Supplemental Information (SI)

Combining Outcome-Based and Preference-Based Matching: A Constrained Priority Mechanism

Avidit Acharya, Kirk Bansak, Jens Hainmueller

A Properties of the Mechanism

A.1 Proof that it is Constrained Efficient

Suppose that φ is not $\bar{g}$-constrained efficient, so that for some preference profile $\succsim$, $\varphi(\succsim)$ is Pareto-dominated by a feasible $\bar{g}$-acceptable matching μ.

For all families i, let $M_i = \{j < i : j \notin N_i\}$ be the families ahead of i that were already assigned a location under $\varphi(\succsim)$, and let $\underline{i} = \min\{i : \mu(i) \succ_i \varphi(\succsim)(i)\}$ be the first family to which μ assigns it a location that it strictly prefers to the one it gets under $\varphi(\succsim)$. (Such a family must exist if μ Pareto-dominates $\varphi(\succsim)$.) By construction $\mu(i) = \varphi(\succsim)(i)$ for all $i \in M_{\underline{i}}$. So for μ to be feasible and $\bar{g}$-acceptable, it must be that $\mu(\underline{i}) \in S_{\underline{i}} \cap L^{\bar{g}}_{\underline{i}}(\alpha_{\underline{i}})$, where $\alpha_{\underline{i}}$ is the completed assignment under $\varphi(\succsim)$ at Step $\underline{i}$. This means that $S_{\underline{i}} \cap L^{\bar{g}}_{\underline{i}}(\alpha_{\underline{i}}) \neq \emptyset$ so $\varphi(\succsim)$ must have assigned the best location $l^*_{\underline{i}}$ in this set to family $\underline{i}$. But since $\mu(\underline{i}) \succ_i \varphi(\succsim)(\underline{i}) = l^*_{\underline{i}}$, this contradicts the assumption that $l^*_{\underline{i}}$ is the best location for $\underline{i}$ in $S_{\underline{i}} \cap L^{\bar{g}}_{\underline{i}}(\alpha_{\underline{i}})$.

A.2 Proof that it is Strategy-proof

Suppose that there is some i for whom reporting a different preference $\succsim'_i$ produces a strictly better location assignment: $\varphi(\succsim'_i, \succsim_{-i}) \succ_i \varphi(\succsim)(i)$.

Let $l'_i = \varphi(\succsim'_i, \succsim_{-i})$ and note that $S_j \cap L^{\bar{g}}_j(\alpha_j)$ is independent of i's reported preference for all $j < i$. Therefore, $N_i = N'_i$ where N_i is the set of families on hold at Step i under the truthfully reported profile $\succsim$ and N'_i are those on hold at Step i under the profile $(\succsim'_i, \succsim_{-i})$. In addition, $\varphi(\succsim'_i, \succsim_{-i})(j) = \varphi(\succsim)(j)$ for all $j \in N_i$. This implies that $\alpha'_i = \alpha_i$, where α'_i is the completed assignment at Step i under preference profile $(\succsim'_i, \succsim_{-i})$ and α_i is the completed assignment at Step i under preference profile $\succsim$. Therefore, $L^{\bar{g}}_i(\alpha_i) = L^{\bar{g}}_i(\alpha'_i) =: L^{\bar{g}}_i$.

Let S'_i be the locations that i ranks strictly under $\succsim'_i$ and S_i the locations that i ranks strictly under $\succsim_i$. If $S_i \cap L^{\bar{g}}_i = \emptyset$, then all of the locations in $L^{\bar{g}}_i$ are ones that i ranks worst, and i is guaranteed to be assigned one of these locations regardless of which location i reports. Therefore it cannot be that $\varphi(\succsim'_i, \succsim_{-i}) \succ_i \varphi(\succsim)(i)$.

On the other hand, if $S_i \cap L^{\bar{g}}_i \neq \emptyset$ then $\varphi(\succsim'_i, \succsim_{-i}) \succ_i \varphi(\succsim)(i)$ and $L^{\bar{g}}_i(\alpha_i) = L^{\bar{g}}_i(\alpha'_i)$ implies that $l'_i \in S_i \cap L_i(\alpha_i)$. But then $l'_i \succ_i \varphi(\succsim)(i) = l^*_i$ contradicts the fact that l^*_i is the unique best location in $S_i \cap L_i(\alpha_i)$ under preference $\succsim_i$.

B Verifying $\bar{g}$-Acceptability

As described in the main text, implementing the $\bar{g}$-constrained priority mechanism involves iteratively verifying that the next assignment of a family to a particular location can be performed without compromising the possibility of a $\bar{g}$-acceptable final matching.

This process requires solving the maximization problem in Equation 2 of the main text:

$$G_i(q^i) := \max_{\beta_i} \sum_{j \in \{i,\ldots,n\} \cup N_i} g_j(\beta_i(j)) \text{ subject to } |\beta_i^{-1}(l)| \leq q_l^i, \forall l \tag{2}$$

This involves computing the maximum possible total outcome score for any remaining set of units and the remaining location capacities.

In implementing the mechanism, Equation 2 can be solved by employing a standard linear sum assignment problem (LSAP) (Burkard et al., 2012). Specifically, the LSAP formulation is applied to an augmented cost matrix, whereby the rows correspond to the remaining units and the columns correspond to location capacity slots (i.e. each column is replicated according to the number of capacity slots belonging to the associated location). Each element $[i, v]$ of the cost matrix corresponds to the complement of the outcome score for the ith unit when assigned to the location to which the vth column pertains.

Various polynomial-time algorithms have been developed for solving the LSAP, beginning with the introduction of the Hungarian algorithm in the 1950s (Kuhn, 1955, Munkres, 1957). We employ the RELAX-IV cost flow solver developed by Bertsekas and Tseng (Bertsekas and Tseng, 1994) and implemented in `R` by the `optmatch` package (Hansen and Klopfer, 2006).

C Simulation Application: Additional Details

The follow describes the data-generating process employed in the simulations.

First a number N is chosen, denoting the number of agents. For simplicity, the same number of locations is also used, each with capacity for one agent. In addition, ρ_p and ρ_{op} are both chosen, denoting the pre-specified correlation between preferences across agents and the correlation between preferences and outcome scores within agents.

Next, N different N-dimensional latent variable vectors are generated, and these vectors are column-bound into an N x N matrix, which we denote by $\mathbf{P}$, representing a simulated preference matrix. Specifically, each vector is a multivariate normal random vector, using a mean vector of 0, and a covariance matrix with 1 for all the diagonal elements and ρ_p for all the off-diagonal elements. Let $\vec{z}_l$ denote the lth N-dimensional latent variable vector, which pertains to the lth location and comprises the lth column of $\mathbf{P}$. For any given vector, the ith element pertains to the ith family.

