Institute for Economic Studies, Keio University

Keio-IES Discussion Paper Series

Fragility of Joint Identification in the Roy Model : A Note on Deterministic Sorting and Stochastic Selection

星野崇宏、篠田和彦、大津泰介

2026 年 5 月 10 日
DP2026-009
https://ies.keio.ac.jp/publications/27620/

Institute for Economic Studies, Keio University
2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan
ies-office-group@keio.jp
10 May, 2026

Fragility of Joint Identification in the Roy Model : A Note on Deterministic Sorting and Stochastic Selection

星野崇宏、篠田和彦、大津泰介

IES Keio DP2026-009
2026 年 5 月 10 日
JEL Classification: C14, C21, C36, J24
キーワード: Roy model, generalized Roy model, joint distribution, selection on gains, partial identification, sensitivity analysis

【要旨】

Heckman and Honoré (1990) identify the joint distribution of latent sector-specific outcomes in a nonparametric Roy model. This note isolates the measurement condition behind that result. The identifying force is not the exclusion restriction alone, but the exclusion restriction combined with the known deterministic observation rule $D = 1\{Y_1 > Y_0\}$, which reveals the side of a moving boundary in the latent plane. The main result shows that this rule is a singular measurement condition. Under a local η-relaxation of deterministic Roy sorting, an unknown stochastic selection kernel opens an absorption face of observationally equivalent latent joint laws whose total-variation diameter is linear in η, and this linear rate is exact on the face. A point-to-set corollary formalizes the boundary: at $\eta = 0$ the Heckman–Honoré restrictions deliver a singleton joint law, whereas every total-variation neighborhood of the exact Roy observed law contains stochastic-Roy laws whose latent identified sets have first-order diameter. The note also records the sharp global geometry that remains after deterministic sorting is relaxed: the latent-law identified set is a submeasure linear program formed from treated and untreated joint submeasures. This global characterization is supporting geometry for the fragility result, not the main message. The results distinguish deterministic Roy tomography from ordinary partial identification u nder stochastic selection.

星野崇宏
慶應義塾大学経済学部
bayesian@keio.jp

篠田和彦
名古屋大学
k-shinoda@nagoya-u.jp

大津泰介
London School of Economics and Political Science
T.Otsu@lse.ac.uk

謝辞：This study is supported by JSPS Kakenhi 25K21651.

Fragility of Joint Identification in the Roy Model

A Note on Deterministic Sorting and Stochastic Selection

Takahiro Hoshino, Kazuhiko Shinoda and Taisuke Otsu

May 10, 2026

KEO Discussion Paper No. 201

Fragility of Joint Identification in the Roy Model

A Note on Deterministic Sorting and Stochastic Selection

Takahiro Hoshino

Faculty of Economics, Keio University; RIKEN Center for Advanced Intelligence Project

hoshino@econ.keio.ac.jp

Kazuhiko Shinoda

Graduate School of Economics, Nagoya University

Taisuke Otsu

Department of Economics, London School of Economics and Political Science

t.otsu@lse.ac.uk

This version: May 9, 2026

Abstract

Heckman and Honoré (1990) identify the joint distribution of latent sector-specific outcomes in a nonparametric Roy model. This note isolates the measurement condition behind that result. The identifying force is not the exclusion restriction alone, but the exclusion restriction combined with the known deterministic observation rule $D = \mathbf{1}\{Y_1 > Y_0\}$, which reveals the side of a moving boundary in the latent plane. The main result shows that this rule is a singular measurement condition. Under a local η-relaxation of deterministic Roy sorting, an unknown stochastic selection kernel opens an absorption face of observationally equivalent latent joint laws whose total-variation diameter is linear in η, and this linear rate is exact on the face. A point-to-set corollary formalizes the boundary: at $\eta = 0$ the Heckman–Honoré restrictions deliver a singleton joint law, whereas every total-variation neighborhood of the exact Roy observed law contains stochastic-Roy laws whose latent identified sets have first-order diameter. The note also records the sharp global geometry that remains after deterministic sorting is relaxed: the latent-law identified set is a submeasure linear program formed from treated and untreated joint submeasures. This global characterization is supporting geometry for the fragility result, not the main message. The results distinguish deterministic Roy tomography from ordinary partial identification under stochastic selection.

Keywords: Roy model; generalized Roy model; joint distribution; selection on gains; partial identification; sensitivity analysis.

JEL codes: C14, C21, C36, J24.

1 Introduction

A Roy economy is a missing-data problem with a special measurement device. Each individual has latent sectoral earnings or potential outcomes (Y_1, Y_0), but the econometrician observes only $Y = DY_1 + (1 - D)Y_0$. What makes the classical model empirically powerful is the deterministic choice rule

$$D = \mathbf{1}\{Y_1 > Y_0\}. \tag{1}$$

The selection indicator D therefore reveals the latent ranking between the two potential outcomes. In this setup, Heckman and Honoré (1990, Theorem 12) consider the latent outcome models

$$Y_1 = g_1(X_1, X_0) + \varepsilon_1, \qquad Y_0 = g_0(X_2, X_0) + \varepsilon_0, \tag{2}$$

where $(\varepsilon_1, \varepsilon_0)$ is independent of $X = (X_0, X_1, X_2)$, the excluded shifters move locations over full support, and suitable location normalizations are imposed. Under (1), the observables reveal probabilities over truncation regions such as

$$\mathbb{P}(Y_1 > Y_0 \mid X), \quad \mathbb{P}(Y_1 \le y,\, Y_1 > Y_0 \mid X), \quad \mathbb{P}(Y_0 \le y,\, Y_0 > Y_1 \mid X),$$

for given y. Varying the shifters moves the boundary $Y_1 = Y_0$ through the latent plane. The deterministic selection therefore converts observed choices into measurements over moving known regions, which identifies the joint distribution of the latent errors.

This note asks what remains when selection is still allowed to depend on both potential outcomes, but the Roy comparison rule is no longer deterministic. In particular, we consider

$$D \mid (Y_1, Y_0, X) \sim \text{Bernoulli}\{\pi(Y_1, Y_0, X)\}, \tag{3}$$

where π is an unknown selection kernel. This formulation allows participation to depend not only on sectoral gains, but also on costs, preferences, information, or other non-pecuniary factors. The indicator D no longer reveals the sign of $Y_1 - Y_0$. The truncation geometry underlying the classical Roy model therefore disappears: observed choices become unknown weighted coarsenings of the latent full-data distribution.

The main result is a quantitative boundary statement. Let η measure a local failure of the Roy ordering measurement: above the Roy boundary the probability of not choosing sector 1 is of order η, while below the boundary the probability of choosing sector 1 is also of order η. Theorem 1 shows that, for every $\eta > 0$, the same observed law can be generated by two latent joint distributions separated by $C\eta$ in total variation. Moreover, this linear order is sharp on the constructed absorption face. Furthermore, Corollary 4 converts this into a point-to-set boundary statement. Under the exact Roy model with $(\eta = 0)$, the latent law is point identified. But every total-variation neighborhood of the exact Roy observed law contains stochastic-Roy economies whose latent identified sets open at first order. Thus the exact Roy rule is not a

regular propensity-score restriction. It is a singular measurement condition.

Two supporting results clarify the mechanism behind the fragility result. First, Lemma 1 and Proposition 3 characterize how an unknown stochastic kernel can absorb perturbations of the latent joint law while preserving the observed treated and untreated outcome distributions. The proposition gives an if-and-only-if characterization on a local absorption slice, showing that the phenomenon is structural rather than a simple dimensional artifact. Second, Theorem 2 records the sharp global object that remains once deterministic sorting is relaxed. The resulting identified set is a submeasure linear program over treated and untreated joint submeasures. These supporting results isolate the geometry underlying the main theorem. The contribution of the note is the fragility of Roy joint identification, not the general observation that stochastic missing-data models are partially identified.

The note is deliberately narrower than a general theory of Roy bounds. Sharp ATE bounds for binary and atomless continuous outcomes are collected in Online Appendix OA.1. Additional numerical checks and an NSW illustration are also reported in the online appendix. These results help calibrate the quantitative content of the fragility theorem and connect the analysis to partial-identification literatures, but they are not needed for the main result. In particular, the NSW exercise is indexed by the overlap parameter κ, not by the local Roy-relaxation parameter η. An empirical η-curve would require additional structural assumptions linking observed choices to deviations from deterministic Roy sorting.

Relation to the literature. The paper contributes to the empirical-content interpretation of the Roy model of Roy (1951) and Heckman and Honoré (1990). Generalized and extended Roy models, including Heckman and Vytlacil (2005), Eisenhauer, Heckman, and Vytlacil (2015), Heckman, Lopes, and Piatek (2014), and Lee and Park (2023), impose structure on choice or latent factors to recover economically meaningful objects. Roy sharp-bounds analyses such as Mourifié and Henry (2020) maintain deterministic or monotone Roy restrictions that carry testable empirical content. By contrast, this note studies the measurement-theoretic gap created when deterministic ordering revelation is replaced by an unrestricted stochastic kernel. Sensitivity analyses such as Masten and Poirier (2018) and Masten and Poirier (2020) relax ignorability or related treatment-assignment assumptions; here η relaxes a different object, the Roy ordering measurement itself. Partial-identification tools in Manski (1990), Bhattacharya, Shaikh, and Vytlacil (2012), Beresteanu, Molchanov, and Molinari (2012), and related work provide the language for the global submeasure set, while the main theorem isolates the singularity of deterministic Roy tomography.

2 Deterministic Roy sorting as known-set measurement

The identification argument in Heckman and Honoré (1990, Theorem 12) can be interpreted as a tomography problem. Under deterministic Roy sorting, the observables reveal probabilities

Literature	Object	Difference here
Heckman–Honoré Roy identification	Joint latent law under deterministic sorting	Shows that the deterministic rule is a known-set measurement and quantifies the loss when it is locally randomized.
Mourifié–Henry and Roy bounds	Sharp bounds under Roy or monotone restrictions	The stochastic kernel here is unrestricted; the Roy measurement restriction is the object being relaxed.
Generalized/extended Roy models	Benefits, costs, or latent factors under structured choice	The note does not replace structure with a new choice model; it studies what happens when the ordering measurement is removed.
Masten–Poirier sensitivity	Treatment effects under relaxations of ignorability	The sensitivity parameter here measures failure of deterministic ordering revelation, not hidden confounding or conditional independence failure.
This note	Joint-law fragility at the Roy boundary	The exact Roy rule is a singular measurement condition; local stochastic leakage opens the joint identified set at linear order.

Table 1: Positioning. The contribution is the fragility of deterministic Roy measurement, not a new general partial-identification calculus.

of moving known regions in the latent outcome plane. Once enough such regions are observed, the latent joint law is pinned down. The following proposition states the abstract measurement principle underlying this logic.

