SIER Working Paper Series No. 112

Somewhere Between Utopia and Dystopia: Choosing From Multiple Incomparable Prospects

By

Gordon Anderson, Thierry Post, and Yoon-Jae Whang

August, 2018

Institute of Economic Research
Seoul National University

Somewhere Between Utopia and Dystopia: Choosing From Multiple Incomparable Prospects

Gordon Anderson*
Department of Economics, University of Toronto

Thierry Post
Graduate School of Business, Nazarbayev University
and
Yoon-Jae Whang
Department of Economics, Seoul National University

August 17, 2018

Abstract

In many fields of decision making, choices have to be made from multiple alternatives, but stochastic dominance rules do not yield a complete ordering due to incomparability of some or all of the prospects. For ranking incomparable prospects, a 'Utopia Index' measuring the proximity to a lower envelope of integrated distribution functions is proposed. Economic interpretations in terms of Expected Utility are provided for the envelope and deviations from it. The analysis generalizes the existing Almost Stochastic Dominance concept from pairwise comparison to a joint analysis of an arbitrary number of prospects. The limit distribution for the empirical counterpart of the index for a general class of dynamic processes is derived together with a consistent and feasible inference procedure based on subsampling techniques. Empirical applications to Chinese household income data and historical investment returns data show that, in every choice set, a single prospect is ranked above all alternatives at conventional significance levels, despite the incomparability problem.

Keywords: Almost Stochastic Dominance, Convex Stochastic Dominance, Subsampling, Wellbeing Analysis, Portfolio Choice.

*The authors thank the editor, associate editor, and two referees, as well as Heim Levy, Oliver Linton, and the seminar participants at the 2016 INET-Cambridge Workshop on Stochastic Dominance Theory and Applications for helpful comments.

1 Introduction

In a considerable variety of situations, choices have to be made from a collection of risky prospects, wellbeing or other outcome quality distributions. A common difficulty for advisers, fiduciaries and policy makers alike is the inherent ambiguity of the preferences of the relevant individuals. Individually specific preferences are difficult to elicit and vary with culture, personal traits and circumstances.

To address this problem, stochastic dominance (SD) criteria (Hadar and Russell (1969), Hanoch and Levy (1969), Rothschild and Stiglitz (1970), Atkinson (1970, 1987), Foster and Shorrocks (1988), Davies and Hoy (1994, 1995)) provide a partial ordering for general classes of preferences.

Unfortunately, all too frequently prospects cannot be ordered because the relevant dominance relation does not prevail. Known as "incomparability" in mathematical order theory, the decision rule fails to identify a unique superior alternative, and potentially leads to indecision or suboptimal choice.

The present study facilitates choice in such cases by introducing a 'Utopia Index' which measures the proximity to the lower envelope of the integrated distribution functions of all feasible prospects. This index admits the possibility that the best and worst prospects may not be unique, whilst identifying the situation of unique superiority or inferiority when it does occur. The analysis provides a formal economic, mathematical and statistical interpretation of the exploratory and informal applications of empirical envelopes in Anderson and Leo (2016).

In Section 2, the basic concepts are introduced. The analysis combines elements of Convex Stochastic Dominance (CSD; Fishburn (1974); Bawa, Bodurtha, Rao and Suri (1985); Post (2017)) and Almost Stochastic Dominance (ASD; Leshno and Levy (2002); Tzeng, Huang and Shih (2013); Tsetlin, Winkler, Huang and Tzeng (2015)) rules.

Through a joint analysis of all feasible alternatives, CSD generally leads to a larger reduction of the choice set than multiple pairwise comparisons. Using the CSD approach, Bawa, Bodurtha, Rao and Suri (1985) shrink an investment opportunity set by about 30 percent compared with Bawa, Lindenberg and Rafsky (1979), who use multiple pairwise comparisons. Anderson and Post (2018) report similar reductions using real and simulated wellbeing distributions. Nevertheless, CSD is not a panacea, because its 'optimal set' still contains at least two elements, if multiple alternatives are not pairwise dominated.

ASD shrinks the choice set by allowing for limited violations of the SD rule. It compares the area that violates the classic 'no crossing rule' of two integrated distribution functions with the total absolute area between them. This approach has an interesting economic interpretation in terms of a set of utility functions with limited variation of the derivatives, but, since it focuses on pairwise comparison, it does not generate the largest possible set reduction, especially in the presence of many alternatives.

To facilitate statistical inference, Section 3 provides derivations of the limit distribution for the empirical counterpart of the Utopia Index for a general class of dynamic processes and proposes consistent and feasible procedures based on resampling techniques in the spirit of Linton, Maasoumi and Whang (2005), Scaillet and Topaloglou (2010) and Arvanitis, Hallam, Post and Topaloglou (2017); see also Whang (2018). This framework also allows for statistical inference about standard ASD relations which arise as special cases of the analysis.

Section 4 presents Monte Carlo experiments investigating the finite-sample properties of the subsampling confidence intervals and the subsampling tests.

To illustrate the wide applicability of these techniques, examples from very different literatures are presented in Sections 5 and 6. Section 5 provides an analysis of comparative economic wellbeing in China before, during and after the Cultural Revolution. Section 6 considers portfolio choice problems drawn from the asset pricing literature.

To facilitate notation and compact presentation of the new concepts a number of simplifications are employed. First, since in most financial applications the maintained assumption of risk aversion is accepted and inequality aversion is ever present in the range of assumptions employed by welfare analysts, attention is focused upon the common second-degree stochastic dominance rule (SSD). Extensions to other degrees of SD are straightforward by changing the order of integration of the distribution functions.

Second, to clarify the economic interpretation the model is formulated and interpreted in terms of Expected Utility (EU), notwithstanding the known applications of SSD outside the realm of EU.

Third, the number of prospects is assumed to be finite. In wellbeing comparisons, combining policy alternatives or sub-populations is generally not feasible or meaningful. Similarly, in Corporate Finance, investment projects are often indivisible and mutually exclusive. In some other applications, linear combinations of the prospects are possible, for example, forming an investment portfolio of liquid securities. In the latter case, the analysis can be extended along the lines of Post (2003) and the SD optimal set then coincides with his 'SD efficient set'.

The online supplementary appendix contains the formal proofs to the propositions of Sections 2 and 3 together with details about an auxiliary regression analysis for the Chinese wellbeing application of Section 5.

2 Concepts

2.1 Preliminaries

Consider a finite choice set of M distinct prospects $\mathcal{P} := \{P_1, \cdots, P_M\}$, $2 \leq M < \infty$. The prospects have random outcomes $(X_1, \cdots, X_M) \in \mathcal{X}^M$, where $\mathcal{X} = [a, b]$, $-\infty < a < b < +\infty$, is a bounded superset of the maximal support of the prospects. Let $\mathcal{G} := \{G_1, \cdots, G_M\}$, where $G_m(x)$ is the cumulative distribution function (CDF) for prospect P_m, $m = 1, \cdots, M$. The second-order integrated cumulative distribution function (ICDF) is defined as follows:

$$G^{(2)}(x) := \int_a^x G(y)dy,\ G \in \mathcal{G}. \tag{1}$$

The ICDF is a non-negative, non-decreasing and convex function with minimum $G^{(2)}(a) = 0$ and maximum $G^{(2)}(b) = b - \mathbb{E}_G[x]$.

The analysis can easily be generalized to N-th degree SD by substituting $G^{(2)}$ by $G^{(N)}(x) = G(x)$ for $N = 1$ or $G^{(N)}(x) := \int_a^x G^{(N-1)}(y)dy$ for $N > 2$.

To discuss the economic interpretation of the analysis, let $\mathcal{U}_2$ denote the set of non-decreasing and concave utility functions. Since the analysis is location invariant, it is assumed without loss of generality that utility is non-positive ($u(x) \leq 0$), to allow for the elementary ramp function $u(x) = (x - y, 0)$.

The prospects are assumed to be non-equivalent, in the sense that, for every pair of prospects (P_m, P_n), $m \neq n$, at least some permissible utility functions are not indifferent between P_m and P_n.

Individual decision makers face the EU optimization problem $\max_{G \in \mathcal{G}} \mathbb{E}_G[u(x)]$, for a given $u \in \mathcal{U}_2$. Absent complete information about decision-maker preferences, one may use the following binary classification scheme:

Definition 1. Prospect P_m, $m = 1, \cdots, M$, is second-degree *inadmissible* if it is strictly dominated by some alternative prospect by SSD:

$$G_m^{(2)}(x) \geq G^{(2)}(x),\ \forall x \in \mathcal{X}, \tag{2}$$

for some $G \in \mathcal{G}$, with $G_m^{(2)}(x) > G^{(2)}(x)$ for at least some $x \in \mathcal{X}$.

The admissible set has an unambiguous mathematical interpretation as the set of 'maximal elements' of the choice set and the dominance relation. The economic interpretation however is more ambiguous. For $M = 2$, it is well-known that P_n dominates P_m if and only if $\mathbb{E}_{G_m}[u(x)] \leq \mathbb{E}_{G_n}[u(x)]$, for all $u \in \mathcal{U}_2$, with at least one strict inequality for some $u \in \mathcal{U}_2$. The EU interpretation however does not carry over to $M > 2$. In this case, an alternative, tighter classification scheme is more relevant:

Definition 2. Prospect P_m, $m = 1, \cdots, M$, is second-degree *non-optimal* if it is strictly dominated by some mixture distribution by SSD:

$$G_m^{(2)}(x) \geq \sum_{n=1}^{M} \lambda_n G_n^{(2)}(x),\ \forall x \in \mathcal{X}, \tag{3}$$

for some mixing weights $\boldsymbol{\lambda} \in \Lambda := \left\{\boldsymbol{\lambda} \in \mathbb{R}_+^n : \boldsymbol{\lambda}'\mathbf{1}_M = 1\right\}$, with $G_m^{(2)}(x) > \sum_{n=1}^{M} \lambda_n G_n^{(2)}(x)$ for at least some $x \in \mathcal{X}$.

Non-optimality occurs if and only if $\mathbb{E}_{G_m}[u(x)] \leq \max_{G \in \mathcal{G}} \mathbb{E}_G[u(x)]$, for all $u \in \mathcal{U}_2$, with at least one strict inequality for some $u \in \mathcal{U}_2$. All non-optimal prospects are economically redundant and can be removed from the choice set for maximization purposes.

The classification scheme based on optimality leads to a larger reduction of the choice set than that based on admissibility. For $M = 2$, optimality and admissibility are equivalent. However, for $M > 2$, admissibility is a necessary but not sufficient condition for optimality. Nevertheless, the optimality criterion cannot reduce the choice set to a singleton, in case of multiple admissible prospects.

2.2 Utopia

In the quest for decision rules that apply in case of multiple admissible prospects, a third, even tighter classification scheme is introduced:

Definition 3. Prospect P_m, $m = 1, \cdots, M$, is second-degree *utopian* if it weakly dominates all alternative mixture distributions by SSD:

$$G_m^{(2)}(x) \leq \sum_{n=1}^{M} \lambda_n G_n^{(2)}(x),\ \forall x \in \mathcal{X}, \tag{4}$$

for all $\boldsymbol{\lambda} \in \Lambda$.

In terms of order theory, the utopian set consists of the 'greatest elements' of the choice set and the dominance relation, which should not be confused with the set of maximal elements, or the admissible set.

The utopian set coincides with the optimal set if the latter is a singleton; else, it is empty. In many relevant applications, multiple prospects are optimal and thus no prospect is utopian. The following construct is introduced as a benchmark for measuring deviations from utopianity:

Definition 4. The second-order *lower envelope* of $\mathcal{G}$ amounts to

$$\underline{\mathcal{G}}_2(x) := \min_{G \in \mathcal{G}} G^{(2)}(x),\ x \in \mathcal{X}. \tag{5}$$

The lower envelope is a non-negative, non-decreasing and piece-wise convex function with minimum $\underline{\mathcal{G}}_2(a) = 0$ and maximum $\underline{\mathcal{G}}_2(b) = b - \max_{G \in \mathcal{G}} \mathbb{E}_G[x]$ (using $G^{(2)}(b) = b - \mathbb{E}_G[x]$ for all $G \in \mathcal{G}$).

The lower envelope consists exclusively of segments of second-order ICDFs of optimal prospects. The envelope is therefore invariant to the exclusion of non-optimal prospects. The ICDF of any admissible but non-optimal prospect will not intersect with the envelope.

By construction, the lower envelope is the tightest envelope that supports from below all ICDF mixtures. For all outcome levels $x \in \mathcal{X}$,

$$\exists G \in \mathcal{G} : G^{(2)}(x) = \underline{\mathcal{G}}_2(x); \tag{6}$$

$$\sum_{n=1}^{M} \lambda_n G_n^{(2)}(x) \geq \underline{\mathcal{G}}_2(x), \ \forall \boldsymbol{\lambda} \in \Lambda. \tag{7}$$

If one of the prospects is utopian, then its second-order ICDF will coincide with the lower envelope. If the utopian set is empty, it seems tempting to view the lower envelope as the second-order ICDF of an infeasible utopian prospect, or an order-theoretical 'upper bound'. That interpretation is not correct, however, as the lower envelope is piece-wise convex, rather than convex, and hence not a second-order ICDF.

Instead the lower envelope is interpreted in terms of EU levels for elementary utility functions. Specifically, $\mathcal{U}_2$ is a convex cone in the space of functions and its extreme rays are given by elementary Russell and Seo (1989) ramp functions:

$$v_y(x) := \min(x - y, 0). \tag{8}$$

These singularity functions play a key role in the SD literature, because a dominance relation occurs if and only the relevant preference relation exists for all these functions.

Proposition 1. *The lower envelope mirrors the maximum expected value of the elementary utility function $v_y(x)$ as a function of the threshold level $y \in \mathcal{X}$:*

$$\underline{\mathcal{G}}_2(y) = -\max_{G \in \mathcal{G}} \mathbb{E}_G[v_y(x)], \ \forall y \in \mathcal{X}. \tag{9}$$

Since all permissible utility functions $u \in \mathcal{U}_2$ are positive linear mixtures of the elementary functions, without loss of generality, attention may be restricted to the following set of convex mixtures:

$$\mathcal{V}_2 := \{u_w(x) : w \in \mathcal{W}\}; \tag{10}$$

$$u_w(x) := \int_a^b w(y) v_y(x) dy,\ w \in \mathcal{W}; \tag{11}$$

$$\mathcal{W} := \left\{ w : \mathcal{X} \to [0,1] : \int_a^b w(y) dy = 1 \right\}. \tag{12}$$

In this formulation, the mixing function $w : \mathcal{X} \to [0,1]$ controls the local risk aversion or poverty aversion of the mixed utility function, with $w(y) = 0$ representing local linearity.

Replacing $\mathcal{U}_2$ with $\mathcal{V}_2$ amounts to using harmless positive linear transformations of utility functions as well as ignoring the shape of the utility functions outside the support superset $\mathcal{X} = [a, b]$. Economic interpretations of the definitions of admissibility, optimality and utopianity are invariant to this replacement.

