^λ must be a positive affine transformation of u specifically on this overlap. Thus, there exist $a(\lambda) > 0$ and $b(\lambda) \in \mathbb{R}$ such that $u^\lambda(y) = u(\lambda y) = a(\lambda)u(y) + b(\lambda)$ for all $y, \lambda y \in \mathcal{Y}$. As established by Aczél (1966) on unbounded domains (see Equation (42), Page 159) and extended in Aczél (2005) for measurable functions on an interval (Corollary p. 3233), the solution to this equation is any increasing affine transformation of $u_\sigma(y)$ as introduced in (25). Note that we cannot consider $u_\sigma(y) = -\ln y$ as solution, because u has to be increasing (A9). The second restriction is $\theta' = \theta/a(\lambda)$. Because $V_{\theta',u',q} = V_{\theta,u^\lambda,q}$, we have $\theta = \theta/a(\lambda)$. If $\theta = 0$, this condition is satisfied and (25) is the global solution. If $\theta > 0$, then $a(\lambda) = 1$, and we must have $u^\lambda(y) = u(\lambda y) = u(y) + b(\lambda)$, whose solution reduces to $u(y) = \ln y$.

Conversely, it is not difficult to establish that u_σ is sufficient to satisfy A9-A10 by observing that, when $u_\sigma(y) = \ln y$ and $\theta \geq 0$, $U(\pi_s^\lambda) = U(\pi_s) + \ln\lambda$ and $V_{\theta,\ln,q}(\lambda f) = V_{\theta,\ln,q}(f) + \ln\lambda$. □

D. Additional Exhibits

Table A.1: Summary statistics for birth cohorts (1940s–1980s)

	Mean	SD	Min	Max
1940s ($N = 484$)				
Male	0.50	0.50	0.00	1.00
Non-Hispanic White	0.81	0.39	0.00	1.00
Age	34.86	0.64	31.00	38.00
Child income y^c ('000 USD)	72.62	43.10	0.11	273.92
Parent income y^p ('000 USD)	59.50	43.74	0.01	300.06
1950s ($N = 2,413$)				
Male	0.49	0.50	0.00	1.00
Non-Hispanic White	0.76	0.43	0.00	1.00
Age	34.70	1.04	31.00	39.00
Child income y^c ('000 USD)	72.37	54.12	0.04	516.78
Parent income y^p ('000 USD)	65.47	50.35	0.00	490.47
1960s ($N = 1,842$)				
Male	0.49	0.50	0.00	1.00
Non-Hispanic White	0.68	0.47	0.00	1.00
Age	34.07	1.00	31.00	38.00
Child income y^c ('000 USD)	76.50	63.74	0.04	777.59
Parent income y^p ('000 USD)	68.17	51.85	0.00	540.95
1970s ($N = 1,713$)				
Male	0.49	0.50	0.00	1.00
Non-Hispanic White	0.60	0.49	0.00	1.00
Age	34.89	0.56	32.00	38.00
Child income y^c ('000 USD)	80.83	67.52	0.01	904.62
Parent income y^p ('000 USD)	67.23	55.78	0.04	684.81
1980s ($N = 2,229$)				
Male	0.50	0.50	0.00	1.00
Non-Hispanic White	0.58	0.49	0.00	1.00
Age	34.04	1.06	32.00	38.00
Child income y^c ('000 USD)	77.49	73.52	0.07	1,141.78
Parent income y^p ('000 USD)	73.94	82.96	0.00	1,413.26

Data: PSID.

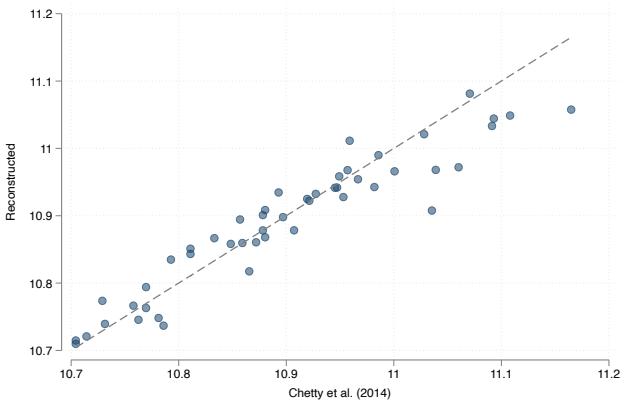
Note: This table reports descriptive statistics for our analysis sample separately for each decadal birth cohort from the 1940s (1940–1949) to the 1980s (1980–1989). Children's income y^c is measured as the average of total family labor income over ages 30–40, and parental income y^p is measured as a weighted average of total family labor income during the child's adolescence. All income values are expressed in 2019 USD. Observations are weighted to match the relative frequencies of sex $\times$ race cells in the corresponding pooled Current Population Survey (CPS, 1969–2023).

Table A.2: Summary statistics for the 50 largest CZs

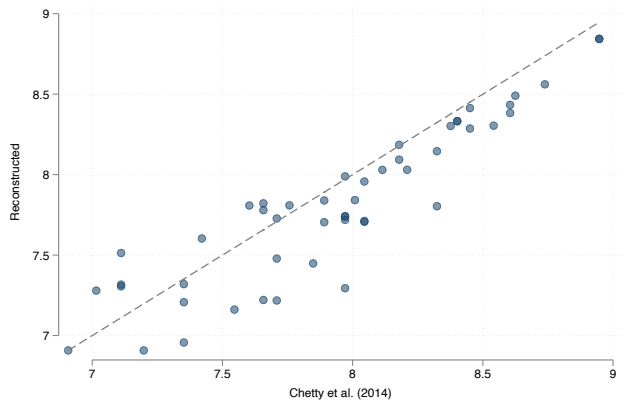
	State	N children ('000)	Share male[a]	Share non-Hispanic White[a]	Mean child income ('000 USD)
Atlanta	GA	206.19	0.51	0.53	47.24
Austin	TX	84.52	0.51	0.55	53.01
Baltimore	MD	181.24	0.51	0.60	56.49
Boston	MA	309.64	0.51	0.75	62.87
Bridgeport	CT	225.77	0.51	0.71	61.90
Buffalo	NY	181.45	0.51	0.80	56.47
Charlotte	NC	97.40	0.51	0.63	48.73
Chicago	IL	576.41	0.51	0.54	56.53
Cincinnati	OH	157.68	0.51	0.81	54.64
Cleveland	OH	202.81	0.51	0.74	52.03
Columbus	OH	120.63	0.51	0.77	52.40
Dallas	TX	241.51	0.51	0.52	54.06
Dayton	OH	94.63	0.51	0.81	51.59
Denver	CO	169.22	0.51	0.68	57.44
Detroit	MI	392.12	0.51	0.70	52.47
Fort Worth	TX	135.27	0.51	0.61	55.55
Fresno	CA	109.07	0.51	0.36	46.14
Grand Rapids	MI	109.43	0.51	0.81	56.08
Houston	TX	345.84	0.51	0.45	55.70
Indianapolis	IN	115.70	0.51	0.76	53.87
Jacksonville	FL	91.83	0.51	0.64	45.01
Kansas City	MO	139.10	0.51	0.74	55.96
Las Vegas	NV	89.04	0.51	0.56	47.74
Los Angeles	CA	1018.87	0.51	0.34	49.88
Manchester	NH	91.28	0.51	0.90	59.28
Miami	FL	236.35	0.51	0.30	46.02
Milwaukee	WI	131.29	0.51	0.72	57.97
Minneapolis	MN	219.66	0.51	0.82	64.95
Nashville	TN	88.52	0.51	0.74	50.76
New Orleans	LA	118.80	0.51	0.49	46.42
New York	NY	660.84	0.51	0.38	54.60
Newark	NJ	361.29	0.51	0.53	63.42
Orlando	FL	114.69	0.51	0.55	44.79
Philadelphia	PA	407.58	0.51	0.65	57.87
Phoenix	AZ	208.32	0.51	0.62	51.19
Pittsburgh	PA	181.13	0.51	0.86	61.15
Port St. Lucie	FL	87.12	0.51	0.58	46.54
Portland	OR	131.47	0.51	0.77	54.23
Providence	RI	104.93	0.51	0.79	57.19
Raleigh	NC	86.70	0.51	0.63	51.96
Sacramento	CA	182.42	0.51	0.56	52.08
Salt Lake City	UT	137.34	0.51	0.84	60.56
San Antonio	TX	138.51	0.51	0.37	47.40
San Diego	CA	175.30	0.51	0.45	53.01
San Francisco	CA	270.01	0.51	0.43	58.22
San Jose	CA	147.56	0.51	0.39	57.99
Seattle	WA	275.14	0.51	0.72	57.25
St. Louis	MO	192.62	0.51	0.72	55.39
Tampa	FL	135.60	0.51	0.66	45.29
Washington DC	DC	303.16	0.51	0.50	62.59

Data: Opportunity Insights.
Note: This table reports descriptive statistics for the 50 largest commuting zones (CZs) in the US, based on the birth cohorts 1980–1985 Chetty *et al.* (2014a). [a] The share of males and non-Hispanic Whites is reported in (Chetty *et al.*, 2026) for the birth cohort 1978–1983. All income values are expressed in 2019 USD.

