informs
http://pubsonline.informs.org/journal/opre

OPERATIONS RESEARCH
Articles in Advance, pp. 1–19
ISSN 0030-364X (print), ISSN 1526-5463 (online)

Crosscutting Areas

Placement Optimization in Refugee Resettlement

Narges Ahani[a], Tommy Andersson[b], Alessandro Martinello[c], Alexander Teytelboym[d], Andrew C. Trapp[e]

[a] Data Science Program, Worcester Polytechnic Institute, Worcester, Massachusetts 01609; [b] Department of Economics, Lund University, Lund SE-220 07, Sweden; [c] Danmarks Nationalbank, 2100 Copenhagen Ø, Denmark; [d] Department of Economics, University of Oxford, Oxford OX1 3UQ, United Kingdom; [e] Foisie Business School and Data Science Program, Worcester Polytechnic Institute, Worcester, Massachusetts 01609

Contact: nahani@wpi.edu, https://orcid.org/0000-0002-6012-1707 (NA); tommy.andersson@nek.lu.se (TA); alem@nationalbanken.dk (AM); alexander.teytelboym@economics.ox.ac.uk, https://orcid.org/0000-0002-6570-1903 (AT); atrapp@wpi.edu, https://orcid.org/0000-0003-0143-9093 (ACT)

Received: February 20, 2019
Revised: March 20, 2020; July 27, 2020
Accepted: September 21, 2020
Published Online in Articles in Advance: March 24, 2021

OR/MS Subject Classifications: programming: integer; statistics: regression; networks/graphs: matchings; information systems: decision support systems
Area of Review: Policy Modeling and Public Sector OR

https://doi.org/10.1287/opre.2020.2093

Abstract. Every year, tens of thousands of refugees are resettled to dozens of host countries. Although there is growing evidence that the initial placement of refugee families profoundly affects their lifetime outcomes, there have been few attempts to optimize resettlement decisions. We integrate machine learning and integer optimization into an innovative software tool, *Annie*™ Matching and Outcome Optimization for Refugee Empowerment (*Annie*™ Moore), that assists a U.S. resettlement agency with matching refugees to their initial placements. Our software suggests optimal placements while giving substantial autonomy to the resettlement staff to fine-tune recommended matches, thereby streamlining their resettlement operations. Initial back testing indicates that *Annie*™ can improve short-run employment outcomes by 22%–38%. We conclude by discussing several directions for future work.

Funding: A. C. Trapp and N. Ahani are grateful for the support of the National Science Foundation (Operations Engineering) [Grant CMMI-Q: 9 1825348] as well as that provided by the Foisie Business School at Worcester Polytechnic Institute. T. Andersson gratefully acknowledges the financial support of the Jan Wallander and Tom Hedelius Foundation [Research Grant P2016-0126:1] and the Ragnar Söderberg Foundation [Grant E8/13]. A. Teytelboym gratefully acknowledges the Skoll Centre for Social Entrepreneurship Research Accelerator Grant (Implementing the Local Refugee Match). This work was supported by the Economic and Social Research Council [Grant ES/R007470/1].
Supplemental Material: The e-companion is available at https://doi.org/10.1287/opre.2020.2093.

Keywords: refugee resettlement • matching • integer optimization • machine learning • humanitarian operations

1. Introduction

In 2018, there were 20.4 million refugees—the highest number ever recorded—under the mandate of the United Nations High Commission for Refugees (UNHCR) (UNHCR 2019b). Of those, the UNHCR considers 1.44 million refugees to be in need of *resettlement*—permanent relocation from their asylum country to a third country (UNHCR 2019a). The number of cases submitted by the UNHCR for resettlement in 2018, however, was just over 81,000, with fewer than 56,000 refugees departing for resettlement (UNHCR 2019a). Refugees in need of resettlement are particularly vulnerable: a quarter are survivors of torture and a third face persecution in their country of origin (UNHCR 2019a, Annex 3). Currently, most refugees departing for resettlement are Syrians who seek asylum in Turkey, Lebanon, and Jordan; but there are also thousands of resettled refugees from the Democratic Republic of the Congo, Iraq, Somalia, and Myanmar.

Dozens of countries, including the United States, Canada, the United Kingdom, Australia, France, Norway, and Sweden, resettle refugees. (For refugee allocation mechanisms *across* countries, see Moraga and Rapoport (2014) and Jones and Teytelboym (2017a).) There is ample empirical evidence that the initial placement of refugees within the host countries determines their lifetime employment, education, and welfare outcomes (Åslund and Rooth 2007, Åslund and Fredriksson 2009, Åslund et al. 2010, 2011, Damm 2014, Ferwerda and Gest 2017). Therefore, ensuring the best initial match between the refugee family and the community is crucial for social, economic, and humanitarian perspectives. Even so, resettlement capacity offered by communities is rarely being used to maximize either the welfare of refugees or of the host population.

This paper integrates machine learning and integer optimization into a software package that we call *Annie*™ Matching and Outcome Optimization for Refugee Empowerment (*Annie*™ Moore), named after Annie Moore, the first immigrant on record at Ellis Island, New York, in 1892. *Annie*™ is, to the best of our knowledge, the first software designed for resettlement agency prearrival staff to recommend data-driven, optimized matches between refugees and local affiliates while respecting refugee needs and affiliate capacities. *Annie*™ was developed in close collaboration with representatives from all levels of the U.S. resettlement agency HIAS (founded as the Hebrew Immigrant Aid Society), where the first version was deployed in May 2018. It was presented in August 2018 to the U.S. Department of State and all staff at HIAS, with new features being regularly added.

We combined techniques from operations research, machine learning, econometrics, and interactive visualization to create *Annie*™. The software is distinctive in that it blends rigorous analysis with careful attention to the detail of the day-to-day resettlement process for resettlement staff. As such, *Annie*™ integrates the generation of data-informed recommendations with substantial end-user autonomy by the end user. This flexibility empowers staff to focus more of their resources on refugee families that might be more challenging to match (for example due to complex medical conditions). Back testing indicates that *Annie*™ would have been able to increase employment outcomes among refugees resettled by HIAS in 2017 by between 22% and 38%, depending on the constraints activated by the agency staff. *Annie*™ also alleviates inefficiencies in the manual matching process, and holds much promise for future impact in refugee resettlement—both domestically and abroad—as well as for new applications, such as asylum matching.

The paper proceeds as follows. Section 2 describes the specific context of refugee resettlement in the United States and places our work in the greater context of humanitarian operations problems. Section 3 sets up the integer optimization model that guides the matching recommendations. In Section 4, we explain how we estimate counterfactual employment probabilities from data. Section 5 discusses the back testing we conducted to validate our approach. Section 6 describes the implementation and features of our software. Section 7 concludes and points to many directions for further work. An e-companion contains detailed data descriptions, estimation procedures and diagnostics, and a variety of experiments using different objective functions and testing the sensitivity of our modeling.

2. Background Context and Prior Work

Because *Annie*™ helps HIAS to resettle refugees in the United States, we briefly describe the U.S. resettlement program.

2.1. Refugee Resettlement in the United States

The United States has historically been, by a wide margin, the world's largest destination of resettled refugees, with 22,491 arriving in 2018, down from 53,716 in 2017 and 84,994 in 2016 (in this manuscript, all references to years of refugee data are presented in terms of the *fiscal* year, that is, from October 1 through September 30) (U.S. Department of State 2020). In terms of refugees resettled per capita (calendar year), the United States trails a number of countries, including Sweden, Norway, Canada, New Zealand, Iceland, and Australia. The U.S. Refugee Admissions Program (USRAP) resettles U.S. refugees and is managed by the Bureau of Population, Refugees, and Migration (PRM) of the U.S. Department of State, with the assistance of the U.S. Citizenship and Immigration Services (USCIS) of the U.S. Department of Homeland Security, and the Office of Refugee Resettlement (ORR) of the U.S. Department of Health and Human Services (HHS). Alongside the UNHCR and the International Organization for Migration, these agencies coordinate identifying refugees, conducting security checks, and arranging for travel funding from the refugees' destinations.

The actual matching of refugees to their initial placements is delegated to nine resettlement agencies, previously known as *voluntary agencies*. In addition to HIAS, these agencies include Church World Service (CWS), Ethiopian Community Development Council (ECDC), Episcopal Migration Ministries (EMM), International Rescue Committee (IRC), Lutheran Immigration and Refugee Service (LIRS), U.S. Conference of Catholic Bishops (USCCB), U.S. Committee for Refugees and Immigrants (USCRI), and World Relief (WR). HIAS handles around 5% of all refugees in the United States, resettling 2,038 refugees in 2017 and 3,844 refugees in 2016. The resettlement agencies are responsible for developing their own networks of *affiliates* in local communities that welcome refugees and help them integrate into a new life in the United States. Affiliates offer resettlement capacity voluntarily, although affiliate capacity is monitored and approved by the U.S. government. There are currently around 360 affiliates in approximately 200 local communities across the United States, and HIAS operates 20 of them at the time of this writing.

Resettlement agencies match refugees to affiliates during the resettlement process largely by hand. Resettlement staff from each agency meets weekly to select, in round-robin fashion, from a pool of "cleared for arrival" refugee *cases*. Each case consists of an immediate *family* of one or more members (we use *case* and *family* interchangeably). A significant portion (roughly one-third) of these are free cases, that is, they have no relatives in the United States. Such cases are especially vulnerable, as the absence of family support exacerbates the challenges of lacking language skills and independent financial means. Thus, the responsible agency must carefully leverage its affiliate network to inform their case selection. After each agency selects their set of weekly cases, staff manually assess—on a one-by-one basis—the feasibility and fit of cases to locations in their network. In addition to integration factors, such as language and nationality feasibility, the fit between the affiliate and the family depends on various community capacities, such as available placement capacity, housing availability, slots for English language instruction, and employment prospects.

This manual process creates multiple inefficiencies that motivated the development of *Annie*™. First, it is organizationally demanding for HIAS staff to keep in mind various support attributes, such as languages, nationalities, family composition, and medical needs for all affiliates. This information overload at times results in not meeting the needs of refugees and in stretching the provision capacity of the affiliates. Second, although established indicators exist to assess the degree to which a refugee has successfully integrated into their new surroundings, estimating and optimizing these welfare outcomes manually is prohibitive. Established indicators include employment and economic sufficiency, developed social networks, and civic engagement activities such as voting (see, e.g., Ager and Strang 2008, Lichtenstein et al. 2016). Hence, refugees are often not placed to the best available affiliate even according to well-defined outcome metrics. Third, inefficiencies arise from processing refugees case by case, in sequential fashion, rather than matching all arriving refugees to affiliates simultaneously. We show that *Annie*™ resolves or mitigates each of these inefficiencies.

2.2. Related Literature

Our work builds on a number of contemporary studies in humanitarian matching systems. One recent example is a tool to match children in state custody to families for adoption used by the Pennsylvania Adoption Exchange (Slaugh et al. 2016). Bansak et al. (2018) first proposed to use machine learning and linear programming for refugee resettlement based on employment data from the United States and Switzerland. Using a similar data set to theirs, we expand on their estimation techniques, while extending their optimization methods. Our integer optimization model extends the multiple multidimensional knapsack model for refugee matching (see also Trapp et al. 2018, Delacrétaz et al. 2019, and Nguyen et al. 2019). However, as we focus on outcome optimization, our work differs substantially from papers that suggested preference-based matching systems for refugee resettlement (Moraga and Rapoport 2014, Andersson and Ehlers 2017, Jones and Teytelboym 2017b, Aziz et al. 2018, Roth 2018, Delacrétaz et al. 2019, Nguyen et al. 2019).