By generating the N x N matrix $\mathbf{P}$ in this way, each row represents a client and each column represents a location. Thus, the ith row, $\mathbf{P}$[i,], denotes a latent preference vector for agent i, with higher (more positive) values corresponding to a higher preference and vice versa. By construction, for any two cleints (rows), the pairwise correlation between the two vectors will be ρ_p in expectation, imposing a correlation of ρ_p across agents' preferences over locations.

Let $\vec{s}_i$ denote the ith cleint's outcome score vector. The outcome score vectors are constructed such that $\vec{s}_i = sign(\rho_{op}) \cdot (\mathbf{P}[i,] + \vec{\epsilon})$, where the elements of $\vec{\epsilon}$ are independently distributed normal with mean 0 and variance σ_ϵ^2. The value of σ_ϵ^2 is determined such that it, in combination with the $sign(\rho_{op})$ operator, produces an expected pairwise correlation of ρ_{op} between $\vec{s}_i$ and $P[i,]$, thereby inducing the correlation of ρ_{op} between a agent's preferences and outcome scores. The outcome score vectors are then row-bound to create an N x N outcome score matrix $\mathbf{S}$, where each row represents a agent and each column represents a location.

In applying our mechanism to the simulated data, the $\mathbf{S}$ matrix is first normalized such that its elements are in the interval $[0, 1]$, and the $\mathbf{P}$ matrix is mapped to preference ranks (i.e. each row $\mathbf{P}[i,]$ is transformed into ranks such that the most positive value becomes 1 and the most negative value becomes N).

For simplicity, the simulations presented in the study employ $N = 100$ (i.e. 100 agents assigned to 100 locations each with one seat). In addition, to mimic reality, in which agents are likely to be able to report only a limited number of location preferences, the preference vectors for each agent are truncated such that only the top 10 ranks are retained and indifference is established among the remaining locations. The simulations vary both the correlation between preference and outcome vectors (three values of ρ_{op}: -0.5, 0, and 0.5) and the correlation between preference vectors across agents (three values of ρ_p: 0, 0.5, and 0.8). This yields nine different scenarios, and in each we apply our mechanism to make the assignment for various values of $\bar{g}$. Figure 1 in the main text displays the results.

In addition, Figure S1 in this SI shows the results of the same simulations when the preference rank vectors are not truncated.

D U.S. Refugee Application

D.1 Background Information on U.S. Resettlement

Resettled refugees in the United States are assigned to locations based on collaboration between the Department of State and nine voluntary resettlement agencies. During a regular draft, refugees are first allocated to one of the nine agencies according to specific quotas. Agencies are then responsible for assigning refugees to locations within their networks. Typically refugees are assigned as cases, where a case is a family. The assignment varies based on whether the refugee has family ties in the United States. Refugees with ties are placed at the location most proximate to the tie. Refugees without such ties, so-called "free cases," are assigned on a case-by-case basis and can be assigned to any location in the network. Placement officers consider special characteristics of the case (nationality, case structure, medical needs) and consult with the local offices on whether they can accommodate a case (e.g. some offices may lack interpreters for particular languages). Among the offices that can accommodate a case, the case is then typically assigned to offices with the smallest proportion of their yearly capacity currently filled. Note that a different process applies to refugees with Special Immigrant Visas (SIVs).

Once a refugee case has been assigned, the local office then provides placement and reception services for 90 days beginning after arrival as mandated by the U.S. Resettlement Program. The duration is 180 days for refugees assigned to the matching grant program. Agencies are mandated to report employment outcomes to the Department of State after the conclusion of the placement and reception period. If a refugee leaves the area before the placement and reception period ends, they may no longer receive the benefits associated with the placement and reception service.

D.2 Registry Data

Our data includes all refugees that were resettled by one of the largest resettlement agencies and arrived between quarter 1, 2011 and quarter 3, 2016. The same data is used in Bansak et al. (2018). We restrict the sample to those aged between 18 and 64 years at the time of arrival (i.e. working age). We also remove a small number of duplicates and locations that have had less than 200 refugees assigned to them over the entire period. In the final data there are 33,782 refugees from 22,144 cases. Of those, 9,506 refugees are from free cases.

Table S1 shows the descriptive statistics for our sample. Below is a list of variables and measures used:

- *Male*: Binary variable coded as 1 for males and 0 for females.
- *Speaks English*: Binary variable coded as 1 for refugees who speak English at the time of arrival and 0 otherwise.
- *Age at arrival*: Age at arrival measured in years.
- *Education*: Highest level of educational attainment at arrival. Categories include: None/Unknown, Less than Secondary, Secondary, Advanced, and University.
- *Country of origin*: Country of origin or nationality.
- *Employed*: Binary variable coded as 1 for refugees who are employed at 90 days after arrival, and 0 otherwise.
- *Year of arrival*: Year of arrival (continuous).
- *Month of arrival*: Month of arrival (continuous).
- *Free case*: Binary variable coded as 1 for refugees who are free cases with no U.S. ties, and 0 otherwise.

D.3 Applying the Mechanism

We applied our mechanism to the data on the refugee families who arrived in the third quarter (Q3) of 2016, specifically focusing on refugees who were free to be assigned to different resettlement locations (561 families, 919 working-age individuals). To generate each family's outcome score vector across each of the locations, we employed the same methodology in Bansak et al. (2018), using the data for the refugees who arrived from 2011 up to (but not including) 2016 Q3 to train gradient boosted tree models that predict the expected employment success of a family (i.e. the mean probability of finding employment among working-age members of the family) at any of the locations, as a function of their background characteristics. These models were then applied to the families who arrived in 2016 Q3 to generate their predicted employment success at each location, which comprise their outcome score vectors.

To generate preference rank vectors, we infer revealed location preferences from secondary migration behavior. Specifically, we use the same modeling procedures used in the outcome score estimation, simply swapping in outmigration in place of employment as the response variable. This allows us to predict for each refugee family that arrived in 2016 Q3 the probability of outmigration at each location as a function of their background characteristics. For each family, we then rank locations such that the location with the lowest (highest) probability of outmigration is ranked first (last).

In applying our mechanism to the 2016 Q3 refugee data, we impose real-world assignment constraints, giving each location capacity for the same number of families as were sent to those locations in actuality. We also truncate each family's preference rank vector such that only the first 10 ranks are retained and indifference is established among the remaining locations. Figure 3 in the main text displays the results. In addition, Figure S2 in this SI shows the results of the same simulations when the preference rank vectors are not truncated.

More details on the procedures used to generate the outcome score and preference rank vectors can be found below.

D.4 Generating Outcome Scores and Preference Ranks

The methods used for estimating the predicted probabilities of employment and outmigration in this study are the same as those employed in Bansak et al. (2018). The following material describes the procedures and is modified directly from the Supplementary Materials document of Bansak et al. (2018).

D.5 Training vs. Prediction Data Designation

Let $\mathbf{T}$ (training data) be the matrix of refugee data, in which the unit of observation is a single refugee, that will be used for model training. The $\mathbf{T}$ matrix contains the data for all working age refugees in our data who arrived starting in 2011 and up to (but not including) the third quarter of 2016. For each refugee we observe her assigned location, response variables of interest (employment for the outcome score and outmigration for the preference rank), and her full set of covariates.