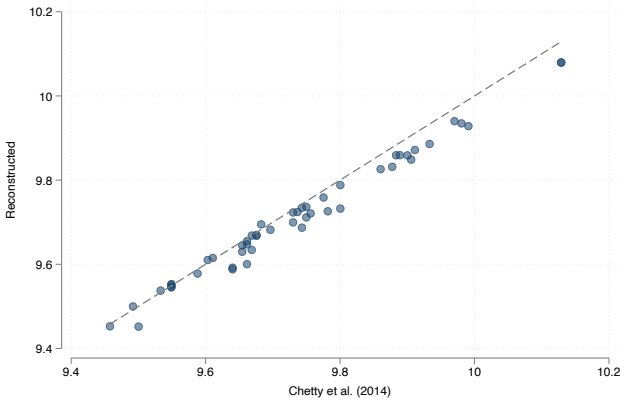
Figure A.1: Validation of marginal distribution against Chetty *et al.* (2014a)



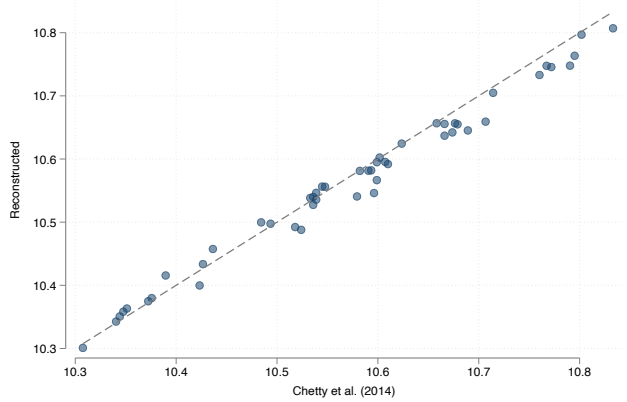
(a) Mean



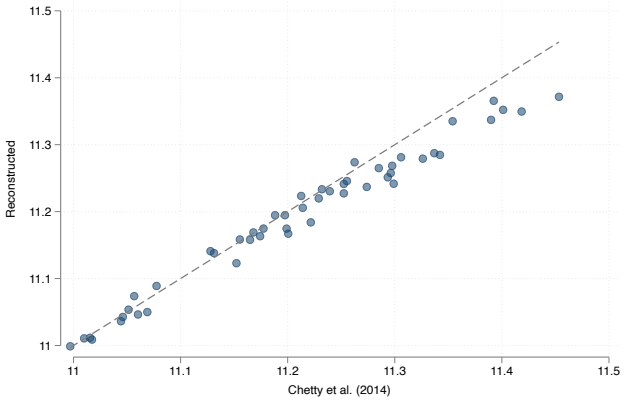
(b) P10



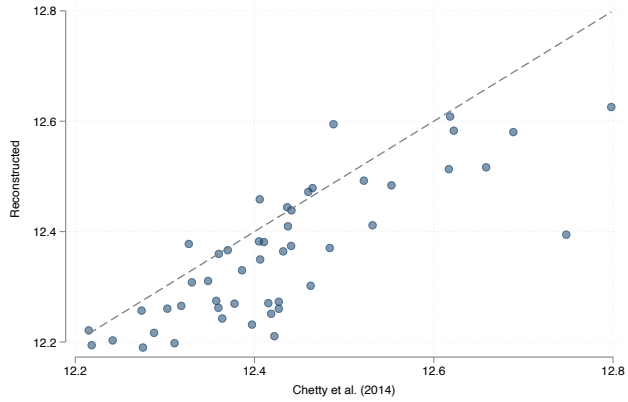
(c) P25



(d) P50



(e) P75

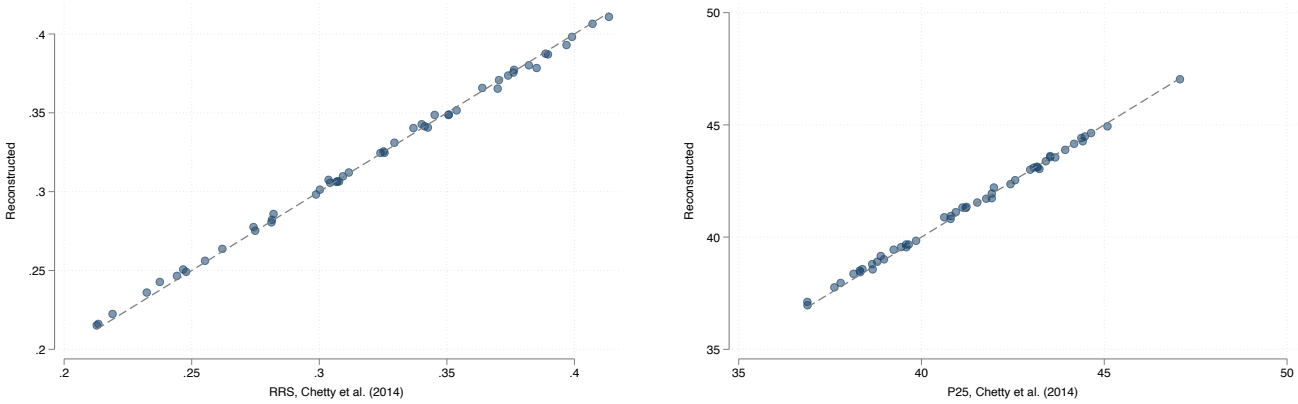


(f) P99

Data: Opportunity Insights.

Note: This figure compares moments of the income distributions of the 50 largest CZs. Panels (a)–(f) show mean income and the 10th, 25th, 50th, 75th, and 99th percentiles, respectively. In each panel, we plot the log income from our reconstructed distribution (y-axis) against the corresponding log income calculated by Chetty *et al.* (2014a) from individual-level microdata. In the case of perfect alignment, all data points would cluster on the dashed 45-degree line. Moments reported in Chetty *et al.* (2014a) are based on the 1980–1982 birth cohort, and reconstructed moments are based on the 1980–1985 cohort.

Figure A.2: Validation of joint distribution against Chetty *et al.* (2014a)



(a) RRS

(b) P25

Data: Opportunity Insights.

Note: This figure compares moments of the joint income distributions of the 50 largest CZs. Panel (a) shows the rank-rank slope (RRS), and Panel (b) shows the mean child rank at the 25th percentile of the parental income distribution (P25). In the case of perfect alignment, all data points would cluster on the dashed 45-degree line. Estimates of Chetty *et al.* (2014a) are based on the 1980–1985 birth cohort.

E. Construction of statistical confidence sets

This section describes how we construct the statistical confidence sets and the combined sets reported in Tables S.4 and S.9. The procedure follows Mogstad *et al.* (2024) closely. While Mogstad *et al.* (2024) rank CZs by an estimate of intergenerational mobility, we rank societies by social welfare $V^j_{\theta,u}$. This has two main implications. First, standard errors are not readily available, and we therefore simulate the sampling distribution of $V^j_{\theta,u}$. Second, since $V^j_{\theta,u}$ varies over θ, the set of pairwise comparisons underlying the inference procedure ranges over triples (j, k, θ) rather than over pairs of societies (j, k). Evaluating dominance relations across pairs *and* over a range of θ values turns statistical confidence sets into combined sets.

Sampling distribution. We first simulate the sampling distribution of $V^j_{\theta,u}$ by generating $B = 2,000$ draws of the estimated inputs to the welfare function. In both applications, we assume independence across the societies of our comparison, which is justified given that children are assigned to only one cohort and CZ, respectively.[29] The resampled object differs across the two applications. In the trends application (Section 8.1), we estimate Equation (8) separately for each cohort j and draw the intercept and slope jointly from a normal distribution centered on the point estimates, using the estimated variance-covariance matrix of the estimators. Note that these draws capture sampling variation in the estimated gradient but not in the type distribution itself. Each draw is then mapped into type utilities through Equation (9). In the spatial application (Section 8.2), we do not observe individual children. Instead, the simulation targets the CZ-specific joint distribution of parental and child income quintiles $p^{CZ}(r^p, k) = q^{CZ}(r^p) \times t^{CZ}(k \mid r^p)$, i.e., the first two factors of Equation (11). For each CZ, we draw the $5 \times 5 = 25$ cell counts from a multinomial distribution with sample size N^{CZ} and cell probabilities p^{CZ} provided in Chetty *et al.* (2014a), which is equivalent to resampling individual children with replacement. We then re-estimate the IRS (Equation (12)) on the resulting frequency weights, holding the national within-quintile placement $h(r^c \mid k(r^c), r^p)$ and the percentile-income schedule $g(y^c \mid r^c)$ fixed across draws. We obtain one welfare estimate $V^{j,b}_{\theta,u}$ per society j and θ for a common draw b.

Standard errors of the pairwise welfare differences. We then convert these draws into standard errors for the pairwise welfare differences $\Delta^{jk}_\theta = V^j_{\theta,u} - V^k_{\theta,u}$. For every pair (j, k) and every θ in the grid $\Theta \cap [\theta_{\text{min}}, \theta_{\text{max}}]$, we estimate $\widehat{se}_{jk}(\theta)$ as the standard deviation of $\hat{\Delta}^{jk,b}_\theta$ across the $b = 1, \ldots, B$ draws.

[29] Technically, draws are only independent conditional on common inputs such as the CPS-calibrated weights in the cohort application, or the national within-quintile placement function $h(r^c \mid k(r^c), r^p)$ and the national percentile-income schedule $g(y^c \mid r^c)$ in the CZ application. Given the sample sizes underlying these national statistics, we treat the dependence induced by these inputs as negligible.

