LEARNING-AUGMENTED STRATEGYPROOF MECHANISMS FOR TWO-FACILITY LOCATION

Anonymous authors
Paper under double-blind review

ABSTRACT

In learning-augmented mechanism design for two-facility location, we examine strategyproof mechanisms that, given access to predictions, elicit facility location preferences of the agents truthfully and determine two facility locations on the line that approximately minimize the utilitarian social cost, defined as the sum of agents' distances to their closest facilities. Because the predictions may be inaccurate, we evaluate the mechanisms' performance in terms of their consistency and robustness, which are the approximation guarantees on the utilitarian social cost when the prediction is correct and when the prediction is arbitrary, respectively. We consider two types of predictions, namely the locations of an optimal pair of facilities and the preferred location of each agent, and investigate the consistency and robustness of strategyproof mechanisms with access to each of the two prediction types. For deterministic strategyproof mechanisms with $n \geq 5$ agents, we characterize the optimal consistency-robustness trade-off for each prediction type and provide mechanisms that attain the trade-off exactly. Predictions of optimal facility locations cannot improve upon the optimal prediction-free approximation ratio, whereas predictions of agent preferences yield a nontrivial trade-off. For randomized mechanisms, we provide mechanisms that establish upper bounds on the consistency-robustness trade-off for each prediction type, and complement them with a uniform lower bound.

1 INTRODUCTION

In recent decades, mechanism design for facility location has attracted significant attention in various communities, including computer science, economics, and operations research (Procaccia & Tennenholtz, 2013; Chan et al., 2021; Moulin, 1980; Drezner & Hamacher, 2002). In a typical setting, a social planner is tasked with determining locations of public facilities (e.g., parks, libraries, or schools) in a region by eliciting the facility location preferences of the agents, who may have incentives to strategically misreport their preferred locations to manipulate the planner's decisions. As a result, much of the work in this area is devoted to designing and analyzing strategyproof mechanisms that elicit agents' preferred locations truthfully and determine facility locations that approximately minimize a given cost objective. Such objectives are defined in terms of the agents' distances to their nearest facility, most commonly the utilitarian social cost (i.e., the sum of these distances) or the egalitarian social cost (i.e., the maximum of these distances). Beyond theoretical interests, this area has practical motivations in social planning (Drezner & Hamacher, 2002) and preference aggregation (Black, 1948; Moulin, 1980).

Mechanism Design for Two-Facility Location. For locating a single facility on the line, it is known that the median mechanism that locates the facility at the median of the agents' reported locations is strategyproof and minimizes the utilitarian social cost (Moulin, 1980; Procaccia & Tennenholtz, 2013). For locating two facilities on the line, Procaccia & Tennenholtz (2013) showed that the TWO-EXTREMES mechanism that locates the facilities at the leftmost and rightmost of agents' reported locations is strategyproof and achieves an approximation ratio of $n - 2$ for the utilitarian social cost, where n is the number of agents. They also showed that no deterministic strategyproof mechanism has an approximation ratio below $\frac{3}{2}$ uniformly over all population sizes. This lower bound was subsequently improved to a value approaching 2 (Lu et al., 2009) and to $\frac{n-1}{2}$ (Lu et al., 2010), and Fotakis & Tzamos (2014) finally proved a lower bound of $n - 2$ for $n \geq 5$, which

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

matches the ratio of TWO-EXTREMES and thus establishes that TWO-EXTREMES is optimal among deterministic strategyproof mechanisms.

In terms of randomized strategyproof mechanisms, Lu et al. (2009) established a lower bound of 1.045 on the approximation ratio attainable uniformly over all population sizes and provided a mechanism that achieves an approximation ratio of $\frac{n}{2}$. Later, Lu et al. (2010) provided the Proportional mechanism, which achieves an approximation ratio of 4 for any metric space.

Very recently, several independent and concurrent works have improved the randomized bounds. These works share the idea of mixing Proportional with a mechanism that samples a pair of reports with probability proportional to their distance. Ma & Peng (2026) introduced a mechanism that achieves an approximation ratio of 3.6667 and improved the lower bound to 1.2071. Jia (2026) and Chan et al. (2026) independently developed mechanisms that further improve the approximation ratio to 3.5186. Chan et al. establish this guarantee not only on the line but throughout Ptolemaic metric spaces.

Learning-Augmented Mechanism Design. To improve upon classical worst-case guarantees when accurate predictions are available, Xu & Lu (2022) studied learning-augmented mechanism design for two-facility location on the line, while Agrawal et al. (2022) studied single-facility location on the line and in the Euclidean plane. In learning-augmented mechanism design for facility location, the mechanism is given access to some predictions and is evaluated in terms of its *consistency* and *robustness*, which are the approximation guarantees when the prediction is correct and when the prediction is arbitrary, respectively. The prediction is external information and remains fixed when an agent changes her report. Strategyproofness is required for every prediction, including incorrect ones.

Given the mechanism's access to the predictions of the agents' preferred locations, Xu & Lu (2022) provided a deterministic strategyproof mechanism that achieves $\left(1+\frac{n}{2}\right)$-consistency and $(2n-1)$-robustness, and asked whether one can achieve $o(n)$-consistency with bounded robustness in this deterministic setting. Lu et al. (2024) showed that the robustness of this mechanism is in fact $n-1$, and answered the question negatively for deterministic mechanisms. That is, for $n \geq 5$, every deterministic strategyproof mechanism with bounded robustness has consistency $\Omega(n)$.

These results leave a more precise question open: for a desired consistency c, what is the smallest robustness $R(c)$ that a strategyproof mechanism can achieve? For deterministic mechanisms, consistency must grow at least linearly in n when robustness is finite, but the tight trade-off is unknown. For randomized mechanisms, while ignoring predictions already gives constant consistency and robustness, the question is how much accurate predictions can improve consistency, possibly bringing it arbitrarily close to 1, and what robustness is required in return. We also distinguish predictions of the agents' preferred locations from predictions of an optimal facility pair. An optimal pair specifies where facilities should be placed, but does not identify the agents assigned to each facility or their positions within each cluster. A full agent profile supplies this additional information. In this paper, we address these questions for both prediction types by analyzing the consistency-robustness trade-offs achievable by deterministic and randomized strategyproof mechanisms for locating two facilities on the line.

1.1 OUR CONTRIBUTIONS

We study full-profile and facility-pair predictions under the utilitarian social-cost objective.

- **Deterministic mechanisms.** For every fixed $n \geq 5$, we characterize the tight consistency-robustness trade-off under both prediction types. Full-profile predictions improve consistency at the cost of robustness (Theorem 1), while facility-pair predictions cannot improve the prediction-free guarantee (Theorem 2).
- **Randomized mechanisms.** We give explicit upper bounds independent of the number of agents, as shown in Figure 1. Full-profile predictions allow consistency arbitrarily close to 1 with finite robustness, and the bounds extend to the prediction-free endpoint (Theorem 3). A simpler directional sampling mechanism handles facility-pair predictions (Theorem 4). A common lower bound rules out perfect consistency with uniformly finite robustness (Theorem 5).

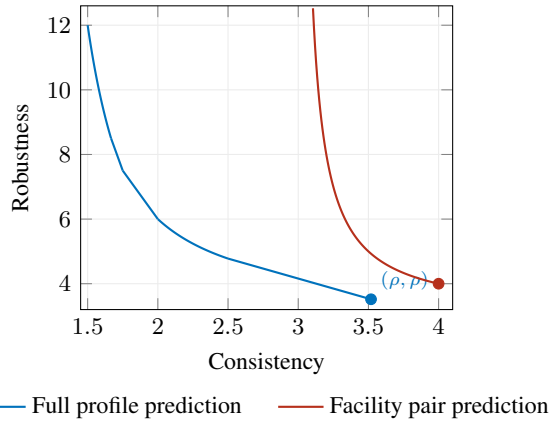

Figure 1: Randomized upper bounds for the two prediction types, with $\rho = \frac{74+4\sqrt{3}}{23}$.

Organization Section 2 introduces the model, prediction types, and performance measures. Section 3 characterizes the tight consistency-robustness trade-offs for deterministic mechanisms. Section 4 presents randomized mechanisms and their guarantees, together with a lower bound under either prediction type. Section 5 concludes with open questions. The remaining proofs and additional technical details appear in Appendices A and B.

1.2 Additional Related Work

Beyond the utilitarian social cost, a widely studied objective in facility location is the egalitarian (maximum) cost. For a single facility on the line, Procaccia & Tennenholtz (2013) showed that the median mechanism achieves an approximation ratio of 2, which is best possible for deterministic strategyproof mechanisms, and that randomization improves this to a best-possible $\frac{3}{2}$. For two facilities, they showed that Two-Extremes achieves the optimal deterministic approximation ratio of 2 and gave a randomized strategyproof mechanism with approximation ratio $\frac{5}{3}$. Other objectives studied for locating a single facility on the line include the L_p norm of the agents' costs (Feigenbaum et al., 2017) and the sum of squared distances (Feldman & Wilf, 2013), for which the median and simple randomized mechanisms achieve constant approximation ratios. Other variants change the agents' preferences or service rules, including heterogeneous facilities (Serafino & Ventre, 2016), capacitated facilities (Aziz et al., 2020b;a), and multiple facilities with concave connection costs (Fotakis & Tzamos, 2013). These differ from our setting of two uncapacitated facilities serving each agent at the nearest location. For a broader account, we refer readers to the survey of Chan et al. (2021).

Regarding learning-augmented mechanisms for single-facility location, Agrawal et al. (2022) showed that on the line, for the egalitarian objective, a prediction of the optimal location yields a deterministic mechanism that is 1-consistent and 2-robust. In the Euclidean plane, they provided the optimal consistency-robustness trade-offs among deterministic anonymous strategyproof mechanisms for both the egalitarian and utilitarian objectives. Balkanski et al. (2024) determined the optimal randomized consistency-robustness trade-off for the maximum-distance objective on the line. Barak et al. (2024) study multi-facility location with mostly approximately correct advice and use fixed-anchor proportional sampling. Their advice model allows arbitrary errors in some predicted agent locations and bounded errors in the rest. In contrast, our robustness requirement imposes no accuracy condition on any part of the prediction. Christodoulou et al. (2024) study mechanism design with output advice, where a recommended outcome replaces predictions of agents' private types. This distinction is closely related to our comparison between a predicted facility pair and a complete labeled profile. Further work considers predictions for candidate locations (Fang et al., 2024) and the envy ratio (Aziz et al., 2025). For two facilities, Walsh (2025) studies predictions of facility locations under maximum-cost and minimum-utility objectives, rather than the utilitarian social cost considered here.

2 PRELIMINARIES

We study the problem of locating two facilities on the real line. Let $N = \{1, \ldots, n\}$ be the set of agents, where agent $i \in N$ has a true location $x_i \in \mathbb{R}$. The vector $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ is called a location profile, and different agents may occupy the same location. In displayed profiles, a^k denotes k agents located at a. When $n = 2$, placing one facility at each reported location is strategyproof and always gives zero social cost under truthful reporting, independently of the prediction. We therefore restrict attention to $n \geq 3$.

Each agent reports a location in $\mathbb{R}$, which need not be her true location. Given the reports and a prediction, a deterministic mechanism selects an unordered pair of facilities $Y = \{y_1, y_2\}$; a randomized mechanism selects a probability distribution over such pairs. The two facilities may coincide and may be placed anywhere on the real line, including outside the interval spanned by the reports. Each agent uses whichever facility is closer to her true location. For a point z, write $d(z, Y) = \min\{|z - y_1|, |z - y_2|\}$, so agent i incurs cost $d(x_i, Y)$. Our objective is to minimize the *utilitarian social cost*, the sum of these individual costs. The social cost of a pair Y and the minimum possible social cost on profile x are

$$\mathrm{SC}(Y; x) = \sum_{i=1}^{n} d(x_i, Y), \qquad \mathrm{OPT}(x) = \min_{y_1, y_2 \in \mathbb{R}} \mathrm{SC}(\{y_1, y_2\}; x).$$

For a randomized mechanism, we evaluate the expected social cost, with the expectation taken over the mechanism's internal randomness. The benchmark $\mathrm{OPT}(x)$ is computed from the true locations, without any strategic or prediction constraint. In particular, randomizing the facilities cannot improve this benchmark.

When sorting a profile, we retain the agents' identities and break ties by a fixed ordering of identities. An optimal solution can be obtained by partitioning the sorted locations into two nonempty consecutive clusters and selecting a median of each. Indeed, any ordered pair of facilities divides the line into two nearest-facility regions, and a median minimizes the sum of distances within each resulting cluster. Neither the optimal partition nor the medians need be unique; whenever a construction needs to choose among them, we fix a deterministic tie-breaking rule. We write P, Q for the two clusters' one-facility optimal costs, so that $P + Q = \mathrm{OPT}(x)$. The optimum is zero exactly when at most two distinct locations are occupied.

Predictions and Strategyproofness. We use π to denote the prediction in either of the following two models. In the *full-profile model*, $\pi = \widehat{x} \in \mathbb{R}^n$ is a labeled vector, and it is correct when $\widehat{x} = x$. Thus $\widehat{x}_i$ predicts the location of the particular agent i, rather than merely specifying an unlabeled collection of locations. In the *facility-pair model*, the prediction is an unordered pair $\pi = \{\widehat{a}, \widehat{b}\}$, and it is correct when $\mathrm{SC}(\pi; x) = \mathrm{OPT}(x)$. Every optimal pair counts as a correct prediction, including when the optimum is not unique. Unlike a full-profile prediction, a facility-pair prediction does not specify the agents' locations or their assignments to the two facilities. We impose no probabilistic assumption on prediction errors: an incorrect prediction may be arbitrary. The prediction is external and is held fixed in every strategic comparison; it is not recomputed from a deviating agent's report.

Write $M(x; \pi)$ for the mechanism's output on reported profile x and prediction π. For an alternative report x_i', the notation (x_i', x_{-i}) means that only agent i changes her report. Strategyproofness requires that such a deviation cannot reduce her cost, measured from her true location x_i.

Definition 1 (Strategyproofness)**.** A possibly randomized mechanism M is strategyproof if, for every profile x, prediction π, agent i, and report $x_i' \in \mathbb{R}$,

$$\mathbb{E}\big[d(x_i, M(x; \pi))\big] \leq \mathbb{E}\big[d(x_i, M(x_i', x_{-i}; \pi))\big].$$

The requirement applies to every fixed prediction, whether correct or incorrect. Thus truthful reporting is a weakly dominant strategy even when an agent knows that the prediction is inaccurate. For randomized mechanisms, the requirement concerns expected cost, and it does not require every deterministic realization to be strategyproof.

We measure performance by the ratio of expected social cost to $\mathrm{OPT}(x)$. Consistency describes the benefit of an accurate prediction, whereas robustness bounds the loss when the prediction is unreliable.

Definition 2 (Consistency and Robustness)**.** For a mechanism M, its consistency C and robustness R are its worst-case approximation ratios under correct and arbitrary predictions, respectively:

$$C = \sup_{\substack{x,\pi \text{ correct} \\ \mathrm{OPT}(x)>0}} \frac{\mathbb{E}[\mathrm{SC}(M(x;\pi);x)]}{\mathrm{OPT}(x)}, \qquad R = \sup_{\substack{x,\pi \\ \mathrm{OPT}(x)>0}} \frac{\mathbb{E}[\mathrm{SC}(M(x;\pi);x)]}{\mathrm{OPT}(x)}.$$

We additionally require zero expected cost on every zero-optimum profile under every prediction whenever a finite robustness guarantee is claimed.