Proposition 1 (Known-set tomography)**.** *Let F be a probability measure on a standard Borel space $\mathcal{E}$. Suppose the observables identify $F(A)$ for every set A in a π-system $\mathcal{A}$ generating the Borel σ-field on $\mathcal{E}$. Then F is identified. Hence any observation rule that reveals probabilities of a separating family of known sets identifies the latent distribution.*

Proof. If two probability measures agree on $\mathcal{A}$, the uniqueness part of the extension theorem implies that they agree on $\sigma(\mathcal{A})$. Since $\sigma(\mathcal{A})$ is the Borel σ-field, the two measures coincide. □

Remark 1 (How the principle reads in the Roy model)**.** *Under exact Roy sorting, the sets generated by the moving boundary $Y_1 = Y_0$, together with the truncated events in (6) and (7), form the relevant separating measurement family after location variation. Proposition 1 is therefore a deliberately abstract version of the tomography logic. The failure under stochastic selection is not that the shifters stop moving the latent plane; it is that the observed probabilities are no longer probabilities of known sets in that plane.*

We now specialize the abstract measurement principle to Heckman–Honoré's model. Fix x_0, and write

$$Y_1 = g_1(X_1, x_0) + \varepsilon_1, \qquad Y_0 = g_0(X_2, x_0) + \varepsilon_0, \tag{4}$$

with $(\varepsilon_1, \varepsilon_0) \perp (X_1, X_2)$. Suppose the support of $(g_1(X_1, x_0), g_0(X_2, x_0))$ is $\mathbb{R}^2$, and impose the usual location normalizations. Under the Roy rule $D = \mathbf{1}\{Y_1 > Y_0\}$, the observed law contains the objects

$$A(x_1, x_2) = \mathbb{P}\{g_1(x_1, x_0) + \varepsilon_1 > g_0(x_2, x_0) + \varepsilon_0\}, \tag{5}$$

$$B(y, x_1, x_2) = \mathbb{P}\{g_1(x_1, x_0) + \varepsilon_1 \leq y, g_1(x_1, x_0) + \varepsilon_1 > g_0(x_2, x_0) + \varepsilon_0\}, \tag{6}$$

$$C(y, x_1, x_2) = \mathbb{P}\{g_0(x_2, x_0) + \varepsilon_0 \leq y, g_0(x_2, x_0) + \varepsilon_0 > g_1(x_1, x_0) + \varepsilon_1\}. \tag{7}$$

These are the probabilities used in the proof of Heckman and Honoré (1990, Theorem 12). The event in (5) is $\varepsilon_1 - \varepsilon_0 > g_0(x_2, x_0) - g_1(x_1, x_0)$. Thus varying (x_1, x_2) moves a known half-space boundary through the joint distribution of $(\varepsilon_1, \varepsilon_0)$. The events in (6) and (7) add truncated marginal information on each side of the boundary.

Proposition 2 (Roy tomography)**.** *Under (4), full support of the two location shifters, independence of $(\varepsilon_1, \varepsilon_0)$ from (X_1, X_2), and the deterministic Roy rule, the observed probabilities (5)–(7) are moving truncation probabilities for the latent joint distribution. The sector indicator supplies the side of the boundary $Y_1 = Y_0$, and the excluded shifters move that boundary. This is the geometric and informational mechanism behind the identification of $(g_1, g_0, G_{\varepsilon_1, \varepsilon_0})$ in Heckman and Honoré's Theorem 12.*

Proof sketch. The equality sets of (5) identify relative locations because

$$\mathbb{P}\{g_1(x_1, x_0) + \varepsilon_1 > g_0(x_2, x_0) + \varepsilon_0\}$$

depends only on the scalar shift $g_1(x_1, x_0) - g_0(x_2, x_0)$. Full support recovers these shifts up to normalizations. Conditional on such a recovered shift, (6) and (7) identify probabilities over regions truncated by $\varepsilon_1 - \varepsilon_0$ exceeding or falling below a known threshold. As the threshold ranges over $\mathbb{R}$, the induced regions sweep the latent plane. □

The proposition is not a new proof of Heckman and Honoré. It is a reinterpretation designed to isolate the informational mechanism behind their identification argument. The deterministic rule does not merely describe behavior. It generates a measurement of the latent ordering, and hence information about the copula of (Y_1, Y_0).

Remark 2 (Multi-sector deterministic Roy sorting)**.** *The same known-set measurement logic is not limited to two sectors. Suppose, for example, that there are three latent outcomes*

$$Y_j = m_j(X_j) + \varepsilon_j, \qquad j = 0, 1, 2,$$

and that the econometrician observes the chosen sector and its outcome under deterministic sorting,

$$D = \arg\max_{j \in \{0,1,2\}} Y_j, \qquad Y = Y_D,$$

with ties having probability zero. Then the event $D = 0$ and $Y \leq y$ is the known polyhedral truncation event

$$\{Y_0 \leq y,\, Y_1 - Y_0 \leq 0,\, Y_2 - Y_0 \leq 0\}.$$

After the location shifts are varied, this gives lower-orthant probabilities for

$$(\varepsilon_0,\, \varepsilon_1 - \varepsilon_0,\, \varepsilon_2 - \varepsilon_0),$$

which is a one-to-one transformation of $(\varepsilon_0, \varepsilon_1, \varepsilon_2)$. Hence, under sector-specific excluded shifters, full support, and the usual location normalizations, deterministic argmax sorting can identify the joint distribution of (Y_0, Y_1, Y_2). The loser–loser ranking need not be directly observed; moving the known polyhedral truncation regions supplies the missing information. If the argmax rule is replaced by an unknown stochastic selection kernel, these known polyhedral probabilities are again replaced by weighted projections, and the fragility mechanism studied below reappears.

3 Stochastic selection and absorption

3.1 Weighted subdensities and copula absorption

We now relax the deterministic Roy measurement rule and allow selection to occur through an unknown stochastic kernel. The key change is that the observables no longer reveal probabilities of known truncation regions. Instead, they reveal only weighted marginal projections of latent subdensities. Replace the deterministic rule by

$$\mathbb{P}(D = 1 \mid Y_1, Y_0, X = x) = \pi_x(Y_1, Y_0), \qquad 0 < \pi_x(Y_1, Y_0) < 1. \tag{8}$$

Then the observed conditional subdistributions are

$$\mathbb{P}(D = 1, Y \leq y \mid X = x) = \mathbb{E}\big[\pi_x(Y_1, Y_0)\mathbf{1}\{Y_1 \leq y\} \mid X = x\big], \tag{9}$$

$$\mathbb{P}(D = 0, Y \leq y \mid X = x) = \mathbb{E}\big[\{1 - \pi_x(Y_1, Y_0)\}\mathbf{1}\{Y_0 \leq y\} \mid X = x\big]. \tag{10}$$

These are weighted marginals of the latent joint distribution. The known truncation sets (5)–(7) have been replaced by marginal projections of unknown subdensities

$$H_{1,x}(y_1, y_0) = \pi_x(y_1, y_0) f_x(y_1, y_0), \qquad H_{0,x}(y_1, y_0) = \{1 - \pi_x(y_1, y_0)\} f_x(y_1, y_0),$$

where f_x denotes the conditional joint density of (Y_1, Y_0) given $X = x$. Before imposing a Roy-boundary relaxation, it is useful to state the sharp local geometry of stochastic selection. To isolate the local geometry, fix the shifters and write

$$H_1(y_1, y_0) = \pi(y_1, y_0) f(y_1, y_0), \qquad H_0(y_1, y_0) = \{1 - \pi(y_1, y_0)\} f(y_1, y_0),$$

Deterministic Roy sorting	**Stochastic selection**
$D = \mathbf{1}\{Y_1 > Y_0\}$ reveals a known side of the boundary $Y_1 = Y_0$. With location shifters, the econometrician observes probabilities over known moving truncation sets.	$D \mid (Y_1, Y_0, X) \sim \text{Bernoulli}\{\pi(Y_1, Y_0, X)\}$ reveals only marginal projections of unknown weighted submeasures $H_{1,x} = \pi_x f_x$ and $H_{0,x} = (1 - \pi_x) f_x$.

Figure 1: Roy tomography versus stochastic weighting. In Heckman–Honoré, the observed probabilities are integrals over known truncation regions. With an unrestricted overlapping stochastic kernel, the observed objects are weighted marginal projections, and the boundary is not observed.

for the treated and untreated joint subdensities of (Y_1, Y_0). The observed treated outcome law is the y_1-marginal of H_1, while the observed untreated outcome law is the y_0-marginal of H_0.

The key observation is that stochastic selection creates local directions along which the latent joint law can change without changing the observed law. The next proposition characterizes these observationally equivalent directions exactly within a local absorption slice.

Proposition 3 (Sharp local absorption slice)**.** *Fix a rectangle B. Consider local alternatives that leave H_0 fixed and perturb the treated subdensity only on B:*

$$H_1^h = H_1 + h, \qquad H_0^h = H_0, \qquad f^h = f + h, \qquad \pi^h = \frac{H_1 + h}{f + h},$$

where h is an integrable signed function supported on B, and where positivity requires $H_1 + h \geq 0$ and $f + h > 0$. Within this local slice, the alternative model generates the same observed law of (D, Y) if and only if *the marginal-vanishing condition*

$$\int h(y_1, y_0)\, dy_0 = 0 \qquad \text{for every } y_1 \tag{11}$$

holds. Hence the observationally equivalent directions in this local slice are exactly the zero-Y_0-marginal perturbations of the treated subdensity; no other direction in the slice is observationally equivalent. If, in addition,

$$\int h(y_1, y_0)\, dy_1 = 0 \qquad \text{for every } y_0, \tag{12}$$

then the perturbation preserves both marginal distributions and changes only the copula. If (12) fails, the same observed law can be generated while the latent Y_0-marginal, and hence the ATE, changes. A symmetric statement holds if H_1 is held fixed and H_0 is perturbed.

The proof is in Appendix A. The proposition also gives the following anatomy of the information loss created by stochastic sorting. This is the positive companion to the negative absorption result: the stochastic model opens different directions, and the Roy measurement rule closes all of them at once.

Corollary 1 (Anatomy of information loss)**.** *The local stochastic-selection slice in Proposition 3 contains two conceptually distinct observationally equivalent directions.*

(i) Copula directions. *If the perturbation satisfies both zero-marginal restrictions (11)–(12), then the observed law and the two latent marginal distributions are unchanged, while functionals depending on the joint dependence structure can change.*

(ii) Marginal directions. *If only the restriction required for the observed marginal is imposed, the same observed law can be generated while one latent marginal, and hence the ATE, changes. The symmetric construction perturbs the other latent marginal.*

Thus deterministic Roy sorting is not merely one additional exclusion restriction. It is a measurement rule that closes both the copula direction and the marginal direction. Once sorting is stochastic and the kernel is unknown, those directions reopen unless extra functional restrictions are imposed.

The anatomy in Corollary 1 characterizes the directions opened by stochastic selection at fixed shifters. The next lemma shows that the same absorption mechanism operates uniformly across the Heckman–Honoré location-shifter environment: a single zero-marginal perturbation of the latent error density induces simultaneously absorbing kernels at every value of X.

Lemma 1 (Copula absorption)**.** *Let*

$$Y_1 = g_1(X_1, X_0) + \varepsilon_1, \qquad Y_0 = g_0(X_2, X_0) + \varepsilon_0, \tag{13}$$

where $(\varepsilon_1, \varepsilon_0) \perp X$ has density f. Assume the support of $(g_1(X_1, x_0), g_0(X_2, x_0))$ is $\mathbb{R}^2$ for each x_0. Let

$$\mathbb{P}(D = 1 \mid \varepsilon_1 = e_1, \varepsilon_0 = e_0, X = x) = \pi_x(e_1, e_0), \qquad \kappa \le \pi_x(e_1, e_0) \le 1 - \kappa$$

for some $\kappa \in (0, 1/2)$. Suppose there is an open rectangle $B \subset \mathbb{R}^2$ on which $f \ge \underline{f} > 0$. Let a be a nonzero bounded function supported on B such that

$$\int a(e_1, e_0)\, de_0 = 0 \quad \text{for all } e_1, \qquad \int a(e_1, e_0)\, de_1 = 0 \quad \text{for all } e_0. \tag{14}$$

Then for all sufficiently small $|\delta| > 0$, there exists another model with the same location functions (g_1, g_0), perturbed joint error density

$$f_\delta(e_1, e_0) = f(e_1, e_0) + \delta a(e_1, e_0),$$

and an overlapping stochastic selection kernel $\pi_{\delta,x}$, such that the conditional observed law of (D, Y) given $X = x$ is exactly the same for every x. The two marginal error distributions are unchanged, but the joint distribution is changed.

Proof sketch. Shift the perturbation a according to the location functions and define

$$H_{1,\delta,x} = H_{1,x} + \delta a_x, \qquad H_{0,\delta,x} = H_{0,x}.$$

The zero-y_1-marginal condition on a_x preserves the observed treated outcome law, while the observed untreated outcome law is unchanged because H_0 is held fixed. The perturbed kernel is recovered from

$$\pi_{\delta,x} = \frac{H_{1,\delta,x}}{H_{1,\delta,x} + H_{0,\delta,x}}.$$

For sufficiently small $|\delta|$, positivity and overlap are preserved. The second zero-marginal condition ensures that the two latent marginal distributions are unchanged, so the perturbation changes only the copula. The full verification is given in Appendix A. □

The lemma is local by construction. It is stronger than a dimensional count because it constructs observationally equivalent models in every regular neighborhood of a baseline stochastic kernel. It also shows why the deterministic Roy rule is not merely a limiting value of a regular propensity score. At the exact Roy rule, the subdensity H_1 is forced to be the restriction of f to a known half-space. Under an unknown stochastic kernel, H_1 is an unknown component of f, and local copula changes can be reallocated to that component.

