Logarithmic Regret in Multisecretary and Online Linear Programs with Continuous Valuations

Robert L. Bray

Kellogg School of Management, Northwestern University

August 28, 2023

Abstract

I study how the shadow prices of a linear program that allocates an endowment of $n\beta \in \mathbb{R}^m$ resources to n customers behave as $n \to \infty$. I show the shadow prices (i) adhere to a concentration of measure, (ii) converge to a multivariate normal under central-limit-theorem scaling, and (iii) have a variance that decreases like $\Theta(1/n)$. I use these results to prove that the expected regret in Li and Ye's (2022) online linear program is $\Theta(\log n)$, both when the customer variable distribution is known upfront and must be learned on the fly. I thus tighten Li and Ye's upper bound from $O(\log n \log \log n)$ to $O(\log n)$, and extend Lueker's (1998) $\Omega(\log n)$ lower bound to the multi-dimensional setting. I illustrate my new techniques with a simple analysis of Arlotto and Gurvich's (2019) multisecretary problem.

Keywords: online linear program; multisecretary problem; network revenue management; dual convergence; regret bounds; empirical process

1 Introduction

Caley (1875) introduced the secretary problem in the nineteenth century. The problem is to hire a man to serve as your secretary (a man because most secretaries were men back then). There are n applicants for the position whom you interview sequentially. But there's a hitch: once you interview a man, you must decide whether or not to hire him before interviewing the next man. So you face an optimal stopping problem, with the objective being to maximize the expected capability of the man you hire or, equivalently, to minimize the expectation of your *regret*, the capability difference between the most competent man and the man you hire.

Arlotto and Gurvich (2019) studied the *multisecretary* problem, which is the same as above except with $n\beta$ posts to fill, for some $\beta \in [0, 1]$. In this version of the problem, your regret is the expected capability difference between the $n\beta$ most capable men and the $n\beta$ men you hire. Arlotto and Gurvich made a striking discovery: If secretary valuations are *i.i.d.* random variables with finite support, then your expected regret is uniformly bounded across $n \in \mathbb{N}$ and $\beta \in [0, 1]$. In other words, you're never expected to make more than some finite number of hiring mistakes, regardless of the number of positions you must fill. For this breakthrough result, Arlotto and Gurvich (2019) won the 2021 Applied Probability Society Best Publication Award.

Citing an example by Robert Kleinberg, however, Arlotto and Gurvich (2019, p. 234, 251) showed that the regret could be large if one valuation had a small probability mass, which suggests that "one cannot generally expect a bound that does not depend on the minimal mass." Nevertheless, they concluded their article by explaining "At this point, it is not clear whether bounded regret is achievable also with continuous [valuation] distributions."

In Section 3, I show that bounded regret is not achievable with continuous valuation distributions. Specifically, I show that if secretary valuations are *i.i.d.* uniform random variables then the expected regret lies between $(\beta/8)(1 - \beta/8)(\log(n)/2 - \log(6))$ and $(\log(n + 1) + 7)/8$ for all $n \geq 2^{20}\beta^{-8}$ and $\beta \in [0, 1/2]$. Moreover, I show that the most obvious heuristic satisfies the upper bound.

Switching from finite to continuous valuations completely changes the mechanics of the model. With finite valuations, the probability of making a period-t hiring mistake decreases exponentially in t, whereas the expected cost of such a mistake remains constant. Hence the total regret grows with n like $\sum_{t=1}^{n} \exp(-t) = \Theta(1)$. With continuous valuations, however, the probability of making a period-t hiring mistake and the expected cost of such a mistake both decrease like $1/\sqrt{t}$. Hence the total cost grows with n like $\sum_{t=1}^{n}(1/\sqrt{t}) \cdot (1/\sqrt{t}) = \Theta(\log n)$.

In Section 4, I extend this logic and $\Theta(\log n)$ regret rate to Li and Ye's (2022) more general online linear programming" (OLP) problem. In this problem, you start with inventory vector $n\beta \in \mathbb{R}_+^m$, and you exchange inventory $a_t \in \mathbb{R}_+^m$ for utility $u_t \geq 0$ if you fulfill the period-t customer's demand. Since none of your stocks can become negative, you must carefully husband each of your m resources. But doing so is difficult, as you have no foreknowledge of the nature of demand; instead, you must learn the demand distribution the old-fashioned way—by serving customers.

The engine underlying my analysis of the online linear program is a set of shadow price convergence results I develop in Section 4.2. Li and Ye (2022, p, 2952) lamented that "there is still a lack of theoretical understanding of the properties of the dual optimal solutions," so I begin by characterizing their limiting behavior. I show that an online linear program's shadow prices (i) conform to a concentration of measure, (ii) converge to a multivariate normal under CLT-like scaling, and (iii) have a covariance matrix whose elements fall like $\Theta(1/t)$. I derive these results with a new approach that hems in the shadow prices with empirical processes.

2 Related Works

2.1 Primary Antecedents

I thought I was the first to resolve Arlotto and Gurvich's (2019) question as to whether bounded regret is achievable with continuous valuation distributions. Unfortunately, after two rounds of review, I became aware of Lueker's (1998) article, which answered the question before Arlotto and Gurvich posed it. (My understanding is that Sid Banerjee found the reference.) Lueker study the "zero-one knapsack problem," in which you successively decide whether to add a given object to your backpack. He proves the expected regret grows like $\Theta(\log n)$ under the optimal policy.

And he establishes this bound with proof unlike any other I have found in the literature. Specifically, he constructs lower and upper envelopes across the entire surface of the offline and online value functions. The induction required to creating these bounds was painstaking because he had to weaken them just so as the inventory level departed the initial resource endowment. In contrast, I use Vera and Banerjee's (2019) simpler *compensated coupling* scheme, which sums up the myopic regret incurred over the random walk traversed by the inventory level under the online policy.[1]

As Lueker's specification generalizes the multisecretary model, it also generalizes Arlotto and Xie's (2020) stochastic knapsack problem, which is the multisecretary's mirror image: one has continuous valuations and fixed capacity consumption, and the other has fixed valuations and continuous capacity consumption. Arlotto and Xie (2020, p. 190) develop an $O(\log n)$ upper bound for the regret in their model, but they do not develop a corresponding lower bound, since "it is well known that the optimal policy often lacks desirable structural properties, so proving [this lower bound] is unlikely to be easy."

Jasin (2014) extend the $O(\log n)$ upper regret bound to the multi-variate setting. Incorporating multiple dimensions enables inventories and shadow prices to veer off course in an *uncoutable* number of ways. Indeed, whereas Lueker has only to prevent these variables from being too high or too low, Jasin has to prevent them from following any unit vector too far. However, Jasin only supports a finite number of consumption bundles, as he considers the network revenue management problem in which you price a set number of products (e.g., flight itineraries), each of which comprises a set number of resources (e.g., flight legs).

Li and Ye (2022) relax the finite-product assumption, allowing a given customer's resource consumption to be any number in a bounded region of $\mathbb{R}^m$. And, more importantly, they incorporate *online learning*, as they suppose you don't know the demand distribution upfront. Not knowing this distribution induces a complex correlation between inventory levels and shadow price estimates. Li and Ye nevertheless managed to establish an $O(\log n \log\log n)$ upper regret bound.

Now, if you compare this $O(\log n \log\log n)$ bound with the prior $O(\log n)$ bounds, you can't help but wonder: Is revenue management with and without learning in the same class of difficulty? I show that they are. Specifically, I establish that the regret is $O(\log n)$ when the demand distribution is unknown and $\Omega(\log n)$ when it is known.

Removing the $\log\log n$ fudge factor from Li and Ye's upper bound requires (i) more sharply characterizing the limiting behavior of the shadow prices and (ii) more tightly controlling the inventory process. Whereas Li and Ye show that the magnitude of the period-t shadow price covariance matrix is $O((\log\log t)/t)$, I show that it is $\Theta(1/t)$. And whereas they constrain inventories for all but the last $O(\log n \log\log n)$ periods, I constrain them for all but the last $O(1)$ periods. New methodological innovations underpin both improvements.

First, I sharpen the shadow price asymptotics by applying empirical process techniques to the subgradient of the dual linear program. Casting this subgradient as an empirical process enables me to create shadow price convergence results that hold uniformly in the inventory level. And this, in turn, allows me to overcome the hopeless entanglement between the current inventory level and the current shadow price estimate.

Second, I create new techniques to constrain the inventory level's random walk. For the upper bound with a known demand distribution, I control the process with a standard martingale concentration inequality. For the upper bound with an unknown demand distribution, I split the process into martingale and drift parts, and I then apply the martingale concentration inequality to the former and inductively bound the latter, showing that the inventory level being "in control" up until period $t+1$ implies that the period-$(t+1)$ shadow price is "in control," which in turn implies that the period-t inventory level is "in control." (This induction wouldn't have been possible without the empirical process' uniform bounds.) Finally, for the lower bound, I regulate the probability of the inventory levels spiraling out of control with the cost of splitting the offline linear program into two separate linear programs. For example, suppose you have 1,000 applicants for 100 secretarial positions and can interview all the men upfront; now, suppose I told you that you can only hire ten of the last 500 candidates. This additional constraint will substantially decrease the value of your hires, with high probability. And your regret conditional on having fewer than ten open positions with 500 remaining applicants is at least as large as the cost imposed by this offline constraint. Hence, the probability of having such a low inventory level must be sufficiently small, or the optimal policy would violate the $O(\log n)$ regret upper bound.

Before discussing the rest of the literature, let me recapitulate this work's primary contributions relative to what came before.

- I extend Lueker's (1998) $\Theta(\log n)$ regret convergence rate to a multivariate setting, both with and without online learning. Neither generalization is trivial (see the discussion at the end of Section 4.1).

- I show that we can tightly regulate an online linear program's shadow prices by casting the subgradient of the dual value function as an empirical process.

- I use this newfound control over shadow prices to establish their asymptotic normality, concentration of measure, and $\Theta(1/t)$ variance. (N.B., the last result does not stem from the first result because convergence in distribution does not imply convergence of variance.)

- I create new methods for constraining the inventory process under the online policy.

2.2 Recent Developments

I will now discuss some noteworthy advancements that emerged since I first circulated my results. First, Balseiro et al. (2023) have produced an insightful and comprehensive survey article that organizes models corresponding to "dynamic pricing with capacity constraints, dynamic bidding with budgets, network revenue management, online matching, and order fulfillment" under a unified umbrella, "dynamic resource-constrained reward collection (DRC^2) problems." The DRC^2 framework is similar to Vera et al.'s (2019) "online resource allocation" framework, except it can accommodate an infinite number of customer types. Accordingly, whereas the online linear program cannot be positioned in Vera et al.'s framework, it can be positioned in Balseiro et al.'s (2023) framework (by making utilities bounded and the demand distribution known). Balseiro et al. explain that their class of problems is especially amenable to the "*certainty-equivalent principle:* replace quantities by their expected values and take the best actions given the current history." And, indeed, Li and Ye (2022) and I employ this approach to upper bound the regret of the online linear program.

An anonymous reviewer brought Balseiro et al.'s article to my attention, highlighting that we both independently devised how to exploit the certainty equivalent heuristic's martingale inventories:

> I want to note that Balseiro, Besbes, and Pizarro (2021) were written after the initial draft of [Bray's] paper. The certainty equivalent (CE) heuristic studied seems like a close cousin of [Bray's] "martingale policy." Specifically, Balseiro, Besbes, and Pizarro (2021) show that under similar assumptions, the resource vector under the CE heuristic evolves like a martingale until it exits some region. ... Both heuristics seem to achieve a Log(T) regret under seemingly similar conditions (the author may want to comment if there are key differences). I want to emphasize that I do not think Balseiro, Besbes, and Pizarro (2021) diminishes the contribution of this paper because a) it is clear that the two papers are developed independent at around the same time, and b) the main contribution of this paper is the lower bound. In fact, Balseiro et al.'s work for a similar class of problems further highlights that a general lower bound is significant to the literature.

Next, Wang and Wang (2022) establish an $\Omega(\log n)$ gap between the expected online value and the fluid approximation value (as opposed to the expected offline value) in Jasin's (2014) network revenue management problem. However, they only establish this result for the one-dimensional version of the problem. For the multi-dimensional version, they show that the optimal policy yields only $O(1)$ more expected value than the policy Jasin used.

Besbes et al. (2022) point out that a multisecretary problem's $O(\log n)$ regret critically hinges on the probability density function being bounded away from zero near the acceptance-rejection threshold. Akshit Kumar explained it to me like this: If you have n applicants for $n\beta$ open positions, then the marginal applicant would have a valuation of $F^{-1}(1-\beta+O_p(\sqrt{n}))$, where F is the utility CDF. Now, if the utility PDF equals $f(u) = |u - u^*|$ in a neighborhood of $u^* \equiv F^{-1}(1-\beta)$ then we would have $F(u) = 1 - \beta + \text{sign}(u - u^*)(u - u^*)^2/2$ and hence $F^{-1}(q) = u^* + \text{sign}(q - 1 + \beta)\sqrt{2|q - 1 + \beta|}$. And in this case, the expected myopic regret would exceed $1/n$ because rather than the usual $n^{-1/2}$ tolerance, we could now only discern the marginal man's utility to within a $n^{-1/4}$ tolerance: $F^{-1}(1 - \beta + O_p(\sqrt{n})) = u^* + \text{sign}(O_p(\sqrt{n}))\sqrt{2|O_p(\sqrt{n})|} = u^* + O_p(n^{-1/4})$. To avoid these low-density regions, Besbes et al. create a version of Balseiro et al.'s (2023) certainty-equivalent principle that is "conservative with respect to gaps." Their algorithm steers the inventory random walk away from regions with high expected myopic regret.

Finally, Jiang et al. (2022) independently developed an $O(1/n)$ bound for the shadow price variance (I posted the result in January 2022, and they posted it in October). They combine this dual convergence result with a technique that's similar in spirit to Besbes et al.'s "conservative with respect to gaps" to establish an $O(\log^2 n)$ regret bound for the network revenue management problem without imposing a non-degenerate fluid limit. In contrast, previous models have assumed the fluid approximation's constraints bind with pressure or are slack, with additional leeway. Assuming extra wiggle room in the fluid model is unreasonable, however, as it implies that some buffer stocks scale *linearly* with demand, which is a way over investment since safety stocks ought to scale with the square root of sales.

3 Multisecretary Problem

I will begin with the simple multi-secretary problem to demonstrate my regret-bounding approach. Lueker (1998) has already established that this model's regret scales like $\Theta(\log n)$, but I will provide a far simpler proof, and my bounds won't have any hidden constants.

3.1 Setup

You have $n \in \mathbb{N}$ applicants for $n\beta \in \mathbb{N}$ positions, where $\beta \in [0, 1/2]$. (It suffices to consider $\beta \in [0, 1/2]$, because the the expected regret with $n\beta$ initial open slots equals that with $n(1-\beta)$ initial open slots.[2]) You interview the candidates sequentially, starting with the nth man and ending with the first man so that the applicant number corresponds with the size of the remaining candidate pool. Interviewing the tth man reveals the utility you would get from hiring him, u_t, a standard uniform random variable independent of the other candidates' utilities. After interviewing this candidate, you must hire him on the spot or reject him for good. You seek to maximize the expected total utility from your hires. Characterizing this utility will take a few steps.

First, let v_t^b denote the utility you receive starting from period t with $tb \in \mathbb{N}$ open positions.

The expectation of this variable satisfies the following Bellman equations:

$$
\begin{aligned}
\mathrm{E}(v_t^b) \equiv& \, \mathrm{E}\left(\max_{x_t \in \{0,1\}} x_t u_t + \mathrm{E}(v_{t-1}^{\psi_t^b(x_t)}) \quad \text{s.t.} \quad x_t \le tb\right), \\
\mathrm{E}(v_0^b) \equiv& 0, \\
\text{and} \quad \psi_t^b(a) \equiv& \begin{cases} (tb-a)/(t-1) & t > 1, \\ 0 & t = 1. \end{cases}
\end{aligned}
$$

I will explain the logic underlying these equations after line (1). But first note that the ψ_t^b function maps the fraction of men you can hire from period t onwards, b, and your period-t hiring decision, x_t, to the fraction of men you can hire from period-$(t-1)$ onwards. For example, if the period-t superscript is b and you hire the period-t applicant—i.e., set $x_t = 1$—then the period-$(t-1)$ superscript is $\underbrace{tb-1}_{\text{positions left}} \;/\; \underbrace{(t-1)}_{\text{applicants left}}$.

The Bellman equations above specify the following optimal action:

$$
\pi_t^b \equiv \arg\max_{x_t \in \{0,1\}} x_t u_t + \mathrm{E}(v_{t-1}^{\psi_t^b(x_t)}) \quad \text{s.t.} \quad x_t \le tb. \tag{1}
$$

The $x_t \le tb$ constraint ensures that you don't extend a job offer if you don't have any positions available—i.e., that you set $x_t = 0$ if $tb = 0$. The expression above states that you hire the period-t man (i.e., set $x_t = 1$) if you have a job opening (i.e., $1 \le tb$) and if the total expected utility conditional on hiring him (i.e., $u_t + \mathrm{E}(v_{t-1}^{\psi_t^b(1)})$) exceeds the total expected utility conditional on rejecting him (i.e., $\mathrm{E}(v_{t-1}^{\psi_t^b(0)})$).

Your corresponding realized value is

$$
v_t^b \equiv \pi_t^b u_t + v_{t-1}^{\psi_t^b(\pi_t^b)} \quad \text{and} \quad v_0^b \equiv 0.
$$

Hence, you garner value v_n^β from your n applicants and $n\beta$ positions under the expected-utility-maximizing policy. But if you could have interviewed every applicant before extending any job offers, then you would have garnered value V_n^β, where

$$
V_t^b \equiv \sum_{s=1}^{tb} h_t^s,
$$

and h_t^s is the sth highest value in $\{u_t, \cdots, u_1\}$. Since the utilities follow a uniform distribution, order statistic h_t^s follows a beta$(t-s+1, s)$ distribution.

The difference between the aggregate utility received in the offline problem and that received in the online problem is your regret:

$$R_n \equiv V_n^\beta - v_n^\beta.$$

The following two propositions bound the expectation of this random variable.

Proposition 1. *The optimal policy of the multisecretary problem yields an expected regret that grows at no more than a* $\log n$ *rate:* $\mathrm{E}(R_n) \leq (\log(n+1)+7)/8$*, for all* $n \in \mathbb{N}$ *and* $\beta \in [0, 1/2]$.

Proposition 2. *The optimal policy of the multisecretary problem yields an expected regret that grows at no less than a* $\log n$ *rate:* $\mathrm{E}(R_n) \geq (\beta/8)(1-\beta/8)(\log(n)/2 - \log(6))$*, for all* $n \geq 2^{20}\beta^{-8}$ *and* $\beta \in [0, 1/2]$.

N.B., these theorems provide non-asymptotic results—i.e., do not rely on big-O notation. Proposition 1's finite-sample bound is especially interesting, as it highlights the near worthlessness of the value of future information. For example, suppose you have a billion applicants for 500 million jobs. In this case, your online value would be around $\underbrace{(1/2+1)/2}_{\text{value of average hire}} \cdot \underbrace{500 \text{ million}}_{\text{number of hires}} = 375$ million and your offline value would exceed your online value by around $(\log(10^9+1)+7)/8 = 3.47$. Hence, knowing the billion worker utilities upfront increases your workforce's value by around 3.47/375 million = .00000093%.

3.2 Upper Bound

I will now prove Proposition 1 by showing that Algorithm 1 honors its bound. The proof has two parts: the first decomposes the total regret into a sum of "myopic regrets," in the fashion of Vera and Banerjee (2019), and the second shows that the expectation of the period t myopic regret is $O(1/t)$ under the myopic-regret-minimizing Algorithm 1, and hence that the expected total regret is $O(\sum_{t=1}^n 1/t) = O(\log n)$.

To derive the policy underling Algorithm 1, suppose that you hire the tth man with tb available positions if and only if his valuation exceeds τ_t^b, where $\{\tau_t^b : t \in [n],\ tb \in \{0, \cdots, t\}\}$ is a collection of thresholds that have yet to be defined. These thresholds will satisfy $\tau_t^0 = 1$ and $\tau_t^1 = 0$ for all

$t \in [n]$, to ensure that $b_t \in [0,1]$ for all $t \in [n]$, where

$$\begin{aligned} b_n &\equiv \beta \\ \text{and} \quad b_{t-1} &\equiv \psi_t^{b_t}(\mathbb{1}\{u_t > \tau_t^{b_t}\}). \end{aligned}$$

In other words, you start period t with $tb_t \in \{0, \cdots, n\}$ open positions under the threshold policy. And you receive corresponding value $\bar{v}_n$, where

$$\begin{aligned} \bar{v}_t &\equiv \mathbb{1}\{u_t > \tau_t^{b_t}\}u_t + \bar{v}_{t-1}, \\ \text{and} \quad \bar{v}_0 &\equiv 0. \end{aligned}$$

Since $\mathrm{E}(\bar{v}_n) \leq \mathrm{E}(v_n^\beta)$, it follows that

$$\begin{aligned} \mathrm{E}(\bar{R}_n) \geq & \mathrm{E}(R_n), \\ \text{where} \quad \bar{R}_t \equiv & V_t^{b_t} - \bar{v}_t. \end{aligned}$$

Accordingly, it will suffice to upper bound $\mathrm{E}(\bar{R}_n)$. To this end, first note that the offline value function satisfies the following recurrence relations:

$$V_t^b = (u_t - h_{t-1}^{tb})^+ + V_{t-1}^{\psi_t^b(0)} \tag{2}$$

$$\text{and} \quad V_{t-1}^{\psi_t^b(0)} = h_{t-1}^{tb} + V_{t-1}^{\psi_t^b(1)}. \tag{3}$$

Line (2) states that if there are tb open positions then the value of increasing the applicant pool from $t-1$ to t applicants equals the option value of replacing the tbth most capable man, out of the first $t-1$ applicants, with the tth man. And line (3) states that if there are $t-1$ remaining applicants then the value of increasing the number of job openings from $(t-1)\psi_t^b(1) = tb - 1$ to $(t-1)\psi_t^b(0) = tb$ positions equals the value of the tbth best man out of the $t-1$ candidates.

Algorithm 1.

1. input n, β, $\{u_t\}_{t=1}^n$,
2. initialize $b_n := \beta$
3. for t from n to 1 do

(a) set $x_t \coloneqq \mathbb{1}\{u_t > 1 - b_t\}$

(b) set $b_{t-1} \coloneqq \psi_t^{b_t}(x_t)$

4. end for

5. output $\{x_t\}_{t=1}^n$

Now suppose that $u_t \le \tau_t^{b_t}$, and hence that $b_{t-1} = \psi_t^{b_t}(0)$ and $\bar{v}_t = \bar{v}_{t-1}$. In this case, (2) implies that

$$\begin{aligned}\bar{R}_t &= V_t^{b_t} - \bar{v}_t \\ &= (u_t - h_{t-1}^{tb})^+ + V_{t-1}^{\psi_t^b(0)} - \bar{v}_{t-1} \\ &= (u_t - h_{t-1}^{tb})^+ + V_{t-1}^{b_{t-1}} - \bar{v}_{t-1} \\ &= (u_t - h_{t-1}^{tb})^+ + \bar{R}_{t-1}.\end{aligned}$$

Next, suppose that $u_t > \tau_t^{b_t}$, and hence that $b_{t-1} = \psi_t^{b_t}(1)$ and $\bar{v}_t = u_t + \bar{v}_{t-1}$. In this case, (2) and (3) imply that

$$\begin{aligned}\bar{R}_t &= V_t^{b_t} - \bar{v}_t \\ &= (u_t - h_{t-1}^{tb})^+ + V_{t-1}^{\psi_t^b(0)} - \bar{v}_t \\ &= (u_t - h_{t-1}^{tb})^+ + (h_{t-1}^{tb} + V_{t-1}^{\psi_t^b(1)}) - (u_t + \bar{v}_{t-1}) \\ &= (u_t - h_{t-1}^{tb})^- + \bar{R}_{t-1}.\end{aligned}$$

Combining these two recurrence relations inductively yields

$$\bar{R}_n = \bar{r}_n + \bar{R}_{n-1} = \sum_{t=1}^{n} \bar{r}_t, \tag{4}$$

$$\begin{aligned}\text{where} \quad \bar{r}_t &\equiv \mathbb{1}\{u_t \le \tau_t^{b_t}\}(u_t - h_{t-1}^{tb_t})^+ + \mathbb{1}\{u_t > \tau_t^{b_t}\}(u_t - h_{t-1}^{tb_t})^- \\ &= (\mathbb{1}\{u_t > h_{t-1}^{tb_t}\} - \mathbb{1}\{u_t > \tau_t^{b_t}\})(u_t - h_{t-1}^{tb_t}).\end{aligned}$$

In the expression above, $\bar{r}_t$ is your "myopic regret," which is the cost of your period-t hiring mistake. Total regret can always be decomposed into a sum of myopic regrets.

Now here's the key: we can integrate over u_t and h_{t-1}^{tb} when taking the expectation of $\bar{r}_t$ because these variables are independent of each other and b_t. To integrate over u_t, we use the fact that this uniform random variable satisfies $\mathrm{E}((\mathbb{1}\{u_t > h\} - \mathbb{1}\{u_t > \tau\})(u_t - h)) = h^2/2 - h\tau + \tau^2/2$, for constants h and τ. And to integrate over h_{t-1}^{tb}, we use the fact that this beta$(t - tb, tb)$ random variable satisfies $\mathrm{E}(h_{t-1}^{tb}) = 1 - b$ and $\mathrm{E}(h_{t-1}^{tb})^2 = \frac{(1-b)+t(1-b)^2}{t+1}$. These properties enable us to express the expected myopic regret in terms of b_t and $\tau_t^{b_t}$:

$$\begin{aligned}
\mathrm{E}(\bar{r}_t) &= \mathrm{E}\left(\mathrm{E}\left((\mathbb{1}\{u_t > h_{t-1}^{tb_t}\} - \mathbb{1}\{u_t > \tau_t^{b_t}\})(u_t - h_{t-1}^{tb_t}) : h_{t-1}^{tb_t} = h,\ b_t = b \right) \right) \\
&= \mathrm{E}\left(\mathrm{E}\left(((h_{t-1}^{tb_t})^2/2 - h_{t-1}^{tb_t}\tau_t^{b_t} + (\tau_t^{b_t})^2/2) : b_t = b \right) \right) \\
&= \mathrm{E}\left(\frac{(1-b_t) + t(1-b_t)^2}{2(t+1)} - \tau_t^{b_t}(1-b_t) + (\tau_t^{b_t})^2/2 \right). \qquad (5)
\end{aligned}$$

I will now minimize the expectation above by setting $\tau_t^b = 1 - b$ (as specified by Algorithm 1), in which case the expression above simplifies to

$$\mathrm{E}(\bar{r}_t) = \frac{\mathrm{E}(b_t(1-b_t))}{2(t+1)}.$$

And, with this, we find that the regret incurred under Algorithm 1 satisfies our logarithmic bound:

$$\begin{aligned}
\mathrm{E}(R_n) &\leq \mathrm{E}(\bar{R}_n) \\
&= \sum_{t=1}^{n} \mathrm{E}(\bar{r}_t) \\
&= \sum_{t=1}^{n} \frac{\mathrm{E}(b_t(1-b_t))}{2(t+1)} \\
&\leq \sum_{t=1}^{n} \sup_{b \in (0,1)} \frac{b(1-b)}{2(t+1)} \\
&= \sum_{t=1}^{n} \frac{1}{8(t+1)} \\
&\leq (\log(n+1) + 7)/8.
\end{aligned}$$

3.3 Lower Bound

I will now prove Proposition 2. The proof has four steps. The first creates an optimal-policy version of the regret decomposition derived in the last section. The decomposition is the same as before, except b_t now denotes the number of open positions under the optimal algorithm rather than under Algorithm 1. The second part of the proof shows that $\Omega(\log n)$ expected regret follows immediately from the regret decomposition, provided that there's an $\Omega(1)$ chance of b_t being bounded away from either endpoint. Finally, the third part of the proof bounds the chance of b_t being too close to one, and the fourth part bounds the chance of it being too close to zero.

To begin the proof, note that the objective in (1) is supermodular in x_t and u_t. Hence, Topkis's theorem implies that there exists threshold collection

$$\{\tau_t^b : t \in [n],\ tb \in \{0, \cdots, t\}\}, \tag{6}$$

such that the optimal policy hires the tth man with tb available positions if and only if $u_t > \tau_t^b$. And as before, these thresholds satisfy $\tau_t^0 = 1$ and $\tau_t^1 = 0$, since the optimal policy always makes exactly n job offers.

