A. Hardness proofs

A.1. Disagreement functions from proofs of Theorems 1 and 2

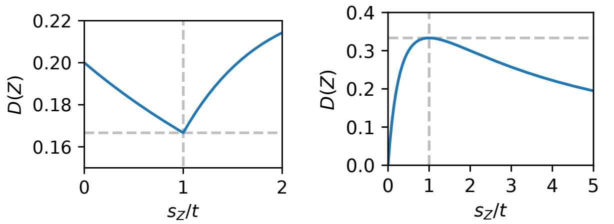

Figure 4. (Left) Plot of $D(Z) = \frac{t}{2t+s_Z} - \frac{3t}{5t+s_Z} + \frac{t}{2t+s_Z} - \frac{2t}{5t+s_Z}$ from the proof of Theorem 1. (Right) Plot of $D(Z) = \frac{2t}{2t+s_Z} - \frac{t/2}{t/2+s_Z}$ from the proof of Theorem 2. Both functions are re-parameterized in terms of the ratio s_Z/t by dividing through by t and achieve local optima at $s_Z/t = 1$ (i.e. $s_Z = t$); this can be verified analytically.

A.2. CDM PROMOTION is hard with $|A| = 2, |C| = 2$

In the main text, we show CDM PROMOTION is NP-hard when $|A| = 1, |C| = 3$ (Theorem 3). Here, we provide an additonal proof for the case when $|A| = 2, |C| = 2$. These are the smallest hard instances of the problem ($|A| = 1, |C| = 2$ is easy to solve: introduce alternatives that increase utility for x^* for than its competitor).

Theorem 7. *In the CDM model,* PROMOTION *is NP-hard, even with just two individuals and two items in C.*

Proof. By reduction from SUBSET SUM. Let S, t be an instance of SUBSET SUM. Let $A = \{a, b\}$, $C = \{x, y\}$, $\overline{C} = S$. Using tuples interpreted entrywise, construct a CDM with the following parameters.

$$
\begin{aligned}
u_a(\langle x^*, y\rangle) &= \langle t+\varepsilon, 0\rangle \\
u_b(\langle x^*, y\rangle) &= \langle \varepsilon, t\rangle \\
u_a(z) = u_b(z) &= -\infty && \forall z \in \overline{C} \\
p_a(z, \langle x^*, y\rangle) &= \langle 0, z\rangle && \forall z \in \overline{C} \\
p_b(z, \langle x^*, y\rangle) &= \langle z, 0\rangle && \forall z \in \overline{C}
\end{aligned}
$$

To promote x^*, we need to add more than $t - \varepsilon$ to b's utility for x^*, but add less than $t + \varepsilon$ to a's utility for x^*. Since all pulls are integral, the only solution is a set Z whose sum of pulls is t. If we could efficiently find such a set, then we could efficiently find the SUBSET SUM solution. $\square$

A.3. Proof of Theorem 4

Proof. By reduction from SUBSET SUM. Let $S = \{z_1, \ldots, z_n\}, t$ be an instance of SUBSET SUM. Let $A = \{a, b\}$, $C = \{x^*, y\}$, $\overline{C} = S$, and $0 < \varepsilon < 1$. The nest structures and utilities are shown in Fig. 5.

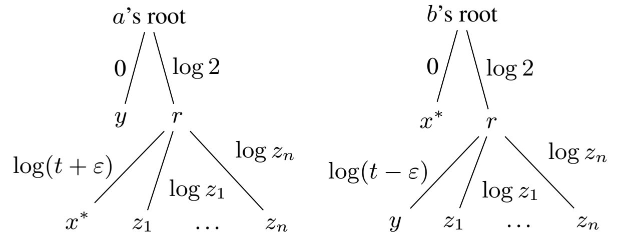

Figure 5. NL trees used in the proof of Theorem 4. The left tree is for individual a and the right tree for individual b.

Notice that x^* and y are swapped in the two trees. We wish to promote x^*. With just the choice set C, a prefers x^* to y, but b does not. To make b prefer x^* to y, we need to cannibalize y by adding z_i items. However, this simultaneously cannibalizes x^* in a's tree, so we need to be careful not to introduce too much additional utility. To ensure a prefers x^*, we need to pick Z such that

$$
\begin{aligned}
&\Pr(a \leftarrow y \mid C \cup Z) < \Pr(a \leftarrow y \mid C \cup Z) \\
&\iff \frac{1}{1+e^{\log 2}} < \frac{e^{\log 2}}{1+e^{\log 2}} \cdot \frac{e^{\log(t+\varepsilon)}}{e^{\log(t+\varepsilon)} + \sum_{z\in Z} e^{\log z}} \\
&\iff \frac{1}{3} < \frac{2}{3} \cdot \frac{t+\varepsilon}{t+\varepsilon+\sum_{z\in Z} z} \\
&\iff \sum_{z\in Z} z < t + \varepsilon.
\end{aligned}
$$

