

# 1 Implementation: Quotas and Refugee Misclassification

Our proposed mechanism is simple to implement given municipality-specific quotas and partitions  $A_m^+(x_i(k))$  and  $A_m^-(x_i(k))$ . The choice of these inputs has two practical implications. First, municipality preferences need to be estimated, and thus municipality-specific partitions will suffer from misclassification error in practical implementations. Second, an assignment agency can aggregate or split quotas over periods of time to affect whether smaller municipalities stays in the economy for multiple shorter spells, or for few, longer ones. This section shows that the algorithm outperforms current assignment strategies even in presence of severe misclassification error, and that with random refugee flows the choice of aggregating or splitting quotas is inconsequential for end-of-period fairness across heterogeneous municipalities.

## 1.1 Misclassification errors in municipality partitions

Estimating municipality-specific partitions is equivalent to predicting a demand matrix such as that shown in Table 1, using refugee characteristics as independent variables and desirable outcomes (such as future employment) as dependent variables. Bansak et al. (2018) use machine learning to estimate such a matrix with US and Swiss refugee data, allowing for municipality-specific synergies with heterogeneous refugees. Yet, out-of-sample prediction implies at least some misclassification error between the real and the estimated municipality-specific partitions. For example, Bansak et al. (2018) report 24% and 14% misclassification error for their preferred models estimated on US and Swiss data respectively.<sup>1</sup>

We assess the sensitivity of our algorithm to misclassification error by simulating random flows of refugees. For each of these simulated flows, we introduce increasingly high amount of misclassification error in the municipality-specific partitions, and compare the evolution of measures of fairness and efficiency through the dynamic refugee assignments. We benchmark the algorithm’s performance against a sequential assignment process, that assigns refugees according to their arrival order to rotating municipalities.<sup>2</sup> Such

---

<sup>1</sup>Note that misclassification error refers to observations wrongly—and not randomly—classified, and thus 50% misclassification error represents a zero-predictive-power scenario.

<sup>2</sup>A sequential assignment mechanism ensures that, in absence of quotas, each municipality receives a similar amount of refugees. Therefore, sequential assignment strategies tend to outperform truly random assignments.

an assignment is the current standard in Sweden, the United States and [Reference!].

We calibrate the flow generating process to match the observed average employment rates of refugees in the US and Sweden. To approximate realistic refugee flows, we simulate three groups of refugees. The first group of refugees ( $AA$ ) is considered acceptable by all municipalities. The second group of refugees ( $A$ ) is acceptable by some, but not all municipalities. These refugees thrive best in some environment and not on others, representing the empirical synergies documented by Bansak et al. (2018). We allow for autocorrelation of preferences across municipalities, so that each refugee of group  $A$  is either acceptable by a large majority or a small minority of municipalities. The third group of refugees ( $NA$ ) is the set of non-demanded refugees, whose assignment depends on  $\sigma(k)$ .

Our proposed algorithm exploits the synergies of refugees with heterogeneous municipalities to perform an efficient and fair assignment. Therefore, relative to the standard sequential assignment, it produces larger fairness and efficiency gains as the proportion of  $A$ -type refugees increases in the sample.<sup>3</sup> We approximate acceptability with employment, and calibrate the simulated refugee flows to minimize the proportion of  $A$ -type refugees given the results of Bansak et al. (2018) for the United States and the employment outcomes of the 2013 refugee wave in Sweden.

Figure 1 compares the performance of our algorithm with increasing degrees of misclassification error in the municipality-specific partitions with that of the standard sequential assignment, in terms of both fairness and efficiency. We use the proportion of municipalities in the sample envying at least another by one and five refugees as measures of fairness. If the municipality-specific partitions are correct, proof NUMBER shows that our algorithm guarantees no envy greater than one. Remarkably, Panel 1a also shows that the share of municipalities envying another by one decreases with the number of assigned refugees and converges towards.

