Positive-unlabeled learning without instruments: identified sets for latent-class moments under selected-at-random labeling

Shaohui Wang[a]

[a] *J. Mack Robinson College of Business, Georgia State University, Atlanta, GA 30303, USA*

Abstract

We study positive-unlabeled data under selected-at-random labeling when no valid instrument is available. We first show that, in this environment, the latent labeling rule and the posterior class probability are not point identified from the observables. We then shift attention from the unidentified latent probability itself to a finite-dimensional latent-class moment target and prove that this target is sharply set identified under a simple positivity bound. Our main result characterizes the exact identified region as a convex set generated by admissible inverse-selection weights and shows that this region depends only on the observable regression. We further derive a support-function representation of the identified set, which delivers sharp bounds on latent prevalence and, for suitable choices of the moment transformation, on positive-class feature means. Finally, we provide a simulation to demonstrate the effectiveness of our method.

Keywords: positive-unlabeled learning, partial identification, selected-at-random labeling, sharp identified set, support function, latent moments
JEL: C14, C21, C51, C63

1. Introduction

Positive-unlabeled (PU) data arise when positive cases can be verified but reliable negative labels are unavailable, so the observed sample consists of

Email address: swang83@gsu.edu (Shaohui Wang)

confirmed positives and a larger unlabeled pool that mixes true negatives with unverified positives (Bekker and Davis, 2020; Liu et al., 2025). This setting appears in medical diagnosis, disease-gene identification, text classification, web-page classification, and related applications (Bekker and Davis, 2020). Under selected-completely-at-random (SCAR) labeling, existing methods can recover class probabilities, estimate class proportions, or construct classification rules using reweighting arguments (du Plessis et al., 2017; Bekker and Davis, 2020). More recent work allows selected-at-random (SAR) or instance-dependent labeling, so the probability that a positive unit is labeled may depend on observables (Gong et al., 2022; Coudray et al., 2023; Liu et al., 2025). Once SAR labeling is allowed, however, identification becomes substantially harder in the absence of excluded variation that separates the labeling rule from the latent class probability (Coudray et al., 2023; Liu et al., 2025).

This note studies the no-instrument SAR-PU case. Let X denote observed features, let $S \in \{0,1\}$ denote the observed positive label, and let $Y^* \in \{0,1\}$ denote the latent class. The data reveal only

$$m(X) = \Pr(S = 1 \mid X) = e(X)p^*(X), \tag{1}$$

where $e(X) = \Pr(S = 1 \mid Y^* = 1, X)$ is the labeling probability for latent positives and $p^*(X) = \Pr(Y^* = 1 \mid X)$ is the latent positive probability. Without an instrument that shifts $e(X)$ independently of $p^*(X)$, the pair (e, p^*) is not point identified. The problem is therefore one of genuine underidentification, closely related to econometric work on contaminated controls, outcome misclassification, and latent binary outcomes (Lancaster and Imbens, 1996; Lewbel, 2000; Meyer and Mittag, 2017).

Rather than trying to recover the full latent posterior $p^*(X)$, we study a lower-dimensional target. For a known transformation $T(X) \in \mathbb{R}^k$, define

$$\lambda_P = E[T(X)\mathbf{1}\{Y^* = 1\}],$$

the unnormalized positive-class moment. This object is economically useful. When $T(X) = 1$, λ_P is latent prevalence. When $T(X) = (1, X)$, the identified set jointly bounds prevalence and positive-class first moments, and therefore positive-class feature means. The question addressed in this note is therefore: what do no-instrument SAR-PU data identify about such latent-class moments?

In this note, we show that, Under no false positives, SAR labeling, and a positivity bound $e(X) \geq \underline{e} > 0$, λ_P is sharply set identified. The identified

set is convex and depends only on the observable regression $m(X)$. We show that the support function can be used to characterize this set exactly and to compute it in practice. We then propose a simple implementation and use simulation evidence to demonstrate the effectiveness of the method.