The first supremum uses $\widehat{x} = x$ in the full-profile model and all optimal facility pairs in the facility-pair model. A guarantee (C, R) means upper bounds on both quantities; a mechanism with a smaller value also satisfies that guarantee. Both ratios are at least one, smaller values are better, and $C \le R$ because correct predictions are included among arbitrary predictions. In particular, consistency one means that the mechanism is optimal whenever the prediction is correct. The deterministic trade-off is defined separately for each fixed n. For randomized results, we take the suprema over all finite n as well, unless explicitly stating a fixed-population inequality. In particular, a randomized lower bound obtained as $n \to \infty$ is not asserted for every fixed population. The prediction-free mechanism TWO-EXTREMES places the facilities at the leftmost and rightmost reports and is strategyproof. With three agents, it is optimal: for sorted locations $a \le b \le c$, both its cost and the optimum equal $\min\{b-a, c-b\}$. Thus $(1,1)$ is attainable in both prediction models for $n = 3$.

3 TIGHT DETERMINISTIC CONSISTENCY-ROBUSTNESS TRADE-OFFS

Throughout this section, $n \ge 5$ is fixed. Let $R_n^*(C)$ be the infimum robustness attainable by deterministic strategyproof mechanisms with full-profile predictions and consistency at most C. Set $h = \lfloor \frac{n}{2} \rfloor$ and $L_n = 1 + \frac{\lfloor (n-4)^2/4 \rfloor}{n-4}$.

Theorem 1. *For every fixed $n \ge 5$ and $C \ge 1$, the tight full-profile trade-off is*

$$R_n^*(C) = \begin{cases} +\infty, & 1 \le C < L_n, \\ n - 2 + \max\left\{1, \frac{h}{C}, \frac{2h(h-1)}{(2h-1)C-(h-1)}\right\}, & L_n \le C < n-2,\ n = 2h, \\ n - 2 + \frac{h}{C-1/2}, & L_n \le C < h,\ n = 2h+1, \\ n - 1, & h \le C < n-2,\ n = 2h+1, \\ n - 2, & C \ge n-2. \end{cases}$$

Every stated finite upper bound is attained.

Theorem 2. *For facility-pair predictions and every fixed $n \ge 5$, the consistency of every deterministic strategyproof mechanism with finite robustness is at least $n-2$. Ignoring the prediction and using* TWO-EXTREMES *attains consistency and robustness both equal to $n-2$.*

Full-profile predictions therefore permit finite robustness exactly when consistency is at least L_n, which equals $\frac{n}{4}$ for even n. In particular, perfect consistency is possible for $n = 5$ with robustness 7, but is impossible with finite robustness for every $n \ge 6$. We prove the base mechanism's guarantees, a structural lemma for the lower bounds, and Theorem 2 below. The remaining full-profile proofs, an equivalent formulation with robustness as the budget, and bounds for $n = 4$ appear in Appendix A.

3.1 UPPER BOUND

We first define ordinary strategyproof mechanisms. The prediction is used only to choose one before the strategic reports are used. Intuitively, we fix one agent as the dictator, place a facility at her report, and choose the other facility by comparing the spans on either side. In the right-biased version, increasing a nonnegative parameter δ favors placing the second facility to the right; reflection gives the left-biased version. The second facility may extend beyond an extreme report, an essential feature of the strategyproofness argument.

Definition 3 (Biased Dictatorial Mechanism)**.** Fix an identity t and a parameter $\delta \ge 0$. On reports x, let $z = x_t$, $A = z - \min_i x_i$, $B = \max_i x_i - z$, and $\ell = |\{i : x_i < z\}|$. If $\ell = 0$, the mechanism

$D^+_{\delta,t}$ returns $\{z, \max_i x_i\}$. Otherwise, let $\tau = 1 + \frac{\delta}{\ell}$ and return

$$D^+_{\delta,t}(x) = \begin{cases} \{z, z + \max\{B, \frac{2A}{\tau}\}\}, & A \le \tau B, \\ \{z, z - \max\{A, 2\tau B\}\}, & A > \tau B. \end{cases}$$

Define $D^-_{\delta,t}$ by reflecting all reports, applying $D^+_{\delta,t}$, and reflecting its output back. For $\delta = 0$, write $D^0_t = D^+_{0,t}$.

Lemma 1. *For every fixed t and $\delta \ge 0$, $D^+_{\delta,t}$ and $D^-_{\delta,t}$ are strategyproof and have robustness exactly $n - 1 + \delta$.*

Proof. We prove the claim for $D^+_{\delta,t}$; reflection gives the result for $D^-_{\delta,t}$. Translate the dictator to 0 for notation only. The dictator and any agent whose true location is 0 have zero cost. For a positive non-dictator agent, first suppose $\ell > 0$ and let B_0 be the largest nonnegative report of the other agents. Reports that remain positive leave A, ℓ, τ unchanged. With $p = \max\{B_0, \frac{2A}{\tau}\}$, the truthful cost equals the distance to $\{0\} \cup [p, +\infty)$. Indeed, below $\frac{p}{2}$ the anchor at 0 is nearest; between $\frac{p}{2}$ and p the positive facility is p; above p the report itself receives a facility. Any positive facility exposed by another positive report is at least p. Reporting a nonpositive location cannot create a closer positive facility: A does not decrease, ℓ does not decrease, and τ does not increase. If $\ell = 0$, let B_0 again be the largest other report. The same argument uses threshold B_0, since every attainable positive facility is at least B_0.

For a negative non-dictator agent, keep the other reports fixed and let A_0 be the largest distance of their negative reports from 0, taking $A_0 = 0$ if there are none. While her report remains negative, τ and B stay fixed, and her truthful cost is the distance to $\{0\} \cup (-\infty, -p]$, where $p = \max\{A_0, 2\tau B\}$. If a nonnegative report leaves other negative reports, it decreases their count and weakly increases τ, while B cannot decrease; any new negative facility is therefore no closer. If no negative reports remain, there is no negative facility. This proves strategyproofness, including the switching ties.

For robustness, fix an optimal consecutive partition and write $T = \mathrm{OPT}(x)$. If $\ell = 0$, the mechanism is TWO-EXTREMES and has cost at most $(n-2)T$. Hence assume $\ell > 0$, so τ is defined. Let α contain the dictator and let β be the other cluster. Each non-dictator in α is within distance T of the dictator, since a cluster's span is at most its one-median cost. If β is to the right, then $A \le T$. When the second facility is the right endpoint, each agent in β is within its cluster span, hence within T. When it extends to $\frac{2A}{\tau}$, every positive report is within $\frac{A}{\tau} \le T$ of one of the two facilities. If the mechanism instead selects the left branch, then $B < \frac{A}{\tau} \le T$, and the dictator suffices. If β is to the left, $B \le T$ and the symmetric argument bounds each strictly negative agent in β by τT. There are at most ℓ such agents, so their extra charge above T is at most $\ell(\tau - 1)T = \delta T$. Thus $\mathrm{SC} \le (n - 1 + \delta)T$ in all cases, including $T = 0$. For tightness, place the dictator at 0, one agent at $-(1 + \delta)$, and the remaining $n - 2$ agents at 1. The output is $\{0, 2\}$, its cost is $n - 1 + \delta$, and the optimum is 1. □

Selection from the Prediction. We now use the prediction to select a mechanism from this family, improving consistency while preserving strategyproofness and the desired robustness guarantee. For a target consistency $L_n \le C < n - 2$, let R be the finite expression for $R^*_n(C)$ in Theorem 1, and set $\delta = R - n + 1 \ge 0$. In the following selection, the shorthand $D^0, D^+_\delta, D^-_\delta$ suppresses the chosen dictator identity. Sort the prediction while retaining identities, and compute an optimal consecutive partition with left and right cluster sizes m, s. For $m, s \ge 2$, consider six mechanisms: the classical mechanisms D^0 with dictators at the predicted upper median of the left cluster and lower median of the right cluster; D^+_δ with dictators at the second report and at the first report of the right cluster; and D^-_δ with dictators at the second-last report and at the last report of the left cluster. Evaluate these six mechanisms on the prediction, select one of minimum predicted social cost, and freeze its dictator, direction, and parameter. If the left cluster is a singleton, use D^-_δ at the upper median of the right cluster; reflect this choice for a right singleton. A zero-optimum prediction may instead select TWO-EXTREMES. For $C \ge n - 2$, use TWO-EXTREMES regardless of the prediction. The selection uses only the prediction, so Lemma 1 gives strategyproofness and robustness at most R in the nontrivial range. The comparison in Lemma 4 bounds the selected mechanism's cost by $C\,\mathrm{OPT}(x)$ on correct predictions; its detailed case analysis is in Appendix A.

3.2 LOWER BOUND

For every fixed prediction, finite robustness allows us to apply the characterization of Fotakis & Tzamos (2014): the mechanism is either TWO-EXTREMES or has a fixed dictator, meaning an identity t whose report x_t is always one of the facilities. This identity may depend on the prediction, but not on the reports. The next lemma locates the second facility in the fixed-dictator case.

Lemma 2 (Location of the Second Facility)**.** *If a deterministic strategyproof mechanism with finite approximation ratio has a fixed dictator at z, its other facility is either z, at or to the left of the minimum report, or at or to the right of the maximum report.*

Proof. Suppose the other facility is at y with $z < y < \max_i x_i$; the other case follows by reflection. Keep the dictator and a nonempty group at the rightmost report fixed, and move every remaining agent to y, one at a time. At each step the moving agent, now truly at y, can obtain the previous facility at y by reporting her previous position. Strategyproofness therefore preserves that facility, and the dictator preserves z. The resulting three nonempty groups have their middle group served. If there were no agent to move, the original two-location profile would already violate the zero-optimum requirement. Now exchange the dictator group and the middle group. The dictator group is served in the middle as well. This contradicts the three-group property of Fotakis & Tzamos (2014): for a fixed partition into three nonempty identity groups, at most one group can receive a facility when placed in the middle. □

Deviations by colocated agents further quantify how far it must extend beyond an endpoint. Three-location instances give the robustness-dependent part of the trade-off, while four-location instances force consistency at least L_n. With only a predicted facility pair, the same information can be correct for profiles with different assignments of identities to locations, which yields a complete proof of Theorem 2.

Proof of Theorem 2. Fix the prediction $\pi = \{\epsilon, 1\}$ with $0 < \epsilon < \frac{1}{n}$. The induced mechanism is either TWO-EXTREMES or has a fixed dictator. In the first case use $\big(0, \epsilon^{n-2}, 1\big)$. In the second case use the same locations but assign the dictator identity to 0. The prediction is optimal and $\mathrm{OPT} = \epsilon$. TWO-EXTREMES has cost $(n-2)\epsilon$. For a dictator at 0, let y be the other facility. If $y < 2\epsilon$, the rightmost agent costs at least $1 - 2\epsilon \geq (n-2)\epsilon$. If $y \geq 2\epsilon$, each middle agent costs at least ϵ. This proves the consistency lower bound; TWO-EXTREMES attains it with robustness $n-2$. □

4 RANDOMIZED MECHANISMS

We now seek guarantees uniform over all population sizes.

4.1 FULL-PROFILE PREDICTIONS

Theorem 3. *For every $0 < \varepsilon \leq \frac{1}{100}$, there is a randomized strategyproof mechanism with full-profile predictions satisfying*

$$C \leq 1 + \varepsilon, \qquad R \leq \frac{2}{3\varepsilon^2} + \frac{18}{5\varepsilon} + 3.$$

In addition, put $\rho = \frac{74+4\sqrt{3}}{23}$. For every target consistency $1 < c \leq \rho$, there is such a mechanism with consistency at most c and robustness at most

$$\mathcal{B}(c) = \begin{cases} 1 + \frac{7}{2(c-1)} + \frac{1}{(c-1)^2}, & 1 < c \leq \frac{5}{3}, \\ \frac{57}{2} - 12c, & \frac{5}{3} \leq c \leq \frac{7}{4}, \\ 18 - 6c, & \frac{7}{4} \leq c \leq 2, \\ 2 + \frac{c^2}{(c-1)^2}, & 2 \leq c \leq \frac{5}{2}, \\ \frac{(\rho-c)(43/9)+(c-5/2)\rho}{\rho-5/2}, & \frac{5}{2} \leq c \leq \rho. \end{cases}$$

In particular, $\mathcal{B}(\rho) = \rho$; for $c \geq \rho$, robustness ρ is attainable by ignoring the prediction. Each guarantee holds for every finite population size.

We compute two optimal clusters of the predicted profile, retaining agent identities and ranks. Label these clusters by $g \in \{0, 1\}$. The prediction fixes anchor probabilities p_i, nominee probabilities q_j within each group, and nonnegative coefficients h_j, L_g. After choosing anchor i from group g and nominee j from the opposite group, we sample a partner ℓ from all agents with probability proportional to

$$\left(1 + \frac{L_g h_j}{q_j} \mathbf{1}_{\{\ell=j\}}\right) |x_i - x_\ell|,$$

and return $\{x_i, x_\ell\}$. Here $\mathbf{1}_{\{\ell=j\}}$ equals one if $\ell = j$ and zero otherwise. Only nominees with $q_j > 0$ are sampled; if all reports coincide, both facilities are placed there. The complete distributions and parameter choices are specified in Appendix B.1.

Lemma 3 (Weighted Proportional Sampling)**.** *Place one facility at a fixed anchor and sample a partner proportionally to its distance from the anchor times a fixed positive identity coefficient, placing the second facility at its report. If all reports equal the anchor, place both facilities there. This rule is strategyproof for every non-anchor agent, even with different fixed positive coefficients on the two sides of the anchor.*

Proof. Fix the anchor location a, a non-anchor agent's true location x, and an alternative report y. If $x = a$, her truthful cost is zero. Otherwise, put $r_0 = |x - a| > 0$ and $e_0 = \min\{r_0, |x - y|\}$. Let $w > 0, w' \geq 0$ be her sampling weights under truthful and deviating reports, respectively. Let T be the other partners' total sampling weight and A the sum of their weights times her distances to the corresponding facility pairs. These quantities are fixed when she deviates, and $0 \leq A \leq r_0 T$. If $T + w' = 0$, all reports are at the anchor and her cost is r_0, so she cannot improve. Otherwise, her truthful and deviating costs are $A/(T + w)$ and $(A + w' e_0)/(T + w')$. If $w' \leq w$, the deviation cannot improve her cost. If $w' > w$, cross-multiplication and $A \leq r_0(T + w)$ show that it suffices to prove $w' e_0 \geq r_0(w' - w)$. This is immediate when $e_0 = r_0$, including any deviation across the anchor. Otherwise both reports are on the same side, with the same coefficient $c > 0$, so $w = c r_0$ and $w' = c|y - a|$. The triangle inequality gives $e_0 = |x - y| \geq |y - a| - r_0$, while $w' > w$ gives $|y - a| > r_0$; hence

$$w' e_0 \geq c|y - a|(|y - a| - r_0) \geq r_0(w' - w).$$

Thus no report y lowers the agent's expected cost. Since the non-anchor agent was arbitrary, the rule is strategyproof for every non-anchor agent. □

Conditional on the anchor and nominee, each partner has a fixed positive sampling coefficient. Lemma 3 gives strategyproofness for every non-anchor, and the anchor's truthful cost is zero. Averaging over prediction-fixed anchor and nominee distributions therefore preserves strategyproofness. For consistency, we analyze gaps between consecutive predicted ranks and the resulting dispersion around each cluster median. For robustness, we compare the sampled locations with the actual optimal clusters, regardless of the predicted partition. Uniform central-rank sampling and fixed mixtures yield $\mathcal{B}(c)$; nonuniform nominee distributions sharpen the bound near $c = 1$.

4.2 Facility-Pair Predictions

With only a predicted facility pair, we use its midpoint to choose a preferred sampling direction. Intuitively, we first place a facility at a uniformly sampled agent's report, called the anchor. If the anchor is at or to the left of the predicted midpoint, we favor partners to its right; otherwise, we favor partners to its left, encouraging the two facilities to cover different regions. Partners are sampled proportionally to their distance from the anchor, with weights on the nonfavored side multiplied by a parameter $\tau \in (0, 1]$. Smaller τ strengthens the directional preference, while $\tau = 1$ recovers ordinary proportional sampling.