Critical values and confidence regions. Given the standard errors, we compute critical values following Section 3.1 of Mogstad *et al.* (2024). For a set of comparisons $S \subseteq \{(j,k,\theta) \in \mathcal{J} \times \mathcal{J} \times \Theta : j \neq k\}$, the critical value $c(S)$ is the $1-\alpha$ quantile of the studentized maximum absolute deviation of $\hat{\Delta}_\theta^{jk,b}$ from $\hat{\Delta}_\theta^{jk}$ over all comparisons in S, which we estimate from the B draws. The main difference to Equation (8) in Mogstad *et al.* (2024) is that the maximum is taken over (j,k,θ) rather than over (j,k). The confidence region $C(1-\alpha, S)$ for the welfare differences in S is then given by the intervals $\hat{\Delta}_\theta^{jk} \pm c(S)\, \widehat{se}_{jk}(\theta)$ (Equation (11) in Mogstad *et al.* (2024)). We set $\alpha = 0.05$ throughout, so that the confidence region covers all welfare differences in S simultaneously with 95% probability.

Defining dominance relations. We next translate the confidence region into dominance relations. Society j dominates k if the confidence band for Δ_θ^{jk} lies entirely above zero at every θ contained in S, i.e., if the smallest lower endpoint across those θ values is strictly positive. In the statistical confidence sets, S contains a fixed θ_0, and only one confidence band is evaluated for each pairwise difference. In the combined sets, dominance also requires that the lower endpoint be positive at every $\theta \in \Theta \cap [\theta_{\text{min}}, \theta_{\text{max}}]$. We then count the numbers n_j^+ and n_j^- of societies that j dominates and that dominate j, and form the set of ranks that j can occupy according to Equation (10) and Equation (13) in Mogstad *et al.* (2024). Tables S.4 and S.9 vary which comparisons enter S and how large $c(S)$ is as a result.

The four types of sets. Setting $c(S) = 0$ at a single θ_0 returns the point-estimate ranking, and setting $c(S) = 0$ while letting θ range over $[\theta_{\text{min}}, \theta_{\text{max}}]$ reproduces the normative dominance sets $R_j^{[\theta_{\text{min}},\theta_{\text{max}}]}$. Allowing $c(S) > 0$ at a single θ_0 gives the statistical confidence set $R_j^{1-\alpha}(\theta_0)$, which we report at $\theta_0 \in \{0, 2, \infty\}$. Allowing $c(S) > 0$ and letting S range over the whole grid $\Theta \cap [\theta_{\text{min}}, \theta_{\text{max}}]$ yields the combined set $R_j^{[\theta_{\text{min}},\theta_{\text{max}}],\, 1-\alpha}$.[30]

Marginal versus simultaneous coverage. Finally, we distinguish marginal from simultaneous sets, following Sections 3.2 and 3.3 of Mogstad *et al.* (2024). Letting S contain only the comparisons involving society j yields the *marginal* sets, which cover the rank of a single society with probability $1-\alpha$; letting it contain all pairs yields the *simultaneous* sets, which jointly cover the entire ranking profile. Since $c(S)$ is weakly increasing in the scope of S, the reported sets nest by construction: Marginal within simultaneous, $R_j^{[0,2]}$ within $R_j^{[0,\infty)}$, and normative and statistical within combined. For simplicity, we do not implement the stepwise multiple-testing procedure in Sections 3.2.2 and 3.3.2 of Mogstad *et al.* (2024), which would yield slightly narrower confidence sets.

[30] We maintain the fixed θ grid used to compute dominance sets when computing combined sets. This grid ensures that pairwise welfare differences $\hat{\Delta}_\theta^{jk}$ do not change sign between grid nodes (Footnote 19). Technically, combined sets require that the lower endpoint of the confidence band $\hat{\Delta}_\theta^{jk} - c(S)\, \widehat{se}_{jk}(\theta)$ remain positive between nodes. We use the maintained grid as an approximation to this stronger requirement.

Supplementary Material

Opportunity-Sensitive Social Welfare

Tim Wienand ⓡ Brice Magdalou ⓡ Richard Nock ⓡ Paul Hufe

Online publication only

S.1. Supplementary Tables and Figures

Table S.1: Dominance sets for birth cohorts under alternative trimming of extreme values

	Baseline			Trim $y^c[p1, p99]$			Trim $y^c[p5, p95]$		
Cohort	Rank	$R^{[0,2]}$	$R^{[0,\infty)}$	Rank	$R^{[0,2]}$	$R^{[0,\infty)}$	Rank	$R^{[0,2]}$	$R^{[0,\infty)}$
1940s	1	[1,1]	[1,1]	1	[1,1]	[1,1]	2	[1,2]	[1,2]
1950s	4	[2,4]	[2,4]	4	[2,4]	[2,4]	4	[2,4]	[2,4]
1960s	3	[3,4]	[3,4]	3	[3,4]	[3,4]	3	[3,4]	[3,4]
1970s	2	[2,3]	[2,3]	2	[2,3]	[2,3]	1	[1,3]	[1,3]
1980s	5	[5,5]	[5,5]	5	[5,5]	[5,5]	5	[5,5]	[5,5]
Average width		1.8	1.8		1.8	1.8		2.2	2.2

Data: PSID.
Note: This table reports dominance sets for the birth cohorts 1940s–1980s conditional on trimming the lower and upper tails of the child income distribution. The baseline uses the full distribution of positive child income y^c. In the other columns, y^c is trimmed at the 1st and 99th percentiles, or the 5th and 95th percentiles, of the cohort-specific distributions of y^c, respectively. Sampling weights are renormalized in each cohort. *Rank* reports the ranking by $V_{0,u}$; rank 1 (5) corresponds to the highest (lowest) $V_{0,u}$.

Table S.2: Dominance sets for birth cohorts under alternative functional forms

		Linear		Quadratic		Cubic		Quintiles, nonparametric		Quintiles, linear	
Cohort	Rank	$R^{[0,2]}$	$R^{[0,\infty)}$	$R^{[0,2]}$	$R^{[0,\infty)}$	$R^{[0,2]}$	$R^{[0,\infty)}$	$R^{[0,2]}$	$R^{[0,\infty)}$	$R^{[0,2]}$	$R^{[0,\infty)}$
1940s	1	[1,1]	[1,1]	[1,1]	[1,1]	[1,1]	[1,1]	[1,1]	[1,1]	[1,1]	[1,1]
1950s	4	[2,4]	[2,4]	[4,4]	[2,4]	[3,4]	[2,4]	[4,4]	[4,4]	[3,4]	[2,4]
1960s	3	[3,4]	[3,4]	[3,3]	[3,4]	[3,4]	[3,4]	[2,3]	[2,3]	[2,3]	[2,4]
1970s	2	[2,3]	[2,3]	[2,2]	[2,3]	[2,2]	[2,3]	[2,3]	[2,3]	[2,4]	[2,4]
1980s	5	[5,5]	[5,5]	[5,5]	[5,5]	[5,5]	[5,5]	[5,5]	[5,5]	[5,5]	[5,5]
Average width		1.8	1.8	1.0	1.8	1.4	1.8	1.4	1.4	1.8	2.2

Data: PSID.
Note: This table reports dominance sets for the birth cohorts 1940s–1980s under different functional forms of Equation (8). The linear specification replicates the baseline results of Figure 3b. The quadratic and cubic specifications contain second- and third-order polynomials in $\ln y^p$, respectively. The remaining two specifications redefine types as parental income quintiles and estimate $U(\pi_s)$ as the weighted mean of $\ln y^c$ within types (nonparametric) and as a linear projection analogous to the baseline specification. *Rank* reports the ranking by $V_{0,u}$; rank 1 (5) corresponds to the highest (lowest) $V_{0,u}$.

Table S.3: Dominance sets for birth cohorts under alternative income definitions

	Family labor income			Family labor income, equivalized			Family disposable income			Family disposable income, equivalized		
Cohort	Rank	$R^{[0,2]}$	$R^{[0,\infty)}$	Rank	$R^{[0,2]}$	$R^{[0,\infty)}$	Rank	$R^{[0,2]}$	$R^{[0,\infty)}$	Rank	$R^{[0,2]}$	$R^{[0,\infty)}$
1940s	1	[1,1]	[1,1]	4	[1,4]	[1,4]	2	[1,2]	[1,2]	4	[3,4]	[1,4]
1950s	4	[2,4]	[2,4]	5	[2,5]	[2,5]	4	[4,4]	[2,4]	5	[4,5]	[4,5]
1960s	3	[3,4]	[3,4]	2	[1,4]	[1,4]	3	[3,3]	[2,4]	2	[2,2]	[1,3]
1970s	2	[2,3]	[2,3]	1	[1,4]	[1,4]	1	[1,2]	[1,4]	1	[1,1]	[1,3]
1980s	5	[5,5]	[5,5]	3	[3,5]	[3,5]	5	[5,5]	[5,5]	3	[3,5]	[3,5]
Average width		1.8	1.8		3.8	3.8		1.4	2.6		1.8	3.0

Data: PSID.
Note: This table reports dominance sets for the birth cohorts 1940s–1980s for two definitions of family income: the baseline measure of pre-tax-and-transfer labor income and post-tax-and-transfer disposable income. For each income definition, we calculate equivalized income as the ratio of total family income to the square root of the number of household members. *Rank* reports the ranking by $V_{0,u}$; rank 1 (5) corresponds to the highest (lowest) $V_{0,u}$.