Let **R** (prediction data) be the matrix of data for the working age, free case refugees who arrived during the third quarter of 2016. This comprises the set of refugees to whom we applied our mechanism in this application. In a real-world application, these **R** matrix data would correspond to new refugee arrivals and must include the same set of covariates as in the model training data. In contrast to the model training data, however, these prediction data need not include refugees' response variables. In fact, in a real-world prospective implementation of the mechanism, refugees belonging to these prediction data will not have yet been assigned to a resettlement location.

Note that when applying our mechanism both the model training and prediction data should be subsetted to the group of refugees for whom the outcomes of interest are relevant. In our application the integration outcome is employment and therefore the population of interest is working-age refugees. In addition, the prediction data should be subsetted only to those refugees who are free to be assigned to different resettlement locations—in contrast to refugees with predetermined geographic destinations due to family ties and other special circumstances—as this is the subset for whom the mechanism is designed to help with the assignment process. That said, the model training data need not be restricted to only free cases. Free-case and non-free-case refugees might be sufficiently dissimilar that forecasting free-case refugees' outcomes with models built using non-free-case data may seem problematic. This issue is addressed, however, by including case type as a predictor variable in the model building process (see below).

D.6 Modeling

The training data is used to build a bundle of learners that predict refugees' probabilities of the response variables (employment and outmigration), and those learned models are then applied to the prediction data to generate their predicted probabilities.

The modeling is implemented on a location-by-location basis. For each resettlement location, the training data are first subsetted to those refugees who were assigned to that location, and a statistical model is then fit that uses those refugees' characteristics to predict the response. That fitted model is then applied to the prediction data (2016 Q3 refugees) to predict the probability of the response for these refugee arrivals if they were hypothetically sent to the location in question. This process is performed separately for each individual location, which yields for each refugee in the prediction data a vector of predicted probabilities, one for each location. Collectively for all refugees in the prediction data, the final result is then a matrix of predicted probabilities (**M** matrix) with rows representing individual refugees and columns representing resettlement locations. Note that there are two **M** matrices: one for probabilities of employment and one for probabilities of outmigration.

More formally, for each refugee $r = 1, ..., n_T$, let the response of interest (e.g. employment) be denoted by $y_r \in \{0, 1\}$ and the location assignment denoted by $w_r \in \{1, ..., k\}$, for a total of k possible resettlement locations. Let $\vec{x}_r$ denote a p-dimensional feature vector comprised of the characteristics of refugee r, and x_{rm} denote the mth feature in $\vec{x}_r$, where $m = 1, ..., p$. The goal of the modeling process is to learn the function $\theta_l(\vec{x}_r) = P(y_r = 1|\vec{x}_r, w_r = l)$. The following describes the steps in the modeling stage.

1. Designate the historical model training data and denote it by the matrix $\mathbf{T}$:

$$\mathbf{T} = \begin{bmatrix} y_1 & w_1 & x_{11} & \cdots & x_{1m} & \cdots & x_{1p} \\ \vdots & \vdots & \vdots & & \vdots & & \vdots \\ y_r & w_r & x_{r1} & \cdots & x_{rm} & \cdots & x_{rp} \\ \vdots & \vdots & \vdots & & \vdots & & \vdots \\ y_{n_T} & w_{n_T} & x_{n_T 1} & \cdots & x_{n_T m} & \cdots & x_{n_T p} \end{bmatrix}$$

2. Train a set of k models, $\mathbf{\Theta} = \{\hat{\theta}_1(\vec{x}_r), ..., \hat{\theta}_l(\vec{x}_r), ..., \hat{\theta}_k(\vec{x}_r)\}$ as follows.
   For $l = 1, ..., k$:

   (a) Subset $\mathbf{T}$ to refugees for whom $w_r = l$ (i.e. refugees assigned to l-th location), and call this $\mathbf{T}_l$:

$$\mathbf{T}_l = \begin{bmatrix} y_1 & x_{11} & \cdots & x_{1m} & \cdots & x_{1p} \\ \vdots & \vdots & & \vdots & & \vdots \\ y_r & x_{r1} & \cdots & x_{rm} & \cdots & x_{rp} \\ \vdots & \vdots & & \vdots & & \vdots \\ y_{n_l} & x_{n_l 1} & \cdots & x_{n_l m} & \cdots & x_{n_l p} \end{bmatrix}_{w=l} = \begin{bmatrix} y_1 & \vec{x}_1 \\ \vdots & \vdots \\ y_r & \vec{x}_r \\ \vdots & \vdots \\ y_{n_l} & \vec{x}_{n_l} \end{bmatrix}_{w=l}$$

   where n_l denotes the number of refugees for whom $w_r = l$.

   (b) Using the data in $\mathbf{T}_l$ (the outcome y_r and feature vector $\vec{x}_r$ for all n_l refugees in $\mathbf{T}_l$), model and estimate the function $\hat{\theta}_l(\vec{x}_r)$.

3. Designate the data on new refugee arrivals and denote them by the matrix $\mathbf{R}$:

$$\mathbf{R} = \begin{bmatrix} \dot{x}_{11} & \cdots & \dot{x}_{1m} & \cdots & \dot{x}_{1p} \\ \vdots & & \vdots & & \vdots \\ \dot{x}_{r1} & \cdots & \dot{x}_{rm} & \cdots & \dot{x}_{rp} \\ \vdots & & \vdots & & \vdots \\ \dot{x}_{n_R 1} & \cdots & \dot{x}_{n_R m} & \cdots & \dot{x}_{n_R p} \end{bmatrix} = \begin{bmatrix} \dot{\vec{x}}_1 \\ \vdots \\ \dot{\vec{x}}_r \\ \vdots \\ \dot{\vec{x}}_{n_R} \end{bmatrix}$$

   where n_R denotes the number of new refugee arrivals.

   The matrix $\mathbf{R}$ corresponds to the 2016 Q3 refugees in this application.

4. For all refugees in $\mathbf{R}$ and all resettlement locations, estimate $P(\dot{y}_r = 1|\dot{\vec{x}}_r, \dot{w}_r = l)$ as follows.
   For $r = 1, ..., n_R$:

For $l = 1, ..., k$:

Estimate $P(\dot{y}_r = 1|\vec{\dot{x}}_r, \dot{w}_r = l)$ by applying lth model in $\mathbf{\Theta}$ to $\vec{\dot{x}}_r$:
$\widehat{P}(\dot{y}_r = 1|\vec{\dot{x}}_r, \dot{w}_r = l) = \hat{\theta}_l(\vec{\dot{x}}_r) \equiv \pi_{rl}$

Arrange the π_{rl} into a vector, $\vec{\pi}_r = [\pi_{r1}, ..., \pi_{rk}]$.