Placement optimization in refugee resettlement shares many common features with other problems in humanitarian operations (Pedraza-Martinez and Van Wassenhove 2016, Besiou et al. 2018). Typical challenges in this sector include a severe lack of resources—financial, labor, time, and data—as well as complex decision environments. The refugee resettlement decision environment includes refugees as well as local communities, nonprofit organizations, donors, and federal, state, and local governments. Hence, similar to other humanitarian operations problems, placement optimization also diverges from the traditional stance of optimizing a single financial metric and may consider alternative objectives such as those based on equity (see, e.g., Çelik et al. 2012, Cao et al. 2016, Besiou and Van Wassenhove 2020). Refugee resettlement is perhaps most differentiated by its particular exposure and sensitivity to shifting political climates and attitudes, both domestic and abroad. This volatility generates significant uncertainty with respect to the operating and planning environments of resettlement agencies.

3. Integer Optimization for Resettlement

We formulate the operational challenge of matching refugee families to local communities, or affiliates, presently solved manually by resettlement agencies, using mathematical optimization.

3.1. Formal Problem Setup

We use i, j, k, and ℓ as indices for family (case), member, service, and affiliate, respectively. For any placement period, let $\mathcal{F} = \{F^1, F^2, \ldots, F^i, \ldots, F^{|\mathcal{F}|}\}$ be the set of refugee families to be placed. Each family F^i is a set of refugees consisting of one or more members, $F^i = \{f^{i,1}, f^{i,2}, \ldots, f^{i,j}, \ldots f^{i,|F^i|}\}$. For clarity of exposition, we refer to member j of family F^i as f^{ij}. Denote as N^i_w the set of working-age refugees in family F^i, where $N^i_w \subseteq F^i$. Denote the set of all refugees as $\mathcal{R} = \bigcup_{i \in \{1,2,\ldots,|\mathcal{F}|\}} \bigcup_{j \in \{1,2,\ldots,|F^i|\}} \{f^{ij}\}$. Moreover, let the set of affiliates (localities) to which families are resettled be $\mathcal{L} = \{L^1, L^2, \ldots, L^\ell, \ldots L^{|\mathcal{L}|}\}$.

A family F^i requires various *capacitated services* from a set $\mathcal{S} = \{S^1, S^2, \ldots, S^k, \ldots S^{|\mathcal{S}|}\}$. The needs of family F^i are summarized by vector s^i, where s^i_k denotes the required units of service k. The set $\mathcal{S}$ may include services such as raw weekly refugee processing capacity at affiliates, slots in English language instruction, school seats for children in the family, and housing availability. For every service S^k of local affiliate L^ℓ, at most $\bar{s}^\ell_k$ units may be filled by families placed in affiliate L^ℓ. There may also be a requirement of at least $\underline{s}^\ell_k$ units of the service S^k to be filled by the families placed in affiliate L^ℓ (we assume $\underline{s}^\ell_k \le \bar{s}^\ell_k$); in practice, nonzero lower bounds exist for certain services, such as ensuring regular, positive refugee placement in affiliates.

For every family F^i and local affiliate L^ℓ, let binary variable z^i_ℓ equal one if family F^i is matched to affiliate L^ℓ, and zero otherwise. Let $a_\ell^i \in \{0,1\}$ indicate whether family F^i can be feasibly placed in affiliate L^ℓ. The value of a_ℓ^i is a priori determined by evaluating the compatibility of family F^i with various binary community support services at affiliate L^ℓ, such as language and nationality, as well as large family and single parent support conditions (should these be present in the family). We denote these community support services as *binary services*.

We attribute to each refugee-affiliate match a single number called the *quality score*. The function $q : \mathcal{R} \times \mathcal{L} \to \mathbb{R}_{\geq 0}$ defines quality score q_ℓ^{ij} for any $f^{ij} \in \mathcal{R}$ and any $L^\ell \in \mathcal{L}$. We are interested in the scenario where q represents the employment outcome of refugee f^{ij} in affiliate L^ℓ and can be estimated from data using observable affiliate and family characteristics. In Section 7, we discuss the sole use of employment data to generate these estimates (indeed, no other data related to integration outcomes are systematically available). We aggregate the refugee-level quality scores q_ℓ^{ij} of each family F^i and affiliate ℓ into a case-level *value* (or weight) v_ℓ^i. The primary means of aggregation that we consider is the *sum* of individual scores over each family $v_\ell^i = \sum_{j=1}^{|F^i|} q_\ell^{ij}$ (SUM). Discussions on alternative interpretations of case-level quality scores can be found in Appendix EC.4 in the e-companion.

3.2. Placement Optimization

We now present integer optimization problem RefMatch, represented by (1a)–(1e):

$$\text{maximize} \quad \sum_{i=1}^{|\mathcal{F}|} \sum_{\ell=1}^{|\mathcal{L}|} v_\ell^i z_\ell^i, \tag{1a}$$

$$\text{subject to} \quad \sum_{\ell=1}^{|\mathcal{L}|} z_\ell^i \leq 1, \ \forall\, i, \tag{1b}$$

$$\underline{s}_k^\ell \leq \sum_{i=1}^{|\mathcal{F}|} s_k^i z_\ell^i \leq \bar{s}_k^\ell, \ \forall\, \ell, \ \forall\, k, \tag{1c}$$

$$z_\ell^i \leq a_\ell^i, \ \forall\, i, \ \forall\, \ell, \tag{1d}$$

$$z_\ell^i \in \{0,1\}, \ \forall\, i, \ \forall\, \ell. \tag{1e}$$

Objective function (1a) maximizes the total value over all matched families to affiliates. Constraint set (1b) ensures that families are placed in at most one affiliate. Constraint set (1c) ensures that lower and upper bounds are respected for all capacitated services and affiliates. Constraint set (1d) ensures that family affiliate matches can only occur when the affiliate can support the needs of the family, that is, the necessary binary services exist. Variable domains are specified in (1e). Finally, let $z^\star$ be the optimized match outcome, that is, the optimal solution representing the assignment of families to affiliates that optimizes objective function (1a).

Although integer optimization problem RefMatch bears similarity to a variety of knapsack-like problem classes, we are unaware of another application of this particular form:

- When $|\mathcal{S}| = 1$, $\underline{s}_k^\ell = 0\ \forall\, \ell$, and $s_k^i = 1\ \forall\, i$, the optimization problem can be solved via linear programming (Bansak et al. 2018).
- When $|\mathcal{S}| = 1$ and $\underline{s}_k^\ell = 0\ \forall\, \ell$, it is the NP-hard *multiple 0–1 knapsack problem*, which features multiple knapsacks and items that consume integer resources for the knapsack in which they are placed (Martello and Toth 1980).
- When $|\mathcal{L}| = 1$ and $\underline{s}_k^\ell = 0\ \forall\, k$, it is the NP-hard *multidimensional 0–1 knapsack problem*, which features knapsack items that consume integer resources along multiple dimensions (Fréville 2004).
- When $\underline{s}_k^\ell = 0\ \forall\, \ell, k$, it is the NP-hard *multiple multidimensional knapsack problem* and combines features of both, that is, multiple knapsacks along multiple dimensions (Song et al. 2008). If in addition, $\sum_{\ell=1}^{|\mathcal{L}|} z_\ell^i = 1\ \forall\, i$, it is the NP-hard *multiple-choice multidimensional knapsack problem* (Hifi et al. 2004); in our setting, there is no requirement (in theory) for every family to be placed in an affiliate.

Integer optimization problem RefMatch generalizes the *multiple multidimensional knapsack problem* of Song et al. (2008), as it allows for positive lower bounds $\underline{s}_k^\ell$ for any services and affiliates. The existence of such lower bounds differentiates it from the multiple multidimensional knapsack problem, as it may lead to infeasibility. The formulation is valid over any operational period. Because of its generality, our model can be customized to specific refugee resettlement settings. Section 5 shows the results of testing the sensitivity of our model under three different scenarios. First, we test the effect of relaxing upper bounds (1c) for the number of total resettled refugees. Second, we test the effects of lower bounds (1c) expressed as distributional requirements (such as minimum average case sizes across affiliates) and as lower bounds on the total number of resettled refugees. We also consider the effects of relaxing the binary service constraints (1d). We discuss alternative models and objective functions and conduct sensitivity checks in Appendices EC.4, EC.5, and EC.6 in the e-companion.

4. Estimation of Counterfactual Employment Probabilities

We use the estimated probability of employment of refugee f^{ij} in each affiliate L^ℓ as a measure of quality score, or

$$q_\ell^{ij} = E[y_{ij} \mid \mathbf{X}_{ij}, \ell],$$

where y_{ij} is (binary) outcome data indicating the employment status of refugee f^{ij} within 90 days of arrival in the United States and $\mathbf{X}_{ij}$ is a set of observable refugee characteristics and quarterly macroeconomic variables. We use national employment ratio and unemployment rate as macroeconomic variables, which are common to all refugees arriving in a given quarter. Further details on the available data appear in Appendix EC.1 in the e-companion.

Using expected potential outcomes rather than stated preferences for our counterfactual analysis creates two challenges. First, y_{ij} is unobserved for incoming refugees. Second, even for past refugees, we only observe $y_{ij}|x^{ij}_{\ell^\star}$, that is, employment status of refugee f^{ij} in $\ell^\star$, the affiliate to which they were actually assigned in the data. We do not observe the corresponding potential outcome distribution $y_{ij} \mid x^{ij}_{\ell}\ \forall\ \ell \neq \ell^\star$. Moreover, the functional form connecting y_{ij}, $\mathbf{X}_{ij}$, and ℓ is unknown. Specific synergies may exist between refugee characteristics and affiliates that affect refugee integration. Following Bansak et al. (2018), we thus exploit machine learning approaches to compute $\hat{q}^{ij}_{\ell}$, the estimated probability of employment of refugee f^{ij} in affiliate L^{ℓ}. Using data on refugees arriving between 2010 and 2016, we estimate both semi- and nonparametric functions $\hat{f}_\ell : \mathcal{R} \rightarrow \mathbb{R}_{\geq 0}$ such that $\hat{q}^{ij}_{\ell} = \hat{f}_\ell(\mathbf{X}_{ij})$. We then test the performance of these models on refugees arriving in 2017.

In the estimation process, we only use free cases, which are those refugees (individuals or families) that the resettlement agency can assign to any of the affiliates. We therefore exclude refugees with preexisting family ties, which are almost always preassigned to the affiliate where their preexisting connection resides. This choice restricts the samples we use to train and test the models to 2,486 and 498 refugees, respectively.

Although it may be tempting to increase the number of available observations for model estimation by including all refugees resettled by HIAS, the additional refugees will likely differ from the free cases to which *Annie*™ will be applied; including them in the estimation might introduce bias and likely overestimate existing synergies for free cases. For example, because of preexisting networks, family reunifications enjoy particular advantages (Edin et al. 2003, Patacchini and Zenou 2012) that would bias our estimates. By restricting our sample to free cases, we align the sample used for estimation with the sample on which *Annie*™ will be applied.

We estimate effects on employment for the 7 (out of 20) affiliates receiving at least 200 refugees up to 2016 and aggregate the remaining affiliates into a single partition ℓ_0. In a parametric approach, it is possible to estimate a fully saturated logit model for employment where flexible transformations of refugee characteristics $\mathbf{X}_{ij}$ are interacted with $\ell - 1$ affiliate dummies. Such an approach would, however, estimate an overly complex model, with poorly identified coefficients, and therefore yield poor predictive properties.