Now, since the optimal policy has a threshold structure, lines (4) and (5) imply that

$$\begin{aligned} \mathrm{E}(R_n) &= \sum_{t=1}^{n} \mathrm{E}\Big(\frac{(1-b_t)+t(1-b_t)^2}{2(t+1)} - \tau_t^{b_t}(1-b_t) + (\tau_t^{b_t})^2/2\Big) \\ &\geq \sum_{t=1}^{n} \mathrm{E}\Big(\min_{\tau}\Big(\frac{(1-b_t)+t(1-b_t)^2}{2(t+1)} - \tau(1-b_t) + \tau^2/2\Big)\Big) \\ &= \sum_{t=1}^{n} \frac{\mathrm{E}(b_t(1-b_t))}{2(t+1)}. \end{aligned} \tag{7}$$

Keep in mind that that b_t now characterizes the number of open positions under the optimal thresholds defined in (6):

$$\begin{aligned} & b_n \equiv \beta \\ \text{and} \quad & b_{t-1} \equiv \psi_t^{b_t}(\mathbb{1}\{u_t > \tau_t^{b_t}\}). \end{aligned} \tag{8}$$

Lower bounding expression (7) will require upper bounding the probability that b_t veers too closely to either endpoint. For this, I will show that $n \geq 2^{20}\beta^{-8}$ and $\sqrt{n} \leq t \leq n/2$ imply

$$\Pr(b_t < \beta/8) \geq \Pr(b_t > 1 - \beta/8) \tag{9}$$

$$\text{and} \quad \Pr(b_t < \beta/8) \leq 1/4. \tag{10}$$

Combining these bounds with line (7) yields Proposition 2:

$$\begin{aligned}
\mathrm{E}(R_n) &\geq \sum_{t=1}^{n} \frac{\Pr(\beta/8 \leq b_t \leq 1-\beta/8)\beta/8(1-\beta/8)}{2(t+1)} \\
&= \sum_{t=1}^{n} \frac{\big(1-\Pr(b_t < \beta/8) - \Pr(b_t > 1-\beta/8)\big)\beta/8(1-\beta/8)}{2(t+1)} \\
&\geq \sum_{t=\lceil\sqrt{n}\rceil}^{\lfloor n/2 \rfloor} \frac{(1-1/4-1/4)\beta/8(1-\beta/8)}{2(t+1)} \\
&\geq \int_{t=2\sqrt{n}}^{n/3} (\beta/8)(1-\beta/8)/(8t)dt \\
&= (\beta/8)(1-\beta/8)(\log(n)/2 - \log(6)).
\end{aligned}$$

Accordingly, it will suffice to establish lines (9) and (10). I will begin with the former, because it is more straightforward. Simply put, the $\{b_t\}_{t=n}^1$ is process is more likely to approach the left endpoint than the right endpoint because it starts at $\beta \leq 1/2$ and is symmetric about 1/2.

I will now formalize this intuition with a coupling argument. First, note that the problem symmetry discussed in Endnote 2 implies that the acceptance thresholds satisfy

$$\tau_t^b = 1 - \tau_t^{1-b}. \tag{11}$$

Basically, this holds because one minus a uniform is also a uniform. Second, consider the following benchmark process:

$$\begin{aligned}
\hat{b}_n &\equiv 1-\beta \\
\text{and} \quad \hat{b}_{t-1} &\equiv \psi_t^{\hat{b}_t}(\mathbb{1}\{u_t > \tau_t^{\hat{b}_t}\}).
\end{aligned}$$

The $\{b_t\}_{t=n}^1$ and $\{\hat{b}_t\}_{t=n}^1$ processes can't jump over one another, because the number of open positions can only decrease by one or remain constant in a given period. And the processes couple whenever they meet, with $b_t = \hat{b}_t$ implying $b_{t-1} = \hat{b}_{t-1}$. Accordingly, $\hat{b}_t < \beta/8$ implies $b_t < \beta/8$, and hence $\Pr(\hat{b}_t < \beta/8) \leq \Pr(b_t < \beta/8)$. Third, since one minus a uniform is also a uniform, the process $\{\hat{b}_t\}_{t=n}^1$ has the same distribution as the process $\{\tilde{b}_t\}_{t=n}^1$, where

$$\begin{aligned}
\tilde{b}_n &\equiv 1-\beta \\
\text{and} \quad \tilde{b}_{t-1} &\equiv \psi_t^{\tilde{b}_t}(\mathbb{1}\{1-u_t > \tau_t^{\tilde{b}_t}\}).
\end{aligned}$$

And with (11) these equations can be rearranged like this:

$$\begin{aligned}
1-\tilde{b}_n &= \beta \\
\text{and} \quad 1-\tilde{b}_{t-1} &= 1-\psi_t^{\tilde{b}_t}(\mathbb{1}\{1-u_t > \tau_t^{\tilde{b}_t}\}) \\
&= 1-\psi_t^{\tilde{b}_t}(\mathbb{1}\{u_t < \tau_t^{1-\tilde{b}_t}\}) \\
&= \frac{t-1-t\tilde{b}_t + \mathbb{1}\{u_t < \tau_t^{1-\tilde{b}_t}\}}{t-1} \\
&= \frac{t(1-\tilde{b}_t) - \mathbb{1}\{u_t \geq \tau_t^{1-\tilde{b}_t}\}}{t-1} \\
&= \psi_t^{1-\tilde{b}_t}(\mathbb{1}\{u_t \geq \tau_t^{1-\tilde{b}_t}\}).
\end{aligned}$$

Compare this system to (8), and you will see that $1-\tilde{b}_t = b_t$, almost surely. Accordingly, $\Pr(b_t > 1-\beta/8) = \Pr(\tilde{b}_t < \beta/8) = \Pr(\hat{b}_t < \beta/8) \leq \Pr(b_t < \beta/8)$, which establishes (9).

Finally, I will establish (10). The argument has three steps. First, I establish that the regret conditional on $b_t < \beta/8$ is at least as high as the value you'd get by replacing the the worst $\lfloor t\beta/8 \rfloor$ men hired before period t with the best $\lfloor t\beta/8 \rfloor$ men rejected after period t, which is at least as high as $\lfloor t\beta/8 \rfloor$ times the difference between the value of the $(tb_t + \lfloor t\beta/8 \rfloor)th$ best man to arrive after t and the $(n\beta - tb_t - \lfloor t\beta/8 \rfloor + 1)$th best man to arrive before t. Second, I use the binomial Chernoff to establish that there's at least a $1-1/12-1/12 = 5/6$ chance that the $(tb_t + \lfloor t\beta/8 \rfloor)th$ best man to arrive after t is at least $\beta/2$ units better than the $(n\beta - tb_t - \lfloor t\beta/8 \rfloor + 1)$th best man to arrive before t. Third, I use these results to show that the optimal policy would violate the $(\log(n+1)+7)/8$ upper regret bound if the event $b_t < \beta/8$ were not sufficiently rare.

Now to begin the proof of line (10), note that conditional on having tb_t open positions at the start of period t, the best the online policy can do is hire the best tb_t men out of the last t applicants and hire the best $n\beta - tb_t$ men out of the first $n-t$ applicants. Thus, the online value must satisfy

$$v_n^\beta \leq \sum_{s=1}^{tb_t} h_t^s + \sum_{s=1}^{n\beta - tb_t} \underleftarrow{h}_t^s,$$

where look-back order statistic $\underleftarrow{h}_t^s$ is the sth largest value in $\{u_n, \cdots, u_{t+1}\}$ (i.e., it equals h_{n-t}^s, but with the order of the applicants reversed). Further, if $b_t < \beta/8$ then the offline policy could hire the best $tb_t + \lfloor t\beta/8 \rfloor$ men out of the last t applicants and the best $n\beta - tb_t - \lfloor t\beta/8 \rfloor$ men out of the first $n-t$ applicants. Hence, the offline value must satisfy the following when $b_t < \beta/8$:

$$V_n^\beta \geq \sum_{s=1}^{tb_t+\lfloor t\beta/8 \rfloor} h_t^s + \sum_{s=1}^{n\beta - tb_t - \lfloor t\beta/8 \rfloor} \underleftarrow{h}_t^s.$$

Differencing the last two inequalities yields the following, for $b_t < \beta/8$:

$$\begin{aligned}
R_n &\geq \sum_{s=tb_t+1}^{tb_t+\lfloor t\beta/8 \rfloor} h_t^s - \sum_{s=n\beta-tb_t-\lfloor t\beta/8 \rfloor+1}^{n\beta-tb_t} \underleftarrow{h}_t^s \\
&\geq \lfloor t\beta/8 \rfloor h_t^{tb_t+\lfloor t\beta/8 \rfloor} - \lfloor t\beta/8 \rfloor \underleftarrow{h}_t^{n\beta-tb_t-\lfloor t\beta/8 \rfloor+1} \\
&\geq \lfloor t\beta/8 \rfloor (h_t^{\lfloor t\beta/4 \rfloor} - \underleftarrow{h}_t^{n\beta-\lfloor t\beta/4 \rfloor}) \\
&\geq \lfloor t\beta/8 \rfloor \mathbb{1}\{h_t^{\lfloor t\beta/4 \rfloor} \geq 1 - 3\beta/8\} \mathbb{1}\{\underleftarrow{h}_t^{n\beta-\lfloor t\beta/4 \rfloor} \leq 1 - 7\beta/8\}(7\beta/8 - 3\beta/8) \\
&\geq \lfloor t\beta^2/16 \rfloor \mathbb{1}\{h_t^{\lfloor t\beta/4 \rfloor} \geq 1 - 3\beta/8\} \mathbb{1}\{\underleftarrow{h}_t^{n\beta-\lfloor t\beta/4 \rfloor} \leq 1 - 7\beta/8\}.
\end{aligned}$$

The first line above states that your regret is at least as large as the benefit you'd get by replacing the worst $\lfloor t\beta/8 \rfloor$ men hired before period t with the best $\lfloor t\beta/8 \rfloor$ men rejected after period t. The second line maintains that the value of this difference is at least as large as $\lfloor t\beta/8 \rfloor$ (i.e., the number of men being exchanged) times the difference between $h_t^{tb_t+\lfloor t\beta/8 \rfloor}$ (i.e., the value of the worst man added) and $\underleftarrow{h}_t^{n\beta-tb_t-\lfloor t\beta/8 \rfloor+1}$ (i.e., the value of the best man removed). The remaining three lines use the fact that h_t^s decreases in its superscript to connect the bound with the following binomial Chernoff results: If $t \geq 48\log(12)/\beta$, $n \geq 336\log(12)/\beta$, and $\sqrt{n} \leq t \leq n/2$ then

$$\begin{aligned}
&\Pr(h_t^{\lfloor t\beta/4 \rfloor} \geq 1 - 3\beta/8) \geq 11/12 \\
\text{and} \quad &\Pr(\underleftarrow{h}_t^{n\beta-\lfloor t\beta/4 \rfloor} \leq 1 - 7\beta/8) \geq 11/12,
\end{aligned}$$

Accordingly, Proposition 1 and Bonferroni's inequality imply the following, for the specified range of n and t:

$$
\begin{aligned}
&(\log(n+1)+7)/8 \\
&\quad \geq \mathrm{E}(R_n) \\
&\quad \geq \lfloor t\beta^2/16 \rfloor \Pr\big(b_t < \beta/8 \ \cap \ h_t^{\lfloor t\beta/4 \rfloor} \geq 1 - 3\beta/8 \ \cap \ \underleftarrow{h}_t^{\lfloor (n-t)\beta \rfloor} \leq 1 - 7\beta/8\big) \\
&\quad \geq \lfloor \sqrt{n}\beta^2/16 \rfloor \big(\Pr(b_t < \beta/8) + 11/12 + 11/12 - 2 \big).
\end{aligned}
$$

Finally, this inequality implies (10) when $n \geq 2^{20}\beta^{-8}$ and $\sqrt{n} \leq t \leq n/2$.

4 Online Linear Programming Problem

4.1 Model

I will now extend the techniques developed in the last section to Li and Ye's (2022) online linear program.[3] See the appendix for a notation guide and the online supplement for the proofs.

As before, I will count backward from period $n \in \mathbb{N}$ to period 1. In each period, a customer arrives, and you must decide whether or not to fulfill their demand from your inventory. You begin in period n with initial inventory endowment $nb_n = n\beta$, for some given $\beta \in \mathbb{R}^m_+$, so that you have $e_j' b_n$ units of the jth resource budgeted for the "average" remaining period, where e_j is the unit vector indicating the jth position. If you satisfy the period–n customer then you exchange inventory bundle $a_n \in \mathbb{R}^m_+$ for utility u_n, so that you begin period $n-1$ with resource vector $b_{n-1} \equiv (nb_n - a_n)/(n-1)$. If, on the other hand, you reject the period–n customer, then you receive no utility and lose no resources, so that you begin period $n-1$ with resource vector $b_{n-1} \equiv nb_n/(n-1)$. And this pattern repeats so that $b_{t-1} \equiv (tb_t - a_t)/(t-1)$ if you satisfy the period–t customer and $b_{t-1} \equiv (tb_t)/(t-1)$ otherwise. The problem is dynamic because you don't observe variables u_t and a_t until the beginning of period t. These variables satisfy the following assumptions:

Assumption 1. *The customers are i.i.d.: vectors $\{(u_t, a_t)\}_{t=1}^n$ are drawn independently of one another, from joint distribution μ.*

Assumption 2. *The utilities and resource requirements are non-negative:* $u_1, a_1 \geq 0$ *almost surely.*

Assumption 3. *The utilities have finite expectation:* $\mathrm{E}(u_1) < \infty$.

Assumption 4. *The resource requirements are bounded: $a_1 \leq \alpha$, almost surely, for some $\alpha \in \mathbb{R}^m_+$.*

N.B. that u_1 can have unbounded support, whereas the other models cited in Section 2—most notably those of Lueker (1998) and Li and Ye (2022)—restrict u_1 to a finite range.

Let v_t^b denote the utility you receive from period t onwards when you begin that period with resource endowment $tb \in \mathbb{R}^m$. Since you follow the expected-utility-maximizing policy, this variable's expectation satisfies the following Bellman equations:

$$\begin{aligned} \mathrm{E}(v_t^b) \equiv& \mathrm{E}\Big(\max_{x_t \in \{0,1\}} x_t u_t + \mathrm{E}(v_{t-1}^{\psi_t^b(x_t a_t)}) \quad \text{s.t.} \quad x_t a_t \le tb\Big), \qquad (12) \\ \mathrm{E}(v_0^b) \equiv& 0, \\ \text{and} \quad \psi_t^b(a) \equiv& \begin{cases} (tb-a)/(t-1) & t > 1, \\ 0 & t = 1. \end{cases} \qquad (13) \end{aligned}$$

To better understand this system, consider the following optimal action:

$$\pi_t^b \equiv \arg\max_{x_t \in \{0,1\}} x_t u_t + \mathrm{E}(v_{t-1}^{\psi_t^b(x_t a_t)}) \quad \text{s.t.} \quad x_t a_t \le tb. \qquad (14)$$

In other words, you accept the period-t customer (i.e., set $x_t = 1$) if you have inventory enough to do so (i.e., $a_t \le tb$) and if the total expected utility conditional on satisfying this customer (i.e., $u_t + \mathrm{E}(v_{t-1}^{\psi_t^b(1)})$) exceeds the total expected utility conditional on turning them away (i.e., $\mathrm{E}(v_{t-1}^{\psi_t^b(0)})$).

Under this policy you garner total value v_n^β from your initial $n\beta$ resource endowment, where

$$v_t^b \equiv \pi_t^b u_t + v_{t-1}^{\psi_t^b(\pi_t^b a_t)} \quad \text{and} \quad v_0^b \equiv 0. \qquad (15)$$

However, if you could have observed all of the customer attributes before deciding which ones to satisfy, then you would have garnered value V_n^β, where

$$V_t^b \equiv \max_{x \in \{0,1\}^t} \sum_{s=1}^t x_s u_s \quad \text{s.t.} \quad \sum_{s=1}^t x_s a_s \le tb. \qquad (16)$$

Your regret is the difference between the utility you extract when you observe all customer variables upfront and the utility you extract when you learn these variables on the fly:

$$R_n \equiv V_n^\beta - v_n^\beta. \qquad (17)$$

Our objective is to show that $\mathrm{E}(R_n) = \Theta(\log n)$ as $n \to \infty$.

Since expanding your choice set from $\{0,1\}$ to $[0,1]$ will not make you worse off, we have

$$\begin{aligned}
\bar{V}_n^\beta \geq & V_n^\beta, \\
\text{where} \quad \bar{V}_t^b \equiv & \max_{x\in[0,1]^t} \sum_{s=1}^t x_s u_s \quad \text{s. t.} \quad \sum_{s=1}^t x_s a_s \leq tb && (18)\\
= & \min_{y\in\mathbb{R}_+^m} t\Lambda_t^b(y), && (19)\\
\Lambda_t^b(y) \equiv & b'y + \sum_{s=1}^t \Delta_s(y)^+/t, \\
\text{and} \quad \Delta_t(y) \equiv & u_t - a_t' y. && (20)
\end{aligned}$$

Line (19) is a reconfiguration of the dual of line (18)'s linear program (see Li and Ye, 2022). (To remember that it is a dual, it helps to think of Λ as an upside down V.) This dual problem has a not-necessarily-unique shadow price minimizer:

$$y_t^b \in \arg\min_{y\in\mathbb{R}_+^m} t\Lambda_t^b(y). \tag{21}$$

Since we initialized $b_n = \beta$, the problem in (19) converges, as $n \to \infty$, to the following deterministic problem:

$$\min_{y\in\mathbb{R}_+^m} \Lambda_\infty^\beta(y) \quad \text{where} \quad \Lambda_\infty^b(y) \equiv b'y + \mathrm{E}(\Delta_1(y)^+). \tag{22}$$

The following assumption endows this problem with a positive shadow price solution.

Assumption 5. *All resources are consumed in the limit: there exists $y_\infty^\beta \in \arg\min_{y\in\mathbb{R}_+^m} \Lambda_\infty^\beta(y)$ such that $y_\infty^\beta > 0$.*

Extending this assumption to accommodate constraints that are strictly slack in the limit is simple. However, it's harder to accommodate constraints that just barely hold in the limit. See Jiang et al.'s (2022) recent work for an interesting analysis of the degenerate-limit case.

The final assumption is the multivariate analog of Lueker's (1998) local restriction. Lueker imposed two critical constraints on the joint distribution of (u_1, a_1): a local restriction that holds in a neighborhood of the $u_1 = a_1' y_\infty^\beta$ level set, and a global restriction that holds across the entire breadth of the distribution. I will need only the former because all the tough calls lie at the margin. For example, the following assumption permits point masses in the distribution, so long as they do not abut the fluid model's accept-reject indifference curve.

Assumption 6. *There's a continuum of marginal customers that strain the resources in a linearly independent fashion: the Jacobian matrix* $\frac{\partial}{\partial y}\mathrm{E}(\mathbb{1}\{\Delta_1(y) > 0\}a_1)$ *exists, is full rank, and is continuous in* y *in a neighborhood of* y_∞^β.

This assumption is more straightforward than the second-order growth condition imposed by Li and Ye (2022) and many others. Indeed, it simply states that shadow prices give us complete control over inventories. To see this, note that $\mathrm{E}(\mathbb{1}\{\Delta_1(y) > 0\}a_1)$ is the mean resource consumption rate when we satisfy all customers with positive surplus utility, under shadow price vector y. Accordingly, Jacobian matrix $\frac{\partial}{\partial y}\mathrm{E}(\mathbb{1}\{\Delta_1(y) > 0\}a_1)$ maps marginal shadow price changes to marginal consumption rate changes. This matrix being full rank ensures that we can control the inventory burn rate in a linearly independent fashion by fine-tuning y. For example, marginally shifting the shadow price in the direction of $(\frac{\partial}{\partial y}\mathrm{E}(\mathbb{1}\{\Delta_1(y) > 0\}a_1))^{-1}e_i$ would marginally decrease the consumption of the ith resource, without changing that of the other resources.

Here's a simple sufficient condition that implies Assumption 6.

Example 1. Suppose that given a_1, utility u_1 has bounded and continuous conditional density function $g(u_1 : a_1)$, which almost surely satisfies $g(a_1' y_\infty^\beta : a_1) > 0$. Further, suppose that $\mathrm{E}(a_1 a_1')$ is non-singular. □

The following lemma is equivalent to Assumption 6, so you can consider it an alternative assumption:

Lemma 1. *The limiting problem's second derivative is positive and continuous at its minimizer: Hessian matrix* $\ddot{\Lambda}_\infty(y) \equiv \frac{\partial^2}{\partial y^2}\Lambda_\infty^b(y) = -\frac{\partial}{\partial y}\mathrm{E}(\mathbb{1}\{\Delta_1(y) > 0\}a_1)$ *exists, is positive definite (and hence full rank), and its elements are continuous in* y *in a neighborhood of* y_∞^β.

Combining Lemma 1 with Assumption 5 yields the following sister lemma via the implicit function theorem.

Lemma 2. *Limiting shadow prices are locally differentiable in the resource vector: if* b *is sufficiently close to* β*, then* Λ_∞^b *has a unique minimizer,* $y_\infty^b > 0$*, which is continuously differentiable—and hence Lipschitz continuous—in* b*, with* $\frac{\partial}{\partial b} y_\infty^b = -\ddot{\Lambda}_\infty(y_\infty^b)^{-1}$.

Together, Lemmas 1 and 2 imply that $\ddot{\Lambda}_\infty(y_\infty^b)$—the Hessian matrix of Λ_∞^b at its minimum—is continuous in b in a neighborhood of β. Accordingly, $\{\omega_i^b\}_{i\in[m]}$ and $\{\sigma_i^b\}_{i\in[m]}$ are likewise continuous in b, where ω_i^b is an eigenvector of $\ddot{\Lambda}_\infty(y_\infty^b)$ with eigenvalue σ_i^b. Further, since $\ddot{\Lambda}_\infty(y_\infty^b)$ is positive definite, we can take $\{\omega_i^b\}_{i\in[m]}$ to be orthonormal and take $\{\sigma_i^b\}_{i\in[m]}$ to be real numbers that satisfy $\sigma_1^b \geq \cdots \geq \sigma_m^b > 0$ (provided that b is sufficiently close to β).

Lemma 1 also implies that

$$\dot{\Lambda}_\infty^b(y) \equiv \tfrac{\partial}{\partial y}\Lambda_\infty^b(y) = b - \mathrm{E}(\mathbb{1}\{\Delta_1(y) > 0\}a_1) \tag{23}$$

exists and is continuous in y a neighborhood of y_∞^β. Unfortunately, the finite analog, Λ_t^b, is not always differentiable, but when it is, its gradient equals subgradient

$$\dot{\Lambda}_t^b(y) \equiv b - \sum_{s=1}^{t} \mathbb{1}\{\Delta_s(y) > 0\}a_s/t. \tag{24}$$

Our model is now fully characterized. Thus, we are now ready for the primary results.

Theorem 1. *The optimal policy of the online linear program* without learning *yields an expected regret that grows at no more than a* $\log n$ *asymptotic rate:* $\mathrm{E}(R_n) = O(\log n)$ *as* $n \to \infty$ *when distribution* μ *is known to the decision maker.*

Theorem 2. *The optimal policy of the online linear program* with learning *yields an expected regret that grows at no more than a* $\log n$ *asymptotic rate:* $\mathrm{E}(R_n) = O(\log n)$ *as* $n \to \infty$ *when distribution* μ *is unknown to the decision maker.*

Theorem 3. *The optimal policy of the online linear program* without learning *yields an expected regret that grows at no less than a* $\log n$ *asymptotic rate:* $\mathrm{E}(R_n) = \Omega(\log n)$ *as* $n \to \infty$ *when distribution* μ *is known to the decision maker.*

Corollary 1. *The optimal policy of the online linear program* with learning *yields an expected regret that grows at no less than a* $\log n$ *asymptotic rate:* $\mathrm{E}(R_n) = \Omega(\log n)$ *as* $n \to \infty$ *when distribution* μ *is unknown to the decision maker.*

Since knowing μ won't decrease your regret, Corollary 1 follows immediately from Theorem 3, and Theorem 1 follows immediately from Theorem 2. However, I don't call Theorem 1 a corollary because I provide an independent proof for it. Indeed, I will use the proof of Theorem 1 as a stepping stone to the proof of Theorem 2.

Also note that the single-dimensional results of Section 3 and Lueker (1998) imply none of the multi-dimensional results above—the previous findings establish that an online linear program *can* exhibit $\log n$ regret, but not that it *must* do so. Naturally, the regret could be larger for the "harder" online linear program, but the regret also be *smaller*. Indeed, while an additional restriction cannot increase the objective value, it can decrease the regret by burdening the offline problem more than the online problem. For instance, Examples 2 and 3 illustrate that adding a second constraint can reduce the expected regret from $\Theta(\log n)$ to $o(1)$ in the multisecretary problem, and from $\Theta(\log n)$ to $O(1)$ in Arlotto and Xie's (2020) stochastic knapsack problem. Hence, some constraints negate the $\log n$ regret rate; I must prove that all such negating constraints violate our assumptions.

Further, increasing the dimesionality can lower the regret by helping the period-t inventory vector, b_t, escape the neighborhood of β for which the expected myopic regret is high. Indeed, since we have imposed only local restrictions on the demand function, we can lower bound the myopic regret only when b_t lies in some ball $B_\delta(\beta)$. And the more dimensions $B_\delta(\beta)$ has, the easier it is for b_t to break free.

Finally, even if we imposed global demand restrictions so that all $\mathbb{R}^m_{>0}$ had uniformly high myopic regrets, we still couldn't easily collapse the state space to a single dimension. The natural state space collapse would project the resource and inventory vectors onto shadow prices, defining $\tilde{a}_t \equiv a_t' y_n^\beta / \|y_n^\beta\|$ and $\tilde{b}_t \equiv b_t' y_n^\beta / \|y_n^\beta\|$. But this projection doesn't neatly transform the model into Lueker's (1998) framework, because the shadow prices we project onto are a function of the data, which makes the resulting $\tilde{a}_t$ values interdependent.

Example 2. Consider the multisecretary problem of Section 3.1, but with an additional payroll budget constraint: now, in addition to the $n\beta$ available positions, you also start with $n\beta/2$ dollars, which you use to pay your workforce. The tth man commands wage u_t, so the applicants all yield the same bang for the buck. By design, you will almost certainly run out of money before you fill all the positions when n is large, both under the optimal online and offline policies. Hence, only your payroll budget constraint is relevant as $n \to \infty$. But you will never regret how you spend this budget because every dollar yields the same marginal utility. Accordingly, the regret must go to zero, almost surely, as $n \to \infty$. □

Example 3. Suppose you sequentially decide which of n items to add to a knapsack with volume $n/4$. The item volumes are independent uniform random variables, but the item values are all one dollar, so you want to add small items to your backpack. Lueker's (1998) theorem indicates that this problem has $\Theta(\log n)$ expected regret. But now suppose you have an A backpack and a B backpack, both with volume $n/4$. And suppose each item comprises an A part, which can be stored only in the A backpack, and a B part, which can be stored only in the B backpack. Furthermore, suppose that the volume of the A part is a uniform random variable and that the volume of the B part is one minus the volume of the A part, and hence also a uniform random variable. Finally, you get a dollar for each part you add to a backpack. Now, if you could put one part of an item in a knapsack and not the other part, your problem would decompose into two independent problems, each with $\Theta(\log n)$ expected regret. But if putting one part of an item into a backpack compels you to put the other part in the other backpack, then your regret is only $O(1)$. To see this, note that your regret will be proportional to the leftover space in one backpack when the other one is filled, and Corollary 4 establishes that this wasted capacity is $O(1)$ under Algorithm 2. □

4.2 Dual Convergence Results

Everything boils down to shadow prices, so we can only make progress once we understand how y_t^b converges to y_∞^b. I will thus begin the analysis by presenting four propositions that crisply characterize the shadow prices' limiting behavior.

Proposition 3. *There exists $\delta > 0$ such that $\sqrt{t}(y_t^b - y_\infty^b) \xrightarrow{d} \mathcal{N}(0, \Sigma^b)$ for all $b \in B_\delta(\beta)$, where $\Sigma^b \equiv \ddot{\Lambda}_\infty(y_\infty^b)^{-1}\,\mathrm{Cov}(\mathbb{1}\{\Delta_1(y_\infty^b) > 0\}a_1)\ddot{\Lambda}_\infty(y_\infty^b)^{-1}$ is full rank and continuous in $b \in B_\delta(\beta)$.*

Unfortunately, this proposition proved less helpful than I had hoped because the rate of convergence could depend on b—i.e., the magnitude of t required to ensure that $\sqrt{t}(y_t^b - y_\infty^b) \approx \mathcal{N}(0, \Sigma^b)$ could be unbounded in any neighborhood of β. This, unfortunately, won't do because I'll need to invoke my convergence results at a random value of b_t. Hence, rather than Proposition 3, I will use the following results, which control the limiting shadow prices *uniformly* across $b \in B_\delta(\beta)$.

Proposition 4. *There exists $\delta > 0$ such that $\mathrm{E}(\sup_{b\in B_\delta(\beta)} \|y_t^b - y_\infty^b\|^2) = O(1/t)$.*

Proposition 5. *There exists $\delta > 0$ such that $\mathrm{E}(\inf_{b\in B_\delta(\beta)} \|y_t^b - y_\infty^b\|^2) = \Omega(1/t)$.*

Corollary 2. *There exists $\delta > 0$ such that the covariance matrix of y_t^b has a $\Theta(1/t)$ spectral norm, for all $b \in B_\delta(\beta)$.*

Note, positioning the $\sup_{b\in B_\delta(\beta)}$ and $\inf_{b\in B_\delta(\beta)}$ terms inside of the expectations makes these results especially strong. We'll need this extra strength to bound the regret when μ is unknown, in which case shadow prices and inventory vectors become tangled. (Jiang et al. (2022) independently developed a $\sup_{b\in B_\delta(\beta)}$-free version of Proposition 4.)