Definition 4 (Directional Proportional Sampling)**.** Let the predicted pair be $\{\widehat{a}, \widehat{b}\}$ with $\widehat{a} \leq \widehat{b}$, and put $z = \frac{\widehat{a}+\widehat{b}}{2}$. Fix $0 < \tau \leq 1$. The mechanism Directional-Proportional$_\tau$ chooses an anchor identity i uniformly. If $x_i \leq z$, it favors reports to the right of x_i; otherwise it favors reports to the left. It selects partner j with probability proportional to $\omega_{ij}|x_i - x_j|$, where $\omega_{ij} = 1$ on the favored side and $\omega_{ij} = \tau$ on the other side, and returns $\{x_i, x_j\}$. If all reports coincide, it returns the common location twice.

The coefficient depends on direction relative to the anchor, not on whether the partner lies beyond the predicted midpoint. Conditional on the anchor's report, both directional coefficients are fixed, so the weighted proportional-sampling argument again establishes strategyproofness.

Theorem 4. *For every $0 < \tau \leq 1$,* DIRECTIONAL-PROPORTIONAL$_\tau$ *has consistency $3 + \tau$ and robustness $3 + \frac{1}{\tau}$. Both bounds are tight for this family of mechanisms as the population size varies.*

Setting $\tau = c - 3$ gives robustness $3 + \frac{1}{c-3}$ at consistency c, for $3 < c \leq 4$. The proof compares sampling costs across the two actual optimal clusters; limiting three-location instances show tightness for this family. See Appendix B.2 for the full argument.

4.3 A LOWER BOUND UNDER EITHER PREDICTION TYPE

We adapt the two-profile comparison of Ma & Peng (2026), fixing an accurate prediction on one of the profiles.

Theorem 5. *Under either prediction model, population-independent consistency and robustness guarantees satisfy*

$$C \geq 1 + \frac{(\sqrt{R} - \sqrt{R-1})^2}{2}.$$

In particular, consistency at most $1 + \varepsilon$, $0 < \varepsilon \leq \frac{1}{2}$, requires $R \geq \frac{(1+2\varepsilon)^2}{8\varepsilon}$. No mechanism with uniformly finite robustness has perfect consistency.

Proof sketch. Consider $A = \left(0^m, \left(\frac{1}{2}\right)^k, 1^m\right)$ and $B = \left(-m, 0^m, \left(\frac{1}{2}\right)^{k-1}, 1^m\right)$ for integers $2 \leq k < m$. Their optimal costs are $k/2$ and m, respectively. Fix the prediction to B, or to its optimal facility pair $\{-m, \frac{1}{2}\}$, and keep it unchanged at A. Only one midpoint agent changes her report. Let a, b be the expected distances from her true location $\frac{1}{2}$ to the outputs at A, B, respectively; strategyproofness gives $a \leq b$.

Write S_A, S_B for the expected social costs and put $t = \frac{k-1}{m-k}$. The two-median costs of $\{0, \frac{1}{2}, 1\}$ and of the weighted locations $(-m, 0^m, 1^m)$ imply $tS_A + (k-1)a \geq tm/2$ and $S_B - (k-1)b \geq m$, using $tk + k - 1 = tm$. Adding and using $a \leq b$ gives $tS_A + S_B \geq m(1 + t/2)$. Robustness gives $S_A \leq Rk/2$, while consistency at B gives $S_B \leq Cm$. Hence

$$C \geq 1 + \frac{k-1}{2(m-k)} - \frac{Rk(k-1)}{2m(m-k)}.$$

Letting $\frac{k}{m} \to \alpha \in (0, 1)$ yields $C \geq 1 + \frac{\alpha(1-R\alpha)}{2(1-\alpha)}$; maximizing over α proves the bound. The full finite-population argument is in Appendix B.3. □

The bounds leave a linear-versus-quadratic gap near perfect consistency with full-profile predictions. The limiting consistency 3 of our facility-pair construction is not a lower bound for general mechanisms.

5 CONCLUSION

We studied strategyproof two-facility location with predictions of either an optimal facility pair or the complete labeled profile. For every fixed $n \geq 5$, we established tight deterministic consistency-robustness trade-offs: full-profile predictions improve the attainable consistency, whereas facility-pair predictions do not improve the optimal prediction-free guarantee. For randomized mechanisms, we gave population-independent upper bounds, including near-perfect consistency with quadratic robustness in the inverse consistency error, an explicit analytic trade-off extending to the prediction-free endpoint, and a general lower bound with linear dependence. Determining the optimal randomized trade-offs remains open, as does establishing a tight deterministic trade-off for four agents. It would also be interesting to study other objectives and higher-dimensional domains, where both the structure of optimal clusters and the incentives to change cluster membership differ from the line.

AI USE STATEMENT

GPT Astra-6 played a crucial role in developing the results for full-profile predictions, including the mechanism designs and the derivation of upper and lower bounds. The authors also used Codex to extensively revise and polish the manuscript. The authors take responsibility for the correctness and content of the final manuscript, including all AI-assisted contributions.

REPRODUCIBILITY STATEMENT

The model, prediction types, and performance measures are specified in Section 2. Proofs are provided in the main text and Appendices A and B, which also specify all sampling distributions and parameter choices; the stated guarantees are established analytically.

REFERENCES

Priyank Agrawal, Eric Balkanski, Vasilis Gkatzelis, Tingting Ou, and Xizhi Tan. Learning-augmented mechanism design: leveraging predictions for facility location. In *Proceedings of the 23rd ACM Conference on Economics and Computation*, pp. 497–528, 2022.

Haris Aziz, Hau Chan, Barton Lee, Bo Li, and Toby Walsh. Facility location problem with capacity constraints: algorithmic and mechanism design perspectives. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, volume 34, pp. 1806–1813, 2020a.

Haris Aziz, Hau Chan, Barton E Lee, and David C Parkes. The capacity constrained facility location problem. *Games and Economic Behavior*, 124:478–490, 2020b.

Haris Aziz, Yuhang Guo, Alexander Lam, and Houyu Zhou. Learning-augmented facility location mechanisms for envy ratio. In *Advances in Neural Information Processing Systems 38: Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2025.

Eric Balkanski, Vasilis Gkatzelis, and Golnoosh Shahkarami. Randomized strategic facility location with predictions. *Advances in Neural Information Processing Systems*, 37:35639–35664, 2024.

Zohar Barak, Anupam Gupta, and Inbal Talgam-Cohen. MAC advice for facility location mechanism design. In *Advances in Neural Information Processing Systems 37: Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2024.

Duncan Black. On the rationale of group decision-making. *Journal of Political Economy*, 56(1): 23–34, 1948.

Hau Chan, Aris Filos-Ratsikas, Bo Li, Minming Li, and Chenhao Wang. Mechanism design for facility location problems: a survey. In *Proceedings of the 30th International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 4356–4365, 2021.

Hau Chan, Jianan Lin, and Chenhao Wang. Randomized strategyproof facility location: Two facilities and beyond. *arXiv preprint arXiv:2608.22484*, 2026.

George Christodoulou, Alkmini Sgouritsa, and Ioannis Vlachos. Mechanism design augmented with output advice. In *Advances in Neural Information Processing Systems 37: Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2024.

Zvi Drezner and Horst W Hamacher (eds.). *Facility location: applications and theory*. Springer Berlin Heidelberg, 2002.

Jiazhu Fang, Qizhi Fang, Wenjing Liu, and Qingqin Nong. Mechanism design with predictions for facility location games with candidate locations. In *Theory and Applications of Models of Computation - 18th Annual Conference, (TAMC)*, volume 14637, pp. 38–49, 2024.

Itai Feigenbaum, Jay Sethuraman, and Chun Ye. Approximately optimal mechanisms for strategyproof facility location: minimizing L_p norm of costs. *Mathematics of Operations Research*, 42 (2):434–447, 2017.

Michal Feldman and Yoav Wilf. Strategyproof facility location and the least squares objective. In *Proceedings of the 14th ACM Conference on Electronic Commerce (EC)*, pp. 873–890, 2013.

Dimitris Fotakis and Christos Tzamos. Strategyproof facility location for concave cost functions. In *Proceedings of the 14th ACM Conference on Electronic Commerce (EC)*, pp. 435–452, 2013.

Dimitris Fotakis and Christos Tzamos. On the power of deterministic mechanisms for facility location games. *ACM Transactions on Economics and Computation (TEAC)*, 2(4):15:1–15:37, 2014.

Jianhao Jia. Product gap mechanisms for multi-facility location. *arXiv preprint arXiv:2608.19633*, 2026.

Pinyan Lu, Yajun Wang, and Yuan Zhou. Tighter bounds for facility games. In *Proceedings of the 5th International Workshop on Internet and Network Economics (WINE)*, pp. 137–148. Springer, 2009.

Pinyan Lu, Xiaorui Sun, Yajun Wang, and Zeyuan Allen Zhu. Asymptotically optimal strategy-proof mechanisms for two-facility games. In *Proceedings of the 11th ACM Conference on Electronic Commerce (EC)*, pp. 315–324, 2010.

Pinyan Lu, Zihan Luo, and Jialin Zhang. Design and characterization of strategy-proof mechanisms for two-facility game on a line. In *International Computing and Combinatorics Conference (COCOON 2024)*, pp. 164–177. Springer, 2024.

Mengfan Ma and Bo Peng. Breaking the 4-approximation barrier in strategyproof two-facility location. *arXiv preprint arXiv:2608.09061*, 2026.

Hervé Moulin. On strategy-proofness and single peakedness. *Public Choice*, 35(4):437–455, 1980.

Ariel D Procaccia and Moshe Tennenholtz. Approximate mechanism design without money. *ACM Transactions on Economics and Computation (TEAC)*, 1(4):18:1–18:26, 2013.

Paolo Serafino and Carmine Ventre. Heterogeneous facility location without money. *Theoretical Computer Science*, 636:27–46, 2016.

Toby Walsh. Mechanism design for facility location using predictions. *arXiv preprint arXiv:2508.03818*, 2025.

Chenyang Xu and Pinyan Lu. Mechanism design with predictions. In *Proceedings of the 31st International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 571–577, 2022.

A Deterministic Mechanisms: Additional Proofs and Details

Throughout this section, $n \geq 5$ is fixed, except in the final proposition. Let $C_n^*(R)$ denote the infimum consistency among deterministic strategyproof mechanisms with full-profile predictions and robustness at most the finite budget R. Set $h = \lfloor \frac{n}{2} \rfloor$ and $L_n = 1 + \frac{\lfloor (n-4)^2/4 \rfloor}{n-4}$. For $q \geq 1$, define

$$\Phi_{2h+1}(q) = \max\left\{L_{2h+1}, \min\left\{h, \frac{1}{2} + \frac{h}{q}\right\}\right\},$$
$$\Phi_{2h}(q) = \max\left\{L_{2h}, \frac{h}{q}, \frac{(h-1)(2h+q)}{(2h-1)q}\right\}.$$

The equivalent consistency-at-robustness formulation of Theorem 1 is

$$C_n^*(R) = \begin{cases} n-2, & n-2 \leq R < n-1, \\ \max\left\{L_n, \min\left\{h, \frac{1}{2} + \frac{h}{R-n+2}\right\}\right\}, & R \geq n-1,\ n = 2h+1, \\ \max\left\{L_n, \frac{h}{R-n+2}, \frac{(h-1)(R+2)}{(n-1)(R-n+2)}\right\}, & R \geq n-1,\ n = 2h. \end{cases}$$

Every stated finite upper bound is attained.

Theorem 2, proved in Section 3, gives the corresponding facility-pair impossibility.

The complete-profile trade-off reaches L_n at the finite budget $R_n^\dagger = n + \frac{2(n-4)}{n^2-3n+4}$ for even n, and $R_n^\dagger = n + \frac{2(n-3)}{n^2-6n+7}$ for odd n. In particular, perfect consistency is possible for $n = 5$ with robustness 7, but is impossible with finite robustness for every $n \geq 6$. The information separation is therefore not simply a distinction between exact and approximate optimization: full predictions improve the guarantee even when strategyproofness still prevents perfect consistency.

A.1 Upper Bound

We use the biased dictatorial mechanisms of Definition 3; Lemma 1 in the main text proves their strategyproofness and exact approximation ratios. Recall that ℓ counts only reports strictly below the dictator; agents at z are not included.

Given a robustness budget $R \geq n-1$, set $q = \min\{R-n+2, R_n^\dagger - n + 2\}$ and $\delta = q - 1$. Use the candidate selection in Section 3 with this value of δ; m, s denote the predicted optimal cluster sizes. For $n-2 \leq R < n-1$, always use Two-Extremes.

Lemma 4. *On a correct prediction, the candidate selection above has cost at most* $\Phi_n(q)\,\mathrm{OPT}$ *for every* $1 \leq q \leq R_n^\dagger - n + 2$.

Proof. Write $T = \mathrm{OPT}$, $q = 1 + \delta$, and $q_n^\dagger = R_n^\dagger - n + 2$. It suffices to consider $1 \leq q \leq q_n^\dagger$, and direct substitution in Φ_n gives $q_n^\dagger \leq 2 + \frac{2}{n-4}$.

Step 1: Candidate Cost Bounds.

A useful comparison is to keep the dictator z and pretend that D_δ^+ always places its second facility to the right, at $z + \max\{B, \frac{2A}{\tau}\}$. This only increases the cost: if the mechanism actually selects the left branch, the hypothetical right facility is beyond $z + 2B$ and is no better than z for any report. The reflected statement holds for D_δ^-. If the strict side count is smaller than a rank-based upper bound, τ is larger and this hypothetical facility is closer to the endpoint, so its cost cannot increase. We will therefore use rank-based counts even when reports coincide.

Let the optimal cluster medians be $0, d$ with $d \geq 0$, cluster sizes m, s, and cluster costs P, Q. Write the endpoints as $-A, d+B$. Let S_a, S_b be the costs of the two classical median-dictator candidates. Charging one cluster to its median and using either the median or the exterior facility for the other gives

$$S_a - T \leq \min\{sd, s\max\{B, 2A-d\} - 2B\}, \quad S_b - T \leq \min\{md, m\max\{A, 2B-d\} - 2A\}. \tag{A.1}$$

For example, $\sum_{i\in R}(x_i - d) \geq 2B - Q$ yields the first exterior-facility bound after subtracting Q.

These two candidates already give consistency at most $h = \lfloor \frac{n}{2} \rfloor$ when both clusters have size at least two. To see this, normalize $T = 1$ and reflect so that $A \geq B$, whence $A+B \leq 1$. If $d \geq 2A-B$, the smaller excess is at most $\min\{(s-2)B, (m-2)A\} \leq h-2$. If $d \leq 2B-A$, then $2d \leq A+B \leq 1$ and the smaller excess is at most $\min\{m, s\}d \leq \frac{h}{2} \leq h-1$. In the remaining interval, the two excess bounds are $(m-2)A$ and $\min\{sd, 2sA - sd - 2B\} \leq sA - B$. Their minimum is at most $\min\{m-2, s\}A \leq h-1$. Singleton clusters are handled below and give a bound of $2 \leq h$. This proves the initial flat segment for odd n; we may henceforth restrict odd n to $q \geq \frac{2h}{2h-1}$.

For the full comparison, suppose $m, s \geq 2$ and write the second and second-last reports as $-u, d+v$. Thus $A \geq u \geq 0$, $B \geq v \geq 0$, and $S = A + B + u + v \leq T$, since four distinct identities contribute these distances to their cluster medians. The chosen upper and lower medians ensure that $m \leq 3$ implies $u = 0$ and $s \leq 3$ implies $v = 0$. If $S = 0$, a classical candidate uses both endpoints and is optimal. Otherwise normalize $S = 1$, reflect so that $B \geq A$, and set $\gamma = \Phi_n(q)$ and $k = \gamma - 1$. It suffices to find a candidate with excess at most $kS = k$; we have not normalized T to one.