2. Sharp identified set

[Table 1 about here]

Table 1 summarizes the main notation used in the remainder of the note. We continue to observe (X, S), where $S = 1$ means that an observation is labeled positive and $Y^* \in \{0, 1\}$ denotes the latent class. Throughout this section, we maintain three assumptions: no false positives, selected-at-random (SAR) labeling, and a positivity bound. Formally,

$$\begin{aligned} &\Pr(S = 1 \mid Y^* = 0, X) = 0, \\ &\Pr(S = 1 \mid Y^* = 1, X) = e(X), \\ &\qquad\qquad e(X) \geq \underline{e} > 0. \end{aligned}$$

These restrictions imply

$$m(X) = \Pr(S = 1 \mid X) = e(X)p^*(X),$$

where $p^*(X) = \Pr(Y^* = 1 \mid X)$.

Proposition 1. *Without an instrument or another exclusion-type restriction,* $(e(\cdot), p^*(\cdot))$ *is not point identified from the joint distribution of* (X, S)*.*

The argument is immediate. Choose any measurable $\tilde{e}(X)$ such that

$$\max\{\underline{e}, m(X)\} \leq \tilde{e}(X) \leq 1.$$

Then

$$\tilde{p}^*(X) = \frac{m(X)}{\tilde{e}(X)}$$

lies in $[0, 1]$ and reproduces the same observables. Hence infinitely many latent decompositions are observationally equivalent, so point identification fails.

Because the full latent pair (e, p^*) is not point identified, we turn to the latent positive-class moment

$$\lambda_P = E[T(X)\mathbf{1}\{Y^* = 1\}].$$

Define the inverse-selection weight

$$w(X) = \frac{1}{e(X)}.$$

Then the maintained assumptions imply the deterministic bounds

$$1 \le w(X) \le \bar{w}(X) := \min\left\{\underline{e}^{-1}, m(X)^{-1}\right\}.$$

Using no false positives and SAR labeling, the latent moment can be rewritten in observable form as

$$\lambda_P = E[S\, w(X)\, T(X)]. \tag{2}$$

Thus λ_P is pinned down by observables and an unknown weight function that is only interval identified.

Theorem 1. *Under no false positives, SAR labeling, and the positivity bound $e(X) \ge \underline{e} > 0$, the sharp identified set for λ_P is*

$$\mathcal{L}_P = \{E[S\, w(X)\, T(X)] : 1 \le w(X) \le \bar{w}(X) \text{ almost surely}\}. \tag{3}$$

The set is nonempty and convex. Moreover, for any $\lambda \in \mathcal{L}_P$, there exists an observationally equivalent SAR-PU model that attains it.

Appendix A proves nonemptiness, convexity, and sharpness. The key idea behind sharpness is constructive. For any admissible weight w, define

$$e_w(X) = \frac{1}{w(X)}, \qquad p_w(X) = m(X)w(X).$$

These objects satisfy

$$e_w(X)p_w(X) = m(X),$$

so they reproduce the observed distribution of (X, S). At the same time, they generate the latent moment

$$E[S\, w(X)\, T(X)].$$

Thus $\mathcal{L}_P$ is not merely an outer bound; it is the exact region implied by the data and the maintained assumptions.

This distinction matters conceptually. The no-instrument SAR-PU problem is not simply a weak-information problem around a point-identified target. The data are compatible with a continuum of latent SAR structures, so the natural inferential object is a set. The gain from focusing on λ_P is that this lower-dimensional target survives the missing-label problem and remains economically interpretable.