Remark 3 (Interior absorption versus boundary fragility)**.** *Lemma 1 and Theorem 1 address different regions of the selection model. Lemma 1 establishes absorption around any regular overlapping stochastic kernel, an interior point of the overlap class. Theorem 1 below establishes absorption around the deterministic Roy rule, which lies on the boundary of every overlap class with $\kappa > 0$. The two results are complementary: Lemma 1 shows that absorption is a generic feature of stochastic selection, while Theorem 1 shows that exact deterministic sorting is a singular boundary point with a linear opening rate.*

4 Main result: Fragility of Roy joint identification

We now study the boundary point that matters for Heckman–Honoré identification: the exact Roy measurement rule. The theorem below considers stochastic kernels that are locally close to the deterministic Roy rule but not identical to it, and shows that the diameter of the nonidentified joint-law set opens at least linearly in the relaxation size. The result is intentionally local. It does not assert a universal Hausdorff upper bound over arbitrary neighborhoods of the deterministic rule. Rather, it shows that the exact Roy rule is a singular measurement condition: once a positive unknown stochastic component is admitted, the same location-shifter exclusions no longer protect joint identification.

Definition 1 (Local η-relaxed Roy class)**.** *Fix constants $0 < c_L < 1 < c_U < \infty$. Let B be a rectangle lying strictly on one side of the Roy boundary. For $\eta > 0$, a stochastic kernel π belongs*

to $\mathcal{M}_\eta(B; c_L, c_U)$ if the following local restriction holds on B:

$$c_L\eta \le 1 - \pi(y_1, y_0) \le c_U\eta, \quad B \subset \{y_1 > y_0\},$$
$$c_L\eta \le \pi(y_1, y_0) \le c_U\eta, \qquad B \subset \{y_1 < y_0\}.$$

Outside B, the kernel is unrestricted except for measurability and values in $[0,1]$. The constants only define an order-η local relaxation of deterministic Roy sorting; they are not substantive. A canonical choice is $c_L = 1/2$ and $c_U = 2$, which gives the symmetric local restriction $\eta/2 \le 1-\pi \le 2\eta$ above the boundary and $\eta/2 \le \pi \le 2\eta$ below it. A global version can be obtained by imposing the same order-η closeness outside B, but the local definition is sufficient for the fragility claim.

Let B be a rectangle strictly above the Roy boundary $Y_1 = Y_0$ and suppose the latent density f is bounded below by $\underline{f} > 0$ on B. The proof for a rectangle below the boundary is obtained by reversing treatment and nontreatment.

Theorem 1 (Main theorem: singular boundary under η-relaxed Roy sorting)**.** *Let a be a nonzero bounded function supported on B with zero marginals as in (14). Define $M = \|a\|_\infty$ and $A = \int |a(e_1, e_0)|\, de_1 de_0$, and fix constants $0 < c_L < 1 < c_U$. Then there exist constants $\bar{\eta} > 0$ and $\rho > 0$, depending only on (c_L, c_U), such that for every $\eta \in (0, \bar{\eta}]$ with $\delta_\eta = M^{-1}\rho\eta\underline{f}$, there are two models in the local class $\mathcal{M}_\eta(B; c_L, c_U)$ with joint densities*

$$f_+(e_1, e_0) = f(e_1, e_0) + \delta_\eta a(e_1, e_0), \qquad f_-(e_1, e_0) = f(e_1, e_0) - \delta_\eta a(e_1, e_0),$$

that generate exactly the same observed law of (D, Y, X). Their marginal distributions of Y_1 and Y_0 are identical, but their joint distributions satisfy

$$d_{TV}(f_+, f_-) = \delta_\eta A = C_a\eta, \qquad \text{where} \quad C_a = M^{-1}\rho\underline{f}A > 0.$$

Consequently, the latent joint identified set has total-variation diameter at least $C_a\eta$ at such observed laws.

Proof sketch. Work above the Roy boundary and take a baseline relaxed kernel with

$$H_1 = (1-\eta)f, \qquad H_0 = \eta f$$

on B. For a zero-marginal perturbation a, define

$$f_\pm = f \pm \delta_\eta a, \qquad H_{1,\pm} = H_1 \pm \delta_\eta a, \qquad H_{0,\pm} = H_0,$$

with $\delta_\eta = M^{-1}\rho\eta\underline{f}$ for sufficiently small $\rho > 0$.

Because a has zero e_0-marginal, the observed treated outcome law is unchanged, while the untreated outcome law is unchanged because $H_{0,\pm} = H_0$. The choice probability is also unchanged

because the zero-marginal condition implies $\int a = 0$. The induced kernels satisfy

$$1 - \pi_{\pm} = \frac{H_0}{f_{\pm}},$$

and for sufficiently small ρ, the local order-η bounds defining $\mathcal{M}_\eta(B; c_L, c_U)$ are preserved. Hence the two latent laws are observationally equivalent, and

$$d_{TV}(f_+, f_-) = \delta_\eta \int |a| = C_a \eta.$$

The case below the boundary is symmetric. The full verification is given in Appendix A. □

Remark 4 (Optimizing the perturbation direction)**.** *The constant $C_a = M^{-1}\rho\underline{f}A$ depends on a only through the scale-invariant ratio $A/M = \|a\|_1/\|a\|_\infty$. For a fixed rectangle B, let $\mathcal{A}_B$ be the class of nonzero bounded zero-marginal perturbations supported on B. The largest lower bound delivered by this construction is*

$$C^*(B) = \sup_{a \in \mathcal{A}_B} C_a = \rho\underline{f} \sup_{a \in \mathcal{A}_B} \frac{\|a\|_1}{\|a\|_\infty}.$$

The constant $C^*(B)$ is positive whenever B contains a rectangle supporting a nonzero product perturbation $r(e_1)s(e_0)$ with zero one-dimensional integrals, and finite whenever B has finite Lebesgue measure. Thus the a-dependence of C_a is not a hidden nuisance; it is the local geometry of the absorption direction.

Corollary 2 (Pathwise sharpness on the absorption face)**.** *Fix the baseline law, rectangle B, and zero-marginal perturbation a used in Theorem 1. Let*

$$\mathcal{P}_\eta(a) = \{f_\gamma = f + \gamma a : |\gamma| \le \delta_\eta\},$$

with the corresponding kernels constructed by holding the untreated subdensity fixed and absorbing γa into the treated subdensity. Conditional on this one-dimensional absorption face, the total-variation diameter is exactly

$$\mathrm{diam}_{TV}\{\mathcal{P}_\eta(a)\} = \delta_\eta \int |a|,$$

and for any bounded linear functional $L_h(f) = \int h(y_1, y_0) f(y_1, y_0) d(y_1, y_0)$, the range over this face is

$$L_h(f) + [-\delta_\eta, \delta_\eta] \int h(y_1, y_0) a(y_1, y_0) dy_1 dy_0.$$

Thus the rate is $\Theta(\eta)$ on the local absorption face. This is not a claim that the entire global identified set under every admissible stochastic kernel has diameter $O(\eta)$.

The proof is immediate from the one-dimensional structure of the absorption face and is given in Appendix A. Appendix B develops two further implications of the absorption geometry: a functional-separation result characterizing which latent functionals remain unidentified

under marginal-neutral perturbations, and a local-universality statement showing that every total-variation neighborhood of the exact Roy observed law contains stochastic-Roy laws with a first-order opening of the latent identified set.

5 Global geometry after relaxing Roy sorting

The fragility theorem in Section 4 is local and boundary-focused. This section records the global object that remains once the deterministic Roy measurement rule is replaced by an unrestricted overlapping stochastic kernel. The result is stated briefly because it is supporting geometry, not the main contribution.

The overlap parameter κ below is distinct from η. The parameter η measures local distance from deterministic Roy sorting; κ is a global pointwise overlap bound on the stochastic kernel. The next theorem characterizes the sharp global identified set under unrestricted stochastic selection.

Theorem 2 (Sharp global submeasure characterization)**.** *Let outcomes take values in a compact set $\mathcal{Y} \subset \mathbb{R}$. Let $P_1(A) = \mathbb{P}(D = 1, Y \in A)$ and $P_0(A) = \mathbb{P}(D = 0, Y \in A)$ be the observed treated and untreated outcome subprobability measures. Under an unrestricted stochastic selection kernel, the sharp identified set for the latent joint law F_{Y_1,Y_0} is*

$$\mathcal{I}_F = \{T + U : T, U \geq 0, T_1 = P_1, U_0 = P_0\},$$

where T and U are finite measures on $\mathcal{Y}^2$, T_1 is the Y_1-marginal of T, and U_0 is the Y_0-marginal of U. If the overlap bound $\kappa \leq \pi(Y_1, Y_0) \leq 1 - \kappa$ is imposed, with $\lambda = \kappa/(1 - \kappa)$, the sharp set is obtained by adding

$$\lambda U(A) \leq T(A) \leq \lambda^{-1} U(A) \qquad \text{for every Borel } A \subset \mathcal{Y}^2.$$

Consequently, for any bounded measurable functional $\varphi(y_1, y_0)$, the sharp identified interval for $\int \varphi \, dF_{Y_1,Y_0}$ is obtained by minimizing and maximizing

$$\int_{\mathcal{Y}^2} \varphi(y_1, y_0) \, d(T + U)$$

over this feasible submeasure set.

Proof. Given any compatible stochastic-selection model, define

$$T(A) = \mathbb{P}((Y_1, Y_0) \in A, D = 1), \qquad U(A) = \mathbb{P}((Y_1, Y_0) \in A, D = 0).$$

Then $T_1 = P_1$, $U_0 = P_0$, and the latent joint law is $F = T + U$. If $\kappa \leq \pi(Y_1, Y_0) \leq 1 - \kappa$, then $T = \pi F$ and $U = (1 - \pi)F$, so $\frac{dT}{dU} = \frac{\pi}{1-\pi}$ on sets where U is positive. Hence $\lambda U \leq T \leq \lambda^{-1} U$, where $\lambda = \kappa/(1 - \kappa)$.

Conversely, take any feasible (T, U). Since $T(\mathcal{Y}^2)+U(\mathcal{Y}^2) = P_1(\mathcal{Y})+P_0(\mathcal{Y}) = 1$, $F = T+U$ is a probability measure. Let $\pi = dT/dF$, defined arbitrarily on F-null sets. Then the model with full-data law F and selection kernel π generates the observed treated and untreated submeasures. The measure inequalities imply $\pi \in [\kappa, 1-\kappa]$. Hence the displayed set is necessary and sufficient, and optimizing any linear functional over it gives sharp lower and upper bounds. □

Remark 5 (Convex geometry of the global set)**.** *The feasible set in Theorem 2 is convex because the restrictions $T, U \geq 0$, $T_1 = P_1$, $U_0 = P_0$, and, when imposed, $\lambda U \leq T \leq \lambda^{-1} U$, are linear in the submeasure pair (T, U). The map $(T, U) \mapsto T + U$ is also linear, so $\mathcal{I}_F$ is convex. Therefore sharp bounds on any bounded linear functional of the latent joint law are attained at extreme points of the feasible submeasure set whenever the usual compactness conditions hold. In finite partitions, these extreme points are the vertices of the linear program; the binary bounds in Online Appendix OA.1.1 are the two-point version of this geometry.*

The local η-relaxed Roy class considered in Section 4 corresponds to adding local odds-ratio restrictions to this global submeasure geometry.