The classical bounds (A.1) imply

$$S_b - T \leq mB - A, \qquad S_a - T \leq \max\{(s-2)B, sA - B\}. \tag{A.2}$$

For the second-report candidates, define

$$K_L = \max\left\{B, \frac{2A}{q} - \left(1 + \frac{2}{q}\right)u - d\right\}, \qquad K_R = \max\left\{A, \frac{2B}{q} - \left(1 + \frac{2}{q}\right)v - d\right\}. \tag{A.3}$$

Their conservative facility pairs are $\{-u, d+K_L\}$ and $\{-K_R, d+v\}$. Moving the left median from 0 to $-u$ raises its cluster cost by at most $(m-4)u$: compared with the Lipschitz bound mu, the two leftmost agents each save $2u$. The right counterpart is $(s-4)v$. Consequently, if $K_L = B$ and if $K_R > A$, respectively, the candidate excesses satisfy

$$\begin{aligned} &S_{2L} - T \leq (s-1)d + (n-5)u - B, \\ &S_{2R} - T \leq -md + \frac{2m}{q}B + Zv - 2(A+u), \qquad Z = s - m - 4 - \frac{2m}{q}. \end{aligned} \tag{A.4}$$

For the first inequality, send the right endpoint to itself and all other agents to $-u$. For the second, send the whole left cluster to $-K_R$ and use $\sum_{i\in L} x_i \leq P - 2A - 2u$. If both $K_L = B$ and $K_R = A$, sending each cluster to the corresponding exterior facility instead gives

$$E_L = (s-2)B + (m-4)u - 2v, \qquad E_R = (m-2)A + (s-4)v - 2u. \tag{A.5}$$

The remaining useful candidate is the first report of the right cluster, at $d - w$, where $0 \leq w \leq d$. Its strict left count is at most m. Let $c = \frac{2m}{m+q-1} \leq 2$. The conservative right facility is $d - w + \max\{B+w, c(d+A-w)\}$. Charging the left cluster to $d-w$, and the right cluster to that point except for its endpoint, bounds the excess by $md + (s-m-4)w + \max\{B+w, c(d+A-w)\} - 2B$. If $s \leq m+3$, both branches are nonincreasing in w, giving

$$E_{\text{out}} \leq md + \max\{B, c(d+A)\} - 2B. \tag{A.6}$$

The right-median classical candidate satisfies the same bound with $c = 2$ and $w = 0$, without a restriction on $s - m$.

Step 2: The Left Candidate Extends Beyond the Endpoint.

If $K_L > B$, (A.3) and $B \geq A$ give $d < \left(\frac{2}{q} - 1\right)A - \left(1 + \frac{2}{q}\right)u$. This is impossible for $q \geq 2$. Otherwise $A \leq \frac{S}{2}$, and the classical candidates have smaller excess at most $hd < \frac{h}{2}\left(\frac{2}{q} - 1\right)S$. The resulting factor $1 + \frac{h}{q} - \frac{h}{2}$ is at most $\frac{h}{q}$ for even n, and at most $\frac{1}{2} + \frac{h}{q}$ for the remaining odd-n range. We may therefore assume $K_L = B$.

Step 3: Both Candidates Reach Their Endpoints.

Suppose also $K_R = A$. Put $L_0 = n - 4$ and $\lambda = L_n - 1 = \frac{\lfloor L_0^2/4 \rfloor}{L_0}$. We compare linear combinations of (A.2) and (A.5); the minimum candidate excess is at most every convex combination

of its available bounds. If $m \le s-2$, combine $mB - A$ and E_R with weights $\frac{s-4}{L_0}$ and $\frac{m}{L_0}$. The B, v coefficients equal $p = \frac{m(s-4)}{L_0} \le \lambda$, the u coefficient is nonpositive, and the A coefficient is at most p because $m(m-s+2) - s + 4 \le 0$. If $|m-s| \le 1$, combine E_L, E_R with weights $\frac{m-2}{L_0}, \frac{s-2}{L_0}$. The A, B coefficients equal $p = \frac{(m-2)(s-2)}{L_0} \le \lambda$. The u coefficient minus p has numerator $(m-2)(m-s) - 2L_0 \le 0$, and the v calculation is symmetric. Finally, if $m \ge s+2$ and the maximum in (A.2) is $(s-2)B$, combine it with E_R using weights $\frac{m-2}{L_0}, \frac{s-2}{L_0}$; all coefficients are at most the same p. If that maximum is $sA - B$, then $sA - B \le \frac{s-1}{2}(A+B) \le \lambda S$. For the last inequality, even n gives $s \le h-1$ and $\lambda = \frac{h-2}{2}$; odd n has $m - s \ge 3$ and gives the same inequality directly. Thus this case has factor at most L_n.

Step 4: Only the Right Candidate Extends Beyond the Endpoint.

It remains that $K_L = B$ and $K_R > A$. Set $W = A + u$, so $W + B + v = 1$ and $u \le \frac{W}{2}$. Combine the second bound in (A.4) first with the classical bound sd, and then with the first bound in (A.4), eliminating d. The minimum excess is bounded by each of

$$F_1 = \alpha B + \beta v - c_0 W, \qquad \alpha = \frac{2ms}{nq}, \qquad \beta = \frac{sZ}{n}, \qquad c_0 = \frac{2s}{n},$$
$$F_2 = K_m B + V_m v + D_m W, \quad K_m = \frac{m[2(s-1)/q - 1]}{n-1}, \quad V_m = \frac{(s-1)Z}{n-1}, \quad D_m = \frac{m-4}{2}. \tag{A.7}$$

The second line uses $u \le \frac{W}{2}$ and $\frac{m(n-5)-4(s-1)}{2(n-1)} = \frac{m-4}{2}$. If $m \le k$, (A.2) already suffices, so assume $m > k$.

First, the v coefficients in (A.7) are at most k, or $v = 0$. If $Z < 0$ this is immediate; otherwise $\beta \ge V_m$. In the region $Z \ge 0$, $\beta(m, q)$ decreases with m and increases with q. Put $k_0 = \frac{n-4}{4} - \frac{1}{4(n-4)}$ and $\bar{q} = 2 + \frac{2}{n-4}$. For $n \ge 12$, $m > k \ge k_0$, $q \le \bar{q}$, and

$$k_0 - \beta(k_0, \bar{q}) = \frac{n^4 - 19n^3 + 105n^2 - 233n + 210}{16n(n-4)^2} \ge 0.$$

The numerator, after $n = 12 + x$, is $x^4 + 29x^3 + 285x^2 + 991x + 438$. For $6 \le n \le 9$, $\beta(2, \bar{q}) \le 0$; for $n = 10$ it is at most $\frac{8}{35} < \frac{3}{2} \le k$; for $n = 11$ it is at most $\frac{45}{44} < \frac{12}{7} \le k$. For $n = 5$ both u and v vanish. We can therefore omit v, prove that one of $f_1 = \alpha B - c_0 W$ and $f_2 = K_m B + D_m W$ is at most $k(W + B)$, and add back v afterward.

If $K_m \le k$ and $D_m \le k$, use f_2. If $K_m \le k < D_m$, set $m_0 = 2k + 4$ and normalize $W + B = 1$. The maximum of the smaller of these two lines is at most $\max\{K_m, G\}$, where their intersection value is

$$G = \frac{\alpha D_m + c_0 K_m}{\alpha - K_m + D_m + c_0} = \frac{2s[m(n-5) - 2q]}{q(n^2 - 3n + 4) + 4m}.$$

For $n = 5, 6$, $D_m \le 0$ and this case does not arise. For $n \ge 7$, $G \le k$ is equivalent to $F(m) \ge 0$, where

$$F(m) = k[q(n^2 - 3n + 4) + 4m] - 2(n-m)[m(n-5) - 2q].$$

This upward quadratic has vertex $m_* = \frac{n}{2} + \frac{q-k}{n-5}$, and $m_0 - m_* \ge \frac{(3n-19)(3n-11)}{4(n-5)(n-4)} > 0$. Also $m_0 \ge h + 2$; K_m decreases for real $m \ge h$, and $K_h \le k$ directly from Φ_n. At m_0, $D_{m_0} = k$ and $K_{m_0} \le k$, so $F(m_0) \ge 0$. Since F increases beyond its vertex, $F(m) \ge 0$ for every relevant $m > m_0$. If $m_0 \ge n - 2$, there is no such cluster size.

Step 5: Remaining Cluster Sizes and Boundary Cases.

Only $K_m > k$ remains. Write $m = h - j$, since $K_h \le k$ and K_m decreases for $m \ge h$. For even n, $K_{h-j} - K_h = \frac{j(q+2-2j)}{(2h-1)q}$, so $q \le \bar{q} \le 3$ leaves only $j = 1, 2$. For odd n in its nonflat range, the difference is $\frac{j(q-2j)}{2hq}$, so $q \le \bar{q} \le 4$ leaves only $j = 1$. The $n = 5$ case $h - 1 = 1$ is not a nonsingleton cluster. In each remaining case $Z \le 0$.

Use (A.6) with $c = \frac{2m}{m+q-1}$ for $m = h-1$, since then $s \le m+3$. For even n and $m = h-2$, use its classical version $c = 2$. If the maximum there is B, averaging with the second bound of (A.4) gives excess at most $\left(\frac{m}{q} - \frac{1}{2}\right)B + \frac{Z}{2}v - (A+u)$, hence factor at most $\max\{1, \frac{1}{2} + \frac{m}{q}\}$. If it is $c(d+A)$, use weights $\frac{m+c}{2m+c}$ and $\frac{m}{2m+c}$ to cancel d. The A, u, v coefficients are nonpositive, since the A numerator is $mc - 2m - 2c \le 0$ and $Z \le 0$. The remaining factor is at most $\max\{1, 1+\kappa_c\}$, where

$$1 + \kappa_c = \begin{cases} \frac{m+1}{q}, & c = \frac{2m}{m+q-1}, \\ \frac{m+1}{q} + \frac{1-1/q}{m+1}, & c = 2. \end{cases}$$

For odd $m = h-1$, both factors are at most $\frac{1}{2} + \frac{h}{q}$. For even $m = h-1$, one factor is $\frac{h}{q}$; the other is at most this for $q \le \frac{h}{h-1}$, and afterward is bounded by the third term of Φ_n, because

$$\frac{(h-1)(2h+q)}{(2h-1)q} - \left(\frac{1}{2} + \frac{h-1}{q}\right) = \frac{2h-2-q}{2(2h-1)q} \ge 0.$$

For even $m = h-2 \ge 2$, $h \ge 4$ and $q \le \bar{q} \le \frac{5}{2} \le h$. Both $\frac{1}{2} + \frac{h-2}{q}$ and $\frac{h-1}{q} + \frac{1-1/q}{h-1}$ are then at most $\frac{h}{q}$. This finishes every nonsingleton cluster case.

If the left cluster is a singleton, let z be the upper median of the right cluster and write the endpoints as $z - A, z + B$. Optimality implies $A \ge B$: replacing the isolated left facility by the right endpoint saves B on the rightmost agent, costs at most A on the isolated agent, and does not harm any other agent. Let r be the strict number of reports above z. If $r = 0$, $B = 0$ and the chosen mechanism is optimal. Otherwise its left facility is $z - \max\{A, cB\}$, where $c = \frac{2r}{r+q-1} \le 2$. Its extra cost above T is at most $(cB - A)_+ \le (c-1)_+ B$, so its factor is at most $\max\{1, c\} \le 2$. For $n \ge 8$, $L_n \ge 2$ suffices. For $n = 6, 7$, $r \le 2$ gives $c \le \frac{4}{q+1} \le \frac{3}{q}$ on the required range $q < 3$; the odd flat segment instead uses the factor 2. For $n = 5$, $r \le 1$ gives $c \le \frac{2}{q} \le \Phi_5(q)$ after taking the maximum with one. Reflect for a right singleton. Tied positions only reduce strict counts, and zero optimum was covered by robustness of the base mechanisms. All comparisons allow equality at the switching boundaries. Thus the candidate comparison holds for every real profile and every required parameter. □

Proposition 1. *This prediction-dependent selection is strategyproof and attains the upper bound in Theorem 1.*

Proof. For each fixed prediction, the selected mechanism is fixed, so Lemma 1 gives strategyproofness and robustness at most $n - 2 + q \le R$. It remains to bound the cost of the selected candidate on a correct prediction. Lemma 4 gives the required bound on every correct prediction. Truncating q is valid because Φ_n has reached its minimum L_n at $R_n^\dagger$. □

A.2 Lower Bound

Fix a prediction throughout the following arguments. The characterization of Fotakis & Tzamos (2014) then applies to the resulting ordinary mechanism: it is either Two-Extremes or has a fixed dictator. Lemma 2, proved in the main text, locates the second facility without an additional invariance assumption. We next quantify the extension beyond an endpoint.

Lemma 5 (Colocated Deviations)**.** *A group of agents at the same true location cannot reduce its common distance by changing its reports jointly. Consequently, if the group is restricted to a common report, its truthful cost is the distance to the set of facility positions attainable by such reports.*

Proof. Change the reports one agent at a time. The agent about to move still reports the common true position, so strategyproofness says that the distance from that position cannot decrease at this step. Adding these comparisons proves the first claim. For the second, if a strictly nearer attainable facility existed, the corresponding common report would improve the group's common cost. The truthful output attains the resulting infimum. □

Lemma 6 (Quantitative Extension Beyond an Endpoint)**.** *Every fixed-dictator mechanism has approximation ratio at least $n-1$. If its ratio is at most $n-1+\delta$, the dictator is at 0, and k other agents are at $-A<0$, then any positive second facility y satisfies*

$$y \geq \frac{2kA}{k+\delta}, \qquad 1 \leq k \leq n-2. \tag{A.8}$$

Proof. First put the remaining $n-1-k$ agents at $b>0$. Suppose the output has no negative facility when the group of k agents is at $-A$. Let p be the infimum absolute value of negative facilities attainable by a common report of this group. The set is nonempty: sending the group arbitrarily far left leaves an optimal cost at most b, and finite approximation must eventually place a facility to the left. By Lemma 5, $p \geq 2A$. For every $a \in \left(\frac{p}{2}, p\right)$, the infimum attainable distance from $-a$ is $p-a<a$. The truthful output must attain this value, so it contains $-p$ as well as the dictator at 0. In particular the boundary p is attained; no continuity assumption has been used. The cost is $k(p-a)+(n-1-k)b$, whereas $\mathrm{OPT} \leq \min\{a,b\}$. Letting $a \downarrow \frac{p}{2}$ gives

$$R \geq \frac{kp/2+(n-1-k)b}{\min\{p/2,b\}} \geq n-1, \qquad \frac{p}{2} \leq \left(1+\frac{\delta}{k}\right)b \quad \text{if } R \leq n-1+\delta.$$

The first comparison applies to every fixed-dictator mechanism: choose $A>0$ sufficiently small that robustness and $\mathrm{OPT} \leq A$ force a positive second facility.

For a general profile in the lemma, Lemma 2 gives $y \geq \max_i x_i$. Fix any $b \in \left(\frac{y}{2}, y\right)$ and move every agent other than the dictator and the negative group to b, one at a time. If the current positive facility is $y_j \in [b,y]$, the newly moved agent can obtain cost $y_j - b < b$ by her old report. The next output must therefore still have a positive facility; Lemma 2 places it at least at b, and strategyproofness places it at most at y_j. At the final three-location profile, the preceding bound gives $A \leq \left(1+\frac{\delta}{k}\right)b$. Let $b \downarrow \frac{y}{2}$ to obtain (A.8). □

Proposition 2. *Every deterministic strategyproof mechanism with full-profile predictions satisfies the lower bound in Theorem 1.*

Proof. The characterization and Lemma 6 show that budgets below $n-2$ are impossible, and budgets in $[n-2, n-1)$ permit only TWO-EXTREMES for each prediction. Now let $R = n-2+q$, $q \geq 1$. Lemma 2 implies that every output costs at least the best solution constrained to include a report endpoint: project an exterior facility to that endpoint, or add an endpoint to a coincident pair. For $n \geq 6$, consider $\left(-A, 0^M, D^S, D+B\right)$, where $M+S=n-2$, $M,S \geq 2$, $A=S-1$, $B=M-1$, and D is sufficiently large. The optimum is $A+B$; the best solution including an endpoint costs $\min\{MA+B, A+SB\}$. Taking $M-1, S-1$ as balanced as possible gives ratio L_n. For $n=5$, $L_n=1$ is immediate.