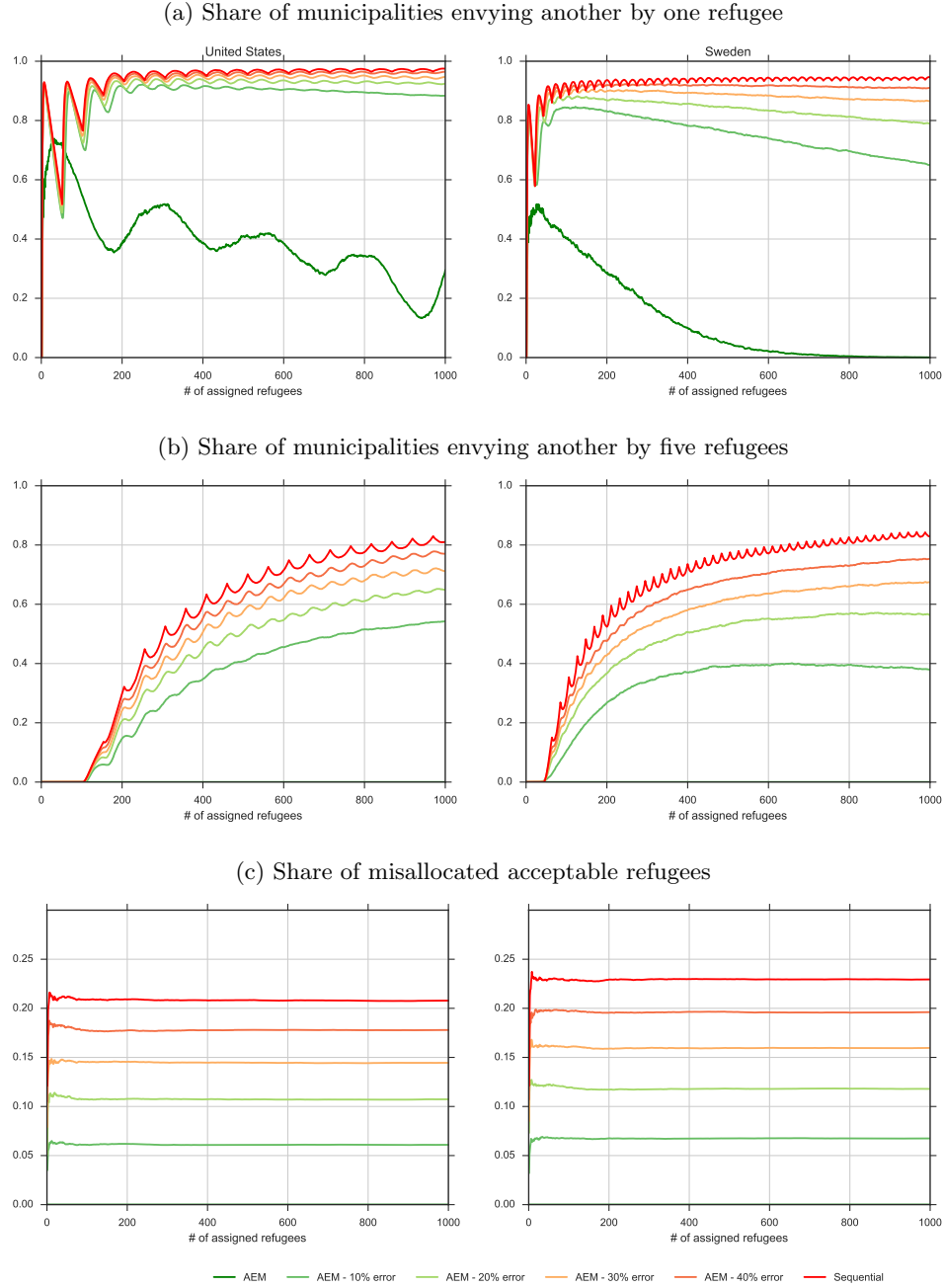
Even with misclassification error, Figure 1 shows that our assignment mechanism always outperforms the standard sequential assignment in terms of fairness. With the 24% misclassification error of Bansak et al. (2018), our algorithm guarantees increases in fairness of between  $N$  and  $N\%$ .