Because $\mathcal{L}_P$ is convex, it can be characterized by its support function, as in the literature on convex identified sets and support-function-based inference (Beresteanu et al., 2011; Kaido, 2016). For any direction vector $q \in \mathbb{R}^k$, define

$$h_P(q) = \sup_{\lambda \in \mathcal{L}_P} q^\top \lambda.$$

The support function is especially convenient here because the maximization problem is linear in the admissible weight $w(X)$, while the constraint $1 \leq w(X) \leq \bar{w}(X)$ is pointwise. Hence, for each direction q, the maximizing weight sets $w(X)$ equal to its upper bound when $q^\top T(X) \geq 0$, and to its lower bound otherwise. This yields

$$h_P(q) = E\Bigg(\Bigg[\min\left\{ \frac{m(X)}{\underline{e}}, 1 \right\} \mathbf{1}\{q^\top T(X) \geq 0\} + m(X)\mathbf{1}\{q^\top T(X) < 0\} \Bigg] q^\top T(X) \Bigg). \tag{4}$$

Equation (4) depends only on the observable regression $m(X)$, not on the latent pair (e, p^*). It therefore provides both a sharp characterization of the identified set and a direct route to computation (Beresteanu et al., 2011; Kaido, 2016; Canay and Shaikh, 2017).

Corollary 1. *If prevalence $\pi = \Pr(Y^* = 1)$ is known, then the negative-class moment*

$$\eta_N = E[T(X) \mid Y^* = 0]$$

is sharply identified through the affine transformation

$$\eta_N = \frac{E[T(X)] - \lambda_P}{1 - \pi}.$$

Hence the sharp identified set for η_N is the affine image of $\mathcal{L}_P$ (Beresteanu et al., 2012; Kline and Tamer, 2023)

3. A simple implementation

The identified-set result in Section 2 admits a direct sample implementation. For an i.i.d. sample $\{(X_i, S_i)\}_{i=1}^n$, let $\hat{m}_i$ be a cross-fitted estimator of the observable regression $m(X_i)$, and define the estimated upper weight bound

$$\hat{\bar{w}}_i = \min\{\underline{e}^{-1}, \hat{m}_i^{-1}\}.$$

A natural raw sample analogue of (3) is

$$\widehat{\mathcal{L}}_{P,n}^{\text{raw}} = \left\{ \frac{1}{n} \sum_{i=1}^n S_i w_i T(X_i) : 1 \le w_i \le \hat{\bar{w}}_i \right\}.$$

This raw set is useful for geometric summaries and visualization because it provides a direct plug-in approximation to the identified region. However, it is not itself a confidence procedure.

For formal inference, it is more convenient to work with the support-function representation in (4). To turn that representation into a practical estimator that is less sensitive to first-stage error, define

$$\rho_q(x, m) = \left[\min\left\{ \frac{m}{\underline{e}}, 1 \right\} \mathbf{1}\{q^\top T(x) \ge 0\} + m \mathbf{1}\{q^\top T(x) < 0\} \right] q^\top T(x),$$

and

$$\dot{\rho}_q(x, m) = \left[\underline{e}^{-1} \mathbf{1}\{m < \underline{e},\ q^\top T(x) \ge 0\} + \mathbf{1}\{q^\top T(x) < 0\} \right] q^\top T(x).$$

The resulting orthogonal support-function estimator is

$$\check{h}_n(q) = \frac{1}{n} \sum_{i=1}^n \left[\rho_q(X_i, \hat{m}_i) + \dot{\rho}_q(X_i, \hat{m}_i)\{S_i - \hat{m}_i\} \right]. \tag{5}$$

This estimator provides a practical way to evaluate the support function of the identified set using only the observable regression and sample data.

Let $\mathcal{Q}_M = \{q_1, \ldots, q_M\} \subset \mathbb{R}^k$ be a finite direction grid, and let $\alpha \in (0, 1)$ be a nominal level. A $(1 - \alpha)$ confidence region is

$$\widehat{\mathcal{C}}_{1-\alpha} = \bigcap_{q \in \mathcal{Q}_M} \left\{ \lambda : q^\top \lambda \le \check{h}_n(q) + \hat{c}_{1-\alpha}/\sqrt{n} \right\},$$

where $\hat{c}_{1-\alpha}$ is the estimated $(1 - \alpha)$ quantile of the multiplier-bootstrap approximation to the support-function process. This finite-grid implementation

follows the standard support-function logic for inference in partially identified models (Kaido, 2016).