We thus use two alternative machine learning models. First, we introduce a Least Absolute Shrinkage and Selection Operator (LASSO) constraint to the interacted logit model to reduce model complexity. The single LASSO hyperparameter disciplines both main and interaction terms with the same weight, biasing them toward zero (and thus biasing predictions toward the mean). Second, we follow Bansak et al. (2018) and estimate a gradient boosted regression tree (GBRT), an iterative ensemble of classification trees. We set the hyperparameters of these models via five-fold cross-validation on our training sample (we internally calibrate constraint strength for LASSO and the learning rate and prepruning level for GBRT). We choose hyperparameter values for each model by maximizing the area under a receiver operating characteristic (ROC) curve.

We benchmark both models against the performance of a naïve constant estimator (see, e.g., Bansak et al. 2018) as well as two second-best standards. The first benchmark model is a standard logit model that includes all variables in $\mathbf{X}_{ij}$ but does not attempt to estimate affiliate-specific effects. The second benchmark model is a logit model with no LASSO constraint, where $\mathbf{X}_{ij}$ interacts with all ℓ affiliates. Table 1 shows that both LASSO and GBRT outperform the second-best benchmarks by over 20% in terms of misclassification error when applied to 2017 refugees. With respect to the constant-logit benchmark used by Bansak et al. (2018), we obtain a 37% and 34% improvement using LASSO and GBRT, respectively, which is comparable to the 28% they obtain in their U.S. data. The area under the ROC is highest for LASSO, but overall both models exhibit similar predictive power.

LASSO, however, produces slightly more stable and well-calibrated predictions, particularly for observations with high predicted employment probabilities. We obtain these results by bootstrapping the distribution of predictions for each data point in the test set given assignment of refugee f^{ij} to $\ell^\star$. In each of 1,000 iterations, we resample with replacement the training data set, re-estimate each model, and compute a new predicted probability of employment. The right panels of Figure 1 show the 5^{th} to 95^{th} percentiles of the prediction distributions for each data point in the test sample. The left panels show the distribution of bootstrapped interquartile ranges for each data point.

LASSO tends to produce more narrow predictions for refugees with high baseline probability of employment, which are highly relevant for the quantification of employment gains. LASSO is also better

Table 1. Model Performance

	Training data	Test data			
	Misc. error	Misc. error	Recall (1)	Precision (1)	AUC-ROC
Constant	0.259	0.319	0.000	0.000	0.500
Logit	0.240	0.263	0.491	0.609	0.790
Logit (by affiliate)	0.177	0.281	0.547	0.561	0.769
LASSO	0.159	0.201	0.453	0.637	0.799
Gradient boosted tree	0.129	0.209	0.396	0.624	0.791

Notes. *Misclassification (Misc.) error* is the proportion of observations incorrectly classified. *Recall* measures the proportion of correctly predicted employed refugees among refugees actually employed (true positives over true positives plus false negatives). *Precision* measures the proportion of correctly predicted employment cases among all predicted employment cases (true positives over true positives plus false positives). All of these measures refer to a binary classification with a threshold set at the standard value of 0.5. Because our measure of quality scores uses predicted probabilities of employment, this specific threshold does not affect optimal allocations. Area under the receiver operating characteristic curve (*AUC-ROC*) measures the area under the receiver operating characteristic curve for each model (ROC curves appear in Appendix EC.2 in the e-companion).

calibrated than GBRT. The sum of predicted employment probabilities given assignment of refugee f^{ij} to $\ell^{\star}$ is 157.93 for LASSO, and only 142.96 for GBRT (the test set consists of 159 refugees; calibration plots appear in Appendix EC.2 in the e-companion). Thus, although using either model has very similar consequences for optimal refugee assignment, in the remainder of the paper we quantify employment gains given the quality scores predicted by LASSO (and in Appendix EC.3 in the e-companion, we replicate employment gains under the predictions of GBRT).

5. Counterfactual Optimization Outcomes

We now describe the counterfactual impact of using our integer optimization problem RefMatch. We create test scenarios that result from varying three constraint sets. To quantify the impact of optimally reassigning refugees to affiliates, we use the employment probabilities for each affiliate estimated in Section 4. We compute the counterfactual gain in employment relative to our prediction from the LASSO model for 2017. Because our prediction is very close to the actual employment values—the LASSO model predicts 157.93 employed refugees versus 159 who were actually employed in the testing data—our optimization is a meaningful counterfactual exercise.

Our objective function (1a) maximizes the total expected number of employed refugees. Our binary service Constraints (1d) are language, nationality, single-parent, and large-family support. We set the capacity Constraints (1c) for each affiliate relative to the observed capacity in 2017. Moreover, we specify minimum average case sizes to enforce distributional constraints via the lower bounds in (1c). We vary the following three factors to create our test scenarios.

5.1. Affiliate Capacity

Affiliate capacity is federally approved but can be exceeded by up to 10% without further preapproval. Moreover, a common aim of the agencies is to fill at least 90% of the approved capacity at each affiliate. In 2017, somewhat unusually, approved capacity was much higher than the observed number of arriving refugees. We therefore use the observed placements at each affiliate to set sensible counterfactual capacities. We test three values: {observed capacity with no lower bound; and 110% of the observed capacity with no lower bound; and 110% of observed capacity with a lower bound of 90% of observed capacity}.

5.2. Binary Service Constraints

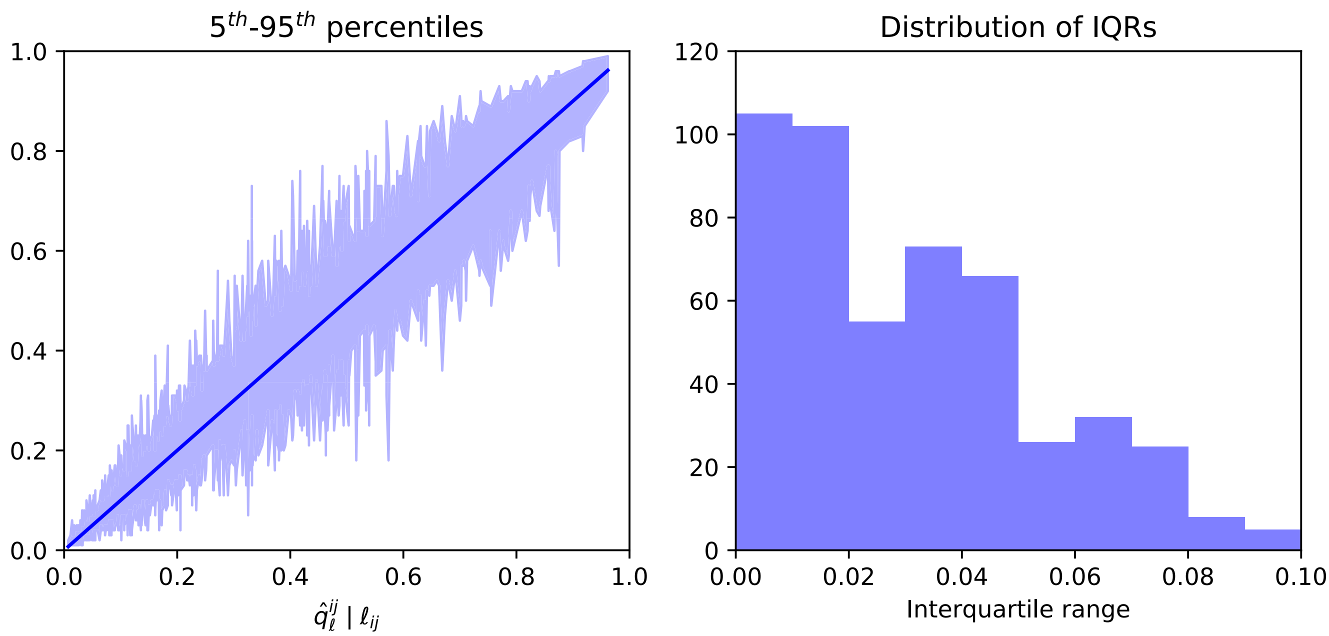
In the observed 2017 placements, binary service constraints were violated 38 times (26 language constraints, 1 nationality constraint, 8 single-parent constraints, and 3 large-family constraints), representing approximately 12% of resettled cases. However, binary service constraints, especially language constraints, can be important to ensure successful refugee integration. We therefore test two values: {binary service constraints are activated (ON) and binary service constraints are not activated (OFF) }.

5.3. Minimum Average Case Size in Each Affiliate

A placement that maximizes the total expected number of employed refugees could potentially pack many single-refugee cases or large-family cases into the same affiliate. This could be seen as unfair by the agencies, reduce support for resettlement, and stymie refugee integration. Therefore, to capture such equity considerations, we experiment with the implementation of a minimum average case size in each affiliate. The average case size in our 2017 test data set across all

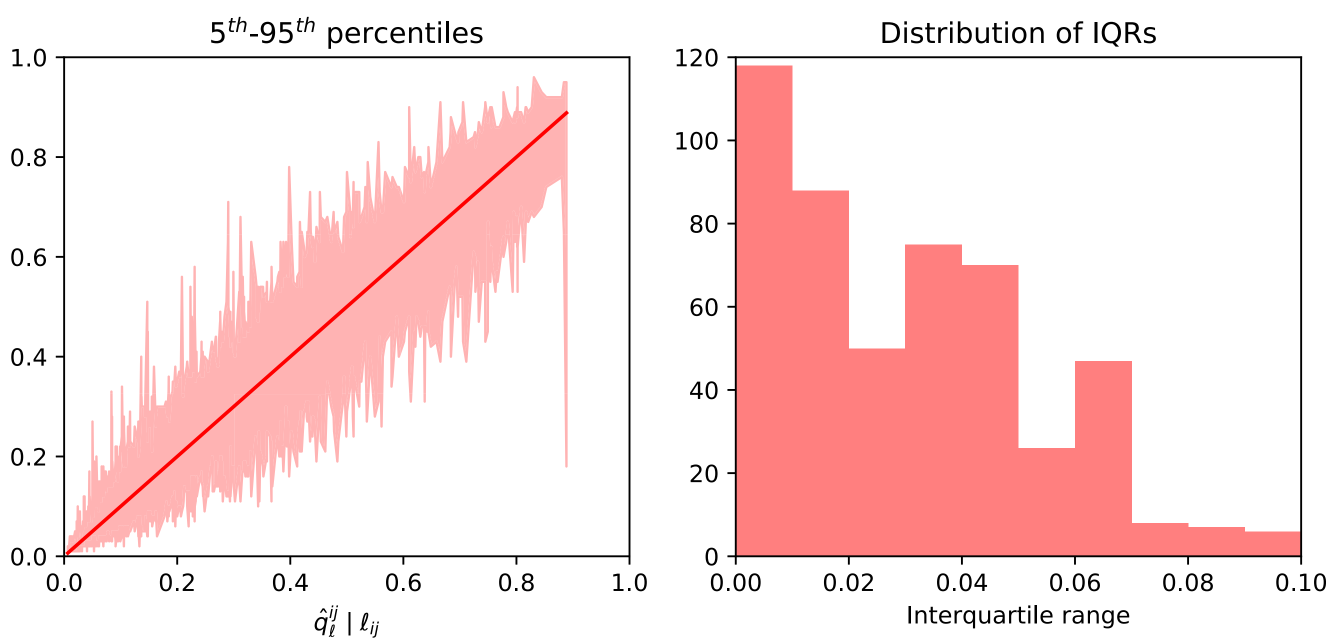
Figure 1. (Color online) Bootstrapped Uncertainty of Predicted Employment Probabilities in 2017 for LASSO and GBRT Model

(a)



LASSO

(b)



Gradient Boosted Regression Tree (GBRT)

Notes. Left panels: prediction distributions (5th–95th percentile) for each data point in test sample. Right panels: distribution of interquartile ranges for each data point in test sample.

affiliates is 2.55. We therefore test five values: {no minimum average case size, observed minimum average case size at each affiliate; 2, 2.5, and 3}.

