Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A Statistical Framework for Pareto-Improving Resource Reallocation: Theory and a Real-Data Health-Insurance Application

houssam boughabi

houssam1404@gmail.com

Institut National de Statistique et d'Economie Appliquée https://orcid.org/0009-0005-3492-1910

Research Article

Keywords: Pareto improvement, resource reallocation, admissible weights, welfare economics, health insurance, Statistical application

Posted Date: September 9th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10964541/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

A Statistical Framework for Pareto-Improving Resource Reallocation: Theory and a Real-Data Health-Insurance Application

June 27, 2026

Abstract

This paper develops a statistical framework for assessing Pareto-improving resource reallocations under weighted allocation schemes. The central objective is to characterize the admissible set of allocation weights that can improve at least one participant's outcome without reducing the outcome of any other participant. The framework formalizes Pareto admissibility through non-deterioration and strict-improvement conditions and interprets these conditions as testable restrictions on weighted reallocations. To illustrate the empirical relevance of the approach, the paper applies the framework to real data from the RAND Health Insurance Experiment, comparing access-related outcomes across cost-sharing regimes and health-status groups. The empirical application shows how the proposed framework can distinguish average improvements from genuine Pareto-admissible reallocations. The paper contributes a methodological tool for welfare analysis, public policy evaluation, and resource-allocation studies where distributional non-deterioration is analytically important.

Keywords: Pareto improvement; resource reallocation; admissible weights; welfare economics; health insurance; Statistical application.

JEL Classification: D60; D61; D63; I13; C18; C21.

1 Introduction

The problem of resource allocation is one of the foundational questions of welfare economics. Since Pareto (1906), economists have used the concept of Pareto improvement to identify reallocations that make at least one participant better off while making no participant worse off. The criterion remains powerful because it avoids direct interpersonal utility comparisons, but it is also demanding because it requires non-deterioration for every participant. Modern welfare economics, social choice theory, and public economics have repeatedly returned to this tension between efficiency, equity, aggregation, and individual protection (Bergson, 1938; Samuelson, 1947; Arrow, 1951; Debreu, 1959; Sen, 1970; Atkinson, 1970).

Traditional approaches usually begin with a specified utility function, social welfare function, planner objective, or mechanism. In such settings, the analyst evaluates allocations after the welfare criterion has already been imposed. This tradition is central to the theory of social welfare and general equilibrium, but it can be restrictive in policy environments where the allocation rule itself is under design and where the admissible set of welfare-preserving weights is not known in advance (Little, 1950; Kaldor, 1939; Hicks, 1939; Harsanyi, 1955; Rawls, 1971; Varian, 1974; Mas-Colell, Whinston, and Green, 1995).

A related literature studies redistribution, taxation, and social insurance. Mirrlees (1971), Diamond and Mirrlees (1971), Atkinson and Stiglitz (1976), and Saez (2001) show that welfare-improving policy design depends on behavioral responses, distributional weights, and feasibility constraints. In parallel, the fair-allocation literature emphasizes envy-freeness, equity, and efficiency as complementary but sometimes conflicting desiderata (Varian, 1974; Fleurbaey and Maniquet, 2011). These literatures demonstrate the importance of allocation weights, but they do not always treat the admissible set of Pareto-preserving weights as the main object of statistical characterization.

The contribution of the present paper is to reverse the usual analytical order. Instead of starting from a pre-specified welfare objective, the paper asks which allocation-weight vectors are admissible because they satisfy Pareto-improvement restrictions. This shift is particularly relevant when resource allocation is implemented through weighted rules, fiscal transfers, health-insurance plans, social expenditure formulas, or institutional assignment mechanisms. In these cases, the economic question is not only whether aggregate welfare increases, but whether the chosen weights preserve non-deterioration for every relevant group.

The paper therefore defines an admissible set of allocation weights and interprets the Pareto-improvement requirement as a statistical constraint. The framework is intentionally general. Participants may be individuals, households, regions, sectors, or policy-relevant groups. Resources may be monetary transfers, public services, insurance generosity, or weighted access to a scarce good. The key requirement is that the post-reallocation outcome can be compared with a baseline outcome and that non-deterioration can be evaluated for each participant or group.