The circle can now be completed by providing an economic interpretation of utopianity:

Proposition 2. *Prospect P_m, $m = 1, \cdots, M$, is second-degree utopian if and only if it is the best solution for all utility functions $v \in \mathcal{V}_2$*

$$\mathbb{E}_{G_m}[u_w(x)] = \max_{G \in \mathcal{G}} \mathbb{E}_G[u_w(x)],\ \forall w \in \mathcal{W}. \tag{13}$$

An economic interpretation can also be provided for violations of utopianity, using a subset of admissible utility functions. For $\mathcal{Y} \subseteq \mathcal{X}$, let

$$\mathcal{V}_2^*(\mathcal{Y}) := \{u_w \in \mathcal{V}_2 : w(y) = 0\ \forall y \in \mathcal{Y}\}; \tag{14}$$

$$\mathcal{X}_{2,m} := \left\{ x \in \mathcal{X} : G_m^{(2)}(x) > \underline{G}_2(x) \right\}. \tag{15}$$

The set $\mathcal{V}_2^*(\mathcal{Y})$ consists of utility functions which are locally linear ($w(y) = 0$) on $\mathcal{Y}$. This set is reminiscent of the piecewise-linear utility functions used in Post (2003, Theorem 2) to model empirical SD relations. The 'violation region' $\mathcal{X}_{2,m}$ consists of the outcomes levels where the ICDF of the evaluated prospect lies above the lower envelope. For utopian prospects, $\mathcal{X}_{2,m} = \emptyset$; for non-optimal prospects, $\mathcal{X}_{2,m} = \mathcal{X}$.

Proposition 3. *Given the violation region $\mathcal{X}_{2,m}$ for prospect P_m, $m = 1, ..., M$, and the usual mean condition* $\mathbb{E}_{G_m}[x] = \max_{n \in \mathcal{P}} \mathbb{E}_{G_n}[x]$,

$$\mathbb{E}_{G_m}[u_w(x)] = \max_{n \in \mathcal{P}} \mathbb{E}_{G_n}[u_w(x)] \ \forall u_w \in \mathcal{V}_2^*(\mathcal{X}_{2,m}). \tag{16}$$

In other words, the evaluated prospect is the best solution for all normalized permissible utility functions which are locally linear in the violation region $\mathcal{X}_{2,m}$. Due to the aforementioned invariance properties, the utility function class $\mathcal{V}_2^*(\mathcal{X}_{2,m})$ may be replaced with its non-normalized equivalent based on $\mathcal{U}_2$, without harm.

On a small interval, any standard utility function can be accurately approximated with a linear line segment. Hence, approximate optimality for standard utility functions can be established in case of minor violations of utopianity. By contrast, for a large violation region, local linearity tends to gives a poor approximation for standard, strictly concave functions and approximate optimality cannot be established.

2.3 Dystopia

Analogous to the definition of utopianity, the 'least element' of the choice set can also be introduced:

Definition 5. Prospect P_m, $m = 1, \cdots, M$, is second-degree *dystopian* if it is weakly dominated by all alternative mixture distributions by SSD:

$$G_m^{(2)}(x) \geq \sum_{n=1}^{M} \lambda_n G_n^{(2)}(x), \ \forall x \in \mathcal{X}, \tag{17}$$

for all $\boldsymbol{\lambda} \in \Lambda$.

Definition 6. The second-order *upper envelope* of $\mathcal{G}$ amounts to

$$\overline{\mathcal{G}}_2(x) := \max_{G \in \mathcal{G}} G^{(2)}(x), \ x \in \mathcal{X}. \tag{18}$$

Like the lower envelope, the upper envelope is a non-negative, non-decreasing and piecewise convex function with minimum $\overline{\mathcal{G}}_2(a) = 0$. However, $\overline{\mathcal{G}}_2(x)$ envelopes $\mathcal{G}$ from above and it achieves a maximum of $\overline{\mathcal{G}}_2(b) = b - \min_{G \in \mathcal{G}} \mathbb{E}_G[x]$.

In a similar fashion to utopianity, the set of prospects that are second-degree dystopian form an inferior set. If only one prospect is dystopian, that is to say the inferior set is a singleton, then its second-order ICDF will coincide with the upper envelope. If no prospect is dystopian, the upper envelope is the tightest envelope that supports from above all second-order ICDF mixtures: for all outcome levels $x \in \mathcal{X}$,

$$\exists G \in \mathcal{G} : G^{(2)}(x) = \overline{\mathcal{G}}_2(x); \tag{19}$$

$$\sum_{n=1}^{M} \lambda_n G_n^{(2)}(x) \leq \overline{\mathcal{G}}_2(x),\ \forall \boldsymbol{\lambda} \in \Lambda. \tag{20}$$

Proposition 4. *The upper envelope mirrors the minimum expected value of the elementary utility function $v_y(x)$ as a function of the threshold level $y \in \mathcal{X}$:*

$$\overline{\mathcal{G}}_2(y) = -\min_{G \in \mathcal{G}} \mathbb{E}_G[v_y(x)],\ \forall y \in \mathcal{X}. \tag{21}$$

Proposition 5. *Prospect P_m, $m = 1, \cdots, M$, is second-degree dystopian if and only if it is the worst solution for all utility functions $v \in \mathcal{V}_2$*

$$\mathbb{E}_{G_m}[u_w(x)] = \min_{G \in \mathcal{G}} \mathbb{E}_G[u_w(x)],\ \forall w \in \mathcal{W}. \tag{22}$$

2.4 Utopia index

A ranking of incomparable prospects may be obtained by comparing individual distributions with the lower envelope. For this purpose, the following measure for violations of utopianity, in the spirit of ASD, is introduced:

Definition 7. The second-order *violation area* for utopianity of prospect P_m, $m = 1, \cdots, M$, amounts to

$$\mathcal{A}_{2,m} := \int_{\mathcal{X}} \left(G_m^{(2)}(x) - \underline{\mathcal{G}}_2(x) \right) dx. \tag{23}$$

This measure captures the deviations of the prospect from the optima for the elementary utility functions. Specifically, it follows from the results in Section 2.2 that

$$\mathcal{A}_{2,m} = \int_{\mathcal{X}} \left(\max_{G \in \mathcal{G}} \mathbb{E}_G[v_y(x)] - \mathbb{E}_{G_m}[v_y(x)] \right) dy. \tag{24}$$

If $\mathcal{A}_{2,m} = 0$, then the prospect is utopian and the unique element of the optimal set. If $\mathcal{A}_{2,m} > 0$, then the prospect is not utopian. The violation area increases with the fraction of the utility functions for which the prospect is suboptimal and the shortfall in EU for these functions.

For normalizing the violation area, the following between-prospect variation measure is introduced:

Definition 8. The second-order *transvariation* of $\mathcal{G}$ amounts to

$$\mathcal{T}_2 := \int_{\mathcal{X}} \left(\bar{\mathcal{G}}_2(x) - \underline{\mathcal{G}}_2(x) \right) dx. \tag{25}$$

This measure extends the existing notions of bi-variate transvariation (Gini (1916, 1959) and multivariate transvariation (Dagum (1968)) to a higher order of integration of distributions.

Whereas the envelopes $\underline{\mathcal{G}}_2$ and $\bar{\mathcal{G}}_2$ capture best-case and worst-case solutions, the transvariation $\mathcal{T}_2$ captures the between-prospect variation in economic goodness for the elementary utility functions:

$$\mathcal{T}_2 = \int_{\mathcal{X}} \left(\max_{G \in \mathcal{G}} \mathbb{E}_G[v_y(x)] - \min_{G \in \mathcal{G}} \mathbb{E}_G[v_y(x)] \right) dy. \tag{26}$$

Since the prospects are non-equivalent, $\mathcal{T}_2 > 0$. By construction, the transvariation is an upper bound for the violation area: $\mathcal{T}_2 \geq \mathcal{A}_{2,m}$, for any $m = 1, \cdots, M$. Motivated by this observation, the following normalized measure for utopianity is proposed:

Definition 9. The second-degree *Utopia Index* for prospect P_m, $m = 1, \cdots, M$, amounts to

$$\mathcal{I}_{2,m} := 1 - \frac{\mathcal{A}_{2,m}}{\mathcal{T}_2}. \tag{27}$$

The Utopia Index $\mathcal{I}_{2,m}$ can readily be shown to satisfy the Continuity, Scale Invariance, Scale Independence and Normalization Axioms familiar in the Inequality Measurement Literature (Shorrocks (1978), Kobos and Milos (2012)). Since it is completely defined by the marginal distribution functions, the index is also Law Invariant.

Since $0 \leq G_m^{(2)}(x) \leq \bar{\mathcal{G}}_2(x)$, for all $m = 1, \cdots, M$, it follows that $0 \leq \mathcal{A}_{2,m} \leq \mathcal{T}_2$. Consequently, $0 \leq \mathcal{I}_{2,m} \leq 1$; $\mathcal{I}_{2,m} = 0$ amounts to dystopianity; $\mathcal{I}_{2,m} = 1$ amounts to utopianity. If $\mathcal{I}_{2,m} < 1$ for all $m = 1, \cdots, M$, then the utopian set is empty and there exist multiple optimal prospects. In this case, the prospect with the highest Utopia Index comes closest to being a unique optimum.

In contrast to inequality-adjusted wellbeing indices (see, for example, Blackorby and Donaldson (1978)) and certainty equivalents, the Utopia Index considers all feasible outcome distributions and all permissible preferences.

Classical examples of inequality-adjusted indices are Atkinson's Index (Atkinson (1983)) and Sen's index (Sen (1987)). Both modify μ, average income in a given distribution, by a factor 'I' which accommodates aversion to inequality in the form: $\mu(1 - I)$. In the case of Atkinson's index, I is based on the Hölder generalized mean; in the case of Sen's index, I corresponds to some power of the Gini coefficient. In contrast to the Utopia Index, these indices are based on a single given wellbeing distribution and a single given inequality aversion level.

One way to interpret the Utopia Index is in terms of sub-indices for individual utility functions:

$$\mathcal{I}_{2,m}(v) := \frac{\mathbb{E}_{G_m}[v(x)] - \min_{G \in \mathcal{G}} \mathbb{E}_G[v(x)]}{\max_{G \in \mathcal{G}} \mathbb{E}_G[v(x)] - \min_{G \in \mathcal{G}} \mathbb{E}_G[v(x)]}, \quad v \in \mathcal{V}_2. \tag{28}$$

It follows from (24) and (26) that the index takes a value between the lowest and highest values which are feasible for individual permissible utility functions:

$$\min_{v \in \mathcal{V}_2} \mathcal{I}_{2,m}(v) \leq \mathcal{I}_{2,m} \leq \max_{v \in \mathcal{V}_2} \mathcal{I}_{2,m}(v). \tag{29}$$

By searching over all permissible preferences, the Utopia Index can identify whether the optimal and inferior sets are singletons and the corresponding prospects are uniquely utopian and dystopian, respectively. Since infinitely many preferences are permitted, this is not possible using a finite collection of sub-indices for individual utility functions.

Changing the choice set by including or excluding prospects generally changes the lower envelope $\underline{\mathcal{G}}_2(x)$ and transvariation $\mathcal{T}_2$ and, consequently, changes the values of the index of the prospects in absolute and relative terms. The mutual ranking of a given subset of prospects is however robust to expansion or contraction of the choice set, because the violation area $\mathcal{A}_{2,m}$ of every prospect changes by the same amount if the lower envelope is shifted. Consequently, $\mathcal{I}_{2,m} \geq \mathcal{I}_{2,n}$ for a given pair of prospects in a given choice set implies $\mathcal{I}_{2,m} \geq \mathcal{I}_{2,n}$ in any superset of the two prospects.

2.5 Special case: ASD

The Utopia Index resembles the 'epsilon' measure of ASD, in the sense that it captures the normalized violation area of utopianity, just as 'epsilon' captures the normalized violation area of pairwise dominance. Indeed, in the case of pairwise analysis ($M = 2$), the proposed framework is equivalent to ASD. Specifically, in this case, 'epsilon' amounts to the anti-index $(1 - \mathcal{I}_{2,m}) = \frac{\mathcal{A}_{2,m}}{\mathcal{T}_2}$.

ASD is typically interpreted in terms of the following subset of admissible utility functions:

$$\mathcal{U}_2^*(\epsilon) := \left\{ u \in \mathcal{U}_2 : \mathcal{C}_u \leq \left(\frac{1}{\epsilon} - 1 \right) \right\}; \tag{30}$$

$$\mathcal{C}_u := \sup\left(-u''(x)\right) / \inf\left(-u''(x)\right). \tag{31}$$

In case of pairwise comparison ($M = 2$), P_1 dominates P_2 by ASD given epsilon value $0 < \epsilon \leq 1/2$ if and only if $\mathbb{E}_{G_1}[u(x)] \geq \mathbb{E}_{G_2}[u(x)]$ for all $u \in \mathcal{U}_2^*(\epsilon)$; see Lesnho and Levy (2002) and Tzeng, Huang and Shih (2013).

It is possible to preserve the interpretation in terms of $\mathcal{U}_2^*(\epsilon)$ in the case of $M > 2$ by making multiple pairwise ASD comparisons. Specifically, $\mathbb{E}_{G_m}[u(x)] = \sup_{n \in \mathcal{P}} \mathbb{E}_{G_n}[u(x)]$ for all $u \in \mathcal{U}_2^*(\epsilon)$ if and only if P_m dominates all alternatives by ASD at an epsilon value of ϵ.

Although our statistical inference procedure allows for pairwise implementation of utopianity tests, it should be pointed out that a series of pairwise tests is less statistically efficient than a joint test for utopianity relative to all prospects.

In addition, the economic interpretation of $\mathcal{U}_2^*(\epsilon)$ seems debatable, because this utility set is biased towards constant risk aversion ($-u''(x) = c$) and does not account for the natural variation of $u''(x)$ of standard utility functions such as the power and the exponential.

As a case in point, the quadratic utility function $u(x) = x - (2b)^{-1} x^2$ features $\mathcal{C}_u = 1$ and is therefore included in $\mathcal{U}_2^*(\epsilon)$ for every $0 < \epsilon \leq 1/2$, despite Increasing Absolute Risk Aversion and exploding Relative Risk Aversion: $\lim_{x \uparrow b} (-u''(x)x/u'(x)) = \infty$.

3 Statistical Inference

3.1 Introduction

In practice, the CDFs $G_m(x)$, $m = 1, \cdots, M$, are latent and the analyst has access to a discrete series of realized outcomes. Let $\mathcal{S}_m := \{X_{m,t} : t = 1, \ldots, T_m\}$ be the observed data from $G_m(x)$. Our analysis considers two types of sampling schemes that typically arise in applications: (i) *Type I (Nonsynchronized) Sampling*: mutually independent samples of T_m independent observations; (ii) *Type II (Synchronized) Sampling*: samples of $T_m = T$ independent or weakly dependent observations

Type I sampling is appropriate in situations where separate random samples are drawn from nonoverlapping populations such as countries or regions, or when random samples are drawn at different points in time for the same population.

Type II sampling is appropriate if the prospects $X_{1,t}, ..., X_{M,t}$ represent measures of the same welfare variable at different points in time (for example, before-tax and after-tax incomes), or different measures of welfare for an individual or a country (for example, income and leisure) at a given time.

Alternatively, the prospects may represent returns of different portfolios at a given time t. In the latter case, it makes sense to assume that observations are (weakly) dependent over t. In any case, unlike the Type I sampling, it is plausible to assume that the observations $X_{1,t}, ..., X_{M,t}$ are mutually dependent in an unknown fashion at a given t.

This section discusses a statistical inference procedure for the Utopia Index $\mathcal{I}_{2,m}$ using the following estimator:

$$\hat{\mathcal{I}}_{2,m} := 1 - \frac{\hat{\mathcal{A}}_{2,m}}{\hat{\mathcal{T}}_2},$$

where $\hat{\mathcal{A}}_{2,m}$ and $\hat{\mathcal{T}}_2$, respectively, are the empirical violation area and empirical transvariation which are based on the empirical equivalent of the ICDF $G_m^{(2)}(x)$:

$$\hat{G}_m^{(2)}(x) := \frac{1}{T_m} \sum_{t=1}^{T_m} (x - X_{m,t}) 1\left(X_{m,t} \leq x\right), \; m = 1, ..., M.$$

The analysis can be extended to the N-th order Utopia Index $\mathcal{I}_{N,m}$ for $N \geq 1$ in a straightforward fashion.