For the budget-dependent terms, use a correct prediction $\left(-a, 0^m, 1^s\right)$ with $m+s=n-1$ and $1 \leq a \leq \min\{m,s\}$, so $\mathrm{OPT}=a$. If the frozen mechanism is TWO-EXTREMES or its dictator is at an endpoint, its cost is at least m. Otherwise its dictator is at 0. Applying (A.8) to the negative singleton or the positive group gives the following lower bounds for a positive and a negative second facility, respectively:

$$a + s\min\left\{1, \left[\frac{2a}{q}-1\right]_+\right\}, \qquad s + \min\left\{a, \left[\frac{2s}{s+q-1}-a\right]_+\right\}. \tag{A.9}$$

A coincident pair also satisfies the needed bounds. For odd $n=2h+1$, take $m=s=h$. For $1 \leq q \leq \frac{2h}{2h-1}$, choose $a=1$, giving ratio at least h. After that choose $a=\frac{2hq}{2h+q}$: both directional costs are at least h, giving ratio $\frac{1}{2}+\frac{h}{q}$ until L_n dominates. For even $n=2h$, take $(m,s)=(h,h-1)$ and $a=\frac{(h-1)q}{h+q-2}$ for $1 \leq q \leq 2$. Both expressions in (A.9) equal $\frac{ha}{q}$, giving $\frac{h}{q}$; for $q \geq 2$ this term is at most L_n. The choice $(m,s)=(h-1,h)$ and $a=\frac{(2h-1)q}{2h+q}$ gives ratio $\frac{(h-1)(2h+q)}{(2h-1)q}$ on $\frac{h}{h-1} \leq q \leq R_n^\dagger - n + 2$. Below this interval it is dominated by $\frac{h}{q}$, and above it by L_n. All chosen a lie in the stated range. Combining the three-location and four-location families proves the result. □

Proposition 3 (Four Agents). *For $n = 4$, no deterministic strategyproof mechanism has robustness below 2. With full-profile predictions, the currently established bounds are*

$$\max\left\{1, \frac{2}{R-1}\right\} \le C_4^*(R) \le \begin{cases} 2, & 2 \le R \le 3, \\ \frac{2}{R-2}, & 3 \le R \le 4, \\ 1, & R \ge 4. \end{cases}$$

The consistency lower bound also applies to facility-pair predictions.

Proof. For the lower bound, fix a prediction correct for the labeled profile $(-1, 0, 0, 1)$. For facility-pair predictions, $\{0, 1\}$ is such a prediction. Use the three-group property of Fotakis & Tzamos (2014), which applies here without the $n \ge 5$ characterization. Either the output is TWO-EXTREMES, with cost 2, or the middle pair receives a facility at 0 for every profile with the same identity-group ordering. The other facility cannot be strictly between 0 and 1: if it were at y, move the left singleton to y. Strategyproofness would require serving this singleton in the middle, contradicting the three-group property. Reflect for the negative side. Thus a noncoincident other facility is at least 1 or at most -1.

Consider a three-location profile $(-A, 0, 0, b)$ on which the other facility is positive. When the left singleton changes her report, 0 always remains served: if she is negative, the middle pair is served; if she is positive, the resulting middle singleton cannot be served and the output is TWO-EXTREMES; two occupied sites follow from zero optimum. Let p be the infimum absolute value of attainable negative facilities. Finite robustness makes this set nonempty, and strategyproofness gives $p \ge 2A$. For a left report $-a$ with $\frac{p}{2} < a < p$, the output must be $\{0, -p\}$, so the cost is $p - a + b$ and $\mathrm{OPT} = \min\{a, b\}$. Letting $a \downarrow \frac{p}{2}$ gives $R \ge 2$ and $\frac{p}{2} + b \le Rb$, hence $A \le (R-1)b$. If the original balanced profile instead uses TWO-EXTREMES, its ratio already gives $R \ge 2$.

Return to $(-1, 0, 0, 1)$ and suppose the other facility is $1 \le y < 2$. For any right report $b \in \left(\frac{y}{2}, 1\right)$, the right agent can obtain cost $y - b < b$ by reporting 1, so the output must still have a positive facility. The preceding bound gives $1 \le (R-1)b$; letting $b \downarrow \frac{y}{2}$ yields $y \ge \frac{2}{R-1}$. The original cost equals y. Coincident facilities or $|y| \ge 2$ cost 2, and the negative case is symmetric. This proves the claimed lower bound without classifying all four-agent mechanisms.

For the upper bound, TWO-EXTREMES gives $(2, 2)$. For $3 \le R \le 4$, set $q = R - 2 \in [1, 2]$ and $\delta = q - 1$. For sorted correct reports, let the consecutive gaps be a, b, c, so $\mathrm{OPT} = \min\{a+b, a+c, b+c\}$. If the optimal split is $2 + 2$, then $a, c \le b$. When $a \le c$, the classical dictator at the second identity returns the second and fourth reports, since $2a \le b + c$, and is optimal; reflect if $c \le a$. If the optimal split is $1 + 3$, then $b, c \le a$. Use D_δ^- at the third identity. Its left span is $A = a + b \ge c$, and if $c > 0$ it has exactly one strictly positive report on the right. The isolated agent's additional cost is at most $\left(\frac{2c}{q} - A\right)_+ \le \left(\frac{2}{q} - 1\right)c \le \left(\frac{2}{q} - 1\right)\mathrm{OPT}$. When $c = 0$, there is no extra cost. Reflect for $3 + 1$ and freeze every choice from the prediction. Lemma 1 applies for four agents as well, giving robustness R and consistency $\frac{2}{q}$. For $R \ge 4$, keep $q = 2$. □

We do not claim that these bounds match between robustness 2 and 4.

B RANDOMIZED MECHANISMS: ADDITIONAL PROOFS AND DETAILS

We now seek guarantees that do not grow with the number of agents. We first give upper bounds with full-profile predictions and with facility-pair predictions, and then prove a lower bound that applies to both models.

B.1 FULL-PROFILE PREDICTIONS

All distributions on identities are determined from the prediction and remain fixed when reports change.

We prove the guarantees stated in Theorem 3.

We first give a common sampling rule and apply the strategyproofness lemma from the main text. We then derive consistency and robustness conditions for its parameters. Uniform central intervals give the trade-off $\mathcal{B}$, and a nonuniform nominee distribution gives the sharper near-perfect guarantee.

B.1.1 MECHANISM AND STRATEGYPROOFNESS

Compute an optimal consecutive two-cluster partition of $\widehat{x}$, breaking ties by a fixed rule. Label the larger group 0 and the smaller group 1, with sizes M, k and ratio $r = \frac{k}{M} \in (0, 1]$. Within each group, sort the predicted locations while retaining identities. For a group of size N_g and width $\alpha_g \in (0, 1]$, its central anchor weights are

$$a_j = \text{length}\left([j-1, j] \cap \left[\frac{(1-\alpha_g)N_g}{2}, \frac{(1+\alpha_g)N_g}{2}\right]\right).$$

These fractional weights are in $[0, 1]$ and sum to $\alpha_g N_g$. They avoid divisibility assumptions about the cluster size. Let $t_0 + t_1 = 1$ be the probabilities of selecting an anchor from the two groups, and set $p_j = \frac{t_g a_j}{\alpha_g N_g}$ for an identity in group g.

For nominations in group g, choose a density φ_g on $[0, \nu_g]$, with integral one, where $0 < \nu_g \leq 1$. Its argument is the normalized rank distance from the median, and its survival function is $S_g(z) = \int_z^1 \varphi_g(w)\, dw$, with density zero beyond ν_g. For rank cell $I_j = \left[\frac{j-1}{N_g}, \frac{j}{N_g}\right]$, define

$$q_j = \int_{I_j} \varphi_g(|2u-1|)\, du, \qquad h_j = N_g \int_{I_j} \mathbf{1}_{\{|2u-1| \leq \nu_g\}}\, du.$$

Then $\sum_j q_j = 1$ and $\sum_j h_j = \nu_g N_g$. Our densities have $q_j > 0$ exactly on cells with $h_j > 0$.

Definition 5 (Prediction-Guided Sampling)**.** Fix the preceding arrays and gains $L_0, L_1 \geq 0$ from the prediction. On actual reports, select an anchor identity i with probability p_i. If i has predicted label g, select a nominee j from the opposite group with probability q_j. Conditional on these choices, select a partner identity ℓ with probability proportional to

$$\left(1 + \frac{L_g h_j}{q_j} \mathbf{1}_{\{\ell = j\}}\right) |x_i - x_\ell|, \tag{B.1}$$

and return $\{x_i, x_\ell\}$. If all reports coincide, return that common location twice.

Only nominees of positive probability are sampled, so the division in (B.1) is well-defined. Every identity retains a positive base coefficient. Thus, on any profile with at most two occupied positions, every positive-distance partner is at the other position and the mechanism has zero cost, regardless of the prediction.

We use the weighted proportional-sampling lemma proved in Section 4.1 (Lemma 3).

Proposition 4. *Every mechanism in Definition 5 is strategyproof.*

Proof. Condition on the anchor and nominee identities, whose probabilities are fixed by the prediction. The coefficients in (B.1) are then positive constants, so Lemma 3 handles every non-anchor. The anchor has zero truthful cost. Averaging these inequalities over the fixed identity lotteries proves the assertion. □

Lemma 7. *Fix an anchor i and two agents j, ℓ, and let c_j, c_ℓ be their distances to a common facility. Writing $d_{uv} = |x_u - x_v|$, we have*

$$d_{ij} \min\{d_{i\ell}, d_{j\ell}\} \leq d_{ij} c_\ell + 2 c_j d_{i\ell}. \tag{B.2}$$

Proof. If $d_{ij} \leq 2 d_{i\ell}$, use $d_{j\ell} \leq c_j + c_\ell$. Otherwise $d_{ij} \leq d_{i\ell} + c_j + c_\ell$ and $d_{i\ell} < c_j + c_\ell$, so $d_{ij} d_{i\ell} \leq 2 d_{i\ell}(c_j + c_\ell) \leq 2 d_{i\ell} c_j + d_{ij} c_\ell$. □

B.1.2 CONSISTENCY AND ROBUSTNESS BOUNDS

With a correct prediction, let A contain anchor i and let B contain nominee j. Write $s = |B|$, $U_i = \sum_{\ell \in A} |x_i - x_\ell|$, $F_j = \sum_{\ell \in B} |x_j - x_\ell|$, $Q = \text{OPT}_B$, and $D_B = \frac{1}{s} \sum_{j,\ell \in B} |x_j - x_\ell|$. The quantity $D_B - Q$ measures how dispersed the cluster is beyond its one-median cost.

Lemma 8 (One-Sided Dispersion). *Let X, X' be independent copies of a nonnegative random variable with finite support. Write $V = \mathbb{E}X$, $Q = \min_z \mathbb{E}|X - z|$, $D = \mathbb{E}|X - X'|$, and $T = \mathbb{E}[X \min\{X', |X - X'|\}]$. Then*

$$Q \le D \le 2Q, \qquad T \le VQ + 2(V - Q)(D - Q). \tag{B.3}$$

Proof. The first two inequalities follow by conditioning on X' and by inserting a median into the triangle inequality, respectively. For the last inequality, put $t(r, s) = r \min\{s, |r-s|\}+s \min\{r, |r-s|\}$ and $\psi_a(r) = r(r - 2a)_+$. For $r, s \ge a \ge 0$, we first establish

$$t(r, s) + (r - s)^2 \le 4a|r - s| + \psi_a(r) + \psi_a(s). \tag{B.4}$$

When $a = 0$, use $t(r, s) \le 2rs$. Otherwise scale to $a = 1$, assume $s = r + d$ with $r \ge 1$ and $d \ge 0$, and note that the left side is $2sd$ if $d \le r$, and s^2 if $d \ge r$. If $s \le 2$, necessarily $d \le r$ and $2sd \le 4d$. If $r \ge 2$ and $d \ge r$, the right side minus the left side is $r^2 - 4r + 2d \ge r(r - 2)$. If $r \ge 2$ and $d \le r$, that difference is $2r(r - 2) + d(2 - d)$, a concave quadratic in d that is nonnegative at 0 and r. Finally, suppose $1 \le r \le 2 \le s$. For $d \ge r$ the difference is $2(d - r)$; for $2 - r \le d \le r$ it is $r(r - 2) + d(2 - d)$, a concave quadratic vanishing at both endpoints. This proves (B.4). Taking half its expectation for independent copies of $Y \ge a$ gives $\mathrm{Var}(Y) + T_Y \le 2aD_Y + \mathbb{E}\psi_a(Y)$.

Split the distribution of X into lower and upper halves at a median, splitting an atom if necessary. Let A, B have these conditional distributions, with means L, U; thus $A \le B$ for every pair in their supports. Then $V = \frac{L+U}{2}$, $Q = \frac{U-L}{2}$, $V - Q = L$, and $D - Q = \frac{D_A+D_B}{4}$. For ordered arguments, $t(A, B) = B^2 - A^2 - B(B - 2A)_+$. Convexity in A therefore gives $\mathbb{E}t(A, B) \le \mathbb{E}B^2 - \mathbb{E}A^2 - \mathbb{E}\psi_L(B)$. Also, the function $F_A(x) = \mathbb{E}|x - A'|$ is 1-Lipschitz, so $\mathrm{Cov}(A, F_A(A)) = \frac{1}{2}\mathbb{E}[(A - A')(F_A(A) - F_A(A'))] \le \mathrm{Var}(A)$. Consequently, $T_A \le \mathbb{E}[AF_A(A)] \le LD_A + \mathrm{Var}(A)$. Apply the preceding consequence of (B.4) to $B \ge L$ to obtain $T_B + \mathrm{Var}(B) - \mathbb{E}\psi_L(B) \le 2LD_B$. Combining these estimates yields

$$4T = T_A + T_B + \mathbb{E}t(A, B) \le U^2 - L^2 + LD_A + 2LD_B \le U^2 - L^2 + 2L(D_A + D_B),$$

which is precisely (B.3). □

Lemma 9. *If the extra sampling coefficient of nominee j is $c \ge 0$, put $b_j = \frac{s}{s+c}$. Then*

$$\mathbb{E}[\mathrm{SC} \mid i, j] \le (1 + b_j)U_i + F_j + 2b_j(D_B - Q). \tag{B.5}$$

Proof. Fix anchor $i \in A$ and nominee $j \in B$, and reflect the line if necessary so that B is to its right. Translate the anchor to zero. Write $s = |B|$, $U = \sum_{\ell \in A} |x_i - x_\ell|$, $V = \sum_{\ell \in B} |x_i - x_\ell|$, $F = \sum_{\ell \in B} |x_j - x_\ell|$, $Q = \mathrm{OPT}_B$, $D = \frac{1}{s}\sum_{k,\ell \in B} |x_k - x_\ell|$, $d = s|x_i - x_j|$, and $G = \sum_{\ell \in B} \min\{|x_i - x_\ell|, |x_j - x_\ell|\}$. Finally, let $T = \sum_{k,\ell \in B} |x_i - x_k| \min\{|x_i - x_\ell|, |x_k - x_\ell|\}$. Rescaling (B.3) from the uniform probability distribution on B gives

$$T \le \min\{V^2, VQ + 2(V - Q)(D - Q)\}, \quad Q \le D \le 2Q, \quad |V - d| \le F \le V + d, \quad G \le \min\{V, F\}.$$

If the nominee's extra coefficient is c, put $b = \frac{s}{s+c}$ and $\theta = 1 - b$. Agents in A have total cost at most U. Separating base partners in A, base partners in B, and the extra nominee weight gives

$$\mathbb{E}[\mathrm{SC} \mid i, j] \le U + \frac{bUV + bT + \theta dG}{b(U + V) + \theta d}.$$

It remains to bound the fraction by $bU + F + 2b(D - Q)$. If $Q = 0$, all of B is colocated, so $F = D = T = G = 0$ and $d = V$; the assertion follows from $\frac{bUV}{bU+V} \le bU$. For $Q > 0$, scale all lengths so $Q = 1$ and write $\lambda = 2D - 1 \in [1, 3]$, $v = V$, $f = F$, and $k = f + b(\lambda - 1)$. The bounds above give $T \le t := \min\{v^2, \lambda v - (\lambda - 1)\}$ and $G \le g := \min\{v, f\}$. After multiplication by the denominator, the desired inequality follows if

$$b^2U^2 + bU[k + \theta(d - v)] + k(bv + \theta d) - bt - \theta dg \ge 0.$$

The coefficient of U is nonnegative: $d - v \ge -f$ implies $k + \theta(d - v) \ge b(f + \lambda - 1)$. For the constant term, if $f \ge v$, then $k \ge v$ and $t \le v^2$, so both $b(kv - t)$ and $\theta d(k - v)$ are nonnegative. If $1 \le f \le v$, the constant term increases with d; substitute $d = v - f$ and divide by b. The result is $(f + \lambda - 1)v - t - \theta(\lambda - 1)f$. When $v \ge \lambda - 1$, use $t \le \lambda v - (\lambda - 1)$ to lower-bound it by $(f - 1)[v - \theta(\lambda - 1)] + b(\lambda - 1) \ge 0$. When $1 \le v \le \lambda - 1$, use $t \le v^2$. The resulting expression is affine in f; at $f = 1$ it is $v(\lambda - v) - \theta(\lambda - 1) \ge b(\lambda - 1)$, by concavity of $v(\lambda - v)$ on this interval, and at $f = v$ it is $b(\lambda - 1)v \ge 0$. This proves (B.5), including its degenerate cases. □

For an anchor in predicted group g, averaging b_j under (B.1) gives $\bar{b}_g = \sum_j \frac{N_{1-g} q_j^2}{N_{1-g} q_j + L_g h_j}$. Integral Cauchy–Schwarz on each rank cell implies

$$\bar{b}_g \le I_{L_g}(\varphi_{1-g}), \qquad I_L(\varphi) = \int_0^\nu \frac{\varphi(z)^2}{\varphi(z) + L}\, dz. \tag{B.6}$$

Indeed, for rank density $d(u)$ and its support indicator $H(u)$, $\frac{(\int_I d)^2}{\int_I (d + LH)} \le \int_I \frac{d^2}{d + LH}$. This also covers cells crossing density or support boundaries, so no large-population approximation is used.

