## Logarithmic Regret in Multisecretary and Online Linear Programs with Continuous Valuations

Robert L. Bray

Kellogg School of Management, Northwestern University

April 15, 2024

#### Abstract

I use empirical processes to 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 secretary from n applicants that you interview sequentially. But there's a hitch: once you interview someone, you must decide whether or not to hire them before interviewing the next candidate. So you face an optimal stopping problem, with the objective being to maximize the expected capability of the secretary you hire or, equivalently, to minimize the expectation of your *regret*, the capability difference between the most competent applicant and the one 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 applicants and the  $n\beta$  applicants you hire. Arlotto and Gurvich made a striking discovery: If secretary valuations are *i.i.d.* random variables with finite support,  $\{v_1, \dots, v_k\}$ , then your expected regret is uniformly bounded across  $n \in \mathbb{N}$  and  $\beta \in [0, 1]$ .

In Section 3, I study the multisecretary problem with secretary valuations drawn from the continuum [0,1], rather than the finite set,  $\{v_1, \dots, v_k\}$ . Specifically, I show that the expected regret lies between  $(\beta/8)(1 - \beta/8)(\log(n)/2 - \log(6))$  and  $(\log(n+1) + 7)/8$  for all  $n \ge 2^{20}\beta^{-8}$  and  $\beta \in [0, 1/2]$  when valuations are *i.i.d.* uniform random variables, and and I derive mirror-image bounds for  $\beta \in [1/2, 1]$ .<sup>1</sup> Further, I show that the most obvious heuristic satisfies the upper regret bound.

In Section 4, I extend this  $\Theta(\log n)$  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  $u_t \ge 0$  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 by hemming in the shadow prices with empirical processes.

#### 2 Related Works

#### 2.1 Primary Antecedents

The online linear program I study in Section 4 is a multi-dimensional extension of Lucker's (1998) "zero-one knapsack problem," in which you successively decide whether to add objects with random valuations and volumes to your backpack. Lucker proves that the expected regret grows like  $\Theta(\log n)$  under the optimal policy, provided that the value-to-volume ratio distribution is sufficiently continuous. He establishes this bound with a 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 create these bounds is painstaking because he has to weaken them just so as the inventory level departs the initial resource endowment. Extending Lueker's value-function-bounding approach to higher dimensions would have been difficult, so I tackled the multi-resource version of his problem with Vera and Banerjee's (2019) *compensated coupling* scheme. Rather than construct multi-dimensional functional envelopes, this more modern approach adds up the *myopic regret* incurred over the inventory level's random walk.

Lueker's specification generalizes the continuous-valuation multisecretary model I study in Section 3, but it does not generalize the finite-valuation multisecretary model that Arlotto and Gurvich (2019) study. Indeed, Arlotto and Gurvich show that they can decrease the regret rate from  $\Theta(\log n)$  to O(1) by replacing Lueker's continuous-support secretary valuation distribution with an analogous finite-support distribution.

The mirror image of the multisecretary model is Arlotto and Xie's (2020) stochastic knapsack problem: the former has random valuations and fixed capacity consumption, and the latter has fixed valuations and random capacity consumption. Arlotto and Xie's model does not fit under Lueker's (1998) framework because it permits an unrestricted knapsack capacity—whereas Lueker make the backpack volume scale linearly with n, Arlotto and Xie set it to a free model parameter. They use this additional degree of freedom to show that Lueker's  $O(\log n)$  upper regret bound holds universally across initial backpack capacities. However, Arlotto and Xie (2020, p. 190) 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) extends the  $O(\log n)$  upper regret bound to the multivariate setting. 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  $\mathbb{R}^m$ . More importantly, they are the first to incorporate an unknown demand distribution. Specifically, they show that the expected regret is at most  $O(\log n \log \log n)$  when the agent starts without knowing the nature of demand. However, Li and Ye (2022) do not provide a corresponding lower regret bound, so when you compare their  $O(\log n \log \log n)$  bound with the prior  $\Theta(\log n)$  results, you can't help but wonder: Is revenue management with and without online 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 is  $\Omega(\log n)$  when it is known—i.e., that it is  $\Theta(\log n)$ , both with and without demand distribution foreknowledge. 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 across inventory levels. 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. 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 means 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 applicants 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.