3.2 Asymptotic distribution

The analysis also requires some additional notation. Let

$$\begin{aligned}
\Delta_{i,j}(x) &:= G_i^{(2)}(x) - G_j^{(2)}(x), 1 \leq i, j \leq M; \\
\Delta_{m,max}(x) &:= \max_{1 \leq j \leq M} \Delta_{m,j}(x); \\
\Delta_{max}(x) &:= \max_{1 \leq i \neq j \leq M} \Delta_{i,j}(x).
\end{aligned}$$

These definitions can be used to reformulate the violation area and transvariation as follows:

$$\begin{aligned}
\mathcal{A}_{2,m} &= \int_{\mathcal{X}} \left[G_m^{(2)}(x) - \underline{\mathcal{G}}_2(x)\right] dx = \int_{\mathcal{X}} \Delta_{m,max}(x) dx; \\
\mathcal{T}_2 &= \int_{\mathcal{X}} \left[\overline{\mathcal{G}}_2(x) - \underline{\mathcal{G}}_2(x)\right] dx = \int_{\mathcal{X}} \Delta_{max}(x) dx.
\end{aligned}$$

The analysis also employs r-enlargements of 'contact sets':

$$B_{m,j}(r) := \left\{x \in \mathcal{X} : |\Delta_{m,j}(x) - \Delta_{m,max}(x)| \leq r\right\}; \tag{32}$$

$$\tilde{B}_{i,j}(r) := \left\{x \in \mathcal{X} : |\Delta_{i,j}(x) - \Delta_{max}(x)| \leq r\right\}. \tag{33}$$

For any subset $A \subset \mathcal{X}$, let $A^0 := A$ and $A^1 := \mathcal{X} \backslash A$. Define the following index sets which consist of vectors of 0 or 1 excluding the unit vector:

$$\mathcal{J}{:=}\{(i_1, ..., i_M) \in \{0,1\}^M : \ i_j = 0 \text{ for some } j\}; \tag{34}$$

$$\mathcal{K}{:=}\{(k_{1,2}, k_{1,3}, ..., k_{M,(M-1)}) \in \{0,1\}^{M(M-1)} : k_{i,j} = 0 \text{ for some } (i,j)\}. \tag{35}$$

For $\mathbf{i} \in \mathcal{J}$, $\mathbf{k} \in \mathcal{K}$, and $0 < r < s$, let

$$A^*_{m,\mathbf{i}}(r,s) := \left(\bigcap_{\{j:i_j=0\}} B^{i_j}_{m,j}(r) \right) \bigcap \left(\bigcap_{\{j:i_j=1\}} B^{i_j}_{m,j}(s) \right); \tag{36}$$

$$\tilde{A}^*_{\mathbf{k}}(r,s) := \left(\bigcap_{\{(i,j):i\neq j, k_{i,j}=0\}} \tilde{B}^{k_{i,j}}_{i,j}(r) \right) \bigcap \left(\bigcap_{\{(i,j):i\neq j, k_{i,j}=1\}} \tilde{B}^{k_{i,j}}_{i,j}(s) \right). \tag{37}$$

It is assumed that the observed data satisfy one of the following assumptions:

Assumption 1-1 (Type I sampling): For each $m = 1, ..., M$, (i) $\{X_{m,t} : t = 1, \ldots, T_m\}$ is an i.i.d. sequence; (ii) the union of supports of $X_{m,t}$, $m = 1, ..., M$ is $\mathcal{X} = [a,b]$, $-\infty < a < b < \infty$, and the distribution of $X_{m,t}$ is absolutely continuous with respect to the Lebesgue measure and has bounded density; (iii) the sampling scheme is such that as $T_m \to \infty$, $T_m/T \to \lambda_m \in (0,1)$, where $T = \sum_{m=1}^M T_m$; (iv) $\{X_{m,t} : t = 1, \ldots, T_m\}$, $m = 1, ..., M$ are mutually independent.

Assumption 1-2 (Type II sampling): For each $m = 1, ..., M$, (i) $\{X_{m,t} : t = 1, \ldots, T\}$ is a strictly stationary and α-mixing sequence with $\alpha(s) = O(s^{-A})$ for some $A > (q-1)(1+q/2)$, where q is an even integer that satisfies $q > 4$; (ii) Assumption 1-1 (ii) holds.

To construct a confidence interval with an asymptotically correct coverage error, an additional assumption is introduced. The assumption involves regularity on the r-enlargement of the contact sets defined in (32) and (33), which plays an important role in the asymptotic approximation; see the proof of Proposition 6 in Appendix A.

Assumption 2: (i) There exists a positive non-random sequence $\{c_T : T \geq 1\}$ that satisfies $c_T \to 0$ and $\sqrt{T} c_T \to \infty$ as $T \to \infty$; (ii) for each decreasing sequence $r_T \to 0$ and

$\varepsilon \in (0,1)$,

$$\begin{aligned} \text{Leb}\left(B_{m,j}((1-\varepsilon)c_T)\backslash B_{m,j}(T^{-1/2}r_T)\right) &\leq r_{T_j}\ \forall j\ ; \\ \text{Leb}\left(\tilde{B}_{i,j}((1-\varepsilon)c_T)\backslash \tilde{B}_{i,j}(T^{-1/2}r_T)\right) &\leq r_T\ \forall i \neq j \end{aligned}$$

for $T > 1/\varepsilon$, where Leb denotes the Lebesgue measure.

Let $\nu := \{\nu_{i,j} : i,j = 1,...,M\}$, where $\nu_{i,j}$ is a zero-mean Gaussian process with covariance kernel

$$C_{i,j}(x,y) = \lim_{T\to\infty} T \cdot \text{Cov}\left[\left(\hat{G}_i^{(2)}(x) - \hat{G}_j^{(2)}(x)\right), \left(\hat{G}_i^{(2)}(y) - \hat{G}_j^{(2)}(y)\right)\right],$$

for $x, y \in \mathbb{R}$. For subsets $\mathcal{C} :=\{\mathcal{C}_{\mathbf{i}} \subset \mathcal{X} : \mathbf{i} \in \mathcal{J}\}$ and $\mathcal{D} :=\{\mathcal{D}_{\mathbf{k}} \subset \mathcal{X} : \mathbf{k} \in \mathcal{K}\}$, define

$$\begin{aligned} \Lambda_{\mathcal{C},\mathcal{D}}(\nu) := &-\frac{1}{\mathcal{T}_2}\sum_{\mathbf{i}\in\mathcal{J}}\int_{\mathcal{C}_{\mathbf{i}}} \max_{j\in\mathcal{M}(\mathbf{i})}\left\{\lambda_j^{-1/2}\nu_{m,j}(x)\right\} dx \\ &+ \frac{\mathcal{A}_{2,m}}{\mathcal{T}_2^2}\sum_{\mathbf{k}\in\mathcal{K}}\int_{\mathcal{D}_{\mathbf{k}}} \max_{(i,j)\in\tilde{\mathcal{M}}(\mathbf{k})}\left\{\lambda_j^{-1/2}\nu_{i,j}(x)\right\} dx, \end{aligned} \tag{38}$$

where $\mathcal{M}(\mathbf{i}) = \{j \in \{1,...,M\} : i_j = 0\}$ and $\tilde{\mathcal{M}}(\mathbf{k}) = \{(i,j) \in \{1,...,M\}^2 : \mathbf{k}_{i,j} = 0, i \neq j\}$.

To control the asymptotic coverage error of the confidence interval described below, the following regularity condition for the Gaussian process ν is introduced (see Linton, Song and Whang (2010) for a related condition):

Definition 10. A Gaussian process ν is *regular* on $(\mathcal{C},\mathcal{D}) \subset \mathcal{X}\times\mathcal{X}$ if for any $\alpha \in (0,1/2]$, there exists $\bar{\varepsilon} > 0$ depending only on α such that

$$P\left\{\Lambda_{\mathcal{C},\mathcal{D}}(\nu) < \bar{\varepsilon}\right\} < 1-\alpha \tag{39}$$

and, for any $c > 0$,

$$\text{limsup}_{\eta\downarrow 0} P\left\{|\Lambda_{\mathcal{C},\mathcal{D}}(\nu) - c| \leq \eta\right\} = 0. \tag{40}$$

The following result establishes the asymptotic approximation of the estimator of the Utopia Index.

Proposition 6. *Suppose that Assumptions 1-1 (or 1-2) and 2 hold. Then, (i)*

$$\sqrt{T}\left(\hat{\mathcal{I}}_{2,m}-\mathcal{I}_{2,m}\right)=\Lambda_{\mathcal{A}_m^*,\tilde{\mathcal{A}}^*}(\nu)+o_p(1),$$

where $\Lambda_{a,b}(\nu)$ *is as defined in (38) (with* $\lambda_j = 1$ *under Assumption 1-2),* $\mathcal{A}_m^* := \{A_{m,\mathbf{i}}^*\left(T^{-1/2}r_T, (1-\varepsilon)c_T\right) : \mathbf{i} \in \mathcal{J}\}$ *and* $\tilde{\mathcal{A}}^* := \{\tilde{A}_{\mathbf{k}}^*\left(T^{-1/2}r_T, (1-\varepsilon)c_T\right) : \mathbf{k} \in \mathcal{K}\}$*; (ii)* $\sqrt{T}\left(\hat{\mathcal{I}}_{2,m}-\mathcal{I}_{2,m}\right) \to_p 0$ *if* $\inf_{x\in\mathcal{X}}\left[G_m(x)-\max_{j\neq m}\{G_j(x)\}\right] > \bar{\eta}$ *or* $\inf_{x\in\mathcal{X}}\left[\min_{j\neq m}\{G_j(x)\}-G_m(x)\right] > \bar{\eta}$ *for some* $\bar{\eta} > 0$.

Proposition 6 implies that the asymptotic distribution of $\sqrt{T}\left(\hat{\mathcal{I}}_{2,m}-\mathcal{I}_{2,m}\right)$ is a functional of the Gaussian process ν, which is is non-degenerate if ν is regular on $\mathcal{A}_m^*$ and $\tilde{\mathcal{A}}^*$. It also shows that the asymptotic distribution may degenerate to zero if G_m corresponds to the upper envelope $\underline{\mathcal{G}}_2$ or the lower envelope $\overline{\mathcal{G}}_2$, sufficiently far away from the other distribution functions.

In general, the asymptotic distribution depends on the latent set of distributions $\mathcal{G}$, so the critical values can not be tabulated and need to be estimated by a simulation or resampling procedure.

3.3 Subsampling approximation

This paper considers a subsampling method (Politis and Romano (1994), Politis, Romano and Wolf (1999)) to approximate the asymptotic distribution of the estimator $\hat{\mathcal{I}}_{2,m}$.

The proposed subsampling procedure for the Type I data is based on the following steps:

(i) Calculate the test statistic $\hat{\mathcal{I}}_{2,m}$ using the original full sample $\mathcal{W}_T := \{X_{m,t} : t = 1,\ldots,T_m; m = 1,...,M\}$.

(ii) Generate subsamples $\mathcal{W}_{b,m,i}$, $i = 1,...,L(T)$, of size b_m from $\{X_{m,t} : t = 1,\ldots,T_m\}$ for $m = 1,...,M$, with $L(T) := \min\{L(T_1),...,L(T_M)\}$ and $L(T_m) := \binom{T_m}{b_m}$.

(iii) Compute test statistics $\hat{\mathcal{I}}_{m,b,i}$ using the subsamples $\{\mathcal{W}_{b,m,i} : m = 1,...,M\}$ for $i = 1,...,L(T)$.

(iv) Approximate the subsampling distribution

$$\widehat{S}_{T,b}(w) := \frac{1}{L(T)} \sum_{i=1}^{L(T)} 1\left(\sqrt{b} \mid \hat{\mathcal{I}}_{m,b,i} - \hat{\mathcal{I}}_{2,m} \mid \leq w\right)$$

with $b = \sum_{m=1}^{M} b_m$.

In practice, it can be computationally too costly to consider all $\binom{T_m}{b_m}$ subsets if T_m is large. In this case, the usual practice is to generate a fixed moderate number, say $L(T_m) = 1000$, of random samples (without replacement) of size b_m from $\{X_{m,t} : t = 1, \ldots, T_m\}$.

Similarly, for the Type II data, the following steps can be used:

(i)* Calculate the test statistic $\hat{\mathcal{I}}_{2,m}$ using the original full sample $\mathcal{W}_T := \{X_{m,t} : t = 1, \ldots, T; m = 1, ..., M\}$.

(ii)* Generate subsamples $\mathcal{W}_{b,i} := \{Y_i, \ldots, Y_{i+b-1}\}$, $i = 1, \ldots, L(T)$, of size b from $\mathcal{W}_T$ with $Y_i = (X_{1,i}, ..., X_{M,i})$ and $L(T) = T - b + 1$.

(iii)* Compute test statistics $\hat{\mathcal{I}}_{m,b,i}$ using the subsamples $\mathcal{W}_{b,i}$ for $i = 1, \ldots, L(T)$.

(iv)* Approximate the subsampling distribution

$$\widehat{S}_{T,b}(w) := \frac{1}{L(T)} \sum_{i=1}^{L(T)} 1\left(\sqrt{b} \mid \hat{\mathcal{I}}_{m,b,i} - \hat{\mathcal{I}}_{2,m} \mid \leq w\right).$$

Let $s_{T,b}(1-\alpha) := \inf\{w : \widehat{S}_{T,b}(w) \geq 1 - \alpha\}$ be the $(1-\alpha)$ sample quantile of $\widehat{S}_{T,b}(\cdot)$ for $\alpha \in (0, 1/2]$. As shown in Proposition 6, there are some cases in which $\sqrt{T}\left(\hat{\mathcal{I}}_{2,m} - \mathcal{I}_{2,m}\right)$ degenerates to zero. In this case, the subsampling distribution $\widehat{S}_{T,b}(\cdot)$ also degenerates to zero and therefore it is hard to ensure that $s_{T,b}(1-\alpha)$ converges to zero more slowly than $\sqrt{T}\left(\hat{\mathcal{I}}_{2,m} - \mathcal{I}_{2,m}\right)$. To avoid the technical difficulty, the critical value is defined as $c_{T,b,\eta}(1-\alpha) = \max\{s_{T,b}(1-\alpha), \eta\}$, where $\eta > 0$ is an arbitrary small fixed constant, say $\eta = 10^{-6}$; see Lee, Song and Whang (2018) for a similar idea in a different context.

The $(1-\alpha)$ level confidence interval for $\mathcal{I}_{2,m}$ is defined as follows:

$$CI_{T,b} := \left\{\mathcal{I} \in [0,1] : \hat{\mathcal{I}}_{2,m} - \frac{c_{T,b,\eta}\,(1-\alpha)}{\sqrt{T}} \leq \mathcal{I} \leq \hat{\mathcal{I}}_{2,m} + \frac{c_{T,b,\eta}\,(1-\alpha)}{\sqrt{T}}\right\}. \tag{41}$$

The following proposition shows that the subsampling confidence interval has asymptotically correct coverage.

Proposition 7. *Suppose that Assumptions 1-1 (or 1-2) and 2 hold. If* $b_m \to \infty$*,* $b_m/b \to \lambda_m$*, and* $b/T \to 0$ *under Assumption 1-1 (or* $b \to \infty$ *and* $b/T \to 0$ *under Assumption 1-2), then*

$$\lim_{T\to\infty} \Pr\left(\mathcal{I}_{2,m} \in CI_{T,b}\right) \geq (1-\alpha),$$

with equality holding if ν *is regular on* $\mathcal{A}_m^*$ *and* $\tilde{\mathcal{A}}^*$*.*

The result of Proposition 7 can be used to test a hypothesis on the Utopia Index by inverting the subsampling confidence interval. In some cases, however, one might be interested in testing the hypothesis directly based on a test statistic. Of particular interest are the following hypotheses $H_0^A : \mathcal{I}_{2,m} = 1$ vs. $H_1^A : \mathcal{I}_{2,m} < 1$, $H_0^B : \mathcal{I}_{2,m} = 0$ vs. $H_1^B : \mathcal{I}_{2,m} > 0$, and $H_0^C : \mathcal{I}_{2,m} = \mathcal{I}_{2,n}$ vs. $H_1^C : \mathcal{I}_{2,m} \neq \mathcal{I}_{2,n}$ for $m \neq n$.