To ensure b prefers x^*, we need

$$
\begin{aligned}
&\Pr(b \leftarrow x^* \mid C \cup Z) > \Pr(b \leftarrow y \mid C \cup Z) \\
&\iff \frac{1}{1+e^{\log 2}} > \frac{e^{\log 2}}{1+e^{\log 2}} \cdot \frac{e^{\log(t-\varepsilon)}}{e^{\log(t-\varepsilon)} + \sum_{z\in Z} e^{\log z}} \\
&\iff \frac{1}{3} > \frac{2}{3} \cdot \frac{t-\varepsilon}{t-\varepsilon+\sum_{z\in Z} z} \\
&\iff \sum_{z\in Z} z > t - \varepsilon.
\end{aligned}
$$

Since the z are all integers, we must then have $\sum_{z\in Z} z = t$. If we could efficiently promote x^*, we could efficiently find such a Z. $\square$

A.4. Proof of Theorem 5

Proof. By reduction from SUBSET SUM. Let S, t be an instance of SUBSET SUM. Let $A = \{a, b\}$, $C = \{x^*, y\}$, $\overline{C} = S$, and $s = \sum_{z\in S} z$. Make aspects χ_z, ψ_z, γ_z for each $z \in S$ as well as two more aspects χ, ψ. The items have aspects as follows:

$$
\begin{aligned}
x^{*\prime} &= \{\chi\} \cup \{\chi_z \mid z \in S\} \\
y' &= \{\psi\} \cup \{\psi_z \mid z \in S\} \\
z' &= \{\chi_z, \psi_z, \gamma_z\} && \forall z \in S
\end{aligned}
$$

The individuals have the following utilities on aspects, where $0 < \varepsilon < 1$:

$$\begin{aligned} &u_a(\chi) = 0 && u_b(\chi) = s - t/2 + \varepsilon && \\ &u_a(\chi_z) = z && u_b(\chi_z) = 0 && \forall z \in S \\ &u_a(\psi) = s - t/2 - \varepsilon && u_b(\psi) = 0 && \\ &u_a(\psi_z) = 0 && u_b(\psi_z) = z && \forall z \in S \\ &u_a(\gamma_z) = s - z && u_b(\gamma_z) = s - z && \forall z \in S \end{aligned}$$

We want to promote x^*. Notice that x^* and y have disjoint aspects. Thus the choice probabilities from C are proportional to the sum of the item's aspects:

$$\begin{aligned} \Pr(a \leftarrow x^* \mid C) &\propto s \\ \Pr(a \leftarrow y \mid C) &\propto s - \frac{t}{2} - \varepsilon \\ \Pr(b \leftarrow x^* \mid C) &\propto s - \frac{t}{2} + \varepsilon \\ \Pr(b \leftarrow y \mid C) &\propto s. \end{aligned}$$

To promote x^*, we need to make b prefer x^* to y. Adding a z item cannibalizes from a's preference for x^* and b's preference for y. As in the NL proof, we want to add just enough z items to make b prefer x^* to y without making a prefer y to x^*.

First, notice that the γ_z aspects have no effect on the individuals' relative preference for x^* and y. If we introduce the alternative z, then if a picks the aspect χ_z, y will be eliminated. The remaining aspects of x^*, namely $x^{*\prime} \setminus \{\chi_z\}$, have combined utility $s - z$, as does γ_z. Therefore a will be equallly likely to pick x^* and z. Symmetric reasoning shows that if b chooses aspect ψ_z, then b will end up picking y with probability 1/2. This means that when we include alternatives $Z \subseteq \overline{C}$, each aspect χ_z, ψ_z for $z \in Z$ effectively contributes $z/2$ to a's utility for x^* and b's utility for y rather than the full z. The optimal solution is therefore a set Z of alternatives whose sum is t, since that will cause a to have effective utility $s - t/2$ on x^*, which exceeds its utility $s - t/2 - \varepsilon$ on y. Meanwhile, b's effective utility on y will also be $s - t/2$, which is smaller than its utility $s - t/2 + \varepsilon$ on x^*. If we include less alternative weight, b will prefer y. If we include more, a will prefer y. Therefore if we could efficiently find the optimal set of alternatives to promote x^*, we could efficiently find a subset of S with sum t. ☐

B. Restrictions on MNL that make AGREEMENT and DISAGREEMENT tractable

As we saw in the proofs of Theorems 1 and 2 that AGREEMENT and DISAGREEMENT are hard in the MNL model even when individuals have identical utilities on alternatives. This is possible because the individuals have different sums of utilities on C; one unit of utility on an alternative has a weaker effect for individuals with higher utility sums on C. To address the issue of identifiability, we assume each individual's utility sum over $\mathcal{U}$ is zero in this section. This allows us to meaningfully compare the sum of utilities of two different individuals.

Definition 1. If an individual a has $\sum_{x \in \mathcal{U}} u_a(x) = 0$, then the *stubbornness* of a is $\sigma_a = \sum_{x \in C} e^{u(x)}$.