Corollary 3 (Global envelope for local η-relaxed Roy sorting)**.** *For the local class $\mathcal{M}_\eta(B; c_L, c_U)$, the sharp global set is obtained from Theorem 2 by adding the local odds-ratio restrictions implied by Definition 1. If $B \subset \{y_1 > y_0\}$, then on B*

$$\frac{1 - c_U\eta}{c_U\eta}\, U(A \cap B) \leq T(A \cap B) \leq \frac{1 - c_L\eta}{c_L\eta}\, U(A \cap B)$$

for every Borel set A. If $B \subset \{y_1 < y_0\}$, the endpoints are $\frac{c_L\eta}{1-c_L\eta}$ and $\frac{c_U\eta}{1-c_U\eta}$ for $0 < \eta < 1/c_U$, so that $c_U\eta < 1$ and the displayed odds ratios are well defined. Theorem 1 gives an explicit closed-form linear lower bound inside this exact global set.

Proof. The result follows from the identity $\frac{dT}{dU} = \frac{\pi}{1-\pi}$ on the local rectangle, together with Theorem 2. □

Theorem 2 is a partial-identification statement. It is useful because it shows where the fragility result lives in the global geometry. It should not be confused with Roy tomography. Under deterministic sorting, choice reveals known sets in the latent plane. Under stochastic selection, the econometrician observes only marginal projections of the two submeasures T and U.

6 Discussion

The central distinction is not exclusion restrictions versus no exclusion restrictions. The distinction is the observation rule. With deterministic Roy sorting, excluded shifters move a known truncation boundary. With an unrestricted stochastic kernel, the same outcome-side shifters move

distributions that are observed only through unknown weights. Theorem 1 quantifies the first-order loss: admitting a positive unknown stochastic component opens the latent joint identified set at linear order along an absorption face.

The result should be read as a fragility theorem for joint-distribution identification. It does not say that every lower-dimensional policy parameter is unidentified under all additional restrictions. In separate work, one may study proxy variables or functional restrictions that recover marginal means directly without recovering the joint distribution. Those functional-identification and estimation questions are outside the scope of this note.

Online Appendix OA.1 gives binary ATE bounds together with an atomless continuous closed-form endpoint formula implied by Theorem 2. Online Appendices OA.2–OA.3 provide numerical and applied calibration exercises that may be of independent interest but are not used in the proof of the main theorem.

Acknowledgments

The authors thank readers of earlier drafts for helpful comments. All remaining errors are our own.

Data and Code Availability

Empirical online appendices are included in the supplementary archive. The NSW data used in Online Appendix OA.3 are publicly available and are described in LaLonde (1986) and Dehejia and Wahba (1999).

References

Beresteanu, Arie, Ilya Molchanov, and Francesca Molinari (2012): "Partial Identification Using Random Set Theory," *Journal of Econometrics*, 166(1), 17–32.

Bhattacharya, Jay, Azeem M. Shaikh, and Edward Vytlacil (2012): "Treatment Effect Bounds: An Application to Swan–Ganz Catheterization," *Journal of Econometrics*, 168(2), 223–243.

Dehejia, Rajeev H. and Sadek Wahba (1999): "Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs," *Journal of the American Statistical Association*, 94(448), 1053–1062.

Eisenhauer, Philipp, James J. Heckman, and Edward Vytlacil (2015): "The Generalized Roy Model and the Cost-Benefit Analysis of Social Programs," *Journal of Political Economy*, 123(2), 413–443.

Hahn, Jinyong (1998): "On the Role of the Propensity Score in Efficient Semiparametric Estimation of Average Treatment Effects," *Econometrica*, 66(2), 315–331.

Heckman, James J. and Bo E. Honoré (1990): "The Empirical Content of the Roy Model," *Econometrica*, 58(5), 1121–1149.

Heckman, James J., Hedibert F. Lopes, and Rémi Piatek (2014): "Treatment Effects: A Bayesian Perspective," *Econometric Reviews*, 33(1–4), 36–67.

Heckman, James J. and Edward Vytlacil (2005): "Structural Equations, Treatment Effects, and Econometric Policy Evaluation," *Econometrica*, 73(3), 669–738.

LaLonde, Robert J. (1986): "Evaluating the Econometric Evaluations of Training Programs with Experimental Data," *American Economic Review*, 76(4), 604–620.

Lee, David S. (2009): "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects," *Review of Economic Studies*, 76(3), 1071–1102.

Lee, Ji Hyung and Byoung Guk Park (2023): "Nonparametric Identification and Estimation of the Extended Roy Model," *Journal of Econometrics*, 235(2), 1087–1113.

Manski, Charles F. (1990): "Nonparametric Bounds on Treatment Effects," *American Economic Review*, 80(2), 319–323.

Manski, Charles F. and John V. Pepper (2000): "Monotone Instrumental Variables: With an Application to the Returns to Schooling," *Econometrica*, 68(4), 997–1010.

Masten, Matthew A. and Alexandre Poirier (2018): "Identification of Treatment Effects under Conditional Partial Independence," *Econometrica*, 86(1), 317–351.

Masten, Matthew A. and Alexandre Poirier (2020): "Inference on Breakdown Frontiers," *Quantitative Economics*, 11(1), 41–111.

Mourifié, Ismaël and Marc Henry (2020): "Sharp Bounds and Testability of a Roy Model of STEM Major Choices," *Journal of Political Economy*, 128(8), 3220–3283.

Roy, A. D. (1951): "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3(2), 135–146.

A Proofs

Proof of Proposition 3. The observed treated outcome subdensity is the y_1-marginal of H_1. Under the alternative it is

$$\int H_1^h(y_1, y_0)\, dy_0 = \int H_1(y_1, y_0)\, dy_0 + \int h(y_1, y_0)\, dy_0.$$

Hence it is unchanged if and only if (11) holds. The observed untreated outcome subdensity is the y_0-marginal of H_0^h, which is fixed by construction. The treatment probability is also unchanged because (11) implies $\int h = 0$. This proves the sharp equivalence within the local slice. The statements about latent marginals follow by integrating $f^h - f = h$; the corresponding ATE can change when the induced change in the Y_0-mean is nonzero. In particular, (11) gives $\Delta\mathbb{E}[Y_1] = 0$, while (12) is exactly preservation of the Y_0-marginal. □

Proof of Corollary 1. Part (i) follows from the zero-marginal restrictions: integrating the perturbation first over y_0 preserves the Y_1-marginal, and integrating first over y_1 preserves the Y_0-marginal. Hence every additive marginal functional is invariant, while joint functionals with nonzero contrast against the perturbation move. Part (ii) is the sharp local-slice statement in Proposition 3: condition (11) is necessary and sufficient for preserving the observed treated outcome law when H_0 is held fixed, but it does not require (12); failing (12) changes the latent Y_0-marginal and therefore can change ATE. □

Proof of Lemma 1. Define the shifted perturbation for a given x by

$$a_x(y_1, y_0) = a\{y_1 - g_1(x_1, x_0),\, y_0 - g_0(x_2, x_0)\}.$$

The zero-marginal property is preserved under shifts. Write the baseline subdensities of (Y_1, Y_0, D) conditional on $X = x$ as

$$H_{1,x} = \pi_x f_x, \qquad H_{0,x} = (1 - \pi_x) f_x,$$

where f_x is the shifted density of (Y_1, Y_0). Set

$$H_{1,\delta,x} = H_{1,x} + \delta a_x, \qquad H_{0,\delta,x} = H_{0,x}.$$

For small $|\delta|$, $H_{1,\delta,x} \geq 0$ because $H_{1,x} \geq \kappa f_x$ on the support of a_x. Moreover

$$H_{1,\delta,x} + H_{0,\delta,x} = f_x + \delta a_x = f_{\delta,x}.$$

Hence define

$$\pi_{\delta,x}(e_1, e_0) = \frac{\pi_x(e_1, e_0) f(e_1, e_0) + \delta a(e_1, e_0)}{f(e_1, e_0) + \delta a(e_1, e_0)}. \tag{15}$$

If $\|a\|_\infty \le M$, take

$$|\delta| \le \frac{\kappa \underline{f}}{2M},$$

with the usual convention that the restriction is void when $M = 0$. On the support of a,

$$f + \delta a \ge (1 - \kappa/2) f > 0, \qquad \pi f + \delta a \ge \kappa f - |\delta| M \ge \kappa f/2.$$

Outside the support of a, the perturbed kernel coincides with the baseline kernel. Moreover

$$1 - \pi_{\delta,x} = \frac{(1 - \pi_x) f}{f + \delta a} \ge \frac{\kappa f}{(1 + \kappa/2) f} \ge \frac{2\kappa}{3},$$

where the final inequality uses $\kappa < 1/2$. Similarly,

$$\pi_{\delta,x} = \frac{\pi_x f + \delta a}{f + \delta a} \ge \frac{\kappa f/2}{(1 + \kappa/2) f} \ge \frac{\kappa}{3}.$$

Thus $\pi_{\delta,x} \in [\kappa/3, 1 - \kappa/3]$ after replacing κ by a smaller overlap constant. This explicit bound is only used to verify that the constructed kernel remains a regular stochastic kernel.

The observed treated outcome density is the y_1-marginal of $H_{1,x}$. Under the perturbed model it is

$$\int H_{1,\delta,x}(y_1, y_0)\, dy_0 = \int H_{1,x}(y_1, y_0)\, dy_0 + \delta \int a_x(y_1, y_0)\, dy_0 = \int H_{1,x}(y_1, y_0)\, dy_0.$$

The observed untreated outcome density is the y_0-marginal of $H_{0,x}$, which is unchanged. The choice probability is also unchanged because the total mass of a_x is zero. Thus the conditional law of (D, Y) given $X = x$ is identical. Since $a \neq 0$ and has zero marginals, the perturbation changes only the copula of the latent error distribution. □

Remark 6 (Zero-marginal perturbations)**.** *Nonzero functions satisfying* (14) *are easy to construct. Let r and s be bounded, compactly supported functions with*

$$\int r(e_1)\, de_1 = 0, \qquad \int s(e_0)\, de_0 = 0.$$

Then $a(e_1, e_0) = r(e_1) s(e_0)$ has zero marginals in both coordinates. Choosing the supports inside a rectangle where the baseline density is bounded away from zero guarantees that $f + \delta a$ is a valid density for sufficiently small $|\delta|$. To support the perturbation inside one side of the Roy boundary, choose the rectangle entirely inside that side.

Proof of Theorem 1. Work above the Roy boundary. Take a baseline relaxed kernel with $\pi = 1 - \eta$ on B, so the local subdensities are

$$H_1 = (1 - \eta) f, \qquad H_0 = \eta f.$$

Choose $\rho > 0$ and $\bar{\eta} > 0$ so small that for all $\eta \le \bar{\eta}$,

$$c_L\eta \le \frac{\eta}{1+\rho\eta} \le \frac{\eta}{1-\rho\eta} \le c_U\eta, \qquad \rho\eta \le 1/2.$$

This is possible because $c_L < 1 < c_U$. One explicit admissible choice is

$$\rho = \min\left\{\frac{1-c_L}{c_U}, \frac{c_U-1}{c_U}, \frac{1}{4}\right\}, \qquad \bar{\eta} = 1.$$

Indeed, this ρ satisfies $\rho \le (1-c_L)/c_L$, $\rho \le (c_U-1)/c_U$, and $\rho \le 1/4$, which imply (A) for every $\eta \le 1$. For the canonical choice $c_L = 1/2, c_U = 2$, this gives $\rho = 1/4$. Set $\delta_\eta = \rho\eta\underline{f}/M$. For the two perturbed models let

$$H_{1,+} = H_1 + \delta_\eta a, \qquad H_{1,-} = H_1 - \delta_\eta a, \qquad H_{0,+} = H_{0,-} = H_0.$$

Since $|\delta_\eta a|/f \le \rho\eta$, the densities $f_\pm = f \pm \delta_\eta a$ are positive. The induced kernels satisfy

$$1 - \pi_\pm = \frac{H_0}{f_\pm} = \frac{\eta f}{f \pm \delta_\eta a}.$$

By (A), $c_L\eta \le 1-\pi_\pm \le c_U\eta$ on B, so $\pi_\pm \in \mathcal{M}_\eta(B; c_L, c_U)$. The observed treated outcome density is the Y_1-marginal of $H_{1,\pm}$, and it is unchanged because a has zero e_0-marginal. The observed untreated outcome density is the Y_0-marginal of H_0, which is fixed. Choice probabilities are also unchanged because a has total integral zero. Therefore the two full-data laws are observationally equivalent. Finally,

$$d_{TV}(f_+, f_-) = \frac{1}{2}\int |2\delta_\eta a| = \delta_\eta A.$$

The case below the boundary is symmetric. □

Proof of Corollary 2. For every $\gamma \in [-\delta_\eta, \delta_\eta]$, the proof of Theorem 1 gives a valid η-relaxed stochastic kernel and the same observed law. Within the defined face, the only free coordinate is γ. Hence the largest total-variation distance is attained at the endpoints $\gamma = \delta_\eta$ and $\gamma = -\delta_\eta$. The formula for L_h follows by linearity. □

B Additional implications of the absorption geometry

B.1 Functional separation on the copula subface

Proposition 4 (Functional separation under marginal-neutral absorption)**.** *In the setting of Lemma 1, consider the observationally equivalent path $f_\delta = f + \delta a$, where a satisfies the zero-marginal restrictions (14). Then:*

(i) For every pair of integrable one-dimensional functions φ_1, φ_0,

$$\int \{\varphi_1(y_1) + \varphi_0(y_0)\}\, f_\delta(y_1, y_0)\, dy_1 dy_0 = \int \{\varphi_1(y_1) + \varphi_0(y_0)\}\, f(y_1, y_0)\, dy_1 dy_0.$$

Hence the perturbation preserves both marginal distributions and every additive marginal functional, including marginal means, marginal distribution functions, and the ATE whenever the two marginal means are otherwise pinned down.

