

Policy Learning by a Magician

Kei Ikegami*

Abstract

We study eligibility learning for sequential programs under an *exact capacity constraint*, which limits treatments to available slots, and *arrival neutrality*, which gives eligible applicants a common treatment probability across arrival orders. We propose *magician-adjusted empirical welfare maximization*, weighting estimated gains by treatment probabilities from the *conservative magician* of Alaei (2012). Per-applicant regret relative to the best rule under this magician decreases at the inverse-square-root rate in experimental sample size. Under certain conditions, we show that the best magician implementation retains nearly all of the welfare attainable under an expected-capacity constraint as capacity grows.

Keywords: treatment choice; empirical welfare maximization; capacity constraints; online contention resolution; prophet inequalities.

1 Introduction

Programs with limited treatment capacity can use experimental or observational data to learn how to allocate scarce slots among applicants who differ in their gains from treatment (Manski, 2004; Kitagawa and Tetenov, 2018; Athey and Wager, 2021). A standard formulation of such a *policy learning* problem chooses a treatment rule to maximize estimated welfare subject to a constraint on expected cost, or, when each treatment uses one slot, the expected number selected for treatment (Bhattacharya and Dupas, 2012). Even when this expectation is below capacity, however, the number selected can exceed the available slots in a particular application: for example, a program with fifty training places may expect forty applicants to qualify but receive sixty eligible applications. So, in practice, the welfare the program can finally deliver also depends on how it rations treatment.

If a program can wait until all applications have arrived, it can allocate its slots after observing the entire pool of eligible applicants. Some programs, however, must commit resources before all applications are known. For example, an employment program with a fixed number of hiring subsidies may need to decide whether to support a job seeker before her current job offer expires, even though other job seekers have yet to apply. In such settings, first-come-first-served respects capacity but makes late eligible applicants bear the risk that treating earlier eligible applicants exhausts the available slots. For programs that regard eligible applicants as having an equal claim to treatment, this gives arrival time an unintended role in determining access. The problem is therefore to give every eligible applicant the same probability of treatment while allocating slots sequentially and never exceeding capacity.

We study how to learn eligibility rules under two requirements on sequential treatment allocation. The number treated must never exceed capacity on any realization; we call this an

*Yokohama City University. Email: ikegami.kei0120@gmail.com.

exact capacity constraint. Every eligible applicant must have the same probability of treatment, regardless of their characteristics or arrival order; we call this requirement *arrival neutrality*. This probability is conditional on the applicant’s own characteristics and averages over earlier applicants’ characteristics and the program’s randomization for each fixed order.

Section 2 formulates the choice of whom to make eligible and how to allocate treatment as applicants arrive. With unit treatment costs, the expected-budget approach plans to treat everyone eligible, keeping their expected number within capacity (Sun, 2026). For a candidate eligibility rule h in our setting, let $B(h)$ be the largest expected number eligible and $J(h)$ the smallest expected gain from treating them all, over the possible combinations of applicant counts across populations. With k slots, this gives

$$\max_{h: B(h) \leq k} J(h).$$

Our approach instead limits actual treatments and chooses a common treatment probability γ for eligible applicants:

$$\sup_{h, \gamma} \gamma J(h),$$

subject to the existence of a sequential allocation rule that treats only eligible applicants, never exceeds capacity, and gives each eligible applicant probability γ in every admissible arrival order.

Section 3 introduces the *conservative magician* of Alaei (2012) as a tool for implementing a common treatment probability under exact capacity. For independent eligibility indicators with known probabilities, its guarantee can be chosen using capacity and a bound on the sum of those probabilities, without forecasting their order. Building on Jiang et al. (2025), we characterize the largest common probability $\Gamma_k(B)$ that can be guaranteed uniformly over all such sequences with total probability at most B , including $B > k$.

Section 4 applies this guarantee to each eligibility rule: using its known eligibility probabilities, the magician implements $\gamma = \Gamma_k(B(h))$, making welfare $\Gamma_k(B(h))J(h)$. Eligibility is therefore chosen to maximize this product within the specified magician family, allowing $B(h) > k$. We also bound its welfare cost relative to optimal online allocation under an expected-capacity constraint with the same eligibility class and arrival information. If that optimum is positive and is attained by treating everyone eligible under a candidate rule with $B(h) \leq k$, the best magician implementation retains nearly all of it as capacity grows (Proposition 1).

Section 5 introduces *magician-adjusted empirical welfare maximization (EWM)*, which multiplies each rule’s estimated gain from treating all eligible applicants by its known common treatment probability and selects a maximizer. With known covariate distributions, the procedure respects exact capacity and arrival neutrality for every training sample. For a finite candidate class fixed before observing the data, we bound its welfare loss relative to the best rule under the same magician with high probability. With N experimental observations per population, the bound per applicant retains the usual $N^{-1/2}$ rate of EWM (Kitagawa and Tetenov, 2018), with no additional dependence on the number of applicants or treatment slots.

Section 6 uses a simulated job-training program with applicants from two regions to separate learning losses from implementation costs. With 64,000 experimental observations per region, magician-adjusted EWM achieves mean welfare of about 99.7 percent of the best welfare under the same numerical magician implementation. Taking treatment effects as known, we distinguish the costs of exact capacity and arrival neutrality and find that this numerical optimum retains

76.5 percent of the expected-capacity optimum with five slots and 97.1 percent with five hundred. In this design, more data improve eligibility choice, while greater capacity reduces the welfare cost of implementation.

Related literature. The paper contributes to statistical treatment choice (Manski, 2004; Hirano and Porter, 2009; Stoye, 2009; Kitagawa and Tetenov, 2018; Athey and Wager, 2021; Mbakop and Tabord-Meehan, 2021), and in particular to work on resource constraints. Bhattacharya and Dupas (2012) characterize optimal assignment under a budget, Kitagawa and Tetenov (2018) discuss capacity constraints enforced by random rationing, and Luedtke and van der Laan (2016) learn individualized treatments under a constraint on the treated proportion. Sun (2026) shows that uniform welfare efficiency and feasibility cannot be achieved simultaneously when an estimated expected budget binds. She proposes a trade-off rule that penalizes overspending. Our setting enforces capacity on every realization and requires equal treatment probabilities across eligible applicants. With known covariate distributions, the magician supplies a known treatment probability for each eligibility rule, which enters the welfare objective as a multiplicative adjustment.

Related work also studies treatment decisions over time. Sakaguchi (2025) studies dynamic treatment regimes under policy constraints, in which the same individuals are treated over several stages. Adusumilli et al. (2024) study dynamically optimal allocation of a budget across sequentially arriving individuals under a model of the arrival process. We instead assume no forecast of the order, restrict attention to arrival-neutral rules, and allow applicants from different populations to arrive in any order.

Fairness constraints in policy learning are studied by Viviano and Bradic (2024) and Kock and Preinerstorfer (2024); arrival neutrality is a fairness requirement across arrival times. Lee et al. (2026) study randomized priority queues in housing allocation, and Haghi et al. (2026) learn allocation policies under long-run capacity constraints.

The magician belongs to the literature on online contention resolution and prophet inequalities (Alaei, 2012; Feldman et al., 2021; Jiang et al., 2025). With exact eligibility probabilities as inputs, the conservative magician gives every eligible applicant the same treatment probability at every arrival position. Feldman et al. (2025) study online contention resolution with sample-based knowledge of the probabilities. Arsenis and Kleinberg (2022) study identity- and time-independent fairness in prophet inequalities, and Manshadi et al. (2021) maximize the minimum expected fill rate across sequentially arriving agents. These papers take the relevant distributions as known or learn about them from samples, without learning eligibility rules from experimental data. We characterize the guarantee curve for any bound on the expected number of eligible applicants and use it to learn eligibility rules from treatment-effect data.

2 Setting

Applicants to a program may come from offices or regions with different distributions of characteristics and treatment effects. The program may know the distribution of characteristics at each office, and how many applicants it will send, without knowing the order in which applicants from different offices will arrive. We first describe this arrival information and the decisions about eligibility and treatment. We then require treatment probabilities to be equal

across eligible applicants and formulate the planner's choice of eligibility, a common treatment probability, and an implementation that respects capacity.

2.1 Applicants, environments, and arrival orders

A program has $k \geq 1$ treatment slots and serves a deployment of $n \geq k$ applicants who arrive one at a time. Applicant $i \in \{1, \dots, n\}$ belongs to an observable environment $s_i \in \mathcal{S} = \{1, \dots, S\}$, such as an office, a region, an application channel, or a period of the application window. She has covariates $X_i \in \mathcal{X}$ and potential outcomes $Y_i(1), Y_i(0) \in [0, 1]$; write $\Delta_i = Y_i(1) - Y_i(0)$. Each treatment uses one slot. Decisions are made upon arrival and cannot be revised.

Assumption 1 (Arrivals). (i) The environment sequence $\sigma = (s_1, \dots, s_n)$ is fixed before the deployment starts. It is unknown to the planner and belongs to the set $\Sigma(\mathcal{C})$ of sequences whose environment counts $c(\sigma) = (\#\{i : s_i = s\})_{s \in \mathcal{S}}$ lie in a known nonempty set $\mathcal{C} \subseteq \{c \in \mathbb{N}_0^S : \sum_s c_s = n\}$. No probability distribution over $\Sigma(\mathcal{C})$ is specified. (ii) Conditional on σ , the vectors $(X_i, Y_i(0), Y_i(1))$, $i = 1, \dots, n$, are independent, and the vector of applicant i has distribution P_{s_i} , which does not depend on i or on σ . (iii) The planner knows the marginal distribution P_s^X of applicant characteristics in each environment. The conditional distributions of potential outcomes are not assumed known.

The planner observes s_i when applicant i arrives. Part (i) allows the order to be chosen adversarially, but not in response to the program's realized decisions. Part (ii) requires the distribution of applicants within an environment to be stable. Taking $\mathcal{C}$ to be all compositions summing to n imposes no restriction on composition; a smaller set encodes institutional knowledge, such as quotas by office. Part (iii) separates information about who applies from information about what treatment achieves: the distribution of characteristics is known, but treatment gains must be learned from data. It does not reveal future applicants' realized characteristics. We maintain this assumption in the main text to focus on introducing the magician to policy learning. Appendix E studies the case in which the covariate distributions are unknown and must also be learned from data.

2.2 Eligibility rules and implementation rules

An *eligibility rule* is a collection $h = (h_s)_{s \in \mathcal{S}}$ of functions $h_s : \mathcal{X} \rightarrow \{0, 1\}$. Applicant i is eligible under h if $Z_i := h_{s_i}(X_i) = 1$. The planner chooses h from a class $\mathcal{H}$. The class contains the empty rule, $h_s \equiv 0$ for all s , under which no applicant is eligible.

An *implementation rule* M for h decides online whom to treat. When applicant i arrives, M observes $(s_1, \dots, s_i)$, the covariates and decisions of earlier applicants, and X_i . It then chooses $A_i \in \{0, 1\}$, possibly at random, subject to

$$A_i \leq Z_i \quad \text{and} \quad \sum_{j=1}^n A_j \leq k \quad \text{on every realization.}$$

Thus only eligible applicants are treated, and the capacity is never exceeded. The rule does not observe the outcome of applicant i when deciding on her. Let $W_i = \sum_{j < i} A_j$ be the number of slots used before applicant i arrives.

2.3 Arrival neutrality

First-come-first-served satisfies the capacity constraint by treating eligible applicants until all slots are used. It leaves later eligible applicants with a smaller chance of treatment. We instead require the program to give every eligible applicant the same treatment probability, whatever her position and whatever the admissible arrival order. The requirement concerns probabilities before earlier applicants and the program's randomization have been realized.

Definition 1. An implementation rule M for h is *arrival-neutral at level* $\gamma \in [0, 1]$ on $\Sigma(\mathcal{C})$ if, for every $\sigma \in \Sigma(\mathcal{C})$, every position i , and $P_{s_i}^X$ -almost every x with $h_{s_i}(x) = 1$,

$$\mathbb{P}(A_i = 1 \mid X_i = x; \sigma) = \gamma.$$

The probability in Definition 1 averages over the applicants who arrived earlier and over the rule's internal randomization. It is the probability of treatment of an eligible applicant who knows her position and environment but not the realized history. A guarantee conditional on every realized history is impossible: once all slots are used, no later applicant can be treated. Arrival neutrality requires that this ex ante probability be the same for every eligible applicant, whatever her position and whatever the order of arrivals.

2.4 The planning problem

The planner's objective is maximizing welfare. The welfare of M along σ is the expected total gain relative to treating nobody,

$$V_\sigma(h, M) = \mathbb{E} \left[\sum_{i=1}^n A_i \Delta_i \mid \sigma \right], \quad V(h, M) = \min_{\sigma \in \Sigma(\mathcal{C})} V_\sigma(h, M).$$

The worst case over arrival sequences reflects the absence of a forecast of the order. We seek an eligibility rule whose implementation delivers high welfare while satisfying capacity and arrival neutrality.

Under arrival neutrality, welfare separates into the gains from treating all eligible applicants and their common probability of treatment. For each environment, define the expected gain per applicant if every eligible applicant were treated by

$$\mu_s(h) := \mathbb{E}_{P_s} [h_s(X) \{Y(1) - Y(0)\}].$$

The smallest expected total gain over the admissible compositions is

$$J(h) := \min_{c \in \mathcal{C}} \sum_{s \in \mathcal{S}} c_s \mu_s(h). \tag{1}$$

This quantity evaluates eligibility before any rationing. If an implementation gives every eligible applicant probability γ of treatment, it delivers the fraction γ of these gains. The following identity holds for any arrival-neutral implementation; its proof is in Appendix A.

Lemma 1 (Welfare under arrival neutrality). *If an implementation rule M for h is arrival-neutral at level γ on $\Sigma(\mathcal{C})$, then*

$$V_\sigma(h, M) = \gamma \sum_{i=1}^n \mu_{s_i}(h) \quad \text{for every } \sigma \in \Sigma(\mathcal{C}), \quad V(h, M) = \gamma J(h).$$

The planner therefore chooses eligibility and the common treatment probability jointly, together with a rule that implements them. The population planning problem is

$$\begin{aligned} & \sup_{h \in \mathcal{H}, \gamma \in [0,1], M} \gamma J(h) \\ & \text{subject to } M \text{ is an online implementation rule for } h, \\ & \quad M \text{ is arrival-neutral at level } \gamma \text{ on } \Sigma(\mathcal{C}). \end{aligned} \tag{2}$$

The first constraint incorporates the information available at each arrival, treatment only of eligible applicants, and capacity on every realization. The second requires one common treatment probability across eligible covariates, positions, and admissible orders. Thus a choice of (h, γ) is feasible only if such an implementation exists. The level γ is a choice variable; it has not yet been tied to a particular allocation procedure.

To make progress on (2), we need to construct implementations and determine which treatment probabilities they can deliver. Section 3 introduces the conservative magician as a tool for this task. Section 4 then uses it to construct feasible choices for the planning problem and selects the best eligibility rule within that family of implementations. The gains $J(h)$ are unknown; Section 5 studies how to learn that rule from experimental data.

3 The magician

The planning problem (2) requires an online implementation that gives eligible applicants a common treatment probability while respecting capacity. The conservative magician provides a tool for constructing such an implementation. We describe it in its original setting before applying it to treatment allocation in Section 4.

Following Alaei (2012), consider a magician with k magic wands who is presented with a sequence of boxes. One box contains a prize, but the magician does not know which one. Using a wand opens a box, but the wand may break; if it survives, it can be used again. Each box states the probability that a wand will break when used to open it. The magician must decide whether to open or skip the current box before seeing the next one. Once all wands have broken, no further box can be opened. Since the prize could be in any box, the magician seeks to make the smallest probability of opening a box as large as possible.

We first describe the conservative magician and the guarantees established by Alaei (2012) and Jiang et al. (2025). We then characterize the largest probability of opening each box that can be guaranteed from the number of wands and a bound on the sum of the break probabilities. This guarantee can be chosen before the future boxes are known. Proofs of the results derived here are in Appendix B.

3.1 The conservative magician and existing guarantees

Fix an integer $k \geq 1$ and a finite sequence of boxes with labels $x = (x_1, \dots, x_n) \in [0, 1]^n$, where x_i is the probability that a wand breaks when used on box i . The entire sequence, including the prize's location, is fixed in advance. The magician sees x_i when box i arrives but need not know the number or labels of future boxes. We take the labels to be the exact break probabilities and the potential breaks to be independent across boxes. Alaei (2012) also permits labels that are upper bounds on these probabilities; exact labels are the case used here.

Let $G_i \in \{0, 1\}$ indicate whether the magician opens box i , and let $Z_i \sim \text{Bernoulli}(x_i)$ indicate whether a wand would break if used on that box. The magician chooses G_i before learning Z_i and observes a break outcome only if the box is opened. A strategy may use the labels seen so far, previous opening decisions and their outcomes, and independent randomization. The actual break indicator is $A_i = G_i Z_i$, and $W_i := \sum_{j < i} A_j$ counts the wands already broken. Opening a box requires a surviving wand, so

$$G_i \leq \mathbf{1}\{W_i \leq k - 1\}, \quad A_i = G_i Z_i, \quad \sum_i A_i \leq k \quad \text{on every realization.}$$

The last inequality follows because each opening can break at most one wand.

A strategy has common opening probability γ if $\mathbb{P}(G_i = 1) = \gamma$ for every box i . An opening-probability *floor* γ requires these probabilities to be at least γ . The probabilities average over earlier break outcomes and the magician's randomization, for the fixed sequence of boxes. A floor γ guarantees a probability of at least γ of finding the prize wherever it is hidden. The requirement includes boxes with $x_i = 0$: opening such a box cannot break a wand, but still requires one.

Opening every box while a wand remains exposes later boxes to the risk that all wands have already broken. The conservative magician of Alaei (2012) sometimes skips a box to preserve that opportunity. Its opening decision depends on how many wands have broken: it opens when this number is small, skips when it is large, and randomizes at a threshold. The threshold is chosen so that the box is opened with the prescribed probability γ .

To compute this threshold, the magician maintains a *reference distribution* π_i for the number of broken wands before box i , starting from the point mass π_1 at 0. For a level $\gamma \in (0, 1]$, it assigns opening probability to total mass γ , starting with the states having the fewest broken wands. With $\theta_i = \min\{w : \pi_i(\{0, \dots, w\}) \geq \gamma\}$, the opening probabilities at each state are

$$\varphi_i(w) = \begin{cases} 1, & w < \theta_i, \\ \frac{\gamma - \pi_i(\{0, \dots, \theta_i - 1\})}{\pi_i(\theta_i)}, & w = \theta_i, \\ 0, & w > \theta_i, \end{cases} \quad \text{so that} \quad \sum_w \pi_i(w) \varphi_i(w) = \gamma.$$

Opening a box moves the count up by one with probability x_i . The reference distribution therefore updates as

$$\pi_{i+1}(w) = \pi_i(w) \{1 - x_i \varphi_i(w)\} + \pi_i(w - 1) x_i \varphi_i(w - 1), \quad \pi_i(-1) = 0.$$

This recursion depends only on the labels and γ . It is defined without imposing the limit of k broken wands. The level is *feasible* on x if the reference opening rule never calls for opening a box after all wands have broken: equivalently, $\theta_i \leq k - 1$ for every box i .