Table S.4: Dominance sets for birth cohorts accounting for statistical uncertainty

		Normative		Statistical (95%)						Combined (95%)			
				Marginal			Simultaneous			Marginal		Simultaneous	
Cohort	Rank	$R^{[0,2]}$	$R^{[0,\infty)}$	$\theta=0$	$\theta=2$	$\theta\to\infty$	$\theta=0$	$\theta=2$	$\theta\to\infty$	$R^{[0,2]}$	$R^{[0,\infty)}$	$R^{[0,2]}$	$R^{[0,\infty)}$
1940s	1	[1,1]	[1,1]	[1,5]	[1,2]	[1,1]	[1,5]	[1,2]	[1,1]	[1,5]	[1,5]	[1,5]	[1,5]
1950s	4	[2,4]	[2,4]	[1,5]	[1,4]	[2,2]	[1,5]	[1,4]	[2,2]	[1,5]	[1,5]	[1,5]	[1,5]
1960s	3	[3,4]	[3,4]	[1,5]	[2,4]	[3,4]	[1,5]	[2,4]	[3,4]	[1,5]	[1,5]	[1,5]	[1,5]
1970s	2	[2,3]	[2,3]	[1,4]	[2,4]	[3,4]	[1,5]	[2,4]	[3,4]	[1,5]	[1,5]	[1,5]	[1,5]
1980s	5	[5,5]	[5,5]	[1,5]	[5,5]	[5,5]	[1,5]	[5,5]	[5,5]	[1,5]	[1,5]	[1,5]	[1,5]
Average width		1.8	1.8	4.8	2.6	1.4	5.0	2.6	1.4	5.0	5.0	5.0	5.0

Data: PSID.
Note: This table reports normative, statistical, and combined sets for the birth cohorts 1940s–1980s. Details of the construction are discussed in Appendix E. The normative dominance sets $R^{[0,2]}$ and $R^{[0,\infty)}$ are the baseline results reported in Figure 3b. Statistical confidence sets account for sampling variation in the parameter estimates in Equation (8) used to compute expected type utilities $U(\pi_s)$. They are reported for the two limits of inequality-of-opportunity aversion θ and an intermediate level ($\theta = 2$). Following Mogstad *et al.* (2024), we report marginal confidence sets for the rankings of individual cohorts and simultaneous confidence sets for the full ranking profile. Combined sets apply the normative dominance evaluation over both $\theta \in [0, 2]$ and $\theta \in [0, \infty)$ to the marginal and simultaneous confidence sets. *Rank* reports the ranking by $V_{0,u}$; rank 1 (5) corresponds to the highest (lowest) $V_{0,u}$.

Table S.5: Dominance sets and sharp sets for birth cohorts

		Dominance sets R_j		Sharp sets S_j	
Cohort	Rank	$R^{[0,2]}$	$R^{[0,\infty)}$	$S^{[0,2]}$	$S^{[0,\infty)}$
1940s	1	[1,1]	[1,1]	[1,1]	[1,1]
1950s	4	[2,4]	[2,4]	[2,4]	[2,4]
1960s	3	[3,4]	[3,4]	[3,4]	[3,4]
1970s	2	[2,3]	[2,3]	[2,3]	[2,3]
1980s	5	[5,5]	[5,5]	[5,5]	[5,5]
Average width		1.8	1.8	1.8	1.8

Data: PSID.
Note: This table compares the dominance sets $R^{[0,2]}$ and $R^{[0,\infty)}$ reported in Figure 3b with the sharp attainable rank sets $S^{[0,2]}$ and $S^{[0,\infty)}$ for the birth cohorts 1940s–1980s. Dominance sets contain every rank a cohort could hold in any complete ranking consistent with the unambiguous pairwise dominance preorder over the respective range of θ. Sharp sets span only the ranks a cohort actually attains across that range. *Rank* reports the ranking by $V_{0,u}$; rank 1 (5) corresponds to the highest (lowest) $V_{0,u}$.

Table S.6: Dominance sets for the 50 largest CZs under alternative trimming of extreme values

	Baseline			Trim $y^c[p1, p99]$			Trim $y^c[p5, p95]$		
Commuting zone	Rank	$R^{[0,2]}$	$R^{[0,\infty)}$	Rank	$R^{[0,2]}$	$R^{[0,\infty)}$	Rank	$R^{[0,2]}$	$R^{[0,\infty)}$
Minneapolis	1	[1,2]	[1,8]	1	[1,1]	[1,5]	1	[1,1]	[1,2]
Salt Lake City	2	[1,2]	[1,2]	2	[2,2]	[1,2]	2	[2,2]	[1,2]
Boston	3	[3,3]	[3,9]	3	[3,3]	[3,8]	3	[3,3]	[3,5]
Newark	4	[4,7]	[4,17]	4	[4,7]	[4,14]	4	[4,7]	[4,12]
Pittsburgh	5	[4,6]	[2,12]	5	[4,6]	[3,10]	5	[4,5]	[4,7]
Washington DC	6	[5,7]	[4,15]	6	[5,7]	[5,13]	6	[5,6]	[5,11]
Bridgeport	7	[7,8]	[7,24]	7	[7,8]	[7,21]	7	[7,8]	[7,18]
Manchester	8	[4,8]	[2,10]	8	[4,8]	[2,9]	8	[6,8]	[3,8]
Milwaukee	9	[9,21]	[9,38]	9	[9,19]	[9,36]	9	[9,16]	[9,32]
Denver	10	[9,10]	[6,16]	10	[9,10]	[6,15]	10	[9,10]	[7,14]
Providence	11	[10,14]	[9,22]	11	[10,14]	[9,19]	11	[10,14]	[9,15]
Buffalo	12	[11,16]	[11,30]	12	[11,16]	[11,26]	13	[12,16]	[12,23]
Grand Rapids	13	[11,16]	[11,33]	13	[11,16]	[11,31]	14	[12,16]	[12,26]
Seattle	14	[10,15]	[5,18]	14	[10,14]	[5,18]	12	[10,12]	[6,15]
San Francisco	15	[11,16]	[3,17]	15	[11,16]	[4,17]	15	[11,15]	[6,16]
San Jose	16	[10,16]	[2,16]	16	[11,16]	[2,16]	16	[12,16]	[4,16]
Philadelphia	17	[17,22]	[16,35]	17	[17,22]	[16,33]	17	[17,19]	[16,32]
Kansas City	18	[16,22]	[15,32]	18	[16,20]	[15,32]	18	[17,18]	[16,29]
Fort Worth	19	[16,20]	[11,24]	19	[16,19]	[12,23]	19	[18,19]	[13,23]
Chicago	20	[20,26]	[19,36]	20	[20,24]	[19,36]	20	[20,23]	[20,33]
Houston	21	[17,23]	[9,24]	21	[19,22]	[11,24]	21	[20,22]	[13,24]
St. Louis	22	[22,30]	[22,43]	22	[22,29]	[22,41]	23	[22,26]	[22,40]
Baltimore	23	[23,31]	[22,41]	23	[23,30]	[22,40]	22	[22,27]	[22,39]
Portland	24	[16,24]	[12,29]	24	[17,24]	[12,28]	24	[20,24]	[13,26]
Cincinnati	25	[25,33]	[25,49]	25	[25,33]	[25,48]	25	[25,30]	[25,44]
Dallas	26	[22,27]	[17,36]	26	[22,27]	[18,36]	26	[23,27]	[19,33]
Indianapolis	27	[26,33]	[26,46]	27	[26,33]	[26,44]	27	[27,32]	[26,44]
San Diego	28	[18,28]	[3,28]	28	[21,28]	[4,28]	28	[22,28]	[8,28]
Austin	29	[24,30]	[18,37]	29	[25,29]	[19,37]	30	[26,30]	[19,34]
New York	30	[26,31]	[15,34]	30	[26,31]	[17,34]	29	[27,30]	[18,33]
Columbus	31	[31,36]	[30,47]	31	[31,35]	[30,45]	31	[31,34]	[29,44]
Cleveland	32	[32,37]	[28,47]	32	[32,36]	[29,45]	33	[33,36]	[30,44]
Dayton	33	[31,36]	[27,45]	33	[31,36]	[27,44]	32	[31,34]	[27,42]
Sacramento	34	[24,34]	[14,36]	34	[26,34]	[16,35]	34	[29,34]	[18,35]
Phoenix	35	[28,35]	[19,40]	35	[28,35]	[20,39]	35	[31,35]	[20,36]
Detroit	36	[36,38]	[32,47]	36	[36,37]	[35,46]	36	[36,37]	[36,45]
Raleigh	37	[37,42]	[36,48]	37	[37,41]	[36,48]	37	[37,39]	[37,46]
Nashville	38	[37,41]	[28,46]	38	[37,40]	[29,46]	38	[37,39]	[31,46]
Los Angeles	39	[29,39]	[6,39]	39	[31,39]	[9,39]	39	[35,39]	[14,39]
San Antonio	40	[38,41]	[20,42]	40	[39,41]	[21,42]	41	[41,41]	[28,42]
Las Vegas	41	[35,41]	[20,43]	41	[38,41]	[21,42]	40	[40,40]	[27,41]
Charlotte	42	[42,47]	[42,50]	42	[42,46]	[42,50]	42	[42,44]	[42,50]
Miami	43	[41,43]	[13,43]	43	[42,43]	[18,43]	43	[42,43]	[23,43]
Atlanta	44	[44,48]	[42,50]	45	[45,48]	[44,50]	46	[46,48]	[45,50]
Port St. Lucie	45	[43,46]	[29,46]	44	[43,45]	[30,45]	44	[44,45]	[33,45]
Fresno	46	[43,46]	[22,46]	46	[43,46]	[26,46]	45	[43,45]	[30,45]
Tampa	47	[45,47]	[31,47]	47	[45,47]	[33,47]	47	[46,47]	[35,47]
Orlando	48	[46,48]	[33,48]	48	[47,48]	[34,48]	48	[47,48]	[36,48]
Jacksonville	49	[49,49]	[37,49]	49	[49,49]	[40,49]	49	[49,49]	[47,49]
New Orleans	50	[50,50]	[41,50]	50	[50,50]	[46,50]	50	[50,50]	[48,50]
Average width		5.6	17.6		4.8	15.7		3.5	12.8

Data: Opportunity Insights.
Note: This table reports dominance sets for the 50 largest CZs conditional on trimming the lower and upper tails of the child income distribution. The baseline uses the full distribution of positive child income y^c. In the other columns, y^c is trimmed at the 1st and 99th percentiles, or the 5th and 95th percentiles, of the CZ-specific distribution, respectively. Sampling weights are renormalized in each CZ. *Rank* reports the ranking by $V_{0,u}$; rank 1 (50) corresponds to the highest (lowest) $V_{0,u}$.