(ii) For every bounded measurable joint functional $\psi(y_1, y_0)$,

$$\int \psi(y_1, y_0) f_\delta(y_1, y_0)\, dy_1 dy_0 - \int \psi(y_1, y_0) f(y_1, y_0)\, dy_1 dy_0 = \delta \int \psi(y_1, y_0) a(y_1, y_0)\, dy_1 dy_0.$$

Thus any nonadditive functional whose contrast with a is nonzero changes while the observed law is unchanged. In particular, by choosing a inside a region on which $y_1 - y_0$ varies, one can change features of the gain distribution $Y_1 - Y_0$, such as probabilities $\mathbb{P}(Y_1 - Y_0 \le t)$ for appropriate thresholds t, without changing the observed law.

The statement separates a positive fact from a negative one: marginal-neutral absorption leaves additive marginal functionals fixed, but it destroys identification of genuinely joint objects. It does not assert that all marginal laws are point identified in the unrestricted stochastic-selection model; it characterizes the absorption directions that replace Roy tomography.

Proof. The first display follows by integrating a first over y_0 for the φ_1 term and first over y_1 for the φ_0 term. The second display is linearity. For the final claim, take a product perturbation $a = r(y_1)s(y_0)$ as in Remark 6; for generic choices of r and s, the inner product with $\mathbf{1}\{y_1 - y_0 \le t\}$ is nonzero for thresholds cutting through the support rectangle. □

Remark 7 (Overlap constants)**.** *The absorption construction may weaken a uniform overlap constant by a fixed factor, as in κ to $\kappa/3$ in Lemma 1. This is harmless for the nonidentification claim: the perturbed model remains in the same qualitative overlapping stochastic-selection class, and the relaxation is in the direction that makes the claim harder to refute.*

B.2 Local universality of the Roy-boundary opening

The following corollary upgrades the directional statement of Theorem 1 into a local-neighborhood universality property: every total-variation neighborhood of the exact Roy observed law contains a stochastic-Roy observed law with a first-order opening of the latent identified set.

Corollary 4 (Local universality of the Roy-boundary opening)**.** *Suppose the exact Roy model satisfies the Heckman–Honoré full-support and normalization conditions so that the latent joint law is identified when $D = \mathbf{1}\{Y_1 > Y_0\}$. Fix a rectangle B strictly on one side of the boundary. Assume that the baseline density is bounded above and below on B, with upper bound $\bar{f}_B < \infty$*

and lower bound $\underline{f} > 0$. Let $\mathcal{A}_B$ denote the class of nonzero bounded zero-marginal perturbations supported on B, and define

$$C^*(B) = \rho \underline{f} \sup_{a \in \mathcal{A}_B} \frac{\|a\|_1}{\|a\|_\infty}.$$

For every $0 < \xi < C^*(B)$, there exist constants $\bar{\varepsilon} > 0$ and $s_B > 0$ such that for every $\varepsilon \in (0, \bar{\varepsilon}]$, there is an observed law P_ε with

$$d_{TV}(P_\varepsilon, P_0) \leq \varepsilon,$$

generated by a model in $\mathcal{M}_{s_B\varepsilon}(B; c_L, c_U)$, for which

$$\operatorname{diam}_{TV}\{\mathcal{I}_{s_B\varepsilon}(P_\varepsilon)\} \geq (C^*(B) - \xi) s_B \varepsilon.$$

Equivalently, every total-variation neighborhood of the exact Roy observed law contains stochastic-Roy observed laws whose latent identified set has first-order diameter. This is an existential, local-neighborhood statement: it does not claim that all nearby observed laws have large identified sets, nor that the global identified-set correspondence has a Hausdorff upper or lower bound of order ε over unrestricted neighborhoods.

Proof. Choose $a_\xi \in \mathcal{A}_B$ such that $C_{a_\xi} \geq C^*(B) - \xi$. The construction in Theorem 1, applied with η, produces an observed law $P_\eta \in \mathcal{M}_\eta(B; c_L, c_U)$ and two observationally equivalent latent laws separated by at least $C_{a_\xi}\eta$. Relative to the exact Roy observed law P_0, the construction changes only the subdensity on B, by moving order-η mass across the two observed regimes. Hence

$$d_{TV}(P_\eta, P_0) \leq \eta |B| \bar{f}_B.$$

Set $s_B = \min\{1, (|B|\bar{f}_B)^{-1}\}/2$ and $\eta = s_B\varepsilon$. Let $\bar{\eta}$ denote the smallness constant in Theorem 1, which depends only on (c_L, c_U), and take $\bar{\varepsilon} = \bar{\eta}/s_B$. Then, for every $\varepsilon \leq \bar{\varepsilon}$, the positivity restrictions used in Theorem 1 hold, P_η lies within ε of P_0, and the diameter lower bound becomes $(C^*(B) - \xi)s_B\varepsilon$. At $\eta = 0$, the Heckman–Honoré assumptions identify the latent joint law, so the limiting exact-Roy identified set is a singleton. □

Remark 8 (Why the slack ξ appears)**.** *The supremum defining $C^*(B)$ need not be attained in $\mathcal{A}_B$. It is therefore written as a supremum, and Corollary 4 uses a near-maximizing perturbation direction a_ξ. The slack $\xi > 0$ is only an approximation device; it is not an additional economic restriction. Since $\|a\|_1/\|a\|_\infty \leq |B|$, the constant is finite whenever B has finite Lebesgue measure.*

Remark 9 (Why this is not a global Hausdorff theorem)**.** *Theorem 1 proves a linear lower bound for the full identified set and an exact linear diameter on the constructed absorption face. A universal upper bound for the full global identified set would require specifying how the stochastic kernel remains close to the Roy rule outside the local rectangle and how the inverse tomography operator is regularized. Without such global restrictions, the identified set may contain other*

directions. This scope is intentional: the theorem identifies the first-order loss created by admitting a positive unknown stochastic component into the Roy measurement rule.

Online Appendix

OA.1 Binary and continuous ATE bounds

OA.1.1 Binary outcomes

Theorem 2 implies explicit sharp ATE bounds in the binary case. The next result is its binary specialization, where the sharp ATE set can be written explicitly. The no-overlap part is the classical binary no-assumption bound of Manski (1990); the new component used here is the overlap-constrained linear program for an unrestricted outcome-dependent stochastic kernel. This differs from Roy sharp-bounds analyses such as Mourifié and Henry (2020), where Roy or monotone selection restrictions carry additional empirical content.

Let $Y_1, Y_0 \in \{0,1\}$. Write

$$a = \mathbb{P}(D=1, Y=1), \quad b = \mathbb{P}(D=1, Y=0), \quad c = \mathbb{P}(D=0, Y=1), \quad d = \mathbb{P}(D=0, Y=0),$$

so $a+b+c+d=1$. Let $f_{ij} = \mathbb{P}(Y_1 = i, Y_0 = j)$ and $\pi_{ij} = \mathbb{P}(D=1 \mid Y_1 = i, Y_0 = j)$. The ATE is

$$\tau = f_{10} - f_{01}.$$

Proposition 5 (Sharp binary ATE bounds)**.** *The following statements hold.*

(i) If the selection kernel π_{ij} is unrestricted apart from $0 \le \pi_{ij} \le 1$, the sharp identified set for $\tau = \mathbb{E}[Y_1 - Y_0]$ is

$$\mathcal{I}_\tau(P_{D,Y}) = [-(b+c),\ a+d]. \tag{16}$$

(ii) If a known overlap bound $\kappa \le \pi_{ij} \le 1-\kappa$, $\kappa \in (0, 1/2)$, is imposed, the sharp identified set is the value range of the finite linear program

$$\begin{aligned} &\textit{optimize} && t_{10} + u_{10} - t_{01} - u_{01} \\ &\textit{over} && t_{ij} \ge 0,\ u_{ij} \ge 0, \qquad i,j \in \{0,1\}, \\ &\textit{subject to} && t_{10} + t_{11} = a, \quad t_{00} + t_{01} = b, \quad u_{01} + u_{11} = c, \quad u_{00} + u_{10} = d, \\ & && \lambda u_{ij} \le t_{ij} \le \lambda^{-1} u_{ij}, \qquad \lambda = \frac{\kappa}{1-\kappa}. \end{aligned}$$

The lower and upper endpoints are obtained by minimizing and maximizing the objective.

(iii) In the balanced case $a = b = c = d = 1/4$, the overlap-constrained set equals

$$\left[-\frac{1-2\kappa}{2(1-\kappa)},\ \frac{1-2\kappa}{2(1-\kappa)} \right]. \tag{17}$$

Proof. Let

$$t_{ij} = \mathbb{P}(Y_1 = i, Y_0 = j, D = 1), \qquad u_{ij} = \mathbb{P}(Y_1 = i, Y_0 = j, D = 0).$$

κ	0.10	0.25	0.40
Upper endpoint	0.444	0.333	0.167
Lower endpoint	-0.444	-0.333	-0.167
Interval length	0.889	0.667	0.333

Figure OA.1: Balanced binary case. The sharp ATE interval under the overlap bound $\kappa \leq \pi_{ij} \leq 1-\kappa$ contracts from the Manski no-overlap interval $[-1/2, 1/2]$ toward the singleton at $\kappa = 1/2$.

The observed law imposes

$$t_{10} + t_{11} = a, \quad t_{00} + t_{01} = b, \quad u_{01} + u_{11} = c, \quad u_{00} + u_{10} = d.$$

Since $f_{ij} = t_{ij} + u_{ij}$,

$$\tau = f_{10} - f_{01} = t_{10} + u_{10} - t_{01} - u_{01}.$$

Without overlap restrictions, (t_{10}, u_{10}) can range independently over $[0, a] \times [0, d]$, and (t_{01}, u_{01}) over $[0, b] \times [0, c]$. This gives the sharp interval (16), and every value in the interval is attained by a feasible allocation.

With $\kappa \leq \pi_{ij} \leq 1-\kappa$, the relation $\pi_{ij} = t_{ij}/(t_{ij} + u_{ij})$ is equivalent to

$$\lambda u_{ij} \leq t_{ij} \leq \lambda^{-1} u_{ij}, \qquad \lambda = \frac{\kappa}{1-\kappa},$$

with the convention that $t_{ij} = u_{ij} = 0$ is allowed. Thus the sharp set is exactly the range of the displayed linear program.