We call this quantity "stubbornness" since it quantifies how reluctant an individual is to change its probabilities on C given a unit of utility on an alternative.

Proposition 1. *In an MNL model where all individuals are equally stubborn* and *have identical utilities on alternatives, the solution to* AGREEMENT *is* $\overline{C}$.

Proof. Assume utilities are in standard form, with $\sum_{x \in \mathcal{U}} u_a(x) = 0$. Let $\sigma = \sum_{x \in C} e^{u(x)}$ be each individual's stubborness and let Z be a set of alternatives. Suppose all individuals have the same utility $u(z)$ for each alternative z. The disagreement between two individuals about a single item x in C is then:

$$\frac{e^{u_a(x)}}{\sigma + \sum_{z \in Z} e^{u(z)}} - \frac{e^{u_b(x)}}{\sigma + \sum_{z \in Z} e^{u(z)}} = \frac{|e^{u_a(x)} - e^{u_b(x)}|}{\sigma + \sum_{z \in Z} e^{u(z)}}.$$

Notice that this strictly decreases if $\sum_{z \in Z} e^{u(z)}$ increases, so we minimize D by including all of the alternatives. ☐

The same reasoning also allows us to solve DISAGREEMENT in this restricted MNL model.

Corollary 3. *The solution to* DISAGREEMENT *in an equal alternative utilities, equal stubbornness MNL model is* $\emptyset$.

C. Approximation algorithm details and extensions

C.1. Proof of Lemma 2

Proof. If a set Z has total exp-utility t_a to individual a, then it is placed in L at position $\lfloor \log_{1+\delta} t_a \rfloor$ in dimension a. So, if two sets Z, Z' with exp-utility totals t_a, t'_a for individual a are mapped to the same cell of L, then for all $a \in A$, $\lfloor \log_{1+\delta} t_a \rfloor = \lfloor \log_{1+\delta} t'_a \rfloor$. We can therefore bound t'_a:

$$\log_{1+\delta} t_a - 1 < \log_{1+\delta} t'_a < \log_{1+\delta} t_a + 1.$$

Exponentiating both sides with base $1 + \delta$ and simplifying yields

$$\frac{t_a}{1+\delta} < t'_a < t_a(1+\delta). \tag{5}$$

With this fact in hand, we proceed by induction on i. When $i = 0$, $\overline{C}_i$ is empty and the lemma holds. Now suppose that

$i > 0$ and that the lemma holds for $i - 1$. Every set in $\overline{C}_i$ was made by adding (a) zero elements or (b) one element to a set in $\overline{C}_{i-1}$. We consider these two cases separately.

(a) For any set $Z \subseteq \overline{C}_i$ that is also contained in $\overline{C}_{i-1}$, we know by the inductive hypothesis that some element in L_{i-1} satisfied the inequality. Since we never overwrite cells, the lemma also holds for Z after iteration i.

(b) Now consider sets $Z' \subseteq \overline{C}_i$ that were formed by adding the new element z to a set $Z \subseteq \overline{C}_{i-1}$. In the inner for loop, we at some point encountered the cell containing the set $Y \in L_{i-1}$ satisfying the lemma for set Z by the inductive hypothesis. Let y_a be the exp-utility totals for Y and t_a for Z. Notice that the exp-utility totals of Z' are exactly $t_a + e_{az}$. Starting with the inductive hypothesis, we see that the exp-utility totals of $Y \cup \{z\}$ satisfy

$$\frac{t_a + e_{az}}{(1+\delta)^{i-1}} < y_a + e_{az} < (t_a + e_{az})(1+\delta)^{i-1}.$$

When we go to place $Y \cup \{z\}$ in a cell, it might be unoccupied, in which case we place it in L_i and the lemma holds for Z'. If it is occupied by some other set, then by applying Eq. (5) we find that the lemma also holds for Z'. $\square$

C.2. Polynomial bound on runtime of Algorithm 1

The runtime of Algorithm 1 is $O((m + kn^2)(1 + \lfloor \log_{1+\delta} s \rfloor)^n)$. We can show that the second part is bounded by a polynomial in k, m, and $\frac{1}{\varepsilon}$:

$$\begin{aligned}
(1 + \lfloor \log_{1+\delta} s \rfloor)^n &\leq \left(1 + \frac{\ln s}{\ln 1+\delta}\right)^n \\
&\leq \left(1 + (1+\delta)\frac{\ln s}{\delta}\right)^n \\
&\quad \text{(since } \ln(1+x) \geq \tfrac{x}{1+x} \text{ for } x > -1) \\
&= \left(1 + \frac{\ln s}{\delta} + \ln s\right)^n \\
&= \left(1 + \frac{2km\binom{n}{2}\ln s}{\varepsilon} + \ln s\right)^n
\end{aligned}$$

C.3. Adapting Algorithm 1 for CDM with guarantees for special cases

We can adapt Algorithm 1 for the CDM model, but we only maintain the approximation error bounds under special cases of the structure of the "pulls". Still, we can use this algorithm as a principled heuristic and it tends to work well in practice, as we saw in Fig. 2.