#### 2.2 Contemporaneous Developments

I will now discuss the 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<sup>2</sup>) problems." The DRC<sup>2</sup> framework is similar to Vera et al.'s (2019) "online resource allocation" framework, except it can accommodate an infinite number of customer types. 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." Indeed, this is how I bound the online linear program's regret, although I learned the technique from Li and Ye (2022).

Next, Jiang and Zhang (2020) extend Arlotto and Xie's (2020) model to allow multiple servers. Specifically, they suppose that you must allocate each customer to one of m servers. They provide an  $O(\log n)$  upper bound, but like Li and Ye (2022) and Arlotto and Xie (2020), do not provide a corresponding lower bound. Neither Jiang and Zhang's multi-server problem nor Li and Ye's OLP problem (which I study) generalize the other. Jiang and Zhang's framework incorporates an additional decision—which server to route a customer to—but Li and Ye's framework incorporates online learning and permits a richer set of restrictions—constraining sales with a general linear program. Moreover, Jiang and Zhang do not make the initial resource endowment scale linear with n, as Li and Ye (2022) and I do.

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 the multisecretary problem's  $O(\log n)$  regret may not hold if the probability density function is near 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  $F^{-1}(1 - \beta + \Omega_p(\sqrt{n}))$ , 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  $F(u) = 1 - \beta + \operatorname{sign}(u - u^*)(u - u^*)^2/2$  and hence  $F^{-1}(q) = u^* + \operatorname{sign}(q - 1 + \beta)\sqrt{2|q - 1 + \beta|}$ . 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:  $F^{-1}(1-\beta+\Omega_p(\sqrt{n})) = u^* + \operatorname{sign}(\Omega_p(\sqrt{n}))\sqrt{2|\Omega_p(\sqrt{n})|} = u^* + \Omega_p(n^{-1/4})$ . 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. 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. But assuming extra wiggle room in the fluid model is unreasonable 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. Lucker (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.<sup>2</sup>) You interview the candidates sequentially, starting with the *n*th applicant and ending with the first applicant, so that the period number corresponds with the size of the remaining candidate pool, with period t - 1 succeeding period t. Interviewing the period-t applicant reveals the utility you would get from hiring them,  $u_t$ , a standard uniform random variable independent of the other candidates' utilities. After interviewing this candidate, you must hire them on the spot or reject them 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:

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

$$E(v_0^b) \equiv 0,$$
  
and  $\psi_t^b(a) \equiv \begin{cases} (tb-a)/(t-1) & t > 1, \\ 0 & t = 1. \end{cases}$ 

I will explain the logic underlying these equations after line (1). But first note that the  $\psi_t^b$  function maps the fraction of applicants you can hire from period t onwards, b, and your period-t hiring decision,  $x_t$ , to the fraction of applicants 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:

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

The  $x_t \leq tb$  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-tapplicant (i.e., set  $x_t = 1$ ) if you have a job opening (i.e.,  $1 \leq tb$ ) and if the total expected utility conditional on hiring them (i.e.,  $u_t + E(v_{t-1}^{\psi_t^b(1)}))$  exceeds the total expected utility conditional on rejecting them (i.e.,  $E(v_{t-1}^{\psi_t^b(0)}))$ .

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

$$V^b_t \equiv \sum_{s=1}^{tb} h^s_t,$$

and  $h_t^s$  is the sth highest value in  $\{u_t, \dots, u_1\}$ . 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:  $E(R_n) \leq (\log(n+1)+7)/8$ , 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:  $E(R_n) \ge (\beta/8)(1 - \beta/8)(\log(n)/2 - \log(6))$ , for all  $n \ge 2^{20}\beta^{-8}$  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  $\underbrace{(1/2+1)/2}_{\text{value of average hire}} \cdot \underbrace{500 \text{ million}}_{\text{number of hires}} = 375 \text{ million}$ 