5. Produce a matrix of predicted probabilities, with rows corresponding to new refugees and columns corresponding to resettlement locations, as follows.
   Arrange vectors $\vec{\pi}_r$ into rows of the matrix $\mathbf{M}$:

$$\mathbf{M} = \begin{bmatrix} \vec{\pi}_1 \\ \vdots \\ \vec{\pi}_r \\ \vdots \\ \vec{\pi}_{n_R} \end{bmatrix} = \begin{bmatrix} \pi_{11} & \cdots & \pi_{1l} & \cdots & \pi_{1k} \\ \vdots & & \vdots & & \vdots \\ \pi_{r1} & \cdots & \pi_{rl} & \cdots & \pi_{rk} \\ \vdots & & \vdots & & \vdots \\ \pi_{n_R1} & \cdots & \pi_{n_Rl} & \cdots & \pi_{n_Rk} \end{bmatrix}$$

   This is the final modeling stage output.

We follow Bansak et al. (2018) and use boosted trees (Friedman et al., 2009, Friedman, 2001) to estimate $\hat{\theta}_l(\vec{x}_r)$ in step 2(b). See Bansak et al. (2018) for more details on the selection criteria and model performance metrics leading to the choice of boosted trees. Specifically, we use stochastic gradient boosted trees (bag fraction of 0.5) with a binomial deviance loss function (Friedman, 2002, Friedman et al., 2009), which we implemented in `R` using the `gbm` package (Ridgeway, 2017). Tuning parameter values, including the interaction depth, learning rate, and number of boosting iterations (the early stopping point) are selected via cross-validation within the training data for each location-specific model.

We use the following predictors: *Free case*, *Speaks English*, *Age at arrival*, *Male*, *Education* (ordered variable differentiating between no/unknown education, less than secondary, secondary, technical/professional, and university), *Country of origin* (one binary variable for each of the largest origin groups including Burma, Iraq, Bhutan, Somalia, Afghanistan, Democratic Republic of Congo, Iran, Eritrea, Ukraine, Syria, Sudan, Ethiopia, and Moldova), *Year of arrival*, and *Month of arrival*.

D.7 Mapping to Case-Level

Since the assignment of refugees typically takes place at the level of the case (typically a family), we need to map the refugee-level predicted probabilities from the modeling process to a case-level metric. For each case-location pair, we apply the mapping function to the refugee-location predicted probabilities for all refugees belonging to that case, yielding a single value for that case-location pair. This results in a new matrix ($\mathbf{M}^*$ matrix) with the same number of columns (locations) as previously but now as many rows as cases rather than refugees.

Formally, let $i = 1, ..., n$ denote the refugee case, with a total of n cases, where $n \leq n_R$. The mapping process then proceeds as follows:

1. Perform mapping of individual predicted probabilities to case-level metric as follows.
   For $i = 1, ..., n$:

   For $l = 1, ..., k$:

   Let $\tilde{\pi}_{il} = \{\pi_{rl} \quad \forall \quad r \in i\}$. (That is, $\tilde{\pi}_{il}$ is the set of all π_{rl} for the lth location and refugees belonging to the ith case.)

   Compute $\gamma_{il} = \psi(\tilde{\pi}_{il})$ where ψ is a predetermined mapping function.

   Arrange the γ_{il} into a vector, $\vec{\gamma}_i = [\gamma_{i1}, ..., \gamma_{ik}]$.

2. Produce a matrix containing the case-level metric for all case-location pairs, with rows corresponding to cases and columns corresponding to resettlement locations, as follows.
   Arrange vectors $\vec{\gamma}_i$ produced in step 1 into rows of the matrix $\mathbf{M}^*$:

$$\mathbf{M}^* = \begin{bmatrix} \vec{\gamma}_1 \\ \vdots \\ \vec{\gamma}_i \\ \vdots \\ \vec{\gamma}_n \end{bmatrix} = \begin{bmatrix} \gamma_{11} & \cdots & \gamma_{1l} & \cdots & \gamma_{1k} \\ \vdots & & \vdots & & \vdots \\ \gamma_{i1} & \cdots & \gamma_{il} & \cdots & \gamma_{ik} \\ \vdots & & \vdots & & \vdots \\ \gamma_{n1} & \cdots & \gamma_{nl} & \cdots & \gamma_{nk} \end{bmatrix}$$

   This is the final mapping stage output.

In step 1, the function ψ must be specified. In our application, we employ the mean for both the predicted probabilities of employment and the predicted probabilities of outmigration (see Bansak et al. (2018) for alternative choices).

D.8 Final Construction of Outcome Scores and Preference Ranks

The $\mathbf{M}^*$ matrix pertaining to the predicted probabilities of employment directly provides the outcome scores for use in the mechanism. However, the $\mathbf{M}^*$ matrix pertaining to the predicted probabilities of outmigration must be further transformed to provide the (inferred) preference ranks. Specifically, for each row (case), we rank locations such that the location with the lowest (highest) average probability of outmigration is ranked first (last), producing a preference rank vector for each case.

D.9 Alternative Method for Estimating Preferences via Structural Adjustment

We also examined an alternative method to estimate location preferences using a model that explicitly corrects for potential bias due to relocation costs. Note that in our data we only observe whether a refugee out-migrates from her initial resettlement location or not. This decision will be a function of both location preferences (i.e. outmigration should be higher in less desirable locations) and the costs of relocation (e.g. varying geographic and economic factors may result in higher costs of relocating from certain locations).

Here we leverage a structural model of outmigration to isolate the component of outmigration that is likely attributable to location preferences, rather than the costs of relocation. In particular, we follow standard models in the literature on immigrant and refugee location choices and estimate the following structural model of outmigration:

$$y_{ijt} = \alpha + \beta X_{jt} + \theta_j + \phi_t + \theta_j \times t + \epsilon_{ijt}$$

where y_{ijt} is the outcome of whether refugee i who arrived in year t out-migrates from her initial resettlement location j, X_{jt} are a set of time-varying location specific characteristics (e.g. rental prices, unemployment rates, ethnic networks, welfare generosity, etc.) that affect the costs of relocation with coefficients β, θ_j is a set of location specific fixed effects that capture all time-invariant factors that affect the costs of relocation (e.g. remote location), ϕ_t is a set of year fixed effects that capture common shocks (e.g. changes in transportation costs), and $\theta_j \times t$ is a set of location-specific linear time trends that capture changes in location-specific relocation costs that have a linear effect on outmigration (e.g. local economic decline, changes in local transportation infrastructure, etc.).

We include in X a set of location characteristics that are commonly included in structural models of location choices (Borjas, 1999, Zavodny, 1999, Damm, 2009, Åslund and Rooth, 2007, Mossad et al., 2020). In particular, we include the local unemployment rate and personal income per capita to proxy for economic opportunities (Åslund and Rooth, 2007, Damm, 2009, Mossad et al., 2020), rental prices to proxy for cost of living (Damm, 2009, Mossad et al., 2020), ethnic shares to proxy for enclave effects (Beaman, 2012, Mossad et al., 2020), and welfare spending per capita to proxy for welfare magnet effects (Damm, 2009, Borjas, 1999). A list of definitions and sources are provided below. To merge in this information, we first identified the county of each resettlement location and then merged in the location-specific characteristics measured at the refugee's time of arrival.