To make the framework empirically meaningful, the paper adds a real-data application using the RAND Health Insurance Experiment. The RAND HIE is one of the most influential randomized social experiments in health economics; it assigned families to insurance plans with different cost-sharing rates and has been widely used to study the effect of insurance generosity on medical utilization (Manning et al., 1987; Newhouse, 1993; Brook et al., 2006). In the present paper, lower cost sharing is interpreted as a more generous resource-allocation regime, and outpatient visits to medical doctors are interpreted as an observed access outcome. This application is not presented as a full causal welfare evaluation. Rather, it illustrates how the proposed admissibility criterion can be applied to real data to check whether a plausible reallocation improves access without deteriorating outcomes for vulnerable health-status groups.

The empirical results are deliberately modest but informative. Raw group comparisons suggest that more generous cost sharing increases average medical visits for excellent-, good-, and fair-health groups, while the poor-health group exhibits a small negative raw difference. Hence, the raw reallocation improves average access but does not satisfy the strict Pareto no-deterioration condition. A covariate-adjusted model produces positive estimated access effects for all health-status groups, although the poor-health margin remains statistically imprecise. This finding strengthens the methodological message: average improvement is not equivalent to Pareto improvement, and the proposed framework helps identify precisely where an apparently beneficial reallocation violates the non-deterioration constraint.

The paper is organized as follows. Section 2 presents the theoretical framework and the admissible-weight set. Section 3 develops the statistical interpretation. Section 4 presents the real-data application using the RAND Health Insurance Experiment. Section 5 concludes with methodological implications and avenues for future research.

2 Theoretical framework

Consider a finite set of participants indexed by i = 1, ..., n. Participant i has a strictly positive resource-cost or need parameter $s_i > 0$ and a baseline outcome x_i^0. A weighted reallocation assigns the post-reallocation outcome x_i^1 through a non-negative allocation weight $\alpha_i \geq 0$.

$$x_i^1 = \alpha_i s_i, \quad i = 1, \ldots, n.$$

The individual gain generated by the reallocation is defined as the difference between the post-reallocation outcome and the baseline outcome:

$$g_i(\alpha) = x_i^1 - x_i^0 = \alpha_i s_i - x_i^0.$$

A reallocation is Pareto-improving when at least one participant experiences a strict gain and no participant experiences deterioration. Formally:

$$g_i(\alpha) \geq 0 \text{ for every } i, \text{ and } g_j(\alpha) > 0 \text{ for at least one } j.$$

The corresponding admissible set of allocation weights is therefore:

$$S(\alpha) = \{\alpha : \alpha_i s_i \geq x_i^0 \text{ for every } i = 1, \ldots, n;\ \alpha_j s_j > x_j^0 \text{ for at least one } j\}.$$

The baseline x_i^0 is defined explicitly at the participant level. It may denote a prior allocation, an observed pre-policy outcome, or a reference outcome for group i. This participant-specific notation makes the admissibility condition directly measurable in empirical applications.

If an aggregate resource constraint is present, the admissible set can be intersected with a feasibility condition:

$$\Sigma_i\ \alpha_i s_i \leq R.$$

where R denotes the total resource envelope. The analytical task is therefore to characterize whether the feasible region contains at least one vector α that satisfies the Pareto-improvement restrictions. The paper does not estimate a single aggregate welfare objective; it defines and operationalizes admissibility conditions for Pareto-preserving reallocations.

3 Statistical characterization

The statistical contribution of the framework is to treat Pareto improvement as an admissibility condition rather than as a final welfare label. In empirical applications, x_i^0 and x_i^1 may be estimated from group means, regression-adjusted predictions, treatment effects, or panel-data comparisons. The observed or estimated gain for group i can be written as:

$$\hat{g}_i = \hat{x}_i^1 - \hat{x}_i^0.$$

A sample-based reallocation is raw-Pareto-admissible when the smallest estimated group gain is non-negative and at least one estimated group gain is strictly positive, that is, $\min_i \hat{g}_i \geq 0$ and $\max_i \hat{g}_i > 0$. A covariate-adjusted reallocation is model-Pareto-admissible if the estimated group-specific effects satisfy the same inequalities after controlling for relevant covariates. Because estimates are uncertain, the empirical conclusion should distinguish between descriptive admissibility, model-adjusted admissibility, and statistical precision.