Definition 2 (γ -conservative magician; Alaei, 2012). At box i , the magician computes φ_i from the labels seen so far and the given level γ . If a wand remains, it opens the box with probability $\varphi_i(W_i)$, using an independent random draw before the break outcome is known. Otherwise it skips the box. At level $\gamma = 0$, it skips every box.

With exact labels, feasibility makes π_i the actual distribution of broken wands. The magician then opens every box with probability exactly γ , and every prescribed opening can be carried out with a surviving wand. This is the exact-label property of the construction in Alaei (2012, Theorem 4 and Section 7). The number of boxes opened can exceed k : a wand is lost only when it breaks.

The guarantee of Alaei (2012) supplies a feasible level whenever the sum of the break probabilities does not exceed the number of wands. This level can be chosen before any boxes arrive.

Theorem 1 (Capacity guarantee; Alaei, 2012, Theorem 4). *Let $x \in [0, 1]^n$ satisfy $\sum_i x_i \leq k$. For every $0 < \gamma \leq 1 - (k+3)^{-1/2}$, the γ -conservative magician is feasible on x . With independent potential breaks and exact labels, it opens every box with probability exactly γ using at most k wands.*

For a particular sequence of boxes, the best feasible level may exceed this bound. Write $\gamma_k^*(x)$ for the supremum of the levels at which the conservative magician is feasible on x . For positive labels, the fixed-sequence optimality result of Jiang et al. (2025) implies that optimizing this level also maximizes the smallest opening probability among all online opening strategies with the same number of wands.

Theorem 2 (Fixed-sequence optimality; Jiang et al., 2025, Theorem 1). *Let $x \in (0, 1]^n$ satisfy $\sum_i x_i \leq k$, with independent potential breaks and exact labels. The largest opening-probability floor attainable by an online strategy with k wands is attained by the conservative magician with an optimally chosen level. This level is $\gamma_k^*(x)$, and the magician opens every box with that same probability.*

3.2 The magician's guarantee curve

The value $\gamma_k^*(x)$ evaluates the best level for a particular sequence of boxes. We now ask which opening probability can be guaranteed for every sequence when only the number of wands and a bound on the sum of the break probabilities are given. For a bound $B \geq 0$, the possible sequences have arbitrary finite length and satisfy $\sum_i x_i \leq B$. The largest uniformly feasible level is the worst-case value of $\gamma_k^*(x)$ over this set. Define

$$\Gamma_k(B) := \inf \left\{ \gamma_k^*(x) : L \in \mathbb{N}, x \in [0, 1]^L, \sum_i x_i \leq B \right\}. \quad (3)$$

We call Γ_k the *guarantee curve*. It uses only k and B ; restrictions on the possible sequences of boxes can permit a higher opening probability. The following result establishes feasibility

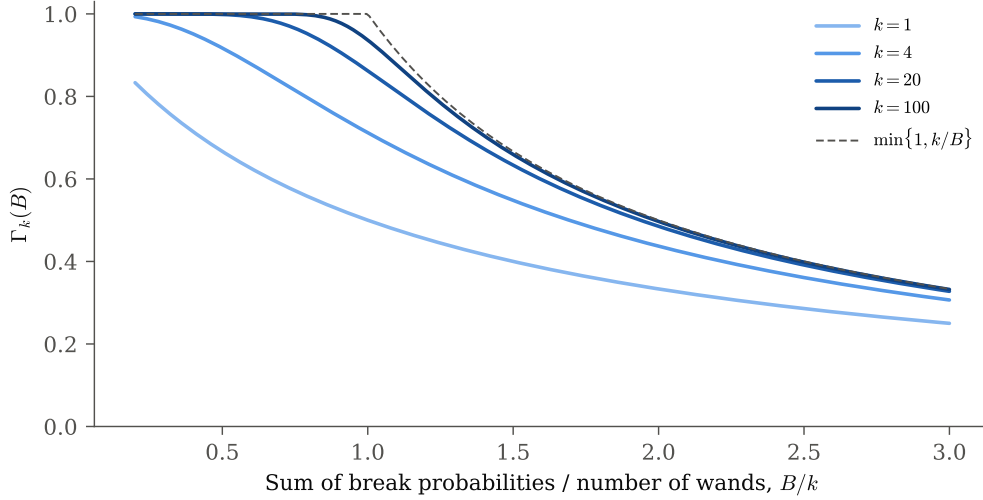


Figure 1: The guaranteed probability of opening each box, $\Gamma_k(B)$, as a function of B/k : the bound on the sum of break probabilities divided by the number of wands. The dashed line is the limit $\min\{1, k/B\}$ as $k \rightarrow \infty$ (Proposition B.2(iii)).

at $\Gamma_k(B)$ and shows that no online opening strategy can uniformly guarantee more. The comparison argument and the extension of fixed-sequence optimality, including zero labels, are in Appendix B.2.

Corollary 1 (Uniform guarantee for any bound B). *Let $B \geq 0$. On every finite sequence of boxes whose break probabilities sum to at most B , the $\Gamma_k(B)$ -conservative magician is feasible and, with independent potential breaks and exact labels, opens every box with probability $\Gamma_k(B)$ using at most k wands. No online opening strategy guarantees a larger opening-probability floor on all such sequences.*

The following Poisson characterization makes $\Gamma_k(B)$ computable. This extends the characterization of Jiang et al. (2025, Theorem 2) from $B = k$ to any $B > 0$. The proof is in Appendix B.3, and numerical details are in Appendix F.

Theorem 3 (Poisson characterization for any bound B). *For $\gamma \in [0, 1]$, let $W_\gamma(t)$ be the unconstrained reference count of broken wands, starting at zero, for the conservative magician at level γ driven by potential breaks from a Poisson process with rate one, as defined in Appendix B.3. For every $B > 0$,*

$$\Gamma_k(B) = \max \left\{ \gamma \in [0, 1] : \mathbb{P}\{W_\gamma(t) < k\} \geq \gamma \text{ for every } t \in [0, B] \right\}.$$

The feasible levels are equivalently characterized by the k -dimensional differential system (B.9).

Figure 1 plots the guarantee curve computed using Theorem 3. For fixed k , the guarantee is nonincreasing in B ; as the number of wands grows, it approaches $\min\{1, k/B\}$. The detail properties are formally established in Proposition B.2.

4 Eligibility Choice and Implementation

We now use the magician to construct feasible choices for the planning problem (2). For each eligibility rule h , we identify the magician’s inputs and choose a common treatment probability

that it can implement while respecting capacity. We then choose eligibility to maximize welfare within this family of implementations. Proofs are in Appendix C.

4.1 Implementing an eligibility rule

Fix an eligibility rule h . Associate each applicant with a box and each treatment slot with a wand. Opening a box permits treatment if the applicant is eligible: writing G_i for the box-opening indicator, set $A_i = G_i h_{s_i}(X_i)$, so each treatment breaks one wand. The label for an applicant from environment s is her eligibility probability,

$$p_s(h) := P_s\{h_s(X) = 1\} = \int h_s(x) dP_s^X(x),$$

which is known because the covariate marginal P_s^X is known. For applicant i , use $x_i = p_{s_i}(h)$ and the potential break indicator $h_{s_i}(X_i)$; these indicators are independent conditional on any fixed environment sequence σ by Assumption 1.

To choose a level that is feasible for every admissible sequence, define the largest expected number of eligible applicants over the admissible compositions,

$$B(h) := \max_{c \in \mathcal{C}} \sum_{s \in \mathcal{S}} c_s p_s(h). \quad (4)$$

For every $\sigma \in \Sigma(\mathcal{C})$, the labels satisfy $\sum_{i=1}^n p_{s_i}(h) = \sum_{s \in \mathcal{S}} c_s(\sigma) p_s(h) \leq B(h)$. We therefore run the magician at level $\gamma(h) := \Gamma_k(B(h))$.

By Corollary 1, this implementation respects hard capacity and satisfies $\mathbb{P}(G_i = 1 \mid \sigma) = \gamma(h)$ at every position in every admissible sequence. The opening decision uses the current environment and past information, but not the current applicant's covariates, so G_i is independent of X_i conditional on σ . Hence, for almost every x with $h_{s_i}(x) = 1$,

$$\mathbb{P}(A_i = 1 \mid X_i = x; \sigma) = \mathbb{P}(G_i = 1 \mid \sigma) = \gamma(h).$$

Denote this arrival-neutral implementation by M_h . The triple $(h, \gamma(h), M_h)$ is therefore feasible in (2). Within this family of implementations, it remains to choose the eligibility rule, which determines both who is eligible and their common treatment probability.

4.2 Welfare and eligibility choice

We compare eligibility rules by the welfare they deliver under these implementations. By Lemma 1,

$$V^{\text{mag}}(h) := V(h, M_h) = \gamma(h) J(h) = \Gamma_k(B(h)) J(h).$$

Here $J(h)$ is the gain from treating all eligible applicants in the least favorable composition, defined in (1). Restricting (2) to the constructed triples $(h, \gamma(h), M_h)$ gives the eligibility-choice problem

$$\sup_{h \in \mathcal{H}} \Gamma_k(B(h)) J(h). \quad (5)$$

This selects the best rule under the specified magician implementation.

This objective accounts for the effect of eligibility on the probability of treatment. Expected-budget EWM instead maximizes $J(h)$ subject to $B(h) \leq k$. Problem (5) permits $B(h) > k$ because the magician enforces capacity on actual treatments.

To distinguish the costs of enforcing capacity and arrival neutrality, we compare three planning problems that differ only in these requirements. All three choose $h \in \mathcal{H}$ and an online assignment rule M that decides which eligible applicants to treat, and maximize the same worst-order welfare $V(h, M)$. They retain the information and eligibility restrictions of Section 2, with capacity imposed as specified below. Treatment gains are taken as known.

Write V^{exp} for the supremum of welfare when $\mathbb{E}[\sum_i A_i \mid \sigma] \leq k$ is required for every $\sigma \in \Sigma(\mathcal{C})$. Write V^{cap} for the supremum when $\sum_i A_i \leq k$ must instead hold on every realization. Finally, let V^{AN} impose both this exact cap and arrival neutrality at some common level γ . Thus V^{AN} is the value of the general planning problem (2). The specified magician family has value $V_*^{\text{mag}} := \sup_{h \in \mathcal{H}} V^{\text{mag}}(h)$, which can be smaller than V^{AN} .

When these values are positive, define the *price of exact cap*, the *price of arrival neutrality*, and the *price of magician*, respectively, by

$$\text{PoC} := \frac{V^{\text{exp}}}{V^{\text{cap}}}, \quad \text{PoAN} := \frac{V^{\text{cap}}}{V^{\text{AN}}}, \quad \text{PoM} := \frac{V^{\text{exp}}}{V_*^{\text{mag}}}. \quad (6)$$

The first two prices measure the successive costs of requiring an exact cap and then equal treatment probabilities. Their product is the total cost of the two requirements. The price of magician also includes any loss from restricting implementation to the specified magician family. The following proposition bounds these prices without assuming that this family solves the general planning problem.

Proposition 1 (Prices of exact cap, arrival neutrality, and magician). *The optimal values satisfy*

$$0 \leq V_*^{\text{mag}} \leq V^{\text{AN}} \leq V^{\text{cap}} \leq V^{\text{exp}}.$$

If $V_^{\text{mag}} > 0$, both PoC and PoAN are at least one, and*

$$\text{PoC} \times \text{PoAN} = \frac{V^{\text{exp}}}{V^{\text{AN}}} \leq \text{PoM}.$$

If the expected-capacity optimum is attained by treating every applicant eligible under some $h^{\text{exp}} \in \mathcal{H}$, with

$$B(h^{\text{exp}}) \leq k, \quad J(h^{\text{exp}}) = V^{\text{exp}} > 0,$$

then

$$\text{PoC} \times \text{PoAN} \leq \text{PoM} \leq \frac{1}{\Gamma_k(k)} \leq \frac{1}{1 - (k+3)^{-1/2}}.$$

Under the stated condition on the eligibility class, a program with many treatment slots can enforce exact capacity and arrival neutrality while retaining nearly all the welfare attainable under an expected-capacity constraint. The relative welfare loss is at most $1/\sqrt{k+3}$: this bound depends only on capacity and vanishes as capacity grows, without requiring a forecast of the arrival order. The proposition does not determine how the combined price is divided between exact capacity and arrival neutrality, or how closely the specified magician approaches the best neutral implementation; Section 6.2 evaluates that decomposition.

4.2.1 Example: Cutoff rules

We now illustrate the eligibility-choice problem in one environment using cutoff rules. To determine how far to expand eligibility, we must compare the welfare gained from newly eligible applicants with the effect on those already eligible. Write $\tau_s(x) := \mathbb{E}_{P_s}[Y(1) - Y(0) \mid X = x]$ for the conditional average treatment effect. In this illustration, we restrict attention to rules $h_t(x) = \mathbf{1}\{\tau(x) \geq t\}$ that make applicants eligible when their effect is at least a cutoff t .

Lowering the threshold admits applicants whose effects are close to t , but also reduces the treatment probability of those already eligible. Let B and J denote the values of $B(h_t)$ and $J(h_t)$ at the current threshold. For a small increase $dB > 0$ in the expected number of eligible applicants, the newly eligible applicants contribute $\Gamma_k(B)t dB$ to welfare to first order. The corresponding change in welfare from those already eligible is $\Gamma'_k(B)J dB$. When $J > 0$, this second term is a loss, so positive effects among the newly eligible applicants alone do not justify expanding eligibility. When the effect distribution has positive density at the threshold and the guarantee curve is differentiable, an interior optimum balances these marginal effects:

$$\Gamma_k(B)t = -\Gamma'_k(B)J.$$

The following proposition expresses this necessary condition within the threshold class using the average effect among eligible applicants and the elasticity of their common treatment probability.

Proposition 2 (Magician-adjusted thresholds). *Let $S = 1$ and $\mathcal{C} = \{n\}$. Consider rules $h_t(x) = \mathbf{1}\{\tau(x) \geq t\}$, and suppose $\tau(X)$ has a density f . Write $p(t) = \mathbb{P}\{\tau(X) \geq t\}$, $\mu(t) = \mathbb{E}[\tau(X)\mathbf{1}\{\tau(X) \geq t\}]$, and $\bar{\tau}(t) = \mu(t)/p(t)$ for the average effect among eligible applicants. Let $\Phi(t) = \Gamma_k(np(t))n\mu(t)$. If t^* is an interior maximizer of Φ , f is continuous at t^* with $f(t^*) > 0$, and Γ_k is differentiable at $B^* = np(t^*)$, then*

$$t^* = \varepsilon_k(B^*)\bar{\tau}(t^*), \quad \varepsilon_k(B) = -\frac{B\Gamma'_k(B)}{\Gamma_k(B)} \in [0, 1]. \quad (7)$$

For $k = 1$, $\varepsilon_1(B) = B/(1 + B)$.

At an interior optimum, the effect at the eligibility boundary equals the average effect among eligible applicants multiplied by $\varepsilon_k(B^*)$. This factor measures the proportional fall in the common treatment probability per proportional increase in the expected number of eligible applicants. Expected-budget EWM instead sets the threshold so that $np(t) = k$ whenever effects are positive at that point, and the magician-adjusted threshold can lie on either side of it.

Example 1 (Eligibility thresholds with a uniform score). Consider one environment with $n = 4k$ applicants, a score $X \sim U(0, 1)$, and treatment effects $\tau(u) = u - u_0$. The expected-budget rule makes the top quarter eligible whenever $u_0 < 3/4$. Figure 2 compares this rule with the magician-adjusted threshold. When u_0 is small, the marginal applicant's effect is close to the average, and the magician-adjusted rule admits more applicants than the expected-budget rule. When u_0 approaches $3/4$, the marginal effect is small relative to the average, and it admits fewer. The switch occurs near $u_0 = 0.6$ for every k , and the deviation shrinks as k grows: at $u_0 = 0$, B/k is 2.00 for $k = 1$ and 1.52, 1.28, 1.13 and 1.06 for $k = 5, 20, 100$ and 500.

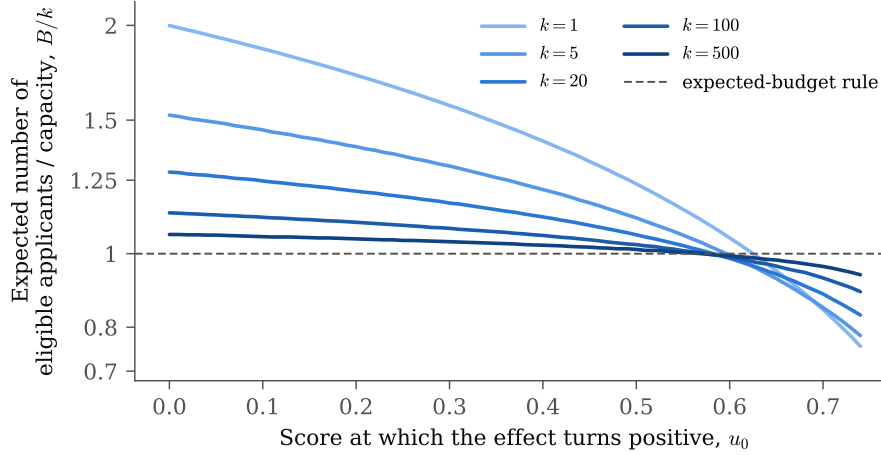


Figure 2: Expected number of eligible applicants relative to capacity under the threshold rule that maximizes $\Gamma_k(B(h))J(h)$ in Example 1. One environment, $n = 4k$ applicants, score $X \sim U(0, 1)$, effects $\tau(u) = u - u_0$. The expected-budget rule has $B = k$. Exact calculations.

5 Magician-Adjusted EWM

The construction in Section 4 fixes an implementation and a treatment probability for each eligibility rule. The remaining task is to learn which rule delivers the greatest welfare under that implementation. Under Assumption 1, the planner knows the covariate marginal P_s^X in each environment. It can therefore calculate $p_s(h)$, $B(h)$, and $\gamma(h) = \Gamma_k(B(h))$ for every candidate rule, without knowing any treatment effects. This section estimates the gains from experimental data, maximizes $\gamma(h)$ times the estimated gain, and then runs the magician with the exact labels $p_s(h)$. We show that the learned rule is feasible in the general planning problem (2) for every training sample, and bound its welfare loss relative to the best rule under the same magician implementation. Proofs are in Appendix D.

Throughout this section, the eligibility class $\mathcal{H}$ is finite and fixed before observing the experimental data.

Assumption 2 (Experimental data). For each environment s , the planner observes N_s independent draws of (X, D, Y) , independently of the deployment and of the samples in the other environments. The covariates have marginal distribution P_s^X , the potential outcomes have the same conditional distribution given X as in the deployment, and $D \mid X \sim \text{Bernoulli}(e_s(X))$, where the known propensity satisfies $\eta \leq e_s(\cdot) \leq 1 - \eta$ for some $\eta \in (0, 1/2]$. Treatment is independent of $(Y(1), Y(0))$ conditional on X , and $Y = DY(1) + (1 - D)Y(0)$.