Suppose B_g is an upper bound on (B.6). For $0 \le z < 1$, define $f_\alpha(z) = \frac{z(\alpha - z)_+}{\alpha(1-z)}$. An internal gap whose smaller rank fraction is p has imbalance $z = 1 - 2p$. Relative to its optimal one-median cost, the anchor cost has multiplier $1 + f_\alpha(z)$, the nominee cost has multiplier $1 + \frac{zS(z)}{1-z}$, and $D_B - Q$ has coefficient z. For example, the probability a symmetric nominee is to the left of the gap is $\frac{1}{2}S(z)$; substituting this in the gap's expected distance gives the nominee multiplier. All costs are nonnegative sums over internal gaps. Consequently, (B.5) yields the following sufficient consistency condition for each group:

$$\mathcal{K}_g(z) = 1 + t_g B_g + t_g(1 + B_g) f_{\alpha_g}(z) + (1 - t_g)\frac{z S_g(z)}{1 - z} + 2(1 - t_g) B_{1-g} z \le C. \tag{B.7}$$

Singleton groups have no internal gaps.

Robustness requires a different partition: fix the actual optimal clusters, not the predicted ones, and write c_ℓ for identity ℓ's distance to her actual optimal center. For an anchor in actual cluster A and nominee j, direct sampling-cost accounting gives

$$\mathbb{E}[\mathrm{SC} \mid i, j] \le 2U_i + 3\,\mathrm{OPT}_B + c_j^{\mathrm{boost}}\big(\mathbf{1}_{\{j \in A\}} |x_i - x_j| + 2\mathbf{1}_{\{j \in B\}} c_j\big). \tag{B.8}$$

Here c_j^{boost} is the extra coefficient, not the optimal distance c_j. To prove (B.8), fix actual optimal clusters A, B, an anchor $i \in A$, and write $U = \sum_{\ell \in A} |x_i - x_\ell|$, $V = \sum_{\ell \in B} |x_i - x_\ell|$, $Q = \mathrm{OPT}_B$. Let the nominee's extra coefficient be c and let $d = |x_i - x_j|$. The product inequality (B.2), summed over ordered pairs in B, bounds the base numerator for serving B by $UV + 3VQ$. If $j \in A$, the extra numerator is at most cdV. Divide by $U + V + cd$: the base part is at most $U + 3Q$, and the extra part is at most cd. Adding the cost U for A gives $2U + 3Q + cd$. If $j \in B$, let c_j denote its distance to the optimal center of B. The same product inequality gives an extra numerator at most $cdQ + 2cc_jV$. The combined numerator is at most $UV + (3V + cd)Q + 2cc_jV$, whose quotient is at most $U + 3Q + 2cc_j$. Again adding U proves (B.8). The identity $q_j \frac{L_g h_j}{q_j} = L_g h_j$ makes the extra coefficient especially simple after averaging nominees. Using $U_i \le |A| c_i + \mathrm{OPT}_A$ and then averaging anchors, the coefficient of c_ℓ, for an identity with predicted label g and actual cluster S, is at most

$$\mathcal{R}_\ell(S) = 2 + 2|S| p_\ell + p_{S^c} + L_g p_\ell \sum_{\substack{j \in S \\ \widehat{L}(j) = 1-g}} h_j + L_{1-g} h_\ell \left(\sum_{\substack{i \in S \\ \widehat{L}(i) = 1-g}} p_i + 2 \sum_{\substack{i \notin S \\ \widehat{L}(i) = 1-g}} p_i \right). \tag{B.9}$$

The same actual cluster S occurs in all terms. Bounding this expression by R for every identity and cluster proves robustness R.

Averaging the extra coefficient over the nominee replaces $q_j c$ by $L_g h_j$. For the base part, insert $U_i \le |A| c_i + \mathrm{OPT}_A$ and then average over anchors. The coefficient of a focal c_ℓ, with actual cluster S, is $2 + 2|S| p_\ell + p_{S^c}$. An extra nominee in the anchor's actual cluster contributes its distance to the anchor, at most $c_i + c_j$; an extra nominee in the other actual cluster contributes $2c_j$. Collecting these terms gives exactly (B.9).

We next give a finite upper bound for this expression that also accommodates fractional rank cells. Put $\kappa_g = \frac{t_g}{\alpha_g N_g}$, so $p_j = \kappa_g a_j$. Because anchor and nominee supports are central intervals, their masks are nested. For label g, the possible corner types (a, h) and their total masses are

$$(1,1) : N_g \min\{\alpha_g, \nu_g\}, \quad (1,0) : N_g(\alpha_g - \nu_g)_+, \quad (0,1) : N_g(\nu_g - \alpha_g)_+, \quad (0,0) : N_g(1 - \max\{\alpha_g, \nu_g\}).$$

Each fractional cell is a convex combination of these corners, with the displayed aggregate masses. For fixed focal weights, (B.9) is affine in the other weights and in their membership in S. Allowing the mass of each corner type to be assigned independently to S or S^c therefore gives an upper bound, even when such an assignment does not correspond to an actual consecutive cluster. Maximizing these affine functions is convex in the focal weights as well, so it suffices to check focal corners.

For a focal corner $(1, h)$ of label g, start with all agents in S. Its value is $2 + 2n\kappa_g + L_g\kappa_g\nu_o N_o + hL_o t_o$, where $o = 1 - g$. Moving an opposite-label corner (a_j, h_j) outside S changes the value per unit mass by $\kappa_o a_j(1 + hL_o) - 2\kappa_g - L_g\kappa_g h_j$. Moving a same-label corner changes it by $\kappa_g a_j - 2\kappa_g \le 0$. Thus the focal value is at most

$$\begin{aligned}\Gamma_g(h) = 2 + 2n\kappa_g + L_g\kappa_g\nu_o N_o + hL_o t_o \\ + N_o \min\{\alpha_o, \nu_o\}\big[\kappa_o(1 + hL_o) - \kappa_g(2 + L_g)\big]_+ \\ + N_o(\alpha_o - \nu_o)_+\big[\kappa_o(1 + hL_o) - 2\kappa_g\big]_+. \end{aligned} \tag{B.10}$$

The $(0, 0)$ and $(0, 1)$ focal corners are bounded by 3 and $3 + 2L_o t_o$, respectively. If $\alpha_g \le \nu_g$, check 3, $3 + 2L_o t_o$, and $\Gamma_g(1)$; if $\alpha_g \ge \nu_g$, check 3, $\Gamma_g(0)$, and $\Gamma_g(1)$. These are sufficient bounds, with absent corners harmlessly retained. In particular, when only label 0 supplies anchors, with anchor width α, opposite nominee width ν, and gain L, they give

$$R \le \max\left\{2 + \frac{2(1 + r) + L\nu r}{\alpha},\ 3 + 2L\right\}. \tag{B.11}$$

The base part alone is at most $2 + 2n \max_i p_i$: in $2 + 2|S|p_\ell + p_{S^c}$, use $p_{S^c} \le (n - |S|) \max_i p_i$.

B.1.3 Uniform Central Sampling

We first give the trade-off $\mathcal{B}$ in Theorem 3 using uniform central intervals. Use a common anchor width α and nominee width $\nu \ge \alpha$ in the two predicted groups, with common gain $L \ge 0$. For a nominee group of size N, uniform sampling gives $q_j = \frac{h_j}{\nu N}$. Thus the extra coefficient is the same $L\nu N$ for every nominee of positive probability, even when its rank cell is only partly included. The background factor is exactly $b = \frac{1}{1+L\nu}$. If a group's anchor probability is t, its gap multiplier is therefore

$$\mathcal{K}(z) = t(1 + b)(1 + f_\alpha(z)) + (1 - t)(1 + f_\nu(z) + 2bz). \tag{B.12}$$

We will bound both components separately, leaving freedom to choose t for robustness.

Lemma 10. *Put $a = \frac{2}{\alpha}$, $A = a - L$, $B = a + \frac{L\nu}{\alpha}$, $D = B - 2 - L$, and $E = a - 1 - 2L$. For anchor-group probabilities $t_0 = t$, $t_1 = 1 - t$, robustness is at most the maximum of $3 + 2L$ and*

$$\begin{aligned} F_0 &= 2 + L + t(A + Br), & F_1 &= 2 + L + (1 - t)\left(A + \frac{B}{r}\right), \\ G_0 &= 3 + 2L + t(E + Dr), & G_1 &= 3 + 2L + (1 - t)\left(E + \frac{D}{r}\right). \end{aligned} \tag{B.13}$$

Moreover, the probability

$$t_{\mathrm{bal}}(r) = \min\left\{1, \frac{B + Ar}{B(1 + r^2) + 2Ar}\right\} \tag{B.14}$$

bounds both F_0, F_1 by $R_{\mathrm{bal}} = 2 + L + \frac{A+B}{2}$.

Proof. Since $\alpha \le \nu$, only the corners $(0, 0)$, $(0, 1)$, and $(1, 1)$ occur. The first two are bounded by 3 and $3 + 2L$. For the full corner of the larger group, (B.10) becomes

$$2 + a(1 + r)t + \frac{L\nu}{\alpha}rt + L(1 - t) + \big[(1 - t)(1 + L) - rt(2 + L)\big]_+.$$

Expanding the positive part gives $\max\{F_0, G_0\}$; exchanging the labels gives $\max\{F_1, G_1\}$. The rank-cell corner argument proves the same bound for fractional cells.

For the last assertion, $B \geq |A|$ and $A + B > 0$, so $J = B\big(1 + r^2\big) + 2Ar$ and $B + Ar$ are positive. If the fraction in (B.14) is at most one, then

$$F_0 = F_1 = 2 + L + \frac{(A + Br)(Ar + B)}{J}, \qquad \frac{A + B}{2} J - (A + Br)(Ar + B) = \frac{B(B - A)}{2}(1 - r)^2 \geq 0.$$

If clipping occurs, $A + Br < 0$ and $t = 1$, so $F_0 \leq 2 + L = F_1 < R_{\mathrm{bal}}$. This proves the assertion without a sign assumption on A. □

Proposition 5. *For $0 < e \leq \frac{2}{3}$, uniform central sampling can achieve consistency $1 + e$ and robustness $1 + \frac{7}{2e} + \frac{1}{e^2}$.*

Proof. Choose

$$\alpha = 1 - \frac{1}{(1 + e)^2}, \qquad \nu = \frac{4e}{1 + 4e}, \qquad b = \frac{e}{2}, \qquad L = \frac{(2 - e)(1 + 4e)}{4e^2}. \tag{B.15}$$

These parameters satisfy $\alpha \leq \nu$ for every $0 < e \leq 2$, and $b = \frac{1}{1 + L\nu}$. The following consistency argument holds on this larger interval as well. The maximum of $1 + f_\alpha(z)$ is $\frac{2}{1 + \sqrt{1 - \alpha}}$, and hence the first component of (B.12) is at most $(1 + b)\frac{2}{1 + \sqrt{1 - \alpha}} = 1 + e$. For the other component, differentiation shows that

$$\max_{0 \leq z < 1} \frac{f_\nu(z)}{1 - z} = \frac{\nu}{4(1 - \nu)} = e, \qquad f_\nu(z) + 2bz \leq e(1 - z) + ez = e.$$

The maximum occurs at $z = \frac{\nu}{2 - \nu}$; beyond the support $f_\nu = 0$. Thus consistency is at most $1 + e$, independently of the anchor-group probability.

For robustness, use the notation of Lemma 10, and put $H = -E = 1 + 2L - a$ and $R_0 = 3 + 2L$. Substitution in (B.15) gives

$$B = \frac{(1 + e)^2}{e^2}, \qquad D = \frac{e + 2}{4e^2}, \qquad H = \frac{4 + 12e - 5e^2 - 6e^3}{2e^2(e + 2)}.$$

We need $H \geq D + 1$ and $H^2 \geq BD$. For the first inequality, $H - D - 1 = \frac{4 + 20e - 19e^2 - 16e^3}{4e^2(e + 2)} > 0$: the numerator is concave on $\big[0, \frac{2}{3}\big]$, and its endpoint values are 4 and $\frac{112}{27}$. For the second,

$$\begin{aligned} 4e^4(e + 2)^2(H^2 - BD) &= 8 + 68e + 66e^2 - 193e^3 - 127e^4 + 59e^5 + 36e^6 \\ &\geq 8 + 68e - \frac{1072}{9} e^2 \geq \frac{32}{81} > 0. \end{aligned}$$

The first inequality uses $e \leq \frac{2}{3}$; the last quadratic is concave and has endpoint values 8 and $\frac{32}{81}$.

Choose $t = 1$ when $r \leq \frac{D}{H}$, and $t = t_{\mathrm{bal}}(r)$ otherwise. The threshold lies in $(0, 1)$ because $H > D + 1$. Below the threshold, (B.13) gives $F_0 = 2 + a + Br \leq 2 + a + \frac{BD}{H} \leq 2 + a + H = R_0$, $F_1 = 2 + L < R_0$, and $G_1 = R_0$. Also $G_0 \leq R_0$, since $E + Dr \leq D - H < 0$. Above the threshold, both $E + Dr$ and $E + \frac{D}{r}$ are nonpositive, so both G bounds are at most R_0 for any t. The inequality $H \geq D + 1$ is equivalent to $R_{\mathrm{bal}} \leq R_0$, and Lemma 10 therefore handles both F bounds, including the clipped case. The remaining corner is also bounded by R_0. Finally, $R_0 = 1 + \frac{7}{2e} + \frac{1}{e^2}$. The threshold depends only on predicted cluster sizes, so switching between these anchor probabilities does not affect strategyproofness. □

Proposition 6. *There is a uniform central-sampling mechanism with consistency $\frac{7}{4}$ and robustness $\frac{15}{2}$.*

Proof. Use (B.15) with $e = \frac{3}{4}$. The consistency argument in Proposition 5 remains valid, while the robustness parameters are

$$\alpha = \frac{33}{49}, \quad \nu = \frac{3}{4}, \quad L = \frac{20}{9}, \quad A = \frac{74}{99}, \quad B = \frac{49}{9}, \quad D = \frac{11}{9}, \quad E = -\frac{245}{99}.$$

For $R = \frac{15}{2}$, choose $t = \min\{1, \frac{649}{148+1078r}\} = \min\{1, \frac{R-2-L}{A+Br}\}$. This directly ensures $F_0 \le R$. Since $R_{\mathrm{bal}} = \frac{161}{22} < R$, this probability is at least t_{bal}. Here $A > 0$, so F_1 decreases with t and is at most R_{bal}. Also $G_0 \le 3 + 2L = \frac{67}{9} < R$, because $E + Dr < 0$ for $r \le 1$.