As a first step, we use the alternative parametrization of the model used by Seshadri et al. (2019, Eq. (1)), which has fewer parameters. In this description of the model, utilities and context effects are merged into a single utility-adjusted pull $q_a(z, x) = p_a(z, x) - u_a(x)$, with the special case $q_a(x, x) = 0$. We then have

$$\Pr(a \leftarrow x \mid C) = \frac{\exp(\sum_{w \in C} q_a(w, x))}{\sum_{y \in C} \exp(\sum_{z \in C} q_a(z, y))}. \tag{6}$$

Refer to Seshadri et al. (2019, Appendix C.1) for details of the equivalence between this formulation and the one we use in the main text.

Matching the notation of the proof of Theorem 6, we use the shorthand $e_{ax} = \exp(\sum_{w \in C} q_a(w, x))$.

To adapt Algorithm 1 to the CDM, we expand L_i to have nk dimensions for each individual-item pair, increasing the runtime to $O((m + kn^2)(1 + \lfloor \log_{1+\delta} s \rfloor)^{nk})$. This is only practical if nk is small, but as we have seen, AGREEMENT, DISAGREEMENT, and PROMOTION are all NP-hard even with $n = 2$ and $k = 2$ or 3. Each individual-item dimension stores e_{ax}, the total exp-utility of that item to that individual given that we have included some set of alternatives. When we include an additional item from $\overline{C}$, we place the new sets in L_i with updated e_{ax} values.

This only preserves the ε-additive approximation if alternatives (items in $\overline{C}$) have zero context effects on each other; however, they may still have context effects on items in C. Formally, we need $q_a(z, z') = 0$ for all $z, z' \in \overline{C}$ and $a \in A$. Although this is a serious restriction, it leaves AGREEMENT, DISAGREEMENT, and PROMOTION NP-hard, as the CDM we used in our proofs had this form (see also Appendix C.5 for how to apply Algorithm 1 to PROMOTION). If this version of the algorithm is applied to a general CDM, it may experience higher error. Nonetheless, our real-data experiments show it to be a good heuristic.

For the following analysis, we assume a CDM with zero context effects between items in $\overline{C}$. We need to verify that if every item's exp-utility is approximated to within factor $(1 + \beta)^{\pm 1}$, then the total disagreement of a set is approximated to within ε as we had in the MNL case. The approximation error guarantee increases to 4ε in the restricted CDM version—to recover the ε-additive approximation, we can make δ smaller by a factor of 4 (that is, we could pick $\delta = \varepsilon/(8km\binom{n}{2})$); we instead keep the old δ for simplicity in the following analysis).

Recall that Z' is the representative in L_m of the optimal set of alternatives Z^*. For compactness, we define T_a to be the denominator of Eq. (6), with T'_a and T^*_a referring to those denominators under the choice sets $C \cup Z'$ and $C \cup Z^*$, respectively. This is where we require zero context effects between alternatives: if alternatives interact, then storing every e_{ax} in the table (from which we can compute T_a) is not enough to determine updated choice probabilities when we add a new alternative.

The difference in the analysis begins when we bound $\Pr(a \leftarrow x \mid C \cup Z')$ on both sides using the fact that each exp-utility sum is approximated within a $1+\beta$ factor (so the probability denomiators T_a are also approximated within this factor):

$$\begin{aligned}\frac{\frac{e^*_{ax}}{1+\beta}}{T^*_a(1+\beta)} &= \frac{1}{(1+\beta)^2}\frac{e^*_{ax}}{T^*_a}\\ &< \frac{e'_{ax}}{T'_a} = \Pr(a \leftarrow x \mid C \cup Z')\\ &< \frac{e^*_{ax}(1+\beta)}{\frac{T^*_a}{1+\beta}} = (1+\beta)^2\frac{e^*_{ax}}{T^*_a}.\end{aligned}$$

Based on the lower bound, the difference between $\Pr(a \leftarrow x \mid C \cup Z^*)$ and $\Pr(a \leftarrow x \mid C \cup Z')$ could be as large as

$$\frac{e^*_{ax}}{T^*_a} - \frac{1}{(1+\beta)^2}\frac{e^*_{ax}}{T^*_a} \le 1 - \frac{1}{(1+\beta)^2}.$$

Now considering the upper bound, the difference between $\Pr(a \leftarrow x \mid C \cup Z^*)$ and $\Pr(a \leftarrow x \mid C \cup Z')$ could be as large as

$$(1+\beta)^2\frac{e^*_{ax}}{T^*_a} - \frac{e^*_{ax}}{T^*_a} \le (1+\beta)^2 - 1.$$