We fit the model with a logistic link function on the training data of refugees that arrived prior to the third quarter of 2016. Note that to fit this model we restrict the training data to only refugees who arrived as free cases. Since free cases do not choose their initial resettlement location but are exogenously placed by the resettlement agencies, this sample restriction limits potential bias due to individuals sorting into initial locations based on unobserved characteristics such as location preferences (see Åslund and Rooth (2007) for a similar identification that leverages a placement policy in Sweden to estimate location preferences).

We then use the fitted model to generate the predicted probabilities for outmigration for each family in the test set of quarter 3 2016 arrivals for each resettlement location. These predictions capture the probabilities of outmigration that we would expect for a given family purely based on the location specific relocation costs as captured by the structural model.

In the next step, we then compute the difference between the predicted probabilities from our previous model that was based on individual-level characteristics and the predicted probabilities from the structural model. The resulting differences can then be interpreted as variation in outmigration that is mostly driven by location preferences because it is adjusted for the variation in outmigration that is driven by structural relocation costs.

For example, consider a family who has a very low predicted probability of outmigration in a given location based on the individual model, but based on the structural model the predicted probability of outmigration in the same location is very high. This would indicate that this family has a strong preference to remain in this location even though based on the structural factors they would be pulled towards relocating. On the flip side, a family that has a very high predicted probability of outmigration based on the individual model but a very low probability of outmigration based on the structural model would suggest that they have a strong preference against living in this location given that the structural factors would pull them towards staying.

Accordingly, as a last step for each family, we rank the locations based on the differences in predicted probabilities such that the location with the most negative (most positive) difference is ranked as most (least) preferred.

The list of geographic factors and data sources is as follows:

- *Annual unemployment rate* in county. Data retrieved from the Local Area Unemployment Statistics (LAUS) from the Bureau of Labor Statistics.
- *Monthly Rental Price Index* in county. Data retrieved from Zillow Rent Index (ZRI) (Time Series Multifamily, SFR, and Condo/Coops). The Zillow Rent Index is a smoothed measure of the typical estimated market rate rent across a given region and housing type. We linearly interpolated for missing values.
- *Personal annual income per capita* in county. Data retrieved from the Bureau of Economic Analysis.
- *Share of co-nationals* in metro area. Share of co-nationals in each county year is estimated based on ACS 5 year and 3 year samples (downloaded from IPUMs, using the BLP and MET2013 variables). For some resettlement locations that were outside a metro area we merge based on city or PUMA instead of metro area.
- *Total annual state and local welfare spending per capita.* Data retrieved from Annual Survey of State and Local Government Finances US Census.

The results of applying our mechanism to the 2016 Q3 refugee data using these new preference estimates is shown in Figure S3.

E Education Application

E.1 Background Information on Application

Here we illustrate our mechanism by applying them to a hypothetical example of choice of elementary schools. We consider a case where a school district might be interested to assign incoming Kindergarten students to elementary schools in the district with the goal to maximize academic achievement as measured by scores on standardized tests that are administered at the end of the Kindergarten grade. Students have preferences over schools and so the goal of the mechanism is to optimize on test scores and preferences subject to the minimum expected average level of test score set by the district.

E.2 Tennesse Star Data

We leverage data from the Tennessee's Student Teacher Achievement Ratio (STAR) project conducted by the Tennessee State Department of Education. This data include student level data on from a longitudinal experiment in Tennessee that began in 1985 and tracked a cohort of students progressing from kindergarten through third grade (for details on the data and sample see Achilles et al. (2008)). The data includes demographic information on the students, indicators for the schools that they attended, as well as information on achievement tests that were administered annually at the end of each grade. We focus on the sample of 1,674 students from 33 schools that are observed for all grades from Kindergarten through 3rd grade and have non-missing data for tests scores and background characteristics.

Table S2 shows the descriptive statistics for our sample. Below is a list of variables and measures used:

- *Month of birth*: This variable is coded with values from 1 to 12.
- *Year of birth*: This variable is coded with values including 1978, 1979, 1980, and 1981
- *Race*: The student's race coded as six categories including White, Black, Asian, Hispanic, Native American, and Other.
- *Free lunch*: Binary variable coded as 1 if the student was eligible for free/reduced lunch in Kindergarten and zero otherwise.
- *Special Education*: Binary variable coded as 1 if the student was eligible for special education status in Kindergarten and zero otherwise.
- *Female*: Binary variable coded as 1 for female students and zero otherwise
- *SAT Score Reading*: Total reading scaled score on the Stanford Achievement Test at the end of Kindergarten.

- *SAT Score Math*: Total math scaled score on the Stanford Achievement Test at the end of Kindergarten.
- *SAT Score Listening*: Total listening scaled score on the Stanford Achievement Test at the end of Kindergarten.
- *Sum of SAT Scores*: Sum of the three SAT scores for Reading, Math, and Total listening scaled score at the end of Kindergarten.
- *Left Kindergarten*: Variable used to measure outmigration from the Kindergarten school. Higher values indicate that the student left the Kindergarten school faster which can be interpreted as a stronger preference for another school. Coded 0 if student remained in the Kindergarten school for 1st, 2nd, and 3rd grade; coded 1 if student stay in Kindergarten school for grade 1 and 2 but left for another school for grade 3; coded 2 if student stay in Kindergarten school for grade 1, but left for another school for grade 2; and coded 3 if students left for another school for grade 1.

E.3 Applying the Mechanism

To generate each student's outcome score vector across each of the schools, we used the same stochastic gradient boosted tree models as in the refugee application to predict the expected tests score of a student at any of the schools, as a function of their background characteristics. The background characteristics included the students' age, gender, race, as well as information on whether they are eligible for free school lunches (a proxy for socioeconomic status) or special education. The test score outcome was defined as the sum of reading, math, and listening scaled SAT scores for the Kindergarten level. Given the small sample size for some schools we used the same data for the training and validation set and increased the bag fraction to 1. We look for the best fitting tree models over interactions depth of 3 to 8 using 5-fold cross-validation with total of 1,500 trees.

To generate the school preferences we inferred revealed school preferences of students from the observed transfers out of the schools. Specifically, we used the same modeling procedure of stochastic gradient boosted tree model as for the test scores but instead used a response variable that measured whether a student had transferred to another school by the first, second, or third grade. Based on these models we can then predict for each student the propensity for leaving each school as a function of their background characteristics. For each student, we then rank schools such that the school with the lowest (highest) propensity for transferring out is ranked first (last). In contrast to the refugee application there is no mapping to a case level since assignments are done at the student level. We impose the constraint that every school can only receive as many students as the did in actuality.