Proposition 4 is a stronger version of Li and Ye's (2022) first theorem, which states that $\mathrm{E}(\|y_t^b - y_\infty^b\|^2) = O((\log\log t)/t)$. I had to shave off the repeated logarithms to derive a sharp $\log n$ upper bound. I did so with a new approach. I first bounded the difference between y_t^b and y_∞^b

with the difference between the limiting gradient, $\dot{\Lambda}^b_\infty(\cdot)$, and its finite analog, $\dot{\Lambda}^b_t(\cdot)$, evaluated at the shadow price midway point, $\hat{y}^b_t \equiv (y^b_t + y^b_\infty)/2$. But $\hat{y}^b_t$ is difficult to work with, so I then bounded the expected value of $\|\dot{\Lambda}^b_t(\hat{y}^b_t) - \dot{\Lambda}^b_\infty(\hat{y}^b_t)\|^2$ with the expected value of $\sup_{y \in B_{2\epsilon}(y^\beta_\infty)} \|\dot{\Lambda}^b_t(y) - \dot{\Lambda}^b_\infty(y)\|^2$. Finally, I bounded the expected value of this supremum with a classic empirical processes result.

I also used empirical processes to prove Proposition 5, which will permit the corresponding lower regret bound. Specifically, I establish this result by showing that $\sqrt{t}(y^b_t - y^b_t)$ is near $\gamma \in \mathbb{R}^m$ if $\sqrt{t}(\dot{\Lambda}^\beta_t(y) - \dot{\Lambda}^\beta_\infty(y))$ is near $\ddot{\Lambda}_\infty(y^\beta_\infty)\gamma$ for all y in a neighborhood of y^β_∞, and this latter condition holds because the mapping $(j, y) \mapsto \sqrt{t}e'_j(\dot{\Lambda}^b_t(y) - \dot{\Lambda}^b_\infty(y))$ converges to a sufficiently well-behaved Gaussian process, indexed by y and j.

While the previous propositions establish that our shadow price variances falls linearly with t, the following proposition and corollary show that their tails falls exponentially with t.

Proposition 6. *For all $p \geq 0$, there exist $\delta, C > 0$ such that* $\mathrm{E}\left(\sup_{b \in B_\delta(\beta)} \mathbb{1}\{y^b_t \notin B_\epsilon(y^b_\infty)\}\|y^b_t - y^b_\infty\|^p\right) \leq \exp(-C\epsilon^2 t)$ *for all sufficiently small $\epsilon > 0$ and sufficiently large t.*

Corollary 3. *There exist $\delta, C > 0$ such that* $\Pr\left(\sup_{b \in B_\delta(\beta)} \|y^b_t - y^b_\infty\| > \epsilon\right) \leq \exp(-C\epsilon^2 t)$ *for all sufficiently small $\epsilon > 0$ and sufficiently large t.*

Whereas Li and Ye's (2022) third proposition establishes a concentration of measure for random subgradient $\dot{\Lambda}^b_t(y^b_\infty)$, Corollary 3 establishes a concentration of measure for random shadow price y^b_t. This latter result is far harder to prove because y^b_t is not a sum of *i.i.d.* random variables, unlike $\dot{\Lambda}^b_t(y^b_\infty)$. I establish the shadow price concentration of measure by projecting the shadow prices onto the subgradient of the dual value function at many points. These projections yield inequalities that describe a small box around y^b_t and y^b_∞. This box has random faces, so its walls don't meet at 90-degree angles, but the angles exhibit a concentration of measure, so the probability that the wall's fluctuations undermine the box's integrity falls exponentially fast with t.

4.3 Upper Bound with Known Distribution

I will now prove Theorem 1 by showing that Algorithm 2 honors its $O(\log n)$ bound. I will begin by showing that the inventory levels follow a martingale under this algorithm. This martingale property concentrates the distribution of b_t to the small neighborhood of β for which our lemmas apply. Next, I will express the values received under Algorithm 2 and those received under the optimal algorithm with Bellman-style recurrence relations. I will then combine these recurrence relations to create an analogous regret recurrence relation, which I will unravel to create a corresponding regret recurrence relation. Finally, I will bound this decomposition's myopic regret with our shadow price convergence results.

Algorithm 2 satisfies the period-t customer if and only if (i) there is inventory enough to do so (i.e., $tb_t \geq a_t$) and (ii) the customer has positive surplus utility under the fluid-approximation shadow prices (i.e., $\Delta_t(y_\infty^{b_t}) > 0$). Under this policy, the inventory vector follows a martingale: for $t > 1$, $b_t \geq \alpha/t$, and b_t sufficiently close to β, we have

$$\begin{aligned}
\mathrm{E}(b_{t-1} : b_t) &= \mathrm{E}(\psi_t^{b_t}(x_t a_t) : b_t) \\
&= (tb_t - \mathrm{E}(\mathbb{1}\{\Delta_t(y_\infty^{b_t}) > 0\}a_t : b_t))/(t-1) \\
&= (tb_t - b_t + \dot{\Lambda}_\infty^{b_t}(y_\infty^{b_t}))/(t-1) \\
&= b_t + \dot{\Lambda}_\infty^{b_t}(y_\infty^{b_t})/(t-1) \\
&= b_t.
\end{aligned}$$

This martingale property implies the following, via the Azuma–Hoeffding inequality.

Lemma 3. *The inventory vector abides by a concentration of measure, under Algorithm 2: for all $\delta > 0$, there exists $C > 0$ such that $\Pr(b_t \notin B_\delta(\beta)) \leq \exp(-Ct)$, for all sufficiently large t.*

This result is stronger than one Li and Ye (2022) used. To see this, let $\tau(\delta)$ represent the first time that b_t leaves $B_\delta(\beta)$:

$$\tau(\delta) \equiv \begin{cases} 0 & \{b_t : t \in [n]\} \subset B_\delta(\beta), \\ \max\{t : b_t \notin B_\delta(\beta)\} & \text{otherwise.} \end{cases} \tag{25}$$

Li and Ye proved that their algorithm yields $\mathrm{E}(\tau(\delta)) = O(\log n \log\log n)$—i.e., that it constrains the resource vector for all but the last $O(\log n \log\log n)$ periods. But I couldn't use this $O(\log n \log\log n)$ result to derive a $O(\log n)$ regret bound, so I had to sharpen their finding. To my surprise, I managed to tighten it $O(1)$, as the following corollary explains.

Corollary 4. *The time remaining after the resource vector leaves a given neighborhood of β is asymptotically independent of n, under Algorithm 2: $\mathrm{E}(\tau(\delta)) = O(1)$ as $n \to \infty$, for all $\delta > 0$.*

Algorithm 2.

1. `input` n, β, $\{u_t\}_{t=1}^n$, $\{a_t\}_{t=1}^n$, μ

2. `initialize` $b_n \coloneqq \beta$

3. `for` t `from` n `to` 1 `do`

   (a) `set` $x_t \coloneqq \mathbb{1}\{\Delta_t(y_\infty^{b_t}) > 0\}\mathbb{1}\{tb_t \geq a_t\}$

   (b) `set` $b_{t-1} \coloneqq \psi_t^{b_t}(x_t a_t)$

4. `end for`

5. `output` $\{x_t\}_{t=1}^n$

Algorithm 2 yields total value $\bar{v}_n$, where

$$\begin{aligned} \bar{v}_t &\equiv \mathbb{1}\{\Delta_t(y_\infty^{b_t}) > 0\}\mathbb{1}\{tb_t \geq a_t\}u_t + \bar{v}_{t-1} \\ \text{and} \quad \bar{v}_0 &\equiv 0. \end{aligned} \tag{26}$$

Since the optimal policy is no worse than our martingale policy, we have

$$\begin{gathered} \mathrm{E}(\bar{R}_n) \geq \mathrm{E}(R_n) \\ \text{where} \quad \bar{R}_t \equiv \bar{V}_t^{b_t} - \bar{v}_t. \end{gathered} \tag{27}$$

Accordingly, it will suffice to show that $\mathrm{E}(\bar{R}_n) = O(\log n)$.

I will now present some basic linear programming identities to express regret $\bar{R}_n$ in terms of shadow prices. First, note that we can express the problem in (18) recursively:

$$\bar{V}_t^b \equiv \begin{cases} \max_{x_t \in [0,1]} x_t u_t + \bar{V}_{t-1}^{\psi_t^b(x_t a_t)} & tb \geq a_t, \\ \bar{V}_{t-1}^{\psi_t^b(0)} & tb < a_t. \end{cases} \tag{28}$$

Second, since the shadow price weakly decreases with the inventory level, we have the following for $x \in [0,1]$ and $tb \geq a_t$:

$$(1-x)a_t' y_{t-1}^{\psi_t^b(0)} \leq \bar{V}_{t-1}^{\psi_t^b(x a_t)} - \bar{V}_{t-1}^{\psi_t^b(a_t)} \leq (1-x)a_t' y_{t-1}^{\psi_t^b(a_t)} \tag{29}$$

$$x a_t' y_{t-1}^{\psi_t^b(0)} \leq \bar{V}_{t-1}^{\psi_t^b(0)} - \bar{V}_{t-1}^{\psi_t^b(x a_t)} \leq x a_t' y_{t-1}^{\psi_t^b(a_t)}. \tag{30}$$

Third, $\Delta_t(y_\infty^{b_t}) > 0$ and $tb_t \geq a_t$ imply $b_{t-1} = \psi_t^{b_t}(a_t)$ and hence $\bar{v}_{t-1} = \bar{V}_{t-1}^{\psi_t^{b_t}(a_t)} - \bar{R}_{t-1}$. Accordingly, lines (26)–(29) yield the following, when $\Delta_t(y_\infty^{b_t}) > 0$ and $tb_t \geq a_t$:

$$\begin{aligned}
\bar{R}_t =& \bar{V}_t^{b_t} - \bar{v}_t \\
=& \max_{x\in[0,1]} u_t x + \bar{V}_{t-1}^{\psi_t^{b_t}(xa_t)} - u_t - \bar{v}_{t-1} \\
=& \max_{x\in[0,1]} u_t(x-1) + \bar{V}_{t-1}^{\psi_t^{b_t}(xa_t)} - \bar{V}_{t-1}^{\psi_t^{b_t}(a_t)} + \bar{R}_{t-1} \\
\leq& \max_{x\in[0,1]} u_t(x-1) + (1-x)a_t' y_{t-1}^{\psi_t^{b_t}(a_t)} + \bar{R}_{t-1} \\
=& (a_t' y_{t-1}^{\psi_t^{b_t}(a_t)} - u_t)^+ + \bar{R}_{t-1} \\
=& \Delta_t(y_{t-1}^{\psi_t^{b_t}(a_t)})^- + \bar{R}_{t-1}.
\end{aligned} \tag{31}$$

Analogously, if $\Delta_t(y_\infty^{b_t}) \leq 0$ and $tb_t \geq a_t$ then (26), (28), and (30) yield

$$\bar{R}_t \leq \Delta_t(y_{t-1}^{\psi_t^{b_t}(0)})^+ + \bar{R}_{t-1}, \tag{32}$$

And if $tb_t < a_t$ then (26) and (28) yield

$$\bar{R}_t = \bar{R}_{t-1}. \tag{33}$$

Finally, since $\bar{V}_t^{b_t}$ can't exceed the sum of the remaining utilities, we must also have

$$\bar{R}_t \leq \sum_{s=1}^t u_s + \bar{R}_{t-1}. \tag{34}$$

Now combining inequalities (31)–(34) inductively yields the following, for sufficiently small $\delta > 0$:

$$\bar{R}_n \leq \sum_{t=1}^n r_t, \tag{35}$$

$$\begin{aligned}
\text{where} \quad r_t \equiv& \mathbb{1}\{b_t \notin B_{\delta/2}(\beta)\} \sum_{s=1}^t u_s \\
&+ \mathbb{1}\{b_t \in B_{\delta/2}(\beta)\}\mathbb{1}\{\Delta_t(y_\infty^{b_t}) > 0\}\Delta_t(y_{t-1}^{\psi_t^{b_t}(a_t)})^- \\
&+ \mathbb{1}\{b_t \in B_{\delta/2}(\beta)\}\mathbb{1}\{\Delta_t(y_\infty^{b_t}) \leq 0\}\Delta_t(y_{t-1}^{\psi_t^{b_t}(0)})^+.
\end{aligned}$$

Note, I condition on $b_t \in B_{\delta/2}(\beta)$, because that implies $\psi_t^{b_t}(0), \psi_t^{b_t}(a_t) \in B_\delta(\beta)$, when t is large.

Finally, combining lines (27) and (35) with the following lemma yields Theorem 1.

Lemma 4. *The expected period-t myopic regret is $O(1/t)$ under Algorithm 2: there exists $C > 0$ such that $\mathrm{E}(r_t) \le C/t$, for all $n \in \mathbb{N}$ and $t \le n$.*

To control the first term of the myopic regret, I use the fact that $\mathrm{E}(\sum_{s=1}^{t} u_s)$ increases linearly in t, whereas $\Pr(b_t \notin B_{\delta/2}(\beta))$ falls exponentially, by Lemma 3. To control the second term, I bound $\mathrm{E}(\mathbb{1}\{\Delta_t(y_\infty^{b_t}) > 0\}\Delta_t(y_{t-1}^{\psi_t^{b_t}(a_t)})^-)$ in terms of $\mathrm{E}(\sup_{b\in B_\delta(\beta)} \|y_{t-1}^b - y_\infty^b\|^2)$, $\mathrm{E}(\sup_{b\in B_\delta(\beta)} \mathbb{1}\{y_{t-1}^b \notin B_\epsilon(y_\infty^b)\}\|y_{t-1}^b - y_\infty^b\|)$, and $\Pr(\sup_{b\in B_\delta(\beta)} \|y_t^b - y_\infty^b\| > \epsilon)$, and then apply Propositions 4 and 6 and Corollary 3. And I control the third term similarly since the argument also holds in the mirror image.

4.4 Upper Bound with Unknown Distribution

I will now prove Theorem 2 by showing that Algorithm 3 honors its $O(\log n)$ bound. The only difference between Algorithms 2 and 3 is that the former uses limiting shadow price $y_\infty^{b_t}$, which requires knowledge of μ, whereas the latter uses *look-back shadow price* $\underleftarrow{y}_t^{b_t}$, which is an estimate of $y_\infty^{b_t}$ given the data observed up until period $t+1$. More specifically $\underleftarrow{y}_t^{b_t}$ is a minimizer of the backwards-looking problem

$$\underleftarrow{\Delta}_t^b(y) \equiv b'y + \sum_{s=t+1}^{n} \Delta_s(y)^+/(n-t). \tag{36}$$

Our shadow price convergence results hold for look-back shadow prices but with $(n-t)$-period scaling rather than t-period scaling. For example, Proposition 4 implies that $\mathrm{E}(\mathbb{1}\{b_t \in B_\delta(\beta)\}\|\underleftarrow{y}_t^{b_t} - y_\infty^{b_t}\|^2) = O(1/(n-t))$. (Note that this would *not* be the case if the proposition positioned the $\sup_{b\in B_\delta(\beta)}$ term outside of the expectation, since b_t correlates with the random map $b \mapsto \underleftarrow{y}_t^b$.)

Algorithm 3.

1. `input` n, β, $\{u_t\}_{t=1}^n$, $\{a_t\}_{t=1}^n$
2. `initialize` $b_n \coloneqq \beta$
3. `for` t `from` n `to` 1 `do`
   (a) `set` $x_t \coloneqq \mathbb{1}\{\Delta_t(\underleftarrow{y}_t^{b_t}) > 0\}\mathbb{1}\{tb_t \ge a_t\}$
   (b) `set` $b_{t-1} \coloneqq \psi_t^{b_t}(x_t a_t)$

```
4. end for

5. output {x_t}_{t=1}^n
```

While the inventory vector does not follow a martingale under Algorithm 3, as it does under Algorithm 2, we can still exercise control over its trajectory for all but $O(1)$ periods, as the following results establish.

Lemma 5. *The inventory vector abides by a concentration of measure, under Algorithm 3: for all* $\delta > 0$*, there exists* $C > 0$ *such that* $\Pr(b_t \notin B_\delta(\beta)) \leq \exp(-C \min(t, \sqrt{n}))$*, for all sufficiently large* $t \leq n$*.*

Corollary 5. *The time remaining after the resource vector leaves a given neighborhood of* β *is asymptotically independent of* n*, under Algorithm 3:* $\mathrm{E}(\tau(\delta)) = O(1)$ *as* $n \to \infty$*, for all* $\delta > 0$*.*

The critical insight underlying Lemma 5 is that b_t can't escape $B_{\delta/2}(\beta)$ in less than $\Omega(n)$ time, and hence without first generating an $\Omega(n)$-sized sample of training data. This means that by the time the $\{b_t\}_{t=n}^1$ process has made it halfway out of $B_\delta(\beta)$—i.e., departed $B_{\delta/2}(\beta)$—our look-back shadow prices are accurate enough to (almost) guarantee that it can't traverse the second half. This property enables us to restrict attention to the periods with accurate look-back shadow prices (i.e., periods after time $\tau(\delta/2)$).

But controlling the evolution of $\{b_t\}_{t=n}^1$ is difficult even when look-back shadow prices are accurate. The problem is that while b_t is independent of the mapping $b \mapsto y_t^b$, it is not independent of the mapping $b \mapsto \underleftarrow{y}_t^b$. Indeed, the inventory vectors and look-back shadow prices intertwine in a complex dance. To extricate b_t from this pas de deux, I decompose it into three parts: $b_{\tau(\delta/2)+1}$, $\sum_{s=t}^{\tau(\delta/2)} b_s - \mathrm{E}(b_s : b_{s+1})$, and $\sum_{s=t}^{\tau(\delta/2)} \mathrm{E}(b_s : b_{s+1}) - b_{s+1}$. By definition, the first part is within $\delta/2$ of β. The second part follows a martingale and thus concentrates about zero. And the third part is small, provided that $\underleftarrow{y}_s^{b_s}$ is near $y_\infty^{b_s}$, for all $s \in \{t+1, \cdots, \tau(\delta/2)+1\}$. Crucially, $\tau(\delta/2)$ will be small enough to ensure that this holds with high probability, provided that b_s is near β for all $s \in \{t+1, \cdots, \tau(\delta/2)+1\}$. And thus, I can inductively establish the result: b_s being near β for $s \in \{t+1, \cdots, \tau(\delta/2)+1\}$ implies that $\underleftarrow{y}_s^{b_s}$ is near $y_\infty^{b_s}$ for $s \in \{t+1, \cdots, \tau(\delta/2)+1\}$, which implies that $\sum_{s=t}^{\tau(\delta/2)} \mathrm{E}(b_s : b_{s+1}) - b_{s+1}$ is small, which implies that b_t is near β.

Having reigned in our inventory vectors, we are ready to decompose our regret. The analysis of the previous section yields the following:

$$
\begin{aligned}
\bar{R}_n &\leq \sum_{t=1}^{n} r_t, && (37)\\
\text{where}\quad r_t &\equiv \mathbb{1}\{b_t \notin B_{\delta/2}(\beta)\} \sum_{s=1}^{t} u_s \\
&\quad + \mathbb{1}\{b_t \in B_{\delta/2}(\beta)\}\mathbb{1}\{\Delta_t(\underleftarrow{y}_t^{b_t}) > 0\}\Delta_t(y_{t-1}^{\psi_t^{b_t}(a_t)})^- \\
&\quad + \mathbb{1}\{b_t \in B_{\delta/2}(\beta)\}\mathbb{1}\{\Delta_t(\underleftarrow{y}_t^{b_t}) \leq 0\}\Delta_t(y_{t-1}^{\psi_t^{b_t}(0)})^+.
\end{aligned}
$$

The only differences between lines (35) and (37) is that $\bar{R}_n$ and b_t now correspond to Algorithm 3, rather than that Algorithm 2, and $\underleftarrow{y}_t^{b_t}$ has replaced $y_\infty^{b_t}$.

Finally, combining lines (27) and (37) with the following lemma yields Theorem 2.

Lemma 6. *The expected period-t myopic regret is $O(1/t) + O(1/(n-t))$ under Algorithm 3: there exists $C > 0$ such that $\mathrm{E}(r_t) \leq C/t + C/(n-t)$, for all $n \in \mathbb{N}$ and $t \leq n$.*

This lemma is the same as Lemma 4, except now both the $O(1/\sqrt{t})$ errors between $y_t^{b_t}$ and $y_\infty^{b_t}$ and the $O(1/\sqrt{n-t})$ errors between $\underleftarrow{y}_t^{b_t}$ and $y_\infty^{b_t}$ contribute to your regret.

4.5 Lower Bound with Known Distribution

I will now prove Theorem 3 by creating a lower-bounding version of the methodology developed in Section 4.3. For example, the lower-bounding decomposition will depend on $\Delta_t(y_{t-1}^{\psi_t^{b_t}(0)})^-$ and $\Delta_t(y_{t-1}^{\psi_t^{b_t}(a_t)})^+$ (as opposed to $\Delta_t(y_{t-1}^{\psi_t^{b_t}(a_t)})^-$ and $\Delta_t(y_{t-1}^{\psi_t^{b_t}(0)})^+$); the lower-bounding version of Lemma 3 will ensure the proximity of b_t and β under the optimal algorithm (as opposed to Algorithm 2); and the lower-bounding version of Lemma 4 will establish that the expected myopic regret is $\Omega(1/t)$ (as opposed to $O(1/t)$).

In this section $\{b_t\}_{t=n}^{1}$ will characterize the inventory levels that correspond to the optimal actions specified in line (14): $b_n = \beta$ and $b_{t-1} = \psi_t^{b_t}(\pi_t^{b_t} a_t)$. Unfortunately, we now have little control over $\{b_t\}_{t=n}^{1}$, because the optimal policy is unknown. Nevertheless, we can still situate b_t near β for a substantial time interval.

Lemma 7. *The inventory vector tends to lie near β under the optimal policy for most of the second half of the horizon: For all $\delta > 0$, if n is sufficiently large then $n^{3/4} \leq t \leq n/2$ implies $\Pr(b_t \notin B_{\delta/2}(\beta)) \leq n^{-1/2}$.*

This lemma was the hardest result in this article to prove because the optimal policy is opaque. Generalizing the technique developed in Section 3.3, I argue that the regret incurred when b_t strays from β is at least as large as the value sacrificed when we chop the linear program into two separate problems, one with horizon t and endowment tb_t and the other with horizon $n-t$ and endowment $n\beta - tb_t$. The concavity of $\bar{V}_t^b$ in b ensures that this division is costly when b_t meaningfuly differs from β.

As before, we will benchmark against the offline linear program, line (18), rather than the offline integer program, line (16). The following result will enable us to do so:

$$\mathrm{E}(\underline{R}_n) = \mathrm{E}(R_n) + O(1), \tag{38}$$
$$\text{where} \quad \underline{R}_t \equiv \bar{V}_t^{b_t} - v_t^{b_t}.$$

Line (38) holds because the linear program has a solution that partially satisfies at most m customers, and thus the integer program must derive at least as much value from resource endowment β as the linear program does from resource endowment $\beta - m\alpha/n$: $V_n^\beta \geq \bar{V}_n^{\beta - m\alpha/n}$. And since the shadow price decreases in the inventory level, this implies that $V_n^\beta \geq \bar{V}_n^\beta - m\alpha' y_n^{\beta - m\alpha/n}$, and hence that $R_n \geq \bar{V}_n^\beta - v_n^\beta - m\alpha' y_n^{\beta - m\alpha/n}$. Finally, Proposition 4 indicates that $\mathrm{E}(y_n^{\beta - m\alpha/n}) = O(1)$ as $n \to \infty$, which establishes the result.

I will now create the lower-bounding version of our regret decomposition. First, note that $\pi_t^{b_t} = 1$ implies $b_{t-1} = \psi_t^{b_t}(a_t)$, and hence $v_{t-1}^{\psi_t^{b_t}(a_t)} = \bar{V}_{t-1}^{\psi_t^{b_t}(a_t)} - \underline{R}_{t-1}$. Also, $\pi_t^{b_t} = 1$ implies $tb_t \geq a_t$, which with lines (28) and (29) yield

$$\begin{aligned}
\underline{R}_t &\equiv \bar{V}_t^{b_t} - v_t^{b_t} \\
&= \max_{x_t \in [0,1]} x_t u_t + \bar{V}_{t-1}^{\psi_t^{b_t}(x_t a_t)} - u_t - v_{t-1}^{\psi_t^{b_t}(a_t)} \\
&= \max_{x_t \in [0,1]} (x_t - 1) u_t + \bar{V}_{t-1}^{\psi_t^{b_t}(x_t a_t)} - \bar{V}_{t-1}^{\psi_t^{b_t}(a_t)} + \underline{R}_{t-1} \\
&\geq \max_{x_t \in [0,1]} (x_t - 1) u_t + (1-x) a_t' y_{t-1}^{\psi_t^{b_t}(0)} + \underline{R}_{t-1} \\
&= (a_t' y_{t-1}^{\psi_t^{b_t}(0)} - u_t)^+ + \underline{R}_{t-1} \\
&= \Delta_t (y_{t-1}^{\psi_t^{b_t}(0)})^- + \underline{R}_{t-1}.
\end{aligned}$$

Analogously, if $\pi_t^{b_t} = 0$ and $tb_t \geq a_t$ then lines (28) and (30) yield

$$\underline{R}_t \geq \Delta_t(y_{t-1}^{\psi_t^{b_t}(a_t)})^+ + \underline{R}_{t-1}.$$

Further, we always trivially have

$$\underline{R}_t \geq \underline{R}_{t-1}.$$

Now choose $\delta > 0$ small enough to ensure that $\delta\iota \leq \beta$, where ι is a vector of ones. In this case, $b_t \in B_{\delta/2}(\beta)$ implies $tb_t \geq a_t$ for $t \geq 2\|\alpha\|/\delta$, which with our previous three inequalities inductively implies

$$\begin{aligned} \underline{R}_n \geq & \sum_{t=\lceil 2\|\alpha\|/\delta \rceil}^{n} r_t, \\ \text{where} \quad r_t \equiv & \mathbb{1}\{b_t \in B_{\delta/2}(\beta)\}\big(\pi_t^{b_t}\Delta_t(y_{t-1}^{\psi_t^{b_t}(0)})^- + (1-\pi_t^{b_t})\Delta_t(y_{t-1}^{\psi_t^{b_t}(a_t)})^+\big). \end{aligned} \tag{39}$$

Finally, combining lines (38) and (39) with the following lemma yields Theorem 3.

Lemma 8. *The expected period-t myopic regret is $\Omega(1/t)$ under the optimal policy for most of the second half of the horizon: There exists $C > 0$ such that $\mathrm{E}(r_t) \geq C/t$, for all sufficiently large t that satisfies $n^{3/4} \leq t \leq n/2$.*

To establish this result, I show that if $b_t \in B_{\delta/2}(\beta)$—which happens with high probability, by Lemma 7—then $\sqrt{t}(y_t^{b_t} - y_\infty^{b_t})$ could be near *any* $\gamma \in \mathbb{R}^m$. Accordingly, both Type I errors—rejecting customers that should have been satisfied—and Type II errors—satisfying customers that should have been rejected—are unavoidable because the shadow price can always be larger or smaller than anticipated. Specifically, I show that there's at least a $\Omega(1/\sqrt{t})$ chance that both the expected Type I and Type II errors are $\Omega(1/\sqrt{t})$.

5 Conclusion

I began this project by investigating what happens when we extend secretary valuations from the finite set $\{v_1, \cdots, v_j\}$ to infinite set $[0, 1]$ in Arlotto and Gurvich's (2019) multisecretary model. I found that this extension invalidates Arlotto and Gurvich's (2019) finite regret bound. Why is this? Well, in the former case, secretaries are interchangeable parts, which means that your hiring errors can be "undone" in the future. For example, you can confidently reject the marginal candidate in Arlotto and Gurvich's setting because you'll almost certainly see his like again (many times over). But no two men are alike when values are drawn from the uniform distribution. So, in this case, your hiring mistakes cannot be undone—they will linger forever, making you ever more regretful.

Alright, fine, so your regret increases. But why does it do so at a $\log(n)$ rate? Well, suppose there are n applicants for $n\beta$ open positions, and the first man you interview has utility u_n. Do you classify him as "hired" or "not hired"? Like all classification problems, you must balance between the threat of a Type I error and a Type II error. You make a Type I error when you extend a job offer and the "shadow price" of an open position exceeds u_n—in this case, you took the slot of a more deserving applicant. And you make a Type II error when you don't extend an offer and u_n exceeds the "shadow price" of an open position—in this case, you reserved the slot for an inferior applicant. The "shadow price" of a position equals the utility of the $(n\beta)$th best man out of the remaining $n-1$ candidates, which is order statistic $h_{n-1}^{n\beta}$. Hence, you must try to anticipate whether or not u_n exceeds $h_{n-1}^{n\beta}$. Fortunately, the standard deviation of $h_{n-1}^{n\beta}$ drops like $1/\sqrt{n}$, so we can think of $h_{n-1}^{n\beta}$ as residing in a "confidence interval" whose length is on the order of $1/\sqrt{n}$. This has two implications. First, you'll only have around a $1/\sqrt{n}$ chance of making a hiring mistake because the optimal action is unambiguous when u_n falls outside the confidence interval. Second, the cost of a hiring mistake is also around $1/\sqrt{n}$ because the capability of the man you should have hired and the capability of the man you did hire both reside in the confidence interval. Thus, the expected regret of your period-n decision will be on the order of $\underbrace{1/\sqrt{n}}_{\text{mistake probability}} \cdot \underbrace{1/\sqrt{n}}_{\text{mistake cost}} = 1/n$, which means that your total regret grows like the harmonic series.