For G_1, either $t = 1$ or $E + \frac{D}{r} \le 0$ gives the same bound immediately. In the remaining case, $t < 1$ implies $r > \frac{501}{1078} > \frac{4}{9}$, whereas $E + \frac{D}{r} > 0$ implies $r < \frac{121}{245} < \frac{1}{2}$. Consequently,

$$t \ge \frac{649}{687} > \frac{9}{10}, \qquad E + \frac{D}{r} \le -\frac{245}{99} + \frac{11}{4} = \frac{109}{396} < \frac{1}{3}, \qquad G_1 \le \frac{67}{9} + \frac{1}{30} = \frac{673}{90} < \frac{15}{2}.$$

The other corners are bounded by $\frac{67}{9}$, completing the proof. □

Proposition 7. *For every $2 \le c \le 3$, there is a uniform central-sampling mechanism with consistency c and robustness $2 + \frac{c^2}{(c-1)^2}$.*

Proof. Use equal anchor and nominee widths $w = 1 - \frac{1}{c^2}$, background factor $b = \frac{c-1}{2}$, and gain $L = \frac{1-b}{wb} = \frac{c^2(3-c)}{(c-1)^2(c+1)}$. The first component of (B.12) is at most $(1+b)\frac{2}{1+\sqrt{1-w}} = c$. For the second, use $\frac{f_w(z)}{1-z} \le \frac{w}{4(1-w)} = \frac{c^2-1}{4} \le c - 1 = 2b$. Thus both components are at most c.

Write $a = \frac{2}{w}$, so $\frac{9}{4} \le a \le \frac{8}{3}$ and $0 \le L \le \frac{a}{2}$. Here $A = a - L$, $B = a + L$, $D = a - 2$, and $E = a - 1 - 2L$. Choose $t = t_{\mathrm{bal}}$, equivalently $t = \frac{1+\eta r}{1+2\eta r+r^2}$ with $\eta = \frac{3c-5}{c+1} \in \left[\frac{1}{3}, 1\right]$. No clipping is needed. Lemma 10 bounds both F expressions by $R = 2 + a + L = 2 + \frac{c^2}{(c-1)^2}$. The remaining constant bound is smaller, since $R - (3 + 2L) = a - 1 - L \ge \frac{a}{2} - 1 > 0$.

For the G expressions, put $J = B\left(1 + r^2\right) + 2Ar$, $T = a + (a-3)L - L^2$, $U = a - (2a-3)L$, and $P(v) = BLv^2 + Uv + T$. Expansion of (B.13) gives

$$(R - G_1)J = P(r), \qquad (R - G_0)J = r^2 P\left(\tfrac{1}{r}\right).$$

It suffices to show $P(v) \ge 0$ for all $v \ge 0$. Because $a < 3$ and $L \le \frac{a}{2}$, we have $T \ge a + (a-3)\frac{a}{2} - \frac{a^2}{4} = \frac{a(a-2)}{4} > 0$. If $U \ge 0$, every coefficient is nonnegative. If $U < 0$, then $L > \frac{a}{2a-3} > 1$ and $a > \frac{5}{2}$. It follows that $BL > \frac{7}{2}$, $T > \frac{5}{16}$, and $U \ge \frac{a(5-2a)}{2} \ge -\frac{4}{9}$. Hence

$$P(v) \ge \frac{7}{2}v^2 - \frac{4}{9}v + \frac{5}{16} \ge \frac{5}{16} - \frac{8}{567} = \frac{2707}{9072} > 0.$$

This proves the robustness bound for every population ratio. □

To complete the full-range guarantee in Theorem 3, we interpolate between the preceding guarantees and the prediction-free guarantee. A fixed mixture of two strategyproof mechanisms remains strategyproof, and its consistency and robustness are bounded by the corresponding convex combinations. Proposition 5 gives $\left(\frac{5}{3}, \frac{17}{2}\right)$, Proposition 6 gives $\left(\frac{7}{4}, \frac{15}{2}\right)$, and Proposition 7 gives $(2, 6)$ and $\left(\frac{5}{2}, \frac{43}{9}\right)$. Mixing the first two points with weight $21 - 12c$ on the first gives robustness $\frac{57}{2} - 12c$ for $c \in \left[\frac{5}{3}, \frac{7}{4}\right]$. Mixing the second and third with weight $8 - 4c$ on the second gives $18 - 6c$ for $c \in \left[\frac{7}{4}, 2\right]$. The straight-line bounds therefore require no additional parameter cases.

For $c \in \left[\frac{5}{2}, \rho\right]$, use the mechanism at $\left(\frac{5}{2}, \frac{43}{9}\right)$ with probability $\theta = \frac{\rho - c}{\rho - 5/2}$, and otherwise ignore the prediction and use the ρ-approximation mechanism of Chan et al. (2026). This gives

$$C \le \theta\frac{5}{2} + (1-\theta)\rho = c, \qquad R \le \theta\frac{43}{9} + (1-\theta)\rho = \frac{(\rho - c)(43/9) + (c - 5/2)\rho}{\rho - 5/2}.$$

At $c = \rho$, the prediction-dependent component disappears and the two guarantees coincide. Together with the propositions above, this proves the full-range guarantee $\mathcal{B}$ in Theorem 3.

B.1.4 NEAR-PERFECT CONSISTENCY

The preceding analysis separates parameter selection from strategic reports. For near-perfect consistency, a useful nominee distribution satisfies $S(z) \le E\frac{(1-z)^2}{z}$. If $\nu \le \frac{4E}{1+4E}$, use the uniform density on $[0,\nu]$. Indeed, its survival function is $S(z) = \max\{1 - \frac{z}{\nu}, 0\}$, and $\max_{0\le z<1} \frac{zS(z)}{(1-z)^2} = \frac{\nu}{4(1-\nu)} \le E$. Otherwise put $z_1 = \frac{2E}{1+2E}$, $z_2 = \frac{\nu}{2-\nu}$, $c_1 = 1 + \frac{1}{4E}$, and $c_2 = E\big(z_2^{-2} - 1\big)$, and use

$$S(z) = \begin{cases} 1 - c_1 z, & 0 \le z \le z_1, \\ E\frac{(1-z)^2}{z}, & z_1 \le z \le z_2, \\ c_2(\nu - z), & z_2 \le z \le \nu, \\ 0, & z \ge \nu. \end{cases} \tag{B.16}$$

The pieces and their first derivatives match at z_1 and z_2, and $S(0) = 1$ and $S(\nu) = 0$. The density is $-S'$. The two straight pieces are tangents to the middle convex curve, so the required survival bound holds throughout.

Proof of the near-perfect guarantee in Theorem 3. Put $e = \varepsilon \in \big(0, \frac{1}{100}\big]$, $\lambda_* = \frac{1}{3} + \frac{9}{5}e$, and $R_* = 3 + \frac{2\lambda_*}{e^2}$. At a boundary between parameter regimes, either adjacent choice can be used. All choices depend on the predicted ratio r only, and hence remain fixed under misreports.

Step 1: Moments of the Nominee Density.

For the curved law (B.16), put $v = \frac{2E(1-\nu)}{\nu}$. Integration of its two constant pieces and its middle density $E\big(z^{-2} - 1\big)$ gives

$$J := \int \varphi^2 = 1 + \frac{1 + 4v^3}{6E}, \qquad K := \int \varphi^3 = 1 + \frac{3/8 + 2v^4}{E} + \frac{3/80 + 4v^5/5}{E^2}.$$

For example, the middle contributions use the antiderivatives $E^2\big(-\frac{1}{3z^3} + \frac{2}{z} + z\big)$ and $E^3\big(-\frac{1}{5z^5} + \frac{1}{z^3} - \frac{3}{z} - z\big)$; the constant pieces contribute their length times the appropriate density power. If $\varphi \le M$, then $\frac{1}{L+x} = \frac{1}{L} - \frac{x}{L(L+x)} \le \frac{1}{L} - \frac{x}{L(L+M)}$ implies $I_L(\varphi) \le \frac{J}{L} - \frac{K}{L(L+M)}$. For $E = e$, $v \le \frac{1}{3}$, and $L = \frac{1+4v^3}{3e^2} + \frac{9}{5e}$, we have $M = 1 + \frac{1}{4e}$, $J - \frac{eL}{2} = \frac{1}{10}$, and

$$K - \frac{L+M}{10} = \frac{9}{10} + \frac{17/100 + 2v^4}{e} + \frac{1/240 + 4v^5/5 - 2v^3/15}{e^2} > 0.$$

To see the last sign, multiply its final numerator by 240 to get $1 + 192v^5 - 32v^3$. For $v \le \frac{1}{4}$ this is at least $\frac{1}{2}$; for $\frac{1}{4} \le v \le \frac{1}{3}$, use $192v^5 \ge 12v^3$ to obtain at least $\frac{7}{27}$. It follows that $I_L(\varphi) \le \frac{e}{2}$.

Step 2: Very Unequal Groups ($0 < r \le 3e$).

Set $t_0 = 1$, $\alpha_0 = \frac{e(2+e)}{(1+e)^2}$, and $\alpha_1 = e$. Use a uniform nominee law of width α_0 in group 0, and (B.16) with $E = e, \nu_1 = 1$ in group 1. Take $L_0 = \frac{\lambda_*}{e^2}$ and $L_1 = \frac{1}{\alpha_0 e}$. The preceding moment bound gives $B_0 = \frac{e}{2}$; the uniform law gives $B_1 \le e$. The maximum of $1 + f_\alpha(z)$ is $\frac{2}{1+\sqrt{1-\alpha}}$, so for group 0, (B.7) is bounded by $\frac{2(1+e/2)}{1+\sqrt{1-\alpha_0}} = 1 + e$. For group 1, the survival bound gives $\frac{zS_1(z)}{1-z} + ez \le e$. Thus both consistency conditions hold. By (B.11), robustness is at most the maximum of R_* and $2 + \frac{2(1+r)+L_0 r}{\alpha_0}$. The latter increases with r. At $r = 3e$, its difference from R_* is $\frac{5-23e-243e^2-276e^3-90e^4}{15e^2(e+2)} > 0$ on the stated interval.

Step 3: Unequal Groups ($3e < r \le \frac{1}{2}$).

Put $u = r - 3e$, $\ell = \lambda_* + \frac{32}{81}u^3$, $s = \frac{r\ell}{1+r\ell}$, and $t_0 = 1 - s$, $t_1 = s$. Choose $\alpha_0 = \frac{2(1+s)e}{1+6e}$, $\alpha_1 = e$, $\nu_0 = \alpha_0$, and $\nu_1 = \frac{3e}{r}$. The nominee law is uniform in group 0, and is (B.16) with $E = e$ in group 1. Take $L_0 = \frac{\ell}{e^2}$ and $L_1 = \frac{1}{\alpha_0 e}$. Here $v = \frac{2u}{3} \le \frac{1}{3}$, so again $B_0 = \frac{e}{2}$ and $B_1 = e$ are valid. We will use $\ell \le \frac{1631}{4050}$, $s < \frac{1}{3}$, and $s \le t_0$.

Table 1: Parameters for the two middle ratio ranges.

Range	g	t_0	A_g	N_g	T_g	μ_g	b_g	ℓ_g
$\left[\frac{1}{2}, \frac{3}{5}\right]$	0	$\frac{3}{4}$	$\frac{9}{4}$	$\frac{9}{4}$	$\frac{9}{8}$	$\frac{1}{2}$	$\frac{11}{20}$	$\frac{1292}{3465}$
	1		$\frac{12}{5}$	6	$\frac{9}{5}$	$\frac{13}{25}$	$\frac{7}{10}$	$\frac{40}{63}$
$\left[\frac{3}{5}, \frac{3}{4}\right]$	0	$\frac{13}{20}$	$\frac{5}{2}$	$\frac{5}{2}$	$\frac{5}{4}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{31}{70}$
	1		2	5	$\frac{3}{2}$	$\frac{9}{20}$	$\frac{4}{5}$	$\frac{1}{2}$

For group 0, the gap multiplier minus one is $\frac{t_0 e}{2} + \left(1 + \frac{t_0 e}{2}\right) f_{\alpha_0}(z) + 2sez$. Outside the support, $\frac{t_0}{2} + 2s \le 1$ suffices. Inside the support use $f_\alpha(z) \le \frac{\alpha}{4(1-\alpha)}$, $z \le \alpha_0$, and divide by the remaining allowance $\frac{e(1+s)}{2}$. The two terms are bounded by $\frac{1+e/2}{1+10e/3}$ and $\frac{8e/3}{1+6e}$; their sum is at most one, since after multiplication by the positive denominators the slack is $e\left(\frac{1}{6} + \frac{73}{9}e\right)$. For group 1 and $z \ge e$, the anchor term vanishes, and the survival bound immediately gives (B.7). For $z \le e$, put $y = \frac{z}{e}$. Since $s \le t_0$, it suffices that

$$\frac{(1+e) f_e(z)}{e} + \frac{z S_1(z)/(1-z) + ez}{e} \le \frac{103}{50} y - \frac{13}{10} y^2 \le \frac{10609}{13000} < 1.$$

For the first inequality use $\frac{1+e}{1-e} \le \frac{21}{20}$ and the first tangent piece of S_1: the two summands are at most $\frac{21}{20} y(1-y)$ and $(1+e)y - \frac{1}{4}y^2$, respectively.

For robustness, the parameters satisfy $L_0 \kappa_0 = L_1 \kappa_1$ and $\frac{1}{2} \le \frac{\kappa_1}{\kappa_0} \le 2$. Indeed, the ratio is $\frac{\ell \alpha_0}{e}$, with lower bound $\frac{100}{159}$ and upper bound $\frac{8}{3}\frac{1631}{4050}$. Anchor masks are contained in nominee masks. For a focal $(1,1)$ corner, moving an opposite-label identity out of its actual cluster changes the coefficient by $\kappa_o a_j - 2\kappa_g + L_o \kappa_o a_j - L_g \kappa_g h_j \le 0$. Thus the all-in-one-cluster value bounds this corner. For the two labels these values are

$$H_0 = 2 + \frac{(1+6e)[2e(1+r) + \ell(3+r)]}{2e^2(1+2r\ell)}, \qquad H_1 = 2 + \frac{2\ell[1 + e(1+r)]}{e^2(1+r\ell)}.$$

Other corners are bounded by 3, $3 + 2L_0 t_0$, or $3 + 2L_1 s$. The inequality $\frac{\ell}{1+r\ell} \le \lambda_*$ follows from $\frac{32}{81}u^3 \le r\lambda_* \ell$, using $u \le r \le \frac{1}{2}$ and $\lambda_*, \ell \ge \frac{1}{3}$. Also $L_1 s \le L_0 t_0$, since $re \le \alpha_0$. Hence these other corners are at most R_*. For H_1, the required scaled slack is

$$\begin{aligned}
&(\lambda_* + \tfrac{1}{2}e^2)(1 + r\ell) - \ell[1 + e(1+r)] \\
&\quad = -\frac{32}{81}u^3 + \frac{e^2}{2} + \ell\left[u\left(\frac{1}{3} + \frac{4}{5}e + \frac{e^2}{2}\right) + \frac{12}{5}e^2 + \frac{3}{2}e^3\right] \ge \frac{u}{81} + \frac{e^2}{2} \ge 0.
\end{aligned}$$

For H_0, multiply $R_* - H_0$ by $2e^2(1+2r\ell)$. Combining $4\lambda_* \ge \frac{4}{3}$ with the terms containing ℓ, and using $8\lambda_* - 1 \ge \frac{5}{3}$, $r \le \frac{1}{2}$, and $e \le \frac{1}{100}$, gives the lower bound

$$2e^2(1+2r\ell) + 4\lambda_*(1+2r\ell) - (1+6e)[2e(1+r) + \ell(3+r)] \ge \frac{4}{3} - \frac{159}{50}\frac{1631}{4050} - \frac{159}{5000} = \frac{2821}{135000} > 0.$$

Step 4: Two Middle Ranges.