For each rule and environment, the planner estimates the gain from treating eligible applicants by

$$\hat{\mu}_s(h) = \frac{1}{N_s} \sum_{j=1}^{N_s} h_s(X_j) \left\{ \frac{D_j Y_j}{e_s(X_j)} - \frac{(1 - D_j) Y_j}{1 - e_s(X_j)} \right\}, \quad \hat{J}(h) = \min_{c \in \mathcal{C}} \sum_s c_s \hat{\mu}_s(h).$$

The estimator is unbiased for $\mu_s(h)$ before taking the minimum over compositions. *Magician-adjusted EWM* chooses

$$\hat{h} \in \arg \max_{h \in \mathcal{H}} \Gamma_k(B(h)) \hat{J}(h). \quad (8)$$

As applicants arrive, the program runs the $\gamma(\hat{h})$ -conservative magician with labels $p_{s_i}(\hat{h})$. The rule and its treatment probability are chosen using the experimental data, but the labels remain exact: once $\hat{h}$ is fixed, $p_s(\hat{h}) = \int \hat{h}_s(x) dP_s^X(x)$ is known. An error in estimating treatment effects can therefore change which rule is selected, but does not change the magician's ability to implement that rule at its specified probability.

We evaluate the learned rule against the best eligibility rule under the same magician implementation. If treatment gains were known, the planner would choose

$$h^* \in \arg \max_{h \in \mathcal{H}} \Gamma_k(B(h))J(h), \quad V_*^{\text{mag}} := V^{\text{mag}}(h^*).$$

The welfare loss from learning is the difference between this population optimum and the welfare delivered by $\hat{h}$.

Because $\hat{h}$ is selected using the experimental data, we need an error bound that holds for every candidate rule at once. The following bound is a direct application of Hoeffding's inequality (Hoeffding, 1963) and a union bound over the finite class.

Proposition 3 (Uniform estimation error). *Under Assumption 2 and the conditions on $\mathcal{H}$ stated above, fix $\delta \in (0, 1)$ and define*

$$\mathcal{E} := \frac{1}{\eta} \max_{c \in \mathcal{C}} \sum_s c_s \sqrt{\frac{2 \log(2S|\mathcal{H}|/\delta)}{N_s}}. \quad (9)$$

Then, with probability at least $1 - \delta$ over the experimental data,

$$\max_{h \in \mathcal{H}} |\hat{J}(h) - J(h)| \leq \mathcal{E}.$$

The bound applies to both $\hat{h}$ and h^* , even though $\hat{h}$ is chosen using the data. It is used to assess performance; neither the selection rule nor the magician requires it.

The welfare-loss argument follows the standard EWM argument of Kitagawa and Tetenov (2018). Here the multiplier $\gamma(h) = \Gamma_k(B(h))$ is known and lies in $[0, 1]$, so Proposition 3 also bounds each candidate's welfare-estimation error by $\mathcal{E}$. Since $\hat{h}$ maximizes estimated welfare, its loss relative to h^* is at most the sum of their two evaluation errors, or $2\mathcal{E}$. The following theorem combines this statistical argument with the magician's capacity and arrival-neutrality guarantees.

Theorem 4 (Policy learning with known covariate distributions). *Under Assumptions 1 and 2 and the conditions on $\mathcal{H}$ stated above, magician-adjusted EWM has the following properties.*

- (i) (Capacity.) $\sum_i A_i \leq k$ for every training sample and every deployment realization.
- (ii) (Exact arrival neutrality.) For every training sample, every $\sigma \in \Sigma(\mathcal{C})$, every position i , and $P_{s_i}^X$ -almost every x with $\hat{h}_{s_i}(x) = 1$,

$$\mathbb{P}(A_i = 1 \mid X_i = x, \text{data}; \sigma) = \gamma(\hat{h}).$$

Consequently, $V(\hat{h}, \widehat{M}) = \gamma(\hat{h})J(\hat{h}) = V^{\text{mag}}(\hat{h})$.

(iii) (Welfare.) With probability at least $1 - \delta$ over the experimental data,

$$0 \leq V_*^{\text{mag}} - V(\hat{h}, \widehat{M}) \leq 2\mathcal{E}.$$

In particular, if $N_s = N$ for all s , then with probability at least $1 - \delta$,

$$\frac{V_*^{\text{mag}} - V(\hat{h}, \widehat{M})}{n} \leq \frac{2}{\eta} \sqrt{\frac{2 \log(2S|\mathcal{H}|/\delta)}{N}}.$$

The bound does not depend on the deployment size n or the capacity k .

Parts (i) and (ii) hold even when the treatment-effect estimates are inaccurate. Conditional on the data, the selected rule is fixed, and its eligibility indicators are independent with the exact probabilities $p_{s_i}(\hat{h})$. The implementation argument in Section 4.1 therefore gives every eligible applicant the same treatment probability in every admissible order. Estimation affects which rule is chosen, while the known covariate distributions preserve the validity of its implementation.

Part (iii) retains the usual $N^{-1/2}$ rate of empirical welfare maximization (Kitagawa and Tetenov, 2018). The known probability adjustment does not enlarge the estimation error, so the bound per applicant has no additional dependence on deployment size or capacity. Appendix D gives sharper bounds that account for each rule’s treatment probability and eligible environments, together with a data-dependent lower bound on the welfare of the selected rule.

6 Simulations

The simulations separate the welfare loss from learning an eligibility rule from the cost of implementing it through the magician. We consider a job-training program with applicants from two regions. Its performance relative to the expected-capacity optimum has two components:

$$\frac{V^{\text{mag}}(\hat{h})}{V^{\text{exp}}} = \underbrace{\frac{V^{\text{mag}}(\hat{h})}{V_*^{\text{mag}}}}_{\text{learning}} \times \underbrace{\frac{V_*^{\text{mag}}}{V^{\text{exp}}}}_{\text{implementation}}.$$

The first experiment asks how closely experimental data allow the program to approach the best eligibility rule under the same magician. The second takes the effects of training as known and asks how much welfare that best rule retains relative to the expected-capacity optimum. This is the comparison in Proposition 1: under its condition on the eligibility class, a program with many places can enforce exact capacity and arrival neutrality with little relative welfare loss. We examine learning as the experimental sample grows, then implementation costs as capacity grows. Appendix G gives further results and computational details.

Design. The program has k places and $n = 4k$ applicants, half from region A ($s = 1$) and half from region B ($s = 2$). For example, a program with 20 places receives 80 applicants, 40 from each region. These regional counts are known, but the order in which applicants arrive is unknown, so $\mathcal{C} = \{(n/2, n/2)\}$. Each applicant has a score that predicts the effect of training. In each region, the score takes 20 equally likely values $u_j = (j - 1/2)/20$; each value accounts

for 5 percent of the regional population. These score distributions are known throughout the main experiments.

The outcome is employment: $Y = 1$ if the applicant finds a job and $Y = 0$ otherwise. Without training, every applicant has a 40 percent chance of employment. Training changes this probability by $\tau_1(u) = 0.05 + 0.35u$ in region A and $\tau_2(u) = -0.30 + 0.90u$ in region B. Training therefore helps applicants throughout region A, with larger gains at higher scores. In region B, it has large positive effects at high scores but negative effects at low scores. Conditional on region s and score u , employment outcomes without training are Bernoulli(0.4) and those with training are Bernoulli($0.4 + \tau_s(u)$).

An eligibility rule chooses a separate score cutoff for each region. It can make the top 0, 5, 10, ..., 100 percent of either region eligible, giving 21 choices per region and $21 \times 21 = 441$ candidate rules in $\mathcal{H}$. For example, one rule makes the top 20 percent of region A and the top 30 percent of region B eligible; another makes the top 40 percent of A and the top 10 percent of B eligible. Choosing zero percent allows the program to exclude either region entirely.

Welfare measures the expected increase in the number of employed applicants due to training, relative to giving nobody training. The rule attaining the expected-budget benchmark V^{EB} maximizes the true gain $J(h)$ from training all eligible applicants, subject to their expected number $B(h)$ being at most k . Among the 441 candidates, this rule makes the top 20 percent of region A and the top 30 percent of region B eligible.¹ For the program with 20 places, this means $40 \times 0.20 = 8$ eligible applicants from A and $40 \times 0.30 = 12$ from B in expectation. More generally, $B = (n/2) \times 0.20 + (n/2) \times 0.30 = n/4 = k$.

The actual number of eligible applicants can exceed the available places. For example, 10 applicants from A and 14 from B could meet these cutoffs even though only 20 places are available. The magician respects this limit while giving every eligible applicant the same probability of receiving training. Because this probability depends on the cutoffs, magician-adjusted EWM accounts for it when choosing whom to make eligible.

6.1 Learning eligibility under the magician

The first experiment measures how closely a learned eligibility rule approaches the best rule under the same implementation. The policymaker learns the effects of training from a separate randomized experiment in each region, with N participants assigned to training with probability one-half. The known regional score distributions give the eligibility probabilities $p_s(h)$ and expected eligible counts $B(h)$ exactly; only the employment gains must be estimated.

We compare two ways of choosing eligibility from these data. Magician-adjusted EWM maximizes $\Gamma_k(B(h))\hat{J}(h)$, whereas *EB-EWM* maximizes $\hat{J}(h)$ subject to $B(h) \leq k$. Both learned rules are implemented by the magician using exact eligibility probabilities. Their comparison therefore isolates the importance of accounting for the eventual treatment probability when learning whom to make eligible.

Figure 3 plots mean employment gains over 500 replications at each of four experimental sample sizes, for programs with $k \in \{5, 20, 100, 500\}$ places. The dotted line at 100 percent is the best rule under the same magician with the true effects of training known. We denote this

¹We solve $\max_{h \in \mathcal{H}: B(h) \leq k} J(h)$ by enumerating all 441 cutoff pairs, using the known score probabilities and the true effects of training to compute $B(h)$ and $J(h)$. The unique maximizing pair makes the four highest score values in region A and the six highest in region B eligible, yielding $B(h) = k$ and $V^{\text{EB}}/n = 0.10625$.

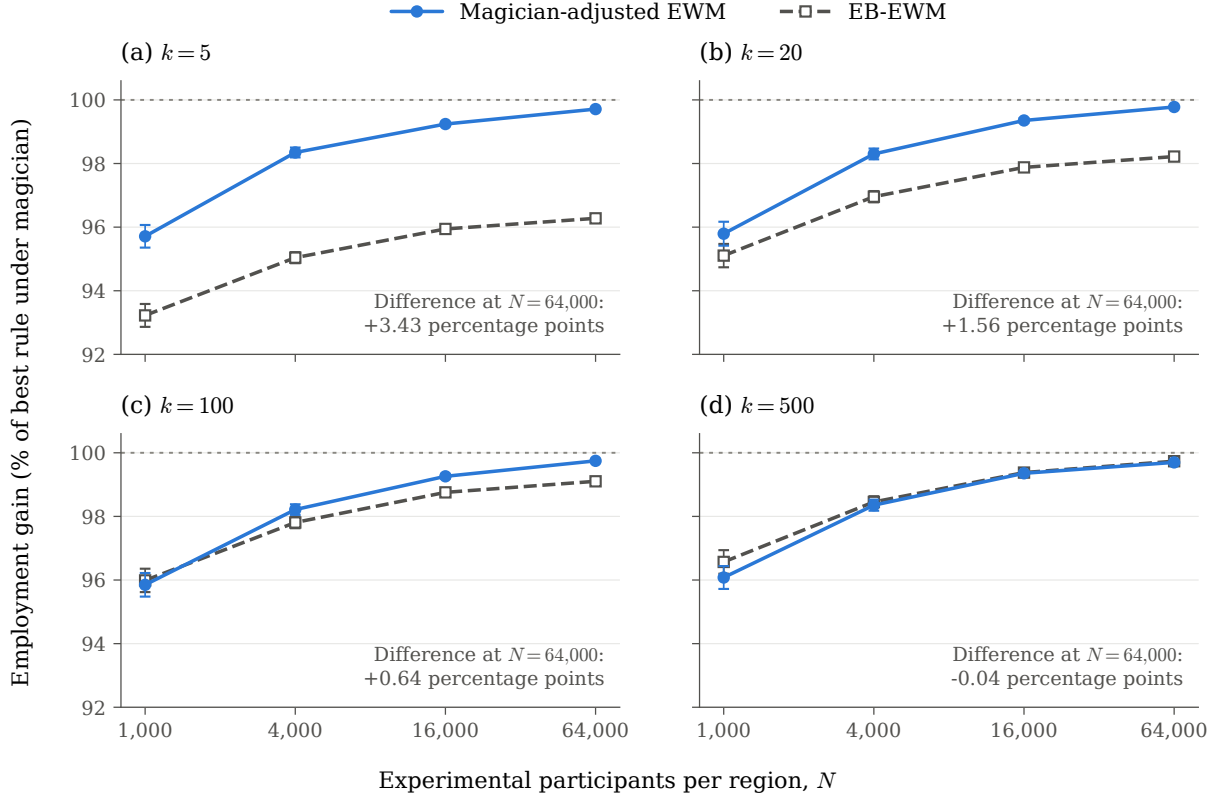


Figure 3: Learning eligibility under the magician: approaching the best rule with known effects
Notes: Job-training programs with k places, $n = 4k$ applicants split equally between regions A and B, known regional score distributions, and 441 pairs of eligibility cutoffs. Both learned rules are implemented by the magician using exact eligibility probabilities. The horizontal axis is the number of experimental participants per region, on a logarithmic scale. Points are mean expected employment gains over 500 experimental samples, expressed as a percentage of $\tilde{V}_*^{\text{mag}}$; the dotted line at 100 is the best rule under the same magician with known effects of training. Error bars are the mean plus or minus 1.96 Monte Carlo standard errors. The annotated differences are magician-adjusted EWM minus EB-EWM at $N = 64,000$, in percentage points. The target uses the numerical guarantee curve and safety adjustment of Appendix F. The two methods use the same experimental samples. Table G.3 reports all settings, including the FCFS comparison.

computed target by $\tilde{V}_*^{\text{mag}}$, the numerical counterpart of V_*^{mag} .² For magician-adjusted EWM, the vertical distance from this line measures the loss from learning; error bars measure Monte Carlo uncertainty in the mean welfare shares.

Larger experimental samples bring magician-adjusted EWM close to its target at every program size. Its mean employment gain rises from about 96 percent of $\tilde{V}_*^{\text{mag}}$ at $N = 1,000$ to about 99.7 percent at $N = 64,000$. The similar learning curves across program sizes are consistent with Theorem 4, whose per-applicant regret bound depends on the experimental sample size but not on the number of training places.

The learning objective matters most in small programs. The advantage of magician-adjusted EWM persists at the largest experimental sample size in the three smaller programs. In the largest program, the curves nearly coincide and EB-EWM performs slightly better on average in these finite samples; the two objectives select the same rule when effects are known. Thus,

²All numerical comparisons use the same computed guarantee curve and safety adjustment described in Appendix F. The learning target and the numerical price of magician both use this implementation; the theoretical bounds concern the exact curve Γ_k .

accounting for the magician improves eligibility choice in the smaller programs in this design, but does not dominate EB-EWM at every sample size and capacity.

This difference reflects the value of broader eligibility under the magician. In this design, training effects decline slowly among applicants newly included by a lower cutoff. The extra gains can outweigh the reduction in the common treatment probability, as described in Proposition 2. With effects known, magician-adjusted EWM therefore chooses broader eligibility than EB-EWM in the three smaller programs. More data allow each method to approach its own objective, so they do not eliminate the difference between those objectives. Table G.1 reports the full comparison with effects known.

Capacity and arrival neutrality are preserved while eligibility is learned. Known score distributions allow the magician to implement the selected rule even when its estimated employment gains are inaccurate; the calculations reproduce the common treatment probability to numerical precision. The experiment therefore concerns how much welfare is lost in choosing the rule. The next experiment asks how much welfare even the best rule under this magician forgoes to obtain the two implementation guarantees.

6.2 The prices of exact cap and arrival neutrality

We now evaluate the target that learning approaches: the best welfare under the specified magician. Taking the effects of training as known removes estimation error and isolates the cost of implementation. All comparisons use the same eligibility class, regional applicant counts, and information available as applicants arrive. Welfare is evaluated in the least favorable fixed regional order, which the program does not know in advance.

The benchmark V^{exp} is the largest employment gain when the expected number trained cannot exceed capacity. In this design, it is attained by training everyone in the top 20 percent of region A and the top 30 percent of region B. This candidate rule has $B(h) = k$ and $J(h) = V^{\text{exp}} > 0$, verifying the condition for the upper bound in Proposition 1. Requiring the program never to exceed capacity gives V^{cap} ; adding a common training probability for all eligible applicants gives V^{AN} . Restricting implementation further to the magician calibrated through $B(h)$ gives V_*^{mag} , the learning target. Appendix G.2 reports the optimization methods, certified numerical bounds, and full values.

Figure 4 shows where the implementation cost comes from. Panel (a) separates the price of exact cap, $V^{\text{exp}}/V^{\text{cap}}$, from the price of arrival neutrality, $V^{\text{cap}}/V^{\text{AN}}$. A ratio of one means no welfare cost. Both prices decline with capacity in this design, but arrival neutrality accounts for more of the remaining cost in larger programs.

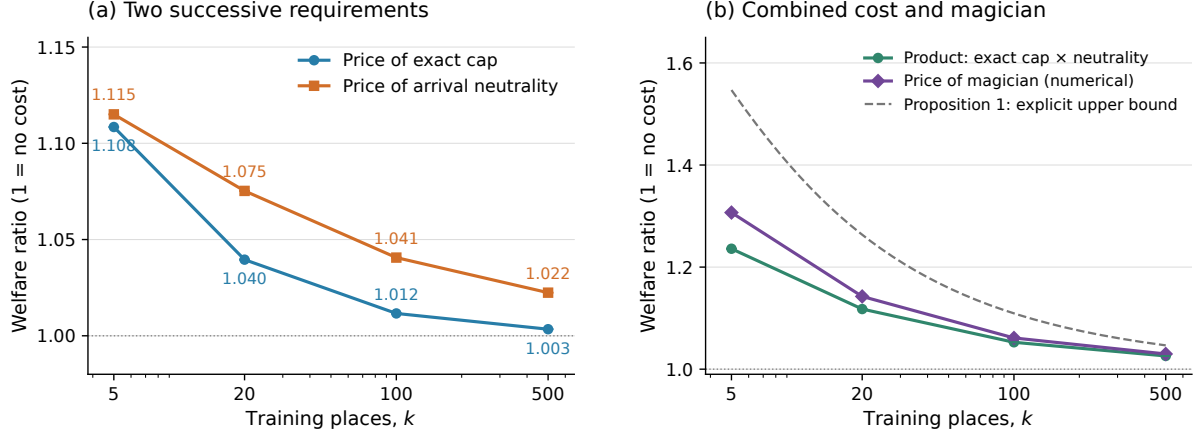


Figure 4: The prices of exact cap, arrival neutrality, and magician. Panel (a) plots $\text{PoC} = V^{\text{exp}}/V^{\text{cap}}$ and $\text{PoAN} = V^{\text{cap}}/V^{\text{AN}}$. Panel (b) compares their product with the numerical price of magician $\widetilde{\text{PoM}}$ and the bound $[1 - (k + 3)^{-1/2}]^{-1}$ from Proposition 1. Each program has $n = 4k$ applicants, half from each region. Certified intervals for the cap and neutral optima are narrower than the plotted markers. Magician values use the same numerical curve and safety adjustment as the learning experiment; the theoretical bound concerns the exact curve.

Panel (b) shows that the magician retains nearly all of the expected-capacity optimum in the largest program. The numerical price of magician is $\widetilde{\text{PoM}} := V^{\text{exp}}/\widetilde{V}_*^{\text{mag}}$, so its inverse is the share of that optimum retained by the implementation. This share rises from 76.5 percent with five places to 97.1 percent with five hundred. The product of the two component prices measures the cost of exact capacity and arrival neutrality themselves. Its gap from the numerical magician price reflects the further restriction to calibration through $B(h)$ and the numerical safety adjustment; this gap is small in the larger programs. The dashed curve gives the explicit upper bound from Proposition 1.