To illustrate the method, we report evidence from a Monte Carlo simulation. Because the latent class Y^* is generated by the data-generating process, simulation provides two useful benchmarks that are unavailable in real PU data: the true latent moment and the oracle identified set obtained by replacing $\hat{m}_i$ with the true regression $m(X_i)$.

Table 2 reports the baseline design with $n = 12{,}000$, 5 cross-fitting folds, $\underline{e} = 0.15$, and 399 multiplier-bootstrap draws. The raw set tracks the oracle identified region closely on average, but it is not a valid confidence procedure: the true latent moment lies in it in 93.3% of replications. The bootstrap region is deliberately wider, and in the reported replications it contains both the true latent moment and the oracle set throughout.

[Table 2 about here]

Figure 1 shows a representative replication from the baseline design. The raw set lies close to the oracle set, while the confidence region expands to absorb first-stage and sampling uncertainty.

[Figure 1 about here]

Figure 2 reports the sensitivity exercise over $\underline{e}$. As the positivity lower bound rises from 0.10 to 0.40, the mean raw area falls from 0.232 to 0.069, and the mean upper-bound weight falls from 5.42 to 2.42. Thus the identified region tightens only when the maintained identifying restriction becomes more informative.

[Figure 2 about here]

In practice, $\underline{e}$ may be motivated by institutional knowledge about minimum revelation rates, or treated as a sensitivity parameter in the spirit of partial-identification analysis more broadly (Kline and Tamer, 2023). The simulation evidence shows that this single restriction has a transparent effect on identification strength: weaker lower bounds permit larger inverse-selection weights and therefore wider latent-moment regions.

4. Conclusion

This note asks what no-instrument SAR-PU data identify. We show that the latent labeling rule and latent class probability are not point identified from (X, S) alone, but positive-class moments remain sharply set identified under no false positives, SAR labeling, and a simple positivity bound. When $T(X) = 1$, this yields sharp bounds on latent prevalence; when $T(X) = (1, X)$, it yields sharp bounds on prevalence and positive-class feature means.

The identified region is convex and admits a support-function representation based only on the observable regression $m(X)$. This leads to a simple implementation based on cross-fitting and bootstrap confidence regions. Simulation evidence shows that the identified region widens when labeling becomes less informative and tightens when the positivity restriction becomes stronger.

The broader implication is that the no-instrument SAR-PU problem should be treated as a partial-identification problem. Rather than forcing point recovery where the data do not support it, the note characterizes exactly which latent-class moments remain identified (Gong et al., 2022; Coudray et al., 2023; Liu et al., 2025).

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The manuscript uses simulated data only. The code used to generate the figures and tables in this submission package is included in the supplementary files and can be shared through a public repository upon request or acceptance.

Declaration of generative AI use

During the preparation of this manuscript, the authors used AI-assisted tools to improve language, and prepare draft LaTeX files. All mathematical arguments, simulation choices, interpretations were provided by authors, who take full responsibility for the submitted content.