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, 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 underlying Algorithm 1, suppose that you hire the period-t applicant with tb available positions if and only if their valuation exceeds  $\tau_t^b$ , where  $\{\tau_t^b \mid t \in [n], tb \in \{0, \dots, t\}\}$  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

$$b_n \equiv \beta$$
  
and  $b_{t-1} \equiv \psi_t^{b_t} (\mathbb{1}\{u_t > \tau_t^{b_t}\}).$ 

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

$$\hat{v}_t \equiv \mathbb{1}\{u_t > \tau_t^{b_t}\}u_t + \hat{v}_{t-1},$$
  
and  $\hat{v}_0 \equiv 0.$ 

Since the value under Algorithm 1 can't exceed the value under the optimal algorithm, we have  $E(\hat{v}_n) \leq E(v_n^\beta)$ , which implies that that

$$\mathbf{E}(\hat{R}_n) \ge \mathbf{E}(R_n),$$
  
where  $\hat{R}_t \equiv V_t^{b_t} - \hat{v}_t.$ 

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

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

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

Line (2) states that if there are tb open positions, then the value of increasing the size of the applicant pool from t-1 to t equals the option value of replacing the tbth most capable person, out of the first t-1 applicants, with the period-t applicant. 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 applicant out of these t-1 candidates.

#### Algorithm 1.

- 1. input n,  $\beta$ ,  $\{u_t\}_{t=1}^n$ , 2. initialize  $b_n \coloneqq \beta$

3. for t from n to 1 do  
(a) set 
$$x_t \coloneqq \mathbb{1}\{u_t > 1 - b_t\}$$
  
(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  $u_t \leq \tau_t^{b_t}$ , and hence that  $b_{t-1} = \psi_t^{b_t}(0)$  and  $\hat{v}_t = \hat{v}_{t-1}$ . In this case, (2) implies that

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

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

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

Combining these two recurrence relations inductively yields

$$\hat{R}_{n} = r_{n} + \hat{R}_{n-1} = \sum_{t=1}^{n} r_{t},$$
(4)  
where  $r_{t} \equiv \mathbb{1}\{u_{t} \leq \tau_{t}^{b_{t}}\}(u_{t} - h_{t-1}^{tb_{t}})^{+} + \mathbb{1}\{u_{t} > \tau_{t}^{b_{t}}\}(u_{t} - h_{t-1}^{tb_{t}})^{-}$ 

$$= (\mathbb{1}\{u_{t} > h_{t-1}^{tb_{t}}\} - \mathbb{1}\{u_{t} > \tau_{t}^{b_{t}}\})(u_{t} - h_{t-1}^{tb_{t}}).$$

In the expression above,  $r_t$  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  $r_t$ 

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  $E((\mathbb{1}\{u_t > h\} - \mathbb{1}\{u_t > \tau\})(u_t - h)) = h^2/2 - h\tau + \tau^2/2$ , for constants h and  $\tau$ . And to integrate over  $h_{t-1}^{tb}$ , we use the fact that this beta(t - tb, tb) random variable satisfies  $E(h_{t-1}^{tb}) = 1 - b$  and  $E(h_{t-1}^{tb})^2 = \frac{(1-b)+t(1-b)^2}{t+1}$ . These properties enable us to express the expected myopic regret in terms of  $b_t$  and  $\tau_t^{b_t}$ :

$$E(r_t) = E\left(E\left(\left(\mathbb{1}\left\{u_t > h_{t-1}^{tb_t}\right\} - \mathbb{1}\left\{u_t > \tau_t^{b_t}\right\}\right)\left(u_t - h_{t-1}^{tb_t}\right) \mid h_{t-1}^{tb_t} = h, \ b_t = b\right)\right)$$
  
=  $E\left(E\left(\left((h_{t-1}^{tb_t})^2/2 - h_{t-1}^{tb_t}\tau_t^{b_t} + (\tau_t^{b_t})^2/2\right) \mid b_t = b\right)\right)$   
=  $E\left(\frac{(1-b_t) + t(1-b_t)^2}{2(t+1)} - \tau_t^{b_t}(1-b_t) + (\tau_t^{b_t})^2/2\right).$  (5)