5.4 Computational Setup

In total, we have $3 \times 2 \times 5 = 30$ counterfactual test scenarios. Akin to Bansak et al. (2018), we conduct our experiments using capacity levels for the period of one year. All experiments were run on a laptop computer with an Intel(R) Core(TM)i5-8365U 1.60GHz processor and 16GB RAM running 64-bit Microsoft Windows 10 Enterprise. The Gurobi Optimizer v9.0.0 (Gurobi 2020) and Python 3.7.4 was used for all counterfactual optimization testing in Section 5, and the optimality gap tolerance parameter MIPGap was set to zero. We summarize our results in Table 2.

5.5 Results and Discussion

First, note that without minimum average case size constraints, the gain in employment from optimization is over 30% in all scenarios. As Figure 2, (a) and (b) show, the employment probability distribution after optimization (almost) first-order stochastically dominates the preoptimized estimated distribution. Therefore, the estimated probabilities of employment increase for nearly all refugees after optimization. Moreover, Figure 3 shows that employment rates rise in nearly two-thirds of the affiliates after optimization. Table 2 further indicates that, if we do not impose binary service constraints, they are violated for around a quarter of the refugees—a rate much higher than in the test data (approximately 12%). However, the presence of binary service constraints and of increasing capacity has a fairly small impact on employment gains. Indeed, because in some cases our model leaves some refugees unplaced (meaning that they would need to be placed manually by agency staff), our employment gain estimates should be even higher.

However, in these scenarios, the optimization suggests rather unequal placement. Figure 4 compares the distribution of average case sizes in each affiliate to the distribution under our second counterfactual optimization, which produces the largest variance in average case sizes. Figure 5(a) shows that without distributional constraints, many single-person cases are placed in just three affiliates that offer a high probability of obtaining employment to many types of refugees. Other affiliates get much larger cases on average. This allocation may not be acceptable to a resettlement agency. Thus, we evaluated the placement optimization by enforcing minimum average case size constraints. At low values (up to 2.5) and at observed 2017 average case size values, the optimization is still able to realize employment gains of well over 20% (see also Figure 5(b)). This is extremely encouraging because it shows that our optimization performs well even under tight distributional constraints. However, at high average case sizes, the constraints bind harder and either reduce the performance of the model substantially (by not placing many refugees) or simply cause infeasibility. It should be noted that these are precisely the instances for which many cases and refugees are unplaced, thus causing reduced optimal objective function values and corresponding gains.

We report the runtime as the time (in seconds) to both build the optimization model and solve it to global optimality using Gurobi (Gurobi 2020). We can see that for the entire fiscal year 2017 (FY17) data set (839 refugees/329 cases/498 working-age refugees/20 affiliates), the combined build and solve times in Gurobi finish in under 30 seconds, with a median combined runtime of less than four seconds.

Overall, our optimization produces a substantial gain in employment, ensures that refugee binary services are better satisfied, and important distributional considerations can be respected. Moreover, the resettlement agency may impose any subset of the binary service constraints or introduce constraints on the number of refugees with certain regional origins (although regional constraints were formerly officially considered in U.S. placements, they are no longer specified).

It is worth emphasizing that the space of objective functions and constraints that the resettlement agency can impose within our model is much richer than what we have presented here. For example, the agency could select a different employment objective function, such as maximizing the sum of *maximum* employment probabilities within every matched case. In Appendix EC.4 in the e-companion, we provide further experiments that optimize over several reasonable (including equity-based) measures based on derived from the individual refugee-level quality scores q_ℓ^{ij}; we find that all perform fairly well with respect to gains in employment. Further details on these experiments can be found in Appendix EC.4 in the e-companion.

We also recognize that there is inherent uncertainty in the modeling environment with respect to estimating the quality score $\hat{q}_\ell^{ij}$ for each f^{ij} in affiliate L^ℓ. In Appendix EC.5 in the e-companion, we investigate how the objective function changes under our optimized placement outcome $z^\star$ when we resample $\hat{q}_\ell^{ij}$ from the estimated distribution. In particular, we observe that the refugee allocation determined by our approach produces stable employment gains and that these employment gains are not artificially inflated by uncertainty in the estimation of employment probabilities. The average expected employment given uncertainty in our predicted probabilities is within 2% of that obtained in our original back testing for almost all of the considered scenarios.

Finally, we note that these outcomes were obtained by optimizing placement of all refugees in FY17 without splitting into multiple periods, that is, over the entire year ($n = 1$), on par with experiments reported in Bansak et al. (2018). Although desirable,

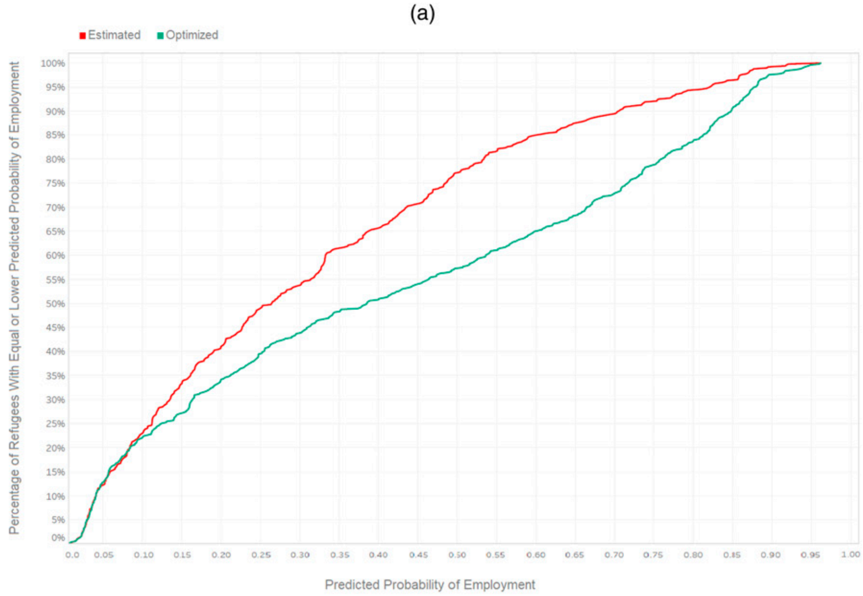
Table 2. Results of counterfactual employment optimization under various scenarios using the SUM objective and LASSO model

Capacity adjustment	Min avg case size	Binary service constraints	Total expected employed refugees	Gains with respect to predicted employed refugees (157.93)	St dev in avg case size across affiliates	No. of unplaced cases/refugees	No. of affiliates violating 90% capacity	No. and % of cases/refugees violating constraints	Build and run time (s)
Observed	None	Off	213.02	34.89%	1.29	0/0	0	70/209 (21.28%/24.91%)	0.48
Observed	None	On	208.25	31.86%	1.35	3/10	1	0/0 (0.00%/0.00%)	0.36
Observed	2	Off	206.28	30.61%	1.04	1/1	0	81/220 (24.62%/26.22%)	0.92
Observed	2	On	202.03	27.92%	1.17	2/9	1	0/0 (0.00%/0.00%)	0.93
Observed	2.5	Off	196.76	24.59%	0.33	1/1	0	97/265 (29.48%/31.59%)	5.43
Observed	2.5	On	192.95	22.17%	0.14	3/7	0	0/0 (0.00%/0.00%)	3.87
Observed	3	Off	172.83	9.43%	0.65	78/86	6	71/217 (21.58%/25.86%)	7.78
Observed	3	On	169.64	7.42%	0.65	79/89	6	0/0 (0.00%/0.00%)	4.97
Observed	Observed	Off	199.34	26.22%	0.84	2/2	0	81/232 (24.62%/27.65%)	5.40
Observed	Observed	On	195.65	23.89%	1.09	4/8	1	0/0 (0.00%/0.00%)	2.80
≤ 110%	None	Off	218.06	38.07%	1.40	0/0	1	71/199 (21.58%/23.72%)	0.78
≤ 110%	None	On	212.96	34.84%	1.42	2/9	2	0/0 (0.00%/0.00%)	0.55
≤ 110%	2	Off	212.39	34.48%	1.16	0/0	2	75/226 (22.80%/26.94%)	1.09
≤ 110%	2	On	207.72	31.53%	0.95	2/9	3	0/0 (0.00%/0.00%)	0.85
≤ 110%	2.5	Off	202.75	28.38%	0.38	0/0	1	87/222 (26.44%/26.46%)	5.19
≤ 110%	2.5	On	198.84	25.90%	0.66	3/7	3	0/0 (0.00%/0.00%)	3.23
≤ 110%	3	Off	177.51	12.40%	0.00	78/86	5	65/191 (19.76%/22.77%)	5.66
≤ 110%	3	On	174.27	10.34%	0.90	79/89	6	0/0 (0.00%/0.00%)	3.58
≤ 110%	Observed	Off	204.27	29.34%	0.83	0/0	3	81/207 (24.62%/24.67%)	6.35
≤ 110%	Observed	On	200.49	26.95%	1.07	3/7	4	0/0 (0.00%/0.00%)	8.03
[90%, 110%]	None	Off	218.06	38.07%	1.36	0/0	0	68/189 (20.67%/22.53%)	1.23
[90%, 110%]	None	On	212.91	34.82%	1.14	1/2	0	0/0 (0.00%/0.00%)	1.01
[90%, 110%]	2	Off	212.39	34.48%	0.95	0/0	0	72/194 (21.88%/23.12%)	1.40
[90%, 110%]	2	On	207.58	31.44%	1.05	2/6	0	0/0 (0.00%/0.00%)	1.67
[90%, 110%]	2.5	Off	202.75	28.38%	0.32	0/0	0	79/198 (24.01%/23.60%)	6.57
[90%, 110%]	2.5	On	198.81	25.89%	0.61	2/3	0	0/0 (0.00%/0.00%)	5.95
[90%, 110%]	3	Off			Infeasible instance				
[90%, 110%]	3	On			Infeasible instance				
[90%, 110%]	Observed	Off	204.26	29.34%	0.86	5/5	0	74/200 (22.49%/23.84%)	24.71
[90%, 110%]	Observed	On	200.36	26.86%	1.10	6/7	0	0/0 (0.00%/0.00%)	12.60

Note. avg, average; Std dev, standard deviation.