Table 1: Main notation

Symbol	Definition
Core model and identification	
X	Observed feature vector
Y^*	Latent positive-class indicator
S	Observed positive label
n	Sample size
k	Dimension of the moment transformation
$m(X)$	Observable regression $\Pr(S = 1 \mid X)$
$e(X)$	Labeling probability for latent positives
$p^*(X)$	Latent class probability $\Pr(Y^* = 1 \mid X)$
$T(X)$	User-chosen moment transformation
$\underline{e}$	Positivity lower bound on $e(X)$
$w(X)$	Inverse-selection weight $1/e(X)$
$\bar{w}(X)$	Upper weight bound $\min\{\underline{e}^{-1}, m(X)^{-1}\}$
λ_P	Positive-class moment $E[T(X)\mathbf{1}\{Y^* = 1\}]$
π	Latent prevalence $\Pr(Y^* = 1)$
q	Direction vector in $\mathbb{R}^k$
$h_P(q)$	Support function of $\mathcal{L}_P$ in direction q
η_N	Negative-class mean $(E[T(X)] - \lambda_P)/(1 - \pi)$ when π is known
$\mathcal{L}_P$	Sharp identified set for λ_P
Implementation and inference	
$\hat{m}_i$	Cross-fitted estimator of $m(X_i)$
$\hat{\bar{w}}_i$	Estimated upper weight bound $\min\{\underline{e}^{-1}, \hat{m}_i^{-1}\}$
$\widehat{\mathcal{L}}_{P,n}^{\text{raw}}$	Raw sample analogue of the sharp identified set
$\rho_q(x, m)$	Direction-specific support-function integrand
$\dot{\rho}_q(x, m)$	Influence-adjustment term for the orthogonal estimator
$\check{h}_n(q)$	Orthogonal support-function estimator
$\mathcal{Q}_M$	Finite direction grid used in computation and inference
α	Nominal significance level
$\hat{c}_{1-\alpha}$	Bootstrap critical value for support-function inference
$\widehat{\mathcal{C}}_{1-\alpha}$	Bootstrap confidence region for the identified set

Table 2: Baseline Monte Carlo summary

Statistic	Value
Raw point containment	0.933
Confidence-set point coverage	1.000
Confidence-set region coverage	1.000
Mean raw area	0.209
Mean oracle area	0.222
Mean confidence-region area	0.310
$\mathrm{Corr}(\hat{m}, m)$	0.675
Mean $\bar{w}(X)$	4.836

Notes: Based on 30 Monte Carlo repetitions under the baseline SAR-PU design reported in the original draft. Because the latent class Y^* is generated in simulation, both the true latent moment and the oracle set are known benchmarks. The oracle set replaces $\hat{m}_i$ with the true regression $m(X_i)$. Raw point containment is the share of replications in which the true latent moment lies in $\widehat{\mathcal{L}}_{P,n}^{\mathrm{raw}}$. Confidence-set point coverage is the share of replications in which the true latent moment lies in the bootstrap confidence region. Confidence-set region coverage is the share of replications in which the oracle set is contained in the bootstrap confidence region. $\mathrm{Corr}(\hat{m}, m)$ is the correlation between estimated and true regression values, and Mean $\bar{w}(X)$ is the average upper-bound weight. The confidence region is computed from the orthogonal support-function estimator with a multiplier bootstrap.