F Tables

Table S1: Descriptive Statistics for United States Refugee Sample

	Mean	SD
Male	0.53	0.50
Speaks English	0.42	0.49
Age:		
18-29	0.44	0.50
30-39	0.28	0.45
40-49	0.16	0.37
50+	0.11	0.31
Education:		
None/Unknown	0.18	0.39
Less than Secondary	0.39	0.49
Secondary	0.21	0.41
Advanced	0.10	0.30
University	0.12	0.33
Origin:		
Burma	0.23	0.42
Iraq	0.20	0.40
Bhutan	0.13	0.34
Somalia	0.11	0.31
Afghanistan	0.07	0.25
Other	0.26	0.44
Employed	0.23	0.42

Sample consists of refugees of working age that were resettled by one of the largest resettlement agencies and arrived in the period from quarter 1, 2011 to quarter 3, 2016. N = 33,782.

Table S2: Descriptive Statistics for Student Sample

	Mean	SD
Month of Birth	6.22	3.45
Year of Birth	1979.74	0.45
Race:		
White	0.82	0.38
Black	0.17	0.38
Asian	0.00	0.05
Hispanic	0.00	0.02
Native American	0.00	0.00
Other	0.00	0.03
Free Lunch	0.34	0.47
Special Education	0.02	0.13
Female	0.51	0.50
SAT Score Reading	445.34	31.17
SAT Score Math	499.43	43.41
SAT Score Listening	547.18	30.38
Sum of SAT Scores	1491.95	88.22
Left Kindergarten	0.06	0.37

Sample consists of students from the Tennessee Star data. N = 1,674. Note that the "Left Kindergarten" variable denotes the number of years during K-3 that a student was in a school that was different from their Kindergarten school.

G Figures

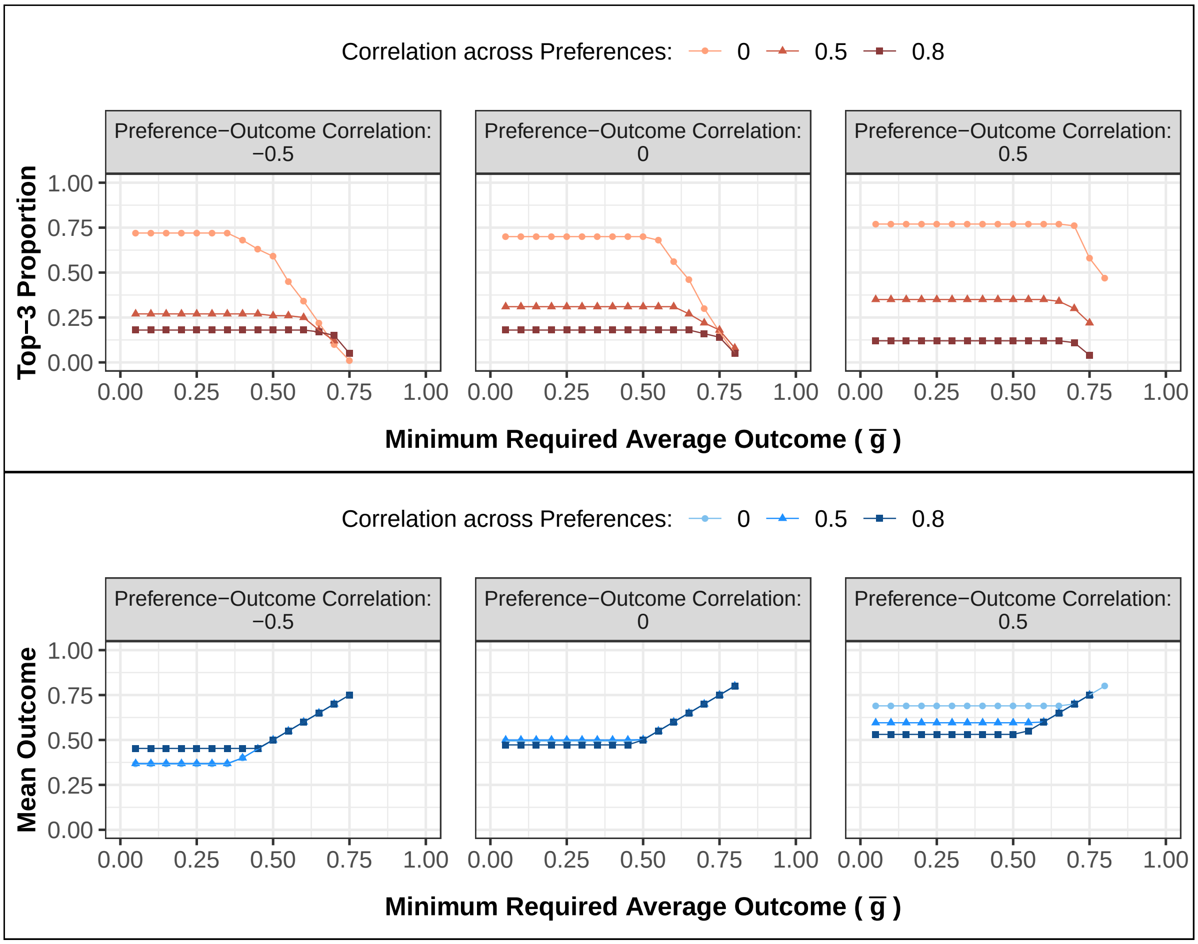

Figure S1: Results from applying our $\bar{g}$-Constrained Priority Mechanism to simulated data (without truncated preferences) that varies the correlations between location preference and integration outcome vectors and the correlations between preference vectors across families. This figure shows the results of the same simulations as in the main text Figure 1, except that the simulated families' preference rank vectors were not truncated in the simulations illustrated here. Upper panel shows the average probability that a family was assigned to one of its top three locations. Lower panel shows the realized average integration outcomes, i.e. the average projected probability of employment. $N = 100$.