For $r \in \left[\frac{1}{2}, \frac{3}{5}\right]$ or $r \in \left[\frac{3}{5}, \frac{3}{4}\right]$, use Table 1. In this table anchor widths are $A_g e$, nominee widths are $N_g e$, and the nominee density has mass μ_g uniformly on $[0, T_g e]$ and mass $1 - \mu_g$ uniformly on $[T_g e, N_g e]$. Gains are $\frac{\ell_g}{e^2}$, and the background upper bounds are $b_g e$. For each row pair, ℓ_g is the opposite nominee law's scaled second moment divided by b_g. Thus $I_{L_g}(\varphi_{1-g}) \le \frac{\int \varphi_{1-g}^2}{L_g} \le b_g e$.

For consistency, put $z = ey$, let $s_g(y)$ be the scaled survival function, and abbreviate $t = t_g$, $A = A_g$, $b = b_g$, $b_o = b_{1-g}$. The gap slack multiplied by $\frac{1-ey}{e}$ is

$$P(y;e) = [1 - tb - 2eb_o(1-t)y](1-ey) - t(1+eb)y\left(1 - \tfrac{y}{A}\right)_+ - (1-t)y s_g(y). \qquad \text{(B.17)}$$

Its derivative in e is $4eb_o(1-t)y^2 - [1 - tb + 2b_o(1-t)]y - tby\left(1 - \frac{y}{A}\right)_+ \le 0$ whenever $e \le \frac{1}{2\max\{A_g, N_g\}}$. Therefore it suffices to check the piecewise quadratics at $e = \frac{1}{100}$. Here is the full

Table 2: Coefficients and minima of $P\big(y;\frac{1}{100}\big) = p_0 + p_1 y + p_2 y^2$ in (B.17). The ranges I and II refer to the two row pairs in Table 1.

Range	g	y interval	p_0	p_1	p_2	Minimum
I	0	$\big[0, \frac{9}{8}\big]$	$\frac{47}{80}$	$-\frac{2027}{2000}$	$\frac{803363}{1800000}$	$\frac{155867}{12800000}$
I	0	$\big[\frac{9}{8}, \frac{9}{4}\big]$	$\frac{47}{80}$	$-\frac{2027}{2000}$	$\frac{803363}{1800000}$	$\frac{38975}{3213452}$
I	1	$\big[0, \frac{9}{5}\big]$	$\frac{33}{40}$	$-\frac{4073}{4000}$	$\frac{64329}{200000}$	$\frac{393527}{20585280}$
I	1	$\big[\frac{9}{5}, \frac{12}{5}\big]$	$\frac{33}{40}$	$-\frac{21911}{28000}$	$\frac{800909}{4200000}$	$\frac{39804069}{1794036160}$
I	1	$\big[\frac{12}{5}, 6\big]$	$\frac{33}{40}$	$-\frac{7431}{14000}$	$\frac{240231}{2800000}$	$\frac{2280}{560539}$
II	0	$\big[0, \frac{5}{4}\big]$	$\frac{27}{40}$	$-\frac{2539}{2500}$	$\frac{100339}{250000}$	$\frac{5219}{160000}$
II	0	$\big[\frac{5}{4}, \frac{5}{2}\big]$	$\frac{27}{40}$	$-\frac{2539}{2500}$	$\frac{100339}{250000}$	$\frac{652723}{20067800}$
II	1	$\big[0, \frac{3}{2}\big]$	$\frac{18}{25}$	$-\frac{2033}{2000}$	$\frac{74293}{200000}$	$\frac{730939}{29717200}$
II	1	$\big[\frac{3}{2}, 2\big]$	$\frac{18}{25}$	$-\frac{12281}{14000}$	$\frac{390051}{1400000}$	$\frac{32228011}{1092142800}$
II	1	$\big[2, 5\big]$	$\frac{18}{25}$	$-\frac{36709}{70000}$	$\frac{143091}{1400000}$	$\frac{94806599}{2003274000}$

calculation. On a piece where $s_g(y) = u - vy$, let $\chi = 1$ if $y \le A$ on that piece and $\chi = 0$ otherwise. Writing $P\big(y;\frac{1}{100}\big) = p_0 + p_1 y + p_2 y^2$, expansion of (B.17) gives

$$p_0 = 1 - tb, \qquad p_1 = -\frac{p_0 + 2b_o(1-t)}{100} - t\left(1 + \frac{b}{100}\right)\chi - (1-t)u,$$
$$p_2 = \frac{b_o(1-t)}{5000} + \frac{t(1 + b/100)\chi}{A} + (1-t)v.$$

The survival coefficients are $(u, v) = \big(1, \frac{\mu_g}{T_g}\big)$ below T_g, and $\big(\frac{(1-\mu_g)N_g}{N_g - T_g}, \frac{1-\mu_g}{N_g - T_g}\big)$ between T_g and N_g; beyond N_g, both are zero. For a piece $[l, h]$, its minimum is the smaller endpoint value, together with $p_0 - \frac{p_1^2}{4p_2}$ if $p_2 > 0$ and the vertex $-\frac{p_1}{2p_2}$ lies in $[l, h]$. Table 2 lists every coefficient and the resulting minimum. For example, the first polynomial decreases up to the endpoint $\frac{9}{8}$, where its value is $\frac{155867}{12800000}$; on the next interval its vertex gives $\frac{38975}{3213452}$. All listed minima are positive. Outside the supports, the condition is $t_g b_g + 2(1 - t_g)b_{1-g} \le 1$, also satisfied by the table.

For robustness, separate the base and extra terms in (B.9). Put $n_0 = 1$, $n_1 = r$, and $k_g = \frac{t_g}{A_g n_g}$. The extra coefficient, multiplied by e^2, is at most the maximum over labels g of

$$2\ell_o t_o, \qquad \ell_g k_g N_o n_o + \ell_o t_o + n_o \min\{A_o, N_o\}[k_o \ell_o - k_g \ell_g]_+ + n_o (A_o - N_o)_+ k_o \ell_o.$$

This follows from the same corner calculation as (B.10), with the base terms removed. For $g = 0$ it is affine plus positive parts in r, and for $g = 1$ in $\frac{1}{r}$; hence the two ratio endpoints suffice. To make the endpoint comparison explicit, write $V_g(r)$ for the second expression above and $W_g = 2\ell_o t_o$ for the first. The endpoint pairs (V_0, V_1) are

$$\left(\frac{9376}{17325}, \frac{127}{220}\right), \quad \left(\frac{10502}{17325}, \frac{646}{1155}\right) \qquad \text{at } r = \frac{1}{2}, \frac{3}{5} \text{ in the first range,}$$
$$\left(\frac{2438}{4375}, \frac{10961}{16800}\right), \quad \left(\frac{8527}{14000}, \frac{1217}{2100}\right) \qquad \text{at } r = \frac{3}{5}, \frac{3}{4} \text{ in the second range.}$$

The corresponding constant pairs (W_0, W_1) are $\big(\frac{20}{63}, \frac{646}{1155}\big)$ and $\big(\frac{7}{20}, \frac{403}{700}\big)$. Thus every value in the first range is at most $\frac{10502}{17325}$, and every value in the second is at most $\frac{2}{3}$. The base bound $2 + 2n \max_i p_i$ gives $2 + \frac{16}{15e}$ and $2 + \frac{1}{e}$, respectively. Thus the total bounds are $\frac{10502}{17325e^2} + \frac{16}{15e} + 2 < R_*$ and $\frac{2}{3e^2} + \frac{1}{e} + 2 < R_*$. The displayed quadratic and endpoint calculations establish the two middle ranges using only rational arithmetic.

Step 5: Nearly Balanced Groups ($\frac{3}{4} \le r \le 1$).

Use $t_0 = t_1 = \frac{1}{2}$, identical uniform anchor and nominee widths $w = \frac{8e}{3(1+3e)}$, and gain $L = \frac{1-b}{wb}$, where $b = \frac{2e}{3}$. The continuous background integral is exactly b, so it is a valid finite-cell upper bound. Within the support, direct substitution in (B.7) gives

$$(1 + e - \mathcal{K}_g(z))(1 - z) = \frac{[(3 + 5e)z - 4e]^2}{24e} \geq 0.$$

Outside the support, use $1 + \frac{b}{2} + bz \leq 1 + e$. The base risk is at most $2 + \frac{7}{3w}$. For the extra risk, a focal $(1, 1)$ corner of the larger label has bound L, while that of the smaller label has bound $\frac{L(1+r)}{2r} \leq \frac{7}{6}L$; the remaining corner has zero extra coefficient. Thus robustness is at most $2 + \frac{7}{3w} + \frac{7}{6}L = \frac{21}{32e^2} + \frac{77}{32e} + \frac{53}{16} < R_*$. The final difference multiplied by $480e^2$ is $5 + 573e - 150e^2 > 0$. All population ratios and all $0 < e \leq \frac{1}{100}$ have now been covered. □

B.2 Facility-Pair Predictions

The full-profile construction uses predicted identities and cluster ranks. With only an optimal facility pair, we instead modify the direction of proportional sampling.

We analyze the directional proportional-sampling mechanism of Definition 4.

The coefficient depends on direction relative to the anchor, not on whether the partner lies beyond the predicted midpoint. Conditional on the anchor's report, both directional coefficients are fixed. The anchor's truthful cost is zero, and Lemma 3 handles every non-anchor, so the mechanism is strategyproof.

We now prove Theorem 4.

Proof. Fix actual optimal consecutive clusters A, B, of sizes m, s and costs P, Q. For anchor $i \in A$, set $U_i = \sum_{j\in A} d_{ij}$, $V_i = \sum_{j\in B} d_{ij}$, and $W_i = \sum_{j\in A} \omega_{ij} d_{ij}$, where $d_{ij} = |x_i - x_j|$. Every partner in B is on the same side of i, so its coefficient is a common $c_i \in \{\tau, 1\}$. Summing (B.2) over $j, \ell \in B$ bounds the cross-cluster sampling numerator by $3V_iQ$. If the partner is in A, charging every agent to the anchor costs $U_i + V_i$. Thus

$$\mathbb{E}[\mathrm{SC} \mid i] \leq U_i + \frac{W_iV_i + 3c_iV_iQ}{W_i + c_iV_i} \leq U_i + \frac{W_i}{c_i} + 3Q. \tag{B.18}$$

The case $V_i = 0$ follows directly.

With a correct prediction, assign agents to their nearest predicted optimal center and put midpoint ties in A. Every anchor favors the opposite cluster, hence $c_i = 1$. For each within-cluster unordered pair at distinct positions, the two directed coefficients sum to $1+\tau$. Consequently $\sum_{i\in A}(U_i+W_i) = (3 + \tau)\sum_{i<j\in A} d_{ij} \leq (3 + \tau)(m - 1)P$. The analogous inequality holds in B. Averaging (B.18), the coefficient of P is at most $\frac{(3+\tau)(m-1)+3s}{n} \leq 3 + \tau$, and likewise for Q. This proves consistency.

With an arbitrary prediction, take $i, j \in A$ with $x_i < x_j$. The report of j and every cross-cluster report lie to the right of i, so $\frac{\omega_{ij}}{c_i} = 1$. At j, the reverse ratio $\frac{\omega_{ji}}{c_j}$ is at most $\frac{1}{\tau}$. Including the two unweighted U terms, each unordered pair contributes at most $3 + \frac{1}{\tau}$ times its distance. The same argument in B is reflected. Use $\sum_{i<j\in A} d_{ij} \leq (m - 1)P$ and $3 \leq 3 + \frac{1}{\tau}$ in (B.18) to obtain robustness $3 + \frac{1}{\tau}$.

For tightness, let $X_m = \left(0^m, e, 1\right)$ with $e = m^{-2}$, so $\mathrm{OPT} = e$ and $\{0, 1\}$ is optimal. Conditional on anchors at 0 and 1, the social costs are $E_0 = \frac{e(2-e)}{1+e}$ and $E_1 = \frac{me(2-e)}{m+1-e}$. For the correct separator $\frac{1}{2}$, the cost conditional on the anchor at e is $E_e^+ = \frac{(1+\tau)me(1-e)}{1-e+\tau me}$. Hence $\frac{mE_0+E_e^++E_1}{(m+2)e} \to 3 + \tau$. For a wrong prediction whose midpoint is below 0, all anchors favor left, and the middle conditional cost becomes $E_e^- = \frac{(1+\tau)me(1-e)}{me+\tau(1-e)}$. The resulting ratio tends to $3 + \frac{1}{\tau}$. These limits establish both supremum equalities. □

B.3 A Lower Bound under Either Prediction Type

The preceding upper bounds leave open whether general randomized mechanisms can do better. The next result applies to all strategyproof mechanisms, not just proportional-sampling constructions. We adapt the two-profile comparison of Ma & Peng (2026), fixing an accurate prediction on one of the two profiles.

We give the full proof of Theorem 5.

Proof. Take integers $2 \le k < m$ and the labeled profiles $A = \big(0^m, \big(\frac{1}{2}\big)^k, 1^m\big)$ and $B = \big(-m, 0^m, \big(\frac{1}{2}\big)^{k-1}, 1^m\big)$, where a designated midpoint identity alone changes her report to $-m$. Their optimal costs are $\frac{k}{2}$ and m, respectively. Fix the full-profile prediction to B, or the facility-pair prediction to $\{-m, \frac{1}{2}\}$, which is optimal for B. Keep this same prediction at A. Let a, b denote the expected distance from the physical point $\frac{1}{2}$ to the two output laws. Strategyproofness of the designated agent gives $a \le b$.

For every pair of facilities, $d(0,Y) + d\big(\frac{1}{2},Y\big) + d(1,Y) \ge \frac{1}{2}$, since this is the optimal two-median cost of these three points. Likewise $d(-m,Y) + md(0,Y) + md(1,Y) \ge m$, the optimum for those three weighted locations. Set $t = \frac{k-1}{m-k}$. The first inequality and $tk + k - 1 = tm$ give $t\,\mathrm{SC}(A) + (k-1)a \ge \frac{tm}{2}$. The second gives $\mathrm{SC}(B) - (k-1)b \ge m$. Adding and using $a \le b$ yields $t\,\mathrm{SC}(A) + \mathrm{SC}(B) \ge m\big(1 + \frac{t}{2}\big)$. Consistency applies at B and robustness at A, so for this finite population

$$C \ge 1 + \frac{k-1}{2(m-k)} - \frac{Rk(k-1)}{2m(m-k)}. \tag{B.19}$$

For population-independent constants, let $\frac{k}{m} \to \alpha \in (0,1)$ to obtain $C \ge 1 + \frac{\alpha(1-R\alpha)}{2(1-\alpha)}$. Maximizing at $\alpha = 1 - \sqrt{1 - \frac{1}{R}}$ gives the stated bound, with the limiting interpretation when $R = 1$. Solving it for R at $C = 1 + \varepsilon$ gives the second assertion, and letting $\varepsilon \downarrow 0$ gives the last. For fixed n, only the finite inequality (B.19) is asserted. □

Together, Theorems 3 and 5 leave a linear-versus-quadratic gap near perfect consistency in the full-profile model. The facility-pair mechanism has limiting consistency 3 as its robustness increases, but this is not a lower bound of 3 for general mechanisms with facility-pair predictions.