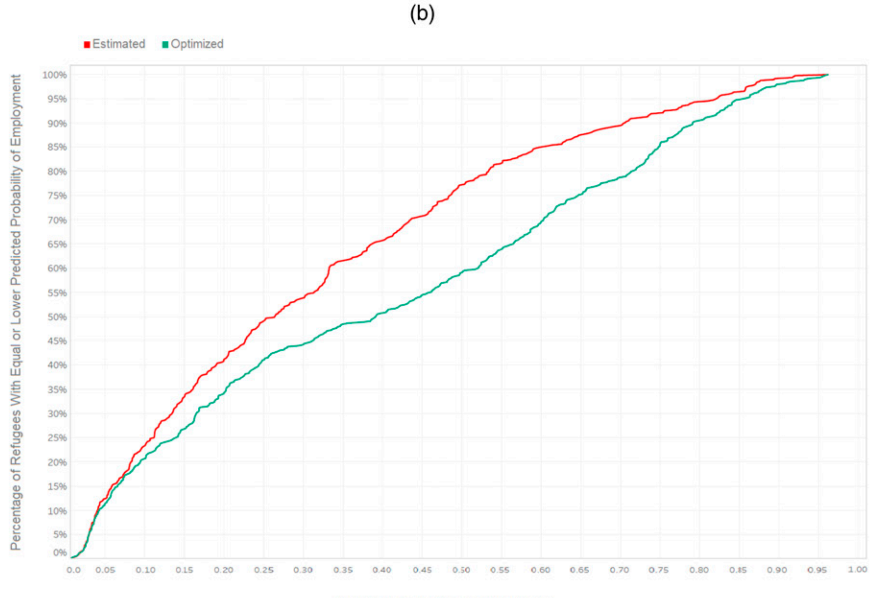
Figure 2. (Color online) Employment Gains from Optimizing Refugee Placement

(a)



Cumulative distribution of employment probabilities. Red: estimated probabilities under HIAS placement. Green: optimized probabilities for {observed capacity, activated binary service constraints, no mininum average case size} scenario.

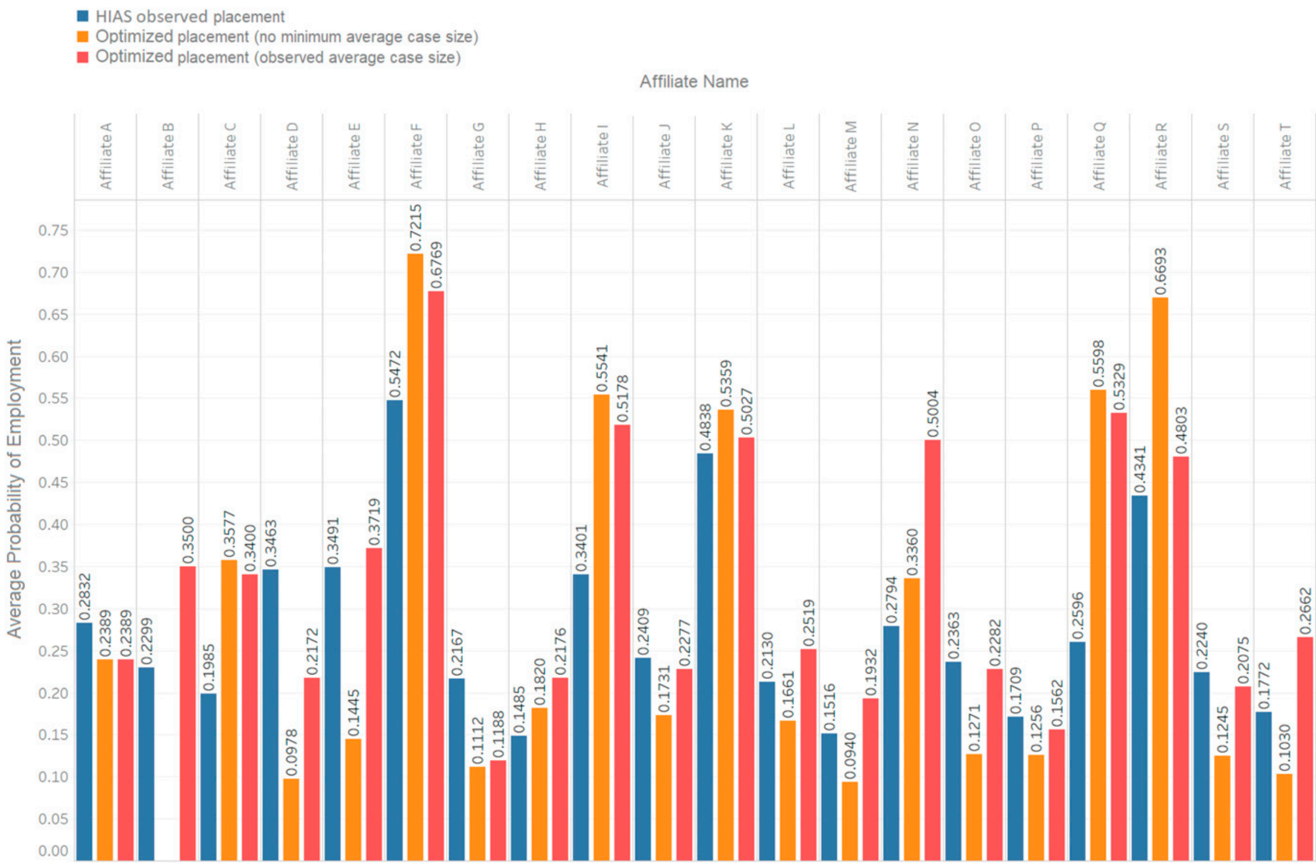
(b)



Cumulative distribution of employment probabilities. Red: estimated probabilities under HIAS placement. Green: optimized probabilities for {observed capacity, activated binary service constraints, at least observed average case size} scenario.

Notes. (a) Cumulative distribution of employment probabilities. Red: estimated probabilities under HIAS placement. Green: optimized probabilities for observed capacity, activated binary service constraints, no minimum average case size scenario. (b) Cumulative distribution of employment probabilities. Red: estimated probabilities under HIAS placement. Green: optimized probabilities for (observed capacity, activated binary service constraints, at least observed case size) scenario.

Figure 3. (Color online) Average Probability of Employment at Each Affiliate



Notes. Blue bar: estimated probabilities under HIAS placement. Orange bar: average probability of employment for observed capacity, activated binary service constraints, no minimum average case size scenario. Red bar: average probability of employment for (observed capacity, activated binary service constraints, at least observed average case size) scenario.

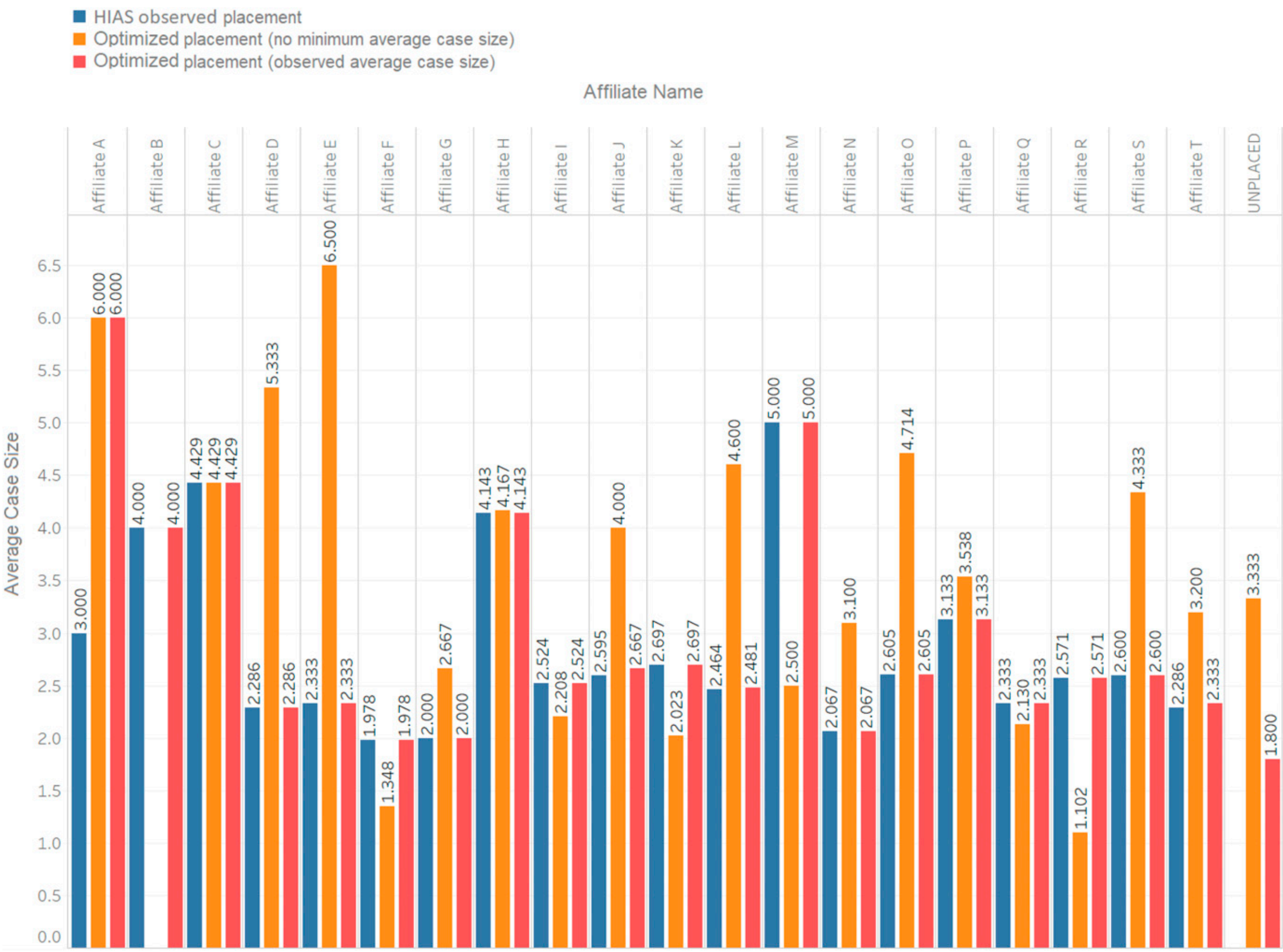
experiments with $n > 1$ placement periods in a given year introduced some additional nuances that required equally detailed implementation strategies. Even so, we present such experiments in Appendix EC.6 in the e-companion. Increasing the number of periods to $n \in \{4, 12, 52\}$ (that is, quarterly, monthly, and weekly) for placing refugees, and thereby allowing for the innate arrival stochasticity present in FY17 data, reveals encouraging results. Although gains are indeed largest for $n = 1$, our methods perform very well for $n = 4$ and $n = 12$ and respectably even for $n = 52$. We refer the reader to Appendix EC.6 in the e-companion for additional details.

6. Operationalizing Placement Software at U.S. Resettlement Agency

Integer optimization and machine learning hold great promise of solving the operational challenge of improving placement outcomes in refugee resettlement. Although these methods offer significant value, expertise is needed for successful implementation. In the private sector, this expertise is readily available. On the other hand, operations research in humanitarian environments, including refugee resettlement, typically feature significant challenges, such as lack of human and financial resources, lack of exposure to technology, and data scarcity. Humanitarian and nonprofit organizations must be responsive to crisis events, immediate needs, and changes in political and donor climates. These realities can make it fairly prohibitive for resettlement agencies to be proactive in pursuing, and implementing, advanced technological innovations.

Successful integration of operations research methods in a humanitarian environment requires cultivating and sustaining partnerships with stakeholders that include both management, as well as practitioners that will use the technology. The authors of this paper worked closely with many dedicated members of staff at HIAS for many months to develop *Annie*™ into an innovative, interactive optimization environment for refugee resettlement. Our close working relationship built a level of rapport that allowed us to understand and remedy real operational challenges faced by resettlement staff. We believe these are key elements for creating a successful software solution for improving humanitarian operations.

Figure 4. (Color online) Average Case Size at Each Affiliate



Notes. Blue bar: observed average case size under HIAS placement. Orange bar: average case size for (observed capacity, activated binary service constraints, no minimum average case size) scenario. Red bar: average case size for (observed capacity, activated binary service constraints, at least observed average case size) scenario.