The same principle applies to the more general online linear program, but extending the tight $O(\log n)$ and $\Omega(\log n)$ bounds to this problem necessitated more refined control over shadow prices. Hence, I precisely characterize the convergence of dual variable y_t^b to its deterministic limit, y_∞^b. Specifically, I develop weak conditions in which

- $\sqrt{t}(y_t^b - y_\infty^b)$ converges to a multivariate normal for all $b \in B_\delta(\beta)$,
- $\Pr(\sup_{b\in B_\delta(\beta)} \|y_t^b - y_\infty^b\| > \epsilon)$ falls exponentially fast in t,
- $\mathrm{E}(\sup_{b\in B_\delta(\beta)} \|y_t^b - y_\infty^b\|^2) = O(1/t)$, and
- $\mathrm{E}(\inf_{b\in B_\delta(\beta)} \|y_t^b - y_\infty^b\|^2) = \Omega(1/t)$.

Further, since the $\sup_{b\in B_\delta(\beta)}$ and $\inf_{b\in B_\delta(\beta)}$ terms lie inside of the expectation, we can apply these bounds to random values of b_t, even if they correlate with shadow prices. This was the key to accommodating the most interesting aspect of the problem: online learning.

Acknowledgments

Many thanks to Siddhartha Banerjee, Itai Gurvich, Ioannis Stamatopoulos, He Wang, and three anonymous reviewers for their very insightful feedback.

Notes

[1] I got the idea for compensated coupling from Arcidiacono and Miller (2011), which used a similar technique to structurally estimate dynamic programs.

[2] To see that the expected regret with $n\beta$ initial open slots equals that with $n(1-\beta)$ initial open slots, note that we can re-express the problem of maximizing the capability of each of the $n\beta$ men you hire to maximizing one minus the capability of the $n(1-\beta)$ men you reject. But one minus a uniform is also a uniform, so this mirror image problem must yield mirror image regrets.

[3] I make three minor changes to the online linear programming model: I impose additional non-negativity constraints, $u_1, a_1 \geq 0$, I do not include constraints that are slack in the limit, and I use a cleaner version of the continuous value assumption, which I inherited from Lueker (1998). The first two modifications are trivial: Accommodating negative u_1 and a_1 would be simple because all that matters is the difference, $\Delta_1(y) = u_1 - a_1' y$. And a simple concentration of measure argument establishes that a constraint that does not bind in the limit has only a $O(1)$ effect on the expected regret because the probability of it binding decreases exponentially fast in n. (I incorporated constraints that are slack in the limit in a previous version of the manuscript.) The third change is more noteworthy, because Assumption 6 is more straightforward and flexible. For example, this assumption permits unbounded shadow prices and hence unbounded utilities, and it extends the model to cover Arlotto and Xie's (2020) specification.

List of Symbols

$[x]$	the set $\{1, \cdots, x\}$	
$x \wedge y$	vector with ith element $\min(x_i, y_i)$	
$x \vee y$	vector with ith element is $\max(x_i, y_i)$	
x^+	$\max(0, x)$	
x^-	$\max(0, -x)$	
e_j	unit vector indicating jth position	
ι	vector of ones	
$\mathbb{1}\{\}$	indicator function	
$B_\delta(b)$	open ball with radius δ about b	
m	number of resources to manage	
n	number of time periods	
t	generic time period	
b_t	inventory vector, defined as period-t inventory holdings divided by t	
b	generic inventory vector	
β	initial inventory vector	
$\tau(\delta)$	first time inventory vector leaves $B_\delta(\beta)$	line (25)
u_t	utility received by satisfying period-t customer	
a_t	resources consumed by satisfying period-t customer	
Δ_t	surplus utility function	line (20)
μ	joint distribution of (u_t, a_t)	Assumption 1
α	upper bound on a_t	Assumption 4
x_t	period-t decision variable	
ψ_t^b	function determining period-$(t-1)$ inventory vector	line (13)
π_t^b	optimal action	line (14)
v_t^b	online objective value	line (15)
$\bar{v}_t$	martingale-policy objective value	line (26)
V_t^b	offline objective value	line (16)
$\bar{V}_t^b$	offline objective value with linear programming relaxation	line (18)
R_n	regret	line (17)
$\bar{R}_t$	upper-bound regret relaxation	line (27)
$\underline{R}_t$	lower-bound regret relaxation	line (38)
r_t	myopic regret	lines (35), (37), and (39)
Λ_t^b	dual objective	line (19)
$\underleftarrow{\Lambda}_t^b$	look-back dual objective	line (36)
$\dot{\Lambda}_t$	dual objective subgradient	line (24)
Λ_∞^b	limiting dual objective	line (22)
$\dot{\Lambda}_\infty$	limiting dual gradient	line (23)
$\ddot{\Lambda}_\infty$	limiting dual Hessian	Lemma 1
ω_i^b	ith orthonormal eigenvector of limiting dual Hessian	below Lemma 2
σ_i^b	ith largest eigenvalue of limiting dual Hessian	below Lemma 2
y_t^b	dual optimal solution	line (21)
$\underleftarrow{y}_t^b$	look-back dual optimal solution	before line (36)
y_∞^b	limiting dual optimal solution	Assumption 5 and Lemma 2
y	generic dual solution	

References

Arcidiacono, Peter, Robert Miller. 2011. Conditional Choice Probability Estimation of Dynamic Discrete Choice Models with Unobserved Heterogeneity. *Econometrica* **79**(6) 1823–1867.

Arlotto, Alessandro, Itai Gurvich. 2019. Uniformly Bounded Regret in the Multisecretary Problem. *Stochastic Systems* **9**(3) 231–260.

Arlotto, Alessandro, Xinchang Xie. 2020. Logarithmic Regret in the Dynamic and Stochastic Knapsack Problem with Equal Rewards. *Stochastic Systems* **10**(2) 170–191.

Balseiro, Santiago R., Omar Besbes, Dana Pizarro. 2023. Survey of Dynamic Resource-Constrained Reward Collection Problems: Unified Model and Analysis. *Operations Research* (Articles in Advance) 1–22.

Besbes, Omar, Yash Kanoria, Akshit Kumar. 2022. The Multi-secretary Problem with Many Types 1146–1147.

Caley, A. 1875. Mathematical questions with their solutions. *The Educational Times* **23** 18–19.

Eaton, Morris L. 1983. Multivariate statistics: a vector space approach. 512.

Jasin, Stefanus. 2014. Reoptimization and self-adjusting price control for network revenue management. *Operations Research* **62**(5) 1168–1178.

Jiang, Jiashuo, Will Ma, Jiawei Zhang. 2022. Degeneracy is OK : Logarithmic Regret for Network Revenue Management with Indiscrete Distributions. *Working Paper* 1–55.

Kosorok, Michael. 2008. *Introduction to Empirical Processes and Semiparametric Inference*. Springer.

Li, Xiaocheng, Yinyu Ye. 2022. Online Linear Programming: Dual Convergence, New Algorithms, and Regret Bounds. *Operations Research* (October 2022).

Lueker, George S. 1998. Average-case analysis of off-line and on-line knapsack problems. *Journal of Algorithms* **29** 277–305.

van der Vaart, Aad W., Jon A. Wellner. 1996. *Weak Convergence and Empirical Processes*. Springer, New York.

Vera, Alberto, Siddhartha Banerjee. 2019. The Bayesian Prophet: A Low-Regret Framework for Online Decision Making. *Working Paper* 1–25.

Vera, Alberto, Siddhartha Banerjee, Itai Gurvich. 2019. Online Allocation and Pricing: Constant Regret via Bellman Inequalities. *Operations Research* **69**(3) 821–840.

Wang, Yining, He Wang. 2022. Constant Regret Resolving Heuristics for Price-Based Revenue Management. *Operations Research* **5463**(3) 1–20.

ONLINE APPENDIX OF PROOFS

Lemma 1 Proof. Assumption 6 implies that the event $\Delta_1(y) = 0$ and $a_1 \neq 0$ has measure zero, for y sufficiently close to y_∞^β. Hence, $\Delta_1(y)^+$ is almost surely differentiable in y, which means that

$$\begin{aligned}
\tfrac{\partial}{\partial y}\Lambda_\infty^b(y) &= \tfrac{\partial}{\partial y}\big(b'y + \mathrm{E}(\Delta_1(y)^+)\big) \\
&= b + \mathrm{E}\left(\tfrac{\partial}{\partial y}\Delta_1(y)^+\right) \\
&= b - \mathrm{E}(\mathbb{1}\{\Delta_1(y) > 0\}a_1).
\end{aligned}$$

Note we can commute the expectation and differentiation because a_1 is bounded. Combining the derivative above with Assumption 6 and the convexity of Λ_∞^b implies the result. □

Lemma 2 Proof. Assumption 5 and Lemma 1 imply that (i) $\dot\Lambda_\infty^\beta(y_\infty^\beta) = 0$, (ii) $\ddot\Lambda_\infty(y_\infty^\beta)$ is non-singular, and (iii) $\dot\Lambda_\infty^b(y)$ is continuously differentiable in y near y_∞^β. Further, $\dot\Lambda_\infty^b(y)$ is continuously differentiable in b, since $\tfrac{\partial}{\partial b}\dot\Lambda_\infty^b(y) = I$ (see the proof of Lemma 1). Accordingly, the implicit function theorem establishes that each b in a neighborhood of β has a corresponding shadow price vector y_∞^b that has continuous derivative $\tfrac{\partial}{\partial b}y_\infty^b = -\tfrac{\partial}{\partial y}\dot\Lambda_\infty^b(y)^{-1}\tfrac{\partial}{\partial b}\dot\Lambda_\infty^b(y)|_{y=y_\infty^b} = -\ddot\Lambda_\infty^b(y_\infty^b)^{-1}$. Further, y_∞^b must be the unique minimizer of Λ_∞^b for b near β, because $\dot\Lambda_\infty^b(y_\infty^b) = 0$ and $\ddot\Lambda_\infty(y)$ is positive definite for y near y_∞^β. □

Corollary 1 Proof. This follow immediately from Theorem 3. □

Proposition 3 Proof. I will first establish that Σ^b is continuous and full rank for all b in a neighborhood of β. Lemmas 1 and 2 imply the continuity, and Lemma 1 implies that Σ^b is full rank if $\mathrm{Cov}(\mathbb{1}\{\Delta_1(y_\infty^b) > 0\}a_1)$ is full rank. If this latter matrix were not full rank, then there would be some $\gamma \neq 0$ that almost surely satisfies $\mathbb{1}\{\Delta_1(y_\infty^b) > 0\}a_1'\gamma = \mathrm{E}(\mathbb{1}\{\Delta_1(y_\infty^b) > 0\}a_1'\gamma)$, which would imply that either (i) $\Delta_1(y_\infty^b) > 0$, almost surely, or (ii) $\mathbb{1}\{\Delta_1(y_\infty^b) > 0\}a_1'\gamma = 0$, almost surely. The former case violates Assumption 6 because it implies that $\mathrm{E}(\mathbb{1}\{\Delta_1(y+dy) > 0\}a_1'\gamma) = \mathrm{E}(\mathbb{1}\{\Delta_1(y) > 0\}a_1'\gamma)$ for $dy \leq 0$, and the latter case violates Assumption 6 because it implies that $\mathrm{E}(\mathbb{1}\{\Delta_1(y+dy) > 0\}a_1'\gamma) = \mathrm{E}(\mathbb{1}\{\Delta_1(y) > 0\}a_1'\gamma)$ for $dy > 0$.

The fact that $\sqrt{t}(y_t^b - y_\infty^b) \xrightarrow{d} \mathcal{N}(0, \Sigma^b)$ follows directly from theorem 2.13 of Kosorok (2008), so it will suffice to show that the conditions of this theorem hold. To use follow Kosorok's notation, define functions

$$\begin{aligned} m_y(u_1, a_1) \equiv& b'y + \Delta_1(y)^+, \\ \dot{m}(a_1) \equiv& \|b\| + \|a_1\|, \\ \text{and} \quad \dot{m}_\infty(u_1, a_1) \equiv& b - \mathbb{1}\{\Delta_1(y^b_\infty) > 0\} a_1. \end{aligned}$$

First, the Hessian matrix of $\mathrm{E}(m_y(u_1, a_1))$ at $y = y^b_\infty$ is $\ddot{\Lambda}_\infty(y^b_\infty)$, which is non-singular when b is sufficiently close to β, by Lemmas 1 and 2. Second, Assumption 4 establishes that $\mathrm{E}(\dot{m}(a_1)^2)$ and $\mathrm{E}(\|\dot{m}_\infty(u_1, a_1)\|^2)$ are finite. Third, functions m_y and $\dot{m}$ satisfy condition (2.18) of Kosorok (2008):

$$\begin{aligned} |m_y(u_1, a_1) - m_z(u_1, a_1)| =& b'y + \Delta_1(y)^+ - b'z - \Delta_1(z)^+ \\ \leq& (\|b\| + \|a_1\|)\|y - z\| \\ =& \dot{m}(a_1)\|y - z\|. \end{aligned}$$

Fourth, Assumption 6 ensures that functions m_y and $\dot{m}_\infty$ satisfy condition (2.19) of Kosorok (2008):

$$\begin{aligned} &\mathrm{E}\left(\left(m_y(u_1, a_1) - m_{y^b_\infty}(u_1, a_1) - \dot{m}_\infty(u_1, a_1)'(y - y^b_\infty)\right)^2\right) \\ &= \mathrm{E}\left(\left(\Delta_1(y)^+ - \Delta_1(y^b_\infty)^+ + \mathbb{1}\{\Delta_1(y^b_\infty) > 0\} a_1'(y - y^b_\infty)\right)^2\right) \\ &= \mathrm{E}\left(\Delta_1(y)^2 |\mathbb{1}\{\Delta_1(y) > 0\} - \mathbb{1}\{\Delta_1(y^b_\infty) > 0\}|\right) \\ &\leq \mathrm{E}\left((a_1'y - a_1'y^b_\infty)^2 |\mathbb{1}\{\Delta_1(y) > 0\} - \mathbb{1}\{\Delta_1(y^b_\infty) > 0\}|\right) \\ &= \|y - y^b_\infty\|^2 \,\mathrm{E}\left(\|\mathbb{1}\{\Delta_1(y) > 0\} a_1 - \mathbb{1}\{\Delta_1(y^b_\infty) > 0\} a_1\|^2\right) \\ &\leq \|\alpha\| \|y - y^b_\infty\|^2 \,\mathrm{E}\left(\mathbb{1}\{\Delta_1(y \wedge y^b_\infty) > 0\} a_1 - \mathbb{1}\{\Delta_1(y \vee y^b_\infty) > 0\} a_1\right) \\ &\leq \|\alpha\| \|y - y^b_\infty\|^2 O(\|y - y^b_\infty\|) \\ &= o(\|y - y^b_\infty\|). \end{aligned}$$

Finally, Proposition 4 establishes that $\|y^b_t - y^b_\infty\| \xrightarrow{p} 0$. □

Proposition 4 Proof. Since Proposition 6 establishes that $\mathrm{E}(\sup_{b \in B_\delta(\beta)} \mathbb{1}\{y^b_t \notin B_\epsilon(y^b_\infty)\} \|y^b_t - y^b_\infty\|^2) = o(1/t)$, it will suffice to show that $\mathrm{E}(\sup_{b \in B_\delta(\beta)} \mathbb{1}\{y^b_t \in B_\epsilon(y^b_\infty)\} \|y^b_t - y^b_\infty\|^2) = O(1/t)$, for sufficiently small $\epsilon > 0$. I will establish this result with Theorems 2.14.2 and 2.14.5 of van der Vaart and Wellner (1996). However, translating the problem into van der Vaart and Wellner's empirical processes framework will take some effort. First, I bound the magnitude of $y^b_t - y^b_\infty$ in terms of the magnitude of $\dot{\Lambda}^b_\infty(\hat{y}^b_t) - \dot{\Lambda}^b_t(\hat{y}^b_t)$, where $\hat{y}^b_t \equiv (y^b_t + y^b_\infty)/2$. Since $\hat{y}^b_t$ lies between the minimizers of Λ^b_∞ and Λ^b_t, the vector $\hat{y}^b_t - y^b_\infty$ projects positively onto gradient $\dot{\Lambda}^b_\infty(\hat{y}^b_t)$ and projects negatively onto subgradient $\dot{\Lambda}^b_t(\hat{y}^b_t)$. I use this fact to show that $(\hat{y}^b_t - y^b_\infty)'(\dot{\Lambda}^b_\infty(\hat{y}^b_t) - \dot{\Lambda}^b_t(\hat{y}^b_t))$ is larger than some fixed multiple of $\|\hat{y}^b_t - y^b_\infty\|^2$, which indicates that $\|\dot{\Lambda}^b_\infty(\hat{y}^b_t) - \dot{\Lambda}^b_t(\hat{y}^b_t)\|$ is larger than some fixed multiple of $\|y^b_t - y^b_\infty\|$. This, in turn, implies that the expectation of the maximum of $\|\dot{\Lambda}^b_t(y) - \dot{\Lambda}^b_\infty(y)\|^2$, across y in some small ball of y^β_∞, is larger than some fixed multiple of the expectation of $\mathbb{1}\{\|y^b_t - y^b_\infty\| \leq \epsilon\}\|y^b_t - y^b_\infty\|^2$. And bounding the expectation of the maximum of $\|\dot{\Lambda}^b_t(y) - \dot{\Lambda}^b_\infty(y)\|^2$ is a classic empirical processes problem.

Now let's get to the proof. First, Lemma 2 establishes that we can choose δ small enough so that $y^b_\infty \in B_\epsilon(y^\beta_\infty)$ for all $b \in B_\delta(\beta)$, in which case $y^b_t \in B_\epsilon(y^b_\infty)$ implies $y^b_t \in B_{2\epsilon}(y^\beta_\infty)$, which in turn implies $\hat{y}^b_t \in B_{3\epsilon/2}(y^b_\infty)$, where $\hat{y}^b_t \equiv (y^b_t + y^b_\infty)/2$.

Second, let σ^b_m denote the smallest singular value of $\ddot{\Lambda}_\infty(y^b_\infty)$. Lemmas 1 and 2 imply that we can set δ small enough so that for all $b \in B_\delta(\beta)$ we have $\sigma^b_m \geq \sigma^\beta_m/2$, and hence

$$(\hat{y}^b_t - y^b_\infty)'\ddot{\Lambda}_\infty(y^b_\infty)(\hat{y}^b_t - y^b_\infty) \geq \sigma^\beta_m\|\hat{y}^b_t - y^b_\infty\|^2/2.$$

Next, note that $\dot{\Lambda}^b_\infty(y^b_\infty) = 0$ implies

$$\begin{aligned}\dot{\Lambda}^b_\infty(\hat{y}^b_t) &= \dot{\Lambda}^b_\infty(\hat{y}^b_t) - \dot{\Lambda}^b_\infty(y^b_\infty) \\ &= \ddot{\Lambda}_\infty(y^b_\infty)(\hat{y}^b_t - y^b_\infty) + o(\|\hat{y}^b_t - y^b_\infty\|),\end{aligned}$$

where the little-o term holds uniformly across $b \in B_\delta(\beta)$. Accordingly, we can set ϵ small enough so that $y^b_t \in B_{2\epsilon}(y^\beta_\infty)$ implies

$$\|\dot{\Lambda}^b_\infty(\hat{y}^b_t) - \ddot{\Lambda}_\infty(y^b_\infty)(\hat{y}^b_t - y^b_\infty)\| \leq \sigma^\beta_m\|\hat{y}^b_t - y^b_\infty\|/4,$$

for all $b \in B_\delta(\beta)$. Now combining these last two results yields the following, for $y^b_t \in B_{2\epsilon}(y^\beta_\infty)$:

$$\begin{aligned}&(\hat{y}^b_t - y^b_\infty)'\dot{\Lambda}^b_\infty(\hat{y}^b_t) \\ &\quad= (\hat{y}^b_t - y^b_\infty)'\ddot{\Lambda}_\infty(y^b_\infty)(\hat{y}^b_t - y^b_\infty) \\ &\qquad+ (\hat{y}^b_t - y^b_\infty)'\big(\dot{\Lambda}^b_\infty(\hat{y}^b_t) - \ddot{\Lambda}_\infty(y^b_\infty)(\hat{y}^b_t - y^b_\infty)\big) \\ &\quad\geq \sigma^\beta_m\|\hat{y}^b_t - y^b_\infty\|^2/2 - \|\hat{y}^b_t - y^b_\infty\|\|\dot{\Lambda}^b_\infty(\hat{y}^b_t) - \ddot{\Lambda}_\infty(y^b_\infty)(\hat{y}^b_t - y^b_\infty)\| \\ &\quad\geq \sigma^\beta_m\|\hat{y}^b_t - y^b_\infty\|^2/4.\end{aligned}$$

And combining this with $(\hat{y}_t^b - y_\infty^b)'\dot{\Lambda}_t^b(\hat{y}_t^b) = (y_t^b - \hat{y}_t^b)'\dot{\Lambda}_t^b(\hat{y}_t^b) \leq 0$, which we get from Lemma 9, yields the following, for $y_t^b \in B_{2\epsilon}(y_\infty^\beta)$:

$$\begin{aligned}
&\|\hat{y}_t^b - y_\infty^b\|\|\dot{\Lambda}_\infty^b(\hat{y}_t^b) - \dot{\Lambda}_t^b(\hat{y}_t^b)\| \\
&\qquad \geq (\hat{y}_t^b - y_\infty^b)'(\dot{\Lambda}_\infty^b(\hat{y}_t^b) - \dot{\Lambda}_t^b(\hat{y}_t^b)) \\
&\qquad \geq \sigma_m^\beta \|\hat{y}_t^b - y_\infty^b\|^2/4.
\end{aligned}$$

Hence, $y_t^b \in B_{2\epsilon}(y_\infty^\beta)$ implies

$$\|\dot{\Lambda}_\infty^b(\hat{y}_t^b) - \dot{\Lambda}_t^b(\hat{y}_t^b)\| \geq \sigma_m^\beta \|\hat{y}_t^b - y_\infty^b\|/4 = \sigma_m^\beta \|y_t^b - y_\infty^b\|/8.$$

And thus, we have

$$\begin{aligned}
&\mathrm{E}\Big(\sup_{b \in B_\delta(\beta)} \mathbb{1}\{y_t^b \in B_\epsilon(y_\infty^b)\}\|y_t^b - y_\infty^b\|^2\Big) \\
&\quad \leq \mathrm{E}\Big(\sup_{b \in B_\delta(\beta)} \mathbb{1}\{y_t^b \in B_{2\epsilon}(y_\infty^\beta)\}\|y_t^b - y_\infty^b\|^2\Big) \\
&\quad \leq (8/\sigma_m^\beta)^2\, \mathrm{E}\Big(\sup_{b \in B_\delta(\beta)} \mathbb{1}\{y_t^b \in B_{2\epsilon}(y_\infty^\beta)\}\|\dot{\Lambda}_t^b(\hat{y}_t^b) - \dot{\Lambda}_\infty^b(\hat{y}_t^b)\|^2\Big) \\
&\quad \leq (8/\sigma_m^\beta)^2\, \mathrm{E}\Big(\sup_{b \in B_\delta(\beta)} \sup_{y \in B_{2\epsilon}(y_\infty^\beta)} \|\dot{\Lambda}_t^b(y) - \dot{\Lambda}_\infty^b(y)\|^2\Big) \\
&\quad = (8/\sigma_m^\beta)^2\, \mathrm{E}\Big(\sup_{y \in B_{2\epsilon}(y_\infty^\beta)} \|\dot{\Lambda}_t^\beta(y) - \dot{\Lambda}_\infty^\beta(y)\|^2\Big),
\end{aligned}$$

where the last line holds because $\dot{\Lambda}_t^b - \dot{\Lambda}_\infty^b$ is independent of b. Finally, Lemma 13 establishes that the expectation in the last line is less than C/t, for some universal constant $C > 0$. □

Proposition 5 Proof. This follows immediately from Lemma 10. □

Corollary 2 Proof. This follows from Proposition 4 and Lemma 10. □

Proposition 6 Proof. The proof hinges on two key results. The first result is that there exists $\delta, C > 0$ such that

$$\Pr\big(\sup_{b \in B_\delta(\beta)} \|y_t^b - y_\infty^b\| > \epsilon\big) \leq 4m^2 \exp(-C\epsilon^2 t), \tag{40}$$

for all $t \in \mathbb{N}$ and sufficiently small $\epsilon > 0$. The second result is that for all sufficiently large $\gamma > 0$

there exists $\delta, C > 0$ such that

$$\mathrm{E}\big(\sup_{b\in B_\delta(\beta)} \mathbb{1}\{y_t^b \notin B_{\gamma^{1/p}}(0)\}\|y_t^b\|^p\big) \le \exp(-Ct), \tag{41}$$

for all sufficiently large t.

The $p = 0$ case follows immediately from the line (40). Deriving the $p > 0$ case from lines (40) and (41) will take a bit more work. To that end, choose γ large enough so that $\gamma \ge \sup_{b\in B_\delta(\beta)} \|y_\infty^b\|^p$, and hence $\|y_t^b - y_\infty^b\| \le \|y_t^b\| + \gamma^{1/p}$ (Lemma 2 establishes that this is possible). And with this, lines (40) and (41) imply that we can choose $C > 0$ so that we have the following for all sufficiently small ϵ and large t:

$$\begin{aligned}
&\mathrm{E}\big(\sup_{b\in B_\delta(\beta)} \mathbb{1}\{y_t^b \notin B_\epsilon(y_\infty^b)\}\|y_t^b - y_\infty^b\|^p\big)\\
&\quad\le \mathrm{E}\big(\sup_{b\in B_\delta(\beta)} \mathbb{1}\{y_t^b \notin B_\epsilon(y_\infty^b)\}\mathbb{1}\{y_t^b \notin B_{\gamma^{1/p}}(0)\}(\|y_t^b\| + \gamma^{1/p})^p\big)\\
&\qquad + \mathrm{E}\big(\sup_{b\in B_\delta(\beta)} \mathbb{1}\{y_t^b \notin B_\epsilon(y_\infty^b)\}\mathbb{1}\{y_t^b \in B_{\gamma^{1/p}}(0)\}(\|y_t^b\| + \gamma^{1/p})^p\big)\\
&\quad\le \mathrm{E}\big(\sup_{b\in B_\delta(\beta)} \mathbb{1}\{y_t^b \notin B_\epsilon(y_\infty^b)\}\mathbb{1}\{y_t^b \notin B_{\gamma^{1/p}}(0)\}2^p\|y_t^b\|^p\big)\\
&\qquad + \mathrm{E}\big(\sup_{b\in B_\delta(\beta)} \mathbb{1}\{y_t^b \notin B_\epsilon(y_\infty^b)\}\mathbb{1}\{y_t^b \in B_{\gamma^{1/p}}(0)\}2^p\gamma\big)\\
&\quad\le 2^p\,\mathrm{E}\big(\sup_{b\in B_\delta(\beta)} \mathbb{1}\{y_t^b \notin B_{\gamma^{1/p}}(0)\}\|y_t^b\|^p\big) + 2^p\gamma\,\Pr\big(\sup_{b\in B_\delta(\beta)} \|y_t^b - y_\infty^b\| > \epsilon\big)\\
&\quad\le 2^p\exp(-Ct) + 2^{p+2}\gamma m^2\exp(-C\epsilon^2 t).
\end{aligned}$$

The inequality above establishes the $p > 0$ case. Hence, proving lines (40) and (41) will complete the argument.

Before getting into the math, let me roughly sketch the proof of line (40). The key tool will be Lemma 11, which is our only means for positioning y_t^b. The lemma corresponds to a set of inequalities that describe a small box, which is roughly aligned with the orthonormal basis $\{\omega_j^b\}_{i=1}^m$; if these inequalities all hold, then the box is intact, and y_t^b resides inside of it. I will use this result to bound the distance between y_t^b and y_∞^b with the distances between $\dot{\Lambda}_t^b(y_\infty^b + \eta k\omega_j^b)$ and $\eta k\sigma_j^b\omega_j^b$, for $j \in m$, $k \in \{-1, 1\}$, and $\eta > 0$ (these latter distances being the constraints that ensure the integrity of the box). This reframing simplifies the problem, because $\dot{\Lambda}_t^b(y_\infty^b + \eta k\omega_j^b)$ is a sum of *i.i.d.* bounded variables. The second part of the proof replaces the $\eta k\sigma_j^b\omega_j^b$ term in our distance measurements with with $\dot{\Lambda}_\infty^b(y_\infty^b + \eta k\omega_j^b)$. This step is useful because $\dot{\Lambda}_t^b(y_\infty^b + \eta k\omega_j^b) - \dot{\Lambda}_\infty^b(y_\infty^b + \eta k\omega_j^b)$ is an empirical process. The final part of the proof invokes a standard empirical process result to establish the desired concentration of measure.