The two experiments illustrate complementary roles for data and capacity. More experimental data bring the learned rule close to the best rule under the magician; a larger capacity makes that target close to the expected-capacity optimum in this design. High welfare can therefore coexist with exact capacity and arrival neutrality when the program has enough places and enough data to learn eligibility well.

Changes in who applies. These conclusions rely on the specified regional counts and score distributions. Appendix G.4 shows that departures in regional shares or eligibility rates can reduce late eligible applicants' chances of receiving training. Calibrating the magician for a wider range of regional shares preserves the common probability across that range, at the cost of a lower promised probability. Uncertainty about who applies therefore affects calibration even though errors in estimating training effects do not invalidate the implementation guarantees.

7 Conclusion

When treatment must be allocated before all applicants are known, the planner chooses eligibility, a common treatment probability, and an online implementation that respects capacity.

Arrival neutrality makes the welfare objective the product $\gamma J(h)$. The conservative magician provides a feasible implementation for each eligibility rule, with treatment probability $\Gamma_k(B(h))$. Choosing the best eligibility rule under this implementation therefore amounts to maximizing $\Gamma_k(B(h))J(h)$. Broader eligibility adds potential gains but can reduce each eligible applicant’s chance of treatment. Other implementations can improve on this family by using further restrictions on the possible applicant sequences.

With known covariate distributions, these probabilities are known for every candidate rule, and experimental data are needed only to learn treatment gains. The resulting learning procedure preserves exact arrival neutrality for every training sample, with per-applicant regret of order $N^{-1/2}$ relative to the best rule under the same magician implementation.

A natural extension is to allow treatment costs to differ across applicants and to impose several resource constraints simultaneously. A training program, for example, may need to respect both an overall budget and separate capacity limits at its facilities. This raises the question of how to learn eligibility rules that deliver high welfare while satisfying every resource constraint on each realization and giving all eligible applicants a common treatment probability.

References

- Karun Adusumilli, Friedrich Geiece, and Claudio Schilter. Dynamically optimal treatment allocation. arXiv:1904.01047, version 5, 2024.
- Saeed Alaei. Bayesian combinatorial auctions: Expanding single buyer mechanisms to many buyers. arXiv:1106.0961, version 4, 2012. Conference version in FOCS 2011.
- Makis Arsenis and Robert Kleinberg. Individual fairness in prophet inequalities. In *Proceedings of the 23rd ACM Conference on Economics and Computation*, 2022.
- Susan Athey and Stefan Wager. Policy learning with observational data. *Econometrica*, 89(1):133–161, 2021.
- Debopam Bhattacharya and Pascaline Dupas. Inferring welfare maximizing treatment assignment under budget constraints. *Journal of Econometrics*, 167(1):168–196, 2012.
- Moran Feldman, Ola Svensson, and Rico Zenklusen. Online contention resolution schemes with applications to Bayesian selection problems. *SIAM Journal on Computing*, 50(2):255–300, 2021.
- Moran Feldman, Ola Svensson, and Rico Zenklusen. Nearly tight sample complexity for matroid online contention resolution. arXiv:2507.09507, 2025.
- Mohammadsaeed Haghi, Mahdi Salmani, and Nima Kelidari. Differentiating through dual prices: End-to-end policy learning under capacity constraints. arXiv:2608.04669, 2026.
- Keisuke Hirano and Jack R. Porter. Asymptotics for statistical treatment rules. *Econometrica*, 77(5):1683–1701, 2009.
- Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963. doi: 10.1080/01621459.1963.10500830.

- Jiashuo Jiang, Will Ma, and Jiawei Zhang. Tight guarantees for multiunit prophet inequalities and online stochastic knapsack. *Operations Research*, 73(3):1703–1721, 2025.
- Toru Kitagawa and Aleksey Tetenov. Who should be treated? Empirical welfare maximization methods for treatment choice. *Econometrica*, 86(2):591–616, 2018.
- Anders Bredahl Kock and David Preinerstorfer. Regularizing discrimination in optimal policy learning with distributional targets. arXiv:2401.17909, 2024.
- JungHo Lee, Johnna Sundberg, Pim Welle, and Bryan Wilder. Learning treatment effects during resource allocation via priority-queue randomization. arXiv:2605.25169, 2026.
- Alexander R. Luedtke and Mark J. van der Laan. Optimal individualized treatments in resource-limited settings. *International Journal of Biostatistics*, 12(1), 2016.
- Vahideh Manshadi, Rad Niazadeh, and Scott Rodilitz. Fair dynamic rationing. arXiv:2102.01240, 2021.
- Charles F. Manski. Statistical treatment rules for heterogeneous populations. *Econometrica*, 72(4):1221–1246, 2004.
- Eric Mbakop and Max Tabord-Meehan. Model selection for treatment choice: Penalized welfare maximization. *Econometrica*, 89(2):825–848, 2021.
- Shosei Sakaguchi. Estimation of optimal dynamic treatment assignment rules under policy constraints. *Quantitative Economics*, 16:981–1022, 2025.
- Jörg Stoye. Minimax regret treatment choice with finite samples. *Journal of Econometrics*, 151(1):70–81, 2009.
- Liyang Sun. Empirical welfare maximization with constraints. *Journal of Econometrics*, 253:106169, 2026. doi: 10.1016/j.jeconom.2025.106169.
- Davide Viviano and Jelena Bradic. Fair policy targeting. *Journal of the American Statistical Association*, 119(545):730–743, 2024.

A Proof for Section 2

Proof of Lemma 1. Fix $\sigma \in \Sigma(\mathcal{C})$ and condition on it throughout. The decision A_i depends on X_i , on the covariates and decisions of earlier applicants, and on independent randomization. By Assumption 1(ii), A_i is independent of $(Y_i(0), Y_i(1))$ given X_i . Hence $\mathbb{E}[A_i \Delta_i \mid X_i] = \mathbb{P}(A_i = 1 \mid X_i) \tau_{s_i}(X_i)$. Arrival neutrality gives $\mathbb{P}(A_i = 1 \mid X_i = x) = \gamma$ for almost every eligible x , and $A_i = 0$ for ineligible x . Therefore

$$\mathbb{E}[A_i \Delta_i] = \gamma \mathbb{E}[h_{s_i}(X_i) \tau_{s_i}(X_i)] = \gamma \mu_{s_i}(h).$$

Summing over i gives $V_\sigma(h, M) = \gamma \sum_i \mu_{s_i}(h) = \gamma \sum_s c_s(\sigma) \mu_s(h)$. The same $\gamma \geq 0$ applies to every $\sigma \in \Sigma(\mathcal{C})$, and these sequences range over all compositions in $\mathcal{C}$. Taking the minimum over σ therefore gives $V(h, M) = \gamma \min_{c \in \mathcal{C}} \sum_s c_s \mu_s(h) = \gamma J(h)$. $\square$

B Proofs for Section 3

Throughout, for a random variable W with values in $\mathbb{N}_0$ and an integer $c \geq -1$, write $\psi(c) = \mathbb{E}[(W - c)_+]$ and $T(c) = \mathbb{P}(W \geq c)$. For integers w and c , $(w + 1 - c)_+ - (w - c)_+ = \mathbf{1}\{w \geq c\}$, so that

$$T(c) = \psi(c - 1) - \psi(c) \quad (c \geq 0), \quad \psi(-1) = \psi(0) + 1. \quad (\text{B.1})$$

The function ψ is nonincreasing and convex in c , and $\psi(c) = 0$ for $c \geq k$ when $W \leq k$.

B.1 The magician's recursion

For a label sequence x and a level $\gamma \in (0, 1]$, let π_i be the reference distributions of Section 3, and let $\Psi_i(c) = \sum_w (w - c)_+ \pi_i(w)$ and $\tau_i(c) = \pi_i(\{c, c + 1, \dots\})$. The mass that moves from state w at step i is $x_i o_i(w)$, where

$$o_i(w) = \pi_i(w) \varphi_i(w) = \min \left\{ \pi_i(w), \left[\gamma - \pi_i(\{0, \dots, w - 1\}) \right]_+ \right\}. \quad (\text{B.2})$$

The second expression follows from the definition of φ_i : the magician opens the box in every state below θ_i , where the cumulative mass is below γ , and randomizes at θ_i to bring the total opening probability to γ . Summing (B.2) over $w < c$ gives

$$\sum_{w < c} o_i(w) = \min \left\{ \gamma, \pi_i(\{0, \dots, c - 1\}) \right\}, \quad \sum_w o_i(w) = \gamma. \quad (\text{B.3})$$

Lemma B.1 (Recursion). *Fix $x \in [0, 1]^n$ and $\gamma \in (0, 1]$.*

(i) *For every i and every integer $c \geq 0$,*

$$\Psi_{i+1}(c) = \Psi_i(c) + x_i [\gamma - 1 + \Psi_i(c - 1) - \Psi_i(c)]_+. \quad (\text{B.4})$$

In particular $\Psi_i(0) = \gamma \sum_{j < i} x_j$, and the coordinates $c = 0, \dots, k - 1$ evolve as a closed system.

(ii) *The magician is feasible on x if and only if $\Psi_i(k - 1) \leq 1 - \gamma$ before every box i .*

(iii) *If the magician is feasible on x and the labels are exact, then $W_i \sim \pi_i$ and $\mathbb{P}(G_i = 1) = \gamma$ for every box i .*

(iv) *$\gamma_k^*(x) \geq 1/(1 + \sum_i x_i)$.*

Proof. (i) Each unit of mass that moves raises $(w - c)_+$ by $\mathbf{1}\{w \geq c\}$, so $\Psi_{i+1}(c) - \Psi_i(c) = x_i \sum_{w \geq c} o_i(w)$. By (B.3), $\sum_{w \geq c} o_i(w) = \gamma - \min\{\gamma, 1 - \tau_i(c)\} = [\gamma - 1 + \tau_i(c)]_+$, and $\tau_i(c) = \Psi_i(c - 1) - \Psi_i(c)$ by (B.1). For $c = 0$ the increment is $x_i \gamma$, because $\tau_i(0) = 1$. The update of coordinate c involves only coordinates $c - 1$ and c , so coordinates $0, \dots, k - 1$ form a closed system.

(ii) Feasibility before box i is $\pi_i(\{0, \dots, k - 1\}) \geq \gamma$, that is, $\tau_i(k) \leq 1 - \gamma$. If the magician is feasible before every earlier box, mass has moved only from states $w \leq \theta_j \leq k - 1$. Hence π_i is supported on $\{0, \dots, k\}$, $\Psi_i(k) = 0$, and $\tau_i(k) = \Psi_i(k - 1)$. Induction over all boxes proves the equivalence. A zero label leaves the state unchanged, but opening that box still requires an unbroken wand and therefore the same feasibility check.

(iii) Suppose $W_i \sim \pi_i$, which holds for $i = 1$. Feasibility gives $\theta_i \leq k - 1$, so $\varphi_i(k) = 0$ and the explicit stop does not bind. Since the latent break indicator Z_i is independent of (W_i, G_i) , $\mathbb{P}(W_{i+1} = w) = \pi_i(w)\{1 - x_i\varphi_i(w)\} + \pi_i(w - 1)x_i\varphi_i(w - 1) = \pi_{i+1}(w)$. Moreover $\mathbb{P}(G_i = 1) = \sum_w \pi_i(w)\varphi_i(w) = \gamma$, including when $x_i = 0$.

(iv) Ψ_i is nonincreasing in c , so $\Psi_i(k - 1) \leq \Psi_i(0) = \gamma \sum_{j < i} x_j \leq \gamma \sum_j x_j$. If $\gamma \leq 1/(1 + \sum_j x_j)$, this is at most $1 - \gamma$, and (ii) gives feasibility. $\square$

For $u, v \in \mathbb{R}$ and $\xi \in [0, 1]$ let

$$F_\xi(u, v) = v + \xi [\gamma - 1 + u - v]_+. \quad (\text{B.5})$$

F_ξ is nondecreasing in each argument: where the bracket is positive its partial derivatives are $(\xi, 1 - \xi)$, and elsewhere they are $(0, 1)$. By (B.4), $\Psi_{i+1}(c) = F_{x_i}(\Psi_i(c - 1), \Psi_i(c))$; for $c = 0$ this reads $\Psi_{i+1}(0) = \Psi_i(0) + x_i\gamma$, which is nondecreasing in $\Psi_i(0)$.

B.2 Comparison and fixed-sequence optimality

For positive labels with $\sum_i x_i \leq k$, Theorem 2 recalls fixed-sequence optimality from Jiang et al. (2025, Theorem 1). Here we give a comparison of cumulative resource use that proves the box-opening result without restricting the sum of break probabilities. We state the comparison more generally: a resource-use decision A_i may depend on the current Bernoulli indicator Z_i as well as the past, provided $A_i \leq Z_i$. This includes a magician's wand breaks, $A_i = G_i Z_i$, and will also allow the treatment comparison in Appendix C.1. The comparison concerns the reference distributions and does not presuppose their feasibility.

Proposition B.1 (Comparison of resource-use distributions). *Fix $\gamma \in (0, 1]$. Let $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \dots \subseteq \mathcal{F}_n$ be a filtration. For $i = 1, \dots, n$ let $Z_i \in \{0, 1\}$ be $\mathcal{F}_i$ -measurable and independent of $\mathcal{F}_{i-1}$, with $\mathbb{P}(Z_i = 1) = x_i$, and let $A_i \in \{0, 1\}$ be $\mathcal{F}_i$ -measurable with $A_i \leq Z_i$ and $\sum_{j \leq i} A_j \leq k$. Let $W_i = \sum_{j < i} A_j$, and let π_i be the reference distributions at level γ for the labels x . Suppose $\mathbb{P}(A_i = 1 \mid Z_i = 1) \geq \gamma$ for every i with $x_i > 0$. Then*

$$\Psi_i(c) = \sum_w (w - c)_+ \pi_i(w) \leq \mathbb{E}[(W_i - c)_+]$$

for every i and every integer c .

The inequalities say that the reference distribution is dominated in increasing convex order by the resource-use distribution of any rule satisfying the conditional floor. The proof compares expected resource use above each threshold and shows that opening boxes in the states with the fewest broken wands minimizes its increase.

For the magician, take $\mathcal{F}_i$ to contain $Z_1, \dots, Z_i$ and the randomization used up to box i . Although these sigma-fields include the latent outcomes of unopened boxes, an actual box-opening strategy uses only outcomes observed when boxes were opened. Its decision G_i precedes the realization of Z_i , and $A_i = G_i Z_i$. The proposition also allows more general decisions made after observing Z_i ; this greater information is used only in the comparison argument.

Proof of Proposition B.1. Write $\psi_i(c) = \mathbb{E}[(W_i - c)_+]$ and $T_i(c) = \mathbb{P}(W_i \geq c)$ for the given rule. Since $W_{i+1} = W_i + A_i$,

$$\psi_{i+1}(c) = \psi_i(c) + \mathbb{P}(A_i = 1, W_i \geq c). \quad (\text{B.6})$$

For $x_i > 0$, because $A_i \leq Z_i$, W_i is $\mathcal{F}_{i-1}$ -measurable, and Z_i is independent of $\mathcal{F}_{i-1}$,

$$\begin{aligned} \mathbb{P}(A_i = 1, W_i \geq c) &= x_i \mathbb{P}(A_i = 1, W_i \geq c \mid Z_i = 1) \\ &\geq x_i [\mathbb{P}(A_i = 1 \mid Z_i = 1) - \mathbb{P}(W_i \leq c - 1)]_+ \geq x_i [\gamma - 1 + T_i(c)]_+. \end{aligned} \quad (\text{B.7})$$

If $x_i = 0$, then $A_i = 0$ almost surely and the same lower bound holds with both sides zero. Combining (B.1), (B.6) and (B.7) gives $\psi_{i+1}(c) \geq F_{x_i}(\psi_i(c-1), \psi_i(c))$ for $c \geq 0$.

We show by induction on i that $\Psi_i(c) \leq \psi_i(c)$ for every integer $c \geq 0$. For $i = 1$ both sides vanish. Suppose the claim holds for i . For every $c \geq 0$, monotonicity of F_{x_i} gives

$$\Psi_{i+1}(c) = F_{x_i}(\Psi_i(c-1), \Psi_i(c)) \leq F_{x_i}(\psi_i(c-1), \psi_i(c)) \leq \psi_{i+1}(c),$$

where for $c = 0$ we use $\Psi_i(-1) = \Psi_i(0) + 1$ and $\psi_i(-1) = \psi_i(0) + 1$. This completes the induction. For $c \leq -1$, $\Psi_i(c) = \Psi_i(0) - c \leq \psi_i(0) - c = \psi_i(c)$. $\square$

For a box-opening strategy, the comparison and the requirement of an unbroken wand imply that the conservative magician is feasible at the same opening-probability floor. This extends fixed-sequence optimality to every finite label sequence, including zero labels and sums of break probabilities exceeding k .

Corollary B.1 (Fixed-sequence optimality for arbitrary label sequences). *For every finite label sequence $x \in [0, 1]^n$, the feasible levels of the conservative magician form the closed interval $[0, \gamma_k^*(x)]$. With independent break indicators and exact labels, $\gamma_k^*(x)$ is both the largest common opening probability and the largest opening-probability floor attainable by an online box-opening strategy with k wands.*

Proof of Corollary B.1. Consider a box-opening strategy with k wands and opening floor $\gamma > 0$, and write $\psi_i(c) = \mathbb{E}[(W_i - c)_+]$. Because G_i is independent of Z_i , its breaks $A_i = G_i Z_i$ satisfy $\mathbb{P}(A_i = 1 \mid Z_i = 1) = \mathbb{P}(G_i = 1) \geq \gamma$ whenever $x_i > 0$. Thus Proposition B.1 applies. In particular $\Psi_i(k) = 0$ because $W_i \leq k$, so the reference distribution is supported on $\{0, \dots, k\}$. Opening any box requires an unbroken wand, including a box whose break probability is zero. Consequently,

$$\gamma \leq \mathbb{P}(G_i = 1) \leq \mathbb{P}(W_i \leq k - 1) = 1 - \psi_i(k - 1) \quad \text{for every } i. \quad (\text{B.8})$$

The comparison also gives $\Psi_i(k - 1) \leq \psi_i(k - 1) \leq 1 - \gamma$ before every box. By Lemma B.1(ii), the γ -conservative magician is feasible, and Lemma B.1(iii) gives $\mathbb{P}(G_i = 1) = \gamma$ for every box.