6.1. Technologies Involved in the Creation of *Annie*™

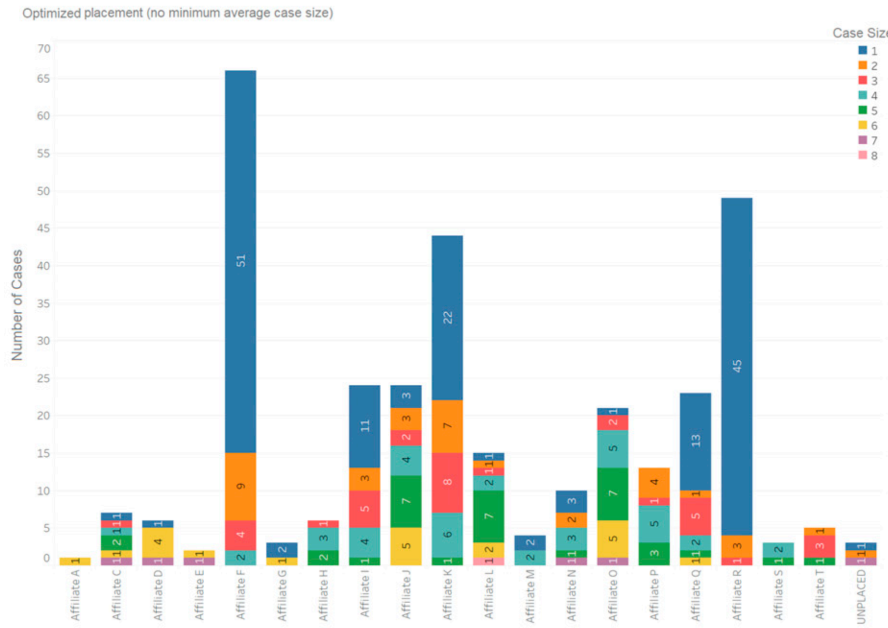
Annie™ represents the confluence of several open-source technologies, critical for this resource-constrained environment. In particular, the integer optimization problem RefMatch is modeled entirely within the PuLP Python modeling environment (Mitchell et al. 2020) and solved using the COIN-OR Branch-and-Cut solver (CBC) (COIN-OR 2020) solver. The machine learning models described in Section 4 are based on the Python scikit-learn package. We chose to develop the interactive environment of *Annie*™ as a web application. The back end is implemented in Python 3 using the Flask framework, with Jinja2 as the templating engine (Ronacher 2020). The front end is a combination of HTML, CSS, and JavaScript. We made this choice of technology because it is modern, stable, accessible, and easy to build upon. The only installation that is needed is (the free) Python 3 and some freely available packages and libraries. Moreover, it is a light technology: The front end operates entirely within a browser rather than as a downloadable, executable file. By combining core open-source integer optimization and machine learning technology within a flexible, modern interface, we were able to achieve a completely free, lightweight software solution for HIAS.

6.2. Interactive Optimization

Representing overall match quality in objective function (1a) is by no means trivial. Employment probabilities for refugees will always be estimated with error margins (see Appendix EC.5 in the e-companion for experiments and related discussions around uncertainty in employment probability estimation and corresponding sensitivity of match outcomes). Even if

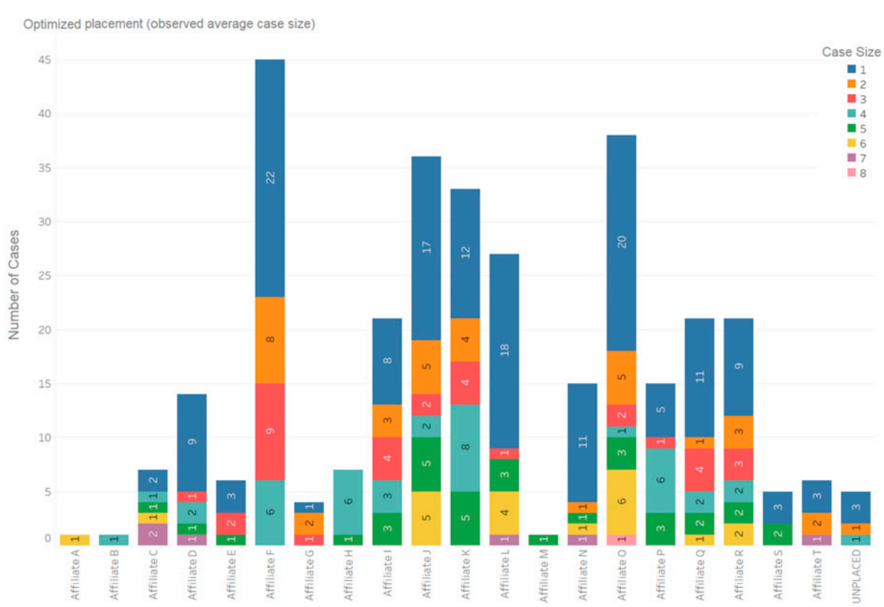
Figure 5. (Color online) Distribution of Case Sizes at Each Affiliate

(a)



Distribution of case sizes for {observed capacity, activated binary service constraints, no minimum average case size} scenario.

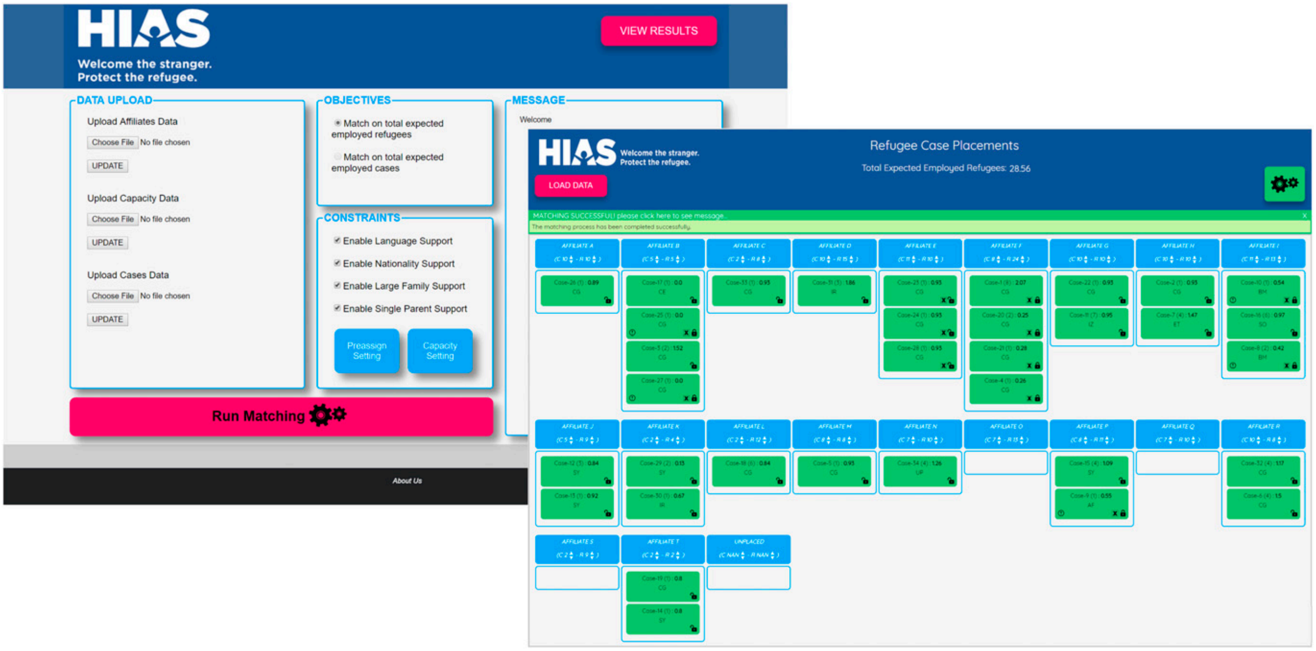
(b)



Distribution of case sizes for {observed capacity, activated binary service constraints, at least observed average case size} scenario.

Notes. (a) Distribution of case sizes for (observed capacity, activated binary service constraints, no minimum average case size) scenario. (b) Distribution of case sizes for (observed capacity, activated binary service constraints, at least observed average case size) scenario.

Figure 6. (Color online) *Annie*™ Interface



the employment probabilities could be perfectly estimated, any algorithmic solution should be carefully evaluated before actual implementation, as the overall livelihoods of refugees are at stake. Therefore, there is a need for an interactive optimization environment, where resettlement staff can interact with various facets of the problem context. Without compromising on the insights afforded by the theory and data, *Annie*™ was designed to accommodate the real needs of the practitioner. The purpose of developing *Annie*™ as an interactive optimization tool is to translate advanced analytical methods into effective decision tools (see, e.g., Meignan et al. 2015). The user of *Annie*™ is intimately involved in the matching process and can fine-tune the optimization results. We believe that *Annie*™ strikes the right balance. Our close interactions with HIAS allow us to iteratively develop and test multiple versions of the software via remote updating. Moreover, our predictive models can be refined as more data on 90-day employment outcomes arrive over time.

6.3. Features of *Annie*™

Operationally, *Annie*™ optimizes for the expected number of employed *refugees* throughout the network of affiliates at HIAS. Alternative objective functions, such as those discussed in Appendix EC.4 in the e-companion, can be easily implemented.

The *Load Data* view is depicted in the rear left of Figure 6, where the optimization environment can be configured for the matching process, including the activation of binary support services. The matching results can be seen in the *View Results* view depicted in the front right of Figure 6, where the total expected number of employed refugees is prominently displayed near the top.

The output of the matching engine results in cases being optimally assigned to affiliates, depicted with user-friendly *tiles*. Figure 7 displays both case and affiliate tiles. Case tiles show language, nationality, and other attributes unique to the family, whereas

Figure 7. (Color online) Expanding Tiles: Refugee and Affiliate Data

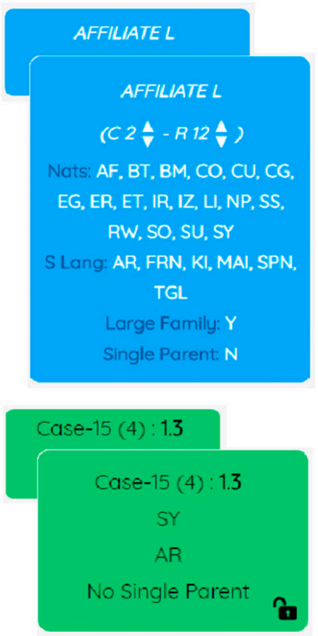
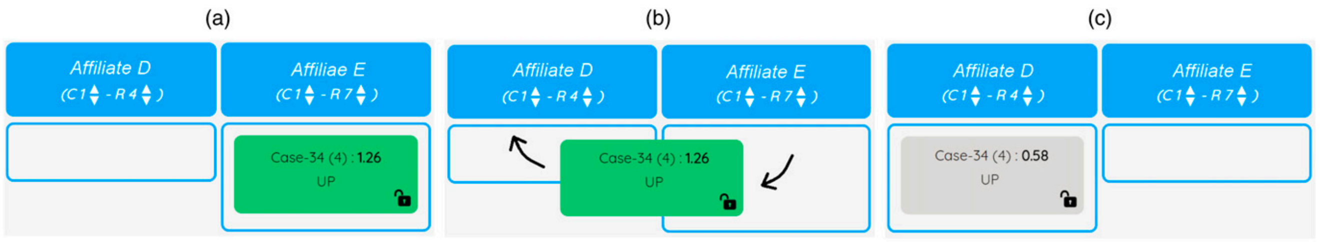


Figure 8. (Color online) Case Tiles Can Be Moved By Dragging to an Alternate Affiliate Tile



Case assigned to Affiliate E. Moving case tile to Affiliate D. Case tile moved to Affiliate D.