Therefore, $|\Pr(a \leftarrow x \mid C \cup Z') - \Pr(b \leftarrow x \mid C \cup Z')|$ can only exceed $|\Pr(a \leftarrow x \mid C \cup Z^*) - \Pr(b \leftarrow x \mid C \cup Z^*)|$ by at most $1 - \frac{1}{(1+\beta)^2} + (1+\beta)^2 - 1 = (1+\beta)^2 - \frac{1}{(1+\beta)^2}$. This is at most 4β:

$$\begin{aligned}4\beta - (1+\beta)^2 + \frac{1}{(1+\beta)^2} &= \frac{\beta^2(2-\beta^2)}{(1+\beta)^2}\\ &> 0. \qquad \text{(for } 0 < \beta < \sqrt{2}\text{)}\end{aligned}$$

So $D(Z')$ and $D(Z^*)$ are within $4\beta\binom{n}{2}k = 4\varepsilon$.

C.4. Adapting Algorithm 1 for NL with full guarantees

We can also adapt Algorithm 1 for the NL model, and unlike the CDM, the ε-additive approximation holds in all parameter regimes. Recall that the NL tree has two types of leaves: choice set items and alternative items. Let P_a be the set of internal nodes of individual a's tree that have at least one alternative item as a child and let $p = \max_{a \in A} |P_a|$. If we know the total exp-utility that alternatives contribute as children of each $v \in P_a$, then we can compute a's choice probabilites over items in C in polynomial time.

With this in mind, we modify Algorithm 1 by having dimensions in L for each individual for each of their nodes in P_a. This results in $\le np$ dimensions. The algorithm then keeps track of the exp-utility sums from alternatives under each node in P_a for each individual. The exponent in the runtime increases to (at most) np, but this remains tractable for some hard instances, such as those in our hardness proofs. In some cases, we can dramatically improve the runtime of the algorithm: if the subtree under an internal node contains only alternatives as leaves in an individuals's tree, then we only need one dimension L for that individual's entire subtree, and it has only two cells: one for sets that contain at least one alternative in that subtree, and one for sets that do not. The only factor that affects the choice probabilities of items in C is whether that subtree is "active" and its root can be chosen.

We now show how the error from exp-utility sums of alternatives propagates to choice probabilities. In the NL model, $\Pr(a \leftarrow x \mid C)$ is the product of probabilities that a chooses each ancestor of x as a descends down its tree. Let $v_1, \dots, v_\ell$ be the nodes in a's tree along the path from the root to x. For compactness, we use $\Pr(x, Z)$ instead of $\Pr(a \leftarrow x \mid C \cup Z)$ in the following analysis.

Pick $\delta \le ([\varepsilon/(2k\binom{n}{2}) + 1]^{1/\ell} - 1)/m$ and recall that $\beta = 2m\delta$. We can use the same analysis as in the proof of Theorem 6 to find that for any set $Z^* \subseteq \overline{C}$, there exists some $Z' \in L$ such that

$$\begin{aligned}\Pr(x, Z^*) &= \Pr(v_1, Z^*) \cdot \dots \cdot \Pr(v_x, Z^*)\\ &< \left(\Pr(v_1, Z') + \frac{\beta}{2}\right) \cdot \dots \cdot \left(\Pr(v_x, Z') + \frac{\beta}{2}\right)\\ &\le \Pr(x, Z') + \left(1 + \frac{\beta}{2}\right)^\ell - 1\\ &\le \Pr(x, Z') + \frac{\varepsilon}{2k\binom{n}{2}}.\end{aligned}$$

Now for the lower bound, pick $\delta \le (1 - [1 - \varepsilon/(2k\binom{n}{2})]^{1/\ell})/m$. Again from the proof of Theorem 6:

$$\begin{aligned}\Pr(x, Z^*) &= \Pr(v_1, Z^*) \cdot \dots \cdot \Pr(v_x, Z^*)\\ &> \left(\Pr(v_1, Z') - \frac{\beta}{2}\right) \cdot \dots \cdot \left(\Pr(v_x, Z') - \frac{\beta}{2}\right)\\ &\ge \Pr(x, Z') + \left(1 - \frac{\beta}{2}\right)^\ell - 1\\ &\ge \Pr(x, Z') - \frac{\varepsilon}{2k\binom{n}{2}}.\end{aligned}$$

Let h be the maximum height of any indivdual's NL tree (so $\ell \le h$). Then, by picking $\delta = \min\{[\varepsilon/(2k\binom{n}{2}) + 1]^{1/h} - 1, 1 - [1 - \varepsilon/(2k\binom{n}{2})]^{1/h}\}/m$, we find that $\Pr(a \leftarrow x \mid C \cup Z^*)$ and $\Pr(a \leftarrow x \mid C \cup Z')$ differ by less than $\varepsilon/(k\binom{n}{2})$ for all $x \in C$ and $a \in A$, meaning that the total disagreement between a and b cannot differ by more than ε as before.