4 Real-data statistical application: RAND Health Insurance Experiment

4.1 Data and empirical interpretation

The empirical application uses the RAND Health Insurance Experiment data distributed with the statsmodels Python package (Statsmodels Developers, 2026). The dataset contains 20,190 observations and includes outpatient visits to medical doctors, coinsurance rates, plan characteristics, health indicators, chronic disease counts, and physical limitation status. The RAND HIE is appropriate for an illustrative application because it explicitly varies the generosity of health-insurance cost sharing, which can be interpreted as a resource-allocation regime affecting access to care.

In this application, health-status groups are treated as the participants i in the theoretical framework. The baseline allocation x_i^0 is the mean number of medical-doctor visits for group i under high cost sharing, while the post-reallocation outcome x_i^1 is the mean number of visits under generous cost sharing. The empirical gain is therefore the difference in mean utilization between generous and high cost-sharing regimes. This is a real-data statistical application, not a claim that medical visits are a complete measure of welfare.

The empirical exercise operationalizes the admissibility condition at the outcome level. It does not claim to estimate a structural allocation weight for each individual; rather, it asks whether the observed change associated with a more generous allocation regime satisfies the same non-deterioration and strict-improvement restrictions defined by the theoretical framework.

Table 1. Variables used in the real-data application

Variable	Interpretation in the paper
mdvis	Observed medical-doctor outpatient visits; empirical access outcome
lncoins	Log of coinsurance plus one; proxy for cost-sharing intensity
coinsurance_pct	Recovered coinsurance rate in percent: exp(lncoins)-1
idp	Indicator for individual deductible plan
lpi	Log of participation incentive payment
fmde	Medical deductible expenditure variable in the RAND data
physlm	Indicator for physical limitation
disea	Number of chronic diseases
hlthg, hlthf, hlthp	Self-rated health categories; excellent is omitted reference category

Table 2. Descriptive statistics by cost-sharing regime

plan_group	observations	mean_md_visits	median_md_visits	mean_coinsurance	fair_or_poor_share
Generous (0-25%)	15062	3.033	2.000	6.747	0.095
High (51-75%)	1401	2.561	1.000	50.000	0.086
Very high (76-100%)	3727	2.276	1.000	96.441	0.085

Note: The moderate category is absent in the loaded RAND HIE sample because the reconstructed coinsurance values fall mainly into 0, 25, 50, and 95-100 percent regimes.

4.2 Raw Pareto-admissibility check

The raw comparison evaluates whether moving from high cost sharing to generous cost sharing improves access without reducing average observed access for any health-status group. Table 3 shows that the strict-improvement condition is satisfied for several groups. However, the no-deterioration condition is not fully satisfied because the poor-health group shows a small negative raw difference. This is substantively important: the framework does not simply celebrate average improvement; it identifies the exact subgroup for which the Pareto constraint is violated.

Table 3. Raw Pareto-admissibility check by health-status group

Health-status group	Mean visits: generous plan	Mean visits: high cost sharing	Raw gain	No-deterioration condition	n generous	n high cost sharing
Excellent	2.779	2.215	0.564	Satisfied	8189	2830
Good	3.124	2.254	0.870	Satisfied	5448	1861
Fair	3.821	3.280	0.541	Satisfied	1189	371
Poor	5.758	5.924	-0.166	Not satisfied	236	66