For $a = b = c = d = s = 1/4$, the maximum is attained by

$$t_{10} = u_{10} = s(1-\lambda), \quad t_{01} = u_{01} = 0,$$

$$t_{11} = u_{00} = s\lambda, \quad u_{11} = t_{00} = s.$$

This allocation satisfies all overlap constraints, with cells $(1, 1)$ and $(0, 0)$ at the lower and upper overlap boundaries. It yields $2s(1-\lambda) = (1-2\kappa)/\{2(1-\kappa)\}$. The minimum follows symmetrically. □

OA.1.2 Continuous-outcome linear-programming formulation

Let $\mathcal{Y} \subset \mathbb{R}$ be compact. Denote the observed treated and untreated subprobability measures by

$$P_1(A) = \mathbb{P}(D = 1, Y \in A), \qquad P_0(A) = \mathbb{P}(D = 0, Y \in A).$$

Let T and U be nonnegative finite measures on $\mathcal{Y}^2$, interpreted as the joint submeasures of (Y_1, Y_0) among treated and untreated units:

$$T(A) = \mathbb{P}((Y_1, Y_0) \in A, D = 1), \qquad U(A) = \mathbb{P}((Y_1, Y_0) \in A, D = 0).$$

Write T_1 for the Y_1-marginal of T, and U_0 for the Y_0-marginal of U.

Theorem 2 proves that these submeasures give the sharp global identified set. This appendix records the closed-form atomless endpoint calculation and the finite-partition computation implied by that theorem.

Proposition 6 (Closed-form continuous bounds under atomless observed marginals)**.** *Let $\mathcal{Y} = [\underline{y}, \bar{y}] \subset \mathbb{R}$ be compact. Suppose that the observed subprobability measures P_1 and P_0 are atomless, with $p_1 = P_1(\mathcal{Y}) = \mathbb{P}(D = 1)$, $p_0 = P_0(\mathcal{Y}) = \mathbb{P}(D = 0) = 1 - p_1$. Fix $\kappa \in (0, \min\{p_1, p_0\})$. Define*

$$q_+(\kappa) = \frac{\kappa(p_0 - \kappa)}{1 - 2\kappa}, \qquad q_-(\kappa) = \frac{\kappa(p_1 - \kappa)}{1 - 2\kappa}.$$

Let y_+^* satisfy $P_1((y_+^*, \bar{y}]) = q_+(\kappa)$, and let y_-^* satisfy $P_0([\underline{y}, y_-^*)) = q_-(\kappa)$. Then the sharp upper bound on $\tau = \mathbb{E}[Y_1 - Y_0]$ under $\kappa \le \pi(Y_1, Y_0) \le 1 - \kappa$ is

$$\begin{aligned} \bar{\tau}_\kappa = &\frac{1}{\kappa}\left\{\int_{(y_+^*, \bar{y}]} y\, dP_1(y) - \int_{[\underline{y}, y_-^*)} y\, dP_0(y)\right\} \\ &+ \frac{1}{1-\kappa}\left\{\int_{[\underline{y}, y_+^*]} y\, dP_1(y) - \int_{[y_-^*, \bar{y}]} y\, dP_0(y)\right\}. \end{aligned} \tag{18}$$

The sharp lower bound is the symmetric expression: let z_-^* satisfy $P_1([\underline{y}, z_-^*)) = q_+(\kappa)$, and let z_+^* satisfy $P_0((z_+^*, \bar{y}]) = q_-(\kappa)$. Then

$$\begin{aligned} \underline{\tau}_\kappa = &\frac{1}{\kappa}\left\{\int_{[\underline{y}, z_-^*)} y\, dP_1(y) - \int_{(z_+^*, \bar{y}]} y\, dP_0(y)\right\} \\ &+ \frac{1}{1-\kappa}\left\{\int_{[z_-^*, \bar{y}]} y\, dP_1(y) - \int_{[\underline{y}, z_+^*]} y\, dP_0(y)\right\}. \end{aligned} \tag{19}$$

As $\kappa \downarrow 0$, the upper and lower endpoints converge to the continuous Manski endpoints. As $\kappa \uparrow \min\{p_1, p_0\}$, they converge to $E_{P_1}[Y]/p_1 - E_{P_0}[Y]/p_0$.

Proof. Consider the upper endpoint. Let G_1 and G_0 denote the latent marginals of Y_1 and Y_0. If a feasible full-data law and kernel generate P_1 and P_0, then

$$P_1(A) = \int_A \alpha_1(y)\, dG_1(y), \qquad P_0(A) = \int_A \alpha_0(y)\, dG_0(y),$$

where $\alpha_1(y) = \mathbb{E}[\pi(Y_1, Y_0) \mid Y_1 = y] \in [\kappa, 1 - \kappa]$ and $\alpha_0(y) = \mathbb{E}[1 - \pi(Y_1, Y_0) \mid Y_0 = y] \in [\kappa, 1 - \kappa]$.

Hence

$$\frac{1}{1-\kappa}\,dP_1 \le dG_1 \le \frac{1}{\kappa}\,dP_1, \qquad \frac{1}{1-\kappa}\,dP_0 \le dG_0 \le \frac{1}{\kappa}\,dP_0,$$

with $G_1(\mathcal{Y}) = G_0(\mathcal{Y}) = 1$. Maximizing $\int y\,dG_1$ subject to these one-dimensional constraints puts density $1/\kappa$ on the upper P_1-tail and density $1/(1-\kappa)$ elsewhere. The total-mass constraint gives

$$P_1((y_+^*, \bar{y}]) = q_+(\kappa).$$

Similarly, minimizing $\int y\,dG_0$ puts density $1/\kappa$ on the lower P_0-tail and density $1/(1-\kappa)$ elsewhere, with $P_0([\underline{y}, y_-^*)) = q_-(\kappa)$. This proves that (18) is an upper bound.

It remains to show achievability. Let G_1^* and G_0^* be the two bang-bang marginal optimizers. The identity

$$q_+(\kappa) + q_-(\kappa) = \kappa$$

implies that the rank cutoffs align:

$$G_1^*(y_+^*) = \frac{p_1 - q_+(\kappa)}{1-\kappa} = \frac{q_-(\kappa)}{\kappa} = G_0^*(y_-^*).$$

Couple G_1^* and G_0^* comonotonically, $Y_1 = (G_1^*)^{-1}(U)$, $Y_0 = (G_0^*)^{-1}(U)$, $U \sim \mathrm{Unif}(0,1)$. Define $\pi^* = 1-\kappa$ on the common lower-rank region $\{Y_1 \le y_+^*\} = \{Y_0 < y_-^*\}$, and $\pi^* = \kappa$ on the complementary upper-rank region. Then $\pi^* \in [\kappa, 1-\kappa]$, and direct substitution gives

$$\pi^* G_1^* = P_1, \qquad (1-\pi^*)G_0^* = P_0,$$

where the equalities are understood as marginal submeasures. Thus the comonotonic law and π^* attain (18). The lower endpoint is obtained by reversing high and low tails. The limiting cases follow by continuity of the displayed formulas. □

Remark 10 (Atoms and the binary case)**.** *If P_1 or P_0 has atoms, the same formulas apply with mass splitting at the threshold atoms. Formally, one may lift an atom into two copies and assign the split pieces to the two bang-bang regions. This convention recovers the fully discrete binary formula in Proposition 5; in the balanced case it gives $(1-2\kappa)/(2(1-\kappa))$, matching (17).*

For computation, choose a finite partition $\Pi = \{I_1, \ldots, I_K\}$ of $\mathcal{Y}$ and replace T and U by their masses on rectangles $I_j \times I_k$. This gives an ordinary finite-dimensional linear program with exactly the same structure as Proposition 5. As the mesh of Π shrinks, the finite programs converge to the measure program under the compactness assumptions above. Thus the binary proposition is not an isolated example; it is the two-atom version of the continuous submeasure problem.

The measure-linear program is related to partial-identification work on treatment effects, but its restriction is different. The no-overlap binary component is the classical no-assumption logic of Manski (1990). Monotone IV and monotone treatment-response bounds (Manski and Pepper, 2000), propensity-score-based semiparametric analysis (Hahn, 1998), and sample-selection

trimming bounds (Lee, 2009) impose different restrictions. The program here does not impose monotonicity, rank invariance, or a Lee-style one-sided selection rule; it constrains the treated and untreated joint submeasures through outcome-dependent stochastic-selection odds.

OA.2 Numerical Illustrations

This online appendix reports numerical checks. They are not used for identification or proof. Their role is to make the main algebraic mechanisms visible and to provide reproducible diagnostics for readers. The replication code is supplied with the package. The first exercise directly visualizes the linear opening in the main theorem; the remaining exercises check the algebraic constructions and the sharp-bound formulas.

OA.2.1 Oracle linear-fragility plot

The main theorem is a population identification statement rather than a sampling claim. The following oracle exercise visualizes its first-order content. We take a truncated bivariate Gaussian baseline latent law for (Y_1, Y_0) with correlation 0.3, choose a rectangle B strictly above the Roy boundary, and use a bounded product perturbation with zero one-dimensional marginals on B. For each η, the treated subdensity is perturbed by $\pm\delta_\eta a$ while the untreated subdensity is held fixed, exactly as in Theorem 1. The two perturbed latent laws are observationally equivalent by construction. Since this is an oracle exercise, we plot a latent object–the correlation gap between the two latent joint laws–that cannot be recovered from the observed data.

Table OA.1: Oracle linear-fragility check. The two latent laws are observationally equivalent for each η. The latent total-variation gap and the latent correlation gap grow linearly in η. The absolute scale is conservative because the perturbation size is normalized by the minimum baseline density on the active rectangle; the point of the table is the first-order rate.

η	$d_{TV}(f_+, f_-)$	corr_+	$\mathrm{corr}_+ - \mathrm{corr}_-$
0.02	5.37×10^{-5}	0.293015	4.42×10^{-6}
0.05	1.34×10^{-4}	0.293018	1.11×10^{-5}
0.10	2.68×10^{-4}	0.293024	2.21×10^{-5}
0.20	5.37×10^{-4}	0.293035	4.42×10^{-5}
0.30	8.05×10^{-4}	0.293046	6.64×10^{-5}

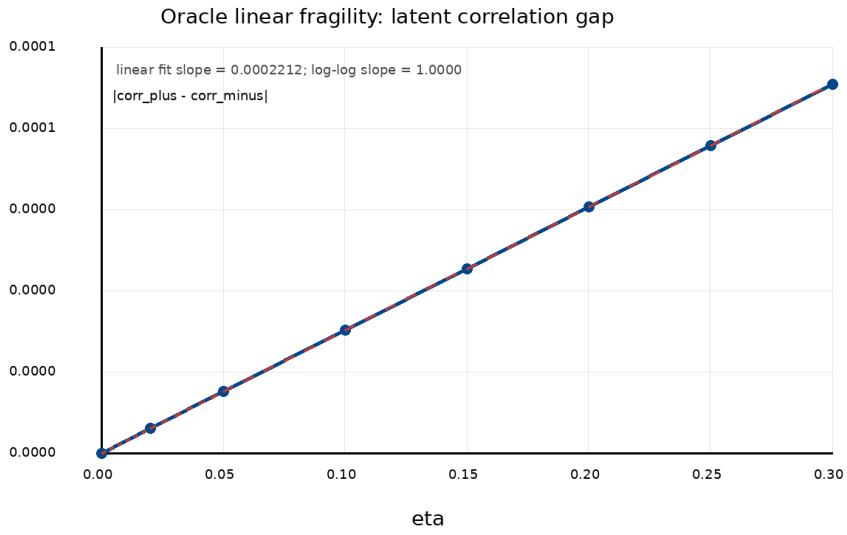

Figure OA.2: Oracle visualization of the main theorem. The vertical axis is the absolute difference in latent correlations between two observationally equivalent latent laws generated by the absorption construction. The dashed line is the least-squares linear fit through the origin; a log–log regression of the positive points gives slope 1.0000.

Figure OA.3 complements the rate check with a projection diagnostic at a fixed $\eta = 0.30$. To make the geometry visible, this diagnostic uses a uniform baseline and a larger zero-marginal product perturbation on a rectangle above the Roy boundary. The two latent densities differ on the local rectangle, with total-variation distance 6.12×10^{-3}. Nevertheless, the observed treated projection is identical up to numerical precision: the maximum grid discrepancy between the two treated marginals is 4.5×10^{-18}. This is the same algebra as Proposition 3: a latent joint direction is open while the observed marginal projection is closed.