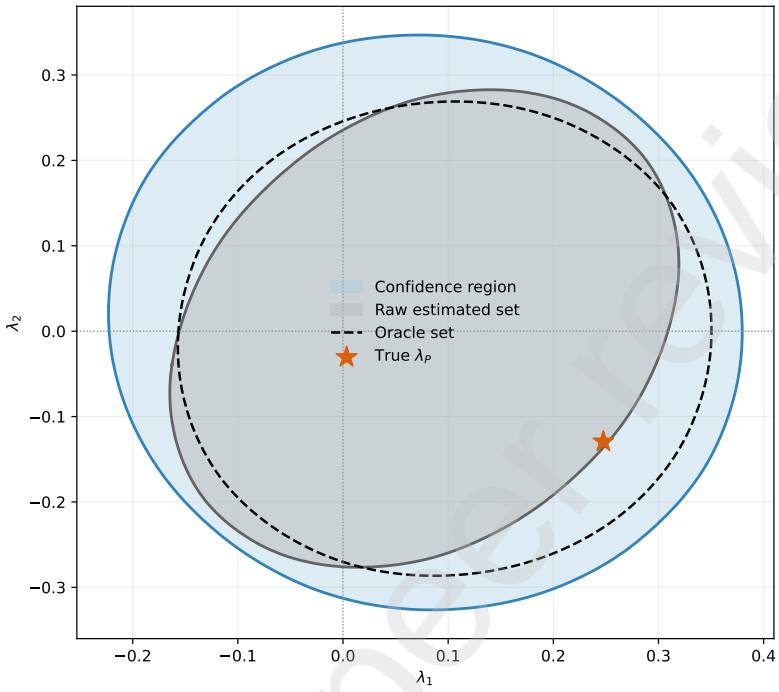

Figure 1: Representative identified-set geometry. The dashed polygon is the oracle set, the filled darker region is the raw estimated set, and the lighter envelope is the bootstrap confidence region. The star marks the true latent moment.

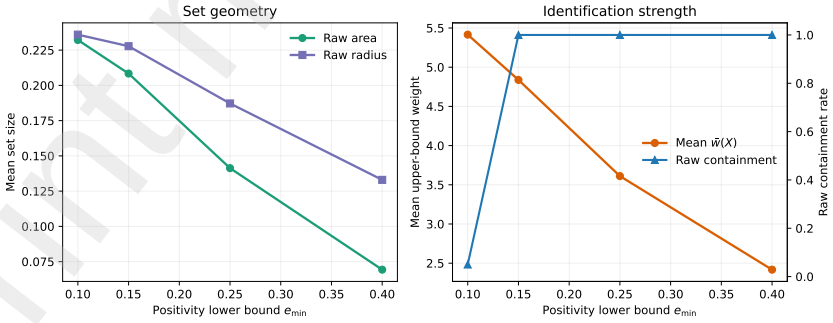

Figure 2: Sensitivity to the positivity lower bound. A larger $\underline{e}$ tightens the admissible inverse-selection weights and shrinks the identified region. Values are drawn from the reported sensitivity exercise in the Monte Carlo study.

References

Bekker, J., Davis, J., 2020. Learning from positive and unlabeled data: A survey. Machine Learning 109, 719–760. doi:10.1007/s10994-020-05877-5.

Beresteanu, A., Molchanov, I., Molinari, F., 2011. Sharp identification regions in models with convex moment predictions. Econometrica 79, 1785–1821.

Beresteanu, A., Molchanov, I., Molinari, F., 2012. Partial identification using random set theory. Journal of Econometrics 166, 17–32. doi:10.1016/j.jeconom.2011.06.003.

Canay, I.A., Shaikh, A.M., 2017. Practical and theoretical advances in inference for partially identified models, in: Advances in Economics and Econometrics. Cambridge University Press, pp. 271–306. doi:10.1017/9781108227223.009.

Coudray, O., Keribin, C., Massart, P., Pamphile, P., 2023. Risk bounds for positive-unlabeled learning under the selected at random assumption. Journal of Machine Learning Research 24, 1–31.

Gong, C., Wang, Q., Liu, T., Han, B., You, J., Yang, J., Tao, D., 2022. Instance-dependent positive and unlabeled learning with labeling bias estimation. IEEE Transactions on Pattern Analysis and Machine Intelligence 44, 4163–4177. doi:10.1109/TPAMI.2021.3061456.

Kaido, H., 2016. A dual approach to inference for partially identified econometric models. Journal of Econometrics 192, 269–290. doi:10.1016/j.jeconom.2015.12.017.

Kline, B., Tamer, E., 2023. Recent developments in partial identification. Annual Review of Economics 15, 125–150. doi:10.1146/annurev-economics-051520-021124.

Lancaster, T., Imbens, G., 1996. Case-control studies with contaminated controls. Journal of Econometrics 71, 145–160. doi:10.1016/0304-4076(94)01698-4.

Lewbel, A., 2000. Identification of the binary choice model with misclassification. Econometric Theory 16, 603–609. doi:10.1017/S0266466600164060.