Figure 1. Mean outpatient visits by health status and cost-sharing regime

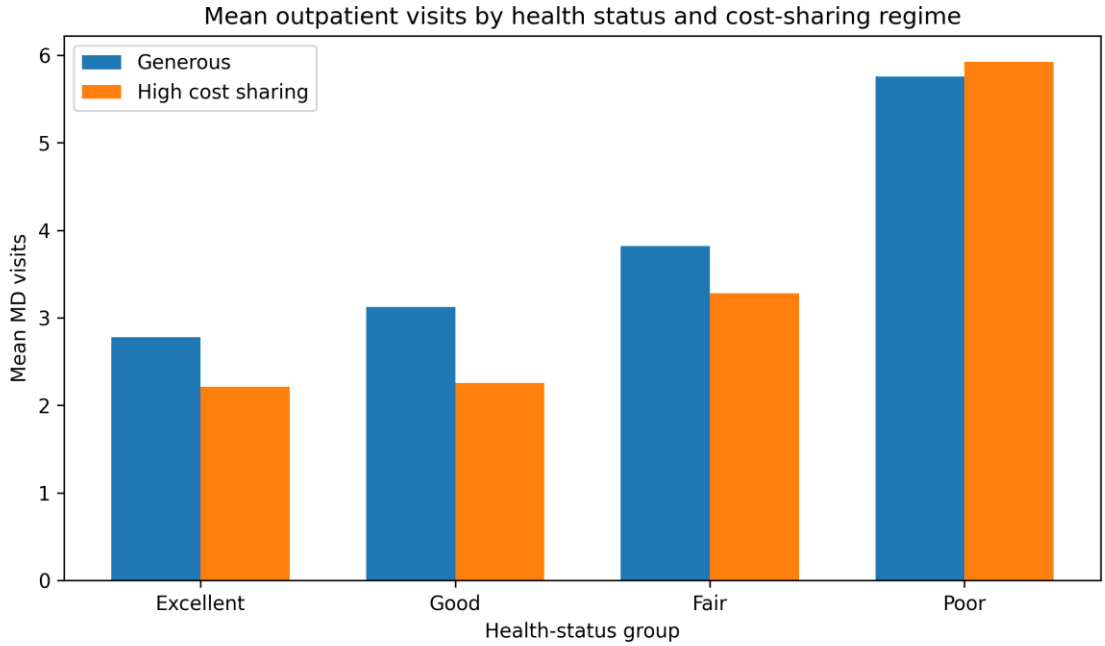
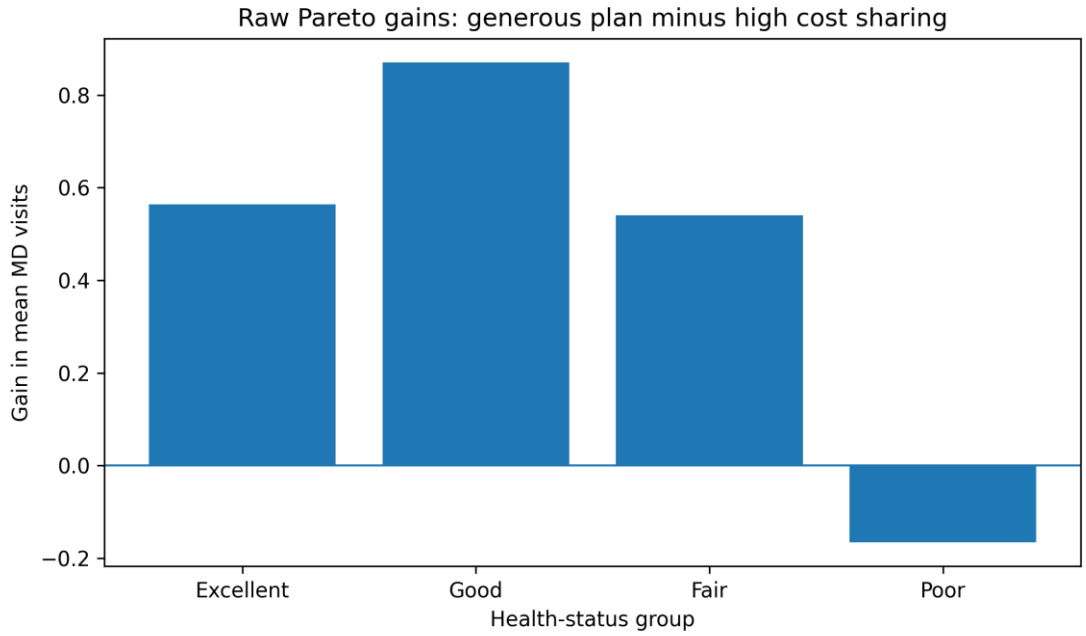


Figure 2. Raw Pareto gains: generous plan minus high cost sharing



4.3 Covariate-adjusted access effects

The raw comparison may be affected by observable differences in health limitations, disease burden, plan type, and participation incentives. To provide a model-adjusted statistical check, the following specification is estimated with heteroskedasticity-robust standard errors:

$$log(1 + mdvis_i) = \beta_0 + \beta_1 Generous_i + \beta_2 Health_i + \beta_3 (Generous_i \times Health_i) + \gamma Z_i + \varepsilon_i.$$

The vector Z_i includes physical limitation, number of chronic diseases, individual deductible plan status, and participation incentive. The group-specific adjusted effect of generous cost sharing is then interpreted as a model-based estimate of the access gain for each health-status group.