Table S.7: Dominance sets for the 50 largest CZs under alternative functional forms

		Linear		Quadratic		Cubic		Non-parametric	
Commuting zone	Rank	$R^{[0,2]}$	$R^{[0,\infty)}$	$R^{[0,2]}$	$R^{[0,\infty)}$	$R^{[0,2]}$	$R^{[0,\infty)}$	$R^{[0,2]}$	$R^{[0,\infty)}$
Minneapolis	1	[1,2]	[1,8]	[1,2]	[1,19]	[1,2]	[1,21]	[1,2]	[1,22]
Salt Lake City	2	[1,2]	[1,2]	[1,2]	[1,3]	[1,2]	[1,3]	[1,2]	[1,3]
Boston	3	[3,3]	[3,9]	[3,3]	[2,11]	[3,3]	[2,12]	[3,3]	[2,11]
Newark	4	[4,7]	[4,17]	[4,6]	[3,15]	[4,6]	[3,15]	[4,6]	[3,15]
Pittsburgh	5	[4,6]	[2,12]	[5,7]	[4,19]	[5,7]	[4,19]	[5,7]	[4,20]
Washington DC	6	[5,7]	[4,15]	[4,7]	[2,12]	[4,7]	[2,12]	[4,7]	[2,13]
Bridgeport	7	[7,8]	[7,24]	[7,8]	[7,27]	[7,8]	[7,27]	[7,8]	[7,27]
Manchester	8	[4,8]	[2,10]	[4,8]	[2,12]	[4,8]	[2,12]	[4,8]	[2,13]
Milwaukee	9	[9,21]	[9,38]	[9,22]	[9,48]	[9,22]	[9,48]	[9,22]	[9,48]
Denver	10	[9,10]	[6,16]	[9,11]	[5,19]	[9,11]	[5,19]	[9,11]	[5,19]
Providence	11	[10,14]	[9,22]	[10,14]	[9,22]	[10,14]	[8,23]	[10,14]	[8,23]
Buffalo	12	[11,16]	[11,30]	[11,16]	[11,33]	[11,17]	[11,33]	[11,17]	[11,34]
Grand Rapids	13	[11,16]	[11,33]	[11,16]	[11,36]	[11,16]	[11,36]	[11,16]	[11,36]
Seattle	14	[10,15]	[5,18]	[10,15]	[5,20]	[10,15]	[5,21]	[10,15]	[5,21]
San Francisco	15	[11,16]	[3,17]	[11,16]	[2,17]	[11,16]	[2,17]	[11,16]	[2,17]
San Jose	16	[10,16]	[2,16]	[9,16]	[1,16]	[9,16]	[1,16]	[9,16]	[1,16]
Philadelphia	17	[17,22]	[16,35]	[16,22]	[15,36]	[16,22]	[15,35]	[16,22]	[15,35]
Kansas City	18	[16,22]	[15,32]	[16,22]	[15,35]	[16,22]	[15,35]	[16,22]	[14,35]
Fort Worth	19	[16,20]	[11,24]	[16,20]	[10,26]	[16,20]	[10,26]	[16,20]	[10,26]
Chicago	20	[20,26]	[19,36]	[20,26]	[17,36]	[20,27]	[18,36]	[20,27]	[18,36]
Houston	21	[17,23]	[9,24]	[17,23]	[5,24]	[17,23]	[6,25]	[17,23]	[5,25]
St. Louis	22	[22,30]	[22,43]	[22,31]	[21,47]	[22,31]	[21,47]	[22,31]	[21,47]
Baltimore	23	[23,31]	[22,41]	[22,31]	[21,41]	[22,31]	[21,39]	[22,31]	[21,39]
Portland	24	[16,24]	[12,29]	[16,24]	[10,29]	[15,24]	[10,29]	[15,24]	[10,29]
Cincinnati	25	[25,33]	[25,49]	[25,35]	[25,50]	[25,35]	[25,50]	[25,35]	[25,50]
Dallas	26	[22,27]	[17,36]	[22,27]	[16,35]	[22,27]	[16,35]	[22,27]	[16,35]
Indianapolis	27	[26,33]	[26,46]	[26,33]	[24,45]	[26,33]	[24,45]	[26,33]	[24,45]
San Diego	28	[18,28]	[3,28]	[18,28]	[3,28]	[18,28]	[3,28]	[18,28]	[3,28]
Austin	29	[24,30]	[18,37]	[24,30]	[18,37]	[24,30]	[18,37]	[24,30]	[18,37]
New York	30	[26,31]	[15,34]	[26,31]	[10,33]	[25,31]	[8,33]	[25,31]	[9,33]
Columbus	31	[31,36]	[30,47]	[31,37]	[29,49]	[31,37]	[30,50]	[31,37]	[30,50]
Cleveland	32	[32,37]	[28,47]	[32,38]	[28,48]	[32,38]	[28,49]	[32,38]	[28,49]
Dayton	33	[31,36]	[27,45]	[30,37]	[28,48]	[30,37]	[28,49]	[30,37]	[28,49]
Sacramento	34	[24,34]	[14,36]	[24,34]	[11,36]	[24,34]	[11,36]	[24,34]	[11,36]
Phoenix	35	[28,35]	[19,40]	[28,35]	[19,40]	[28,35]	[19,40]	[28,35]	[19,40]
Detroit	36	[36,38]	[32,47]	[34,38]	[28,46]	[34,38]	[29,47]	[34,38]	[29,47]
Raleigh	37	[37,42]	[36,48]	[37,42]	[33,49]	[37,42]	[31,49]	[37,42]	[31,49]
Nashville	38	[37,41]	[28,46]	[37,42]	[30,49]	[37,42]	[30,49]	[37,42]	[30,49]
Los Angeles	39	[29,39]	[6,39]	[29,39]	[5,39]	[29,39]	[5,39]	[29,39]	[5,39]
San Antonio	40	[38,41]	[20,42]	[38,41]	[20,42]	[38,41]	[18,42]	[38,41]	[17,42]
Las Vegas	41	[35,41]	[20,43]	[33,41]	[20,42]	[33,41]	[20,42]	[33,41]	[19,42]
Charlotte	42	[42,47]	[42,50]	[42,47]	[41,50]	[42,47]	[39,50]	[42,47]	[39,50]
Miami	43	[41,43]	[13,43]	[40,43]	[9,43]	[40,43]	[7,43]	[40,43]	[6,43]
Atlanta	44	[44,48]	[42,50]	[44,48]	[32,49]	[44,48]	[32,49]	[44,48]	[32,49]
Port St. Lucie	45	[43,46]	[29,46]	[43,46]	[27,46]	[43,46]	[27,46]	[43,46]	[27,46]
Fresno	46	[43,46]	[22,46]	[43,46]	[23,46]	[43,46]	[23,46]	[43,46]	[23,46]
Tampa	47	[45,47]	[31,47]	[45,47]	[29,47]	[45,47]	[29,47]	[45,47]	[29,47]
Orlando	48	[46,48]	[33,48]	[46,48]	[30,48]	[46,48]	[30,48]	[46,48]	[30,48]
Jacksonville	49	[49,49]	[37,49]	[49,49]	[36,49]	[49,49]	[37,49]	[49,49]	[37,49]
New Orleans	50	[50,50]	[41,50]	[50,50]	[40,50]	[50,50]	[39,50]	[50,50]	[39,50]
Average width		5.6	17.6	6.0	19.9	6.1	20.2	6.1	20.3

Data: Opportunity Insights.
Note: This table reports dominance sets for the 50 largest CZs under different functional forms of Equation (12). The linear specification replicates the baseline results of Figure 5b. The quadratic and cubic specifications contain second- and third-order polynomials in r^p, respectively. The nonparametric specification estimates $U(\pi_s)$ as the weighted mean of $\ln y^c$ within types. *Rank* reports the ranking by $V_{0,u}$; rank 1 (50) corresponds to the highest (lowest) $V_{0,u}$.