To begin the proof of line (40), note that Lemmas 1 and 2 imply that we can choose $\delta > 0$ and $\epsilon > 0$ small enough to ensure the existence and continuity of $\ddot{\Lambda}_\infty$ between y^b_∞ and $y^b_\infty + \eta k \omega^b_j$, and small enough to ensure that $\sigma^b_1 \le 2\sigma^\beta_1$, $\sigma^b_m \ge \sigma^\beta_m/2 > 0$, and $y^b_\infty + \eta k \omega^b_j \ge 0$, for all $j \in [m]$, $k \in \{-1, 1\}$, $b \in B_\delta(\beta)$, and $\eta \equiv \epsilon/(1 + 8\sqrt{m}\sigma^\beta_1/\sigma^\beta_m)$. Now, with these conditions, we can use Lemma 11 to bound the left-hand side of (40) in terms of more amenable subgradients:

$$
\begin{aligned}
&\Pr\big(\sup_{b\in B_\delta(\beta)} \|y^b_t - y^b_\infty\| > \epsilon\big)\\
&\quad\le \Pr\big(\sup_{b\in B_\delta(\beta)} \|y^b_t - y^b_\infty\| > \eta(1 + 2\sqrt{m}\sigma^b_1/\sigma^b_m)\big)\\
&\quad\le \Pr\big(\sup_{b\in B_\delta(\beta)} \max_{j\in[m]} \max_{k\in\{-1,1\}} \|\dot{\Lambda}^b_t(y^b_\infty + \eta k\omega^b_j) - \eta k\sigma^b_j\omega^b_j\| - \eta\sigma^b_m/(2\sqrt{m}) > 0\big)\\
&\quad\le \Pr\big(\sup_{b\in B_\delta(\beta)} \max_{j\in[m]} \max_{k\in\{-1,1\}} \|\dot{\Lambda}^b_t(y^b_\infty + \eta k\omega^b_j) - \eta k\sigma^b_j\omega^b_j\| - \eta\sigma^\beta_m/(4\sqrt{m}) > 0\big). \qquad (42)
\end{aligned}
$$

Now I will frame the last expression above as an empirical process by replacing the $\eta k\sigma^b_j\omega^b_j$ term with $\dot{\Lambda}^b_\infty(y^b_\infty + \eta k\omega^b_j)$. To this end, note that the mean value theorem indicates that there exists $\xi \in (0, \eta)$ for which

$$
\begin{aligned}
\dot{\Lambda}^b_\infty(y^b_\infty + \eta k\omega^b_j) &= \dot{\Lambda}^b_\infty(y^b_\infty + \eta k\omega^b_j) - 0\\
&= \dot{\Lambda}^b_\infty(y^b_\infty + \eta k\omega^b_j) - \dot{\Lambda}^b_\infty(y^b_\infty)\\
&= \eta k\ddot{\Lambda}_\infty(y^b_\infty + \xi k\omega^b_j)\omega^b_j\\
&= \eta k\ddot{\Lambda}_\infty(y^b_\infty)\omega^b_j + \eta k(\ddot{\Lambda}_\infty(y^b_\infty + \xi k\omega^b_j) - \ddot{\Lambda}_\infty(y^b_\infty))\omega^b_j\\
&= \eta k\sigma^b_j\omega^b_j + o(\eta),
\end{aligned}
$$

where the little-o term holds uniformly across $b \in B_\delta(\beta)$. Accordingly, we can set ϵ small enough so that $\sup_{b\in B_\delta(\beta)} \|\dot{\Lambda}^b_\infty(y^b_\infty + \eta k\omega^b_j) - \eta k\sigma^b_j\omega^b_j\| \le \eta\sigma^\beta_m/(8\sqrt{m})$, in which case we have

$$
\begin{aligned}
&\|\dot{\Lambda}^b_t(y^b_\infty + \eta k\omega^b_j) - \eta k\sigma^b_j\omega^b_j\|\\
&\quad\le \|\dot{\Lambda}^b_t(y^b_\infty + \eta k\omega^b_j) - \dot{\Lambda}^b_\infty(y^b_\infty + \eta k\omega^b_j)\| + \|\dot{\Lambda}^b_\infty(y^b_\infty + \eta k\omega^b_j) - \eta k\sigma^b_j\omega^b_j\|\\
&\quad\le \|\dot{\Lambda}^b_t(y^b_\infty + \eta k\omega^b_j) - \dot{\Lambda}^b_\infty(y^b_\infty + \eta k\omega^b_j)\| + \eta\sigma^\beta_m/(8\sqrt{m}).
\end{aligned}
$$

And, finally, combining this with line (42) and the fact that $\dot\Lambda_t^b - \dot\Lambda_t^b = \dot\Lambda_t^\beta - \dot\Lambda_t^\beta$ yields the following:

$$
\begin{aligned}
&\Pr\big(\sup_{b\in B_\delta(\beta)} \|y_t^b - y_\infty^b\| > \epsilon\big)\\
&\quad\le \Pr\big(\sup_{b\in B_\delta(\beta)} \max_{j\in[m]} \max_{k\in\{-1,1\}} \|\dot\Lambda_t^b(y_\infty^b + \eta k\omega_j^b) - \eta k\sigma_j^b\omega_j^b\| > \eta\sigma_m^\beta/(4\sqrt{m})\big)\\
&\quad\le \Pr\big(\sup_{b\in B_\delta(\beta)} \max_{j\in[m]} \max_{k\in\{-1,1\}} \|\dot\Lambda_t^b(y_\infty^b + \eta k\omega_j^b) - \dot\Lambda_\infty^b(y_\infty^b + \eta k\omega_j^b)\| > \eta\sigma_m^\beta/(8\sqrt{m})\big)\\
&\quad\le \sum_{i=1}^m\sum_{j=1}^m\sum_{k\in\{-1,1\}} \Pr\big(\sup_{b\in B_\delta(\beta)} |e_i'\dot\Lambda_t^b(y_\infty^b + \eta k\omega_j^b) - e_i'\dot\Lambda_\infty^b(y_\infty^b + \eta k\omega_j^b)| \ge \eta\sigma_m^\beta/(8m)\big)\\
&\quad\le \sum_{i=1}^m\sum_{j=1}^m\sum_{k\in\{-1,1\}} \Pr\big(\sup_{y\in B_\nu(y_\infty^\beta)} |e_i'\dot\Lambda_t^\beta(y) - e_i'\dot\Lambda_\infty^\beta(y)| \ge \eta\sigma_m^\beta/(8m)\big),
\end{aligned}
$$

where $\nu > 0$ is a constant that's large enough to ensure that $y_\infty^b + \eta k\omega_j^b \in B_\nu(y_\infty^\beta)$ for all $b \in B_\delta(\beta)$. Finally, Theorem 2.14.9 of van der Vaart and Wellner (1996) implies that this last expression falls exponentially fast in t (see the proof of lemma 13 for confirmation of this theorem's hypothesis.)

This establishes line (40), which establishes the $p = 0$ case. I will now prove line (41), assuming $p > 0$. The proof will proceed as follows: First, I will bound the probability that $\|y_t^b\|^p$ exceeds some $\gamma > 0$ with the probability that $e_j'\dot\Lambda_t^b(e_j\gamma^{1/p}/\sqrt{m})$ is negative. This latter random variable is easier to work with because it is a sum of *i.i.d.* random variables. Second, I will lower bound $e_j'\dot\Lambda_t^b(e_j\gamma^{1/p}/\sqrt{m})$ with a binomial random variable, with success probability $\rho_\gamma \equiv \Pr(u_1 > \eta\gamma^{1/p}/(2\sqrt{m}))$. This characterization will enable me to use the binomial Chernoff bound to establish that $\Pr(\|y_t^b\|^p > \gamma)$ falls exponentially fast in t. And finally, I will integrate over this tail bound to create a corresponding expectation bound.

To begin the proof, I will show that $e_j'y_t^b > \omega$ implies $e_j'\dot\Lambda_t^b(\omega e_j) \le 0$, for $\omega \in \mathbb{R}$. To see this, take $e_j'y_t^b \geq \omega$ and $\hat{y} \equiv y_t^b - e_j(e_j'y_t^b - \omega)/2$, and apply Lemma 9:

$$\begin{aligned}
0 \geq& (y_t^b - \hat{y})'\dot{\Lambda}_t^b(\hat{y}) \\
=& ((e_j'y_t^b - \omega)/2)e_j'\dot{\Lambda}_t^b(\hat{y}) \\
=& ((e_j'y_t^b - \omega)/2)e_j'\big(b - \sum_{s=1}^{t} \mathbb{1}\{\Delta_s(\hat{y}) > 0\}a_s/t\big) \\
\geq& ((e_j'y_t^b - \omega)/2)e_j'\big(b - \sum_{s=1}^{t} \mathbb{1}\{u_s > a_s'e_je_j'\hat{y}\}a_s/t\big) \\
\geq& ((e_j'y_t^b - \omega)/2)e_j'\big(b - \sum_{s=1}^{t} \mathbb{1}\{u > a'e_j\omega\}a_s/t\big) \\
=& ((e_j'y_t^b - \omega)/2)e_j'\dot{\Lambda}_t^b(\omega e_j).
\end{aligned}$$

Since $e_j'y_t^b - \omega$ is positive, by assumption, it follows that $e_j'\dot{\Lambda}_t^b(\omega e_j)$ must be non-positive.

And now, I'll use this result to replace the shadow price with a simpler subgradient:

$$\begin{aligned}
\Pr\big(\sup_{b \in B_\delta(\beta)} \|y_t^b\|^p > \gamma\big) &\leq \sum_{j=1}^{m} \Pr\big(\sup_{b \in B_\delta(\beta)} e_j'y_t^b > \gamma^{1/p}/\sqrt{m}\big) \\
&\leq \sum_{j=1}^{m} \Pr\big(\sup_{b \in B_\delta(\beta)} e_j'\dot{\Lambda}_t^b(e_j\gamma^{1/p}/\sqrt{m}) \leq 0\big).
\end{aligned}$$

Next, we will bound the complex random variable in the last probability above with a simple binomial random variable. To that end, choose $\delta, \eta > 0$ so that $\eta \leq e_j'b$ for all $b \in B_\delta(\beta)$, in which case we have the following:

$$\begin{aligned}
&\sup_{b \in B_\delta(\beta)} e_j'\dot{\Lambda}_t^b(e_j\gamma^{1/p}/\sqrt{m}) \\
&\qquad = \sup_{b \in B_\delta(\beta)} e_j'b - \sum_{s=1}^{t} \mathbb{1}\{u_s > a_s'e_j\gamma^{1/p}/\sqrt{m})\}e_j'a_s/t \\
&\qquad \geq \sup_{b \in B_\delta(\beta)} e_j'b - \sum_{s=1}^{t} \Big(\mathbb{1}\{e_j'a_s \leq e_j'b/2\}(e_j'b/2)/t + \mathbb{1}\{e_j'a_s > e_j'b/2\}\mathbb{1}\{u_s > a_s'e_j\gamma^{1/p}/\sqrt{m})\}\alpha/t\Big) \\
&\qquad \geq \sup_{b \in B_\delta(\beta)} e_j'b/2 - \sum_{s=1}^{t} \mathbb{1}\{u_s > e_j'b\gamma^{1/p}/(2\sqrt{m})\}\alpha/t \\
&\qquad \geq \eta/2 - \xi_t\alpha/t,
\end{aligned}$$

where $\xi_t \equiv \sum_{s=1}^{t} \mathbb{1}\{u_s > \eta\gamma^{1/p}/(2\sqrt{m})\}$ is a binomial(t, ρ_γ), with $\rho_\gamma \equiv \Pr(u_1 > \eta\gamma^{1/p}/(2\sqrt{m}))$. Further, since $\mathrm{E}(u_1) \leq \infty$, we must have $\rho_\gamma \leq \gamma^{-1/p}$, for sufficiently large γ. Hence, combining the previous two results with the binomial Chernoff bound yields the following for sufficiently large γ:

$$
\begin{aligned}
\Pr\big(\sup_{b\in B_\delta(\beta)} \|y_t^b\|^p > \gamma\big) &\leq \sum_{j=1}^{m} \Pr(\eta/2 - \xi_t\alpha/t \leq 0) \\
&= m \Pr(\xi_t \geq t\eta/(2\alpha)) \\
&\leq m \exp\Big(-\frac{t\eta}{2\alpha}\Big(\log\frac{\eta}{2\alpha\rho_\gamma} - 1\Big)\Big) \\
&\leq m \exp\Big(-\frac{t\eta}{4\alpha}\log\frac{\eta\gamma^{1/p}}{2\alpha}\Big) \\
&= m(\eta/(2\alpha))^{\frac{-t\eta}{4\alpha}}\gamma^{\frac{-t\eta}{4p\alpha}},
\end{aligned}
$$

where the penultimate line supposes that γ is large enough to satisfy $\log(\frac{\eta\gamma^{1/p}}{2\alpha})/2 \geq 1$. Now choosing γ large enough to satisfy the previous result and large enough to ensure that $\|y_t^b\|^p \geq \gamma$ implies $y_t^b \notin B_\epsilon(y_\infty^b)$ yields the following:

$$
\begin{aligned}
&\mathrm{E}\big(\sup_{b\in B_\delta(\beta)} \mathbb{1}\{y_t^b \notin B_{\gamma^{1/p}}(0)\}\|y_t^b\|^p\big) \\
&\leq \gamma \Pr\big(\sup_{b\in B_\delta(\beta)} \|y_t^b\|^p \geq \gamma\big) + \int_{x=\gamma}^{\infty} \Pr\big(\sup_{b\in B_\delta(\beta)} \|y_t^b\|^p > x\big)dx \\
&\leq \gamma \Pr\big(\sup_{b\in B_\delta(\beta)} \|y_t^b - y_\infty^b\| \geq \epsilon\big) + \int_{x=\gamma}^{\infty} m(\eta/(2\alpha))^{\frac{-t\eta}{4\alpha}} x^{\frac{-t\eta}{8\alpha}} dx \\
&\leq \gamma \Pr\big(\sup_{b\in B_\delta(\beta)} \|y_t^b - y_\infty^b\| \geq \epsilon\big) + \frac{m(\eta/(2\alpha))^{\frac{-t\eta}{4\alpha}}\gamma^{1-\frac{t\eta}{8\alpha}}}{\frac{t\eta}{8\alpha} - 1}.
\end{aligned}
$$

The last expression above falls exponentially fast in t, by line (40), so this establishes line (41). $\square$

Corollary 3 Proof. This is the $p = 0$ case of Proposition 6. $\square$

Lemma 3 Proof. Consider an alternative martingale $\{\hat{b}_t\}_{t=n}^{1}$ in which $\hat{b}_n \equiv b_n$ and

$$
\hat{b}_t \equiv \begin{cases} b_t & \hat{b}_{t+1} \in B_\delta(\beta), \\ \hat{b}_{t+1} & \hat{b}_{t+1} \notin B_\delta(\beta). \end{cases}
$$

In other words, $\hat{b}_t$ tracks b_t until the first time that b_t departs $B_\delta(\beta)$, at which point $\hat{b}_t$ remains frozen in place. By design, $\hat{b}_t \in B_\delta(\beta)$ implies $b_t \in B_\delta(\beta)$, and hence $\Pr(b_t \notin B_\delta(\beta)) \leq \Pr(\hat{b}_t \notin B_\delta(\beta))$.

And, with this, the result follows from the Azuma–Hoeffding inequality, since $\|\hat{b}_t - \hat{b}_{t+1}\| \leq (\|\beta\| + \delta + \|\alpha\|)/t$:

$$\begin{aligned}
&\Pr(b_t \notin B_\delta(\beta)) \\
&\quad \leq \sup_{N \geq t} \Pr(\hat{b}_t \notin B_\delta(\beta))) \\
&\quad \leq \sup_{N \geq t} \sum_{j=1}^{m} \Pr\left(|e_j'\hat{b}_t - e_j' b_n| \geq \delta/\sqrt{m}\right) \\
&\quad \leq \sup_{N \geq t} 2m \exp\Big(-\frac{\delta^2/m}{2\sum_{s=t}^{N-1}(\|\beta\| + \delta + \|\alpha\|)^2/s^2}\Big) \\
&\quad < 2m \exp\Big(-\frac{\delta^2}{2m(\|\beta\| + \delta + \|\alpha\|)^2 \int_{s=t-1}^{\infty} dss^2}\Big) \\
&\quad < 2m \exp\Big(-\frac{\delta^2(t-1)}{2m(\|\beta\| + \delta + \|\alpha\|)^2}\Big). \qquad (43)
\end{aligned}$$

□

Corollary 4 Proof. Let $\{\hat{b}_t\}_{t=n}^{1}$ be the alternative martingale defined in the proof of Lemma 3. Note that we have $\hat{b}_t \notin B_\delta(\beta)$ if and only if $t \leq \tau(\delta)$. And, with this, line (43) implies the result:

$$\begin{aligned}
\mathrm{E}(\tau(\delta)) &= \sum_{s=1}^{n} \Pr(\tau(\delta) \geq s) \\
&= \sum_{t=1}^{n} \Pr(\hat{b}_t \notin B_\delta(\beta)) \\
&< \sum_{t=1}^{\infty} 2m \exp\Big(-\frac{\delta^2(t-1)}{2m(\|\beta\| + \delta + \|\alpha\|)^2}\Big) \\
&= O(1).
\end{aligned}$$

□

Lemma 4 Proof. Bounding the first term of the myopic regret is simple: Since $\mathbb{1}\{b_t \notin B_{\delta/2}(\beta)\}$ is independent of $\sum_{s=1}^t u_s$, Lemma 3 indicates that there exists $C > 0$ for which

$$\begin{aligned}
&\mathrm{E}\left(\mathbb{1}\{b_t \notin B_{\delta/2}(\beta)\}\sum_{s=1}^t u_s\right) \\
&\quad = \Pr(b_t \notin B_{\delta/2}(\beta))\sum_{s=1}^t \mathrm{E}(u_s) \\
&\quad \leq \exp(-Ct)t\,\mathrm{E}(u_1) \\
&\quad = o(1/t).
\end{aligned}$$

I will now bound the second term of the myopic regret with Propositions 4 and 6, and Corollary 3. Doing so will take several steps. First, the independence of b_t from the random mapping $b \mapsto \mathbb{1}\{\Delta_t(y_\infty^b) > 0\}\Delta_t(y_{t-1}^{\psi_t^b(a_t)})^-$ establishes that

$$\begin{aligned}
&\mathrm{E}\left(\mathbb{1}\{b_t \in B_{\delta/2}(\beta)\}\mathbb{1}\{\Delta_t(y_\infty^{b_t}) > 0\}\Delta_t(y_{t-1}^{\psi_t^{b_t}(a_t)})^-\right) \\
&\quad = \mathrm{E}\left(\mathbb{1}\{b_t \in B_{\delta/2}(\beta)\}\,\mathrm{E}\left(\mathbb{1}\{\Delta_t(y_\infty^b) > 0\}\Delta_t(y_{t-1}^{\psi_t^b(a_t)})^-\right)\Big|_{b=b_t}\right).
\end{aligned}$$

Accordingly, it will suffice to show that

$$\sup_{b \in B_{\delta/2}(\beta)} \mathrm{E}\left(\mathbb{1}\{\Delta_t(y_\infty^b) > 0\}\Delta_t(y_{t-1}^{\psi_t^b(a_t)})^-\right) = O(1/t). \tag{44}$$

Next, I will remove the random a_t in the superscript of $y_{t-1}^{\psi_t^b(a_t)}$ by showing that the following holds, for all $b \in B_{\delta/2}(\beta)$ and large t:

$$\mathbb{1}\{\Delta_t(y_\infty^b) > 0\}\Delta_t(y_{t-1}^{\psi_t^b(a_t)})^- \leq \mathbb{1}\{\Delta_t(y_\infty^b) > 0\}\Delta_t(y_\infty^b + \xi_{t-1}\iota)^-, \tag{45}$$

where $\xi_{t-1} \equiv \sup_{b \in B_\delta(\beta)} \|y_{t-1}^b - y_\infty^b\| + C/t$, for some constant C that has yet to be defined. Note that the first term of ξ_{t-1} appears compatible with our shadow price convergence results. Moreover, ξ_{t-1}, unlike $y_{t-1}^{\psi_t^b(a_t)}$, is independent of the random function Δ_t, which makes the right-hand side of (45) congruent with Lemma 12. This lemma will convert ξ_{t-1} into something like $\sup_{b \in B_\delta(\beta)} \|y_t^b - y_\infty^b\|^2$, which is precisely what we need to invoke Proposition 4. But there will also be a few stray terms, which we'll bound with Proposition 6 and Corollary 3.

I will now derive line (45). First, note that Assumption 4 and Lemma 2 imply there exits $C > 0$

that satisfies the following, almost surely:

$$\sup_{b \in B_{\delta/2}(\beta)} \|y_\infty^{\psi_t^b(a_t)} - y_\infty^b\| \leq C/t.$$

Accordingly, if we choose t large enough so that $b \in B_{\delta/2}(\beta)$ implies $\psi_t^b(a_t) \in B_{\delta/2}(\beta)$, then we have the following:

$$\begin{aligned}
\|y_{t-1}^{\psi_t^b(a_t)} - y_\infty^b\| &\leq \|y_{t-1}^{\psi_t^b(a_t)} - y_\infty^{\psi_t^b(a_t)}\| + \|y_\infty^{\psi_t^b(a_t)} - y_\infty^b\| \\
&\leq \sup_{b \in B_\delta(\beta)} \|y_{t-1}^b - y_\infty^b\| + C/t \\
&\equiv \xi_{t-1}.
\end{aligned}$$

And, by design, this new variable we've defined satisfies $y_{t-1}^{\psi_t^b(a_t)} \leq y_\infty^b + \xi_{t-1}\iota$, for $b \in B_{\delta/2}(\beta)$ and large t, which establishes line (45).

Next, I will apply Lemma 12 to the right-hand side of (45). I can only do so when y_∞^b and $y_\infty^b + \xi_{t-1}\iota$ reside in the ϵ-ball of y_∞^β, for some sufficiently small ϵ. Lemma 2 guarantees that the first vector can be thus situated, when $b \in B_{\delta/2}(\beta)$ and δ is small, but we have no such guarantee for the second vector. So we will have to consider the $y_\infty^b + \xi_{t-1}\iota \in B_\epsilon(y_\infty^\beta)$ and $y_\infty^b + \xi_{t-1}\iota \notin B_\epsilon(y_\infty^\beta)$ cases separately:

$$\begin{aligned}
&\mathrm{E}\left(\mathbb{1}\{\Delta_t(y_\infty^b) > 0\}\Delta_t(y_\infty^b + \xi_{t-1}\iota)^-\right) \\
&\quad = \mathrm{E}\left(\mathrm{E}\left(\mathbb{1}\{y_\infty^b + \xi_{t-1}\iota \in B_\epsilon(y_\infty^\beta)\}\mathbb{1}\{\Delta_t(y_\infty^b) > 0\}\Delta_t(y_\infty^b + \xi_{t-1}\iota)^- : \xi_{t-1}\right)\right) \\
&\qquad + \mathrm{E}\left(\mathbb{1}\{y_\infty^b + \xi_{t-1}\iota \notin B_\epsilon(y_\infty^\beta)\}\mathbb{1}\{\Delta_t(y_\infty^b) > 0\}\Delta_t(y_\infty^b + \xi_{t-1}\iota)^-\right) \\
&\quad \leq \mathrm{E}\left(\mathbb{1}\{y_\infty^b + \xi_{t-1}\iota \in B_\epsilon(y_\infty^\beta)\}2\sigma_1^\beta\|\xi_{t-1}\iota\|^2\right) \\
&\qquad + \mathrm{E}\left(\mathbb{1}\{y_\infty^b + \xi_{t-1}\iota \notin B_\epsilon(y_\infty^\beta)\}(\|\alpha\|\|y_\infty^b\| + \|\alpha\|\|\xi_{t-1}\iota\|)\right) \\
&\quad \leq 2\sigma_1^\beta\|\iota\|^2\,\mathrm{E}(\xi_{t-1}^2) + \|\alpha\|\|y_\infty^b\|\Pr(y_\infty^b + \xi_{t-1}\iota \notin B_\epsilon(y_\infty^\beta)) + \\
&\qquad + \|\alpha\|\|\iota\|\,\mathrm{E}\left(\mathbb{1}\{y_\infty^b + \xi_{t-1}\iota \notin B_\epsilon(y_\infty^\beta)\}\xi_{t-1}\right) \\
&\quad \leq 2\sigma_1^\beta\|\iota\|^2\,\mathrm{E}(\xi_{t-1}^2) + 2\|\alpha\|\|y_\infty^\beta\|\Pr(\xi_{t-1} > \epsilon/(2\sqrt{m})) + \\
&\qquad + \|\alpha\|\|\iota\|\,\mathrm{E}\left(\mathbb{1}\{\xi_{t-1} > \epsilon/(2\sqrt{m})\}\xi_{t-1}\right).
\end{aligned}$$

For the last line, I suppose that δ is small enough so that (i) $\sup_{b \in B_{\delta/2}(\beta)} \|y_\infty^b\| \leq 2\|y_\infty^\beta\|$ and (ii) $\{y_\infty^b : b \in B_{\delta/2}(\beta)\} \subset B_{\epsilon/2}(y_\infty^\beta)$, in which case $y_\infty^b + \xi_{t-1}\iota \notin B_\epsilon(y_\infty^\beta)$ implies $\xi_{t-1} > \epsilon/(2\sqrt{m})$.

Further, Proposition 4 establishes that the first term in the last line above is $O(1/t)$, Corollary 3 establishes that the second term is $o(1/t)$, and Proposition 6 establishes that the third term is $o(1/t)$. Hence, the last expression—which holds for all $b \in B_{\delta/2}(\beta)$—is $O(1/t)$, which with line (45) establishes line (44).

Finally, the same argument yields the same bound for the third term of myopic regret. $\square$

Lemma 5 Proof. Let $\{\hat{b}_t\}_{t=n}^{1}$ denote the inventory process defined in the proof of Lemma 3, but derived from from Algorithm 3's b_t values. Just to remind you, the $\{\hat{b}_t\}_{t=n}^{1}$ process tracks the $\{b_t\}_{t=n}^{1}$ process until time $\tau(\delta)$—i.e., until Algorithm 3's b_t values first depart $B_\delta(\beta)$—at which point the process freezes in place. The $\{\hat{b}_t\}_{t=n}^{1}$ process will be easier to study because a constant multiple of t bounds its innovations. And since $\hat{b}_t \in B_\delta(\beta)$ implies $b_t \in B_\delta(\beta)$, it will suffice to establish the concentration of measure for $\hat{b}_t$.

I will bound the distance between $\hat{b}_t$ and β with the following inequality:

$$\|\hat{b}_t - \beta\| \leq \|\hat{b}_{\tau(\delta/2)+1} - \beta\| + \|\xi_t\| + \sum_{s=t}^{\tau(\delta/2)} \| \mathrm{E}(\hat{b}_s : \hat{b}_{s+1}) - \hat{b}_{s+1}\|, \tag{46}$$

$$\text{where} \quad \xi_t \equiv \sum_{s=t}^{\tau(\delta/2)} \hat{b}_s - \mathrm{E}(\hat{b}_s : \hat{b}_{s+1}).$$

I cap the sums at time $\tau(\delta/2)$ to give our look-back shadow prices a sufficiently large sample. Indeed, a sample with $n - \tau(\delta/2)$ observations will comprise enough data to ensure that the look-back shadow prices—and hence the $\hat{b}_t$ values—are well-behaved. More specifically, I will show that $n - \tau(\delta/2) = \Theta(n)$ by showing that there exists $\gamma < 1$ that satisfies

$$\tau(\delta/2) + 1 \leq \gamma n. \tag{47}$$

To see this, note that period-t's resource vector satisfies

$$(n\beta - (n-t)\alpha)/t \leq \underbrace{\big(n\beta - \sum_{s=t}^{n} x_s a_s\big)/t}_{=b_t} \leq n\beta/t,$$

where the lower bound is within $\delta/2$ of β unless $t \leq \frac{n}{1+\delta/(2\|\alpha-\beta\|)}$, and the upper bound is within $\delta/2$ of β unless $t \leq \frac{n}{1+\delta/(2\|\beta\|)}$. Hence, if $\|\alpha - \beta\| \geq \|\beta\|$, which we can suppose without loss of generality, then $b_t \notin B_{\delta/2}(\beta)$ implies $t \leq \frac{n}{1+\delta/(2\|\alpha-\beta\|)}$.

I will now use (46) to inductively prove that there exists $C > 0$ such that

$$\Pr\big(\max_{s=t}^{n} \|\hat{b}_s - \beta\| > \delta\big) \leq (\tau(\delta/2) + 1 - t)\big(2\exp(-Ct) + 2n\exp(-C(1-\gamma)\sqrt{n})\big), \tag{48}$$

for all sufficiently large $t \leq n$. Initializing our induction will be simple: by definition, we have $\Pr(\hat{b}_t \in B_\delta(\beta)) = 1$ for $t \geq \tau(\delta/2) + 1$, which establishes the base case. However, establishing the inductive step will require unraveling the knotty relationship between look-back shadow prices and inventory vectors. Specifically, showing that $\|\hat{b}_t - \beta\|$ is small for $t \leq \tau(\delta/2)$ will require showing that $\|\operatorname{E}(\hat{b}_s : \hat{b}_{s+1}) - \hat{b}_{s+1}\|$ is small for all $s \in \{t, \cdots, \tau(\delta/2)\}$, which in turn will require showing that $\|\underleftarrow{y}_s^{\hat{b}_s} - y_\infty^{\hat{b}_s}\|$ is small for all $s \in \{t+1, \cdots, \tau(\delta/2)+1\}$, which in turn will require showing that $\|\hat{b}_s - \beta\|$ is small for all $s \in \{t+1, \cdots, \tau(\delta/2)+1\}$.