Liu, S., Yeh, C.K., Zhang, X., Tian, Q., Li, P., 2025. Positive and unlabeled data: Model, estimation, inference, and classification. Journal of the American Statistical Association 120, 2547–2558. doi:10.1080/01621459.2025.2474266.

Meyer, B.D., Mittag, N., 2017. Misclassification in binary choice models. Journal of Econometrics 200, 295–311. doi:10.1016/j.jeconom.2017.06.012.

du Plessis, M.C., Niu, G., Sugiyama, M., 2017. Class-prior estimation for learning from positive and unlabeled data. Machine Learning 106, 463–492. doi:10.1007/s10994-016-5604-6.

Appendix A. Appendix: Proofs and Simulation Details

Throughout this appendix, all expectations are assumed finite so that λ_P and the support function $h_P(q)$ are well defined. This appendix provides proofs of Proposition 1, Theorem 1, Corollary 1, and the support-function formula, and then records the main simulation design used in the note.

Appendix A.1. Proof of Proposition 1

Proof. Recall that the observables satisfy

$$m(X) = e(X)p^*(X) \qquad \text{a.s.}$$

under the maintained assumptions. To show failure of point identification, fix any measurable function $\tilde{e}(X)$ such that

$$\max\{\underline{e}, m(X)\} \le \tilde{e}(X) \le 1 \qquad \text{a.s.}$$

and define

$$\tilde{p}^*(X) = \frac{m(X)}{\tilde{e}(X)}.$$

Because $\tilde{e}(X) \ge m(X)$, it follows that $\tilde{p}^*(X) \in [0,1]$ almost surely. Moreover,

$$\tilde{e}(X)\tilde{p}^*(X) = m(X) = \Pr(S = 1 \mid X).$$

Hence the pair $(\tilde{e}, \tilde{p}^*)$ generates exactly the same observables as the true latent pair (e, p^*). Since there are infinitely many admissible choices of $\tilde{e}$, point identification fails. □

Appendix A.2. Proof of Theorem 1

Proof. We verify nonemptiness, convexity, and sharpness in turn.

Step 1: Nonemptiness and inclusion of the true moment.. Let

$$w_0(X) = \frac{1}{e(X)}$$

denote the true inverse-selection weight. Since $e(X) \in [\underline{e}, 1]$ and

$$m(X) = e(X)p^*(X) \leq e(X),$$

we have

$$1 \leq w_0(X) \leq \min\{\underline{e}^{-1}, m(X)^{-1}\} = \bar{w}(X).$$

Thus w_0 is admissible. Under no false positives and SAR labeling,

$$E\left[\frac{S}{e(X)}T(X) \,\middle|\, X, Y^*\right] = Y^*T(X).$$

Taking expectations gives

$$\lambda_P = E[T(X)\mathbf{1}\{Y^* = 1\}] = E[S\, w_0(X)\, T(X)] \in \mathcal{L}_P.$$

Therefore $\mathcal{L}_P$ is nonempty and contains the true latent moment.

Step 2: Convexity.. Take any $\lambda_1, \lambda_2 \in \mathcal{L}_P$. By definition, there exist admissible weights w_1 and w_2 such that

$$\lambda_j = E[S\, w_j(X)\, T(X)], \qquad j = 1, 2.$$

For any $a \in [0, 1]$, define

$$w_a(X) = aw_1(X) + (1 - a)w_2(X).$$

Because the admissible weight class is determined by pointwise interval restrictions,

$$1 \leq w_a(X) \leq \bar{w}(X) \qquad \text{a.s.}$$

Hence w_a is admissible, and

$$a\lambda_1 + (1 - a)\lambda_2 = E[S\, w_a(X)\, T(X)] \in \mathcal{L}_P.$$

Therefore $\mathcal{L}_P$ is convex.

Step 3: Sharpness.. Take any admissible w in (3). Define

$$e_w(X) = \frac{1}{w(X)}, \qquad p_w(X) = m(X)w(X).$$

Because $1 \le w(X) \le \bar{w}(X)$, we have

$$e_w(X) \in (0, 1], \qquad p_w(X) \in [0, 1].$$

Thus e_w and p_w define valid conditional probabilities.