Unfortunately, this means we need to make δ exponentially (in h) smaller in the NL model. Put another way, our error bound gets exponentially worse as h increases if we keep δ constant. However, we have seen that there are NP-hard families of NL instances in which h is a small constant (e.g., $h = 2$ in our hardness proof), so once again this algorithm is an exponential improvement over brute force. Moreover, the error bound here is often far from tight, since we use the very loose bounds $\Pr(v_i, Z') \le 1$ in the analysis. This means the algorithm will tend to outperform the worst-case guarantee by a significant margin.

C.5. Adapting Algorithm 1 for PROMOTION

C.5.1. CDM PROMOTION WITH SPECIAL CASE GUARANTEES

Algorithm 1 can be applied to PROMOTION in the (restricted) CDM model with only a small modification to the CDM version described in Appendix C.3: at the end of the algorithm, we return the set that results in the maximum number of individuals having x^* as an ε-favorite item. Additionally, we choose $\delta = \varepsilon/(10m)$ (we don't need the factors $\binom{n}{2}$ or k since we aren't optimizing $D(Z)$).

Following the analysis in Appendix C.3 (with $\beta = 2m\delta = \varepsilon/5$), we find that $\Pr(a \leftarrow x \mid C \cup Z^*)$ and $\Pr(a \leftarrow x \mid C \cup Z')$ differ by at most $\max\{1 - \frac{1}{(1+\varepsilon/5)^2}, (1+\varepsilon/5)^2 - 1\}$ for all x. On the interval $[0, 1]$, this is bounded by $\varepsilon/2$. Thus, if x^* is the favorite item for a given the optimal choice set $C \cup Z^*$, then it must be an ε-favorite of individual a given $C \cup Z'$ (as always, Z' is the representative of Z^* in L_m). This is because when we go from $C \cup Z^*$ to $C \cup Z'$, the choice probability of x^* can shrink by at most $\varepsilon/2$ and the choice probability for any other item can grow by at most $\varepsilon/2$. Thus, including Z' makes at least as many individuals have x^* as an *ε-favorite* item as including Z^* makes have x^* as a *favorite* item.

This is exactly what it means for Algorithm 1 to ε-approximate PROMOTION in the CDM (when items in $\overline{C}$ do not exert context effects on each other). Moreover, not having to compute $D(Z)$ makes the runtime of Algorithm 1 $O(m(1 + \lfloor \log_{1+\delta} s \rfloor)^{nk})$ when applied to PROMOTION in the CDM. In the general CDM, this algorithm is only a heuristic.

C.5.2. NL PROMOTION WITH FULL GUARANTEES

A very similar idea allows us to apply the NL version of Algorithm 1 from Appendix C.4 to PROMOTION and retain an approximation guarantee. As before, use the NL version and return the set that results in the maximum number of individuals having x^* as an ε-favorite item. However, we instead use $\delta = \min\{(\varepsilon/4+1)^{1/h} - 1, 1 - (1-\varepsilon/4)^{1/h}\}/m$, which by the analysis in Appendix C.4 results in $\Pr(a \leftarrow x \mid C \cup Z^*)$ and $\Pr(a \leftarrow x \mid C \cup Z')$ differing by at most $\varepsilon/2$. As in the CDM case, this guarantees that if x^* is the favorite item for a given the optimal choice set $C \cup Z^*$, then it must be an ε-favorite of a given $C \cup Z'$. Therefore this version of Algorithm 1 ε-approximates PROMOTION in the NL model with runtime $O(m(1 + \lfloor \log_{1+\delta} s \rfloor)^{np})$.

D. Mixed-integer bilinear programs for MNL agreement and disagreement optimization

D.1. AGREEMENT

Let x_i be a decision variable indicating whether we add in the ith item in $\overline{C}$. Let $e_{ya} = e^{u_a(y)}$ and $e_{Ca} = \sum_{y \in C} e_{ya}$. We can write AGREEMENT as the following 0-1 optimization problem.

$$\begin{aligned} \min_{x} \quad & \sum_{a,b \in A} \sum_{y \in C} \left| \frac{e_{ya}}{e_{Ca} + \sum_{i \in \overline{C}} x_i e_{ia}} - \frac{e_{yb}}{e_{Cb} + \sum_{i \in \overline{C}} x_i e_{ib}} \right| \\ \text{s.t.} \quad & x_i \in \{0, 1\} \end{aligned}$$