Notes. Upon moving, the match scores dynamically update. The background of the case tile changes to gray to indicate a nonoptimized state.

affiliate tiles show support features offered by affiliates, along with the ability to quickly adjust capacity. Clicking on the tiles expands their size to reveal detailed information at a quick glance. Case tiles can be moved to other affiliates as desired. Figure 8 illustrates the ability to dynamically view changes in the match scores as refugee case tiles are moved from one affiliate to the next. Moreover, the total expected number of employed refugees is also dynamically updated at the top of the *View Results* view. Hence, at a glance, the effect of moving cases to alternative affiliates is easily and clearly visualized.

Perhaps the most important feature of *Annie*™ is its ability for interactive optimization. Resettlement staff may interact with intermediate solver output in a manner that progresses toward eventual convergence of a finalized assignment of refugee cases to affiliates. This is facilitated through a lock icon on the case tile that resettlement staff can click, which locks desired case-affiliate matches. Figure 9 depicts this capability.

When locked, that case is temporarily "assigned" to that affiliate and is literally unable to be moved elsewhere until unlocked. After locking certain case-affiliate matches (this essentially assigns $z^i_\ell = 1$ for family F^i and affiliate L^ℓ), any remaining unlocked cases may be rematched, with affiliate capacities adjusted down from any locked cases, via a color-coded gray reoptimize button that indicates the nonoptimized state (see Figure 9). Thus, any "final" matches can be locked, and all remaining cases can be rematched using the remaining available capacity.

We also enable cross-referencing. Cross-referencing occurs when refugee cases are linked to other cases that (a) have previously been resettled to a specific local affiliate or (b) are among the pool of cases that are presently to be resettled to the same affiliate (note that these are cases with U.S. ties, as previously described in Section 4). In either case, *Annie*™ visually depicts cases that are associated with (a) an affiliate or (b) other cases via unique yellow borders upon hovering over a large, boxed X icon, for associated case tiles. For any two cases i, i' that are cross-referenced, *Annie*™ sets $z^i_\ell = z^{i'}_\ell$ for all local communities L^ℓ; and if i, i' are cross-referenced to a particular local community ℓ', *Annie*™ sets $z^i_{\ell'} = z^{i'}_{\ell'}$ only for local community $L^{\ell'}$. Figure 10 depicts an example where two cases are cross-referenced not only to one another (e.g., adult siblings) but also to an affiliate.

If a case tile is moved into an affiliate but there is a lack of compatibility between this case and the new affiliate in terms of binary community support services, the background color of the case becomes red as an indication and an exclamation mark icon appears

Figure 9. (Color online) Locking Case Tiles and Reoptimizing

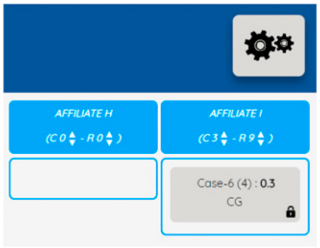


Figure 10. (Color online) Cross-Referencing Cases to Affiliate F

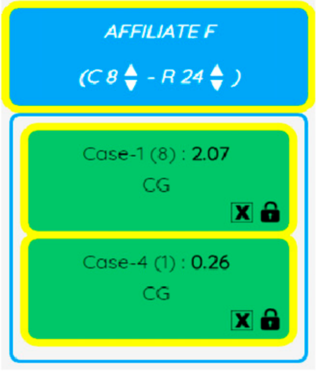
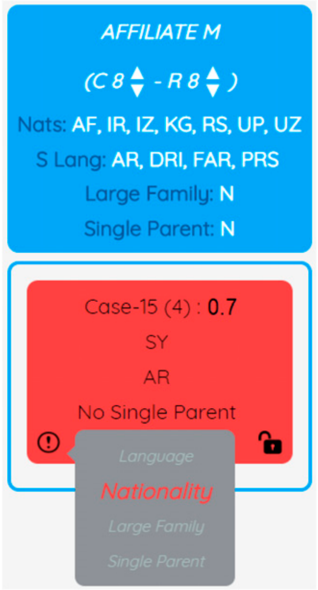


Figure 11. (Color online) Case Tile Changes Color When Placed into Affiliate that Violates Binary Service Constraints



Note. Hovering over exclamation point reveals additional details.

in the bottom left of the case tile (see Figure 11). Hovering over this exclamation mark icon displays a new list that shows the unsupported needs for that particular case-community match.

Throughout the development process, we have firmly maintained that *Annie*™ is a tool that augments the perspective of resettlement staff at HIAS. That is, matches generated by *Annie*™ are suggestive in nature. HIAS has complete discretion to match and rematch cases according to their expert judgment. In this way, we allow for the best of both worlds: leveraging the strengths of modern computational technology—machine learning, integer optimization, and interactive visualization—while arming human decision makers with all available information to facilitate the decision-making process.

7. Conclusion

Refugee resettlement is a complex humanitarian challenge that requires insights from a number of disciplines, including operations research, statistics, economics, political science, and sociology. Much work is urgently needed to improve the livelihoods of resettled refugees and the communities into which they integrate. In this paper, we show how combining tools from machine learning, integer optimization, and interactive visualization can improve refugee outcomes within the United States. We introduce the innovative software tool, *Annie*™ Moore, that assists the U.S. resettlement agency HIAS with matching refugees to their initial placements. Our software suggests optimal placements while giving substantial autonomy for the resettlement staff to fine-tune recommended matches. Because *Annie*™ matches on refugee employment outcomes, we expect refugees to more quickly integrate economically into each affiliate as well as make more productive economic and societal contributions, such as creating new jobs and generating tax revenues, benefitting local communities.

Annie™ has analytically enhanced the placement decision-making process at HIAS, having largely eliminated the inefficiencies of the former manual placement process. The operational process of placing refugees has improved considerably, enabling resettlement staff to place greater emphasis on cases that need greater attention, such as those with severe medical conditions.

Technological solutions, including machine learning and integer optimization, have enormous potential to help tackle humanitarian operations problems, such as placement optimization in refugee resettlement. Although the humanitarian sector offers many opportunities for impact, any solution must properly account for the severe lack of resources—including financial, labor, time, and data. These factors must be carefully considered in designing solutions, to afford the best opportunity of effecting change. Particular solution design features that we advocate include being lightweight, open-source, and designed with the end user in mind by incorporating important aspects of their regular operations.

There are several directions for further work. First, as is often the case in the humanitarian context, data have been difficult to obtain because of a severely resource-constrained environment. Indeed, data collection appears to be underprioritized across the resettlement agencies. We used the only existing outcome data from previous U.S. placements, namely, a refugee-specific binary indicator for employment measured 90 days after arrival. Although we went through great efforts to make the most out of the available data, the relative lack thereof necessarily hampered our prediction ability. Further work could apply our techniques to data on other outcomes, such as longer-term employment, physical and mental health, education, and household earnings. Unfortunately, at the time of writing, no data on these objectives for resettled refugees arriving in the United States appear to be systematically available. However, we anticipate to be able to better process other constraints like free-form text fields to discern whether refugees require medical accommodations, such as wheelchair access.

Second, although annual approved arrival capacities exist for affiliates, refugees arrive stochastically over the course of a year. Therefore, it is important to schedule the arrival of refugees given the partial information about future arrival over the course of the whole year. Andersson et al. (2018) tackle this problem in the Swedish context.

Third, it is interesting to consider which features of local areas offer the best potential to host refugees. For example, we could analyze to what extent local unemployment or community demographics affect refugee outcomes. This could help refugee agencies target areas for new affiliates.

Fourth, the social objective considered in this paper is to maximize employment. Even if it can be argued that "there is no single, generally accepted definition, theory or model of immigrant and refugee integration" (Castles et al. 2002, p.114), it is also clear that there are key aspects of integration beyond employment. Ager and Strang (2008), for example, argue that there are 10 established integration indicators, including health, housing, and education. This additional information—such as housing information, social networks, or new job opportunities—likely exists to at least some degree at the local community level and could prove very useful in supplementing the decision process. Moreover, regular and sustained engagement of local communities and associated stakeholders can also produce valuable insights that augment decision outcomes (see, e.g., Johnson et al. 2018). So, although 90-day binary employment outcomes are at present the only data available to estimate future integration outcomes, additional integration factors may be possible to integrate in the future; we thus leave the analysis of such models for future research.

Fifth, recent theoretical work on refugee matching (Jones and Teytelboym 2017b, Aziz et al. 2018, Delacrétaz et al. 2019, Olberg and Seuken 2019) suggests that preferences of refugees should be explicitly taken into account, because refugee families themselves know best where they would like to settle and where they are most likely to thrive. Refugee preferences could ideally be collected during the refugee prearrival orientation using a questionnaire that elicits how refugees might trade off features of areas (such as climate, urban versus rural, crime, amenities, and quality of schools). Unfortunately, refugee preferences are not elicited either by UNHCR or by the U.S. Department of State. In any case, the consideration of refugee preferences should be handled with care. Including preferences while optimizing for a particular observable outcome can in itself be problematic (Biró and Gudmundsson 2020), and it is also unclear how preferences should be elicited based on the reported information. Allowing refugees to report complete preferences may also be overly challenging. Hence, although it is clear that the approach we adopt has room for improvement, we believe it to be a reasonable approach in line with the growing evidence that the initial placement of refugee families greatly affects lifetime employment, which, in turn, profoundly alters lifetime welfare (Åslund and Rooth 2007, Damm 2014, Ferwerda and Gest 2017).

A final challenge with eliciting refugees' preferences—and a main theme in the book by Roth (2016)—is that agents often find it "unsafe" to report true information. Rather than strategically misrepresenting information to "game" the system, agents may be reluctant to report complete information simply because a lack of knowledge on how the information will used, how it will be spread, or trust that the reported information will be used in their best interest. This is surprisingly often the case even in applications where the outcome has a large impact on future welfare and life quality, such as in school choice (Abdulkadiroğlu and Sönmez 2003) and kidney exchange (Roth et al. 2004). It is in general difficult to design systems where all agents find it "safe" to report true and complete preferences (see, e.g., Roth 1982, Alcalde and Barberà 1994, Barberà and Jackson 1995, Sönmez 1999). Recent work on refugee matching with preferences also indicates that it can be difficult to design matching systems in which refugees have no incentive to misreport their preferences (e.g., Andersson and Ehlers 2017, Delacrétaz et al. 2019).

Annie™ has primarily been developed to assist HIAS in their initial refugee placements; there is significant potential to expand to additional resettlement contexts, both within the United States as well as beyond—the most direct being with other U.S. resettlement agencies that face analogous placement decision challenges. *Annie*™ could be used to help improve placement in the (Syrian) Vulnerable Persons Resettlement Scheme operated by the British government between 2015 and 2020. A recent report by the UK Independent Chief Inspector of Borders and Immigration recommended that the Home Office "improve the geographical matching process" of refugees in this resettlement scheme (Bolt 2018, p. 12). In Sweden, asylum seekers who enter are temporarily placed at Migration Board accommodation facilities in anticipation of either a deportation order or a residence permit. If a residence permit is granted, the legal responsibility for asylum seekers (such as finding housing and schooling) is transferred from the Migration Board facility to one of the 290 municipalities in Sweden (43,745 such transfers were made in 2018). This system is, in a sense, a version of refugee resettlement in which asylum seekers are resettled within Sweden. Although the current Swedish system is not based on sophisticated matching techniques, a recent report by the Swedish Government (SOU 2018) recommends that carefully designed optimization and matching systems should be adopted. (Indeed, *Annie*™ could be adapted for the Swedish context; the authors of this paper have already presented the first version of *Annie*™ at the Swedish Ministry of Finance for potential adoption.) Finally, *Annie*™ may, for example, be adapted for distributing asylum seekers who are currently at reception centers in host countries (such as Germany or the southern border of the United States).