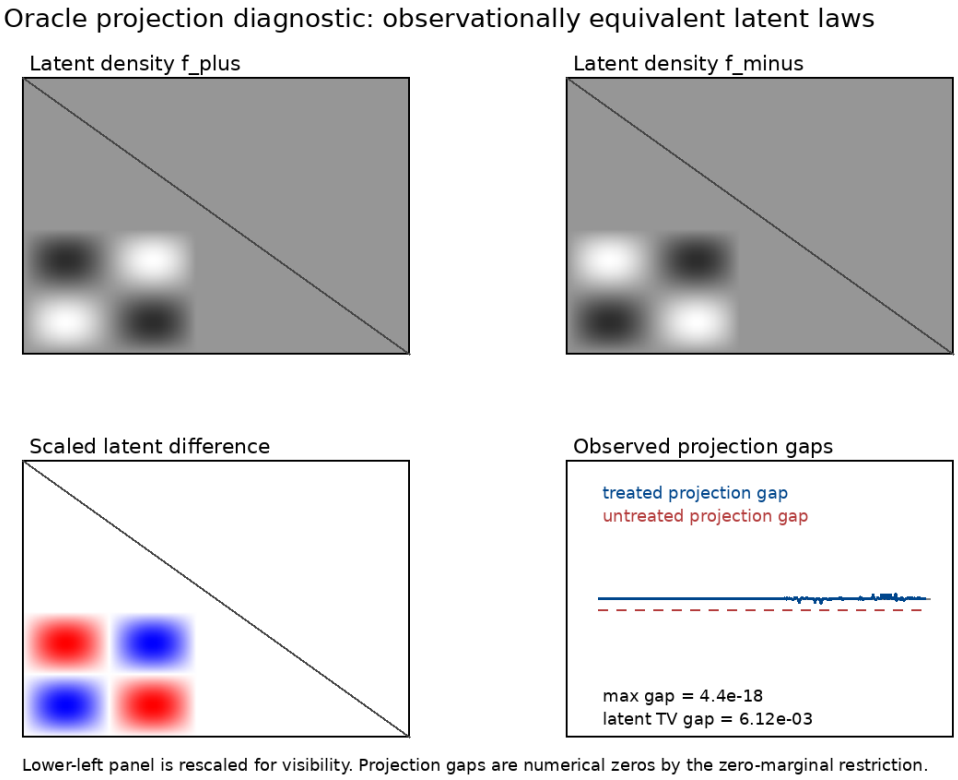

Figure OA.3: Oracle projection diagnostic. The upper panels show two observationally equivalent latent laws, f_+ and f_-, generated by the local absorption construction at $\eta = 0.30$. The lower-left panel rescales the signed latent difference for visibility. The lower-right panel shows that the corresponding observed projection gaps are numerical zeros. The figure visualizes how stochastic sorting with an unknown kernel can hide a local joint-law perturbation behind the same observed marginal projection.

OA.2.2 Copula absorption on a grid

The first illustration discretizes the construction in Theorem 1. The baseline latent density is uniform on $[0,1]^2$. The zero-marginal perturbation is

$$a(y_1, y_0) = \sin\left(2\pi \frac{y_1 - 0.55}{0.45}\right) \sin\left(2\pi \frac{y_0}{0.45}\right)$$

inside the rectangle $[0.55, 1] \times [0, 0.45]$, and zero outside. This rectangle lies strictly above the Roy boundary. For each η, the perturbation size is $\delta_\eta = \eta/(4\|a\|_\infty)$, as in Theorem 1. The treated subdensity is perturbed by $\pm\delta_\eta a$, while the untreated subdensity is held fixed.

Table OA.2: Numerical check of copula absorption. The latent joint density changes at linear order in η, while the maximum discrepancy in the observed law is numerical zero. The final two columns show that the latent copula changes sign through the correlation of (Y_1, Y_0).

η	$d_{TV}(f_+, f_-)$	max observed-law diff	corr(f_+)	corr(f_-)
0.05	1.026×10^{-3}	2.2×10^{-18}	1.558×10^{-4}	-1.558×10^{-4}
0.10	2.052×10^{-3}	3.3×10^{-18}	3.116×10^{-4}	-3.116×10^{-4}
0.15	3.078×10^{-3}	4.4×10^{-18}	4.675×10^{-4}	-4.675×10^{-4}
0.20	4.104×10^{-3}	4.4×10^{-18}	6.233×10^{-4}	-6.233×10^{-4}
0.25	5.130×10^{-3}	4.4×10^{-18}	7.791×10^{-4}	-7.791×10^{-4}

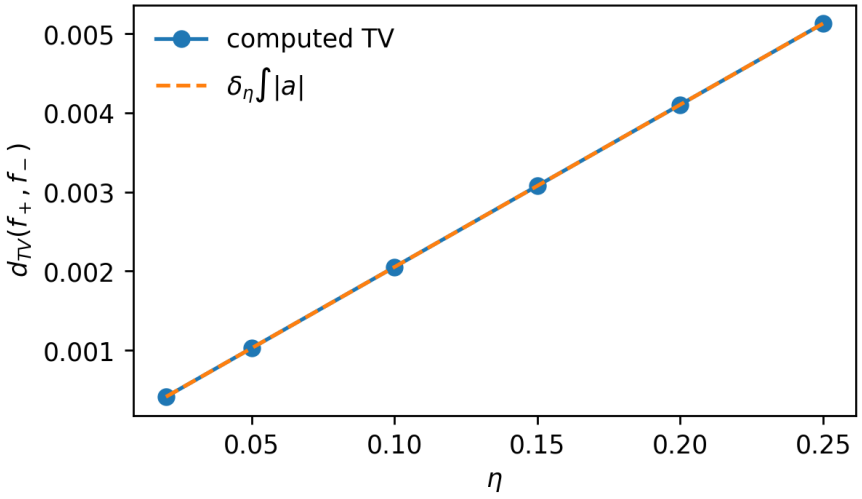

Figure OA.4: The total-variation distance between the two latent joint densities grows linearly with the stochastic-relaxation size η. The dashed line is the theoretical value $\delta_\eta \int |a|$.

OA.2.3 Opposite sharp binary endpoints with the same observed law

The second illustration simulates the balanced binary case with $a = b = c = d = 1/4$ and $\kappa = 0.25$. Proposition 5 gives the sharp interval $[-1/3, 1/3]$. We simulate two extremal full-data laws: one attaining $\tau = 1/3$, and the symmetric one attaining $\tau = -1/3$. Both imply the same observed law of (D, Y). In each design, 200 samples of size 20,000 are generated.

Table OA.3: Monte Carlo confirmation of the binary endpoint construction. The observed shares (a, b, c, d) are nearly $(1/4, 1/4, 1/4, 1/4)$ in both designs, but the latent ATEs are at opposite endpoints of the sharp identified set.

Design	true τ	mean sample τ	mean a	mean b	mean c	mean d
upper endpoint	0.333	0.333	0.250	0.250	0.250	0.250
lower endpoint	-0.333	-0.333	0.250	0.250	0.250	0.250

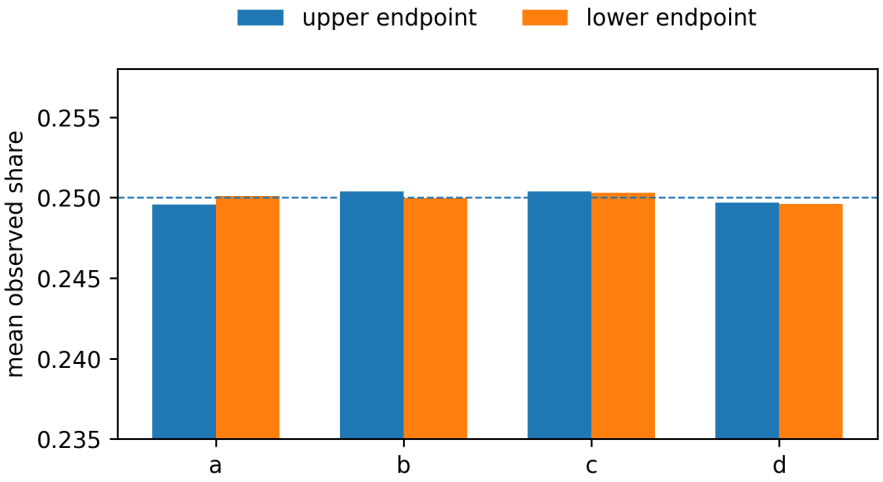

Figure OA.5: The two endpoint laws have the same observed cell probabilities (a, b, c, d) up to Monte Carlo error, even though their latent ATEs are $1/3$ and $-1/3$.

OA.2.4 LP verification of the balanced binary sharp bound

The third illustration checks the linear-program characterization of Proposition 5. For the balanced case $a = b = c = d = 1/4$, the closed-form sharp interval is

$$\left[-\frac{1-2\kappa}{2(1-\kappa)}, \frac{1-2\kappa}{2(1-\kappa)}\right].$$

We solve the finite-dimensional submeasure LP directly, using variables (T_{ij}, U_{ij}) for treated and untreated joint submeasures of $(Y_1, Y_0) = (i, j)$. The overlap constraints are

$$\lambda U_{ij} \leq T_{ij} \leq \lambda^{-1} U_{ij}, \qquad \lambda = \frac{\kappa}{1-\kappa}.$$

Table OA.4: LP verification of the balanced binary sharp bound. The linear-program endpoints match the closed-form formula up to numerical solver tolerance.

κ	LP lower	formula lower	LP upper	formula upper	max error
0.00	-0.500	-0.500	0.500	0.500	0.0e+00
0.10	-0.444	-0.444	0.444	0.444	5.6e-17
0.25	-0.333	-0.333	0.333	0.333	5.6e-17
0.40	-0.167	-0.167	0.167	0.167	0.0e+00
0.45	-0.091	-0.091	0.091	0.091	5.6e-17

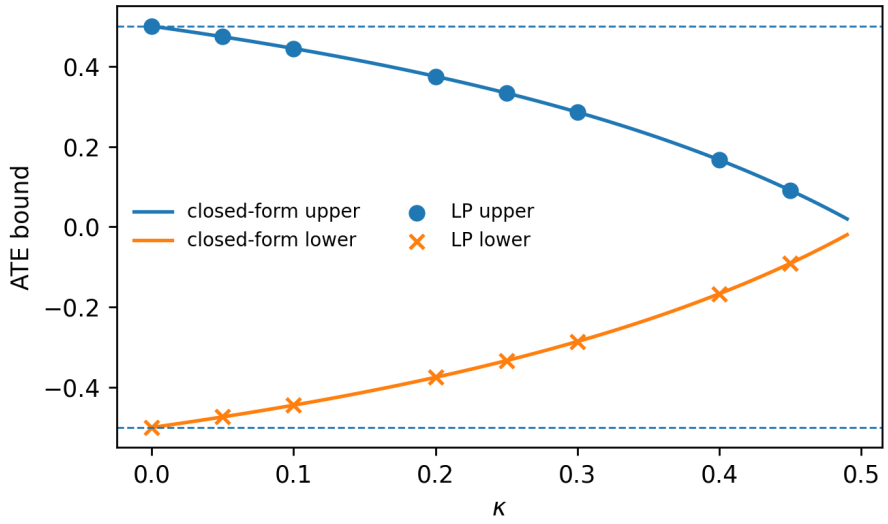

Figure OA.6: The finite-dimensional LP endpoints coincide with the closed-form balanced binary sharp bounds. The dashed horizontal lines are the no-overlap Manski endpoints $[-1/2, 1/2]$.

OA.2.5 Continuous closed form versus discretized measure LP

The final numerical check verifies Proposition 6. We compare the closed-form endpoints with the finite-partition LP from Theorem 2. The uniform symmetric case aligns exactly with the grid, so the two endpoints agree at machine precision. The asymmetric-uniform and beta-shaped cases differ only by the finite-bin approximation of the continuous objective; the discrepancy decreases as the partition is refined. The script `continuous_closed_form_verification_v15.py` is included in the replication archive.