We can rewrite this with no absolute values by introducing new variables δ_{yab} that represent the absolute disagreement about item y between individuals a and b. We then use the standard trick for minimizing an absolute value in linear programs:

$$\begin{aligned} \min_{x} \quad & \sum_{a,b \in A} \sum_{y \in C} \delta_{yab} \\ \text{s.t.} \quad & \frac{e_{ya}}{e_{Ca} + \sum_{i \in \overline{C}} x_i e_{ia}} - \frac{e_{yb}}{e_{Cb} + \sum_{i \in \overline{C}} x_i e_{ib}} \le \delta_{yab} \\ & \qquad \forall y \in C, \{a, b\} \subset A \qquad , \\ & \frac{e_{yb}}{e_{Cb} + \sum_{i \in \overline{C}} x_i e_{ib}} - \frac{e_{ya}}{e_{Ca} + \sum_{i \in \overline{C}} x_i e_{ia}} \le \delta_{yab} \\ & \qquad \forall y \in C, \{a, b\} \subset A \qquad , \\ & x_i \in \{0, 1\} \quad \forall i \in \overline{C}, \\ & \delta_{yab} \in \mathbb{R} \quad \forall y \in C, \{a, b\} \subset A \end{aligned}$$

To get rid of the fractions, we introduce the new variables $z_a = \frac{1}{e_{Ca} + \sum_i x_i e_{ia}}$ for each individual a and add corresponding constraints enforcing the definition of z_a:

$$\begin{aligned} \min_{x} \quad & \sum_{a,b \in A} \sum_{y \in C} \delta_{yab} \\ \text{s.t.} \quad & \\ z_a e_{ya} - z_b e_{yb} &\le \delta_{yab} \quad && \forall y \in C, \{a, b\} \subset A, \\ z_b e_{yb} - z_a e_{ya} &\le \delta_{yab} \quad && \forall y \in C, \{a, b\} \subset A, \\ z_a e_{Ca} + z_a \sum_{i \in \overline{C}} x_i e_{ia} &= 1 && \forall a \in A, \\ x_i &\in \{0, 1\} && \forall i \in \overline{C}, \\ \delta_{yab} &\in \mathbb{R} && \forall y \in C, \{a, b\} \subset A, \\ z_a &\in \mathbb{R} && \forall a \in A \end{aligned}$$

This is a mixed-integer bilinear program (MIBLP) with m binary variables, $n + k\binom{n}{2}$ real variables, $2k\binom{n}{2}$ linear constraints, and n bilinear constraints. We plug this form of the problem directly into a branch-and-bound solver (we use Gurobi).

D.2. DISAGREEMENT

A similar technique works for DISAGREEMENT, but maximizing an absolute value is slightly trickier than minimizing. In addition to the variables δ_{yab} that we used before, we also add new binary variables g_{yab} indicating whether each difference in choice probabilities is positive or negative. With these new variables (and following the same steps as above), DISAGREEMENT can be written as the following MIBLP:

$$\max_{x} \quad \sum_{a,b \in A} \sum_{y \in C} \delta_{yab}$$

$$\text{s.t.}$$

$$\begin{aligned}
z_a e_{ya} - z_b e_{yb} &\leq \delta_{yab} & \forall y \in C, \{a,b\} \subset A,\\
z_b e_{yb} - z_a e_{ya} &\leq \delta_{yab} & \forall y \in C, \{a,b\} \subset A,\\
2g_{yab} + z_a e_{ya} - z_b e_{yb} &\geq \delta_{yab} & \forall y \in C, \{a,b\} \subset A,\\
2(1-g_{yab}) + z_b e_{yb} - z_a e_{ya} &\geq \delta_{yab} & \forall y \in C, \{a,b\} \subset A,\\
z_a e_{Ca} + z_a \sum_{i \in \overline{C}} x_i e_{ia} &= 1 & \forall a \in A,\\
x_i &\in \{0,1\} & \forall i \in \overline{C},\\
g_{yab} &\in \{0,1\} & \forall y \in C, \{a,b\} \subset A,\\
\delta_{yab} &\in \mathbb{R} & \forall y \in C, \{a,b\} \subset A,\\
z_a &\in \mathbb{R} & \forall a \in A
\end{aligned}$$

E. Additional experiment details

E.1. Simple example of poor performance for Greedy

As we saw in experimental data, Greedy can perform poorly even in small instances of AGREEMENT. Below we provide an MNL instance with $n = m = k = 2$ for which the error of the greedy solution is approximately 1. With only two individuals, $0 \leq D(Z) \leq 2$, so an error of 1 is very large.

In the bad instance for greedy, $A = \{a, b\}$, $C = \{x, y\}$, $\overline{C} = \{p, q\}$, and the utilities are as follows.

$$\begin{aligned}
u_a(x) &= 8 & u_b(x) &= 8\\
u_a(y) &= 2 & u_b(y) &= 8\\
u_a(p) &= 10 & u_b(p) &= 0\\
u_a(q) &= 0 & u_b(q) &= 15
\end{aligned}$$

In this instance of AGREEMENT, the greedy solution is $D(\emptyset) \approx 0.9951$ (including either p or q alone increases disagreement), while the optimal solution is $D(\{p, q\}) \approx 0.0009$.