Acknowledgments
The authors are grateful to many for helping make this research successful. Mike Mitchell and Karen Monken and other members of HIAS staff have patiently worked with the authors since early 2017 to build and implement *Annie*™. Emily Aiken provided invaluable software development assistance for an early prototype. Comments and suggestions from Barbara Day, David Delacrétaz, Pat Flaherty, Daniel Freund, Carina Hausladen, Scott Duke Kominers, Will Jones, Amin Khademi, Kayse Lee Maass, Karen Monken, Ariel Procaccia, Hillel Rapoport, Alvin Roth, Andrew Schaefer, Vince Slaugh, Tayfun Sönmez, and M. Utku Ünver significantly improved this paper.

References
Abdulkadiroğlu A, Sönmez T (2003) School choice: A mechanism design approach. *Amer. Econom. Rev.* 93(3):729–747.

Ager A, Strang A (2008) Understanding integration: A conceptual framework. *J. Refugee Stud.* 21(2):166–191.

Alcalde J, Barberà S (1994) Top dominance and the possibility of strategy-proof stable solutions to the marriage problem. *Econom. Theory* 4(3):417–435.

Andersson T, Ehlers L (2017) Assigning refugees to landlords in Sweden: Efficient, stable, and maximum matchings. *Scandinavian J. Econom.* 122(3):937–965.

Andersson T, Ehlers L, Martinello A (2018) Dynamic refugee matching. Technical report, Lund University, Sweden.

Åslund O, Fredriksson P (2009) Peer effects in welfare dependence quasi-experimental evidence. *J. Human Resources* 44(3):798–825.

Åslund O, Rooth DO (2007) Do when and where matter? Initial labour market conditions and immigrant earnings. *Econom. J. (London)* 117(518):422–448.

Åslund O, Östh J, Zenou Y (2010) How important is access to jobs? Old question, improved answer. *J. Econom. Geography* 10(3):389–422.

Åslund O, Edin PA, Fredriksson P, Grönqvist H (2011) Peers, neighborhoods, and immigrant student achievement: Evidence from a placement policy. *Amer. Econom. J. Appl. Econom.* 3(2): 67–95.

Aziz H, Chen J, Gaspers S, Sun Z (2018) Stability and Pareto optimality in refugee allocation matchings. *Proc. 17th Internat. Conf. Autonomous Agents MultiAgent Systems* (International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC), 964–972.

Bansak K, Ferwerda J, Hainmueller J, Dillon A, Hangartner D, Lawrence D, Weinstein J (2018) Improving refugee integration through data-driven algorithmic assignment. *Science* 359(6373): 325–329.

Barberà S, Jackson M (1995) Strategy-proof exchange. *Econometrica* 63(1):51–87.

Besiou M, Van Wassenhove LN (2020) Humanitarian operations: A world of opportunity for relevant and impactful research. *Manufacturing Service Oper. Management* 22(1):135–145.

Besiou M, Pedraza-Martinez AJ, Van Wassenhove LN (2018) OR applied to humanitarian operations. *Eur. J. Oper. Res.* 269(2): 397–405.

Biró P, Gudmundsson J (2020) Complexity of finding Pareto-efficient allocations of highest welfare. *Eur. J. Oper. Res.*, ePub ahead of print, March 13, https://doi.org/10.1016/j.ejor.2020.03.018.

Bolt D (2018) An inspection of the vulnerable persons resettlement scheme. Technical report, Independent Chief Inspector of Borders and Immigration, London, UK.

Cao W, Çelik M, Ergun Ö, Swann J, Viljoen N (2016) Challenges in service network expansion: An application in donated breastmilk banking in South Africa. *Socio-Econom. Planning Sci.* 53:33–48.

Castles S, Korac M, Vasta E, Vertovec S (2002) Integration: Mapping the field. Technical report, Centre for Migration and Policy Research and Refugee Studies Centre, University of Oxford, Oxford, UK.

Çelik M, Ergun Ö, Johnson B, Keskinocak P, Lorca Á, Pekgün P, Swann J (2012) Humanitarian logistics. Mirchandani PB, ed. New Directions in Informatics, Optimization, Logistics, and Production, INFORMS Tutorials in Operations Research (INFORMS, Cantonsville, MD), 18–49.

COIN-OR (2020) CBC user guide. Accessed March 18, 2020, http://www.coin-or.org/cbc.

Damm AP (2014) Neighborhood quality and labor market outcomes: Evidence from quasi-random neighborhood assignment of immigrants. *J. Urban Econom.* 79:139–166.

Delacrétaz D, Kominers SD, Teytelboym A (2019) Matching mechanisms for refugee resettlement. Technical report, University of Oxford, Oxford, UK.

Edin PA, Fredriksson P, Åslund O (2003) Ethnic enclaves and the economic success of immigrants—evidence from a natural experiment. *Quart. J. Econom.* 118(1):329–357.

Ferwerda J, Gest J (2017) Location, location: Refugee resettlement and employment outcomes in the United States. Technical report, Dartmouth College, Hanover, NH.

Fréville A (2004) The multidimensional 0–1 knapsack problem: An overview. *Eur. J. Oper. Res.* 155(1):1–21.

Gurobi (2020) *Gurobi Optimizer 9.0.0 Reference Manual* (Gurobi Optimization, Inc., Houston, TX).

Hifi M, Michrafy M, Sbihi A (2004) Heuristic algorithms for the multiple-choice multidimensional knapsack problem. *J. Oper. Res. Soc.* 55(12):1323–1332.

Johnson MP, Midgley G, Chichirau G (2018) Emerging trends and new frontiers in community operational research. *Eur. J. Oper. Res.* 268(3):1178–1191.

Jones W, Teytelboym A (2017a) The international refugee match: A system that respects refugees' preferences and the priorities of states. *Refugee Survey Quart.* 36(2):84–109.

Jones W, Teytelboym A (2017b) The local refugee match: Aligning refugees' preferences with the capacities and priorities of localities. *J. Refugee Stud.* 31(2):152–178.

Lichtenstein G, Puma J, Engelman A, Miller M (2016) The Refugee Integration Survey Evaluation Project (RISE) year five: Final report. Technical Report, Colorado Office of Economic Security, Denver, CO.

Martello S, Toth P (1980) Solution of the zero-one multiple knapsack problem. *Eur. J. Oper. Res.* 4(4):276–283.

Meignan D, Knust S, Frayret JM, Pesant G, Gaud N (2015) A review and taxonomy of interactive optimization methods in operations research. *ACM Trans. Interactive Intelligent Systems* 5(3): 1–43 (TiiS).

Mitchell S, O'Sullivan M, Dunning I (2020) PuLP: A linear programming toolkit for Python. Accessed March 18, 2020, https://github.com/coin-or/pulp.

Moraga JFH, Rapoport H (2014) Tradable immigration quotas. *J. Public Econom.* 115:94–108.

Nguyen T, Nguyen H, Teytelboym A (2019) Stability in matching markets with complex constraints. *Proc. 2019 ACM Conf. Econom. Comput.* (ACM, New York), 61–61 .

Olberg N, Seuken S (2019) Enabling trade-offs in machine learning-based matching for refugee resettlement. Technical report, University of Zurich, Zurich.

Patacchini E, Zenou Y (2012) Ethnic networks and employment outcomes. *Regional Sci. Urban Econom.* 42(6):938–949.

Pedraza-Martinez AJ, Van Wassenhove LN (2016) Empirically grounded research in humanitarian operations management: The way forward. *J. Oper. Management* 45(1):1–10.

Ronacher A (2020) Flask (Python microframework). Accessed March 18, 2020, https://palletsprojects.com/p/flask/.

Roth AE (1982) The economics of matching: Stability and incentives. *Math. Oper. Res.* 7(4):617–628.

Roth AE (2016) *Who Gets What—And Why: Understand the Choices You Have, Improve the Choices You Make* (William Collins, London, UK).

Roth AE (2018) Marketplaces, markets, and market design. *Amer. Econom. Rev.* 108(7):1609–1658.

Roth AE, Sönmez T, Ünver MU (2004) Kidney exchange. *Quart. J. Econom.* 119:457–488.

Slaugh VW, Akan M, Kesten O, Ünver MU (2016) The Pennsylvania Adoption Exchange improves its matching process. *Interfaces* 46(2):133–153.

Song Y, Zhang C, Fang Y (2008) Multiple multidimensional knapsack problem and its applications in cognitive radio networks. MILCOM 2008-2008 IEEE Military Comm. Conf. (IEEE, Piscataway, NJ), 1–7.

Sönmez T (1999) Strategy-proofness and essentially singled-valued cores. *Econometrica* 67(3):677–689.

SOU (2018) Ett ordnat mottagande–Gemensamt ansvar för etablering eller återvändande. *Statens Offentliga Utredningar, SOU.* 2018, report 22 (Stockholm).

Trapp AC, Teytelboym A, Ahani N, Andersson T (2018) Refugee resettlement via machine learning and integer optimization. *OR60 Annual Conf.–Keynote Papers Extended Abstr.* (The OR Society Lancaster University, Lancaster, UK), 21–26.

UNHCR (2019a) Global resettlement needs 2020. Technical report, United Nations High Commissioner for Refugees, Geneva, Switzerland.

UNHCR (2019b) Global trends: Forced displacement in 2018. Technical report, United Nations High Commissioner for Refugees, United Nations High Commissioner for Refugees/Headquarters.

U.S. Department of State (2020) Refugee arrivals by state and nationality as of March 6, 2020. Accessed March 18, 2020, http://www.wrapsnet.org/admissions-and-arrivals.

Narges Ahani is a doctoral student in data science at the Worcester Polytechnic Institute. Her primary research interests are in the areas of decision making under uncertainty, integer optimization, and machine learning, with applications in humanitarian operations.

Tommy Andersson is a professor of economics at Lund University and an affiliated professor of economics at Stockholm School of Economics. His is one of the founders of the Scandiatransplant Paired Kidney Exchange Program; his research interests are in the areas of mechanism design, market design, and game theory.

Alessandro Martinello is currently a senior data scientist at the Danish central bank but performed most of the work for this paper while at Lund University, Department of Economics. His area of expertise is applied microeconometrics and causal inference. He particularly enjoys applying machine learning tools in economic research.

Alexander Teytelboym is an associate professor at the Department of Economics, Tutorial Fellow at St Catherine's College, and a senior research fellow at the Institute for New Economic Thinking at the Oxford Martin School, University of Oxford. His research interests lie in market design and the economics of networks (as well as their applications to environmental economics and energy markets). He co-founded Refugees.AI, an organisation that is developing new technology for refugee resettlement.

Andrew C. Trapp is an associate professor of operations and industrial engineering, data science, and mathematical sciences at Worcester Polytechnic Institute. His primary research interests are mathematical optimization theory and its application in prescriptive analytics, such as stochastic, integer linear, and nonlinear optimization, in conjunction with predictive analytics, such as machine learning, to improve societal systems, in particular those that focus on vulnerable peoples.