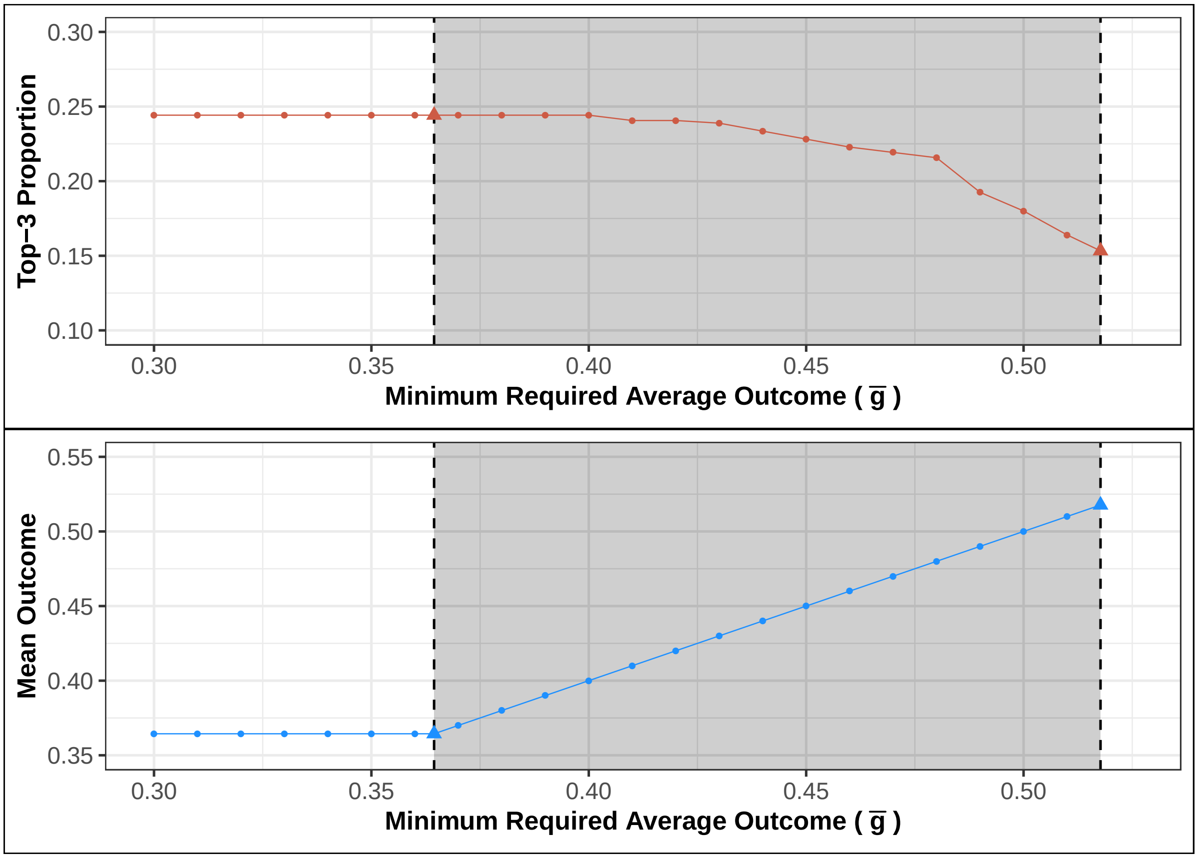

Figure S2: Results of applying our $\bar{g}$-Constrained Priority Mechanism to refugee families in the United States (without truncated preferences) for various specified thresholds for the expected minimum level of average integration outcomes ($\bar{g}$). This figure shows the results of applying the mechanism to the same data as in the main text Figure 3, except that the families' preference rank vectors were not truncated in the application illustrated here. Upper panel shows the average probability that a refugee got assigned to one of their top three locations. Lower panel shows the realized average integration outcomes, i.e. the average projected probability of employment. $N = 561$ families who arrived in Q3 of 2016.

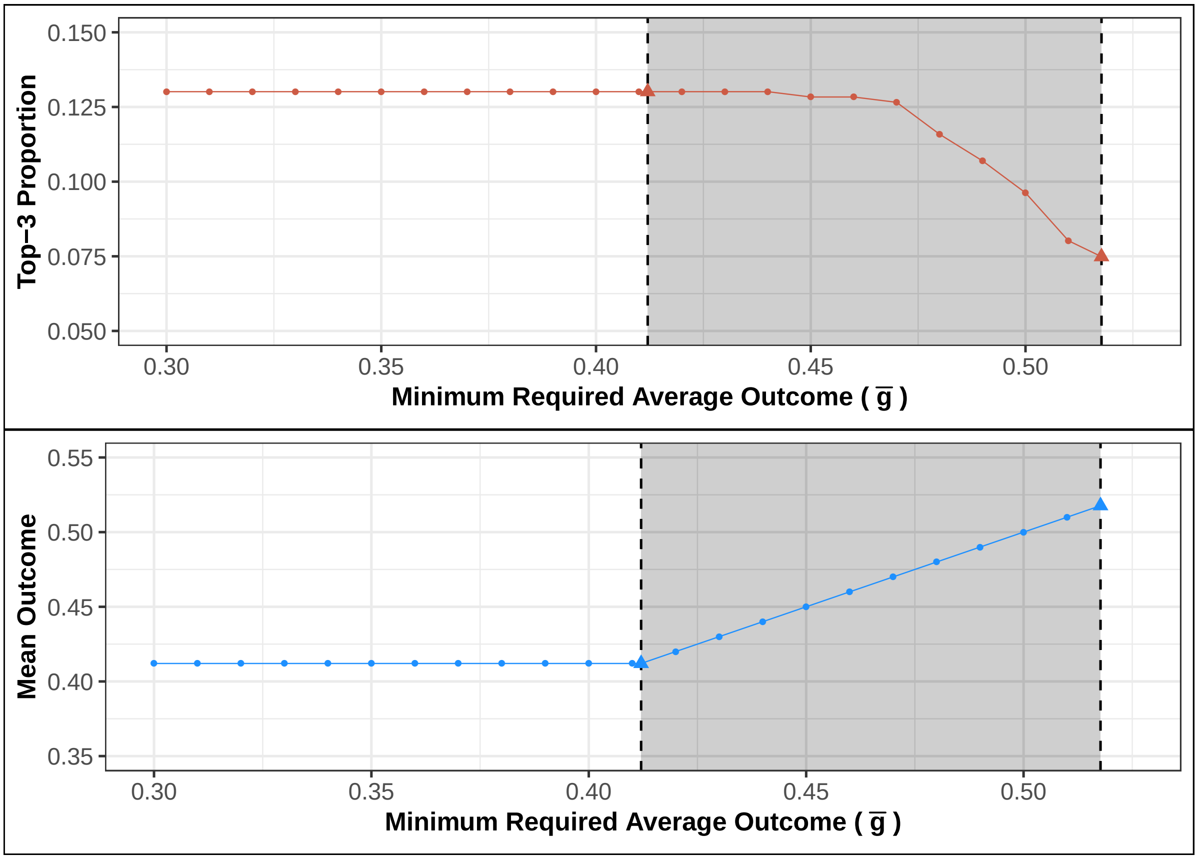

Figure S3: Results of applying our $\bar{g}$-Constrained Priority Mechanism to refugee families in the United States for various specified thresholds for the expected minimum level of average integration outcomes ($\bar{g}$), with structurally adjusted preference estimates. This figure shows the results of applying the mechanism to the same data as in the main text Figure 3, except that the families' preference rank vectors were adjusted using a structural model design to account for differential relocation costs across locations. Upper panel shows the average probability that a refugee got assigned to one of their top three locations. Lower panel shows the realized average integration outcomes, i.e. the average projected probability of employment. $N = 561$ families who arrived in Q3 of 2016.