The level $\gamma = 0$ is feasible. Suppose the magician is feasible at level γ and let $0 < \gamma' < \gamma$. With exact labels, the γ -magician is a box-opening strategy with floor γ' , so the preceding argument shows that the γ' -magician is feasible. The feasible levels therefore form an interval containing 0. By (B.4), each $\Psi_i(c)$ is a continuous function of γ , and by Lemma B.1(ii) feasibility is a finite collection of weak inequalities $\Psi_i(k - 1) \leq 1 - \gamma$. The set of feasible levels is therefore closed, and its maximum $\gamma_k^*(x)$ exists. A common opening probability is an opening floor; any floor is at most $\gamma_k^*(x)$ by the first part of this proof; and the $\gamma_k^*(x)$ -magician attains that same opening probability for every box. $\square$

B.3 The guarantee curve

Proof of Corollary 1. Let x have total at most B . By (3), $\Gamma_k(B) \leq \gamma_k^*(x)$, so the $\Gamma_k(B)$ -magician is feasible on x by Corollary B.1 and opens every box with probability $\Gamma_k(B)$ by Lemma B.1(iii). The level $\Gamma_k(B)$ is positive by Lemma B.1(iv). If $\gamma > \Gamma_k(B)$, there is a label sequence x with total at most B and $\gamma_k^*(x) < \gamma$, and by Corollary B.1 no box-opening strategy has floor γ on x . $\square$

Preparation for Theorem 3. Write $\mathcal{T}_\xi$ for the update map $\Psi \mapsto \Psi'$ on $\mathbb{R}^k$ (coordinates $c = 0, \dots, k-1$) defined by $\Psi'(c) = F_\xi(\Psi(c-1), \Psi(c))$ with $\Psi(-1) = \Psi(0) + 1$. For $\xi \in [0, 1]$ it is order-preserving. By Lemma B.1, $\gamma_k^*(x)$ is the largest γ such that the iterates $\Psi_{i+1} = \mathcal{T}_{x_i} \Psi_i$, $\Psi_1 = 0$, satisfy $\Psi_i(k-1) \leq 1 - \gamma$ before every box.

Lemma B.2 (Monotonicity and splitting). *If $x \leq x'$ coordinatewise, then $\gamma_k^*(x) \geq \gamma_k^*(x')$. If x' is obtained from x by replacing one label x_i with two consecutive labels $y, z \geq 0$, $y + z = x_i$, then $\gamma_k^*(x') \leq \gamma_k^*(x)$.*

Proof. Monotonicity. Let $x \leq x'$ and let γ be feasible on x' . The increment in (B.4) is nondecreasing in x_i , and $\mathcal{T}_\xi$ is order-preserving, so by induction $\Psi_i \leq \Psi'_i$ for every i . Thus $\Psi_i(k-1) \leq \Psi'_i(k-1) \leq 1 - \gamma$ before every box, and γ is feasible on x .

Splitting. Let Ψ be the state before position i , which is common to x and x' . After the step with label y , $\Psi'(c) = \Psi(c) + y[\gamma - 1 + \tau(c)]_+$, where $\tau(c) = \Psi(c-1) - \Psi(c)$, and

$$\tau'(c) = \Psi'(c-1) - \Psi'(c) = \tau(c) + y([\gamma - 1 + \tau(c-1)]_+ - [\gamma - 1 + \tau(c)]_+) \geq \tau(c),$$

because $\tau(c-1) \geq \tau(c)$. After the step with label z ,

$$\Psi''(c) = \Psi'(c) + z[\gamma - 1 + \tau'(c)]_+ \geq \Psi(c) + (y+z)[\gamma - 1 + \tau(c)]_+ = (\mathcal{T}_{x_i} \Psi)(c).$$

Later steps are order-preserving, so the state of x' dominates that of x at every common later position, and x' has one additional feasibility check. Hence any γ feasible on x' is feasible on x . $\square$

The Poisson reference process. A Poisson process with rate one supplies potential breaks. Time is measured in cumulative break probability, so the interval $[0, B]$ contains B potential breaks in expectation. Fix a level $\gamma \in [0, 1]$. Starting with no broken wands, let $W_\gamma(t)$ be the reference count at time t , with distribution π_t^γ . The magician uses the same conservative opening rule as in Section 3.1: it assigns opening probability to total mass γ of this distribution, starting with the states with the fewest broken wands. Write $\varphi_t^\gamma(w)$ for the prescribed opening probability in state w . At a potential break, the count increases by one with probability $\varphi_t^\gamma(W_\gamma(t-))$, using an independent random draw; otherwise it remains unchanged. The opening probabilities are determined before the potential break occurs. At $\gamma = 0$ the count stays at zero, and at $\gamma = 1$ every potential break is taken.

As with the discrete reference distribution, we define this process without stopping it at k broken wands. The opening rule can be carried out with k wands precisely when

$$\mathbb{P}\{W_\gamma(t) < k\} \geq \gamma \quad \text{for every } t \in [0, B].$$

This condition leaves enough probability mass on states with a surviving wand to supply the prescribed opening probability. We call the reference process with this opening rule the *Poisson magician*. Its count $W_\gamma(t)$ records only the potential breaks taken by the rule and generally is not itself Poisson distributed.

Write $\pi_t^\gamma(w) := \mathbb{P}\{W_\gamma(t) = w\}$. The mass assigned to opening in state w is

$$o_t^\gamma(w) := \min \left\{ \pi_t^\gamma(w), \left[\gamma - \sum_{v < w} \pi_t^\gamma(v) \right]_+ \right\}, \quad o_t^\gamma(-1) := 0.$$

The probability law of the Poisson magician satisfies the forward equations

$$\frac{d}{dt} \pi_t^\gamma(w) = o_t^\gamma(w-1) - o_t^\gamma(w), \quad \pi_0^\gamma(w) = \mathbf{1}\{w=0\}, \quad w \geq 0.$$

These equations are triangular and Lipschitz on each finite prefix of states. They determine a unique law represented by a pure birth process with jump rate $\varphi_t^\gamma(w) = o_t^\gamma(w)/\pi_t^\gamma(w)$ when $\pi_t^\gamma(w) > 0$, and rate zero otherwise. Each rate is at most one, so a Poisson clock with rate one and independent uniform marks constructs the process without explosion. Under this construction, $W_\gamma(t)$ is at most the number of clock events by time t . The total opening mass is $\sum_w o_t^\gamma(w) = \gamma$. In particular, W_0 stays at zero and W_1 takes every clock event.

Let $\psi(t, c) := \mathbb{E}[(W_\gamma(t) - c)_+]$. Each jump from state w raises $(w - c)_+$ by $\mathbf{1}\{w \geq c\}$. Therefore, without imposing capacity,

$$\begin{aligned} \frac{d}{dt} \psi(t, c) &= \sum_{w \geq c} o_t^\gamma(w) = [\gamma - \mathbb{P}\{W_\gamma(t) < c\}]_+ \\ &= [\gamma - 1 + \psi(t, c-1) - \psi(t, c)]_+. \end{aligned}$$

Thus the first k coordinates of these expected excess counts form the following closed system.

The differential system. For $\gamma \in [0, 1]$, let $\psi(t) = (\psi(t, c))_{c=0}^{k-1}$ solve

$$\frac{d}{dt} \psi(t, c) = [\gamma - 1 + \psi(t, c-1) - \psi(t, c)]_+, \quad \psi(0, c) = 0, \quad c = 0, \dots, k-1, \quad (\text{B.9})$$

with the convention $\psi(t, -1) = \psi(t, 0) + 1$. The variable t is cumulative label mass. The right side, $\mathcal{F}^\gamma(\psi)$, is Lipschitz with constant 2 in the supremum norm, so (B.9) has a unique solution on $[0, \infty)$. Each coordinate is nondecreasing with derivative in $[0, 1]$, and $\psi(t, 0) = \gamma t$. We say that (B.9) is *feasible on* $[0, B]$ if $\psi(t, k-1) \leq 1 - \gamma$ for all $t \in [0, B]$. Let $\gamma_k^\infty(B)$ be the supremum of the γ for which it is feasible; $\gamma = 0$ is feasible because then $\psi \equiv 0$.

This condition is equivalent to feasibility of the Poisson magician. If the differential system is feasible, then

$$\mathbb{P}\{W_\gamma(t) \geq k\} \leq \mathbb{E}[(W_\gamma(t) - k + 1)_+] = \psi(t, k-1) \leq 1 - \gamma$$

throughout $[0, B]$. Conversely, if $\mathbb{P}\{W_\gamma(t) < k\} \geq \gamma$ throughout this interval, then $o_t^\gamma(w) = 0$ for every $w \geq k$. The process therefore cannot jump beyond k , and $\psi(t, k-1) = \mathbb{P}\{W_\gamma(t) = k\} \leq 1 - \gamma$. It follows that $\gamma_k^\infty(B)$ is also the supremum of the feasible Poisson levels.

Proof of Theorem 3. Let $x^L = (B/L, \dots, B/L)$ have L coordinates, with integer $L \geq B$. We show that

$$\Gamma_k(B) = \gamma_k^\infty(B) = \lim_{L \rightarrow \infty} \gamma_k^*(x^L). \quad (\text{B.10})$$

Together with the equivalence above and attainment of the supremum, this gives the stated Poisson characterization. For the equal-label sequence, the total number of potential breaks has distribution Binomial($L, B/L$) and converges to Poisson(B). Measuring time by cumulative label mass gives the Poisson clock on $[0, B]$. The argument below also tracks the opening rule and establishes that these sequences approach the worst-case guarantee.

Step 1: the magician is an Euler scheme. For a label sequence x , let $t_i = \sum_{j < i} x_j$. By (B.4), $\Psi_{i+1} = \Psi_i + x_i \mathcal{F}^\gamma(\Psi_i)$: the magician's recursion is the explicit Euler scheme for (B.9) with step sizes x_i and grid times t_i .

Step 2: refinement raises the scheme. Let $x^{(j)}$ be obtained from x by splitting every label into 2^j equal labels. Each splitting raises the state at the end of the split step (proof of Lemma B.2), and the update maps are order-preserving. By induction, $\Psi_i \leq \Psi^{(j)}(t_i)$ for every i , where $\Psi^{(j)}(t_i)$ denotes the state of the refined scheme at the grid time t_i .

Step 3: convergence. The solution of (B.9) satisfies $|\psi(t) - \psi(s)|_\infty \leq |t - s|$, so its derivative $\mathcal{F}^\gamma(\psi(t))$ is Lipschitz in t with constant 2. The local error of an Euler step of size ξ is therefore at most ξ^2 , and the standard error recursion gives, for any scheme whose steps are at most $\bar{x}$,

$$\max_i |\Psi^{\text{Euler}}(t_i) - \psi(t_i)|_\infty \leq \frac{\bar{x}}{2} (e^{2t_{\max}} - 1).$$

In particular $\Psi^{(j)}(t_i) \rightarrow \psi(t_i)$ as $j \rightarrow \infty$, uniformly over the grid times of x .

Step 4: $\Gamma_k(B) \geq \gamma_k^\infty(B)$. Let (B.9) be feasible on $[0, B]$ at level γ , and let x satisfy $\sum_i x_i \leq B$. By Steps 2 and 3, $\Psi_i(k-1) \leq \lim_j \Psi^{(j)}(t_i, k-1) = \psi(t_i, k-1) \leq 1 - \gamma$ for every i , because $t_i \leq B$. By Lemma B.1(ii), $\gamma_k^*(x) \geq \gamma$. Taking the supremum over feasible γ and the infimum over x gives the claim.

Step 5: $\limsup_L \gamma_k^*(x^L) \leq \gamma_k^\infty(B)$ for $x^L = (B/L, \dots, B/L)$. Let $\gamma > \gamma_k^\infty(B)$. Then (B.9) is infeasible on $[0, B]$ at level γ : there is $t_0 \in (0, B]$ with $\psi(t_0, k-1) - (1 - \gamma) =: \varrho > 0$, and by continuity we may take $t_0 < B$. Let $i_L = \lfloor Lt_0/B \rfloor + 1 \leq L$, so that the grid time $t_{i_L} = (i_L - 1)B/L$ satisfies $t_0 - B/L < t_{i_L} \leq t_0$. Then $\psi(t_{i_L}, k-1) \geq \psi(t_0, k-1) - B/L$, and by Step 3 the Euler state satisfies $\Psi_{i_L}(k-1) \geq \psi(t_{i_L}, k-1) - \frac{B}{2L}(e^{2B} - 1)$. For L large, $\Psi_{i_L}(k-1) > 1 - \gamma$, so by Lemma B.1(ii) the γ -magician is infeasible on x^L and $\gamma_k^*(x^L) < \gamma$.

Conclusion. Restrict L to integers with $L \geq B$, so every x^L is admissible in (3). Then

$$\gamma_k^\infty(B) \leq \Gamma_k(B) \leq \inf_L \gamma_k^*(x^L) \leq \liminf_L \gamma_k^*(x^L) \leq \limsup_L \gamma_k^*(x^L) \leq \gamma_k^\infty(B).$$

Finally, ψ depends continuously on γ , so the set of levels at which (B.9) is feasible on $[0, B]$ is closed and contains its supremum $\Gamma_k(B)$. $\square$

The following proposition collects the properties of the guarantee curve used in the main text.

Proposition B.2 (Properties of the guarantee curve). *(i) (One wand.)* $\Gamma_1(B) = 1/(1+B)$ for every $B \geq 0$.

(ii) (Scaling.) Γ_k is nonincreasing and continuous on $[0, \infty)$, and for $0 < B \leq B'$,

$$\frac{B}{B'} \Gamma_k(B) \leq \Gamma_k(B') \leq \Gamma_k(B).$$

In particular $B \mapsto B\Gamma_k(B)$ is nondecreasing.

(iii) (Many wands.) For every $B > 0$,

$$0 \leq \min\{1, k/B\} - \Gamma_k(B) \leq (k+3)^{-1/2}.$$

Proof of Proposition B.2(i). For $k = 1$, the recursion gives

$$\Psi_i(0) = \gamma \sum_{j < i} x_j.$$

By Lemma B.1(ii), feasibility requires $\gamma(1 + \sum_{j < i} x_j) \leq 1$ before every box. The binding check is therefore at the last box, whether or not its label is positive, and $\gamma_k^*(x) = 1/(1 + \sum_{j < n} x_j)$. In particular $\Gamma_1(0) = 1$, and $\sum_i x_i \leq B$ implies $\gamma_k^*(x) \geq 1/(1 + B)$. Conversely, take any finite sequence of labels in $[0, 1]$ summing to B and append a zero label. Its last-box check gives $\gamma_k^*(x) = 1/(1 + B)$. Hence $\Gamma_1(B) = 1/(1 + B)$. $\square$

Proof of Proposition B.2(ii). Γ_k is nonincreasing because the infimum in (3) is over a set that grows with B . Let $0 < B \leq B'$, $\beta = B/B'$, and let x satisfy $\sum_i x_i \leq B'$. Run an auxiliary $\Gamma_k(B)$ -magician with labels βx , whose sum is at most B , and write H_i for its opening decision. Independently draw $\xi_i \sim \text{Bernoulli}(\beta)$ before deciding whether to open box i , and open the actual box only when $G_i = H_i \xi_i = 1$. The actual break is $H_i \xi_i Z_i$, so the auxiliary process has independent latent break indicators $\xi_i Z_i$ with probabilities βx_i . When $H_i = 1$ and $\xi_i = 0$, its break is known to be zero without opening the actual box. Thus the construction uses only observable past outcomes and independent randomization, and it is a valid box-opening strategy. The auxiliary magician is feasible by Corollary 1; it has the same broken-wand count as the actual process and opens with probability $\Gamma_k(B)$ before every box. Therefore $\mathbb{P}(G_i = 1) = \beta \Gamma_k(B)$ for every i , including zero labels. By Corollary B.1, $\gamma_k^*(x) \geq \beta \Gamma_k(B)$. Taking the infimum over x gives $\Gamma_k(B') \geq (B/B') \Gamma_k(B)$, and $B\Gamma_k(B) \leq B'\Gamma_k(B')$ is the same inequality. Continuity on $(0, \infty)$ follows from $0 \leq \Gamma_k(B) - \Gamma_k(B') \leq \Gamma_k(B)(1 - B/B') \leq (B' - B)/B'$; at 0 it follows from $1 = \Gamma_k(0) \geq \Gamma_k(B) \geq 1/(1 + B)$, by Lemma B.1(iv). $\square$

Proof of Proposition B.2(iii). Fix $B > 0$ and let $x = (B/L, \dots, B/L)$ with $L \geq B$. The $\gamma_k^*(x)$ -magician breaks $\sum_i \mathbb{P}(G_i = 1)x_i = \gamma_k^*(x)B$ wands in expectation, and this cannot exceed k . Hence $\Gamma_k(B) \leq \gamma_k^*(x) \leq k/B$; also $\Gamma_k \leq 1$. Theorem 1 gives $\Gamma_k(k) \geq 1 - (k+3)^{-1/2}$. Write $\beta = B/k$. For $\beta \geq 1$, part (ii) with the sum of break probabilities bounded by k gives $\Gamma_k(k)/\beta \leq \Gamma_k(\beta k) \leq 1/\beta$, so $0 \leq 1/\beta - \Gamma_k(\beta k) \leq (1 - \Gamma_k(k))/\beta \leq (k+3)^{-1/2}$. For $\beta < 1$, $\Gamma_k(k) \leq \Gamma_k(\beta k) \leq 1$, giving $0 \leq 1 - \Gamma_k(\beta k) \leq (k+3)^{-1/2}$. $\square$

C Proofs for Section 4

C.1 Implementation of an eligibility rule

Fix h and let $\gamma = \Gamma_k(B(h))$. Every $\sigma \in \Sigma(\mathcal{C})$ has $\sum_i p_{s_i}(h) = \sum_s c_s(\sigma) p_s(h) \leq B(h)$, so by Corollary 1 the γ -magician with labels $x_i = p_{s_i}(h)$ is feasible along σ . Under Assumption 1(ii) the eligibility indicators $Z_i = h_{s_i}(X_i)$ are independent with $\mathbb{P}(Z_i = 1) = p_{s_i}(h)$. They serve as latent break indicators: opening the corresponding box authorizes treatment, and a wand breaks, consuming one slot, exactly when the applicant is eligible and the box is opened. The labels are exact, and Lemma B.1(iii) gives $\mathbb{P}(G_i = 1) = \gamma$ at every position. Conditional on σ , this opening decision is independent of the current covariates, so $\mathbb{P}(A_i = 1 \mid X_i = x; \sigma) = \gamma$ for almost every eligible x . The magician's decision at position i depends only on $(s_1, \dots, s_i)$, on earlier treatments and on its own randomization, so it is an implementation rule for h , and it is arrival-neutral at level γ on $\Sigma(\mathcal{C})$.

For comparison with arbitrary implementation rules, take $\mathcal{F}_i$ in Proposition B.1 to be generated by $X_1, \dots, X_i$ and an implementation rule's randomization up to step i . The eligibility indicator $Z_i = h_{s_i}(X_i)$ is independent of $\mathcal{F}_{i-1}$. A treatment-probability floor $a \in [0, 1]$ requires $\mathbb{P}(A_i = 1 \mid X_i = x; \sigma) \geq a$ for every position i and almost every eligible x . Integrating over eligible covariates then gives $\mathbb{P}(A_i = 1 \mid Z_i = 1) \geq a$. At every position with $x_i > 0$, capacity therefore gives

$$a \leq \mathbb{P}(A_i = 1 \mid Z_i = 1) \leq \mathbb{P}(W_i \leq k - 1) = 1 - \mathbb{E}[(W_i - k + 1)_+].$$

Together with Proposition B.1, this implies $\Psi_i(k - 1) \leq 1 - a$ at every positive label. Let x^+ be the subsequence obtained by deleting the zero labels. Zero-label steps leave the reference state unchanged, so these inequalities make the a -magician feasible on x^+ . When x^+ is nonempty, any treatment floor is consequently at most $\gamma_k^*(x^+)$, and the magician run on that subsequence attains it. If there are no positive labels, treatment is almost surely absent and the floor restriction is vacuous.