I will now that if (48) holds for sufficiently large $t + 1 \leq \tau(\delta/2) + 1$, then there a suitably high probability that

$$\begin{aligned} \|\hat{b}_{\tau(\delta/2)+1} - \beta\| &\leq \delta/2, \\ \|\xi_t\| &\leq \delta/4, \\ \text{and} \qquad \sum_{s=t}^{\tau(\delta/2)} \|\operatorname{E}(\hat{b}_s : \hat{b}_{s+1}) - \hat{b}_{s+1}\| &\leq \delta/4, \end{aligned} \tag{49}$$

which with line (46) will establish induction. Note that the first inequality in display (49) holds by the definition of $\tau(\delta/2)$, so we will only have to concern ourselves with the latter two inequalities.

I will now show that the second inequality in display (49) holds with high probability, conditional on $\hat{b}_s \in B_\delta(\beta)$ for all $s \in \{t+1, \cdots, \tau(\delta/2)+1\}$. Since $\{\xi_t\}_{t=\tau(\delta/2)}^{1}$ is a martingale that satisfies $\|\xi_t - \xi_{t+1}\| = \|\hat{b}_t - \operatorname{E}(\hat{b}_t : \hat{b}_{t+1})\| \leq (\|\beta\| + \delta + \|\alpha\|)/t$, by design, the argument underlying line (43) analogously implies that there exists $C > 0$ such that $\Pr(\|\xi_t\| > \delta/4) \leq \exp(-Ct)$, for all sufficiently large t. And since $\Pr(A|B) = \Pr(A \cap B)/\Pr(B) \leq \Pr(A)/\Pr(B)$, it follows that

$$\begin{aligned} &\Pr\big(\|\xi_t\| > \delta/4 : \max_{s=t+1}^{\tau(\delta/2)+1} \|\hat{b}_s - \beta\| \leq \delta\big) \\ &\leq \frac{\Pr(\|\xi_t\| > \delta/4)}{\Pr\big(\max_{s=t+1}^{\tau(\delta/2)+1} \|\hat{b}_s - \beta\| \leq \delta\big)} \\ &\leq 2\exp(-Ct). \end{aligned} \tag{50}$$

Note, the last line holds because $\Pr(\max_{s=t+1}^{\tau(\delta/2)+1} \|\hat{b}_s - \beta\| \leq \delta) \geq 1/2$, by our inductive hypothesis.

I will now show that the third inequality in display (49) holds with high probability, conditional on $\hat{b}_s \in B_\delta(\beta)$ for all $s \in \{t+1, \cdots, \tau(\delta/2)+1\}$. This step will take more work. First note that $\hat{b}_{s+1} \in B_\delta(\beta)$ implies $\hat{b}_{s+1} = b_{s+1}$ and $\hat{b}_s = b_s$, and thus implies

$$
\begin{aligned}
&\mathrm{E}(\hat{b}_s : \hat{b}_{s+1}) - \hat{b}_{s+1} \\
&\quad = \mathrm{E}(b_s : b_{s+1}) - b_{s+1} \\
&\quad = ((s+1)b_{s+1} - \mathrm{E}(\mathbb{1}\{\Delta_{s+1}(\underleftarrow{y}_{s+1}^{b_{s+1}}) > 0\}a_{s+1} : b_{s+1}))/s - b_{s+1} \\
&\quad = ((s+1)b_{s+1} - b_{s+1} + \dot{\Lambda}_\infty^{b_{s+1}}(\underleftarrow{y}_{s+1}^{b_{s+1}}))/s - b_{s+1} \\
&\quad = \dot{\Lambda}_\infty^{b_{s+1}}(\underleftarrow{y}_{s+1}^{b_{s+1}})/s \\
&\quad = \ddot{\Lambda}_\infty(y_\infty^{b_{s+1}})(\underleftarrow{y}_{s+1}^{b_{s+1}} - y_\infty^{b_{s+1}})/s + o(\|\underleftarrow{y}_{s+1}^{b_{s+1}} - y_\infty^{b_{s+1}}\|)/s \\
&\quad = \ddot{\Lambda}_\infty(y_\infty^{\hat{b}_{s+1}})(\underleftarrow{y}_{s+1}^{\hat{b}_{s+1}} - y_\infty^{\hat{b}_{s+1}})/s + o(\|\underleftarrow{y}_{s+1}^{\hat{b}_{s+1}} - y_\infty^{\hat{b}_{s+1}}\|)/s,
\end{aligned}
$$

where the penultimate line holds by Lemma 1, since $\dot{\Lambda}_\infty^{b_{s+1}}(y_\infty^{b_{s+1}}) = 0$ when $b_{s+1} \in B_\delta(\beta)$ and δ is small. Thus, we can choose n sufficiently small so that $\max_{s=t+1}^{\tau(\delta/2)+1} \|\hat{b}_s - \beta\| \leq \delta$ and $\max_{s=t+1}^{\tau(\delta/2)+1} \|\underleftarrow{y}_s^{\hat{b}_s} - y_\infty^{\hat{b}_s}\| \leq n^{-1/4}$ imply

$$
\begin{aligned}
&\sum_{s=t}^{\tau(\delta/2)} \|\mathrm{E}(\hat{b}_s : \hat{b}_{s+1}) - \hat{b}_{s+1}\| \\
&\quad \leq \sum_{s=t}^{\tau(\delta/2)} \|\ddot{\Lambda}_\infty(y_\infty^{\hat{b}_{s+1}})(\underleftarrow{y}_{s+1}^{\hat{b}_{s+1}} - y_\infty^{\hat{b}_{s+1}})\|/s + o(\|\underleftarrow{y}_{s+1}^{\hat{b}_{s+1}} - y_\infty^{\hat{b}_{s+1}}\|)/s \\
&\quad \leq \sum_{s=t}^{\tau(\delta/2)} 2\sigma_1^\beta \|\underleftarrow{y}_{s+1}^{\hat{b}_{s+1}} - y_\infty^{\hat{b}_{s+1}}\|/s + o(\|\underleftarrow{y}_{s+1}^{\hat{b}_{s+1}} - y_\infty^{\hat{b}_{s+1}}\|)/s \\
&\quad \leq \sum_{s=t}^{\tau(\delta/2)} 3\sigma_1^\beta n^{-1/4}/s \\
&\quad \leq \delta/4. \qquad (51)
\end{aligned}
$$

Note, the third line holds because the largest singular value of $\ddot{\Lambda}_\infty(y_\infty^{\hat{b}_{s+1}})$ is less than twice the largest singular value of $\ddot{\Lambda}_\infty(y_\infty^\beta)$ when $\hat{b}_{s+1} \in B_\delta(\beta)$ and δ is small, and the fourth line holds because the little-o term is less than $\sigma_1^\beta n^{-1/4}$ when $\|\underleftarrow{y}_{s+1}^{\hat{b}_{s+1}} - y_\infty^{\hat{b}_{s+1}}\| \leq n^{-1/4}$ and n is large.

Further, we can use Corollary 3 upper bound the probability that $\max_{s=t+1}^{\tau(\delta/2)+1} \|\underleftarrow{y}_s^{\hat{b}_s} - y_\infty^{\hat{b}_s}\| > n^{-1/4}$, conditional on $\max_{s=t+1}^{\tau(\delta/2)+1} \|\hat{b}_s - \beta\| \leq \delta$. Specifically, combining this corollary with line (47) and our inductive hypothesis yields the following, for some $C > 0$ and all sufficiently large n:

$$
\begin{aligned}
&\Pr\big(\max_{s=t+1}^{\tau(\delta/2)+1} \| \underleftarrow{y}_s^{\hat{b}_s} - y_\infty^{\hat{b}_s}\| > n^{-1/4} : \max_{s=t+1}^{\tau(\delta/2)+1} \|\hat{b}_s - \beta\| \leq \delta\big) \\
&\quad \leq \sum_{s=t+1}^{\tau(\delta/2)+1} \Pr\big(\sup_{b \in B_\delta(\beta)} \| \underleftarrow{y}_s^{b} - y_\infty^{b}\| > n^{-1/4} : \max_{s=t+1}^{\tau(\delta/2)+1} \|\hat{b}_s - \beta\| \leq \delta\big) \\
&\quad \leq \sum_{s=t+1}^{\tau(\delta/2)+1} \frac{\Pr\big(\sup_{b \in B_\delta(\beta)} \| \underleftarrow{y}_s^{b} - y_\infty^{b}\| > n^{-1/4}\big)}{\Pr\big(\max_{s=t+1}^{\tau(\delta/2)} \|\hat{b}_s - \beta\| \leq \delta\big)} \\
&\quad \leq \sum_{s=t+1}^{\tau(\delta/2)+1} \frac{\exp(-Cn^{-1/2}(n-s))}{1/2} \\
&\quad \leq 2n \exp(-Cn^{-1/2}(n - \tau(\delta/2) - 1)) \\
&\quad \leq 2n \exp(-C(1-\gamma)\sqrt{n}). \qquad (52)
\end{aligned}
$$

And now, finally, we can combine lines (46), (50), (51), and (52) to establish that

$$
\begin{aligned}
&\Pr(\hat{b}_t \notin B_\delta(\beta) : \max_{s=t+1}^{n} \|\hat{b}_s - \beta\| \leq \delta) \\
&\quad \leq \Pr\big(\|\xi_t\| > \delta/4 : \max_{s=t+1}^{n} \|\hat{b}_s - \beta\| \leq \delta\big) \\
&\qquad + \Pr\big(\sum_{s=t}^{\tau(\delta/2)} \| \mathrm{E}(\hat{b}_s : \hat{b}_{s+1}) - \hat{b}_{s+1}\| > \delta/4 : \max_{s=t+1}^{n} \|\hat{b}_s - \beta\| \leq \delta\big) \\
&\quad \leq 2\exp(-Ct) + \Pr\big(\max_{s=t+1}^{\tau(\delta/2)+1} \| \underleftarrow{y}_s^{\hat{b}_s} - y_\infty^{\hat{b}_s}\| > n^{-1/4} : \max_{s=t+1}^{n} \|\hat{b}_s - \beta\| \leq \delta\big) \\
&\quad \leq 2\exp(-Ct) + 2n\exp(-C(1-\gamma)\sqrt{n}).
\end{aligned}
$$

And with our inductive hypothesis, this implies that

$$
\begin{aligned}
&\Pr\big(\max_{s=t}^{n} \|\hat{b}_s - \beta\| > \delta\big) \\
&\quad = \Pr\big(\max_{s=t+1}^{n} \|\hat{b}_s - \beta\| > \delta\big) + \Pr(\hat{b}_t \notin B_\delta(\beta) : \max_{s=t+1}^{n} \|\hat{b}_s - \beta\| \leq \delta) \\
&\quad \leq (\tau(\delta/2) + 1 - t - 1)(2\exp(-C(t+1)) + 2n\exp(-C(1-\gamma)\sqrt{n})) \\
&\qquad + 2\exp(-Ct) + 2n\exp(-C(1-\gamma)\sqrt{n}) \\
&\quad \leq (\tau(\delta/2) + 1 - t)\big(2\exp(-Ct) + 2n\exp(-C(1-\gamma)\sqrt{n})\big).
\end{aligned}
$$

□

Lemma 6 Proof. This proof will closely follow the Lemma 4 proof. That proof's argument establishes that the first term of the myopic regret is $o(1/t)$, so I will begin with the second term of the myopic regret. First, define $\xi_{t-1} \equiv \sup_{b \in B_\delta(\beta)} \|y^b_{t-1} - y^b_\infty\| + C/t$, for the $C > 0$ used in the proof of Lemma 4, and analogously define $\underleftarrow{\xi}_{t+1} \equiv \sup_{b \in B_{\delta/2}(\beta)} \|\underleftarrow{y}^{b}_{t} - y^b_\infty\|$. Note, I subscript this latter variable with $t+1$ because $\underleftarrow{y}^{b}_{t}$ is determined by that time. Next, the argument underlying line (45) establishes the following, for large t:

$$
\begin{aligned}
&\mathrm{E}\left(\mathbb{1}\{b_t \in B_{\delta/2}(\beta)\}\mathbb{1}\{\Delta_t(\underleftarrow{y}^{b_t}_{t}) > 0\}\Delta_t(y^{\psi^{b_t}_t(a_t)}_{t-1})^-\right) \\
&\quad\le \mathrm{E}\left(\mathbb{1}\{b_t \in B_{\delta/2}(\beta)\}\mathbb{1}\{\Delta_t(y^{b_t}_\infty - \underleftarrow{\xi}_{t+1}\iota) > 0\}\Delta_t(y^{b_t}_\infty + \xi_{t-1}\iota)^-\right) \\
&\quad= \mathrm{E}\left(\mathbb{1}\{b_t \in B_{\delta/2}(\beta)\}\,\mathrm{E}(\mathbb{1}\{\Delta_t(y) > 0\}\Delta_t(\bar{y})^-)\big|_{y = y^{b_t}_\infty - \underleftarrow{\xi}_{t+1}\iota,\ \bar{y} = y^{b_t}_\infty + \xi_{t-1}\iota}\right).
\end{aligned}
$$

The last line holds because the random function Δ_t is independent of b_t, $\underleftarrow{\xi}_{t+1}$, and ξ_{t-1}.

Now let $\mathcal{E}$ represent the event in which $y^b_\infty + \xi_{t-1}\iota \in B_\epsilon(y^\beta_\infty)$ and let $\underleftarrow{\mathcal{E}}$ represent the event in which $y^b_\infty - \underleftarrow{\xi}_t\iota \in B_\epsilon(y^\beta_\infty)$, where ϵ is the constant defined in Lemma 12. Further, let δ be small enough so that (i) $\sup_{b \in B_{\delta/2}(\beta)} \|y^b_\infty\| \le 2\|y^\beta_\infty\|$ and (ii) $\{y^b_\infty : b \in B_{\delta/2}(\beta)\} \subset B_{\epsilon/2}(y^\beta_\infty)$, in which case $\mathcal{E}^c$ implies $\xi_{t-1} > \epsilon/(2\sqrt{m})$ and $\underleftarrow{\mathcal{E}}^c$ implies $\underleftarrow{\xi}_{t+1} > \epsilon/(2\sqrt{m})$. And with this, we can use Lemma 12 to continue where we left off:

$$
\begin{aligned}
&\mathrm{E}\left(\mathbb{1}\{b_t \in B_{\delta/2}(\beta)\}\mathbb{1}\{\Delta_t(\underleftarrow{y}^{b_t}_{t}) > 0\}\Delta_t(y^{\psi^{b_t}_t(a_t)}_{t-1})^-\right) \\
&\quad\le \mathrm{E}\left(\mathbb{1}\{b_t \in B_{\delta/2}(\beta)\}\mathbb{1}\{\mathcal{E} \cap \underleftarrow{\mathcal{E}}\}\,\mathrm{E}(\mathbb{1}\{\Delta_t(y) > 0\}\Delta_t(\bar{y})^-)\big|_{y = y^{b_t}_\infty - \underleftarrow{\xi}_{t+1}\iota,\ \bar{y} = y^{b_t}_\infty + \xi_{t-1}\iota}\right) \\
&\qquad+ \mathrm{E}\left(\mathbb{1}\{b_t \in B_{\delta/2}(\beta)\}\mathbb{1}\{\mathcal{E}^c \cup \underleftarrow{\mathcal{E}}^c\}\,\mathrm{E}(\mathbb{1}\{\Delta_t(y) > 0\}\Delta_t(\bar{y})^-)\big|_{y = y^{b_t}_\infty - \underleftarrow{\xi}_{t+1}\iota,\ \bar{y} = y^{b_t}_\infty + \xi_{t-1}\iota}\right) \\
&\quad\le \mathrm{E}\left(\mathbb{1}\{b_t \in B_{\delta/2}(\beta)\}\mathbb{1}\{\mathcal{E} \cap \underleftarrow{\mathcal{E}}\}2\sigma^\beta_1\|\xi_{t-1}\iota + \underleftarrow{\xi}_{t+1}\iota\|^2\right) \\
&\qquad+ \mathrm{E}\left(\mathbb{1}\{b_t \in B_{\delta/2}(\beta)\}\mathbb{1}\{\mathcal{E}^c \cup \underleftarrow{\mathcal{E}}^c\}\|\alpha\|\|y^{b_t}_\infty + \xi_{t-1}\iota\|\right) \\
&\quad\le 4m\sigma^\beta_1\,\mathrm{E}(\xi^2_{t-1} + \underleftarrow{\xi}^{2}_{t+1}) \\
&\qquad+ 2\|\alpha\|\|y^\beta_\infty\|\Pr\left(\xi_{t-1} > \epsilon/(2\sqrt{m}) \cup \underleftarrow{\xi}_{t+1} > \epsilon/(2\sqrt{m})\right) \\
&\qquad+ 2\sqrt{m}\|\alpha\|\|y^\beta_\infty\|\,\mathrm{E}\left(\mathbb{1}\{\xi_{t-1} > \epsilon/(2\sqrt{m}) \cup \underleftarrow{\xi}_{t+1} > \epsilon/(2\sqrt{m})\}\xi_{t-1}\right),
\end{aligned}
$$

where the last line holds because $(x+y)^2 \le 2x^2 + 2y^2$ and $\|\iota\| = \sqrt{m}$. Finally, Proposition 4, Proposition 6, and Corollary 3 imply that this last expression is less than $C/t + C/(n-t)$, for some $C > 0$.

Finally, the same argument yields the same bound for the third term of myopic regret. $\square$

Corollary 5 Proof. Copy the proof of Corollary 4. $\square$

Lemma 7 Proof. I will begin with a high-level plan of attack. The main idea of the proof is that allocating $t(\beta+\xi)$ units of inventory to the first t periods leaves us with only $(n-t)(\beta - \frac{t}{n-t}\xi)$ units for the last $n-t$ periods, so $b_t = \beta + \xi$ must imply

$$\begin{aligned} v_n^\beta &\leq \bar{V}_t^{\beta+\xi} + \underleftarrow{\bar{V}}_t^{\beta - \frac{t}{n-t}\xi} \\ &= t\Lambda_t^{\beta+\xi}(y_t^{\beta+\xi}) + (n-t)\underleftarrow{\Lambda}_t^{\beta-\frac{t}{n-t}\xi}(\underleftarrow{y}_t^{\beta-\frac{t}{n-t}\xi}) \\ &\equiv \bar{\Lambda}_t^\xi \end{aligned} \tag{53}$$

where $\underleftarrow{\bar{V}}_t^b$, $\underleftarrow{\Lambda}_t^b$, and $\underleftarrow{y}_t^b$ and are equivalent to $\bar{V}_{n-t}^b$, Λ_{n-t}^b, and y_{n-t}^b, but with the order of the customers reversed. Accordingly, it follows that

$$\underline{R}_n \geq \bar{V}_n^\beta - \bar{\Lambda}_t^{b_t-\beta} = n\Lambda_n^\beta(y_n^\beta) - \bar{\Lambda}_t^{b_t-\beta}. \tag{54}$$

And the expression on the right should be large when b_t meaningfully deviates from β since, in the limit, we have

$$n\Lambda_\infty^\beta(y_\infty^\beta) - \bar{\Lambda}_\infty^\xi \geq Ct\min(\|\xi\|^2, 1), \tag{55}$$

for some $C > 0$, where[4]

$$\bar{\Lambda}_\infty^\xi \equiv t\Lambda_\infty^{\beta+\xi}(y_\infty^{\beta+\xi}) + (n-t)\Lambda_\infty^{\beta-\frac{t}{n-t}\xi}(y_\infty^{\beta-\frac{t}{n-t}\xi}). \tag{56}$$

Line (55) suggests that reserving $t(\beta+\xi)$ units of inventory for the first t periods should sacrifice $\Omega(t)$ units of value when ξ non-negligible. I will leverage this fact to show that $\underline{R}_n$ is almost always large when $\|b_t - \beta\|$ is non-negligible. And since $\mathrm{E}(\underline{R}_n)$ is relatively small, by Theorem 1, it follows that $\|b_t - \beta\|$ is usually negligible.

Before delving into the details, I will provide a more thorough proof sketch. The proof will have five steps. The first derives limiting bound (55), our only tool for establishing the cost of b_t diverging from β. Now with (55), it's relatively easy to lower bound the regret when $b_t = \beta + \xi$, for some specific ξ that lies outside of a ball of the origin. But that's not enough, as we must lower bound this regret when $b_t = \beta + \xi$, for *any* ξ that lies outside of a ball of the origin. To create such a uniform result, the second part of the proof bounds $\bar{\Lambda}_t^{\xi}$ for all $\xi \in \mathbb{R}^m$ in terms of $\bar{\Lambda}_t^{\zeta}$, $y_t^{\beta+\zeta}$, and $\underleftarrow{y}_t^{\beta-\frac{t}{n-t}\zeta}$, for some given given $\zeta \in \mathbb{R}^m$, and the third part uses this bound to show that $\sup_{\xi \notin B_{2\sqrt{m}\epsilon}(0)} \bar{\Lambda}_t^{\xi}$ is usually smaller than $\max_{k\in\{-1,1\}} \max_{j\in[m]} \bar{\Lambda}_t^{k\epsilon\omega_j^{\beta}}$, where ϵ is a small positive number and $\{\omega_i^{\beta}\}_{i\in[m]}$ are the orthonormal eigenvectors of $\ddot{\Lambda}_\infty(y_\infty^{\beta})^{-1}$. Hence, the second and third steps of the proof collapse the relevant domain of $b_t - \beta$ from the infinite set $\mathbb{R}^m \setminus B_{2\sqrt{m}\epsilon}(0)$ to the finite set $\{k\epsilon\omega_j^{\beta} : k \in \{-1,1\} \times j \in [m]\}$. The fourth step of the proof uses our shadow price convergence results to show that $n\Lambda_n^{\beta}(y_n^{\beta}) - \max_{k\in\{-1,1\}} \max_{j\in[m]} \bar{\Lambda}_t^{k\epsilon\omega_j^{\beta}}$ is usually very large, and the last step combines this with the previous results to establish that $\xi = b_t - \beta$ must rarely fall outside of $B_{2\sqrt{m}\epsilon}(0)$.

To begin the proof, note that Lemmas 1 and 2 imply that $\Lambda_\infty^b(y_\infty^b)$ is concave in b, since $\frac{\partial^2}{\partial b^2}\Lambda_\infty^b(y_\infty^b) = \frac{\partial}{\partial b} y_\infty^b = -\ddot{\Lambda}_\infty(y_\infty^b)^{-1}$ is negative definite. This concavity implies that we can restrict attention to small ξ vectors, since $\bar{\Lambda}_\infty^{\xi}$ decreases in the magnitude of ξ. But, more importantly, the concavity implies line (55), as I will now show.

Let $i \in [m]$ denote the index of the largest element of ξ, so that either $e_i'\xi = \|\xi\|_\infty$ or $-e_i'\xi = \|\xi\|_\infty$. Since the minus sign doesn't meaningfully affect the analysis, I will henceforth suppose $e_i'\xi = \|\xi\|_\infty \equiv \gamma$, in which case

$$\bar{\Lambda}_\infty^{\xi} \le \sup_{\{\zeta\in\mathbb{R}^m \,:\, e_i'\zeta=\gamma\}} \bar{\Lambda}_\infty^{\zeta}. \tag{57}$$

The solution to this optimization problem satisfies the following first-order conditions for some Lagrange multiplier λ:

$$\begin{aligned} 0 =& \tfrac{\partial}{\partial\zeta}\Big(t\Lambda_\infty^{\beta+\zeta}(y_\infty^{\beta+\zeta}) + (n-t)\Lambda_\infty^{\beta-\frac{t}{n-t}\zeta}(y_\infty^{\beta-\frac{t}{n-t}\zeta}) - \lambda(e_i'\zeta-\gamma)\Big) \\ =& t(y_\infty^{\beta+\zeta} - y_\infty^{\beta-\frac{t}{n-t}\zeta}) - \lambda e_i. \end{aligned}$$

Now we'll use Lemma 2 to differentiate this with respect to γ:

$$\begin{aligned} 0 &= \tfrac{\partial}{\partial\gamma}0\big|_{\gamma=0} \\ &= \tfrac{\partial}{\partial\gamma}\Big(t(y_\infty^{\beta+\zeta} - y_\infty^{\beta-\frac{t}{n-t}\zeta}) - \lambda e_i\Big)\big|_{\gamma=0} \\ &= -\Big(t\ddot{\Lambda}_\infty(y_\infty^{\beta+\zeta})^{-1} + \frac{t^2}{n-t}\ddot{\Lambda}_\infty(y_\infty^{\beta-\frac{t}{n-t}\zeta})^{-1}\Big)\tfrac{\partial}{\partial\gamma}\zeta - e_i\tfrac{\partial}{\partial\gamma}\lambda\big|_{\gamma=0} \\ &= -\frac{nt}{n-t}\ddot{\Lambda}_\infty(y_\infty^{\beta})^{-1}\tfrac{\partial}{\partial\gamma}\zeta - e_i\tfrac{\partial}{\partial\gamma}\lambda\big|_{\gamma=0}. \end{aligned}$$

The last line holds because $\gamma = 0$ implies $\zeta = 0$. Combining the $e_i'\zeta = \gamma$ constraint with the expression above yields

$$-\frac{n-t}{nt} e_i' \ddot{\Lambda}_\infty(y_\infty^\beta) e_i \tfrac{\partial}{\partial\gamma}\lambda\big|_{\gamma=0} = \tfrac{\partial}{\partial\gamma} e_i'\zeta\big|_{\gamma=0} = \tfrac{\partial}{\partial\gamma}\gamma\big|_{\gamma=0} = 1,$$

which implies that

$$\tfrac{\partial}{\partial\gamma}\lambda\big|_{\gamma=0} = \frac{-nt}{(n-t)e_i'\ddot{\Lambda}_\infty(y_\infty^\beta)e_i}.$$

Further, since $\lambda = 0$ when $\gamma = 0$, by the concavity of $\Lambda_\infty^b(y_\infty^b)$ in b, it follows that for sufficiently small γ we have

$$\lambda \le \frac{-nt\gamma}{2(n-t)e_i'\ddot{\Lambda}_\infty(y_\infty^\beta)e_i}.$$

By definition, our Lagrange multiplier also satisfies $\frac{\partial}{\partial\gamma}\sup_{\{\zeta\in\mathbb{R}^m\,:\,e_i'\zeta=\gamma\}} \bar{\Lambda}_\infty^\zeta = \lambda$, which with the result above yields the following, for sufficiently small γ:

$$\begin{aligned}
n\Lambda_\infty^\beta(y_\infty^\beta) - &\sup_{\{\zeta\in\mathbb{R}^m\,:\,e_i'\zeta=\gamma\}} \bar{\Lambda}_\infty^\zeta \\
&= -\Big(\sup_{\{\zeta\in\mathbb{R}^m\,:\,e_i'\zeta=\gamma\}} \bar{\Lambda}_\infty^\zeta - \sup_{\{\zeta\in\mathbb{R}^m\,:\,e_i'\zeta=0\}} \bar{\Lambda}_\infty^\zeta\Big) \\
&= -\int_{g=0}^{\gamma} \tfrac{\partial}{\partial g} \sup_{\{\zeta\in\mathbb{R}^m\,:\,e_i'\zeta=g\}} \bar{\Lambda}_\infty^\zeta \, dg \\
&\ge -\int_{g=0}^{\gamma} \frac{-ntg}{2(n-t)e_i'\ddot{\Lambda}_\infty(y_\infty^\beta)e_i} dg \\
&= \frac{nt\gamma^2}{4(n-t)e_i'\ddot{\Lambda}_\infty(y_\infty^\beta)e_i} \\
&\ge \frac{t(\|\xi\|/\sqrt{m})^2}{4\max_{j\in[m]} e_j'\ddot{\Lambda}_\infty(y_\infty^\beta)e_j}.
\end{aligned}$$

Note, the first line above holds because the concavity of $\Lambda_\infty^b(y_\infty^b)$ in b implies that $n\Lambda_\infty^\beta(y_\infty^\beta) = \sup_{\{\zeta\in\mathbb{R}^m\,:\,e_i'\zeta=0\}} \bar{\Lambda}_\infty^\zeta$, and the last line holds because $\gamma = \|\xi\|_\infty \ge \|\xi\|/\sqrt{m}$. Finally, combining the result above with line (57) yields line (55).