We measure efficiency as the share of refugees that, while acceptable for some municipalities, end up in a municipality that considers them not

---

<sup>3</sup>For example, in the extreme case in which all refugees are always accepted ( $AA$ ), the sequential assignment also produces pareto-efficient results. The same applies to the case in which all refugees are never accepted ( $NA$ )

Figure 1: Algorithm's sensitivity to misclassification error



NOTES:—We simulate a thousand refugee waves of a thousand refugee each, distributed along 51 locations in the United States and 21 in Sweden. We compute measures of efficiency and equity at each refugee assignment<sup>3</sup>, comparing the performance of our algorithm with increasing degrees of misclassification error with that of the standard sequential assignment.

acceptable. This measure reflects the preferences of refugees of achieving the best possible outcomes for their future, and in a Pareto-efficient assignment is always equal to zero. Even with misclassification error, our algorithm produces strong and stable efficiency gains with respect to the standard sequential assignment. With 24% misclassification error, our algorithm decreases the share of misallocated acceptable refugees by X%.

## 1.2 Choice of quotas

The presence of heterogeneous quotas across geographical areas introduces a further seasonality-related issue. Each area has the obligation of reserving housing units to asylum seekers. With yearly quotas, smaller areas will fill up all their housing obligation in a few months, and thus will be obliged to provide housing units only at a given time of the year. Larger municipalities will instead have to provide housing throughout the year. A solution to this disparity would be splitting yearly quotas in e.g. trimester quotas, and dynamically assign refugees within each trimester. this section shows that this strategy is feasible, as with random refugee flows the losses in fairness and efficiency due to splitting are minimal.

We simulate a thousand random refugee flows calibrated as in Figure 1 for Sweden, and allocate them to the 21 Swedish regional entities (län) using 2017 quotas.<sup>4</sup> For each refugee flow, we split quotas at yearly, semestral, trimestral and monthly interval. At each interval the assignment algorithm restarts either blindly (i.e. resetting the envy matrix at zero at the start of the period) or taking into account dynamic legacies (i.e. initiating the algorithm with the computed envy matrix at the end of the previous period). We compute measures of fairness and efficiency once the full refugee flow has been assigned, and benchmark our results to those obtained through a standard sequential assignment strategy.

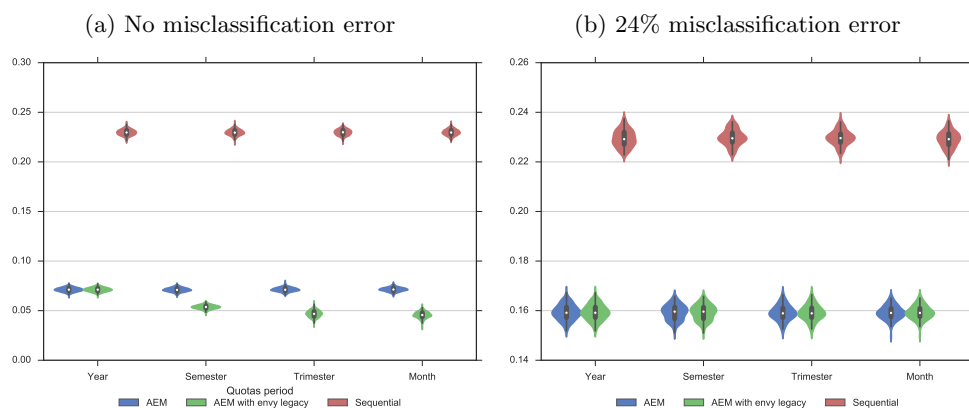
Figure 2 shows

- Show distribution of shares at refugee N - outperforms sequential - if allow for legacy, smaller is better - intuition: If flows of refugee random, smaller areas "in the game" for longer the smaller the quotas periods are
- results for fairness similar with or without legacy, results appear in the appendix.

---

<sup>4</sup>In 2017, the sum of all quotas was 23600. For computational simplicity we round each regional quota to a number divisible by 12, and we thus simulate flows of 23568 refugees each.

Figure 2: Share of misallocated acceptable refugees, by assignment and quotas split



NOTES:—We ...

## References

**Bansak, Kirk, Jeremy Ferwerda, Jens Hainmueller, Andrea Dillon, Dominik Hangartner, Duncan Lawrence, and Jeremy Weinstein,** “Improving refugee integration through data-driven algorithmic assignment,” *Science*, January 2018, *359* (6373), 325–329.