Table OA.5: Closed-form continuous bounds versus discretized measure-LP bounds. The LP uses 40 equal-width bins on $[0, 1]$.

Case	p_1	κ	CF upper	LP upper	max error
Uniform symmetric	0.50	0.10	0.4000	0.4000	2.8×10^{-16}
Uniform symmetric	0.50	0.20	0.3000	0.3000	1.7×10^{-16}
Uniform asymmetric	0.40	0.20	0.2778	0.2775	2.8×10^{-4}
Beta symmetric	0.50	0.20	-0.2443	-0.2441	2.0×10^{-4}
Beta asymmetric	0.30	0.20	-0.3112	-0.3113	1.6×10^{-4}
Beta asymmetric	0.30	0.30	-0.4286	-0.4286	3.0×10^{-7}

OA.3 Empirical Illustration: NSW Training Data

This appendix illustrates the binary sharp bounds of Proposition 5 on the National Supported Work (NSW) data. The illustration has a specific role: it calibrates the global sharp set of Theorem 2 in a Roy-style program-evaluation dataset where an experimental benchmark is available. The bounds are computed as if random assignment were not available; the experiment is used only to benchmark containment and to give an interpretable scale for the overlap parameter κ. We use the cleaned Dehejia–Wahba sample of Dehejia and Wahba (1999) together with the CPS-1 nonexperimental controls of LaLonde (1986).

The NSW experimental sample contains $n = 445$ observations (185 treated, 260 controls) of disadvantaged workers randomly assigned to a training program. Outcome $re78$ is 1978 earnings. Because assignment is random in the experiment, the experimental ATE is point identified by a difference in means. We treat this as an external benchmark, not as an input to the bounds.

In the empirical results below, $D = 1$ denotes participation in NSW training and Y is a binarized version of $re78$ under three thresholds: $Y > 0$ (any positive earnings), $Y >$ median, and $Y >$ 5,000. Proposition 5 is then applied without using random assignment; the sharp bounds are computed from the observed cells (a, b, c, d) for a grid of overlap parameters κ. This isolates the ATE empirical content of the unrestricted stochastic-selection model. The exercise also provides a feasibility diagnostic: because a global lower bound κ on treatment probabilities cannot exceed the smaller observed treatment or control mass, the admissible κ-range reveals whether a pooled comparison has enough common support to make an overlap-based stochastic-selection analysis meaningful.

OA.3.1 Illustration 1: NSW experimental sample

Figure OA.7 reports the sharp ATE identified set as a function of κ for each of the three binarizations. Bootstrap pointwise 90% confidence intervals ($B = 400$) summarize sampling variability. In all three cases the experimental ATE (dashed red line) lies inside the identified set for every feasible

κ. As κ approaches the marginal propensity $P(D = 1) \approx 0.42$, the identified set collapses toward the difference-in-means estimator, recovering the experimental ATE. This is Proposition 5 (iii) in finite samples.

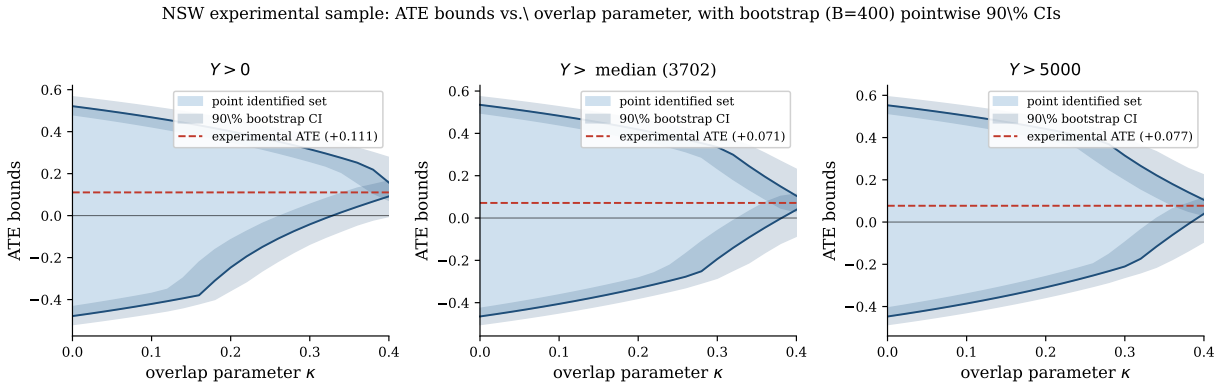

Figure OA.7: Sharp ATE bounds on the NSW experimental sample as a function of the overlap parameter κ. Dark blue curves are the point-estimated lower and upper endpoints of the identified set in Proposition 5 (ii). The light blue band is the point identified set. The narrower shaded ribbon is the pointwise 90% bootstrap confidence interval over $B = 400$ bootstrap replications. The dashed red line is the experimentally identified ATE. For each $\kappa < P(D = 1)$ the experimental ATE is contained in the identified set.

Two features of Figure OA.7 are worth recording. First, at $\kappa = 0$ the identified set is the Manski no-assumption interval and has width one for binary outcomes, regardless of the true ATE. Second, the bounds are markedly asymmetric around the experimental ATE. For the threshold $Y >$ median, for example, the identified set at $\kappa = 0.25$ is approximately $[-0.29, +0.38]$, while the experimental ATE is $+0.07$. The width 0.67 is the empirical content the analyst would dispense with by giving up either the deterministic Roy measurement of (1) or the random-assignment measurement of the experiment.

Table OA.6 summarizes the headline numbers.

Table OA.6: Sharp ATE bounds on NSW data under different stochastic-selection overlap restrictions. The threshold $Y >$ median uses the NSW $re78$ median. $\kappa = 0.25$ corresponds to a propensity-score interval $[0.25, 0.75]$. The experimental ATE is point-identified by random assignment.

Sample / threshold	n	$P(D = 1)$	exp. ATE	Manski bd	$\kappa = 0.25$ bd
NSW exp. ($Y > 0$)	445	0.416	$+0.111$	$[-0.479, +0.521]$	$[-0.129, +0.362]$
NSW exp. ($Y >$ median (3702))	445	0.416	$+0.071$	$[-0.465, +0.535]$	$[-0.287, +0.380]$
NSW exp. ($Y > 5000$)	445	0.416	$+0.077$	$[-0.447, +0.553]$	$[-0.263, +0.404]$
NSW exp. stratified	445		$+0.071$	$[-0.465, +0.535]$	$[-0.255, +0.378]$
LaLonde trimmed	676	0.229	$+0.071$	$[-0.466, +0.534]$	infeasible

OA.3.2 Illustration 2: NSW stratified by covariates

If the analyst has covariates that are believed to render selection ignorable within strata, the binary bounds can be applied stratum-by-stratum and aggregated. Panel (a) of Figure OA.8 reports the result of stratifying NSW by $(black, \mathbf{1}\{re75 > 0\})$, giving four cells. The within-stratum bounds are weighted by stratum sizes to form an aggregate identified set. Stratification narrows the identified set modestly: at $\kappa = 0.25$ the stratified set is approximately $[-0.26, +0.38]$, compared with $[-0.29, +0.38]$ without stratification. The narrowing is real but modest, because the four NSW strata still have substantial residual heterogeneity.

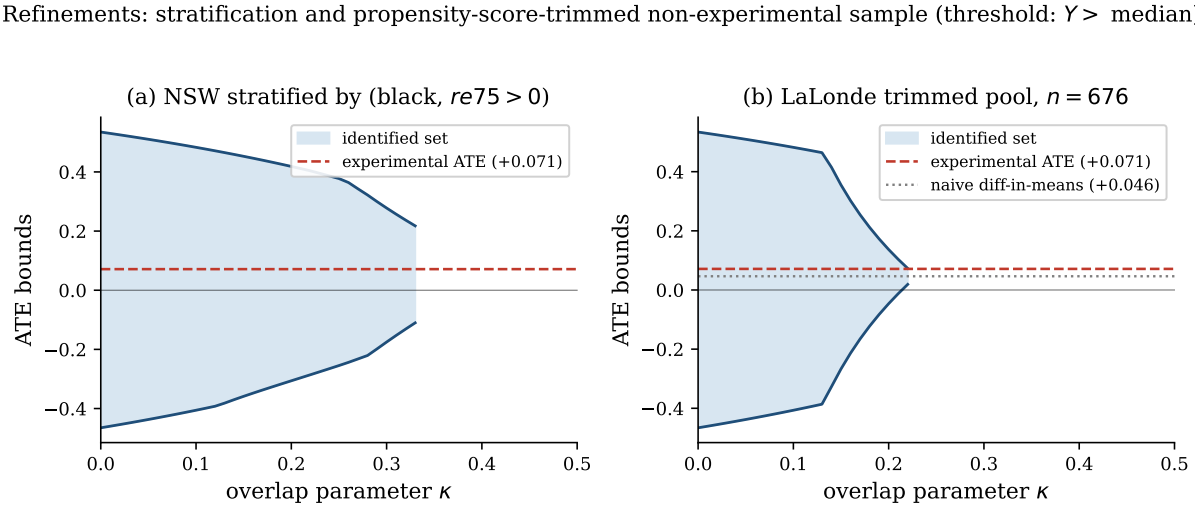

Figure OA.8: Refinements of the binary bounds on NSW with threshold $Y >$ median. Panel (a): bounds computed within strata defined by $(black, \mathbf{1}\{re75 > 0\})$, then aggregated by stratum mass. Panel (b): bounds on the LaLonde nonexperimental sample (NSW treated plus CPS-1 controls), trimmed to estimated propensity scores in $[0.05, 0.95]$. Both panels show the experimental ATE benchmark. Panel (b) also reports the naive difference in means in the trimmed sample.

OA.3.3 Illustration 3: LaLonde nonexperimental sample

LaLonde (1986) constructed a nonexperimental version of the same evaluation by combining NSW treated workers with controls drawn from the Current Population Survey. In the untrimmed NSW-treated/CPS-control comparison, the treated share is about one percent; hence any global overlap restriction with κ above that mass is infeasible before outcomes are even used. This is not a numerical nuisance but a partial-identification diagnostic: a global stochastic-selection model with nontrivial overlap must be imposed on a population with common support. Panel (b) of Figure OA.8 therefore reports the bounds after trimming to estimated propensity scores in $[0.05, 0.95]$; see Dehejia and Wahba (1999). The propensity score is estimated by a logistic model using age, education, race indicators, marital status, no-degree status, and pre-program earnings

$re74, re75$. After trimming, $n = 676$, $P(D = 1) = 0.23$, and the naive difference in means in the trimmed sample is +0.046, already biased downward relative to the experimental ATE of +0.071. The unrestricted bounds at $\kappa = 0$ are essentially Manski, and the constrained bounds at κ close to $P(D = 1)$ again collapse toward the difference in means. Without further structural restrictions, the experimental ATE is contained in the identified set whenever the latter is feasible, but the naive estimator alone misses the experimental benchmark.

OA.3.4 Mapping to the theoretical sections

Three mappings are useful for reading these figures. First, the empirical illustration is an implementation of Theorem 2 through its binary face Proposition 5: it computes the sharp identified set under κ-overlap on observed cells. Second, the curves provide a sensitivity interpretation of κ: moving κ from zero toward the smaller treatment/control mass traces how much empirical content is gained by strengthening the stochastic-selection kernel from unrestricted missingness toward near-observed conditional means. Third, the local fragility result Theorem 1 concerns kernel-distance from the deterministic Roy rule, while κ is a global overlap parameter. Both are aspects of the same broader phenomenon: under unrestricted stochastic selection the data identify only the marginal projections of weighted subdensities, and the empirical content of point identification is recovered only at boundary regimes of either kernel restriction.

The LaLonde comparison gives the empirical exercise a second use. In the untrimmed NSW-treated/CPS-control sample, the treated share is so small that nontrivial values of κ are infea-sible. This is the submeasure characterization in Theorem 2 detecting lack of common support. Propensity-score trimming restores a feasible range of κ and makes the sensitivity curve inter-pretable. Thus the applied analysis illustrates both endpoint contraction and feasibility diagnos-tics.