As test statistics for the hypotheses H_0^A, H_0^B, and H_0^C, one may consider the Wald-type statistics $W_T^A = \sqrt{T}\left(\hat{\mathcal{I}}_{2,m} - 1\right)$, $W_T^B = \sqrt{T}\hat{\mathcal{I}}_{2,m}$, and $W_T^C = \sqrt{T}\left|\,\hat{\mathcal{I}}_{2,m} - \hat{\mathcal{I}}_{2,n}\,\right|$, respectively. Then, one may reject H_0^A, H_0^B, and H_0^C if $W_T^A < \min\{\bar{s}_{T,b}(\alpha), -\eta\}$, $W_T^B > \max\{\bar{s}_{T,b}(1-\alpha), \eta\}$, and $W_T^C > \max\{\tilde{s}_{T,b}(1-\alpha), \eta\}$, respectively, where $\bar{s}_{T,b}(\gamma)$ and $\tilde{s}_{T,b}(\gamma)$ denote the $\gamma \in (0, 1/2]$ sample quantiles of

$$\bar{S}_{T,b}(w) := \frac{1}{L(T)} \sum_{i=1}^{L(T)} 1\left(\sqrt{b}\left(\hat{\mathcal{I}}_{m,b,i} - \hat{\mathcal{I}}_{2,m}\right) \leq w\right);$$

$$\tilde{S}_{T,b}(w) := \frac{1}{L(T)} \sum_{i=1}^{L(T)} 1\left(\sqrt{b}\,\left|\,\left(\hat{\mathcal{I}}_{m,b,i} - \hat{\mathcal{I}}_{n,b,i}\right) - \left(\hat{\mathcal{I}}_{2,m} - \hat{\mathcal{I}}_{2,n}\right)\,\right| \leq w\right),$$

respectively.

Likewise, the subsampling p-values for the tests W_T^A, W_T^B, and W_T^C can be computed by

$$
\begin{aligned}
p_A &:= \frac{1}{L(T)} \sum_{i=1}^{L(T)} 1\left(\min\left\{\sqrt{b}\left(\hat{\mathcal{I}}_{m,b,i} - \hat{\mathcal{I}}_{2,m}\right), -\eta\right\} \leq W_T^A\right); \\
p_B &:= \frac{1}{L(T)} \sum_{i=1}^{L(T)} 1\left(\max\left\{\sqrt{b}\left(\hat{\mathcal{I}}_{m,b,i} - \hat{\mathcal{I}}_{2,m}\right), \eta\right\} \geq W_T^B\right); \qquad (42) \\
p_C &:= \frac{1}{L(T)} \sum_{i=1}^{L(T)} 1\left(\max\left\{\sqrt{b}\,\left|\,\left(\hat{\mathcal{I}}_{m,b,i} - \hat{\mathcal{I}}_{n,b,i}\right) - \left(\hat{\mathcal{I}}_{2,m} - \hat{\mathcal{I}}_{2,n}\right)\right|, \eta\right\} \geq \left|\, W_T^C \,\right|\right),
\end{aligned}
$$

respectively.

The following proposition shows that the tests have asymptotically correct size and are consistent against the alternatives:

Proposition 8. *Suppose that Assumptions 1-1 (or 1-2) and 2 hold. If $b_m \to \infty$, $b_m/b \to \lambda_m$, and $b/T \to 0$ under Assumption 1-1 (or $b \to \infty$ and $b/T \to 0$ under Assumption 1-2). Then, for a=A, B, and C, we have*

$$
\begin{aligned}
&\text{(i)} \lim_{T\to\infty} \Pr(\textit{Reject}\, H_0^a) \leq \alpha \quad \textit{under}\, H_0^a; \\
&\text{(ii)} \lim_{T\to\infty} \Pr(\textit{Reject}\, H_0^a) = 1 \quad \textit{under}\, H_1^a.
\end{aligned}
$$

4 Monte Carlo Experiments

This section investigates the finite sample properties of the subsampling confidence intervals and the subsampling tests using simulation experiments.

In the experiments, T observations of X_1, X_2, X_3, X_4, X_5 and X_6 were generated independently from $U(0, 0.75)$, $U(0, 0.8)$, $U(0, 1)$, $U(0, 1)$, $U(0.2, 1)$, and $U(0.25, 1)$, respectively ($M = 6$). The true Utopia Index values are given by $\mathcal{I}_{2,1} = 0$, $\mathcal{I}_{2,2} = 0.0967$, $\mathcal{I}_{2,3} = \mathcal{I}_{2,4} = 0.4167$, $\mathcal{I}_{2,5} = 0.8967$, and $\mathcal{I}_{2,6} = 1$.

The sample size is $T \in \{100, 500, 2000, 8000, 32000\}$. The subsample size was selected using the fixed rule $b = \lfloor T^\gamma \rfloor$ for $\gamma \in \{0.70, 0.75, 0.80\}$. The number of subsamples and the number of simulation repetitions were fixed to be 1000. Furthermore, $\eta = 10^{-6}$ was used, but the simulation results were not sensitive to this choice.

Table 1: Coverage probabilities of subsampling 95% confidence intervals

γ	T	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=6$
0.70	100	0.953	0.859	0.860	0.853	0.825	0.979
	500	0.994	0.881	0.913	0.892	0.859	1.000
	2000	1.000	0.874	0.927	0.928	0.909	1.000
	8000	1.000	0.924	0.932	0.942	0.937	1.000
	32000	1.000	0.916	0.942	0.953	0.946	1.000
0.75	100	0.959	0.844	0.839	0.841	0.841	0.985
	500	0.993	0.833	0.915	0.894	0.855	1.000
	2000	1.000	0.883	0.924	0.927	0.899	1.000
	8000	1.000	0.907	0.928	0.945	0.926	1.000
	32000	1.000	0.943	0.938	0.940	0.928	1.000
0.80	100	0.946	0.848	0.820	0.854	0.836	0.984
	500	0.997	0.846	0.891	0.888	0.847	1.000
	2000	1.000	0.863	0.898	0.897	0.877	1.000
	8000	1.000	0.900	0.930	0.920	0.927	1.000
	32000	1.000	0.937	0.933	0.933	0.937	1.000

Table 1 presents the empirical coverage probabilities of the subsampling confidence intervals at level 0.95. As expected, for $m \in \{2,3,4,5\}$, the simulated probabilities generally tend to the nominal confidence level as the sample size increases. On the other hand, for $m \in \{1,6\}$, the subsampling confidence intervals are conservative and the coverage probabilities tend to 1 rather quickly as the sample size increases. These instances correspond to the 'boundary' or 'non-regular' case where the asymptotic distributions degenerate to zero, as discussed in Propositions 6 (ii) and 7.

Table 2 gives the rejection probabilities of the subsampling tests. For brevity, the results with $\gamma = 0.75$ and $T \in \{100, 500, 2000\}$ are reported; other results are analogous. For each $m \neq n$, the proposed subsampling tests were applied to the following 4 hypotheses: $H_0^1 : \mathcal{I}_{2,m} = 0$ vs. $H_1^1 : \mathcal{I}_{2,m} > 0$, $H_0^2 : \mathcal{I}_{2,m} = 1$ vs. $H_1^2 : \mathcal{I}_{2,m} < 1$, $H_0^3 : \mathcal{I}_{2,m} = \mathcal{I}_{2,m}^0$ vs. $H_1^3 : \mathcal{I}_{2,m} \neq \mathcal{I}_{2,m}^0$ with $\mathcal{I}_{2,m}^0 = 0.4167$, and $H_0^4 : \mathcal{I}_{2,m} = \mathcal{I}_{2,n}$ vs. $H_1^4 : \mathcal{I}_{2,m} \neq \mathcal{I}_{2,n}$. The analysis examines both the size and power performances of the subsampling tests.

As the sample size increases, the results show that the Type I errors quickly go to zero when $\mathcal{I}_{2,m} = 0$ or 1, while they tend to the nominal significance level 0.05 when $\mathcal{I}_{2,m} = \mathcal{I}_{2,m}^0$ or $\mathcal{I}_{2,m} = \mathcal{I}_{2,n}$. This is expected from Proposition 8, because the former corresponds to the boundary case where the asymptotic distributions degenerate to zero, while the latter case has a regular Gaussian process ν, leading to a test with an asymptotically exact size.

On the other hand, the rejection probabilities increase as the distributions deviate from the null hypotheses. They tend to 1 under the fixed alternative hypotheses as the sample size increases, confirming consistency of the proposed tests.

5 Wellbeing of Chinese Households

Evolving government policy after the Communist Revolution changed fundamentally the nature of Chinese families and the extent to which household wellbeing is reflective of such change seems of interest.

At the outset, around 1950, households were classified into ordered social classes ('*chéng fèn*') according to employment status, income source, and political loyalty. Farmland was redistributed from landlord to landless peasant and class labels became a primary criteria for individual advancement.

Later, the Cultural Revolution (CR), an 'equalizing' movement designed to curtail intergenerational transfer of social and educational advantage by elites, saw school closures at all levels. When higher education institutions reopened, young people from lower classes were given educational and occupational preference.

The end of the CR saw renewed educational recruitment on merit, a profound growth spurt precipitated by Economic Reform and the introduction of the One Child Policy which ultimately increased investment in child education (Anderson and Leo (2009)). The extent to which these events affected household wellbeing in terms of their a vintage-equivalent, adult-equivalised income distribution is at question.

The CR affected the average number of years and quality of schooling experienced by birth cohort. Those born between 1948-1955 possibly missed senior high school and those born between 1956-1963 missed part of primary school and junior high school.

Thus, three household types (vintages) can be distinguished by the era in which their head of household was educated ($M = 3$). 'Pre-CR household' heads (older than 51) completed their education before the CR, an educationally immobile era with low levels of education and equality of opportunity. 'CR household' heads (age between 39 and 51) educational opportunities were equalized by suspending education for most. 'Post-CR

Table 2: Rejection probabilities of tests with 5% nominal level

T	H_0	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=6$
100	$\mathcal{I}_{2,m}=0$	0.041	0.348	0.957	0.946	1.000	1.000
	$\mathcal{I}_{2,m}=1$	1.000	1.000	1.000	1.000	0.457	0.015
	$\mathcal{I}_{2,m}=\mathcal{I}^0_{2,m}$	0.995	0.884	0.161	0.159	1.000	1.000
	$\mathcal{I}_{2,m}=\mathcal{I}_{2,n}$						
	$n=1$	-	0.223	0.868	0.878	1.000	1.000
	$n=2$	0.223	-	0.687	0.676	1.000	1.000
	$n=3$	0.868	0.687	-	0.142	0.972	0.998
	$n=4$	0.878	0.676	0.142	-	0.965	0.997
	$n=5$	1.000	1.000	0.972	0.965	-	0.317
	$n=6$	1.000	1.000	0.998	0.997	0.317	-
500	$\mathcal{I}_{2,m}=0$	0.007	0.530	1.000	1.000	1.000	1.000
	$\mathcal{I}_{2,m}=1$	1.000	1.000	1.000	1.000	0.817	0.000
	$\mathcal{I}_{2,m}=\mathcal{I}^0_{2,m}$	1.000	1.000	0.087	0.106	1.000	1.000
	$\mathcal{I}_{2,m}=\mathcal{I}_{2,n}$						
	$n=1$	-	0.415	1.000	1.000	1.000	1.000
	$n=2$	0.415	-	0.998	0.996	1.000	1.000
	$n=3$	1.000	0.998	-	0.085	1.000	1.000
	$n=4$	1.000	0.996	0.085	-	1.000	1.000
	$n=5$	1.000	1.000	1.000	1.000	-	0.741
	$n=6$	1.000	1.000	1.000	1.000	0.741	-
2000	$\mathcal{I}_{2,m}=0$	0.000	0.912	1.000	1.000	1.000	1.000
	$\mathcal{I}_{2,m}=1$	1.000	1.000	1.000	1.000	0.997	0.000
	$\mathcal{I}_{2,m}=\mathcal{I}^0_{2,m}$	1.000	1.000	0.076	0.073	1.000	1.000
	$\mathcal{I}_{2,m}=\mathcal{I}_{2,n}$						
	$n=1$	-	0.852	1.000	1.000	1.000	1.000
	$n=2$	0.852	-	1.000	1.000	1.000	1.000
	$n=3$	1.000	1.000	-	0.062	1.000	1.000
	$n=4$	1.000	1.000	0.062	-	1.000	1.000
	$n=5$	1.000	1.000	1.000	1.000	-	0.994
	$n=6$	1.000	1.000	1.000	1.000	0.994	-

household' heads (age less than 39), completed their education after the CR and were unaffected by it in an era with more investment in education for larger portions of the population.

Direct comparison of the household income distributions of the three household types is difficult, because households of identical circumstances but different vintages would have different incomes due to the life cycle (Friedman (1957), Ando and Modigliani (1963), Modigliani (1976)).

Using a sample of 6,822 urban households from The Chinese Household Income Project (Li, Luo, Wei and Yue (2008)), vintage-equivalent wellbeing distributions for the 3 household types were generated (see Appendix B for details). The Pre-CR cohort contained 2,400 households, the CR 3,118 and the Post-CR 1,306 households. Household Wellbeing was based upon total adult-equivalised income and measured as the log of total household income deflated by the square root of household size (Brady and Barber (1952)). Income is measured in Chinese yuan and refers to the year of 2002.

This survey is the only available data set to hand linking households through the child-parent-grand parent relationship to the social class designation accorded grandparents at the time of the Communist Revolution.

The three income distributions are analyzed using the foregoing classification schemes and inference methods. Figure 1 and Figure 2 show kernel estimates of the probability density functions and, according with Daltonian social preferences, empirical ICDF $\hat{G}_m^{(2)}$ for each of the three cohorts, $m = 1, 2, 3$. The sample range is $\widehat{\mathcal{X}} = [8.38, 11.41]$. Visual inspection of Figure 2 indicates that the income distribution of the Pre-CR cohort is second-degree dystopian. The other two cohorts are incomparable, because the income distribution of the CR cohort is more equal than that of the Post-CR cohort. The optimal set thus consists of the CR and Post-CR cohorts and neither cohort is utopian. The Utopia Index facilitates ranking of these incomparable cohorts.