Deleting zero labels is essential for this comparison. A zero-probability applicant requires no treatment, whereas opening a zero-break-probability box still requires an unbroken wand. For example, with $k = 1$ and $x = (1, 0)$, the largest common box-opening probability is $\gamma_k^*(x) = 1/2$, while the first applicant can be treated surely and the second is never eligible, giving treatment floor $\gamma_k^*(x^+) = 1$. This distinction does not change the uniform curve $\Gamma_k(B)$. Its lower guarantee holds for all boxes by Corollary 1; when $B > 0$, its upper bound is approached by the strictly positive equal-label sequences in (B.10). The comparison above applies to those sequences even when decisions use current covariates. Thus $\Gamma_k(B)$ is also the largest treatment floor that can be guaranteed using only k and the bound B over all such independent sequences. This does not imply optimality when further restrictions in $\mathcal{C}$ are used.

C.2 Gate-based implementation

Lemma C.1. *Fix h and σ . Suppose that M has gate form: $A_i = Z_i G_i$, where the gate $G_i \in \{0, 1\}$ depends only on the observed environments $(s_1, \dots, s_i)$, the history before i , and independent randomization. Then $\mathbb{P}(A_i = 1 \mid X_i = x; \sigma) = \mathbb{P}(G_i = 1 \mid \sigma)$ for almost every eligible x , and $V_\sigma(h, M) = \sum_{i=1}^n \mathbb{P}(G_i = 1 \mid \sigma) \mu_{s_i}(h)$.*

Proof. Conditional on σ , the gate G_i depends only on earlier applicants and independent randomization, so it is independent of $(X_i, Y_i(0), Y_i(1))$. Hence $\mathbb{P}(A_i = 1 \mid X_i = x) = h_{s_i}(x) \mathbb{P}(G_i = 1)$ and $\mathbb{E}[A_i \Delta_i] = \mathbb{P}(G_i = 1) \mathbb{E}[h_{s_i}(X_i) \Delta_i] = \mathbb{P}(G_i = 1) \mu_{s_i}(h)$. Summing over i gives the welfare identity. $\square$

C.3 Welfare comparisons

Proof of Proposition 1. The empty rule gives $V_*^{\text{mag}} \geq 0$. Each magician implementation is an admissible arrival-neutral implementation, so $V_*^{\text{mag}} \leq V^{\text{AN}}$. Dropping neutrality gives $V^{\text{AN}} \leq V^{\text{cap}}$, and replacing the realization-wise cap by its conditional expectation for every σ gives $V^{\text{cap}} \leq V^{\text{exp}}$. If $V_*^{\text{mag}} > 0$, all four values are positive. Their ordering implies $\text{PoC}, \text{PoAN} \geq 1$ and

$$\text{PoC} \times \text{PoAN} = \frac{V^{\text{exp}}}{V^{\text{AN}}} \leq \frac{V^{\text{exp}}}{V_*^{\text{mag}}} = \text{PoM}.$$

Now suppose the stated rule h^{exp} exists. Since Γ_k is nonincreasing and $J(h^{\text{exp}}) = V^{\text{exp}} > 0$,

$$V_*^{\text{mag}} \geq \Gamma_k(B(h^{\text{exp}}))J(h^{\text{exp}}) \geq \Gamma_k(k)V^{\text{exp}} > 0.$$

Dividing gives $\text{PoM} \leq 1/\Gamma_k(k)$, and Theorem 1 bounds this by $[1 - (k + 3)^{-1/2}]^{-1}$. $\square$

Proof of Proposition 2. With $S = 1$ and $\mathcal{C} = \{n\}$, $B(h_t) = np(t)$ and $J(h_t) = n\mu(t)$, because $\mathbb{E}[\mathbf{1}\{\tau(X) \geq t\}(Y(1) - Y(0))] = \mathbb{E}[\mathbf{1}\{\tau(X) \geq t\}\tau(X)]$. Since f is continuous at t^* , $p'(t^*) = -f(t^*)$ and $\mu'(t^*) = -t^*f(t^*)$. At an interior maximizer,

$$0 = \Phi'(t^*) = -nf(t^*)[\Gamma'_k(B^*)n\mu(t^*) + \Gamma_k(B^*)t^*].$$

Dividing by $nf(t^*)\Gamma_k(B^*) > 0$ and using $n\mu(t^*) = B^*\bar{\tau}(t^*)$ gives (7). Since Γ_k is nonincreasing, $\varepsilon_k \geq 0$. Since $B\Gamma_k(B)$ is nondecreasing (Proposition B.2(ii)), its derivative $\Gamma_k(B) + B\Gamma'_k(B)$ is nonnegative at B^* , so $\varepsilon_k \leq 1$. For $k = 1$, $\Gamma_1(B) = 1/(1 + B)$ gives $\varepsilon_1(B) = B/(1 + B)$. $\square$

D Proofs for Section 5

Error bounds. Fix $\delta \in (0, 1)$. For each environment, define

$$\ell_Y = \log(2S|\mathcal{H}|/\delta), \quad \epsilon_s = \frac{1}{\eta} \sqrt{\frac{2\ell_Y}{N_s}}.$$

The common bound in (9) can be written as $\mathcal{E} = \max_{c \in \mathcal{C}} \sum_s c_s \epsilon_s$. For a rule with $h_s \equiv 0$, both $\hat{\mu}_s(h)$ and $\mu_s(h)$ are zero. Set $\epsilon_s(h) = 0$ in that case and $\epsilon_s(h) = \epsilon_s$ otherwise, and let

$$\mathcal{E}(h) = \max_{c \in \mathcal{C}} \sum_s c_s \epsilon_s(h) \leq \mathcal{E}.$$

Let $\mathcal{G}_Y$ be the event that $|\hat{\mu}_s(h) - \mu_s(h)| \leq \epsilon_s(h)$ for every environment s and rule h .

Lemma D.1. *Under Assumption 2 and the conditions on $\mathcal{H}$ in Section 5, $\mathbb{P}(\mathcal{G}_Y) \geq 1 - \delta$. On $\mathcal{G}_Y$, $|\hat{J}(h) - J(h)| \leq \mathcal{E}(h) \leq \mathcal{E}$ for every $h \in \mathcal{H}$.*

Proof of Proposition 3 and Lemma D.1. For each s and h , the summands of $\hat{\mu}_s(h)$ have mean $\mu_s(h)$ by random assignment and lie in $[-1/\eta, 1/\eta]$. Hoeffding's inequality (Hoeffding, 1963)

and a union bound over at most $S|\mathcal{H}|$ pairs give $\mathbb{P}(\mathcal{G}_Y) \geq 1 - \delta$, using the exact zero error when $h_s \equiv 0$.

On $\mathcal{G}_Y$, for every h and $c \in \mathcal{C}$,

$$\left| \sum_s c_s \hat{\mu}_s(h) - \sum_s c_s \mu_s(h) \right| \leq \sum_s c_s \epsilon_s(h) \leq \mathcal{E}(h).$$

Taking minima over the same set $\mathcal{C}$ gives $|\hat{J}(h) - J(h)| \leq \mathcal{E}(h)$ simultaneously for all h . $\square$

Rule-specific welfare bounds. Write $\Gamma^* = \gamma(h^*)$. On $\mathcal{G}_Y$, the learned rule satisfies the welfare certificate

$$V(\hat{h}, \widehat{M}) \geq \gamma(\hat{h})\{\hat{J}(\hat{h}) - \mathcal{E}(\hat{h})\} \quad (\text{D.1})$$

and the refined regret bound

$$\begin{aligned} 0 \leq V_*^{\text{mag}} - V(\hat{h}, \widehat{M}) &\leq \gamma(\hat{h})\mathcal{E}(\hat{h}) + \Gamma^*\mathcal{E}(h^*) \\ &\leq (1 + \Gamma^*)\mathcal{E} \leq 2\mathcal{E}. \end{aligned} \quad (\text{D.2})$$

The certificate is a lower bound on the welfare of the selected rule computed from the data; the regret bound compares that rule with the population optimum. The proof below establishes both inequalities and Theorem 4.

Proof of Theorem 4. (i) The explicit stop in Definition 2 enforces the capacity on every realization.

(ii) Condition on any realized training sample and fix $\sigma \in \Sigma(\mathcal{C})$. The rule $\hat{h}$ is now fixed, and the deployment is independent of the training data. The eligibility indicators are therefore independent with probabilities $x_i = p_{s_i}(\hat{h})$. These labels are exact because each P_s^X is known, and

$$\sum_i x_i = \sum_s c_s(\sigma) p_s(\hat{h}) \leq B(\hat{h}).$$

By Corollary 1, the $\gamma(\hat{h})$ -conservative magician is feasible and opens each relevant gate with probability $\gamma(\hat{h})$. Its gate depends only on earlier applicants, the environment sequence observed so far, and independent randomization; it is independent of the current covariates conditional on the data and σ . The explicit stop never binds when an eligible applicant's gate opens. Lemma C.1 thus gives the stated conditional treatment probability, and

$$V_\sigma(\hat{h}, \widehat{M}) = \gamma(\hat{h}) \sum_s c_s(\sigma) \mu_s(\hat{h}).$$

Minimizing over σ yields $V(\hat{h}, \widehat{M}) = \gamma(\hat{h})J(\hat{h})$. No concentration event is used in this argument.

(iii) By Lemma D.1, on $\mathcal{G}_Y$ we have $|\hat{J}(h) - J(h)| \leq \mathcal{E}(h)$ for every h . Part (ii) then gives (D.1).

By the optimality of $\hat{h}$ in (8),

$$\begin{aligned} V_*^{\text{mag}} - V(\hat{h}, \widehat{M}) &= \Gamma^* J(h^*) - \gamma(\hat{h}) J(\hat{h}) \\ &= \Gamma^* \{J(h^*) - \hat{J}(h^*)\} + \{\Gamma^* \hat{J}(h^*) - \gamma(\hat{h}) \hat{J}(\hat{h})\} \\ &\quad + \gamma(\hat{h}) \{\hat{J}(\hat{h}) - J(\hat{h})\} \\ &\leq \Gamma^* \mathcal{E}(h^*) + \gamma(\hat{h}) \mathcal{E}(\hat{h}). \end{aligned}$$

This regret is nonnegative because h^* maximizes the true objective and the deployed welfare equals that objective at $\hat{h}$. The remaining inequalities in (D.2) use $\mathcal{E}(h) \leq \mathcal{E}$ and $\gamma(h) \leq 1$. Since $\mathbb{P}(\mathcal{G}_Y) \geq 1 - \delta$, this proves part (iii). With $N_s = N$, every composition has total n , so $\mathcal{E} = n\eta^{-1}\sqrt{2\ell_Y/N}$. Divide the bound by n and apply Lemma D.1. $\square$

E Unknown Covariate Distributions

When the covariate distributions are unknown, the planner must estimate how often each rule makes an applicant eligible as well as how much treatment helps those applicants. We replace part (iii) of Assumption 1 with a covariate sample, while retaining the independent arrivals and stable environment distributions in parts (i)–(ii). The procedure uses upper confidence bounds on eligibility rates to keep later applicants' treatment probabilities from falling below the announced level. Its implementation guarantee is a common lower bound: unlike Theorem 4, it need not give every eligible applicant exactly the same probability.

The reason for using upper bounds concerns the magician's use of capacity. If a label understates the true probability of eligibility, more slots may be used than its reference process predicts. A label that overstates this probability leaves more slots available and preserves the promised floor. The following lemma, proved in Appendix E.1, makes this comparison precise.

Lemma E.1 (Overstated labels). *Let $u \in [0, 1]^S$, let $\sigma \in \Sigma(\mathcal{C})$, and write $u(\sigma) = (u_{s_1}, \dots, u_{s_n})$. Run the γ -conservative magician with labels $u(\sigma)$, where $\gamma \leq \gamma_k^*(u(\sigma))$. Suppose that applicant i is eligible with probability $x_i \leq u_{s_i}$, independently across applicants; x_i may vary with i . Then every eligible applicant is treated with probability $\mathbb{P}(G_i = 1) \in [\gamma, 1]$. The capacity holds on every realization, whatever the true eligibility probabilities.*

The proof couples the true eligibility process with the labelled one so that the number of used slots is never larger under the true process. Because the magician's gate is nonincreasing in the number of used slots, it opens at least as often under the true process. The extra opening probability can differ across positions, which is why the lemma gives a floor rather than arrival neutrality. We now construct labels that overstate all candidate rules' eligibility rates with high probability, and use those labels both to select a rule and to implement it.

The class $\mathcal{H}$ is finite, fixed before observing the data, contains the empty rule, and is closed under environment masks: if $h \in \mathcal{H}$ and $A \subseteq \mathcal{S}$, then $h^A \in \mathcal{H}$, where $h_s^A = h_s \mathbf{1}\{s \in A\}$.

Assumption E.1 (Data). For each environment s the planner observes, independently of each other and of the deployment:

- (i) a covariate sample of m_s independent draws of X from the unknown marginal P_s^X ;
- (ii) an experimental sample of N_s independent draws of (X, D, Y) , where X has covariate marginal P_s^X , the potential outcomes have the same conditional distribution given X as

in the deployment, $D \mid X \sim \text{Bernoulli}(e_s(X))$ with a known propensity $\eta \leq e_s(\cdot) \leq 1 - \eta$ for some $\eta \in (0, 1/2]$, $(Y(1), Y(0))$ is independent of D given X , and $Y = DY(1) + (1 - D)Y(0)$.

We keep the covariate and experimental samples separate in the stated procedure. The concentration argument can also use the experimental covariates to estimate eligibility probabilities, because it only requires a union bound across the two types of estimation error. Covariate samples, such as the records of past applicants, can be much larger than the experiment because they require no outcome data. For each rule and environment, the planner estimates the eligibility rate and the gain by

$$q_s(h) = \frac{1}{m_s} \sum_{j=1}^{m_s} h_s(X_j), \quad \hat{\mu}_s(h) = \frac{1}{N_s} \sum_{j=1}^{N_s} h_s(X_j) \left\{ \frac{D_j Y_j}{e_s(X_j)} - \frac{(1 - D_j) Y_j}{1 - e_s(X_j)} \right\}.$$

For $\delta \in (0, 1)$ let $\ell_U = \log(4S|\mathcal{H}|/\delta)$, $r_s = \sqrt{\ell_U/(2m_s)}$, and $\epsilon_s^U = \eta^{-1} \sqrt{2\ell_U/N_s}$. For a rule with $h_s \equiv 0$, both estimates are exactly zero; set $r_s(h) = \epsilon_s^U(h) = 0$ in that case and $r_s(h) = r_s$, $\epsilon_s^U(h) = \epsilon_s^U$ otherwise. Let

$$\mathcal{R} = \max_{c \in \mathcal{C}} \sum_s c_s r_s, \quad \mathcal{E}_U = \max_{c \in \mathcal{C}} \sum_s c_s \epsilon_s^U, \quad \mathcal{E}_U(h) = \max_{c \in \mathcal{C}} \sum_s c_s \epsilon_s^U(h),$$

and let $\mathcal{G}_U$ be the event that $|q_s(h) - p_s(h)| \leq r_s(h)$ and $|\hat{\mu}_s(h) - \mu_s(h)| \leq \epsilon_s^U(h)$ for all $s \in \mathcal{S}$ and $h \in \mathcal{H}$. By Hoeffding's inequality and a union bound, $\mathbb{P}(\mathcal{G}_U) \geq 1 - \delta$ (Lemma E.3, proved in Appendix E.1). Everything below uses only the event $\mathcal{G}_U$, so any other widths for which $\mathcal{G}_U$ has probability at least $1 - \delta$, such as empirical Bernstein bounds or a uniform band over a class of thresholds, can be used instead.

With these estimates, the planner computes for each $h \in \mathcal{H}$

$$U_s(h) = \min\{1, q_s(h) + r_s(h)\}, \quad B_U(h) = \max_{c \in \mathcal{C}} \sum_s c_s U_s(h), \quad \hat{\Gamma}(h) = \Gamma_k(B_U(h)),$$

$$\hat{J}(h) = \min_{c \in \mathcal{C}} \sum_s c_s \hat{\mu}_s(h), \quad \hat{\Phi}(h) = \hat{\Gamma}(h) \hat{J}(h).$$

In this extension, $\hat{h}$ denotes a maximizer of $\hat{\Phi}$ over $\mathcal{H}$ with $\hat{\mu}_s(\hat{h}) \geq 0$ for every s . Such a maximizer exists: if $\hat{\Phi}(h) > 0$, masking the environments with $\hat{\mu}_s(h) < 0$ weakly raises $\hat{J}(h)$ and weakly lowers $B_U(h)$, and if $\max_h \hat{\Phi}(h) \leq 0$, the empty rule is a maximizer. As applicants arrive, the program runs the $\hat{\Gamma}(\hat{h})$ -conservative magician with labels $U_{s_i}(\hat{h})$, and it treats applicant i if and only if the gate opens, she is $\hat{h}$ -eligible, and a slot remains. During deployment the magician tracks a distribution on $\{0, \dots, k\}$.

The procedure inherits the capacity from the magician's explicit stop, the guarantee from Lemma E.1, and its welfare from the comparison of $\hat{\Phi}$ with the planner's objective. Let $V_*^{\text{mag}} = \max_{h \in \mathcal{H}} V^{\text{mag}}(h)$, let h^* be a maximizer, and write $\Gamma^* = \Gamma_k(B(h^*))$.

Theorem E.1 (Policy learning with estimated eligibility rates). *Under parts (i)–(ii) of Assumption 1, Assumption E.1, and the conditions on $\mathcal{H}$ stated above:*

- (i) (Capacity.) $\sum_i A_i \leq k$ for every training sample and every deployment realization.

(ii) (*Common treatment-probability floor.*) On $\mathcal{G}_U$, for every composition in $\mathcal{C}$, every order, every position i , and $P_{s_i}^X$ -almost every x with $\hat{h}_{s_i}(x) = 1$,

$$\mathbb{P}(A_i = 1 \mid X_i = x, \text{data}; \sigma) \geq \hat{\Gamma}(\hat{h}).$$

(iii) (*Welfare.*) On $\mathcal{G}_U$,

$$\begin{aligned} V(\hat{h}, \widehat{M}) &\geq \hat{\Gamma}(\hat{h})\hat{J}(\hat{h}) - \mathcal{E}_U(\hat{h}) \\ &\geq V_*^{\text{mag}} - (1 + \Gamma^*)\mathcal{E}_U - 2\Gamma^*\mathcal{R} \\ &\geq V_*^{\text{mag}} - 2\mathcal{E}_U - 2\mathcal{R}. \end{aligned}$$

In particular, with $m_s = m$ and $N_s = N$ for all s ,

$$\frac{V_*^{\text{mag}} - V(\hat{h}, \widehat{M})}{n} \leq \frac{2}{\eta} \sqrt{\frac{2 \log(4S|\mathcal{H}|/\delta)}{N}} + 2\sqrt{\frac{\log(4S|\mathcal{H}|/\delta)}{2m}}$$

with probability at least $1 - \delta$. The bound does not depend on the deployment size n or the capacity k .