Table 4. Covariate-adjusted access effects of generous cost sharing

Health-status group	Adjusted effect of generous plan on log(1+visits)	Robust SE	Lower 95% CI	Upper 95% CI
Excellent	0.207	0.018	0.172	0.241
Good	0.261	0.022	0.219	0.304
Fair	0.111	0.051	0.010	0.212
Poor	0.023	0.141	-0.254	0.300

Figure 3. Covariate-adjusted access effects with 95 percent confidence intervals

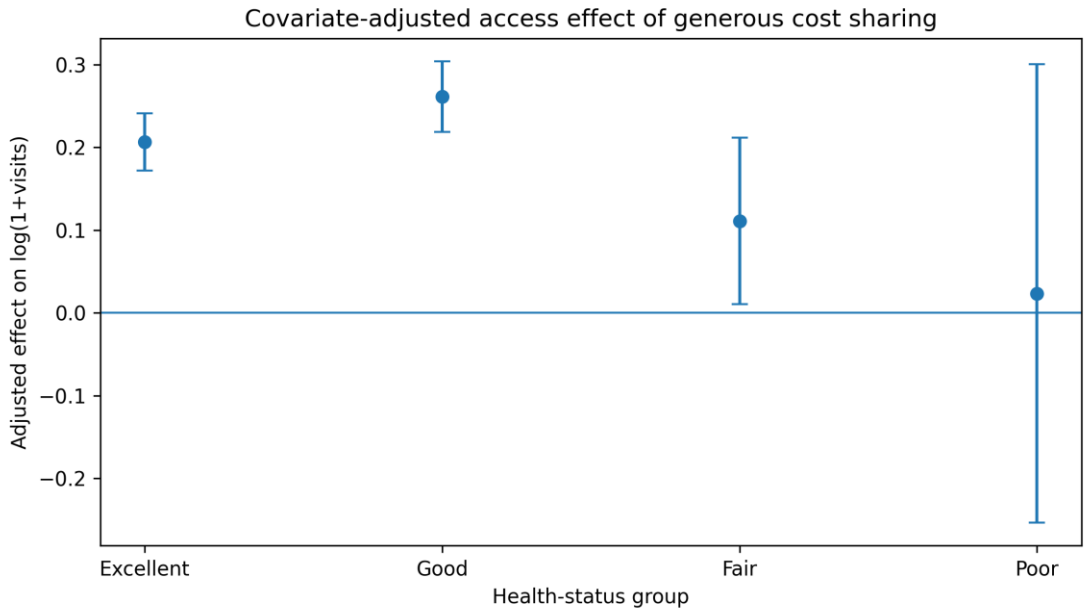


Table 5. Interpretation of the Pareto-admissibility conditions

Condition	Result
Strict improvement for at least one group	Yes: at least three groups show positive raw gains.
No deterioration for all groups	No: the poor-health group has a small negative raw mean difference.
Overall Pareto admissibility of raw reallocation	Not fully satisfied in raw group means; the framework correctly detects a constraint violation.
Covariate-adjusted interpretation	After controlling for health and plan covariates, the estimated generous-plan effect is positive for all groups, although the poor-health margin remains imprecisely estimated.

4.4 Discussion of empirical defensibility

This empirical section is more defensible than a simulated illustration because it uses real observations from a recognized social experiment. It nevertheless remains an illustrative application of the admissibility framework rather than a complete welfare evaluation. The outcome variable captures utilization/access, not utility. Moreover, a higher number of visits is not always welfare-improving if some visits are medically unnecessary. For this reason, the empirical claims are carefully limited: the section shows how the proposed Pareto-admissibility restrictions can be operationalized using real data, and it demonstrates that average gains can coexist with subgroup-level deterioration.