Table S.8: Dominance sets for the 50 largest CZs under alternative percentile-income schedules

	Baseline (Chetty et al., 2014)			Alternative (Chetty et al., 2026)		
Commuting zone	Rank	$R^{[0,2]}$	$R^{[0,\infty)}$	Rank	$R^{[0,2]}$	$R^{[0,\infty)}$
Minneapolis	1	[1,2]	[1,8]	1	[1,2]	[1,8]
Salt Lake City	2	[1,2]	[1,2]	2	[1,2]	[1,2]
Boston	3	[3,3]	[3,9]	3	[3,3]	[3,9]
Newark	4	[4,7]	[4,17]	4	[4,7]	[4,17]
Pittsburgh	5	[4,6]	[2,12]	5	[4,7]	[2,12]
Washington DC	6	[5,7]	[4,15]	6	[4,7]	[4,15]
Bridgeport	7	[7,8]	[7,24]	7	[7,8]	[7,24]
Manchester	8	[4,8]	[2,10]	8	[4,8]	[2,10]
Milwaukee	9	[9,21]	[9,38]	9	[9,21]	[9,38]
Denver	10	[9,10]	[6,16]	10	[9,11]	[6,17]
Providence	11	[10,14]	[9,22]	11	[10,14]	[9,22]
Buffalo	12	[11,16]	[11,30]	12	[11,15]	[11,29]
Grand Rapids	13	[11,16]	[11,33]	13	[12,16]	[12,34]
Seattle	14	[10,15]	[5,18]	14	[10,15]	[5,18]
San Francisco	15	[11,16]	[3,17]	15	[11,16]	[3,18]
San Jose	16	[10,16]	[2,16]	16	[9,16]	[2,16]
Philadelphia	17	[17,22]	[16,35]	17	[17,22]	[16,35]
Kansas City	18	[16,22]	[15,32]	18	[16,22]	[15,33]
Fort Worth	19	[16,20]	[11,24]	19	[16,20]	[11,24]
Chicago	20	[20,26]	[19,36]	20	[20,29]	[19,36]
Houston	21	[17,23]	[9,24]	21	[17,23]	[9,24]
St. Louis	22	[22,30]	[22,43]	22	[22,30]	[22,43]
Baltimore	23	[23,31]	[22,41]	23	[23,31]	[22,41]
Portland	24	[16,24]	[12,29]	24	[16,24]	[12,29]
Cincinnati	25	[25,33]	[25,49]	25	[25,35]	[25,49]
Dallas	26	[22,27]	[17,36]	26	[22,28]	[17,36]
Indianapolis	27	[26,33]	[26,46]	27	[26,33]	[26,46]
San Diego	28	[18,28]	[3,28]	28	[18,28]	[3,28]
Austin	29	[24,30]	[18,37]	29	[24,30]	[18,37]
New York	30	[26,31]	[15,34]	30	[25,31]	[15,33]
Columbus	31	[31,36]	[30,47]	31	[31,36]	[30,47]
Cleveland	32	[32,37]	[28,47]	32	[32,39]	[28,47]
Dayton	33	[31,36]	[27,45]	33	[30,37]	[27,45]
Sacramento	34	[24,34]	[14,36]	34	[23,34]	[14,36]
Phoenix	35	[28,35]	[19,40]	35	[27,36]	[20,40]
Detroit	36	[36,38]	[32,47]	36	[35,39]	[32,47]
Raleigh	37	[37,42]	[36,48]	37	[37,42]	[36,49]
Nashville	38	[37,41]	[28,46]	38	[37,42]	[28,46]
Los Angeles	39	[29,39]	[6,39]	39	[27,39]	[5,39]
San Antonio	40	[38,41]	[20,42]	40	[36,41]	[20,42]
Las Vegas	41	[35,41]	[20,43]	41	[33,41]	[20,43]
Charlotte	42	[42,47]	[42,50]	42	[42,47]	[42,50]
Miami	43	[41,43]	[13,43]	43	[40,43]	[12,43]
Atlanta	44	[44,48]	[42,50]	44	[44,49]	[41,50]
Port St. Lucie	45	[43,46]	[29,46]	45	[43,46]	[28,46]
Fresno	46	[43,46]	[22,46]	46	[43,46]	[22,46]
Tampa	47	[45,47]	[31,47]	47	[45,47]	[31,47]
Orlando	48	[46,48]	[33,48]	48	[46,48]	[32,48]
Jacksonville	49	[49,49]	[37,49]	49	[48,49]	[37,49]
New Orleans	50	[50,50]	[41,50]	50	[50,50]	[41,50]
Average width		5.6	17.6		6.2	17.7

Data: Opportunity Insights.
Note: This table reports dominance sets for the 50 largest CZs using the different percentile-income schedules reported by Chetty *et al.* (2014a) (baseline) and Chetty *et al.* (2026). *Rank* reports the ranking by $V_{0,u}$; rank 1 (50) corresponds to the highest (lowest) $V_{0,u}$.

Table S.9: Dominance sets for the 50 largest CZs accounting for statistical uncertainty

		Normative		Statistical (95%)						Combined (95%)			
				Marginal			Simultaneous			Marginal		Simultaneous	
Commuting zone	Rank	$R^{[0,2]}$	$R^{[0,\infty)}$	$\theta=0$	$\theta=2$	$\theta\to\infty$	$\theta=0$	$\theta=2$	$\theta\to\infty$	$R^{[0,2]}$	$R^{[0,\infty)}$	$R^{[0,2]}$	$R^{[0,\infty)}$
Minneapolis	1	[1,2]	[1,8]	[1,1]	[1,2]	[5,10]	[1,1]	[1,2]	[4,12]	[1,2]	[1,11]	[1,2]	[1,12]
Salt Lake City	2	[1,2]	[1,2]	[2,2]	[1,2]	[1,1]	[2,2]	[1,2]	[1,1]	[1,2]	[1,2]	[1,2]	[1,2]
Boston	3	[3,3]	[3,9]	[3,3]	[3,4]	[7,10]	[3,3]	[3,4]	[5,12]	[3,4]	[2,12]	[3,4]	[2,12]
Newark	4	[4,7]	[4,17]	[4,5]	[5,7]	[12,20]	[4,5]	[5,7]	[12,20]	[4,7]	[4,20]	[4,7]	[4,20]
Pittsburgh	5	[4,6]	[2,12]	[4,5]	[5,7]	[4,10]	[4,7]	[4,7]	[4,11]	[4,7]	[2,14]	[4,8]	[2,15]
Washington DC	6	[5,7]	[4,15]	[5,7]	[5,7]	[11,18]	[5,7]	[5,7]	[10,20]	[4,8]	[4,20]	[4,8]	[4,21]
Bridgeport	7	[7,8]	[7,24]	[6,7]	[8,8]	[22,28]	[5,8]	[8,10]	[21,28]	[5,9]	[5,28]	[5,10]	[5,28]
Manchester	8	[4,8]	[2,10]	[8,8]	[3,4]	[3,6]	[7,8]	[3,5]	[2,9]	[3,8]	[2,12]	[3,8]	[2,12]
Milwaukee	9	[9,21]	[9,38]	[9,13]	[20,22]	[36,43]	[9,13]	[19,22]	[36,43]	[9,22]	[9,43]	[9,22]	[9,44]
Denver	10	[9,10]	[6,16]	[9,13]	[9,12]	[10,18]	[9,13]	[8,12]	[8,19]	[8,16]	[4,22]	[8,16]	[3,24]
Providence	11	[10,14]	[9,22]	[9,14]	[13,13]	[13,21]	[9,14]	[13,13]	[12,21]	[9,16]	[6,25]	[9,16]	[6,26]
Buffalo	12	[11,16]	[11,30]	[9,14]	[14,17]	[24,28]	[9,15]	[14,17]	[22,28]	[9,18]	[8,30]	[9,18]	[8,30]
Grand Rapids	13	[11,16]	[11,33]	[9,14]	[14,17]	[29,34]	[9,16]	[14,17]	[29,34]	[9,18]	[9,35]	[9,18]	[9,35]
Seattle	14	[10,15]	[5,18]	[11,15]	[9,12]	[7,12]	[11,16]	[9,12]	[7,15]	[9,16]	[3,19]	[9,16]	[3,21]
San Francisco	15	[11,16]	[3,17]	[14,17]	[9,12]	[4,8]	[12,18]	[9,12]	[4,9]	[9,17]	[2,19]	[9,18]	[2,20]
San Jose	16	[10,16]	[2,16]	[15,18]	[9,12]	[2,3]	[13,18]	[8,12]	[2,4]	[9,18]	[2,19]	[8,18]	[2,19]
Philadelphia	17	[17,22]	[16,35]	[15,18]	[21,22]	[33,34]	[15,18]	[21,23]	[31,34]	[14,23]	[13,35]	[14,23]	[13,35]
Kansas City	18	[16,22]	[15,32]	[16,18]	[18,21]	[29,33]	[15,19]	[16,21]	[29,33]	[15,22]	[14,35]	[14,22]	[13,35]
Fort Worth	19	[16,20]	[11,24]	[19,21]	[14,20]	[11,20]	[18,22]	[14,20]	[10,20]	[14,23]	[9,27]	[14,24]	[8,28]
Chicago	20	[20,26]	[19,36]	[19,21]	[24,28]	[35,35]	[19,22]	[23,28]	[35,36]	[19,29]	[18,38]	[19,29]	[18,39]
Houston	21	[17,23]	[9,24]	[19,23]	[17,20]	[10,15]	[19,23]	[16,21]	[7,16]	[16,25]	[6,28]	[16,25]	[5,28]
St. Louis	22	[22,30]	[22,43]	[21,23]	[30,31]	[37,47]	[19,23]	[29,32]	[36,47]	[20,31]	[18,47]	[19,32]	[18,49]
Baltimore	23	[23,31]	[22,41]	[21,24]	[30,32]	[36,43]	[21,25]	[30,34]	[36,46]	[21,33]	[20,46]	[21,34]	[20,47]
Portland	24	[16,24]	[12,29]	[23,26]	[14,18]	[15,21]	[23,27]	[14,20]	[13,23]	[14,27]	[10,33]	[14,28]	[9,35]
Cincinnati	25	[25,33]	[25,49]	[24,27]	[32,36]	[46,50]	[23,27]	[31,37]	[45,50]	[24,36]	[24,50]	[23,38]	[22,50]
Dallas	26	[22,27]	[17,36]	[24,27]	[23,25]	[24,28]	[24,27]	[23,26]	[23,28]	[21,30]	[15,37]	[21,30]	[15,37]
Indianapolis	27	[26,33]	[26,46]	[25,29]	[31,36]	[41,48]	[24,30]	[30,36]	[39,49]	[24,36]	[23,49]	[22,36]	[21,49]
San Diego	28	[18,28]	[3,28]	[27,30]	[16,20]	[2,4]	[27,30]	[16,20]	[2,4]	[16,30]	[2,30]	[16,30]	[2,30]
Austin	29	[24,30]	[18,37]	[27,31]	[23,27]	[22,28]	[27,31]	[22,28]	[22,28]	[22,32]	[16,38]	[21,33]	[16,38]
New York	30	[26,31]	[15,34]	[28,30]	[24,28]	[15,20]	[27,31]	[24,28]	[14,20]	[23,33]	[12,36]	[23,33]	[10,36]
Columbus	31	[31,36]	[30,47]	[30,33]	[32,38]	[41,48]	[29,35]	[32,40]	[39,49]	[28,39]	[26,49]	[27,40]	[25,49]
Cleveland	32	[32,37]	[28,47]	[31,35]	[34,40]	[37,47]	[31,35]	[33,40]	[36,47]	[31,40]	[26,48]	[30,40]	[26,48]
Dayton	33	[31,36]	[27,45]	[31,36]	[32,36]	[36,43]	[31,37]	[31,38]	[36,46]	[29,38]	[26,47]	[28,39]	[25,48]
Sacramento	34	[24,34]	[14,36]	[32,37]	[23,26]	[12,20]	[31,37]	[23,27]	[11,20]	[23,37]	[10,39]	[23,37]	[10,39]
Phoenix	35	[28,35]	[19,40]	[32,37]	[26,29]	[22,28]	[31,37]	[25,29]	[22,28]	[24,38]	[17,42]	[24,38]	[17,42]
Detroit	36	[36,38]	[32,47]	[33,37]	[35,40]	[41,47]	[33,37]	[34,40]	[39,48]	[31,41]	[27,49]	[30,41]	[27,49]
Raleigh	37	[37,42]	[36,48]	[34,38]	[39,42]	[42,49]	[33,38]	[39,42]	[41,49]	[34,42]	[29,49]	[33,42]	[28,50]
Nashville	38	[37,41]	[28,46]	[37,38]	[37,42]	[36,41]	[37,39]	[35,42]	[35,43]	[34,42]	[28,48]	[34,42]	[27,48]
Los Angeles	39	[29,39]	[6,39]	[39,39]	[28,29]	[4,8]	[38,39]	[28,30]	[4,8]	[28,39]	[4,39]	[28,39]	[4,39]
San Antonio	40	[38,41]	[20,42]	[40,42]	[37,42]	[19,23]	[40,42]	[35,42]	[19,24]	[35,43]	[18,44]	[35,43]	[18,44]
Las Vegas	41	[35,41]	[20,43]	[40,42]	[32,38]	[21,26]	[40,42]	[32,40]	[20,28]	[32,42]	[18,44]	[32,42]	[17,44]
Charlotte	42	[42,47]	[42,50]	[40,42]	[45,49]	[49,50]	[40,42]	[45,49]	[49,50]	[40,49]	[39,50]	[40,49]	[38,50]
Miami	43	[41,43]	[13,43]	[43,43]	[39,42]	[11,18]	[43,44]	[39,42]	[10,18]	[39,43]	[11,44]	[39,44]	[10,44]
Atlanta	44	[44,48]	[42,50]	[44,46]	[47,49]	[45,49]	[44,46]	[47,49]	[44,49]	[43,49]	[38,50]	[43,49]	[37,50]
Port St. Lucie	45	[43,46]	[29,46]	[44,46]	[43,44]	[29,33]	[43,47]	[43,45]	[29,33]	[43,46]	[28,48]	[42,47]	[27,48]
Fresno	46	[43,46]	[22,46]	[44,47]	[43,44]	[21,26]	[44,49]	[43,44]	[20,27]	[43,47]	[21,48]	[43,49]	[20,49]
Tampa	47	[45,47]	[31,47]	[46,49]	[45,47]	[29,33]	[45,50]	[44,47]	[29,34]	[44,49]	[28,49]	[43,50]	[28,50]
Orlando	48	[46,48]	[33,48]	[47,50]	[45,47]	[29,33]	[46,50]	[45,47]	[29,34]	[44,50]	[28,50]	[44,50]	[28,50]
Jacksonville	49	[49,49]	[37,49]	[47,50]	[47,49]	[36,42]	[46,50]	[47,49]	[36,43]	[45,50]	[34,50]	[44,50]	[33,50]
New Orleans	50	[50,50]	[41,50]	[48,50]	[50,50]	[36,46]	[47,50]	[50,50]	[36,47]	[48,50]	[35,50]	[47,50]	[34,50]
Average width		5.6	17.6	3.7	3.8	6.3	4.6	4.7	7.8	8.9	21.8	9.6	22.7