Part (ii) combines Corollary 1 with Lemma E.1: along every admissible order the labels U sum to at most $B_U(\hat{h})$, and on $\mathcal{G}_U$ they overstate the true eligibility probabilities. For part (iii), the magician offers each eligible applicant a probability between $\hat{\Gamma}(\hat{h})$ and one, so welfare is at least $\hat{\Gamma}(\hat{h})$ times the expected gain, up to the negative part of the gains, which is at most $\mathcal{E}_U(\hat{h})$ for a rule whose estimated gains are nonnegative. The first inequality in (iii) is therefore a data-dependent certificate: on $\mathcal{G}_U$, the deployed rule attains at least $\hat{\Gamma}(\hat{h})\hat{J}(\hat{h}) - \mathcal{E}_U(\hat{h})$ in every order. The comparison with h^* uses Proposition B.2(ii), which gives $\hat{\Gamma}(h^*) \geq \{B(h^*)/B_U(h^*)\}\Gamma^*$ when $B(h^*) > 0$.

The bound is free of n and k because the magician's guarantee responds to the *relative* error in the expected number of eligible applicants. By Proposition B.2(ii), overstating this expectation by a factor B_U/B lowers the guarantee by at most the same factor, and the resulting welfare loss is at most the absolute overstatement, $2\mathcal{R}$, times Γ^* . A deployment of twice the size doubles the expected number of eligible applicants and doubles the absolute error, but leaves the relative error unchanged.

The guarantee in part (ii) relies on parts (i)–(ii) of Assumption 1. If the realized composition lies outside $\mathcal{C}$, or the distribution of applicants within an environment drifts, the labels may understate eligibility probabilities and the guarantee may fail, although the capacity in part (i) still holds. Lemma E.1 shows how to restore the guarantee when the departure is bounded. If eligibility probabilities may drift by a factor of at most $1 + d$ within each environment, labels $\min\{1, (1 + d)U_s(h)\}$ and the level $\Gamma_k((1 + d)B_U(h))$ deliver part (ii) for every order, and a wider envelope $\mathcal{C}$ plays the same role for compositions. Appendix G.4 quantifies both departures.

E.1 Proofs

Overstated labels. Lemma E.1 follows from a coupling that applies to every rule whose gate probability depends on the number of used slots and is nonincreasing in it.

Lemma E.2 (Monotone coupling). *Fix σ and a gate-form rule with $\mathbb{P}(G_i = 1 \mid \text{past}) = \varphi_i(W_i)$, where $\varphi_i : \{0, \dots, k\} \rightarrow [0, 1]$ may depend on i and on $(s_1, \dots, s_i)$ but not on data, and is*

nonincreasing with $\varphi_i(k) = 0$. Let $g_i(x) = \mathbb{P}(G_i = 1)$ when $Z_1, \dots, Z_n$ are independent with $\mathbb{P}(Z_i = 1) = x_i$. If $x \leq x'$ coordinatewise, then $g_i(x) \geq g_i(x')$ for every i .

Proof. Let $(V_i, U_i)_{i \leq n}$ be independent uniform random variables. Set $Z_i = \mathbf{1}\{V_i \leq x_i\}$, $Z'_i = \mathbf{1}\{V_i \leq x'_i\}$, $G_i = \mathbf{1}\{U_i \leq \varphi_i(W_i)\}$, $G'_i = \mathbf{1}\{U_i \leq \varphi_i(W'_i)\}$, $W_{i+1} = W_i + Z_i G_i$ and $W'_{i+1} = W'_i + Z'_i G'_i$, starting from $W_1 = W'_1 = 0$. Each system has the law of the rule under the corresponding probabilities, because $\varphi_i(k) = 0$ enforces the capacity. We show by induction that $W'_i \geq W_i$. If $W'_i = W_i$, then $G'_i = G_i$ and $Z'_i \geq Z_i$, so $W'_{i+1} \geq W_{i+1}$. If $W'_i \geq W_i + 1$, then $W'_{i+1} \geq W'_i \geq W_i + 1 \geq W_{i+1}$. Since φ_i is nonincreasing, $g_i(x') = \mathbb{E}\varphi_i(W'_i) \leq \mathbb{E}\varphi_i(W_i) = g_i(x)$. $\square$

Proof of Lemma E.1. If $\gamma = 0$, the claim follows by never opening a gate. For $\gamma > 0$, the magician's gate functions are determined by the labels $u(\sigma)$ and γ , not by data, and $\pi_i(\theta_i) > 0$, so φ_i equals 1 below θ_i , lies in $[0, 1]$ at θ_i , and vanishes above it; it is nonincreasing. Because $\gamma \leq \gamma_k^*(u(\sigma))$, the magician is feasible on $u(\sigma)$, so $\theta_i \leq k - 1$ and $\varphi_i(k) = 0$ at every position, including those with zero labels. By Lemma B.1(iii), $g_i(u(\sigma)) = \gamma$. Lemma E.2 with $x \leq u(\sigma)$ gives $g_i(x) \geq \gamma$, and Lemma C.1 identifies $g_i(x)$ with the treatment probability of an eligible applicant. The explicit stop enforces the capacity. $\square$

Learning guarantees.

Lemma E.3. Under Assumption E.1 and the conditions on $\mathcal{H}$ stated above, $\mathbb{P}(\mathcal{G}_U) \geq 1 - \delta$.

Proof. Fix s and h with $h_s \neq 0$. The summands of $q_s(h)$ lie in $[0, 1]$ and have mean $p_s(h)$, so Hoeffding's inequality gives $\mathbb{P}\{|q_s(h) - p_s(h)| > r_s\} \leq 2e^{-2m_s r_s^2} = 2e^{-\ell_U}$. The summands of $\hat{\mu}_s(h)$ lie in $[-1/\eta, 1/\eta]$, an interval of length $2/\eta$, and have mean $\mu_s(h)$ by unconfoundedness and $Y \in [0, 1]$. Hence $\mathbb{P}\{|\hat{\mu}_s(h) - \mu_s(h)| > \epsilon_s^U\} \leq 2\exp\{-2N_s(\epsilon_s^U)^2/(2/\eta)^2\} = 2e^{-\ell_U}$. If $h_s \equiv 0$, both estimates equal their targets. A union bound over the at most $2S|\mathcal{H}|$ events gives $\mathbb{P}(\mathcal{G}_U^c) \leq 4S|\mathcal{H}|e^{-\ell_U} = \delta$. $\square$

Proof of Theorem E.1. (i) The explicit stop in Definition 2 enforces the capacity on every realization.

(ii) Work on $\mathcal{G}_U$. Then $p_s(h) \leq q_s(h) + r_s(h)$ and $p_s(h) \leq 1$, so $p_s(h) \leq U_s(h)$ for all s and h . Let $\sigma \in \Sigma(\mathcal{C})$. The labels $U(\sigma, \hat{h}) = (U_{s_1}(\hat{h}), \dots, U_{s_n}(\hat{h}))$ sum to $\sum_s c_s(\sigma)U_s(\hat{h}) \leq B_U(\hat{h})$, so $\gamma_k^*(U(\sigma, \hat{h})) \geq \Gamma_k(B_U(\hat{h})) = \hat{\Gamma}(\hat{h})$ by (3). Lemma E.1, applied with $u = U(\hat{h})$ and true probabilities $p_{s_i}(\hat{h}) \leq U_{s_i}(\hat{h})$, gives $\mathbb{P}(G_i = 1 \mid \text{data}; \sigma) \geq \hat{\Gamma}(\hat{h})$, and Lemma C.1 turns this into the claim. Note that $\hat{h}$ and the labels depend only on the training data, which are independent of the deployment.

(iii) *Step 1: the certificate.* Work on $\mathcal{G}_U$ and fix $\sigma \in \Sigma(\mathcal{C})$. By part (ii), the gate probabilities $g_i = \mathbb{P}(G_i = 1 \mid \text{data}; \sigma)$ lie in $[\hat{\Gamma}, 1]$, where $\hat{\Gamma} = \hat{\Gamma}(\hat{h})$. By Lemma C.1, $V_\sigma(\hat{h}, \widehat{M}) = \sum_i g_i \mu_{s_i}(\hat{h})$. For a nonnegative gain, $g_i \mu \geq \hat{\Gamma} \mu$; for a negative gain, $g_i \mu \geq \mu = \hat{\Gamma} \mu - (1 - \hat{\Gamma})|\mu|$. Hence

$$V_\sigma(\hat{h}, \widehat{M}) \geq \hat{\Gamma} \sum_i \mu_{s_i}(\hat{h}) - (1 - \hat{\Gamma}) \sum_i \mu_{s_i}(\hat{h})_-,$$

where $\mu_- = \max\{-\mu, 0\}$. On $\mathcal{G}_U$, $\mu_s(\hat{h}) \geq \hat{\mu}_s(\hat{h}) - \epsilon_s^U(\hat{h})$, and $\hat{\mu}_s(\hat{h}) \geq 0$ by the choice of $\hat{h}$, so $\mu_s(\hat{h})_- \leq \epsilon_s^U(\hat{h})$. Therefore

$$\begin{aligned} V_\sigma(\hat{h}, \widehat{M}) &\geq \widehat{\Gamma} \sum_s c_s(\sigma) \{\hat{\mu}_s(\hat{h}) - \epsilon_s^U(\hat{h})\} - (1 - \widehat{\Gamma}) \sum_s c_s(\sigma) \epsilon_s^U(\hat{h}) \\ &= \widehat{\Gamma} \sum_s c_s(\sigma) \hat{\mu}_s(\hat{h}) - \sum_s c_s(\sigma) \epsilon_s^U(\hat{h}). \end{aligned}$$

The right side is at least $\widehat{\Gamma} \hat{J}(\hat{h}) - \mathcal{E}_U(\hat{h})$. Minimizing over σ gives the first inequality.

Step 2: comparison with h^ .* Write $J^* = J(h^*)$ and $B^* = B(h^*)$. Since the empty rule has value zero, $V_*^{\text{mag}} \geq 0$ and $J^* \geq 0$. By the choice of $\hat{h}$, $\hat{\Phi}(\hat{h}) \geq \hat{\Phi}(h^*)$ and $\hat{\Phi}(\hat{h}) \geq \hat{\Phi}(\emptyset) = 0$. If $J^* \leq \mathcal{E}_U$, then $V_*^{\text{mag}} = \Gamma^* J^* \leq \Gamma^* \mathcal{E}_U$, and Step 1 gives $V(\hat{h}, \widehat{M}) \geq -\mathcal{E}_U(\hat{h}) \geq V_*^{\text{mag}} - (1 + \Gamma^*) \mathcal{E}_U$. Suppose instead that $J^* > \mathcal{E}_U$. On $\mathcal{G}_U$, for every $c \in \mathcal{C}$, $\sum_s c_s \hat{\mu}_s(h^*) \geq \sum_s c_s \mu_s(h^*) - \sum_s c_s \epsilon_s^U$, so

$$\hat{J}(h^*) \geq J^* - \mathcal{E}_U > 0.$$

Also $p_s(h^*) \leq U_s(h^*) \leq q_s(h^*) + r_s(h^*) \leq p_s(h^*) + 2r_s(h^*)$, so $B^* \leq B_U(h^*) \leq B^* + 2\mathcal{R}$. Moreover $B^* \geq J^* > 0$, because $|\mu_s| \leq p_s$. By Proposition B.2(ii),

$$\widehat{\Gamma}(h^*) = \Gamma_k(B_U(h^*)) \geq \frac{B^*}{B_U(h^*)} \Gamma_k(B^*) = \frac{B^*}{B_U(h^*)} \Gamma^*.$$

Hence

$$\hat{\Phi}(h^*) \geq \Gamma^*(J^* - \mathcal{E}_U) \frac{B^*}{B_U(h^*)} = \Gamma^*(J^* - \mathcal{E}_U) - \Gamma^*(J^* - \mathcal{E}_U) \frac{B_U(h^*) - B^*}{B_U(h^*)}.$$

Since $0 < J^* - \mathcal{E}_U \leq J^* \leq B^* \leq B_U(h^*)$, the last term is at most $\Gamma^*\{B_U(h^*) - B^*\} \leq 2\Gamma^*\mathcal{R}$. Therefore $\hat{\Phi}(\hat{h}) \geq \Gamma^* J^* - \Gamma^* \mathcal{E}_U - 2\Gamma^*\mathcal{R}$, and Step 1 gives

$$V(\hat{h}, \widehat{M}) \geq \hat{\Phi}(\hat{h}) - \mathcal{E}_U(\hat{h}) \geq V_*^{\text{mag}} - (1 + \Gamma^*) \mathcal{E}_U - 2\Gamma^*\mathcal{R}.$$

The last inequality in (iii) uses $\Gamma^* \leq 1$.

Balanced samples. With $m_s = m$ and $N_s = N$, $\mathcal{E}_U = n\epsilon^U$ and $\mathcal{R} = nr$, where $\epsilon^U = \eta^{-1} \sqrt{2\ell_U/N}$ and $r = \sqrt{\ell_U/(2m)}$. Dividing by n and using Lemma E.3 completes the proof. $\square$

F Computation of the guarantee curve

Table F.1 uses the magician's recursion with L and $2L$ equal labels, where L is at least 4000 and at least $200B$. For each B , feasibility is checked on a grid of γ that is refined around the boundary. The two discretizations differ by less than 2×10^{-4} , and the table reports the Richardson extrapolation $2\gamma_{2L} - \gamma_L$. For $k = 1$ the computed values agree with $1/(1 + B)$ to within 10^{-5} .

The simulations require Γ_k at many values of B . There we compute, for a grid of 1,000 values of γ , the label mass $T(\gamma)$ at which the Euler scheme with equal labels first violates feasibility, interpolating linearly within the crossing step. The labels have size $k/800$ for $k \leq 100$ and $k/3200$ for larger k ; we repeat with labels of half that size, extrapolate $T = 2T_{1/2} - T_1$, and invert the monotone map $\gamma \mapsto T(\gamma)$. For $k \leq 100$ the result agrees with Table F.1 to within 3×10^{-5} , and for $k = 500$ successive refinements change it by less than 10^{-4} . The procedures

Table F.1: The guarantee curve $\Gamma_k(B)$

k	B/k						Alaei (2012)
	0.5	0.75	1	1.25	1.5	2	$1 - (k + 3)^{-1/2}$
1	0.6667	0.5714	0.5000	0.4444	0.4000	0.3333	0.5000
2	0.8117	0.7043	0.6148	0.5418	0.4824	0.3933	0.5528
3	0.8791	0.7722	0.6741	0.5913	0.5230	0.4209	0.5918
4	0.9173	0.8147	0.7120	0.6224	0.5480	0.4370	0.6220
5	0.9412	0.8444	0.7389	0.6442	0.5651	0.4476	0.6464
10	0.9859	0.9179	0.8093	0.6997	0.6067	0.4715	0.7226
20	0.9987	0.9651	0.8621	0.7389	0.6333	0.4851	0.7915
50	> 0.9999	0.9940	0.9112	0.7708	0.6523	0.4939	0.8626
100	> 0.9999	0.9994	0.9366	0.7842	0.6593	0.4969	0.9015

Notes: Values solve (B.9) numerically by Euler steps with L and $2L$ equal labels ($L \geq \max\{4000, 200B\}$) and Richardson extrapolation. For $k = 1$ the values equal $1/(1 + B)$. The column $B/k = 1$ reproduces the tight constants of Jiang et al. (2025) for $k \leq 5$.

use the computed curve minus a safety margin of 2×10^{-4} . Every deployment evaluation records whether the magician’s reference process is feasible along each evaluated order; in the correctly specified designs it always was.

G Simulation details and further results

G.1 Design and evaluation

Table G.1: Magician and FCFS implementations: employment gain (percent of V^{EB}) and minimum training probability, known score distributions and effects of training

k	Magician (arrival-neutral)				First come, first served			
	Magician-adjusted rule		EB rule		EB rule		Best rule under FCFS	
	B/k	prob.	gain	gain	gain	min. prob.	gain	min. prob.
5	1.40	0.595	76.5	73.9	83.5	0.454	86.3	0.260
20	1.20	0.762	87.5	86.2	91.6	0.477	92.9	0.283
100	1.10	0.876	94.2	93.6	96.2	0.490	96.3	0.120
500	1.00	0.971	97.1	97.1	98.3	0.495	98.3	0.495

Notes: Job-training program with $n = 4k$ applicants, half from each region. “Gain” is the expected increase in employed applicants relative to no training, expressed as a percentage of V^{EB} ; “prob.” refers to receiving training. The magician implements a rule at level $\Gamma_k(B)$, which is also its minimum training probability; for the EB rule ($B = k$) this is $\Gamma_k(k)$. FCFS welfare is evaluated in the worst order. V^{EB} equals 0.106 per applicant.

Main design. Regions A and B correspond to $s = 1$ and $s = 2$. In each region, the score takes the value $u_j = (j - 1/2)/20$ with known probability $1/20$ for each $j \in \{1, \dots, 20\}$. The effects of training on employment are $\tau_1(u_j) = 0.05 + 0.35u_j$ and $\tau_2(u_j) = -0.30 + 0.90u_j$; conditional employment outcomes are $Y(0) \sim \text{Bernoulli}(0.4)$ and $Y(1) \sim \text{Bernoulli}(0.4 + \tau_s(u_j))$. A rule is a pair of thresholds $(t_1, t_2) \in \{0, \dots, 20\}^2$, with $h_s(j) = \mathbf{1}\{j > t_s\}$; $t_s = 20$ excludes all applicants from region s . In the main learning experiment, experimental data consist of N experimental participants per region, with score index j , training assignment $D \sim \text{Bernoulli}(1/2)$, and employment outcome Y ; the simulation generates multinomial counts of these observations. No separate

sample of scores is needed because the score distribution is known. The welfare estimate uses the score $2DY - 2(1 - D)Y$. With $S = 2$, $|\mathcal{H}| = 441$ and $\delta = 0.05$, $\ell_Y = \log(2S|\mathcal{H}|/\delta) = 10.47$, giving $\epsilon = 0.289, 0.145, 0.072, 0.036$ for $N = 1,000, 4,000, 16,000$ and $64,000$. The population benchmarks per applicant are $V^{\text{EB}}/n = 0.106$ and $V_*^{\text{mag}}/n = 0.081, 0.093, 0.100, 0.103$ for $k = 5, 20, 100, 500$. Under the expected-budget rule, the average effect among eligible applicants is 0.365 in region A and 0.465 in region B, so by Lemma G.2 the worst order for FCFS places region A first.

Orders. For regional applicant counts (c_1, c_2) , the eight test orders are the two block orders (all applicants from A before B, or all from B before A), the two alternating orders (which end with a block when $c_1 \neq c_2$), and four uniformly random permutations drawn once with a fixed seed. By Lemma G.2 the worst order for FCFS welfare is a block order, and the minimum training probability under FCFS is attained at the last position of a block order; FCFS results are therefore exact worst cases. For the main learning experiment, finite-state recursions for every distinct selected rule on all eight orders confirm both the welfare identity $\Gamma_k(B(h))J(h)$ and equality of eligible applicants' training probabilities. The maximum absolute probability discrepancy is 9×10^{-16} . These are numerical checks of the implementation; the every-order guarantee follows from Theorem 4.

Computation. Random numbers come from the seed-sequence generator of NumPy, with base seed 20261023 and keys specific to each setting. The main learning experiment uses keys $(20261023, 22, N, \text{rep})$ and reuses each experimental sample across the four deployment sizes. It computes magician welfare from the identity above and checks each distinct selected rule by recursion, taking less than a minute. The numerical guarantee curve uses the 2×10^{-4} safety adjustment of Appendix F; this adjustment is also used in the reported oracle values. The main experiment records source hashes and settings alongside its results; all experiments save their results in machine-readable files.