E.2. All-pairs agreement results for MIBLP

Figure 6 shows the comparison in performance between Algorithm 1 and the MIBLP approach for the all-pairs AGREEMENT and DISAGREEMENT experiment. The methods perform nearly identically on both ALLSTATE and YOOCHOOSE. The MIBLP approach performs marginally better in some cases of YOOCHOOSE AGREEMENT. As noted in the paper, the MIBLP heuristic is considerably faster (12x and 240x on YOOCHOOSE and ALLSTATE, respectively), but provides no a priori performance guarantee and cannot be applied to CDM or NL. Nonetheless, we can see that it performs very competitively and would be a good approach to use in practice for MNL AGREEMENT and DISAGREEMENT.

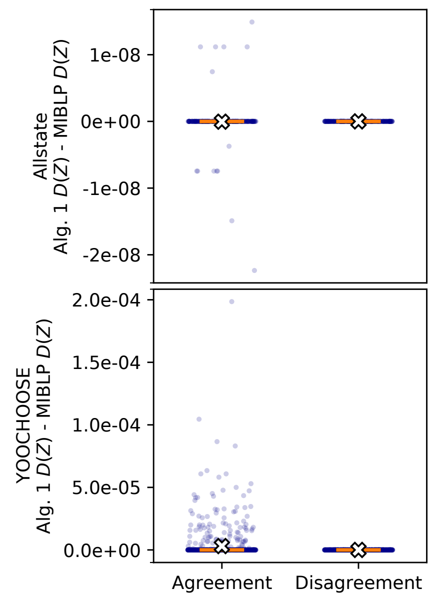

Figure 6. MIBLP vs. Algorithm 1 performance box plots when applied to all 2-item choice sets in ALLSTATE and YOOCHOOSE under MNL. Each point is the difference in $D(Z)$ when MIBLP and Algorithm 1 are run on a choice set, and Xs mark means.

E.3. Choice sets sampled from data

We repeated the all-pairs agreement experiment with 500 choice sets of size up to 5 sampled uniformly from each dataset, allowing us to evaluate the performance of Algorithm 1 on realistic choice sets. We limited the size of sampled choice sets since the CDM version of Algorithm 1 scales poorly with $|C|$ (see Appendix C.3). For this version of the experiment, we fixed larger values of ε (2 for MNL, 500 for CDM) to handle larger choice sets and to keep running time down. Again, Algorithm 1 has better mean performance in every case (Fig. 7), showing that it performs well on real choice sets.

F. A note on ethical considerations

Influencing the preferences of decision-makers has the potential for malicious applications, so it is important to address the ethical context of this work.

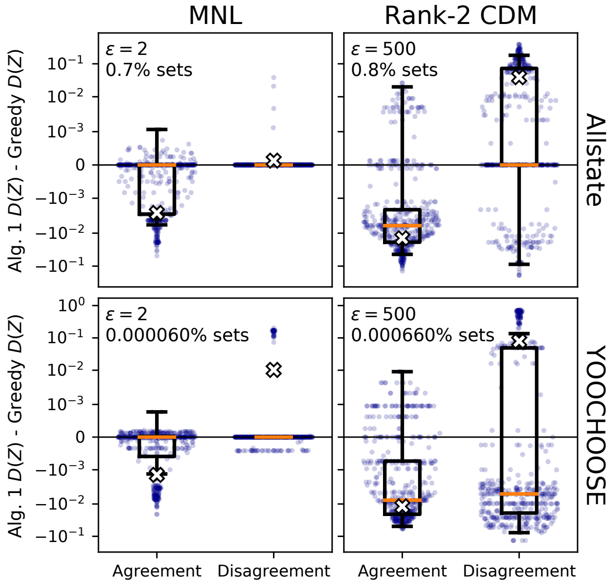

Figure 7. Results of the agreement experiment with 500 choice sets sampled uniformly from each dataset. Compare with Fig. 2 in the main text. Again, Algorithm 1 has better mean performance in all cases. The larger values of ε result in slightly worse performance on the margins than in Fig. 2, but also fewer sets computed.

Any problem with positive social applications (e.g., AGREEMENT: encouraging consensus, PROMOTION: promoting environmentally-friendly transportation options, DISAGREEMENT: increasing diversity of opinions) has the potential to be used for ill. This should not prevent us from seeking methods to acheive these positive ends, but we should certainly be cognizant of the possibility of unintended applications. In a different vein, understanding when a group is susceptible to undesired interventions (or detecting such interventions) makes problems like DISAGREEMENT worth studying from an adversarial perspective. Along these lines, our hardness results are encouraging since optimal malicious interventions are difficult.

Finally, we note that all of the theoretical problems we study presuppose access to choice data from which preferences can be learned and the ability to influence choice sets. Any entity which has both of these (such as an online retailer, a government deciding transportation policy, etc.) already has significant power to influence choosers. If such an entity had malicious intent, then near-optimal DISAGREEMENT solutions would be the least of our concerns.

To summarize, these problems are worth studying because of (1) their purely theoretical value in furthering the field of discrete choice, (2) their potential for positive applications, (3) insight into the potential for harmful manipulation by an adversary, and (4) the minimal additional risk from undesired use of our methods.