To illustrate the construction of the index, Figure 3 shows the deviations from the lower envelope, or $\hat{G}_m^{(2)} - \underline{\hat{\mathcal{G}}}_2$, $m = 1, 2, 3$, which are used to compute the transvariation $\hat{\mathcal{T}}_2$ and violation areas $\hat{\mathcal{A}}_{2,m}$, $m = 1, 2, 3$. The transvariation $\hat{\mathcal{T}}_2 = 0.3087$ is the total area below the supremum of the three cohorts, in this case the graph of the Pre-CR cohort. In the

Figure 1: Income distributions of three cohorts

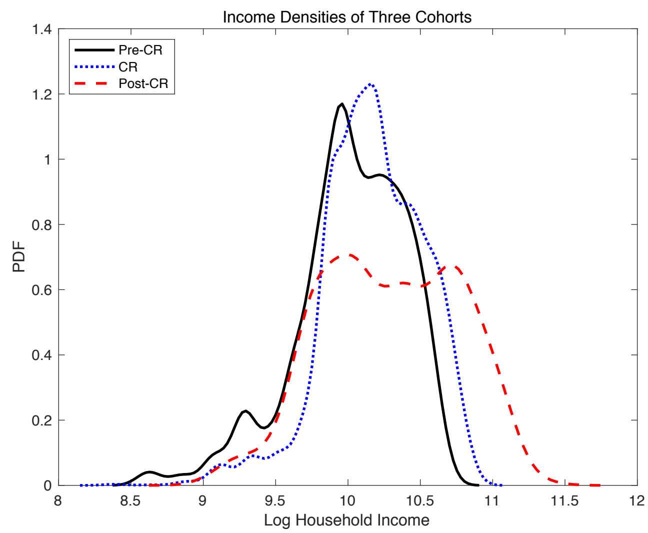


Figure 2: Empirical ICDFs of three cohorts

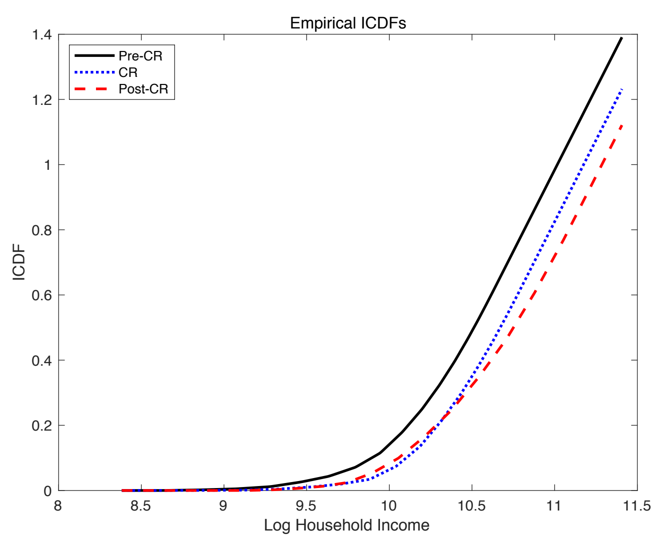
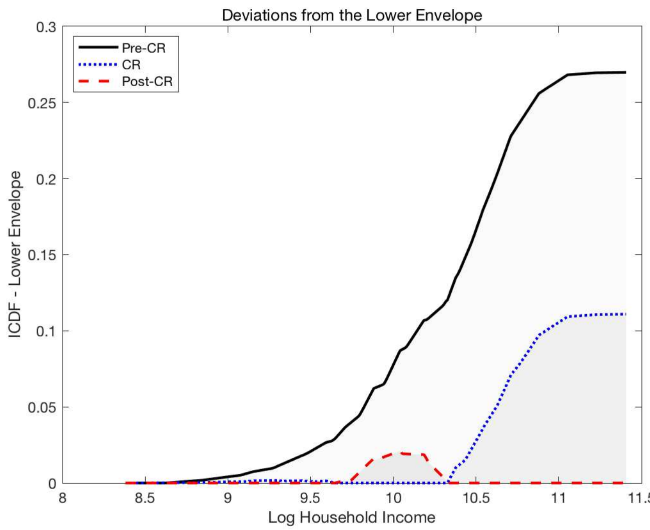


Figure 3: Deviations from the lower envelope



figure, this area is marked in light gray color. The violation area $\widehat{\mathcal{A}}_{2,m}$ for a given cohort is the total area below the graph for that cohort. Clearly, the violation area is smallest for the Post-CR cohort; $\hat{\mathcal{A}}_{2,\text{Post-CR}} = 0.0082$, corresponding to the dark gray area in Figure 3. The Utopia Index normalizes the violation area by the transvariation: $\hat{\mathcal{I}}_{2,\text{Post-CR}} = 0.9734$. The CR cohort has a lower score $\hat{\mathcal{I}}_{2,\text{CR}} = 0.7262$ and the Pre CR cohort has the lowest score $\hat{\mathcal{I}}_{2,\text{Pre-CR}} = 0.0001$.

Table 3 summarizes various statistical inference results. Throughout the empirical applications, the results were not very sensitive to the choice of the tuning parameters, and we fixed $\gamma = 0.75$ for subsample size $b_m = \lfloor T_m^{\gamma} \rfloor$, $m = 1, 2, 3$ and $\eta = 10^{-6}$.

The subsampling standard errors for $\hat{\mathcal{I}}_{2,\text{Pre-CR}}$, $\hat{\mathcal{I}}_{2,\text{CR}}$, and $\hat{\mathcal{I}}_{2,\text{Post-CR}}$ are 0.0002, 0.0350, and 0.0134, respectively, implying a high level accuracy. The subsampling 95% confidence intervals $\left[CI^L_{0.95},\, CI^H_{0.95}\right]$ and the p-values of various hypothesis tests are also provided.

It is evident from the reported results that households whose head was educated in the Post-CR era have a wellbeing distribution that is very close to utopian, a finding consistent with the dominant judgment about the socioeconomic significance of the historical events.

Since FSD plays an important role in welfare economics in addition to SSD, the analysis was repeated for FSD. The FSD results at the bottom of Table 3 further support the aforementioned conclusions about the rankings of the three cohorts. A noteworthy differ-

Table 3: Statistical inference results

	m = Pre-CR	m = CR	m = Post-CR
Utopia Index $\hat{\mathcal{I}}_{2,\mathrm{m}}$	0.0001	0.7262	0.9734
Rank	3	2	1
Standard Error	0.0002	0.0350	0.0134
$CI^L_{0.95}$	0.0000	0.6552	0.9454
$CI^H_{0.95}$	0.0003	0.7973	1.0000
p-value $\mathcal{I}_{2,m} = 0$	0.2510	0.0000	0.0000
p-value $\mathcal{I}_{2,m} = 1$	0.0000	0.0000	0.0690
p-value $\mathcal{I}_{2,m} = \mathcal{I}_{2,n}$			
n = Pre-CR	-	0.0000	0.0000
n = CR	0.0000	-	0.0000
n = Post-CR	0.0000	0.0000	-
Utopia Index $\hat{\mathcal{I}}_{1,\mathrm{m}}$	0.0027	0.5413	0.9176
Rank	3	2	1
Standard Error	0.0031	0.0245	0.0171

ence is the substantial decrease in the relative wellbeing measure for the CR cohort: under first-order comparison, the index falls from 0.7262 (0.0350) to 0.5413 (0.0245), reflecting that first-order comparison does not reward the intense equalizing measures invoked during the CR era.

6 Active Investment Strategies

The second application analyzes the investment returns to two distinct sets of stock benchmark portfolios from the empirical asset pricing literature.

The first benchmark set consists of six active portfolios ($M = 6$) of US common stocks that are formed, and periodically re-balanced, based on the market capitalization of equity ('size') and book-to-market equity ratio ('valuation'). The six portfolios are labeled as Small Growth (SG), Small Blend (SB), Small Value (SV), Large Growth (LG), Large Blend (LB) and Large Value (LV). These portfolios are of particular interest, because a wealth of empirical research, starting with Banz (1981) and Basu (1983), suggests that the low historical average returns to SG stocks and high average for SV stocks defy rational explanations based on investment risk.

The second benchmark set consist of six portfolios that are based on market capitalization and recent past return: Small Loser (SL), Small Neutral (SN), Small Winner (SW), Large Loser (LL), Large Neutral (LN) and Large Winner (LW). Past return is measured using a one-month lagged trailing window of 11 months, to avoid the short-term reversal effect (Jegadeesh (1990)) and exploit the intermediate-term momentum effect (Jegadeesh and Titman (1993)). The historical average return to SL stocks seems abnormally low and that of SW stocks appears exceptionally high.

Choice from these benchmark portfolios is considered without allowing for portfolio mixtures, since many active money managers specialize on security selection for a given market segment or investment style in order to exploit economies of scale and specialization.

Monthly, percentage, value-weighted portfolio returns from January 1927 to August 2016 ($T = 1,076$) from Kenneth French's data library are analyzed. The sample range is $\widehat{\mathcal{X}} = [-35.14, 83.55]$ for the first data set and $\widehat{\mathcal{X}} = [-39.34, 92.17]$ for the second data set.

It is assumed that the investor possesses no conditioning information and focuses on the unconditional distribution. The use of the empirical distribution as a CDF estimator is justified by the long time series ($T = 1,076$), low return frequency, small number of benchmark portfolios ($M = 6$) and the careful stock selection and portfolio construction rules used in building the benchmark portfolios.

Bali, Demirtas, Levy and Wolf (2009) and Bali, Brown and Demirtas (2013) analyze similar data sets using pairwise ASD comparisons. The proposed Utopia Index framework allows for a joint analysis of a cross-section of benchmark portfolios and consistent statistical inference.

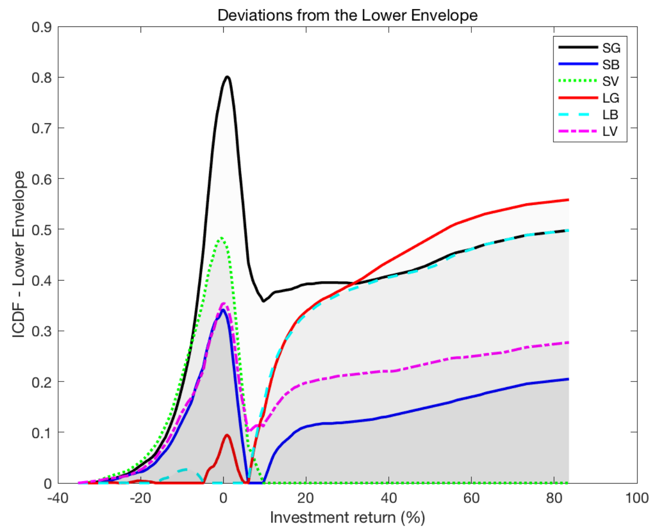
6.1 Portfolios formed on size and valuation

In the first data set, four out of six portfolios (SB, SV, LG, LB) are classified as second-dergee optimal. Thus, the optimality criterion leads to indecision in this case.

Figure 4 shows the deviations from the second-order lower envelope for each of the six portfolios. The violation area is smallest for the SV portfolio-the dark gray area in the figure. Investing in SV stocks thus comes closest to being the optimal investment style for all risk-averse investors.

The Utopia Index for the SV portfolio takes the value $\hat{\mathcal{I}}_{2,\text{SV}} = 0.8502$; SV stocks are

Figure 4: Deviations from the lower envelope for portfolios based on size and valuation



ranked above each of the other market segments at conventional significance levels. By contrast, investing in SG stocks comes closest to being dystopian; $\hat{\mathcal{I}}_{2,\mathrm{SG}} = 0.0612$; the null of dystopianity cannot be rejected for this market segment. The Utopia Index thus confirms the notion that SV stocks offer a return premium over SG stocks.

Table 4 shows further details about the test results. It includes the subsampling confidence intervals and the p-values for the hypotheses of dystopianity and utopianity, for each portfolio, and the hypothesis that two portfolios have the same score, for every pair of portfolios. Significant differences occur between the SG, SB and SV portfolios, consistent with the existing evidence about the book-to-market effect being concentrated in the small-cap stock market segment.

6.2 Portfolios formed on size and past return

In the second data set, the optimality criterion again leads to indecision. Four of the six portfolios (SN, SW, LN, LW) are optimal in the sample.

Figure 5 shows the deviations from the lower envelope for each of the portfolios. The violation area is smallest for the SW portfolio and largest for the LL portfolio. The Utopia Index therefore identifies the former as the most appealing investment style and the latter as the least appealing: $\hat{\mathcal{I}}_{2,\mathrm{SW}} = 0.9477$ and $\hat{\mathcal{I}}_{2,\mathrm{LL}} = 0.0702$. The violation area is relatively

Table 4: Inference results for portfolios based on size and valuation

	$m = SG$	$m = SB$	$m = SV$	$m = LG$	$m = LB$	$m = LV$
Utopia Index $\hat{\mathcal{I}}_{2,\mathrm{m}}$	0.0612	0.6765	0.8502	0.2660	0.3304	0.5215
Rank	6	2	1	5	4	3
Standard Error	0.0833	0.0631	0.0559	0.0917	0.0623	0.0842
$CI^{L}_{0.95}$	0.0000	0.5550	0.7504	0.0614	0.1648	0.3529
$CI^{H}_{0.95}$	0.2716	0.7981	0.9501	0.4706	0.4960	0.6900
p-value $\mathcal{I}_{2,m} = 0$	0.4139	0.0000	0.0000	0.0302	0.0000	0.0000
p-value $\mathcal{I}_{2,m} = 1$	0.0000	0.0000	0.0268	0.0000	0.0000	0.0000
p-value $\mathcal{I}_{2,m} = \mathcal{I}_{2,n}$						
$n = SG$	-	0.0000	0.0000	0.1633	0.0190	0.0000
$n = SB$	0.0000	-	0.0145	0.0000	0.0000	0.0000
$n = SV$	0.0000	0.0145	-	0.0000	0.0000	0.0000
$n = LG$	0.1633	0.0000	0.0000	-	0.4351	0.0123
$n = LB$	0.0190	0.0000	0.0000	0.4351	-	0.0034
$n = LV$	0.0000	0.0000	0.0000	0.0123	0.0034	-

small and SW stocks are even closer to being utopian than SV stocks in Section 6.1. The top ranking of the SW strategy is statistically significant at every conventional significance level. By contrast, for the SL and LL strategies, the null of dystopianity cannot be rejected. These results are again consistent with existing conclusions in the Investments literature.

Table 5 contains more detailed test results. Among other things, it shows that the momentum effect is strong in both the small-cap segment and the large-cap segment. Large winner stocks therefore appear utopian within the universe of large-cap stocks.

7 Concluding Remarks

The applications illustrate the insights obtained from using the Utopia Index. In both applications, the choice sets contain a single prospect which ranks above all alternatives at conventional significance levels, despite the incomparability problem.

In the analysis of Chinese household income distributions, that unique prospect is the Post-CR educated cohort, consistent with the dominant views about the socioeconomic developments in Chinese modern history.

In the Investments application, the unique prospect is the Small Value stock portfolio in one investment universe and the Small Winner portfolio in another universe, consistent

Figure 5: Deviations from the lower envelope for portfolios based on size and past return

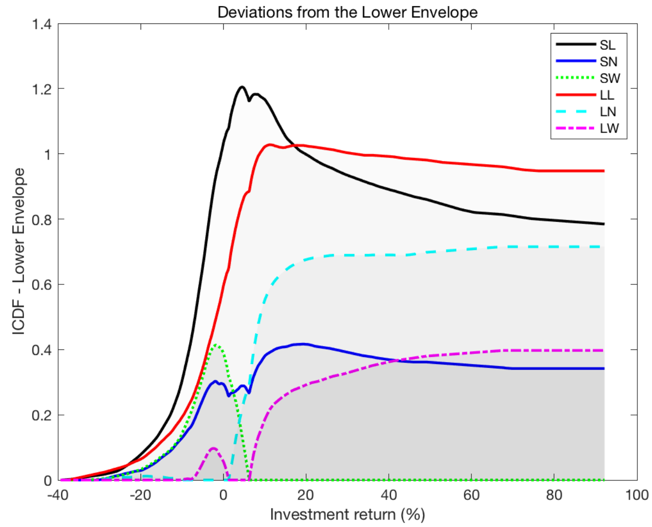


Table 5: Inference results for portfolios based on size and past return

	$m = SL$	$m = SN$	$m = SW$	$m = LL$	$m = LN$	$m = LW$
Utopia Index $\hat{\mathcal{I}}_{2,\mathrm{m}}$	0.0866	0.6418	0.9477	0.0702	0.4116	0.7004
Rank	5	3	1	6	4	2
Standard Error	0.0658	0.0288	0.0117	0.0759	0.0529	0.0559
$CI^L_{0.95}$	0.0000	0.5872	0.9290	0.0000	0.3167	0.5942
$CI^H_{0.95}$	0.2429	0.6965	0.9664	0.2916	0.5065	0.8066
p-value $\mathcal{I}_{2,m} = 0$	0.2193	0.0000	0.0000	0.1879	0.0000	0.0000
p-value $\mathcal{I}_{2,m} = 1$	0.0000	0.0000	0.0067	0.0000	0.0000	0.0000
p-value $\mathcal{I}_{2,m} = \mathcal{I}_{2,n}$						
$n = SL$	-	0.0000	0.0000	0.8920	0.0000	0.0000
$n = SN$	0.0000	-	0.0000	0.0000	0.0000	0.3926
$n = SW$	0.0000	0.0000	-	0.0000	0.0000	0.0000
$n = LL$	0.8920	0.0000	0.0000	-	0.0000	0.0000
$n = LN$	0.0000	0.0000	0.0000	0.0000	-	0.0000
$n = LW$	0.0000	0.3926	0.0000	0.0000	0.0000	-

with the conventional wisdom among investment practitioners and numerous studies in the empirical asset pricing literature.