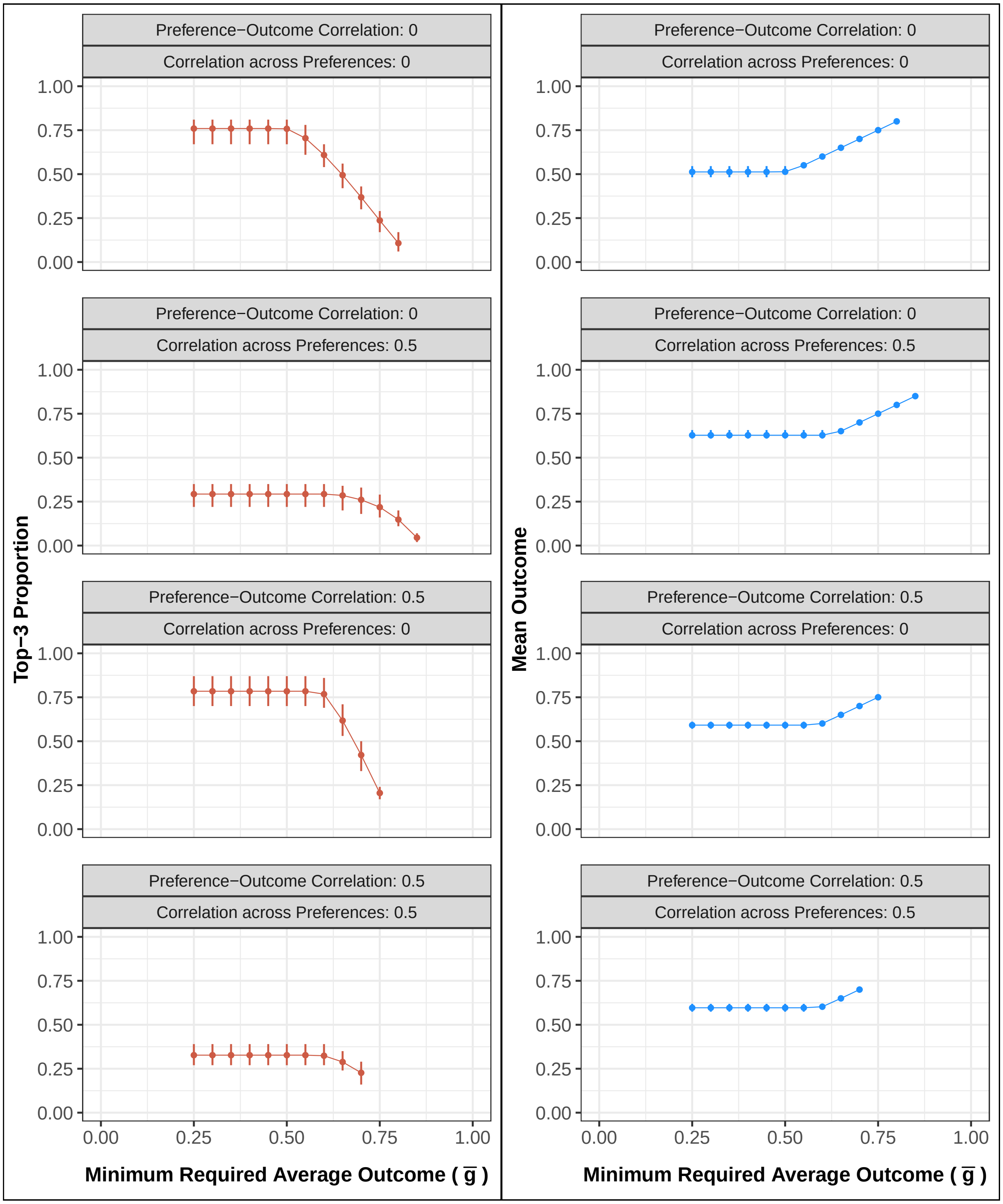

Figure S4: Results from re-running simulations discussed in the main text Simulation Data Section and shown in Figure 1, where at each level of $\bar{g}$, the mechanism is applied 100 separate times and the order of the agents is re-randomized each time. The dots denote the results—the proportion assigned to a top-3 location in the panels on the left, and the mean outcome score on the right—averaged across the 100 re-orderings, and the intervals denote the maximum and minimum results obtained across the 100 re-orderings.

References

Achilles, C., H. P. Bain, F. Bellott, J. Boyd-Zaharias, J. Finn, J. Folger, J. Johnston, and E. Word (2008): "Tennessee's Student Teacher Achievement Ratio (STAR) project," Https://doi.org/10.7910/DVN/SIWH9F, Harvard Dataverse.

Åslund, O. and D.-O. Rooth (2007): "Do when and where matter? Initial labour market conditions and immigrant earnings," *The Economic Journal*, 117, 422–448.

Bansak, K., J. Ferwerda, J. Hainmueller, A. Dillon, D. Hangartner, D. Lawrence, and J. Weinstein (2018): "Improving refugee integration through data-driven algorithmic assignment," *Science*, 359, 325–329.

Beaman, L. A. (2012): "Social networks and the dynamics of labour market outcomes: Evidence from refugees resettled in the US," *The Review of Economic Studies*, 79, 128–161.

Bertsekas, D. P. and P. Tseng (1994): "RELAX-IV: A faster version of the RELAX code for solving minimum cost flow problems," Tech. rep., Massachusetts Institute of Technology, Laboratory for Information and Decision Systems Cambridge, MA.

Borjas, G. J. (1999): "Immigration and welfare magnets," *Journal of Labor Economics*, 17, 607–637.

Burkard, R., M. Dell'Amico, and S. Martello (2012): *Assignment Problems*, Society for Industrial and Applied Mathematics.

Damm, A. P. (2009): "Determinants of recent immigrants' location choices: Quasi-experimental evidence," *Journal of Population Economics*, 22, 145–174.

Friedman, J. H. (2001): "Greedy function approximation: A gradient boosting machine," *Annals of Statistics*, 29, 1189–1232.

——— (2002): "Stochastic gradient boosting," *Computational Statistics & Data Analysis*, 38, 367–378.

Friedman, J. H., T. Hastie, and R. Tibshirani (2009): *The Elements of Statistical Learning, 2nd ed.*, Springer.

Hansen, B. B. and S. O. Klopfer (2006): "Optimal full matching and related designs via network flows," *Journal of Computational and Graphical Statistics*, 15, 609–627.

Kuhn, H. W. (1955): "The Hungarian method for the assignment problem," *Naval Research Logistics (NRL)*, 2, 83–97.

Mossad, N., J. Ferwerda, D. Lawrence, J. M. Weinstein, and J. Hainmueller (2020): "In search of opportunity and community: The secondary migration of refugees in the United States," *Science Advances*, forthcoming.

Munkres, J. (1957): "Algorithms for the assignment and transportation problems," *Journal of the Society for Industrial and Applied Mathematics*, 5, 32–38.

Ridgeway, G. (2017): "Package 'gbm'," Tech. rep., Comprehensive R Archive Network (CRAN), https://cran.r-project.org/web/packages/gbm/gbm.pdf.

Zavodny, M. (1999): "Determinants of recent immigrants' locational choices," *International Migration Review*, 33, 1014–1030.