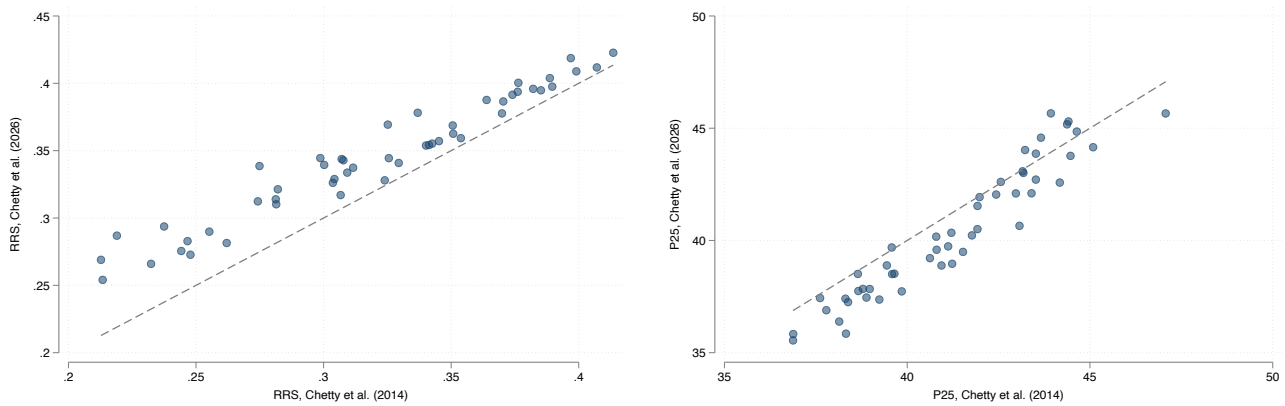
Data: Opportunity Insights.
Note: This table reports normative, statistical, and combined sets for the 50 largest CZs. Details of the construction are discussed in Appendix E. The normative dominance sets $R^{[0,2]}$ and $R^{[0,\infty)}$ are the baseline results reported in Figure 5b. Statistical confidence sets account for sampling variation in the CZ-specific frequency-weight matrices (Equation 11). They are reported for the two limits of inequality-of-opportunity aversion θ and an intermediate level ($\theta = 2$). Following Mogstad *et al.* (2024), we report marginal confidence sets for the rankings of individual CZs and simultaneous confidence sets for the full ranking profile. Combined sets apply the normative dominance evaluation over both $\theta \in [0, 2]$ and $\theta \in [0, \infty)$ to the marginal and simultaneous confidence sets. *Rank* reports the ranking by $V_{0,u}$; rank 1 (50) corresponds to the highest (lowest) $V_{0,u}$.

Table S.10: Dominance sets and sharp sets for the 50 largest CZs

		Dominance sets R_j		Sharp sets S_j	
Commuting zone	Rank	$R^{[0,2]}$	$R^{[0,\infty)}$	$S^{[0,2]}$	$S^{[0,\infty)}$
Minneapolis	1	[1,2]	[1,8]	[1,2]	[1,8]
Salt Lake City	2	[1,2]	[1,2]	[1,2]	[1,2]
Boston	3	[3,3]	[3,9]	[3,3]	[3,9]
Newark	4	[4,7]	[4,17]	[4,7]	[4,17]
Pittsburgh	5	[4,6]	[2,12]	[4,5]	[4,9]
Washington DC	6	[5,7]	[4,15]	[6,7]	[5,14]
Bridgeport	7	[7,8]	[7,24]	[7,8]	[7,24]
Manchester	8	[4,8]	[2,10]	[4,8]	[2,8]
Milwaukee	9	[9,21]	[9,38]	[9,21]	[9,38]
Denver	10	[9,10]	[6,16]	[9,10]	[9,12]
Providence	11	[10,14]	[9,22]	[10,13]	[10,19]
Buffalo	12	[11,16]	[11,30]	[12,15]	[12,28]
Grand Rapids	13	[11,16]	[11,33]	[13,14]	[13,32]
Seattle	14	[10,15]	[5,18]	[10,14]	[7,14]
San Francisco	15	[11,16]	[3,17]	[12,15]	[5,15]
San Jose	16	[10,16]	[2,16]	[10,16]	[2,16]
Philadelphia	17	[17,22]	[16,35]	[17,22]	[17,34]
Kansas City	18	[16,22]	[15,32]	[17,20]	[17,30]
Fort Worth	19	[16,20]	[11,24]	[16,19]	[14,19]
Chicago	20	[20,26]	[19,36]	[20,26]	[20,35]
Houston	21	[17,23]	[9,24]	[18,21]	[11,21]
St. Louis	22	[22,30]	[22,43]	[22,30]	[22,43]
Baltimore	23	[23,31]	[22,41]	[23,31]	[23,40]
Portland	24	[16,24]	[12,29]	[16,24]	[13,24]
Cincinnati	25	[25,33]	[25,49]	[25,33]	[25,49]
Dallas	26	[22,27]	[17,36]	[23,26]	[23,28]
Indianapolis	27	[26,33]	[26,46]	[27,32]	[27,45]
San Diego	28	[18,28]	[3,28]	[18,28]	[3,28]
Austin	29	[24,30]	[18,37]	[24,29]	[22,29]
New York	30	[26,31]	[15,34]	[27,30]	[18,30]
Columbus	31	[31,36]	[30,47]	[31,36]	[31,46]
Cleveland	32	[32,37]	[28,47]	[32,37]	[32,43]
Dayton	33	[31,36]	[27,45]	[32,34]	[32,39]
Sacramento	34	[24,34]	[14,36]	[24,34]	[16,34]
Phoenix	35	[28,35]	[19,40]	[28,35]	[22,35]
Detroit	36	[36,38]	[32,47]	[36,38]	[36,44]
Raleigh	37	[37,42]	[36,48]	[37,42]	[37,47]
Nashville	38	[37,41]	[28,46]	[38,40]	[36,41]
Los Angeles	39	[29,39]	[6,39]	[29,39]	[6,39]
San Antonio	40	[38,41]	[20,42]	[39,41]	[21,41]
Las Vegas	41	[35,41]	[20,43]	[35,41]	[21,41]
Charlotte	42	[42,47]	[42,50]	[42,47]	[42,50]
Miami	43	[41,43]	[13,43]	[41,43]	[13,43]
Atlanta	44	[44,48]	[42,50]	[44,48]	[44,49]
Port St. Lucie	45	[43,46]	[29,46]	[43,45]	[30,45]
Fresno	46	[43,46]	[22,46]	[43,46]	[22,46]
Tampa	47	[45,47]	[31,47]	[45,47]	[31,47]
Orlando	48	[46,48]	[33,48]	[46,48]	[33,48]
Jacksonville	49	[49,49]	[37,49]	[49,49]	[37,49]
New Orleans	50	[50,50]	[41,50]	[50,50]	[41,50]
Average width		5.6	17.6	5.0	14.3