In the wellbeing application, the Pre-CR cohort is ranked below all alternatives and, moreover, dystopianity cannot be rejected; by contrast, in the financial application, two investments portfolios are tied for the bottom rank, in both investment universes.

Our test procedure is feasible and consistent for a general class of dynamic processes. The procedure can also be applied for statistical inference about standard ASD relations which arise as the special case of pairwise comparison ($M = 2$). The applications in this study involve a favorable combination of a small number of prospects ($M = 3$ cohorts of households or $M = 6$ stock portfolios) and a large number of observations ($T > 1,000$ individual households or investment returns). To deal with less favorable data dimensions, further research could focus on developing small-sample bias-correction methods and consistent bootstrap methods.

References

[1] Anderson, GJ and TW Leo, 2009, Child Poverty and Generational Mobility in Urban China, Review of Income and Wealth 55, 607-629.

[2] Anderson GJ and TW Leo, 2016, On Providing a Complete Ordering of Non-Combinable Alternative Prospects, Mimeo University of Toronto Discussion Paper.

[3] Anderson GJ and GT Post, 2018, On Shrinking the Planner's Choice Set in Multiple Wellbeing Comparisons, forthcoming in Social Choice and Welfare.

[4] Ando, A. and F. Modigliani, 1963, The 'life-cycle' hypothesis of saving: aggregate implications and tests, American Economic Review 53, 55–84.

[5] Andrews, DWK, 2000, Inconsistency of the Bootstrap When a Parameter is on the Boundary of the Parameter Space, Econometrica 68, 399-405.

[6] Andrews, DWK and P. Guggenberger, 2009, Validity of subsampling and plug-in asymptotic inference for parameters defined by moment inequalities, Econometric Theory 25, 669-709.

[7] Arvanitis, S. and M. Hallam and T. Post, 2017, Stochastic Spanning, forthcoming in Journal of Business and Economic Statistics.

[8] Atkinson, AB, 1983, Social Justice and Public Policy, MIT press.

[9] Atkinson, AB, 1970, On the measurement of inequality, J. Econ. Theory 2, 244-263.

[10] Atkinson, AB, 1987, On the Measurement of Poverty, Econometrica 55, 749-764.

[11] Bali, TG, KO Demirtas, H Levy and A Wolf, 2009, Bonds versus stocks: Investors' age and risk taking, Journal of Monetary Economics 56, 817-830.

[12] Bali, TG, SJ Brown, and KO Demirtas, 2013, Do Hedge Funds Outperform Stocks and Bonds?, Management Science 59, 1887-1903.

[13] Banz, RW, 1981, The relationship between return and market value of common stocks, Journal of Financial Economics 9, 3-18.

[14] Basu, S., 1983, The relationship between earnings' yield, market value and return for NYSE common stocks: Further evidence, Journal of Financial Economics 12, 129-156.

[15] Bawa, VS, JN Bodurtha Jr., MR Rao and HL Suri, 1985, On Determination of Stochastic Dominance Optimal Sets, Journal of Finance 40, 417–431.

[16] Bawa, VS, EB Lindenberg and LC Rafsky, 1979, An Efficient Algorithm to Determine Stochastic Dominance Admissible Sets, Management Science 25, 609-622.

[17] Blackorby, C. and D. Donaldson. 1978. Measures of Relative Equality and Their Meaning in Terms of Social Welfare, Journal of Economic Theory 18, 59-80.

[18] Brady, DS and HA Barber, 1948, The Pattern of Food Expenditures, Review of Economics and Statistics 30, 198–206.

[19] Chernozhukov, V. and I. Fernandez-Val, 2005, Subsampling inference on quantile regression processes, Sankhya 67, 253-276.

[20] Chernozhukov, V. and C. Hansen, 2005, An IV model of quantile treatment effects, Econometrica 73, 245-261.

[21] Cho, YH, O. Linton, and YJ Whang, 2007, Are there Monday Effects in Stock Returns?: A Stochastic Dominance Approach, Journal of Empirical Finance 14, 736-755.

[22] Dagum, C., 1968, Multivariate transvariation theory among several distributions and its economic applications, Econometric Research Program, Princeton University, Research Memorandum 100.

[23] Davies, J. and M. Hoy, 1994, The normative significance of using third-degree stochastic dominance in comparing income distributions, Journal of Economic Theory 64,520-530.

[24] Davies, J. and M. Hoy, 1995, Making Inequality Comparisons When Lorenz Curves Intersect, The American Economic Review 85, 980-986.

[25] Fishburn, PC, 1974, Convex stochastic dominance with continuous distribution functions, Journal of Economic Theory 7, 143-158.

[26] Foster, J., J. Greer and E. Thorbecke, 1984, A class of decomposable poverty measures, Econometrica 52, 761–766.

[27] Foster JE and AF Shorrocks, 1988, Poverty Orderings, Econometrica 56, 173-177.

[28] Friedman, M., 1957, A Theory of the Consumption Function, Princeton University Press.

[29] Gini, C., 1916, Il concetto di transvariazione e le sue prime applicazioni, Giornale degli Economisti e Rivista di Statistica, in C Gini (1959): 1-55.

[30] Gini, C., 1959, Transvariazione, Libreria Goliardica, Roma.

[31] Hadar, J. and WR Russell, 1969, Rules for ordering uncertain prospects, Amer. Econom. Rev. 59, 25–34.

[32] Hanoch, G. and H. Levy, 1969, The efficiency analysis of choices involving risk, Rev. Econom. Stud. 36, 335–346.

[33] Jegadeesh N., 1990, Evidence of predictable behavior of security returns. J. Finance 45, 881–898.

[34] Kobos M. and P. Milos, 2012, Inequality Decomposition by Population Subgroups for Ordinal Data, Journal of Health Economics 31, 15-21.

[35] Lee, S., K. Song and YJ Whang, 2018, Testing for a General Class of Functional Inequalities, forthcoming in Econometric Theory.

[36] Leshno, M. and H. Levy, 2002, Preferred by "All" and Preferred by "Most" Decision Makers: Almost Stochastic Dominance, Management Science 48, 1074–1085.

[37] Levy, H., 2016, Stochastic Dominance: Investment Decision Making Under Uncertainty, Kluwer Academic Publishers, Boston, MA, Third Edition.

[38] Li, S., C. Luo, Z. Wei and X. Yue, 2008, "Appendix: The 1995 and 2002 Household Surveys: Sampling Methods and Data Description," in B Gustafsson, S Li and T Sicular, eds., Inequality and Public Policy in China, 337-353, New York: Cambridge University Press.

[39] Linton, O., E. Maasoumi and YJ Whang, 2005, Consistent Testing for Stochastic Dominance under General Sampling Schemes, Review of Economic Studies 72, 735-765.

[40] Linton, O., T. Post, and YJ Whang, 2014, Testing for the Stochastic Dominance Efficiency of a Given Portfolio, Econometrics Journal 17, 59-74.

[41] Linton, O., K. Song and YJ Whang, 2014, An Improved Bootstrap Test of Stochastic Dominance, Journal of Econometrics 154, 186-202.

[42] Modigliani, F., 1976, Life-cycle, individual thrift, and the wealth of nations, American Economic Review 76, 297–313.

[43] Politis, DN and JP Romano, 1994, Large Sample Confidence Regions based on Subsamples under Minimal Assumptions. Annals of Statistics 22, 2031-2050.

[44] Politis, DN, JP Romano and M. Wolf, 2000, Subsampling, Springer-Verlag, New York.

[45] Post, T., 2003, Empirical Tests for Stochastic Dominance Efficiency, Journal of Finance 58, 1905-1932.

[46] Post, T., 2017, Empirical Tests for Stochastic Dominance Optimality, Review of Finance 21, 793–810.

[47] Rothschild, M. and JE Stiglitz, 1970, Increasing Risk: I. A Definition, Journal of Economic Theory 2, 225-243.

[48] Romano, JP and M. Wolf, 2001, Subsampling Intervals in Autoregressive Models with Linear Time Trend, Econometrica 69, 1283-1314.

[49] Russell, WR and TK Seo, 1989, Representative sets for stochastic dominance rules, in: TB Fomby and TK Seo (eds.), Studies in the economics of uncertainty: In honor of Josef Hadar, New York: Springer Verlag, 59-76.

[50] Scaillet, O. and N. Topaloglou, 2010, Testing for stochastic dominance efficiency, Journal of Business and Economic Statistics 28, 169-180.

[51] Shorrocks AF, 1978, Income Inequality and Income Mobility, Journal of Economic Theory 45, 376-393.

[52] Sen, AK, 1987, Choice, Welfare and Measurement, MIT press.

[53] Tsetlin, I., RL Winkler, RJ Huang and LY Tzeng, 2015, Generalized Almost Stochastic Dominance, Operations Research 63, 363-377.

[54] Tzeng, LY, RJ Huang and PT Shih, 2013, Revisiting Almost Second-Degree Stochastic Dominance, Management Science 59, 1250-1254

[55] Whang, YJ, 2018, Econometric Analysis of Stochastic Dominance, Cambridge University Press.

ONLINE SUPPLEMENTARY MATERIAL

for

Somewhere Between Utopia and Dystopia: Choosing From Multiple Incomparable Prospects

by

Gordon Anderson, Thierry Post, and Yoon-Jae Whang

Appendix A

Proof of Proposition 1: For every $G \in \mathcal{G}$ and $y \in \mathcal{X}$, $G^{(2)}(y) = \int_a^x G(y)dy = \mathbb{E}_G[\max(y - x, 0)] = -\mathbb{E}_G[\min(x-y,0)] = -\mathbb{E}_G[v_y(x)]$. By minimizing $G^{(2)}(y)$ over $\mathcal{G}$ for every $y \in \mathcal{X}$, the lower envelope thus results from maximizing expected value for every elementary utility function, or (9). ■

Proof of Proposition 2:

$$\begin{aligned}\max_{G\in\mathcal{G}} \mathbb{E}_G[u_w(x)] &= \max_{G\in\mathcal{G}} \int_a^b w(y)\mathbb{E}_G[v_y(x)]dy \\ &= -\min_{G\in\mathcal{G}} \int_a^b w(y)G^{(2)}(y)dy \\ &\leq -\int_a^b w(y)\underline{\mathcal{G}}_2(y)dy \\ &= -\int_a^b w(y)G_m^{(2)}(y)dy \\ &= \int_a^b w(y)\mathbb{E}_{G_m}[v_y(x)]dy = \mathbb{E}_{G_m}[u_w(x)].\end{aligned}$$

The first equality follows from $u_w(x) = \int_a^b w(y)v_y(x)dy$ and the second one from $G^{(2)}(y) = -\mathbb{E}_G[v_y(x)]$. The inequality follows from $G_n^{(2)}(x) \geq \underline{\mathcal{G}}_2(x)$, for all $n = 1, \cdots, M$. The third equality follows from utopianity of prospect P_m; the last two equalities are based on the same insights as the first two inequalities. ■

Proof of Proposition 2.3: The proof is based on the following chain of arguments for any $u_w \in \mathcal{V}_2^*(\mathcal{X}_{2,m})$:

$$
\begin{aligned}
\mathbb{E}_{G_m}\left[u_w(x)\right] &= \int w(y)\mathbb{E}_{G_m}\left[v_y(x)\right]dy \\
&= -\int w(y)G_m^{(2)}(y)dy \\
&= -\int w(y)\underline{G}_2(y)dy \qquad (1) \\
&= -\int w(y)\left(\min_{n\in\mathcal{P}} G_n^{(2)}(y)\right)dy \\
&\geq -\min_{n\in\mathcal{P}}\int w(y)G_n^{(2)}(y)dy \qquad (2) \\
&= \max_{n\in\mathcal{P}}\int w(y)\mathbb{E}_{G_n}\left[v_y(x)\right]dy \\
&= \max_{n\in\mathcal{P}}\mathbb{E}_{G_n}\left[u_w(x)\right].
\end{aligned}
$$

Equality (1) is based on $w(y) = 0$ for all $y \in \mathcal{X}_{2,m}$; inequality (2) is based on the concavity of the pointwise minimum. The condition on the mean $\mathbb{E}_{G_m}[x]$ is needed to deal with the risk neutral preferences $u_b(x) = x - b$ which are always permissible regardless of $\mathcal{X}_{2,m}$. ■

Proof of Proposition 2.4: As in the proof of Proposition 2.1, note that $G^{(2)}(y) = -\mathbb{E}_G[v_y(x)]$, for every $G \in \mathcal{G}$ and $y \in \mathcal{X}$. By maximizing $G^{(2)}(y)$ over $\mathcal{G}$ for every $y \in \mathcal{X}$, the upper envelope thus minimizes expected value for every elementary utility function, or (20). ■

Proof of Proposition 2.5: By analogy to the proof of Proposition 2.2, the following arguments are used:

$$
\begin{aligned}
\min_{G\in\mathcal{G}}\mathbb{E}_G[u_w(x)] &= -\max_{G\in\mathcal{G}}\int_a^b w(y)G^{(2)}(y)dy \\
&\geq -\int_a^b w(y)\overline{\mathcal{G}}_2(y)(y)dy \\
&= -\int_a^b w(y)G_m^{(2)}(y)dy = \mathbb{E}_{G_m}[u_w(x)].
\end{aligned}
$$

■

Proof of Proposition 3.1: We shall establish Proposition 3.1 under Assumption 1-2, because the corresponding proof under Assumption 1-1 is similar.