Second, I will now bound the difference between $\bar{\Lambda}_t^\xi$ and $\bar{\Lambda}_t^\zeta$ in terms of $y_t^{\beta+\zeta}$ and $\underleftarrow{y}_t^{\beta-\frac{t}{n-t}\zeta}$, which will enable us to invoke our shadow price convergence results. To this end, first note that

$$\begin{aligned}
\bar{\Lambda}_t^{\xi} = \max_{x\in[0,1]^n} & \sum_{s=1}^{n} x_s u_s \\
\text{s. t.} \quad & \sum_{s=1}^{t} x_s a_s \leq t(\beta+\xi), \\
& \sum_{s=t+1}^{n} x_s a_s \leq (n-t)\big(\beta - \frac{t}{n-t}\xi\big).
\end{aligned}$$

Since this linear program is concave in its constraints, $\bar{\Lambda}_t^{\xi}$ must be concave in ξ. Accordingly, the $\bar{\Lambda}_t^{\xi}$ function lies below the hyperplane characterized by supergradient $\frac{\partial}{\partial\zeta}\bar{\Lambda}_t^{\zeta} \equiv t y_t^{\beta+\zeta}\frac{\partial}{\partial\zeta}(\beta+\zeta) + (n-t)\,\underleftarrow{y}_t^{\beta-\frac{t}{n-t}\zeta}\frac{\partial}{\partial\zeta}\big(\beta - \frac{t}{n-t}\zeta\big) = t(y_t^{\beta+\zeta} - \underleftarrow{y}_t^{\beta-\frac{t}{n-t}\zeta})$:

$$\bar{\Lambda}_t^{\xi} - \bar{\Lambda}_t^{\zeta} \leq t(\xi-\zeta)'\big(y_t^{\beta+\zeta} - \underleftarrow{y}_t^{\beta-\frac{t}{n-t}\zeta}\big). \tag{58}$$

Third, I will use the preceding inequality to establish that

$$\Pr\big(\sup_{\xi\notin B_{2\sqrt{m}\epsilon}(0)} \bar{\Lambda}_t^{\xi} \leq \max_{k\in\{-1,1\}}\max_{j\in[m]} \bar{\Lambda}_t^{k\epsilon\omega_j^{\beta}}\big) \geq 3/4, \tag{59}$$

for all sufficiently small $\epsilon > 0$ and large t. This bound is crucial, as it enables us to replace the infinite continuum of $\bar{\Lambda}_t^{\xi}$ values for all $\xi \notin B_{2\sqrt{m}\epsilon}(0)$, with the largest of the $2m$ values of $\bar{\Lambda}_t^{k\epsilon\omega_j^{\beta}}$. To begin, note that Lemma 2 yields the following, for $k \in \{-1,1\}$ and small $\epsilon > 0$:

$$y_\infty^{\beta+k\epsilon\omega_i} - y_\infty^{\beta} = -k\epsilon\ddot{\Lambda}_\infty(y_\infty^{\beta})^{-1}\omega_i^{\beta} + o(\epsilon) = -k\epsilon\omega_i^{\beta}/\sigma_i^{\beta} + o(\epsilon).$$

Combining this with (58) yields the following, for $t \leq n/2$:

$$\begin{aligned}
\bar{\Lambda}_t^{\xi} - \bar{\Lambda}_t^{k\epsilon\omega_i^{\beta}} &\leq t(\xi - k\epsilon\omega_i^{\beta})'\big(y_t^{\beta+k\epsilon\omega_i^{\beta}} - \underleftarrow{y}_t^{\beta-\frac{t}{n-t}k\epsilon\omega_i^{\beta}}\big) \\
&= t(\xi - k\epsilon\omega_i^{\beta})'\big(y_\infty^{\beta+k\epsilon\omega_i^{\beta}} - y_\infty^{\beta} - y_\infty^{\beta-\frac{t}{n-t}k\epsilon\omega_i^{\beta}} + y_\infty^{\beta}\big) \\
&\quad + t(\xi - k\epsilon\omega_i^{\beta})'(y_t^{\beta+k\epsilon\omega_i^{\beta}} - y_\infty^{\beta+k\epsilon\omega_i^{\beta}}) - t(\xi - k\epsilon\omega_i^{\beta})'(\underleftarrow{y}_t^{\beta-\frac{t}{n-t}k\epsilon\omega_i} - y_\infty^{\beta-\frac{t}{n-t}k\epsilon\omega_i}) \\
&= \frac{-ntk\epsilon}{n-t}(\xi - k\epsilon\omega_i^{\beta})'\omega_i^{\beta}/\sigma_i^{\beta} + t\|\xi - k\epsilon\omega_i^{\beta}\|(o(\epsilon) + O_p(t^{-1/2})),
\end{aligned} \tag{60}$$

where the last line holds because $\|y_t^{\beta+\epsilon\omega_i^{\beta}} - y_\infty^{\beta+\epsilon\omega_i^{\beta}}\|$ and $\|\underleftarrow{y}_t^{\beta-\frac{t}{n-t}\epsilon\omega_i} - y_\infty^{\beta-\frac{t}{n-t}\epsilon\omega_i}\|$ are $O_p(t^{-1/2})$

when $t \leq n/2$, by Proposition 4.

Now, to derive (59) from (60), let $\gamma_i \equiv \xi' \omega_i^\beta$, so that $\xi = \sum_{i=1}^m \gamma_i \omega_i^\beta$, and let $j = \arg\max_{i \in [m]} |\gamma_i|$ and $k = \text{sign}(\gamma_j)$, so that $k\gamma_j \geq \|\xi\|/\sqrt{m}$. Further, choose $\xi \notin B_{2\sqrt{m}\epsilon}$, in which case $\epsilon \leq \|\xi\|/(2\sqrt{m})$, and hence

$$\begin{aligned} k(\xi - k\epsilon\omega_j^\beta)'\omega_j^\beta &= k\gamma_j - \epsilon \geq \|\xi\|/(2\sqrt{m}) \\ \text{and} \quad \|\xi - k\epsilon\omega_i^\beta\| &\leq 2\|\xi\|. \end{aligned}$$

Finally, set ϵ small enough so that the $o(\epsilon)$ term in (60) is less than $\max_{i\in[m]} \epsilon/(16\sqrt{m}\sigma_i^\beta)$, and choose t large enough so that the $O_p(t^{-1/2})$ term is less than $\max_{i\in[m]} \epsilon/(16\sqrt{m}\sigma_i^\beta)$, with at least three-quarters probability. When this last event happens, the previous two inequalities and line (60) yield the following, for all $\xi \notin B_{2\sqrt{m}\epsilon}(0)$:

$$\begin{aligned} \bar{\Lambda}_t^\xi - \bar{\Lambda}_t^{k\epsilon\omega_j^\beta} &\leq \frac{-nt\epsilon}{n-t}\|\xi\|/(2\sqrt{m}\sigma_j^\beta) + 2t\|\xi\|(o(\epsilon) + O_p(t^{-1/2})) \\ &\leq -t\epsilon\|\xi\|/(2\sqrt{m}\sigma_j^\beta) + 2t\|\xi\|(\epsilon/(16\sqrt{m}\sigma_i^\beta) + \epsilon/(16\sqrt{m}\sigma_j^\beta)) \\ &\leq -t\epsilon\|\xi\|/(4\sqrt{m}\sigma_j^\beta) \\ &\leq 0. \end{aligned}$$

This establishes line (59).

Fourth, I will use (55) to show that

$$\Pr\left(n\Lambda_n^\beta(y_n^\beta) - \max_{k\in\{-1,1\}} \max_{j\in[m]} \bar{\Lambda}_t^{k\epsilon\omega_j^\beta} \geq n^{2/3}\right) \geq 3/4. \tag{61}$$

This expression implies that there is probably at least one combination of $k \in \{-1, 1\}$ and $j \in [m]$ for which allocating $t(\beta + k\epsilon\omega_j^\beta)$ units of inventory to the first t periods is very costly. And with (59), this will imply that there's a decent chance that allocating $t(\beta + \xi)$ units of inventory to the first t periods will be very costly, for *any* $\xi \notin B_{2\sqrt{m}\epsilon}(0)$.

To begin, a nasty series of triangle inequalities yields the following, for $\zeta \equiv k\epsilon\omega_j^\beta$:

$$
\begin{aligned}
n\Lambda_n^\beta(y_n^\beta) - \bar{\Lambda}_t^\zeta \geq\ & n\Lambda_\infty^\beta(y_\infty^\beta) - t\Lambda_\infty^{\beta+\zeta}(y_\infty^{\beta+\zeta}) + (n-t)\Lambda_\infty^{\beta-\frac{t}{n-t}\zeta}(y_\infty^{\beta-\frac{t}{n-t}\zeta}) \\
& - n|\Lambda_\infty^\beta(y_n^\beta) - \Lambda_\infty^\beta(y_\infty^\beta)| \\
& - t|\Lambda_\infty^{\beta+\zeta}(y_t^{\beta+\zeta}) - \Lambda_\infty^{\beta+\zeta}(y_\infty^{\beta+\zeta})| \\
& - (n-t)|\Lambda_\infty^{\beta-\frac{t}{n-t}\zeta}(\underleftarrow{y}_t^{\beta-\frac{t}{n-t}\zeta}) - \Lambda_\infty^{\beta-\frac{t}{n-t}\zeta}(y_\infty^{\beta-\frac{t}{n-t}\zeta})| \\
& - t\Big|\Lambda_t^\beta(y_n^\beta) - \Lambda_\infty^\beta(y_n^\beta) - \Lambda_t^{\beta+\zeta}(y_t^{\beta+\zeta}) + \Lambda_\infty^{\beta+\zeta}(y_t^{\beta+\zeta})\Big| \\
& - (n-t)\Big|\underleftarrow{\Lambda}_t^\beta(y_n^\beta) - \Lambda_\infty^\beta(y_n^\beta) - \underleftarrow{\Lambda}_t^{\beta-\frac{t}{n-t}\zeta}(\underleftarrow{y}_t^{\beta-\frac{t}{n-t}\zeta}) + \Lambda_\infty^{\beta-\frac{t}{n-t}\zeta}(\underleftarrow{y}_t^{\beta-\frac{t}{n-t}\zeta})\Big|.
\end{aligned}
$$

Each term on the right is bounded in probability: First, line (55) establishes that $n\Lambda_\infty^\beta(y_\infty^\beta) - t\Lambda_\infty^{\beta+\zeta}(y_\infty^{\beta+\zeta}) + (n-t)\Lambda_\infty^{\beta-\frac{t}{n-t}\zeta}(y_\infty^{\beta-\frac{t}{n-t}\zeta}) = \Omega(t)$. Second, since Λ_∞^β is differentiable and since $\|y_n^\beta - y_\infty^\beta\| = O_p(n^{-1/2})$, by Proposition 4, we have $n|\Lambda_\infty^\beta(y_n^\beta) - \Lambda_\infty^\beta(y_\infty^\beta)| = O_p(n^{1/2})$. Likewise, $t|\Lambda_\infty^{\beta+\zeta}(y_t^{\beta+\zeta}) - \Lambda_\infty^{\beta+\zeta}(y_\infty^{\beta+\zeta})|$ and $(n-t)|\Lambda_\infty^{\beta-\frac{t}{n-t}\zeta}(\underleftarrow{y}_t^{\beta-\frac{t}{n-t}\zeta}) - \Lambda_\infty^{\beta-\frac{t}{n-t}\zeta}(y_\infty^{\beta-\frac{t}{n-t}\zeta})|$ are $O_p(t^{1/2})$ and $O_p((n-t)^{1/2})$, respectfully. Finally, Proposition 4 and Lemma 15 imply that the last two terms are also $O_p(t^{1/2})$ and $O_p((n-t)^{1/2})$. Accordingly, we can set n large enough so that if $n^{3/4} \leq t \leq n/2$ then there is at least a 75% chance that (i) the $\Omega(t)$ term exceeds $2n^{2/3}$ and (ii) the sum of the $O_p(n^{1/2})$, $O_p(t^{1/2})$, and $O_p((n-t)^{1/2})$ terms are no more than $n^{2/3}$, for all combinations of $k \in \{-1,1\}$ and $j \in [m]$. And this establishes (61).

Finally, combining (54), (59), and (61) yields the following, for sufficiently small ϵ, sufficiently large n, and $n^{3/4} \leq t \leq n/2$:

$$
\begin{aligned}
&\Pr(\underline{R}_n \geq n^{2/3} : b_t \notin B_{2\sqrt{m}\epsilon}(\beta)) \\
&\quad \geq \Pr(n\Lambda_n^\beta(y_n^\beta) - \bar{\Lambda}_t^{b_t-\beta} \geq n^{2/3} : b_t \notin B_{2\sqrt{m}\epsilon}(\beta)) \\
&\quad \geq \Pr\big(n\Lambda_n^\beta(y_n^\beta) - \sup_{\xi \notin B_{2\sqrt{m}\epsilon}(0)} \bar{\Lambda}_t^\xi \geq n^{2/3}\big) \\
&\quad \geq \Pr\big(n\Lambda_n^\beta(y_n^\beta) - \max_{k\in\{-1,1\}} \max_{j\in[m]} \bar{\Lambda}_t^{k\epsilon\omega_j^\beta} \geq n^{2/3} \ \cap \sup_{\xi \notin B_{2\sqrt{m}\epsilon}(0)} \bar{\Lambda}_t^\xi \leq \max_{k\in\{-1,1\}} \max_{j\in[m]} \bar{\Lambda}_t^{k\epsilon\omega_j^\beta}\big) \\
&\quad \geq \Pr\big(n\Lambda_n^\beta(y_n^\beta) - \max_{k\in\{-1,1\}} \max_{j\in[m]} \bar{\Lambda}_t^{k\epsilon\omega_j^\beta} \geq n^{2/3}\big) + \Pr\big(\sup_{\xi \notin B_{2\sqrt{m}\epsilon}(0)} \bar{\Lambda}_t^\xi \leq \max_{k\in\{-1,1\}} \max_{j\in[m]} \bar{\Lambda}_t^{k\epsilon\omega_j^\beta}\big) - 1 \\
&\quad \geq 3/4 + 3/4 - 1 \\
&\quad = 1/2.
\end{aligned}
$$

And now, since $\mathrm{E}(\underline{R}_n) = O(\log n)$, by Theorem 1 and line (38), we have

$$\begin{aligned} O(\log n) &= \mathrm{E}(\underline{R}_n) \\ &\geq n^{2/3} \Pr(b_t \notin B_{2\sqrt{m}\epsilon}(\beta)) \Pr(\underline{R}_n \geq n^{2/3} : b_t \notin B_{2\sqrt{m}\epsilon}(\beta)) \\ &\geq n^{2/3} \Pr(b_t \notin B_{2\sqrt{m}\epsilon}(\beta))/2, \end{aligned}$$

which implies the result. $\square$

Lemma 8 Proof. I will show that there exists $C > 0$ that satisfies

$$\inf_{b \in B_{\delta/2}(\beta)} \mathrm{E}\left(\pi_t^b \Delta_t(y_{t-1}^{\psi_t^b(0)})^- + (1-\pi_t^b)\Delta_t(y_{t-1}^{\psi_t^b(\alpha)})^+\right) \geq Cm\sigma_m^\beta/(2t), \tag{62}$$

for all sufficiently large t. Combining this result with Lemma 7 yields the desired result:

$$\begin{aligned} \mathrm{E}(r_t) &= \mathrm{E}\left(\mathbb{1}\{b_t \in B_{\delta/2}(\beta)\}\left(\pi_t^{b_t}\Delta_t(y_{t-1}^{\psi_t^{b_t}(0)})^- + (1-\pi_t^{b_t})\Delta_t(y_{t-1}^{\psi_t^{b_t}(a_t)})^+\right)\right) \\ &= \mathrm{E}\left(\mathbb{1}\{b_t \in B_{\delta/2}(\beta)\}\, \mathrm{E}\left(\pi_t^b\Delta_t(y_{t-1}^{\psi_t^b(0)})^- + (1-\pi_t^b)\Delta_t(y_{t-1}^{\psi_t^b(a_t)})^+\right)\big|_{b=b_t}\right) \\ &\geq \Pr(b_t \in B_{\delta/2}(\beta)) \inf_{b \in B_{\delta/2}(\beta)} \mathrm{E}\left(\pi_t^b\Delta_t(y_{t-1}^{\psi_t^b(0)})^- + (1-\pi_t^b)\Delta_t(y_{t-1}^{\psi_t^b(\alpha)})^+\right) \\ &\geq (1-n^{-1/2})Cm\sigma_m^\beta/(2t), \end{aligned}$$

where the second line holds because b_t is independent of the random mapping $b \mapsto \pi_t^b\Delta_t(y_{t-1}^{\psi_t^b(0)})^- + (1-\pi_t^b)\Delta_t(y_{t-1}^{\psi_t^b(a_t)})^+$, and the last line holds because we're given a large t that satisfies $n^{3/4} \leq t \leq n/2$.

Let me briefly outline how we will establish line (62). First, Lemma 10 implies that there's a $O(1)$ chance that we underestimate the shadow price by at least $4\iota/\sqrt{t}$. I use this fact to establish that there exists some constant $C > 0$ that satisfies the following for sufficiently large t:

$$\begin{aligned} &\mathrm{E}\left((1-\pi_t^b)\Delta_t(y_{t-1}^{\psi_t^b(a_t)})^+\right) \\ &\quad \geq C\,\mathrm{E}\left((1-\pi_t^b)\left(\mathbb{1}\{\Delta_t(y_\infty^b - \iota/\sqrt{t}) > 0\} - \mathbb{1}\{\Delta_t(y_\infty^b + \iota/\sqrt{t}) > 0\}\right)a_t'\iota/\sqrt{t}\right). \end{aligned}$$

This lower bound looks nasty, but it's almost exactly in the form we need to apply our one remaining tool: Assumption 6. However, before applying this assumption, I must eliminate the pesky $1-\pi_t^b$ term. Fortunately, $\mathrm{E}\left(\pi_t^b\Delta_t(y_{t-1}^{\psi_t^b(0)})^-\right)$ honors the same bound, except with π_t^b replacing $1-\pi_t^b$, which means that $\mathrm{E}\left(\pi_t^b\Delta_t(y_{t-1}^{\psi_t^b(0)})^- + (1-\pi_t^b)\Delta_t(y_{t-1}^{\psi_t^b(\alpha)})^+\right)$ has a corresponding π_t^b-free bound, which makes it amenable to Assumption 6. Finally, the last part of the proof combines this assumption with the fundamental theorem of calculus and Lemma 1 to express this expectation as an integral over $\ddot{\Lambda}_\infty$.

To begin the proof, let t be large enough so that $b \in B_{\delta/2}(\beta)$ implies $\psi_t^b(a_t) \in B_\delta(\beta)$. In this case Lemma 10 establishes that there exists $C > 0$ that satisfies the following, for all $b \in B_{\delta/2}(\beta)$:

$$\begin{aligned}
&\Pr\big(\|\sqrt{t}(y_{t-1}^{\psi_t^b(a_t)} - y_\infty^{\psi_t^b(a_t)}) + 4\iota\| \leq 1\big) \\
&\quad \geq \Pr\big(\sup_{b \in B_\delta(\beta)} \|\sqrt{t}(y_{t-1}^b - y_\infty^b) + 4\iota\| \leq 1\big) \\
&\quad = \Pr\big(\sup_{b \in B_\delta(\beta)} \|\sqrt{t}(y_{t-1}^b - y_\infty^b) + 4\iota\| \leq 1\big) \\
&\quad > C.
\end{aligned}$$

Furthermore, $\|\sqrt{t}(y_{t-1}^{\psi_t^b(a_t)} - y_\infty^{\psi_t^b(a_t)}) + 4\iota\| \leq 1$ implies the following, when t is large:

$$\begin{aligned}
y_{t-1}^{\psi_t^b(a_t)} - y_\infty^b =& (y_{t-1}^{\psi_t^b(a_t)} - y_\infty^{\psi_t^b(a_t)}) + (y_\infty^{\psi_t^b(a_t)} - y_\infty^b) \\
&\leq -3\iota/\sqrt{t} + \iota/\sqrt{t} \\
&= -2\iota/\sqrt{t},
\end{aligned}$$

where the second line follows because $\|y_\infty^{\psi_t^b(a_t)} - y_\infty^b\| = o(1/\sqrt{t})$, by Assumption 4 and Lemma 2. Now combining the previous two results yields the following, for $b \in B_{\delta/2}(\beta)$ and t large:

$$\begin{aligned}
&\mathrm{E}\big((1-\pi_t^b)\Delta_t(y_{t-1}^{\psi_t^b(a_t)})^+\big) \\
&\quad \geq \mathrm{E}\big(\mathbb{1}\{\|\sqrt{t}(y_{t-1}^{\psi_t^b(a_t)} - y_\infty^{\psi_t^b(a_t)}) + 4\iota\| \leq 1\}(1-\pi_t^b)\Delta_t(y_\infty^b - 2\iota/\sqrt{t})^+\big) \\
&\quad \geq C\,\mathrm{E}\big((1-\pi_t^b)\Delta_t(y_\infty^b - 2\iota/\sqrt{t})^+\big) \\
&\quad \geq C\,\mathrm{E}\big((1-\pi_t^b)\mathbb{1}\{\Delta_t(y_\infty^b - \iota/\sqrt{t}) \geq 0\}\Delta_t(y_\infty^b - 2\iota/\sqrt{t})^+\big) \\
&\quad \geq C\,\mathrm{E}\big((1-\pi_t^b)\mathbb{1}\{\Delta_t(y_\infty^b - \iota/\sqrt{t}) \geq 0\}a_t'\iota/\sqrt{t}\big) \\
&\quad \geq C\,\mathrm{E}\big((1-\pi_t^b)\big(\mathbb{1}\{\Delta_t(y_\infty^b - \iota/\sqrt{t}) > 0\} - \mathbb{1}\{\Delta_t(y_\infty^b + \iota/\sqrt{t}) > 0\}\big)a_t'\iota/\sqrt{t}\big).
\end{aligned}$$

Note, the third line above holds because $y_{t-1}^{\psi_t^b(a_t)}$ is independent of Δ_t and π_t, and the fifth line holds because $\Delta_t(y_\infty^b - \iota/\sqrt{t}) \geq 0$ implies $u_t \geq a_t'y_\infty^b - a_t'\iota/\sqrt{t}$ and hence implies $\Delta_t(y_\infty^b - 2\iota/\sqrt{t})^+ \geq a_t'\iota/\sqrt{t}$.

Next, an analogous argument implies that we can set C small enough to satisfy the following, for $b \in B_{\delta/2}(\beta)$ and large t:

$$\begin{aligned}
&\mathrm{E}\left(\pi_t^b \Delta_t(y_{t-1}^{\psi_t^b(0)})^-\right)\\
&\quad \geq C\,\mathrm{E}\left(\pi_t^b\big(\mathbb{1}\{\Delta_t(y_\infty^b - \iota/\sqrt{t}) > 0\} - \mathbb{1}\{\Delta_t(y_\infty^b + \iota/\sqrt{t}) > 0\}\big)a_t'\iota/\sqrt{t}\right).
\end{aligned}$$

Finally, adding our two bounds establishes line (62):

$$\begin{aligned}
&\mathrm{E}\left(\pi_t^b \Delta_t(y_{t-1}^{\psi_t^b(0)})^- + (1-\pi_t^b)\Delta_t(y_{t-1}^{\psi_t^b(\alpha)})^+ : b \in B_{\delta/2}(\beta)\right)\\
&\quad = C\,\mathrm{E}\left(\big(\mathbb{1}\{\Delta_t(y_\infty^b - \iota/(2\sqrt{t})) > 0\} - \mathbb{1}\{\Delta_t(y_\infty^b + \iota/(2\sqrt{t})) > 0\}\big)a_t'\iota/\sqrt{t}\right)\\
&\quad = C/\sqrt{t}\int_{\gamma=-1}^{1} \iota' \tfrac{\partial}{\partial\gamma}\,\mathrm{E}\left(\mathbb{1}\{\Delta_1(y_\infty^b - \gamma\iota/(2\sqrt{t})) > 0\}a_1\right)d\gamma\\
&\quad = C/\sqrt{t}\int_{\gamma=-1}^{1} \iota'\ddot{\Lambda}_\infty(y_\infty^b - \gamma\iota/\sqrt{t})\iota/(2\sqrt{t})d\gamma\\
&\quad \geq C/\sqrt{t}\int_{\gamma=-1}^{1} \iota'\iota\sigma_m^\beta/(4\sqrt{t})d\gamma\\
&\quad \geq Cm\sigma_m^\beta/(2t).
\end{aligned}$$

The penultimate line above holds because the smallest singular value of $\ddot{\Lambda}_\infty(y_\infty^b - \gamma\iota/\sqrt{t})$ is at least half of the smallest singular value of $\ddot{\Lambda}_\infty(y_\infty^\beta)$, when b is near β and t is large. □

Lemma 9. *$(y_t^b - y)'\dot{\Lambda}_t^b(y) \leq 0$ for all $t \in \mathbb{N}$, $b \in \mathbb{R}_+^m$, and $y \in \mathbb{R}_+^m$.*

Proof. Since $\dot{\Lambda}_t^b$ is a subgradient, it satisfies $\Lambda_t^b(y_t^b) - \Lambda_t^b(y) \geq (y_t^b - y)'\dot{\Lambda}_t^b(y)$. And since $\Lambda_t^b(y_t^b) \leq \Lambda_t^b(y)$, this yields the result. □

Lemma 10. *For all $\gamma \in \mathbb{R}^m$ and $\epsilon > 0$ there exist $\delta, C > 0$ such that* $\Pr(\sup_{b\in B_\delta(\beta)} \|\sqrt{t}(y_t^b - y_\infty^b) - \gamma\| \leq \epsilon) \geq C$ *for all sufficiently large t.*

Proof. I will begin with a brief proof sketch. Our primary tool for positioning y_t^b is is Lemma 11, which maintains that y_t^b will be close to $y_\infty^b + \gamma/\sqrt{t}$ for all $b \in B_\delta(\beta)$ if $\dot{\Lambda}_t^b(y_\infty^b + (\gamma + \eta k\omega_j^b)/\sqrt{t})$ is close to $\eta k\sigma_j^b\omega_j^b$, for all $b \in B_\delta(\beta)$, $j \in [m]$, and $k \in \{-1, 1\}$. And with a few triangle inequalities and some basic calculus, I show that this condition holds when $\sqrt{t}(\dot{\Lambda}_t^\beta(y) - \dot{\Lambda}_\infty^\beta(y))$ is near $\ddot{\Lambda}_\infty(y_\infty^\beta)\gamma$ for all y in the ν-ball of y_∞^β, for some $\nu > 0$. Finally, I use Lemma 16 to show that there's an $\Theta(1)$ chance of this happening. This lemma maintains that the mapping $(j, y) \mapsto \sqrt{t}e_j'(\dot{\Lambda}_t^\beta(y) - \dot{\Lambda}_\infty^\beta(y))$ converges to a Gaussian process whose mean is near $\ddot{\Lambda}_\infty(y_\infty^\beta)\gamma$ when we condition on $\sqrt{t}(\dot{\Lambda}_t^\beta(y_\infty^\beta) - \dot{\Lambda}_\infty^\beta(y_\infty^\beta))$ being near $\ddot{\Lambda}_\infty(y_\infty^\beta)\gamma$.

Lemmas 1 and 2 imply that we can choose δ small enough so that $\sigma_1^b \leq 2\sigma_1^\beta$ and $\sigma_m^b \geq \sigma_m^\beta/2 > 0$ for all $b \in B_\delta(\beta)$. And these lemmas also imply that we can choose t large enough to ensure that $y_\infty^b + (\gamma + \eta k \omega_j^b)/\sqrt{t} \geq 0$ for a given $\eta > 0$ and all $j \in [m]$, $k \in \{-1, 1\}$, and $b \in B_\delta(\beta)$. With this, Lemma 11 indicates that $\sup_{b \in B_\delta(\beta)} \|\sqrt{t}(y_t^b - y_\infty^b) - \gamma\| \leq \epsilon$ if

$$\sup_{b \in B_\delta(\beta)} \max_{j \in [m]} \max_{k \in \{-1,1\}} \|\sqrt{t}\dot{\Lambda}_t^b(y_\infty^b + (\gamma + \eta k \omega_j^b)/\sqrt{t}) - \eta k \sigma_j^b \omega_j^b\| \leq \kappa, \tag{63}$$

where $\eta \equiv \epsilon/(1 + 8\sqrt{m}\sigma_1^\beta/\sigma_m^\beta)$ and $\kappa \equiv \eta\sigma_m^\beta/(4\sqrt{m})$. Further, this inequality holds when the following inequalities hold for all $b \in B_\delta(\beta)$, $j \in [m]$, and $k \in \{-1, 1\}$:

$$\|\sqrt{t}(\dot{\Lambda}_t^\beta(y_\infty^b + (\gamma + \eta k \omega_j^b)/\sqrt{t}) - \dot{\Lambda}_\infty^\beta(y_\infty^b + (\gamma + \eta k \omega_j^b)/\sqrt{t})) + \ddot{\Lambda}_\infty(y_\infty^\beta)\gamma\| \leq \kappa/3, \tag{64}$$

$$\|\sqrt{t}\dot{\Lambda}_t^b(y_\infty^b + (\gamma + \eta k \omega_j^b)/\sqrt{t}) - \ddot{\Lambda}_\infty(y_\infty^b)\gamma - \eta k \sigma_j^b \omega_j^b\| \leq \kappa/3, \tag{65}$$

$$\text{and} \quad \|\ddot{\Lambda}_\infty(y_\infty^b)\gamma - \ddot{\Lambda}_\infty(y_\infty^\beta)\gamma\| \leq \kappa/3. \tag{66}$$

Lines (64)–(66) imply line (63) by the triangle inequality, and by the fact that $\dot{\Lambda}_t^b(y) - \dot{\Lambda}_\infty^b(y)$ is independent of b, which enables me to change the superscripts in line (64) from b to β. I will now show that there's a non-negligible chance that these inequalities hold universally across b, j, and k in their respective domains.

First, since $B_\delta(\beta)$ is compact and $\ddot{\Lambda}_\infty(y_\infty^b)$ is continuous in b, by Lemmas 1 and 2, it follows that we can set δ small enough to make inequality (66) hold universally.