Data: Opportunity Insights.
Note: This table compares the dominance sets $R^{[0,2]}$ and $R^{[0,\infty)}$ reported in Figure 5b with the sharp attainable rank sets $S^{[0,2]}$ and $S^{[0,\infty)}$ for the 50 largest CZs. Dominance sets contain every rank a CZ could hold in any complete ranking consistent with the unambiguous pairwise dominance preorder over the respective range of θ. Sharp sets span only the ranks a CZ actually attains across that range. *Rank* reports the ranking by $V_{0,u}$; rank 1 (50) corresponds to the highest (lowest) $V_{0,u}$.

Figure S.1: Comparison between Chetty *et al.* (2014a) and Chetty *et al.* (2026)



(a) RRS

(b) P25

Data: Opportunity Insights.

Note: This figure compares the data of Chetty *et al.* (2014a) and Chetty *et al.* (2026) for relative and absolute mobility. Panel (a) plots, for each of the 50 largest CZs, the rank-rank slope (RRS) reported by Chetty *et al.* (2014a) against the RRS based on Chetty *et al.* (2026), which is approximated by the slope of the secant through children's mean income rank at the 25th and 75th percentiles of the parental income distribution. Panel (b) plots the mean child rank at the 25th percentile of parental income (P25), based on Chetty *et al.* (2014a) and Chetty *et al.* (2026). Estimates of Chetty *et al.* (2014a) are based on the 1980–1982 birth cohort.

S.2. Normative weights and absolute mobility

Normative weights $p^{\star}_{\theta,u}(f,q;s)$ redistribute attention away from the population average toward types with lower expected utility (Figure 1). This provides a bridge to the intergenerational mobility literature, in which various studies calculate measures of absolute mobility by anchoring their analyses to specific parental income percentiles. For instance, following the influential work of Chetty *et al.* (2014a, 2026), a common benchmark is the expected income of children whose parents sit at the 25th percentile.

To illustrate this bridge, recall the expression for normative weights (Equation (6)):

$$p^{\star}_{\theta,u}(f,q;s) = \frac{q(s)e^{-\theta U(\pi_s)}}{\sum_{s\in\mathcal{S}} q(s)\left(e^{-\theta U(\pi_s)}\right)}. \tag{S.1}$$

Now, suppose that types $s \in \mathcal{S}$ are defined by parental income percentiles $r^p \in \{1, ..., 100\}$, so that $q(s) = 1/100$ for all s, and that expected type utilities are estimated using the linear specification of the income–rank slope of Equation (12).

Then, we can express the normative focus of the social planner by calculating the average parental income rank considered in their evaluation:

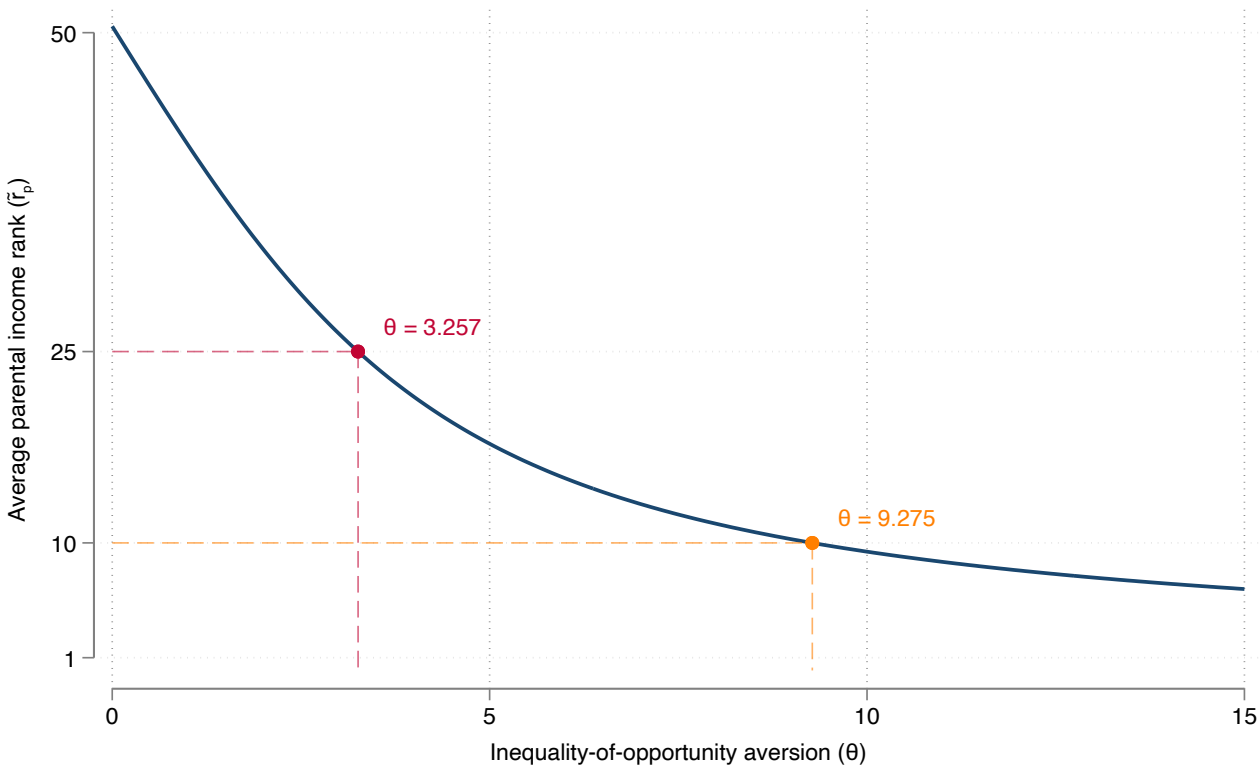
$$\tilde{r}^p = \sum_{s\in\mathcal{S}} p^{\star}_{\theta,u}(f,q;s) \times r^p_s = \sum_{s\in\mathcal{S}} \frac{e^{-\theta(\hat{\beta}^{IRS}\times r^p_s)}}{\sum_{s\in\mathcal{S}}\left(e^{-\theta(\hat{\beta}^{IRS}\times r^p_s)}\right)} \times r^p_s, \tag{S.2}$$

where the intercept $\hat{\alpha}$ cancels as it is uniform across s. Note that $\tilde{r}^p$ is strictly decreasing if and only if $\hat{\beta}^{IRS} > 0$.

There are two polar cases. At $\theta = 0$ the weights are uniform, $p^{\star}_0(s) = q(s)$, so $\tilde{r}^p = 50.5$: The planner weights all parental percentiles equally. As $\theta \to \infty$ the weights concentrate on the type with the lowest utility, which is the lowest percentile if $\hat{\beta}^{IRS} > 0$, so $\tilde{r}^p \to 1$: The planner focuses on the worst-off only. Between the limiting cases, $\tilde{r}^p$ traces out a monotonically decreasing curve based on Equation (S.2).

Figure S.2 illustrates the relationship between θ and $\tilde{r}^p$ using national US data. At the given level of intergenerational mobility in the US ($\hat{\beta}^{IRS} = 0.011$), an average normative focus on the 25th percentile corresponds to $\theta \approx 3.3$, whereas a focus on the 10th percentile requires $\theta \approx 9.3$. Importantly, while measures of absolute mobility focus on a single percentile, $\tilde{r}^p$ provides an average equivalent that allows for weights across neighboring parental income percentiles.

Figure S.2: Mapping parental income ranks and θ



Data: Opportunity Insights.

Note: This figure shows the weighted average parental income rank as a function of θ calibrated on intergenerational mobility at the national level in the US based on our reconstructed dataset ($\hat{\beta}^{RRS} = 0.32$ and $\hat{\beta}^{IRS} = 0.011$). Dashed lines and markers indicate the values of θ at which the planner's average focus equals the 25th and 10th percentiles of parental income, respectively.