Now construct a latent variable $Y_w^* \in \{0, 1\}$ such that

$$\Pr(Y_w^* = 1 \mid X) = p_w(X),$$

and define the observed label rule by

$$\Pr(S = 1 \mid Y_w^* = 1, X) = e_w(X), \qquad \Pr(S = 1 \mid Y_w^* = 0, X) = 0.$$

Then

$$\Pr(S = 1 \mid X) = e_w(X)p_w(X) = m(X),$$

so the constructed model is observationally equivalent to the original one.

Its implied positive-class moment is

$$\begin{aligned} E[T(X)\mathbf{1}\{Y_w^* = 1\}] &= E[p_w(X)T(X)] \\ &= E[m(X)w(X)T(X)] \\ &= E[S\, w(X)\, T(X)]. \end{aligned}$$

Therefore every element of $\mathcal{L}_P$ is attainable under some observationally equivalent SAR-PU model. This proves sharpness. □

Appendix A.3. Proof of Corollary 1

Proof. If prevalence $\pi = \Pr(Y^* = 1)$ is known, then

$$E[T(X)] = \pi E[T(X) \mid Y^* = 1] + (1 - \pi)E[T(X) \mid Y^* = 0].$$

Using the definitions of λ_P and η_N,

$$E[T(X)] = \lambda_P + (1 - \pi)\eta_N.$$

Solving for η_N gives

$$\eta_N = \frac{E[T(X)] - \lambda_P}{1 - \pi}.$$

Hence the sharp identified set for η_N is the affine image of $\mathcal{L}_P$. □

Appendix A.4. Derivation of the support-function formula

For any $q \in \mathbb{R}^k$, define the support function

$$h_P(q) = \sup_{\lambda \in \mathcal{L}_P} q^\top \lambda.$$

Using the representation in (3),

$$h_P(q) = \sup_{1 \le w \le \bar{w}} E[S\, w(X)\, q^\top T(X)].$$

By iterated expectations,

$$E[S\, w(X)\, q^\top T(X)] = E[m(X) w(X) q^\top T(X)].$$

Because the admissible restriction on w is pointwise and the objective is linear in w, the maximization can be solved pointwise. The maximizing weight is therefore

$$w_q(X) = \bar{w}(X)\mathbf{1}\{q^\top T(X) \ge 0\} + \mathbf{1}\{q^\top T(X) < 0\}.$$

This function is measurable and admissible. Substituting it into the objective yields

$$\begin{aligned} h_P(q) = E\Big(& m(X)\bar{w}(X)\mathbf{1}\{q^\top T(X) \ge 0\} q^\top T(X) \\ & + m(X)\mathbf{1}\{q^\top T(X) < 0\} q^\top T(X) \Big). \end{aligned}$$

Finally, since

$$\bar{w}(X) = \min\{\underline{e}^{-1}, m(X)^{-1}\},$$

we have

$$m(X)\bar{w}(X) = \min\left\{ \frac{m(X)}{\underline{e}}, 1 \right\}.$$

Substituting this identity yields equation (4).

Appendix A.5. Simulation design

The baseline design in the simulations sets

$$n = 12{,}000, \qquad d = 2, \qquad \underline{e} = 0.15,$$

and uses a two-dimensional moment transformation $T(X)$, five-fold cross-fitting, and 399 multiplier-bootstrap draws. The data-generating process is logistic in both the latent positive probability and the labeling rule. Figure 1 reports a representative replication from this baseline design. Figure 2 reports the sensitivity exercise over

$$\underline{e} \in \{0.10, 0.15, 0.25, 0.40\}.$$

These simulations are used to compare the raw identified set, the oracle identified set, and the bootstrap confidence region, and to illustrate how identification strength changes with the positivity lower bound.