Second, since $\dot{\Lambda}_\infty^b(y_\infty^b) = 0$ and $\ddot{\Lambda}_\infty$ is locally continuous near y_∞^β, the mean value theorem indicates that there exists $\xi_{tjk}^b \in (0, 1)$ for which

$$\begin{aligned}
&\sqrt{t}\dot{\Lambda}_\infty^b(y_\infty^b + (\gamma + \eta k \omega_j^b)/\sqrt{t}) \\
&\qquad = \sqrt{t}(\dot{\Lambda}_\infty^b(y_\infty^b + (\gamma + \eta k \omega_j^b)/\sqrt{t}) - \dot{\Lambda}_\infty^b(y_\infty^b)) \\
&\qquad = \ddot{\Lambda}_\infty(y_\infty^b + \xi_{tjk}^b(\gamma + \eta k \omega_j^b)/\sqrt{t})(\gamma + \eta k \omega_j^b) \\
&\qquad = \ddot{\Lambda}_\infty(y_\infty^b)\gamma + \eta k \sigma_j^b \omega_j^b + \zeta_{tjk}^b(\gamma + \eta k \omega_j^b), \\
\text{where} \quad &\zeta_{tjk}^b \equiv \ddot{\Lambda}_\infty(y_\infty^b + \xi_{tjk}^b(\gamma + \eta k \omega_j^b)/\sqrt{t}) - \ddot{\Lambda}_\infty(y_\infty^b).
\end{aligned}$$

And the continuity of $\ddot{\Lambda}_\infty$ implies that we can set δ small enough so that

$$\sup_{b \in B_\delta(\beta)} \max_{j \in [m]} \max_{k \in \{-1,1\}} \|\zeta_{tjk}^b(\gamma + \eta k \omega_j^b)\| \leq \kappa/3,$$

for all sufficiently large t. Hence, inequality (65) will hold universally for all sufficiently large t and small δ.

Finally, I will show that for all sufficiently large t the probability that inequality (64) holds universally across $b \in B_\delta(\beta)$, $j \in [m]$, and $k \in \{-1, 1\}$ exceeds some $C > 0$. Since y^b_∞ and ω^b_j are continuous in b, by Lemmas 1 and 2, it will suffice to show that there exist $\nu > 0$ such that

$$\liminf_{t\to\infty} \Pr\Big(\sup_{y\in B_\nu(y^\beta_\infty)} \|\sqrt{t}(\dot\Lambda^\beta_t(y) - \dot\Lambda^\beta_\infty(y)) + \ddot\Lambda_\infty(y^\beta_\infty)\gamma\| \le \kappa/3\Big) > 0.$$

I will prove this inequality with Lemma 16, which with proposition 3.13 of Eaton (1983) implies that conditional on $\zeta_t \equiv \sqrt{t}(\dot\Lambda^\beta_t(y^\beta_\infty) - \dot\Lambda^\beta_\infty(y^\beta_\infty))$, the random map $(j, y) \mapsto \sqrt{t}e'_j(\dot\Lambda^\beta_t(y) - \dot\Lambda^\beta_\infty(y))$ weakly converges to a Gaussian process with domain $[m] \times B_\nu(y^\beta_\infty)$, mean function ρ^{ζ_t}, and covariance function Ξ, where

$$\begin{aligned}
\rho^{\zeta_t}_j(y) &\equiv e'_j\Omega(y, y^\beta_\infty)\Omega(y^\beta_\infty, y^\beta_\infty)^{-1}\zeta_t,\\
\Xi_{j\bar{j}}(y, \bar{y}) &\equiv e'_j\Omega(y, \bar{y})e_{\bar{j}} - e'_j\Omega(y, y^\beta_\infty)\Omega(y^\beta_\infty, y^\beta_\infty)^{-1}\Omega(y^\beta_\infty, \bar{y})e_{\bar{j}},\\
\text{and}\quad \Omega(y, \bar{y}) &\equiv \mathrm{E}(\mathbb{1}\{\Delta_1(y) > 0\}\mathbb{1}\{\Delta_1(\bar{y}) > 0\}a_1a'_1) - \mathrm{E}(\mathbb{1}\{\Delta_1(y) > 0\}a_1)\,\mathrm{E}(\mathbb{1}\{\Delta_1(\bar{y}) > 0\}a'_1).
\end{aligned}$$

Since Ξ is independent of ζ_t, the random map $(j, y) \mapsto \sqrt{t}e'_j(\dot\Lambda^\beta_t(y) - \dot\Lambda^\beta_\infty(y)) - e'_j\rho^{\zeta_t}_j(y)$ is assymptotically independent of ζ_t, and hence assymptotically independent of the random map $y \mapsto \rho^{\zeta_t}(y)$. Accordingly, for sufficiently large t, we have

$$\begin{aligned}
&\Pr\Big(\sup_{y\in B_\nu(y^\beta_\infty)} \|\sqrt{t}(\dot\Lambda^\beta_t(y) - \dot\Lambda^\beta_\infty(y)) + \ddot\Lambda_\infty(y^\beta_\infty)\gamma\| \le \kappa/3\Big)\\
&\quad\ge \Pr\Big(\sup_{y\in B_\nu(y^\beta_\infty)} \|\sqrt{t}(\dot\Lambda^\beta_t(y) - \dot\Lambda^\beta_\infty(y)) - \rho^{\zeta_t}(y)\| \le \kappa/6\\
&\qquad\qquad \cap \sup_{y\in B_\nu(y^\beta_\infty)} \|\rho^{\zeta_t}(y) - \ddot\Lambda_\infty(y^\beta_\infty)\gamma\| \le \kappa/6\Big)\\
&\quad\ge p^1_tp^2_t/2,\\
\text{where}\quad p^1_t &\equiv \Pr\big(\sup_{y\in B_\nu(y^\beta_\infty)} \|\sqrt{t}(\dot\Lambda^\beta_t(y) - \dot\Lambda^\beta_\infty(y)) - \rho^{\zeta_t}(y)\| \le \kappa/6\big)\\
\text{and}\quad p^2_t &\equiv \Pr\big(\sup_{y\in B_\nu(y^\beta_\infty)} \|\rho^{\zeta_t}(y) - \ddot\Lambda_\infty(y^\beta_\infty)\gamma\| \le \kappa/6\big).
\end{aligned}$$

I will now lower bound probability p^2_t. It is straightforward to confirm that $\|\zeta_t - \rho^{\zeta_t}(y)\| = O(\|y - y^\beta_\infty\|)O(\|\zeta_t\|)$, which implies that we can choose ν small enough so that $\|\zeta_t - \ddot\Lambda_\infty(y^\beta_\infty)\gamma\| \le$

$\kappa/12$ implies $\|\zeta_t - \rho^{\zeta_t}(y)\| \le \kappa/12$ for all $y \in B_\nu(y^\beta_\infty)$. And this, in turn, implies that

$$\begin{aligned} p^2_t &\ge \Pr\Big(\|\zeta_t - \ddot\Lambda_\infty(y^\beta_\infty)\gamma\| \le \kappa/12 \ \cap \sup_{y\in B_\nu(y^\beta_\infty)} \|\zeta_t - \rho^{\zeta_t}(y)\| \le \kappa/12\Big) \\ &= \Pr\Big(\|\zeta_t - \ddot\Lambda_\infty(y^\beta_\infty)\gamma\| \le \kappa/12\Big). \end{aligned}$$

Finally, the limit inferior of this last probability is strictly positive, as $t \to \infty$, because ζ_t converges to a multivariate normal with a full-rank covariance matrix, by Lemma 16.

I will now lower bound probability p^1_t. First, $\|\zeta_t - \rho^{\zeta_t}(y)\| = O(\|y - y^\beta_\infty\|)O(\|\zeta_t\|)$ implies that for a given $M > 0$ we can set ν small enough so that $\|\zeta_t\| \le M$ implies $\|\zeta_t - \rho^{\zeta_t}(y)\| \le \kappa/12$ for all $y \in B_\nu(y^\beta_\infty)$. And since ζ_t converges to a multivariate normal, we can choose M large enough so that the last equality below holds for all sufficiently large t:

$$\begin{aligned} &\Pr\big(\sup_{y\in B_\nu(y^\beta_\infty)} \|\sqrt{t}(\dot\Lambda^\beta_t(y) - \dot\Lambda^\beta_\infty(y)) - \rho^{\zeta_t}(y)\| \le \kappa/6\big) \\ &\quad \ge \Pr\big(\sup_{y\in B_\nu(y^\beta_\infty)} \|\sqrt{t}(\dot\Lambda^\beta_t(y) - \dot\Lambda^\beta_\infty(y)) - \zeta_t\| \le \kappa/12 \ \cap \sup_{y\in B_\nu(y^\beta_\infty)} \|\zeta_t - \rho^{\zeta_t}(y)\| \le \kappa/12\big) \\ &\quad \ge \Pr\big(\sup_{y\in B_\nu(y^\beta_\infty)} \|\sqrt{t}(\dot\Lambda^\beta_t(y) - \dot\Lambda^\beta_\infty(y)) - \zeta_t\| \le \kappa/12 \ \cap \ \|\zeta_t\| < M\big) \\ &\quad \ge \Pr\big(\sup_{y\in B_\nu(y^\beta_\infty)} \|\sqrt{t}(\dot\Lambda^\beta_t(y) - \dot\Lambda^\beta_\infty(y)) - \zeta_t\| \le \kappa/12\big)/2. \end{aligned}$$

Further, Lemma 14 implies that we can set ν small enough so that the last inequality in the expression below holds:

$$\begin{aligned} &\Pr\big(\sup_{y\in B_\nu(y^\beta_\infty)} \|\sqrt{t}(\dot\Lambda^\beta_t(y) - \dot\Lambda^\beta_\infty(y)) - \zeta_t\| > \kappa/12\big) \\ &\quad = \Pr\big(\sup_{y\in B_\nu(y^\beta_\infty)} \|\sqrt{t}(\dot\Lambda^\beta_t(y) - \dot\Lambda^\beta_\infty(y)) - \zeta_t\|^2 > \kappa^2/144\big) \\ &\quad \le \mathrm{E}\big(\sup_{y\in B_\nu(y^\beta_\infty)} \|\sqrt{t}(\dot\Lambda^\beta_t(y) - \dot\Lambda^\beta_\infty(y)) - \zeta_t\|^2\big)/(\kappa^2/144) \\ &\quad \le 1/2. \end{aligned}$$

Accordingly, we can set ν small enough so that $p^1_t \ge 1/4$ for all sufficiently large t. $\square$

Lemma 11. *If b is close enough to β to ensure that $\{\omega^b_i\}_{i\in[m]}$ and $\{\sigma^b_i\}_{i\in[m]}$ exist, and if $y \in \mathbb{R}^m_{>0}$ and $\eta > 0$ satisfy $y + \eta k\omega^b_j \ge 0$ and $\|\dot\Lambda^b_t(y + \eta k\omega^b_j) - \eta k\sigma^b_j\omega^b_j\| \le \eta\sigma^b_m/(2\sqrt{m})$ for all $j \in [m]$ and*

$k \in \{-1, 1\}$ *then* $y_t^b \in B_{\eta(1+2\sqrt{m}\sigma_1^b/\sigma_m^b)}(y)$.

Proof. Combining Lemma 9 with the hypotheses of the current lemma implies the following:

$$\begin{aligned}
0 &\geq (y_t^b - y - \eta k\omega_j^b)'\dot{\Lambda}_t^b(y + \eta k\omega_j^b) \\
&= (y_t^b - y - \eta k\omega_j^b)'\eta k\sigma_j^b\omega_j^b + (y_t^b - y - \eta k\omega_j^b)'(\dot{\Lambda}_t^b(y + \eta k\omega_j^b) - \eta k\sigma_j^b\omega_j^b) \\
&\geq \eta k\sigma_j^b(y_t^b - y)'\omega_j^b - \eta^2 k^2\sigma_j^b\omega_j^{b\prime}\omega_j^b - \|y_t^b - y - \eta k\omega_j^b\|\|\dot{\Lambda}_t^b(y + \eta k\omega_j^b) - \eta k\sigma_j^b\omega_j^b\| \\
&\geq \eta k\sigma_m^b(y_t^b - y)'\omega_j^b - \eta^2\sigma_1^b - (\|y_t^b - y\| + \eta)\eta\sigma_m^b/(2\sqrt{m}).
\end{aligned}$$

And since $\omega_1^b, \cdots, \omega_m^b$ are orthonormal, there must be at least one $j \in [m]$ and one $k \in \{-1, 1\}$ for which $k(y_t^b - y)'\omega_j^b \geq \|y_t^b - y\|/\sqrt{m}$. And thus, we must have

$$0 \geq \eta\sigma_m^b\|y_t^b - y\|/\sqrt{m} - \eta^2\sigma_1^b - \big(\|y_t^b - y\| + \eta\big)\eta\sigma_m^b/(2\sqrt{m}).$$

Finally, rearranging the terms yields the result. $\square$

Lemma 12. *There exists* $\epsilon > 0$ *such that if* $y, \bar{y} \in B_\epsilon(y_\infty^\beta)$ *then* $\mathrm{E}\big(\mathbb{1}\{\Delta_1(y) > 0\}\big)\Delta_1(\bar{y})^-) \leq 2\sigma_1^\beta\|\bar{y} - y\|^2$.

Proof. I will first consider the case in which $\bar{y} \geq y$. To begin, note that $\mathbb{1}\{\Delta_1(y + dy) > 0\} \neq \mathbb{1}\{\Delta_1(y) > 0\}$ implies that $u_1 = a_1'(y + O(dy))$, in which case $\Delta_1(\bar{y})^- = a_1'(\bar{y} - y + O(dy))$. And with this, Assumption 6 and Lemma 1 imply the following, for y near y_∞^β:

$$\begin{aligned}
&\mathrm{E}\left(\big(\mathbb{1}\{\Delta_1(y + dy) > 0\} - \mathbb{1}\{\Delta_1(y) > 0\}\big)\Delta_1(\bar{y})^-\right) \\
&\quad = \mathrm{E}\left(\big(\mathbb{1}\{\Delta_1(y + dy) > 0\} - \mathbb{1}\{\Delta_1(y) > 0\}\big)a_1\right)'(\bar{y} - y + O(dy)) \\
&\quad = \left(\tfrac{\partial}{\partial y}\mathrm{E}(\mathbb{1}\{\Delta_1(y) > 0\}a_1)dy + o(dy)\right)'(\bar{y} - y + O(dy)) \\
&\quad = (\bar{y} - y)'\tfrac{\partial}{\partial y}\mathrm{E}(\mathbb{1}\{\Delta_1(y) > 0\}a_1)dy + o(dy) \\
&\quad = (y - \bar{y})'\ddot{\Lambda}(y)dy + o(dy).
\end{aligned}$$

Accordingly, for y near y_∞^β we have $\frac{\partial}{\partial y}\mathrm{E}\big(\mathbb{1}\{\Delta_1(y) > 0\}\big)\Delta_1(\bar{y})^-) = (y - \bar{y})'\ddot{\Lambda}(y)$. And thus, for y

and $\bar{y}$ sufficiently close to y^{β}_{∞} we have

$$\begin{aligned}
&\mathrm{E}\big(\mathbb{1}\{\Delta_1(y) > 0\}\big)\Delta_1(\bar{y})^-\big) \\
&\quad = \mathrm{E}\big(\mathbb{1}\{\Delta_1(y) > 0\}\big)\Delta_1(\bar{y})^-\big) - \mathrm{E}\big(\mathbb{1}\{\Delta_1(\bar{y}) > 0\}\big)\Delta_1(\bar{y})^-\big) \\
&\quad = \int_{\gamma=0}^{1} \tfrac{\partial}{\partial\gamma}\,\mathrm{E}\big(\mathbb{1}\{\Delta_1(\bar{y} + \gamma(y-\bar{y})) > 0\}\Delta_1(\bar{y})^-\big)d\gamma \\
&\quad = \int_{\gamma=0}^{1} (y-\bar{y})'\ddot{\Lambda}(\bar{y} + \gamma(y-\bar{y}))(y-\bar{y})d\gamma \\
&\quad \leq \int_{\gamma=0}^{1} 2\sigma^{\beta}_1\|y-\bar{y}\|^2 d\gamma \\
&\quad = 2\sigma^{\beta}_1\|y-\bar{y}\|^2,
\end{aligned}$$

where the penultimate line holds because the largest singular value of $\ddot{\Lambda}(\bar{y} + \gamma(y - \bar{y}))$ is smaller than twice the largest singular value of y^{β}_{∞}, when y and $\bar{y}$ are sufficiently close to y^{β}_{∞}, by Lemma 1.

Now we can use what we've just established to prove the $\bar{y} \not\geq y$ case, since

$$\begin{aligned}
&\mathrm{E}\big(\mathbb{1}\{\Delta_1(y) > 0\}\big)\Delta_1(\bar{y})^-\big) \\
&\quad \leq \mathrm{E}\big(\mathbb{1}\{\Delta_1(y) > 0\}\big)\Delta_1(\bar{y} \vee y)^-\big) \\
&\quad \leq 2\sigma^{\beta}_1\|y - \bar{y} \vee y\|^2 \\
&\quad \leq 2\sigma^{\beta}_1\|y-\bar{y}\|^2.
\end{aligned}$$

□

Lemma 13. *There exists $C > 0$ such that $\mathrm{E}\left(\sup_{b\in\mathbb{R}^m_+}\sup_{y\in\mathbb{R}^m_+}\|\dot{\Lambda}^b_t(y) - \dot{\Lambda}^b_\infty(y)\|^2\right) \leq C/t$ for all $t \in \mathbb{N}$.*

Proof. I will show that the conditions of van der Vaart and Wellner's (1996) theorems 2.14.2 and 2.14.5 are satisfied. First, to translate the problem into van der Vaart and Wellner's format, note that

$$\begin{aligned}
e_j'(\dot{\Lambda}^b_t(y) - \dot{\Lambda}^b_\infty(y)) &= \sum_{s=1}^{t} \lambda^y_j(u_s, a_s)/t - \mathrm{E}(\lambda^y_j(u_1, a_1)), \\
\text{where}\quad \lambda^y_j(u_1, a_1) &\equiv \mathbb{1}\{\Delta_1(y) > 0\}e_j' a_1.
\end{aligned}$$

The λ^y_j functions lie under an upper envelope—since $|\lambda^y_j(u_1, a_1)| \leq \|\alpha\|$—and so it will suffice to show that van der Vaart and Wellner's (1996) bracketing integral is finite for the set $\{\lambda_j^y\}_{y\in\mathbb{R}_+^m, j\in[m]}$.

To streamline the argument, I will suppose that $a_1 > 0$, almost surely, and that the conditional distribution of u_1 given a_1 is characterized by density function f, which is bounded by some $M \in \mathbb{N}$. These assumptions are not necessary, but the argument is messy without them.

For a given $\nu > 0$, define m-dimensional grid $G \equiv \gamma\mathbb{Z}^m$, where $\gamma \equiv \nu/(M\|\alpha\|_1^2)$. Next, let $\ell(y) \equiv \max\{g \in G : g \leq y\}$ represent the largest gridpoint that's weakly less than $y \in \mathbb{R}_+^m$ and let $h(y) \equiv \min\{g \in G : g > \ell(y)\}$ represent the smallest gridpoint that's strictly larger than $\ell(y)$, so that $h(y) - \ell(y) = \gamma\iota$. Finally, define $U \equiv F_u^{-1}(1 - \nu/(2\|\alpha\|))$, $A \equiv F_a^{-1}(\nu/(2\|\alpha\|))$, and $Y \equiv U/A$, where F_u is the CDF of u_1 and F_a is the CDF of the smallest element of a_1 (which we've assumed to be larger than zero).

I will now show that the pair $(\mathbb{1}\{\|y\|_\infty \leq Y\}\lambda_j^{h(y)}, \lambda_j^{\ell(y\wedge Y)})$ is a ν-bracket that contains λ_j^y. First, if $\|y\|_\infty \leq Y$ then

$$
\begin{aligned}
\mathrm{E}\big((\lambda_j^{\ell(y\wedge Y)}&(u_1, a_1) - \mathbb{1}\{\|y\|_\infty \leq Y\}\lambda_j^{h(\bar{y})}(u_1, a_1))^2\big)^{1/2} \\
&\leq \|\alpha\| \Pr\big(u_1 \in (\ell(y)'a_1, \ell(y)'a_1 + \gamma\iota' a_1]\big) \\
&\leq \|\alpha\| \,\mathrm{E}\Big(\Pr\big(u_1 \in (\ell(y)'a_1, \ell(y)'a_1 + \gamma\|\alpha\|_1] : a_1\big)\Big) \\
&\leq \gamma M \|\alpha\|_1^2 \\
&= \nu.
\end{aligned}
$$

Next, if $e_i' y > Y$ then

$$
\begin{aligned}
\mathrm{E}\big((\lambda_j^{\ell(y\wedge Y)}&(u_1, a_1) - \mathbb{1}\{\|y\|_\infty \leq Y\}\lambda_j^{h(\bar{y})}(u_1, a_1))^2\big)^{1/2} \\
&\leq \|\alpha\| \Pr\big(u_1 > a_1'(y \wedge Y)\big) \\
&\leq \|\alpha\| \Pr(u_1 > Y a_1' e_i) \\
&\leq \|\alpha\| \big(\Pr(a_1' e_i \leq A) + \Pr(u_1 > YA)\big) \\
&\leq \nu.
\end{aligned}
$$

Finally, the set $\{(\mathbb{1}\{\|y\|_\infty \leq Y\}\lambda_j^{h(y)}, \lambda_j^{\ell(y\wedge Y)})\}_{y\in\mathbb{R}_+^m}$ has $N_\nu = (\lfloor Y/\gamma \rfloor + 1)^m$ elements. Note that $E(u_1) < \infty$ implies $U < 2\|\alpha\|/\nu$, for all sufficiently small ν. Hence, for small enough ν we have $N_\nu < (\lfloor 2\|\alpha\|/(\gamma\nu A)\rfloor + 1)^m \leq (\lfloor 2\|\alpha\|_1^3 M/(\nu^2 A)\rfloor + 1)^m \leq C/\nu^{2m}$, which implies that $\int_{\nu=0}^1 \sqrt{\log N_\nu}\, d\nu < \infty$. $\square$

Lemma 14. *For all $\delta > 0$ there exist $\epsilon > 0$ such that* $\mathrm{E}\left(\sup_{y \in B_\epsilon(y^\beta_\infty)} \|\dot{\Lambda}^b_t(y) - \dot{\Lambda}^b_\infty(y) - \dot{\Lambda}^b_t(y^\beta_\infty) + \dot{\Lambda}^b_\infty(y^\beta_\infty)\|^2\right) \leq$ δ/t *for all* $t \in \mathbb{N}$.

Proof. The proof is similar to the proof of Lemma 13, except with

$$\lambda^y_j(u_1, a_1) \equiv (\mathbb{1}\{\Delta_1(y) > 0\} - \mathbb{1}\{\Delta_1(y^\beta_\infty) > 0\})e_j' a_1.$$

Modifying the proof of Lemma 13 establishes that the bracketing integral of $\{\lambda^y_j\}_{y \in B_\epsilon(y^\beta_\infty), j \in [m]}$ is uniformly bounded in $\epsilon \in [0, 1]$. Further, λ^y_j is bounded by envelope function $\bar{\lambda}^y_j$, where

$$\begin{aligned} \bar{\lambda}^y_j(u_1, a_1) &\equiv (\mathbb{1}\{\Delta_1(\underline{y}) > 0\} - \mathbb{1}\{\Delta_1(\bar{y}) > 0\})e_j' a_1, \\ \underline{y} &\equiv y \wedge y^\beta_\infty, \\ \text{and} \quad \bar{y} &\equiv y \vee y^\beta_\infty. \end{aligned}$$

I will now show that the second moment of this envelope can be made arbitrarily small. First, Assumption 6 and Lemma 1 yield the following:

$$\begin{aligned} \mathrm{E}(\bar{\lambda}^y_j(u_1, a_1)^2) &= \mathrm{E}\left((\mathbb{1}\{\Delta_1(\underline{y}) > 0\} - \mathbb{1}\{\Delta_1(\bar{y}) > 0\})(e_j' a_1)^2\right) \\ &\leq \|\alpha\| e_j' \,\mathrm{E}\left((\mathbb{1}\{\Delta_1(\underline{y}) > 0\} - \mathbb{1}\{\Delta_1(\bar{y}) > 0\})a_1\right) \\ &= \|\alpha\| e_j' \int_{\gamma=0}^1 \frac{\partial}{\partial\gamma} \mathrm{E}\left(\mathbb{1}\{\Delta_1(\bar{y} + \gamma(\underline{y} - \bar{y})) > 0\} a_1\right) d\gamma \\ &= -\|\alpha\| e_j' \int_{\gamma=0}^1 \ddot{\Lambda}_\infty(\bar{y} + \gamma(\underline{y} - \bar{y}))(\underline{y} - \bar{y}) d\gamma \\ &= -\|\alpha\| e_j' \ddot{\Lambda}_\infty(\bar{y} + \bar{\gamma}(\underline{y} - \bar{y}))(\underline{y} - \bar{y}), \end{aligned}$$

for some $\bar{\gamma} \in [0, 1]$. And since we constrain $y \in B_\epsilon(y^\beta_\infty)$, it follows that $\ddot{\Lambda}_\infty(\bar{y} + \bar{\gamma}(\underline{y} - \bar{y})) \to \ddot{\Lambda}_\infty(y^\beta_\infty)$ and $(\underline{y} - \bar{y}) \to 0$ as $\epsilon \to 0$. Accordingly, $\mathrm{E}(\bar{\lambda}^y_j(u_1, a_1)^2) \to 0$ as $\epsilon \to 0$, which with theorems 2.14.2 and 2.14.5 of van der Vaart and Wellner (1996) establishes the result. □

Lemma 15. *For any compact $\Omega \subset \mathbb{R}^m_+$, there exists $C > 0$ such that*
$\mathrm{E}\left(\sup_{y, \bar{y} \in \Omega} |\Lambda^b_t(y) - \Lambda^b_\infty(y) - \Lambda^{\bar{b}}_t(\bar{y}) + \Lambda^{\bar{b}}_\infty(\bar{y})|^2\right) \leq C/t$ *for all* $t \in \mathbb{N}$, $b, \bar{b} \in \mathbb{R}^m_+$.

Proof. Like in the proofs of Lemmas 13 and 14, I will use theorems 2.14.2 and 2.14.5 of van der Vaart and Wellner's (1996). As before, I will cast the problem as an empirical process with a new set of functions:

$$\Lambda_t^b(y) - \Lambda_\infty^b(y) - \Lambda_t^{\bar{b}}(\bar{y}) + \Lambda_\infty^{\bar{b}}(\bar{y}) = \sum_{s=1}^t \lambda_y^{\bar{y}}(u_s, a_s)/t - \mathrm{E}(\lambda_y^{\bar{y}}(u_1, a_1)),$$
$$\text{where} \quad \lambda_y^{\bar{y}}(u_1, a_1) \equiv (u_1 - a_1' y)^+ - (u_1 - a_1' \bar{y})^+.$$

Note that $|\lambda_y^{\bar{y}}(u_1, a_1)| \leq \|\alpha\| \mathrm{diam}(\Omega) < \infty$. Accordingly, it will suffice to show that the bracketing integral of $\{\lambda_j^y\}_{y \in \Omega, \bar{y} \in \Omega}$ is finite.

To bound this bracketing integral, define m-dimensional grid $G \equiv \gamma \mathbb{Z}^m$, where $\gamma \equiv \nu/(4\|\alpha\|\|\iota\|)$. Next, let $\ell(y) \equiv \max\{g \in G : g \leq y\}$ represent the largest gridpoint that's weakly less than $y \in \mathbb{R}_+^m$ and let $h(y) \equiv \min\{g \in G : g > \ell(y)\}$ represent the smallest gridpoint that's strictly larger than $\ell(y)$. By design, we have

$$\begin{aligned}
\lambda_{\ell(y)}^{h(\bar{y})} - \lambda_{h(y)}^{\ell(\bar{y})} &\leq (u_1 - a_1'(y - \gamma\iota))^+ - (u_1 - a_1'(\bar{y} + \gamma\iota))^+ \\
&\quad - (u_1 - a_1'(y + \gamma\iota))^+ + (u_1 - a_1'(\bar{y} - \gamma\iota))^+ \\
&\leq 4\gamma\|\alpha\|\|\iota\| \\
&= \nu.
\end{aligned}$$

Accordingly, the pair $(\lambda_{h(y)}^{\ell(\bar{y})}, \lambda_{\ell(y)}^{h(\bar{y})})$ is a ν-bracket that contains $\lambda_y^{\bar{y}}$. Finally, there are only $O(\nu^{2m})$ such brackets; hence, the bracketing integral is finite. $\square$

Lemma 16. *For all sufficiently small $\epsilon > 0$, the random mapping $(j, y) \mapsto \sqrt{t} e_j' (\dot{\Lambda}_t^b(y) - \dot{\Lambda}_\infty^b(y))$, with $j \in [m]$ and $y \in R_+^m$, weakly converges, as $t \to \infty$, to a mean-zero Gaussian process with domain $[m] \times R_+^m$ and covariance function $\Sigma_{j\bar{j}}(y, \bar{y}) \equiv e_j' \,\mathrm{E}(\mathbb{1}\{\Delta_1(y) > 0\} \mathbb{1}\{\Delta_1(\bar{y}) > 0\} a_1 a_1') - \mathrm{E}(\mathbb{1}\{\Delta_1(y) > 0\} a_1) \,\mathrm{E}(\mathbb{1}\{\Delta_1(\bar{y}) > 0\} a_1') e_{\bar{j}}$.*

Proof. This is a direct application of theorem 2.3 of Kosorok, a classical empirical processes result. The proof of Lemma 13 establishes that the corresponding bracketing integral is finite. $\square$