Let $\hat{\Delta}_{m,j} = \hat{G}_m^{(2)} - \hat{G}_j^{(2)}$ denote the estimator of $\Delta_{m,j}$ and $B_m(r) = \bigcup_{j=1}^{M} B_{m,j}(r)$. The proof first writes

$$
\begin{aligned}
& \sqrt{T}\left(\hat{\mathcal{A}}_{2,m} - \mathcal{A}_{2,m}\right) \qquad (3) \\
= & \int_{\mathcal{X}} \sqrt{T}\left[\max_{1\le j\le M}\left\{\hat{\Delta}_{m,j}(x)\right\} - \max_{1\le j\le M}\left\{\Delta_{m,j}(x)\right\}\right] dx \qquad (4) \\
= & \int_{\mathcal{X}} \sqrt{T}\left[\max_{1\le j\le M}\left\{\hat{\Delta}_{m,j}(x) - \Delta_{m,\max}(x)\right\}\right] dx \\
= & \int_{\mathcal{X}} \left[\max_{1\le j\le M}\left\{\sqrt{T}\left(\hat{\Delta}_{m,j}(x) - \Delta_{m,j}(x)\right) + \sqrt{T}\left(\Delta_{m,j}(x) - \Delta_{m,\max}(x)\right)\right\}\right] dx \\
= & \int_{\mathcal{X}\backslash B_m((1-\varepsilon)c_T)} \left[\max_{1\le j\le M}\left\{\sqrt{T}\left(\hat{\Delta}_{m,j}(x) - \Delta_{m,j}(x)\right) + \sqrt{T}\left(\Delta_{m,j}(x) - \Delta_{m,\max}(x)\right)\right\}\right] dx \\
& + \int_{B_m((1-\varepsilon)c_T)} \left[\max_{1\le j\le M}\left\{\sqrt{T}\left(\hat{\Delta}_{m,j}(x) - \Delta_{m,j}(x)\right) + \sqrt{T}\left(\Delta_{m,j}(x) - \Delta_{m,\max}(x)\right)\right\}\right] dx \\
:= & A_{1T} + A_{2T}.
\end{aligned}
$$

As for A_{1T}, observe that for any positive sequence $r'_T \to 0$,

$$
\begin{aligned}
& \Pr\left(\int_{\mathcal{X}\backslash B_m((1-\varepsilon)c_T)} \left[\max_{1\le j\le M}\left\{\sqrt{T}\left(\hat{\Delta}_{m,j}(x) - \Delta_{m,j}(x)\right) + \sqrt{T}\left(\Delta_{m,j}(x) - \Delta_{m,\max}(x)\right)\right\}\right] dx > r'_T\right) \\
\le & \Pr\left(\int_{\mathcal{X}\backslash B_m((1-\varepsilon)c_T)} \left[\max_{1\le j\le M}\left\{\sqrt{T}\left(\hat{\Delta}_{m,j}(x) - \Delta_{m,j}(x)\right) - \sqrt{T}(1-\varepsilon)c_T\right\}\right] dx > r'_T\right) \\
\le & \sum_{j=1}^{M} \Pr\left(\sup_{x\in\mathcal{X}} \sqrt{T}\left(\hat{\Delta}_{m,j}(x) - \Delta_{m,j}(x)\right) > \sqrt{T}(1-\varepsilon)c_T\right) \to 0. \qquad (5)
\end{aligned}
$$

The first inequality holds because $(\Delta_{m,j}(x) - \Delta_{m,\max}(x)) < -(1-\varepsilon)c_T$ for $x \in \mathcal{X}\backslash B_m((1-\varepsilon)c_T)$; the second inequality follows from the inequality $\Pr(\cup A_i) \le \sum_i P(A_i)$; the final convergence to 0 holds because $\sqrt{T}\left(\hat{\Delta}_{m,j} - \Delta_{m,j}\right)$ is asymptotically tight (Lemma A.1 of Linton, Post and Whang (2014)).

As for A_{2T}, write

$$\begin{aligned}
&\int_{B_m((1-\varepsilon)c_T)}\left[\max_{1\leq j\leq M}\left\{\sqrt{T}\left(\hat{\Delta}_{m,j}(x)-\Delta_{m,j}(x)\right)+\sqrt{T}\left(\Delta_{m,j}(x)-\Delta_{m,\max}(x)\right)\right\}\right]dx \qquad (6)\\
=&\int_{B_m((1-\varepsilon)c_T)\backslash B_m(\frac{r_T}{\sqrt{T}})}\left[\max_{1\leq j\leq M}\left\{\sqrt{T}\left(\hat{\Delta}_{m,j}(x)-\Delta_{m,j}(x)\right)+\sqrt{T}\left(\Delta_{m,j}(x)-\Delta_{m,\max}(x)\right)\right\}\right]dx\\
&+\int_{B_m(\frac{r_T}{\sqrt{T}})}\left[\max_{1\leq j\leq M}\left\{\sqrt{T}\left(\hat{\Delta}_{m,j}(x)-\Delta_{m,j}(x)\right)+\sqrt{T}\left(\Delta_{m,j}(x)-\Delta_{m,\max}(x)\right)\right\}\right]dx.
\end{aligned}$$

The first term on the right hand side of (6) is bounded by

$$\begin{aligned}
&\left|\int_{B_m((1-\varepsilon)c_T)\backslash B_m(\frac{r_T}{\sqrt{T}})}\left[\max_{1\leq j\leq M}\left\{\sqrt{T}\left(\hat{\Delta}_{m,j}(x)-\Delta_{m,j}(x)\right)\right\}\right]dx\right|\\
\leq\;&\max_{1\leq j\leq M}\left\{\sup_{x\in\mathcal{X}}\left|\sqrt{T}\left(\hat{\Delta}_{m,j}(x)-\Delta_{m,j}(x)\right)\right|\right\}\cdot\int_{B_m((1-\varepsilon)c_T)\backslash B_m(\frac{r_T}{\sqrt{T}})}dx\\
=\;&O_p(r_T)=o_p(1), \qquad (7)
\end{aligned}$$

where the first equality holds by Assumption 2(ii) and asymptotic tightness of $\sqrt{T}\left(\hat{\Delta}_{m,j}-\Delta_{m,j}\right)$.

Let $\mathcal{J}:=\{(i_1,...,i_M)\in\{0,1\}^M:\ i_j=0 \text{ for some } 1\leq i\leq M\}$. For $\mathbf{i}\in\mathcal{J}$ and $0<r<s$, let

$$A_{m,\mathbf{i}}(r,s):=A^{*}_{m,\mathbf{i}}(r,s)\bigcup A^{**}_{m,\mathbf{i}}(r,s), \qquad (8)$$

where

$$\begin{aligned}
A^{*}_{m,\mathbf{i}}(r,s)&:=\left(\bigcap_{\{j:i_j=0\}}B^{i_j}_{m,j}(r)\right)\bigcap\left(\bigcap_{\{j:i_j=1\}}B^{i_j}_{m,j}(s)\right),\\
A^{**}_{m,\mathbf{i}}(r,s)&:=\left(\bigcap_{\{j:i_j=0\}}B^{i_j}_{m,j}(r)\right)\bigcap\left(\bigcap_{\{j:i_j=1\}}\left(B^{i_j}_{m,j}(r)\backslash B^{i_j}_{m,j}(s)\right)\right),
\end{aligned}$$

with $B^{0}_{m,j}(r):=B_{1,j}(r)$ and $B^{1}_{m,j}(r):=\mathcal{X}\backslash B_{m,j}(r)$. Then, we may write $B_m(r)$ as a union of the disjoint intervals, that is,

$$B_m(r) := \bigcup_{j=1}^{M} B_{m,j}(r) = \bigcup_{\mathbf{i}\in\mathcal{J}} A_{m,\mathbf{i}}(r,r). \tag{9}$$

We have

$$\begin{aligned}
&\int_{B_m\left(\frac{r_T}{\sqrt{T}}\right)} \max_{1\le j\le M}\left\{\sqrt{T}\left(\hat{\Delta}_{m,j}(x)-\Delta_{m,\max}(x)\right)\right\}dx \\
=&\sum_{\mathbf{i}\in\mathcal{J}}\int_{A_{m,\mathbf{i}}\left(\frac{r_T}{\sqrt{T}},\frac{r_T}{\sqrt{T}}\right)} \max_{1\le j\le M}\left\{\sqrt{T}\left(\hat{\Delta}_{m,j}(x)-\Delta_{m,\max}(x)\right)\right\}dx \qquad (10)\\
=&\sum_{\mathbf{i}\in\mathcal{J}}\int_{A^{*}_{m,\mathbf{i}}\left(\frac{r_T}{\sqrt{T}},(1-\varepsilon)c_T\right)} \max_{1\le j\le M}\left\{\sqrt{T}\left(\hat{\Delta}_{m,j}(x)-\Delta_{m,\max}(x)\right)\right\}dx \qquad (11)\\
&+\sum_{\mathbf{i}\in\mathcal{J}}\int_{A^{**}_{m,\mathbf{i}}\left(\frac{r_T}{\sqrt{T}},(1-\varepsilon)c_T\right)} \max_{1\le j\le M}\left\{\sqrt{T}\left(\hat{\Delta}_{m,j}(x)-\Delta_{m,\max}(x)\right)\right\}dx \\
=&\sum_{\mathbf{i}\in\mathcal{J}}\int_{A^{*}_{m,\mathbf{i}}\left(\frac{r_T}{\sqrt{T}},(1-\varepsilon)c_T\right)} \max_{1\le j\le M}\left\{\sqrt{T}\left(\hat{\Delta}_{m,j}(x)-\Delta_{m,\max}(x)\right)\right\}dx \qquad (12)\\
&+o_p(1)\\
=&\sum_{\mathbf{i}\in\mathcal{J}}\int_{A^{*}_{m,\mathbf{i}}\left(\frac{r_T}{\sqrt{T}},(1-\varepsilon)c_T\right)} \max_{j\in\mathcal{M}(\mathbf{i})}\left\{\sqrt{T}\left(\hat{\Delta}_{m,j}(x)-\Delta_{m,\max}(x)\right)\right\}dx+o_p(1) \qquad (13)\\
=&\sum_{\mathbf{i}\in\mathcal{J}}\int_{A^{*}_{m,\mathbf{i}}\left(\frac{r_T}{\sqrt{T}},(1-\varepsilon)c_T\right)} \max_{j\in\mathcal{M}(\mathbf{i})}\left\{\sqrt{T}\left(\hat{\Delta}_{m,j}(x)-\Delta_{m,j}(x)\right)\right\}dx+o_p(1), \qquad (14)
\end{aligned}$$

where $\mathcal{M}(\mathbf{i}) = \{j \in \{1, ..., M\} : i_j = 0\}$. In the above equations, the equalities (10) and (11) hold by (8) and (9), respectively. The equality (12) holds because

$$\begin{aligned}
&\left|\sum_{\mathbf{i}\in\mathcal{J}}\int_{A^{**}_{m,\mathbf{i}}\left(\frac{r_T}{\sqrt{T}},(1-\varepsilon)c_T\right)} \max_{1\le j\le M}\left\{\sqrt{T}\left(\hat{\Delta}_{m,j}(x)-\Delta_{m,j}(x)\right)+\sqrt{T}\left(\Delta_{m,j}(x)-\Delta_{m,\max}(x)\right)\right\}dx\right|\\
\le&\max_{1\le j\le M}\left\{\sup_{x\in\mathcal{X}}\left|\sqrt{T}\left(\hat{\Delta}_{m,j}(x)-\Delta_{m,j}(x)\right)\right|\right\}\cdot\sum_{\mathbf{i}\in\mathcal{J}}\int_{\bigcap_{\{j:i_j=1\}}\left(B_{m,j}((1-\varepsilon)c_T)\setminus B_{m,j}(\frac{r_T}{\sqrt{T}})\right)}dx\\
=&\,o_p(1)
\end{aligned}$$

by Assumption 2(ii) and asymptotic tightness of $\sqrt{T}\left(\hat{\Delta}_{m,j}-\Delta_{m,j}\right)$.

On the other hand, the equality (13) holds because: for each $\mathbf{i} \in \mathcal{J}$ and for any positive constant $\eta > 0$,

$$\begin{aligned}
&\Pr\left(\int_{A^*_{m,\mathbf{i}}\left(\frac{r_T}{\sqrt{T}},(1-\varepsilon)c_T\right)} \left[\max_{1\le j\le M} \left\{ \sqrt{T}\left(\hat{\Delta}_{m,j}(x)-\Delta_{m,\max}(x)\right)\right\} \right.\right. \\
&\qquad \left.\left. - \max_{j\in\mathcal{M}(\mathbf{i})} \left\{ \sqrt{T}\left(\hat{\Delta}_{m,j}(x)-\Delta_{m,\max}(x)\right)\right\} \right] dx > \eta \right) \\
&\le \Pr\left(\sup_{x\in\mathcal{X}} \left[\max_{j\notin M(\mathbf{i})} \left\{ \sqrt{T}\left(\hat{\Delta}_{m,j}(x)-\Delta_{m,\max}(x)\right)\right\} \right.\right. \\
&\qquad \left.\left. - \max_{j\in\mathcal{M}(\mathbf{i})} \left\{ \sqrt{T}\left(\hat{\Delta}_{m,j}(x)-\Delta_{m,\max}(x)\right)\right\} \right] > 0 \right) \\
&\le \Pr\left(\sup_{x\in\mathcal{X}} \left[\max_{j\notin \mathcal{M}(\mathbf{i})} \left\{ \sqrt{T}\left(\hat{\Delta}_{m,j}(x)-\Delta_{m,j}(x)\right)\right\} \right.\right. \\
&\qquad \left.\left. - \max_{j\in\mathcal{M}(\mathbf{i})} \left\{ \sqrt{T}\left(\hat{\Delta}_{m,j}(x)-\Delta_{m,j}(x)\right)\right\} \right] > (1-\varepsilon)\sqrt{T}c_T \right) \\
&\to 0,
\end{aligned}$$

where the second inequality holds since $-\frac{r_T}{\sqrt{T}} \le (\Delta_{m,j}(x)-\Delta_{m,\max}(x)) \le 0$ for $j \in \mathcal{M}(\mathbf{i})$ and $(\Delta_{m,j}(x)-\Delta_{m,\max}(x)) < -(1-\varepsilon)c_T$ for $j \notin \mathcal{M}(\mathbf{i})$ and the last convergence to zero holds by Assumption 2(i) and the asymptotic tightness of $\sqrt{T}\left(\hat{\Delta}_{m,j}-\Delta_{m,j}\right)$ for $j = 1,...,M$.

Similarly, the last equality (14) holds using $-\frac{r_T}{\sqrt{T}} \le (\Delta_{m,j}(x)-\Delta_{m,\max}(x)) \le 0$ for $j \in \mathcal{M}(\mathbf{i})$.