G.2 Computing the price decomposition

This subsection derives the population values used in Section 6.2. We first solve the expected-capacity problem, then bound the optimum under an exact cap, and finally compute the optimum with arrival neutrality. Throughout the numerical calculations, each region supplies $m = n/2 = 2k$ applicants, and the regional score distributions and training effects are those of the main design.

Expected capacity. The expected-capacity optimum is attained by a candidate eligibility rule that trains everyone eligible. To see this, fix any regional order and any assignment rule satisfying the expected-capacity constraint. Let z_{sj} be the expected number trained from region s at score u_j . Independence and the information restriction imply that its expected employment gain is $\sum_{s,j} \tau_s(u_j) z_{sj}$, with

$$0 \leq z_{sj} \leq \frac{n}{40}, \quad \sum_{s,j} z_{sj} \leq k.$$

Maximizing this linear objective over these constraints gives an upper bound for every online assignment rule. Since $k = 10(n/40)$, the optimum fills the ten region–score categories with the

largest effects. These are the top four scores in region A and the top six in region B: the smallest selected effect is 0.33875, whereas the largest excluded effect is 0.32125. Training everyone in these categories belongs to $\mathcal{H}$, has expected training count k in every order, and attains the upper bound. Consequently,

$$V^{\text{exp}} = V^{\text{EB}} = \frac{17}{160}n. \quad (\text{G.1})$$

Exact capacity. We enclose V^{cap} between the welfare guaranteed by one feasible online rule and an upper bound obtained by revealing a particular regional order in advance. Since $\mathcal{H}$ contains the rule that makes everyone eligible, eligibility places no further restriction on this optimization.

For a continuation-value vector f indexed by remaining places r , define

$$(\mathcal{B}_s f)(r) := \frac{1}{20} \sum_{j=1}^{20} \max\{f(r), \tau_s(u_j) + f(r-1)\}, \quad r \geq 1,$$

and $(\mathcal{B}_s f)(0) := 0$. The two alternatives are rejecting and training the current applicant after observing the score. With a applicants from A and b from B remaining, initialize $L_{0,0}(r) = L_{a,b}(0) = 0$ and compute

$$L_{a,b}(r) := \lfloor \min\{(\mathcal{B}_1 L_{a-1,b})(r), (\mathcal{B}_2 L_{a,b-1})(r)\} \rfloor_\varepsilon, \quad (\text{G.2})$$

omitting a term when its region has no applicants left. Here $\lfloor \cdot \rfloor_\varepsilon$ rounds down to the grid $\varepsilon\mathbb{Z}$, with $\varepsilon = (800 \cdot 10^{12})^{-1}$. After observing the current region and score, the assignment rule chooses the maximizing alternative in $\mathcal{B}_s$ using the corresponding next-count vector L . For either possible next region, (G.2) is bounded above by that rule's expected current gain plus its continuation lower bound. Induction on $a+b$ therefore shows that this one rule guarantees at least $L_{a,b}(r)$ in every fixed regional order. In particular, $L_{m,m}(k) \leq V^{\text{cap}}$.

For the upper bound, fix the order with every applicant from A preceding every applicant from B, and let the planner know this order. Its online optimum is obtained by applying $\mathcal{B}_2$ m times and then $\mathcal{B}_1$ m times to the zero vector. Rounding each update upward on the same grid gives a value $U(k)$ at least as large as this known-order optimum. Every rule in the original problem has worst-order welfare no greater than its welfare on this particular order, so

$$L_{m,m}(k) \leq V^{\text{cap}} \leq U(k). \quad (\text{G.3})$$

This argument requires neither an equivalence between fixed and adaptive order choice nor a claim that the chosen block order is worst for every rule. All grid operations use exact integers: the training effects are $(47 + 14j)/800$ in A and $(-222 + 36j)/800$ in B for $j = 0, \dots, 19$. A separate rational-arithmetic calculation of the unrounded recursions at $k = 5$ gives coincident lower and upper values,

$$V^{\text{cap}} = \frac{1256374970631536438975171}{6553600000000000000000000}.$$

At the other capacities we retain the intervals in (G.3); their widths include the gap between the two comparison problems, not just rounding error.

Arrival neutrality. To compute V^{AN} , we use the actual regional counts and eligibility probabilities. The following ordering result reduces the calculation to one label sequence per composition. Zero-probability positions are deleted because no eligible applicant arrives there, so the treatment-probability restriction is vacuous.

Lemma G.1 (Descending labels). *For a finite multiset of positive labels, let $x^\downarrow$ list them in nonincreasing order. The γ -conservative magician is feasible on every permutation of these labels if and only if it is feasible on $x^\downarrow$. Its largest such level is therefore $\gamma_k^*(x^\downarrow)$.*

Proof. Fix $\gamma > 0$. Write T_x for the stop-loss update (B.4), so

$$(T_x v)_0 = v_0 + x\gamma, \quad (T_x v)_c = v_c + x[\gamma - 1 + v_{c-1} - v_c]_+ \quad (c \geq 1).$$

The map is nondecreasing in each state coordinate by (B.5), and also in x . For a reference stop-loss vector v , put $d_c = \gamma - 1 + v_{c-1} - v_c$, $q_c = [d_c]_+$ for $c \geq 1$, and $q_0 = \gamma$. The tail probabilities $v_{c-1} - v_c$ decrease in c , so $q_{c-1} \geq q_c$. For adjacent labels $a \geq b \geq 0$, the difference between applying a then b and applying b then a , in coordinate $c \geq 1$, is

$$(a - b)[d]_+ + b[d + at]_+ - a[d + bt]_+ \geq 0, \quad d = d_c, \quad t = q_{c-1} - q_c \geq 0.$$

If $d \geq 0$, the expression is zero. If $d < 0$ and $b > 0$, the inequality follows from $[d/a + t]_+ \geq [d/b + t]_+$ after multiplying by ab ; the case $b = 0$ is immediate. Coordinate zero is unchanged by the exchange. Subsequent updates preserve this coordinatewise inequality. Thus sorting a fixed prefix in descending order maximizes its terminal stop-loss vector.

By Lemma B.1(ii), feasibility requires $v_{k-1} \leq 1 - \gamma$ before each label. Every coordinate is nondecreasing along the sequence, so it suffices to check the prefix preceding the last label. Among prefixes obtained by omitting one label, keeping the largest labels gives the largest state: sort each prefix and use monotonicity in each label. Combining these two observations, the largest value to be checked occurs before the last label of $x^\downarrow$. This proves the equivalence. The level $\gamma = 0$ is feasible, and existence of the maximal level follows from Corollary B.1. $\square$

For $h \in \mathcal{H}$ and a composition c , let $x^\downarrow(c, h)$ contain c_s copies of each positive $p_s(h)$, sorted as in the lemma. Define

$$\gamma_{\mathcal{C}}^*(h) := \min_{c \in \mathcal{C}} \gamma_k^*(x^\downarrow(c, h)),$$

with $\gamma_k^*(\emptyset) := 1$. This is the largest common training probability achievable for h over all admissible fixed orders. Indeed, any such rule must satisfy the fixed-sequence upper bound on each descending order; Appendix C.1 establishes that bound even for decisions that depend on current covariates. Conversely, the conservative magician at level $\gamma_{\mathcal{C}}^*(h)$ is feasible on every order by Lemma G.1. It uses only the observed prefix and therefore gives a single admissible rule for all those orders. Combining this result with Lemma 1, and allowing level zero when $J(h) < 0$, yields

$$V^{\text{AN}} = \sup_{h \in \mathcal{H}} [\gamma_{\mathcal{C}}^*(h) J(h)]_+. \quad (\text{G.4})$$

For the 441 candidate cutoff pairs, this supremum is a maximum. If only one region has positive eligibility probability, its zero-label counterpart is simply skipped; if neither does, welfare is zero.

Numerical certification. For each candidate rule, bisection uses the finite recursion (B.4) on the descending sequence, checking $\Psi(k-1) \leq 1-\gamma$ immediately before its last positive label. Each arithmetic operation is rounded outward to enclose the reference state. The reported lower endpoint passes the feasibility test using the upper state bound; the upper endpoint either fails using the lower state bound or equals the universal bound one. Monotonicity of (B.5) in both state arguments justifies propagating these enclosures. Multiplying by the exact rational $J(h)$ and comparing all candidates encloses the maximum in (G.4). Candidates with $J(h) \leq 0$ can be omitted; a positive-gain candidate can also be omitted when

$$\min\{1, k/B(h)\}J(h)$$

is below an already attained lower bound, because neutrality implies an expected training count $\gamma B(h) \leq k$ in this fixed composition. These comparisons use rational arithmetic, and every candidate is either evaluated or excluded by this bound.

The resulting intervals for V^{AN} have width below 4.3×10^{-9} in total employment-gain units. The intervals for V^{cap} have relative width below 3.6×10^{-6} . We propagate these intervals through the price ratios and report the ratios to digits shared by both endpoints; Table G.2 gives interval midpoints and widths for V^{cap} . These are bounds on population optima, with no Monte Carlo uncertainty. As in the learning experiment, the reported $\tilde{V}_*^{\text{mag}}$ instead uses the numerical guarantee curve and its 2×10^{-4} safety adjustment from Appendix F; it is not an interval certification of the exact curve Γ_k . For each selected magician rule, its numerical level is below the certified lower endpoint for $\gamma_C^*(h)$, directly verifying feasibility in this design.

Table G.2: Population values underlying the price decomposition

k	V^{exp}	V^{cap}	Interval width	V^{AN}	$\tilde{V}_*^{\text{mag}}$
5	2.125	1.917076	1.00×10^{-14}	1.719300	1.626039
20	8.500	8.176589	1.43×10^{-5}	7.604231	7.439692
100	42.500	42.012327	1.48×10^{-4}	40.372443	40.049516
500	212.500	211.776325	1.85×10^{-4}	207.136324	206.353332

Notes: Values measure the expected increase in employed applicants. The V^{cap} column displays interval midpoints; the next column gives the widths of the certified intervals before display rounding. For $k=5$, an independent rational calculation establishes the exact optimum. The certified interval for V^{AN} has width below 5×10^{-9} in every row. The optimal eligible shares in regions A and B are (0.35, 0.35), (0.30, 0.30), (0.25, 0.30), and (0.20, 0.30) for increasing k . Magician values retain the numerical approximation and safety adjustment used in the learning experiment.

G.3 First come, first served

The simulations evaluate first-come-first-served exactly in its worst orders. The following lemma identifies them.

Lemma G.2. *Fix h and a composition c , and consider first-come-first-served along orders σ with $c(\sigma) = c$. Let $\bar{\tau}_s = \mu_s(h)/p_s(h)$ for regions with $p_s(h) > 0$.*

- (i) *The training probability of an eligible applicant at position i is $\mathbb{P}\{\sum_{j < i} Z_j \leq k-1\}$. It is nonincreasing in i .*
- (ii) *Welfare is minimized over such orders by any block order that places the regions in increasing order of $\bar{\tau}_s$; regions with $p_s(h) = 0$ may be placed anywhere.*

(iii) *The minimum over such orders and over positions with $p_{s_i}(h) > 0$ of the training probability of an eligible applicant is attained at the last position of an order whose last applicant belongs to a region with the smallest positive $p_s(h)$.*

Proof. (i) Under first-come-first-served, $W_i = \min\{k, \sum_{j < i} Z_j\}$, and the gate $\mathbf{1}\{W_i \leq k - 1\}$ depends only on earlier applicants. The probability $\mathbb{P}\{\sum_{j < i} Z_j \leq k - 1\}$ is nonincreasing in i .

(ii) Consider adjacent positions $i, i + 1$ occupied by applicants from regions a and b . Since $W_{i+2} = \min\{k, W_i + Z_i + Z_{i+1}\}$ is symmetric in the two applicants, exchanging them leaves the law of W_j unchanged for $j \neq i + 1$, and hence leaves the welfare of every other position unchanged. With the order (a, b) , the welfare at the two positions is $\mu_a \mathbb{P}(W_i \leq k - 1) + \mu_b \{\mathbb{P}(W_i \leq k - 1) - p_a \mathbb{P}(W_i = k - 1)\}$. Subtracting the corresponding expression for (b, a) gives

$$V_{(a,b)} - V_{(b,a)} = \mathbb{P}(W_i = k - 1) \{\mu_a p_b - \mu_b p_a\},$$

which is nonpositive when $\bar{\tau}_a \leq \bar{\tau}_b$, and zero when $p_a = 0$ or $p_b = 0$ (then the corresponding gain is also zero). Starting from any order, exchanging adjacent applicants whose regions are out of increasing order of $\bar{\tau}$ therefore weakly lowers welfare, and finitely many exchanges reach a block order sorted by $\bar{\tau}$. Block orders that differ only in the placement of regions with $p_s = 0$, or in the order of regions with equal $\bar{\tau}$, have equal welfare by the same identity.

(iii) By (i), for a fixed order the minimum is attained at the last position i° with $p_{s_{i^\circ}} > 0$, and its value is $\mathbb{P}\{\sum_{j < i^\circ} Z_j \leq k - 1\}$. This probability is smallest when the applicants before i° include every applicant except one from a region with positive p , and that one has the smallest such p : a sum of independent Bernoulli variables is stochastically increasing in each success probability and in the set of summands. $\square$

G.4 Robustness to the arrival model

Both checks use the main design with known score distributions and effects of training and the magician-adjusted rule of Table G.1, implemented by the magician; under this rule region B has the larger eligibility rate (0.35 against 0.25).

Regional shares outside the assumed range. Figure G.1 (left) calibrates the magician to the composition $(n/2, n/2)$ and varies the realized share of region A, for $k = 20$. A smaller share of region A raises the expected number of eligible applicants. The promised level 0.762 holds whenever the realized share is at least 0.45. When the share is 0.25, so that this expected number is 8 percent above the calibrated value, the minimum training probability falls to 0.35 because late applicants find no slot left. Calibrating instead to the envelope of shares in $[1/4, 3/4]$ lowers the level to 0.716 and restores the guarantee for every composition in it. For $k = 100$ the minimum falls from 0.876 to 0.25, and the envelope costs a level of 0.845 instead of 0.876.

Changing eligibility rates within each region. Figure G.1 (right) lets eligibility rates drift linearly from $(1 - d)p_s$ to $(1 + d)p_s$ over the arrival sequence, which keeps their average fixed but violates Assumption 1(ii). The magician's minimum training probability declines from 0.762 to 0.620 at $d = 0.3$ (from 0.876 to 0.682 for $k = 100$). FCFS gives the last eligible applicants a probability between 0.12 and 0.15 throughout, because the magician-adjusted rule has an expected number of eligible applicants above capacity.

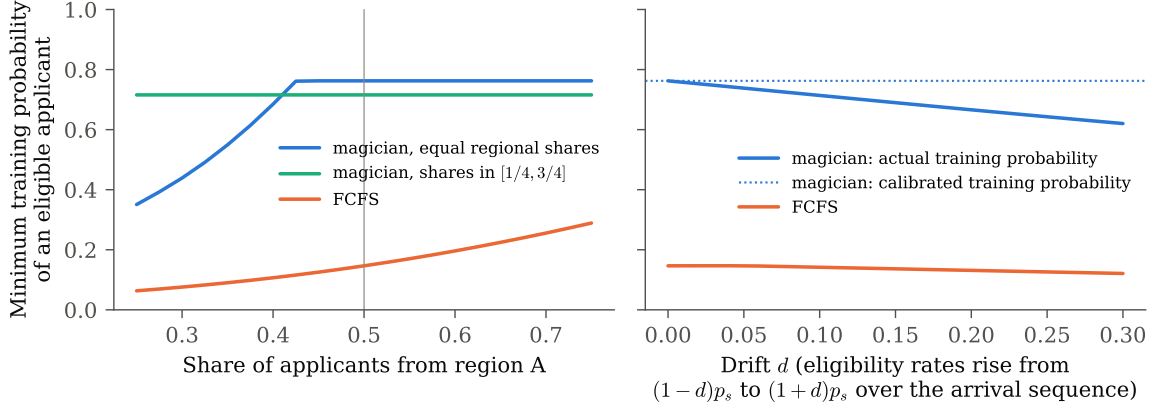


Figure G.1: Minimum training probability of an eligible applicant when region A’s share differs from the assumed one-half (left) and when eligibility rates rise over the arrival sequence within each region (right); job-training program, magician-adjusted rule, $k = 20$, $n = 80$. The dotted line in the right panel is the training probability calibrated under no drift.

G.5 All learning settings with known regional score distributions

Table G.3 includes all four experimental sample sizes. Magician-adjusted EWM attains 95.7 to 96.1 percent of V_*^{mag} at $N = 1,000$ and 99.2 to 99.4 percent at $N = 16,000$. The Monte Carlo standard error of its welfare share is at most 0.20 percentage points in any setting, below 0.09 points at $N = 4,000$ and below 0.017 points at $N = 64,000$. The uniform gain event and the welfare certificate hold in every replication. These event frequencies concern the statistical welfare bound; exact arrival neutrality follows from using known eligibility probabilities and does not require that event.

Table G.3: Learning effects of training with known regional score distributions: all settings (means over 500 replications)

n	k	Magician-adjusted EWM			EB-EWM, magician		EB-EWM, FCFS	
		gain	B/k	prob.	gain	B/k	gain	min. prob.
$N = 1,000$ <i>per region</i>								
20	5	95.7	1.31	0.632	93.2	1.00	105.7	0.445
80	20	95.8	1.17	0.781	95.1	1.00	101.1	0.473
400	100	95.8	1.09	0.879	96.0	1.00	98.6	0.489
2000	500	96.1	1.07	0.924	96.6	1.00	97.8	0.496
$N = 4,000$ <i>per region</i>								
20	5	98.3	1.35	0.616	95.0	1.00	107.6	0.448
80	20	98.3	1.19	0.770	97.0	1.00	103.0	0.474
400	100	98.2	1.10	0.877	97.8	1.00	100.5	0.489
2000	500	98.4	1.05	0.936	98.5	1.00	99.7	0.495
$N = 16,000$ <i>per region</i>								
20	5	99.2	1.38	0.604	95.9	1.00	108.5	0.452
80	20	99.4	1.18	0.771	97.9	1.00	104.0	0.476
400	100	99.3	1.09	0.883	98.8	1.00	101.5	0.489
2000	500	99.4	1.03	0.948	99.4	1.00	100.6	0.495
$N = 64,000$ <i>per region</i>								
20	5	99.7	1.38	0.602	96.3	1.00	108.9	0.454
80	20	99.8	1.19	0.766	98.2	1.00	104.4	0.477
400	100	99.7	1.10	0.877	99.1	1.00	101.9	0.490
2000	500	99.7	1.03	0.953	99.7	1.00	101.0	0.495

Notes: Main design, $k = n/4$, known regional score distributions, and no separate sample of scores. “Gain” is the expected increase in employed applicants relative to no training, as a percentage of V_*^{mag} . “Prob.” is the common training probability under magician-adjusted EWM; “min. prob.” is the minimum under FCFS. B/k is the expected number of eligible applicants relative to capacity. V_*^{mag} per applicant is 0.081, 0.093, 0.100 and 0.103 for $n = 20, 80, 400$ and 2,000. All magician calculations use the numerical curve with the 2×10^{-4} safety adjustment in Appendix F. The same learned EB rule is implemented by the magician and FCFS. FCFS is not arrival-neutral, so its welfare can exceed V_*^{mag} .