The main empirical conclusion is therefore methodological. The framework is useful because it prevents the analyst from confusing average improvement with Pareto improvement. In the RAND HIE application, a generous insurance allocation increases access on average and in adjusted models, but the raw subgroup comparison shows that strict Pareto admissibility is not automatically guaranteed. This result supports the value of explicitly characterizing admissible weights and vulnerable-group constraints before describing a reallocation as welfare-improving.

5 Conclusion

This paper has developed a statistical framework for identifying Pareto-improving resource reallocations through admissible allocation weights. The framework defines gains relative to a baseline, imposes strict improvement for at least one participant, and requires non-deterioration for all participants. The real-data application using the RAND Health Insurance Experiment demonstrates how these conditions can be implemented in practice. The findings show that a reallocation can improve average access while still failing a strict Pareto condition for a vulnerable subgroup. This is precisely the type of distinction that the proposed framework is designed to reveal.

Future research can extend the framework in three directions. First, the admissible set can be estimated under uncertainty using confidence regions or Bayesian posterior probabilities. Second, the approach can be generalized to panel data and dynamic reallocations. Third, the framework can be applied to fiscal transfers, education spending, social insurance, regional policy, and public health resource allocation, where non-deterioration constraints are normatively and politically central.

Acknowledgements: The author thanks colleagues and seminar participants for helpful discussions on welfare economics, allocation theory, and empirical policy evaluation.

References

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

Atkinson, A. B. (1970). On the measurement of inequality. Journal of Economic Theory, 2(3), 244-263.

Atkinson, A. B., and Stiglitz, J. E. (1976). The design of tax structure: Direct versus indirect taxation. Journal of Public Economics, 6(1-2), 55-75.

Bergson, A. (1938). A reformulation of certain aspects of welfare economics. Quarterly Journal of Economics, 52(2), 310-334.

Brook, R. H., Keeler, E. B., Lohr, K. N., Newhouse, J. P., and Ware, J. E. (2006). The Health Insurance Experiment: A Classic RAND Study Speaks to the Current Health Care Reform Debate. RAND Corporation.

Debreu, G. (1959). Theory of Value: An Axiomatic Analysis of Economic Equilibrium. Yale University Press.

Diamond, P. A., and Mirrlees, J. A. (1971). Optimal taxation and public production I: Production efficiency. American Economic Review, 61(1), 8-27.

Fleurbaey, M., and Maniquet, F. (2011). A Theory of Fairness and Social Welfare. Cambridge University Press.

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

Hicks, J. R. (1939). The foundations of welfare economics. Economic Journal, 49(196), 696-712.

Kaldor, N. (1939). Welfare propositions of economics and interpersonal comparisons of utility. Economic Journal, 49(195), 549-552.

Little, I. M. D. (1950). A Critique of Welfare Economics. Oxford University Press.

Manning, W. G., Newhouse, J. P., Duan, N., Keeler, E. B., Leibowitz, A., and Marquis, M. S. (1987). Health insurance and the demand for medical care: Evidence from a randomized experiment. American Economic Review, 77(3), 251-277.

Mas-Colell, A., Whinston, M. D., and Green, J. R. (1995). Microeconomic Theory. Oxford University Press.

Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. Review of Economic Studies, 38(2), 175-208.

Newhouse, J. P. (1993). Free for All? Lessons from the RAND Health Insurance Experiment. Harvard University Press.

Pareto, V. (1906). Manuale di Economia Politica. Societa Editrice Libraria.

Rawls, J. (1971). A Theory of Justice. Harvard University Press.

Saez, E. (2001). Using elasticities to derive optimal income tax rates. Review of Economic Studies, 68(1), 205-229.

Samuelson, P. A. (1947). Foundations of Economic Analysis. Harvard University Press.

Sen, A. K. (1970). Collective Choice and Social Welfare. Holden-Day.

Statsmodels Developers. (2026). RAND Health Insurance Experiment Data. statsmodels.datasets.randhie documentation and Python package dataset.

Varian, H. R. (1974). Equity, envy, and efficiency. Journal of Economic Theory, 9(1), 63-91.

Appendix A. Reproducibility note

The accompanying Python file pareto_real_data_code_pycharm2022.py reproduces all empirical tables and figures using real RAND HIE data from statsmodels. The script does not require manually downloaded CSV files. It creates a results folder containing CSV tables and PNG figures.