Combining (4), (6), (7) and (14), it follows that

$$\sqrt{T}\left(\hat{\mathcal{A}}_{2,m}-\mathcal{A}_{2,m}\right) = \sum_{\mathbf{i}\in\mathcal{J}} \int_{A^*_{m,\mathbf{i}}\left(\frac{r_T}{\sqrt{T}},(1-\varepsilon)c_T\right)} \max_{j\in\mathcal{M}(\mathbf{i})} \left\{ \sqrt{T}\left(\hat{\Delta}_{m,j}(x)-\Delta_{m,j}(x)\right)\right\} dx + o_p(1). \tag{15}$$

By weak convergence of $\sqrt{T}\left(\hat{\Delta}_{m,j}-\Delta_{m,j}\right)$ to the tight Gaussian process $\nu_{1,j}$, one can conclude that

$$\sqrt{T}\left(\hat{\mathcal{A}}_{2,m}-\mathcal{A}_{2,m}\right) = \sum_{\mathbf{i}\in\mathcal{J}} \int_{A^*_{m,\mathbf{i}}\left(\frac{r_T}{\sqrt{T}},(1-\varepsilon)c_T\right)} \max_{j\in\mathcal{M}(\mathbf{i})} \left\{ \nu_{m,j}(x)\right\} dx + o_p(1). \tag{16}$$

Similarly, for $\mathbf{k} \in \mathcal{K}$ and $0 < r < s$, let

$$\tilde{A}_{\mathbf{k}}(r,s) := \tilde{A}^{*}_{\mathbf{k}}(r,s) \bigcup \tilde{A}^{**}_{\mathbf{k}}(r,s), \tag{17}$$

where

$$\tilde{A}^{*}_{\mathbf{k}}(r,s) := \left(\bigcap_{\{(i,j):i\neq j, k_{i,j}=0\}} \tilde{B}^{k_{i,j}}_{i,j}(r) \right) \bigcap \left(\bigcap_{\{(i,j):i\neq j, k_{i,j}=1\}} \tilde{B}^{k_{i,j}}_{i,j}(s) \right),$$
$$\tilde{A}^{**}_{\mathbf{k}}(r,s) := \left(\bigcap_{\{(i,j):i\neq j, k_{i,j}=0\}} \tilde{B}^{k_{i,j}}_{i,j}(r) \right) \bigcap \left(\bigcap_{\{(i,j):i\neq j, k_{i,j}=1\}} \left(\tilde{B}^{k_{i,j}}_{i,j}(r) \backslash \tilde{B}^{k_{i,j}}_{i,j}(s) \right) \right),$$

with $\tilde{B}^{0}_{i,j}(r) := \tilde{B}_{i,j}(r)$ and $\tilde{B}^{1}_{i,j}(s) := \mathcal{X} \backslash \tilde{B}_{i,j}(s)$. Then, we may write

$$\tilde{B}(r) := \bigcup_{i\neq j} \tilde{B}_{i,j}(r) = \bigcup_{\mathbf{k}\in\mathcal{K}} \tilde{A}_{\mathbf{k}}(r,r). \tag{18}$$

By replacing $\Delta_{m,j}$, $\max_{1\leq j\leq M}$, $\Delta_{m,\max}$, $B_{m,j}(r)$, $B_m(r)$, $A_{\mathbf{i}}$, $A^{*}_{\mathbf{i}}$, $A^{**}_{\mathbf{i}}$, and $\mathcal{M}(\mathbf{i})$ by $\Delta_{i,j}$, $\max_{1\leq i\neq j\leq M}, \Delta_{\max}(x)$, $\tilde{B}_{i,j}(r)$, $\tilde{B}(r)$, $\tilde{A}_{\mathbf{k}}$, $\tilde{A}^{*}_{\mathbf{k}}$, $\tilde{A}^{**}_{\mathbf{k}}$, and $\tilde{\mathcal{M}}(\mathbf{k})$, respectively, in the above arguments, one can establish

$$\begin{aligned} & \sqrt{T}\left(\hat{\mathcal{T}}_2 - \mathcal{T}_2\right) && (19) \\ = & \sum_{\mathbf{k}\in K} \int_{\tilde{A}^{*}_{\mathbf{k}}\left(\frac{r_T}{\sqrt{T}}, (1-\varepsilon)c_T\right)} \max_{(i,j)\in\tilde{\mathcal{M}}(\mathbf{k})} \left\{ \sqrt{T}\left(\hat{\Delta}_{i,j}(x) - \Delta_{i,j}(x)\right) \right\} dx + o_p(1) \\ = & \sum_{\mathbf{k}\in K} \int_{\tilde{A}^{*}_{\mathbf{k}}\left(\frac{r_T}{\sqrt{T}}, (1-\varepsilon)c_T\right)} \max_{(i,j)\in\tilde{\mathcal{M}}(\mathbf{k})} \{\nu_{i,j}(x)\}\, dx + o_p(1). && (20) \end{aligned}$$

Therefore, from (16), (19) and using the delta method, it follows that

$$\begin{aligned} \sqrt{T}\left(\hat{\mathcal{I}}_{2,m} - \mathcal{I}_{2,m}\right) &= -\sqrt{T}\left(\frac{\hat{\mathcal{A}}_{2,m}}{\hat{\mathcal{T}}_2} - \frac{\mathcal{A}_{2,m}}{\mathcal{T}_2} \right) \\ &= -\frac{1}{\mathcal{T}_2} \sum_{\mathbf{i}\in\mathcal{J}} \int_{A^{*}_{m,\mathbf{i}}\left(\frac{r_T}{\sqrt{T}}, (1-\varepsilon)c_T\right)} \max_{j\in\mathcal{M}(\mathbf{i})} \{\nu_{m,j}(x)\}\, dx \qquad (21) \\ &\quad + \frac{\mathcal{A}_{2,m}}{\mathcal{T}_2^2} \sum_{\mathbf{k}\in\mathcal{K}} \int_{\tilde{A}^{*}_{\mathbf{k}}\left(\frac{r_T}{\sqrt{T}}, (1-\varepsilon)c_T\right)} \max_{(i,j)\in\tilde{\mathcal{M}}(\mathbf{k})} \{\nu_{i,j}(x)\}\, dx + o_p(1). \end{aligned}$$

This establishes part (i) of Proposition 3.1.

We now prove part (ii). Suppose that $\inf_{x\in\mathcal{X}}\left[G_m(x)-\max_{j\neq m}\{G_j(x)\}\right]>\bar{\eta}$ for some $\bar{\eta}>0$. Then, G_m is dystopian and $\mathcal{A}_{2,m}=\mathcal{T}_{2,m}$, so that $\mathcal{I}_{2,m}=0$. In this case, we have

$$
\begin{aligned}
&\sqrt{T}\left(\hat{\mathcal{I}}_{2,m}-\mathcal{I}_{2,m}\right)\\
=&\frac{1}{\hat{\mathcal{T}}_2}\sqrt{T}\left(\hat{\mathcal{A}}_{2,m}-\hat{\mathcal{T}}_2\right)\\
=&\frac{1}{\mathcal{T}_2}\int_{\mathcal{X}}\min_{1\leq j\leq M}\left\{\sqrt{T}\left(\hat{\Delta}_{m,j}(x)-\Delta_{m,j}(x)\right)+\sqrt{T}\Delta_{m,j}(x)\right\}dx\ +\ o_p(1).
\end{aligned}
\tag{22}
$$

Now, given $\hat{\Delta}_{m,m}=\Delta_{m,m}=0$ and $\inf_{x\in\mathcal{X}}\Delta_{m,j}(x)>\bar{\eta}$, asymptotic tightness of $\sqrt{T}\left(\hat{\Delta}_{m,j}(x)-\Delta_{m,j}(x)\right)$ implies that the right hand side of (22) is $o_p(1)$.

On the other hand, suppose that $\inf_{x\in\mathcal{X}}\left[\min_{j\neq m}\{G_j(x)\}-G_m(x)\right]>\bar{\eta}$ for some $\bar{\eta}>0$. Then, G_m is utopian and $\mathcal{A}_{2,m}=0$, so that $\mathcal{I}_{2,m}=1$. In this case,

$$
\begin{aligned}
&\sqrt{T}\left(\hat{\mathcal{I}}_{2,m}-\mathcal{I}_{2,m}\right)\\
=&-\frac{1}{\hat{\mathcal{T}}_2}\sqrt{T}\left(\hat{\mathcal{A}}_{2,m}-\mathcal{A}_{2,m}\right)\\
=&-\frac{1}{\mathcal{T}_2}\int_{\mathcal{X}}\max_{1\leq j\leq M}\left\{\sqrt{T}\left(\hat{\Delta}_{m,j}(x)-\Delta_{m,j}(x)\right)+\sqrt{T}\Delta_{m,j}(x)\right\}dx\ +\ o_p(1).
\end{aligned}
\tag{23}
$$

Given $\hat{\Delta}_{m,m}=\Delta_{m,m}=0$ and $\sup_{x\in\mathcal{X}}\Delta_{m,j}(x)<-\bar{\eta}$, asymptotic tightness of $\sqrt{T}\left(\hat{\Delta}_{m,j}(x)-\Delta_{m,j}(x)\right)$ implies that the right hand side of (23) is again $o_p(1)$ as was to be shown. ■

Proof of Proposition 3.2: If the Gaussian process ν is regular on $\mathcal{A}_1^*$ and $\tilde{\mathcal{A}}^*$, then Proposition 3.1 implies that $\sqrt{T}\left(\hat{\mathcal{I}}_{2,m}-\mathcal{I}_{2,m}\right)$ converges weakly to a non-degenerate limiting law. Therefore, the asymptotic coverage probability is $(1-\alpha)$ by Lemma 3.2.1 of Politis, Romano and Wolf (2000).

Otherwise, the coverage probability

$$
\begin{aligned}
\Pr\left(\mathcal{I}_{2,m}\in CI_{T,b}\right)&=\Pr\left(\sqrt{T}\left|\ \hat{\mathcal{I}}_{2,m}-\mathcal{I}_{2,m}\right|\leq\max\{s_{T,b}(1-\alpha/2),\eta\}\right)\\
&=\Pr\left(o_p(1)\leq\max\{o_p(1),\eta\}\right)\to 1.
\end{aligned}
$$

This establishes Proposition 3.2. ■

Proof of Proposition 3.3: We shall prove the result for the case $a = C$, because the proofs of the case $a = A, B$ are similar.

For subsets $\mathcal{C}_{m,n} = \{(\mathcal{C}_{m,\mathbf{i}}, \mathcal{C}_{n,\mathbf{i}}) \subset \mathcal{X} \times \mathcal{X} : \mathbf{i} \in \mathcal{J}\}$ and $\mathcal{D} := \{\mathcal{D}_{\mathbf{k}} \subset \mathcal{X} : \mathbf{k} \in \mathcal{K}\}$, define

$$\tilde{\Lambda}_{\mathcal{C},\mathcal{D}}(\nu) := \frac{1}{\mathcal{T}_2} \sum_{\mathbf{i} \in \mathcal{J}} \left[\int_{\mathcal{C}_{n,\mathbf{i}}} \max_{j \in \mathcal{M}(\mathbf{i})} \left\{ \lambda_j^{-1/2} \nu_{n,j}(x) \right\} dx - \int_{\mathcal{C}_{m,\mathbf{i}}} \max_{j \in \mathcal{M}(\mathbf{i})} \left\{ \lambda_j^{-1/2} \nu_{m,j}(x) \right\} dx \right] \quad (24)$$
$$- \frac{(\mathcal{A}_{2,n} - \mathcal{A}_{2,m})}{\mathcal{T}_2^2} \sum_{\mathbf{k} \in \mathcal{K}} \int_{\mathcal{D}_{\mathbf{k}}} \max_{(i,j) \in \tilde{\mathcal{M}}(\mathbf{k})} \left\{ \lambda_j^{-1/2} \nu_{i,j}(x) \right\} dx.$$

Then, by an argument similar to the proof of Proposition 3.1, we can establish the following asymptotic approximation:

$$\sqrt{T} \left[\left(\hat{\mathcal{I}}_{2,m} - \mathcal{I}_{2,m} \right) - \left(\hat{\mathcal{I}}_{2,n} - \mathcal{I}_{2,n} \right) \right] = \tilde{\Lambda}_{\mathcal{A}^*_{m,n}, \tilde{\mathcal{A}}^*}(\nu) + o_p(1),$$

where $\mathcal{A}^*_{m,n} = \{(\mathcal{A}^*_{m,\mathbf{i}} \left(T^{-1/2} r_T,\, (1-\varepsilon) c_T\right),\, \mathcal{A}^*_{n,\mathbf{i}} \left(T^{-1/2} r_T,\, (1-\varepsilon) c_T\right)) \subset \mathcal{X} \times \mathcal{X} : \mathbf{i} \in \mathcal{J}\}$, $\tilde{\mathcal{A}}^*$ is as defined in Proposition 3.1, and $\mathcal{A}^*_{m,\mathbf{i}}$ and $\mathcal{A}^*_{n,\mathbf{i}}$ are as defined in (32) with $B_{1,j}$ replaced by $B_{m,j}$ and $B_{n,j}$, respectively. If ν is regular on in the sense that it satisfies (39) and (40) with Λ replaced by $\tilde{\Lambda}$, then the result of Proposition 3.3 (i) holds with equality under the null hypothesis H_0^C. Otherwise, the rejection probability converges to 0.

To prove (ii), consider the subsampling distribution (not scaled by $\sqrt{b}$)

$$\hat{S}^0_{T,b}(w) := \frac{1}{LT)} \sum_{i=1}^{L(T)} 1 \left(\left| \left(\hat{\mathcal{I}}_{m,b,i} - \hat{\mathcal{I}}_{2,m} \right) - \left(\hat{\mathcal{I}}_{n,b,i} - \hat{\mathcal{I}}_{2,n} \right) \right| \leq w \right).$$

The distribution $\hat{S}^0_{T,b}(\cdot)$ converges in distribution to a point mass at 0. It follows that $c^B_{T,b,\eta}(1-\alpha/2)/\sqrt{b}$ converges to 0. Since $\liminf_{T \to \infty}(T/b) > 1$ and $\left| \hat{\mathcal{I}}_{2,m} - \hat{\mathcal{I}}_{2,n} \right| \to \left| \mathcal{I}_{2,m} - \mathcal{I}_{2,n} \right|$ in probability (with $\left| \mathcal{I}_{2,m} - \mathcal{I}_{2,n} \right| > 0$ under H_1^A), it follows by Slutsky's theorem that the asymptotic rejection probability is one. ■

Appendix B

Household wellbeing is measured as some monotonic non decreasing function of adult-equivalised household income. Lifecycle theory (Friedman (1957)) predicts that, households identical in every respect except for their vintage or lifecycle stage (determined by the age of the head of household), will have different incomes. Thus in order to examine income based wellbeing distributions of households of Pre-CR, CR and Post-CR cohorts at a common vintage, an adult-equivalent household income function was estimated and adult-equivalent household incomes were projected for each cohort at a common vintage. Because vintage also affected household size (which in turn affects household wellbeing) a household size equation was also generated. After some experimentation the preferred specifications are reported in Table B-1.

Aside from the age (vintage) of the household the primary drivers of household income was assumed to be the aggregate level of education (the sum of husbands and wives education levels) and the size of the household. Income turned out to be an increasing concave quadratic form in the aggregate level of education (modified by the vintage of the household) which, because of increasing returns to education, becomes convex in the Post-CR period. Household size has a negative effect on household wellbeing the magnitude of which is diminished in the CR era. Generally household income increases with vintage though the impact is diminished in the Post-CR era.

Household size was modeled as a function of household vintage, social class and aggregate years of education and turned out to be a concave quadratic function of the household vintage in Pre-CR and Post-CR eras which becomes convex in the CR with an overall increase in family size during the CR. Family size is negatively related to education years though this effect is diminished in the CR and Post-CR eras. Social class has a negative effect on family size (higher class means smaller families).

These equations were used to project adult-equivalent household income distributions for households of vintage 50 in each cohort. Of 6,822 households 2,400 were in the Pre-CR cohort, 3,116 in the CR cohort and 1,306 in the Post-CR cohort.

Table B-1: Equivalised Household Income and Family Size equations

Ln(Income/$\sqrt{\text{Size}}$)	Coefficient (|t-stat|)	$\sqrt{\text{Size}}$	Coefficient (|t-stat|)
Intercept	8.802 (93.27)	Intercept	0.332 (1.80)
Education	0.227 (5.77)	Vintage	0.010 (1.76)
Vintage * Education	0.001 (0.44)	Vintage^2	0.000 (2.30)
Vintage^2 * Education	0.000 (0.24)	Education	-0.002 (5.53)
Education^2	-0.007 (7.97)	Social Class (Class)	-0.003 (2.82)
$\sqrt{\text{Size}}$	-0.725 (18.47)	CR	0.790 (1.88)
CR dummy (CR)	-0.280 (3.88)	Post-CR dummy (Post)	-0.211 (0.65)
Vintage^2 * Education * CR	0.000 (3.69)	Education * CR	0.001 (1.30)
Vintage * Education * Post	-0.010 (4.32)	Class * CR	0.003 (1.81)
Vintage^2 * Education * Post	0.000 (3.83)	Vintage * CR	-0.031 (1.72)
Education^2 * Post	0.008 (4.90)	Vintage^2 * CR	0.000 (1.50)
$\sqrt{\text{Size}}$ * Post	0.352 (3.44)	Education * Post	0.001 (0.67)
		Class * Post	0.002 (0.96)
		Vintage * Post	0.016 (0.90)
		Vintage^2 * Post	0.000 (0.95)