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

$$\mathbf{E}(r_t) = \frac{\mathbf{E}(b_t(1-b_t))}{2(t+1)}$$

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

$$E(R_n) \le E(\hat{R}_n)$$
  
=  $\sum_{t=1}^{n} E(r_t)$   
=  $\sum_{t=1}^{n} \frac{E(b_t(1-b_t))}{2(t+1)}$   
 $\le \sum_{t=1}^{n} \sup_{b \in (0,1)} \frac{b(1-b)}{2(t+1)}$   
=  $\sum_{t=1}^{n} \frac{1}{8(t+1)}$   
 $\le (\log(n+1)+7)/8$ 

#### 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

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

such that the optimal policy hires the period-t applicant 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

$$E(R_n) = \sum_{t=1}^{n} E\left(\frac{(1-b_t)+t(1-b_t)^2}{2(t+1)} - \tau_t^{b_t}(1-b_t) + (\tau_t^{b_t})^2/2\right)$$
  

$$\geq \sum_{t=1}^{n} E\left(\min_{\tau}\left(\frac{(1-b_t)+t(1-b_t)^2}{2(t+1)} - \tau(1-b_t) + \tau^2/2\right)\right)$$
  

$$= \sum_{t=1}^{n} \frac{E(b_t(1-b_t))}{2(t+1)}.$$
(7)

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

$$b_n \equiv \beta$$
  
and  $b_{t-1} \equiv \psi_t^{b_t} (\mathbb{1}\{u_t > \tau_t^{b_t}\}).$  (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  $n \ge 2^{20}\beta^{-8}$  and  $\sqrt{n} \le t \le n/2$  imply

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

and 
$$\Pr(b_t < \beta/8) \le 1/4.$$
 (10)

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

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

Accordingly, it will suffice to establish lines (9) and (10). I will begin with the former because it is more straightforward. Simply put, the  $\{b_t\}_{t=n}^1$  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

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

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

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

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

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

And with (11), we can rearrange these equations like this:

$$1 - \tilde{b}_n = \beta$$
  
and  $1 - \tilde{b}_{t-1} = 1 - \psi_t^{\tilde{b}_t} (\mathbbm{1}\{1 - u_t > \tau_t^{\tilde{b}_t}\})$ 
$$= 1 - \psi_t^{\tilde{b}_t} (\mathbbm{1}\{u_t < \tau_t^{1 - \tilde{b}_t}\})$$
$$= \frac{t - 1 - t\tilde{b}_t + \mathbbm{1}\{u_t < \tau_t^{1 - \tilde{b}_t}\}}{t - 1}$$
$$= \frac{t(1 - \tilde{b}_t) - \mathbbm{1}\{u_t \ge \tau_t^{1 - \tilde{b}_t}\}}{t - 1}$$
$$= \psi_t^{1 - \tilde{b}_t} (\mathbbm{1}\{u_t \ge \tau_t^{1 - \tilde{b}_t}\}).$$

Compare this system to (8), and you will see that  $1 - \tilde{b}_t = b_t$ , almost surely. Accordingly,  $\Pr(b_t > 1 - \beta/8) = \Pr(\tilde{b}_t < \beta/8) = \Pr(\hat{b}_t < \beta/8) \le \Pr(b_t < \beta/8)$ , 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  $\lfloor t\beta/8 \rfloor$  applicants hired before period t with the best  $\lfloor t\beta/8 \rfloor$  applicants 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 applicant interviewed after period t and the  $(n\beta - tb_t - \lfloor t\beta/8 \rfloor + 1)$ th best applicant interviewed before period 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 applicant interviewed after period 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  $tb_t$  out of the last t applicants

and hire the best  $n\beta - tb_t$  out of the first n - t applicants. Thus, the online value must satisfy

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

where look-back order statistic  $\underline{h}_{t}^{s}$  is the *s*th largest value in  $\{u_{n}, \dots, u_{t+1}\}$  (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  $tb_{t} + \lfloor t\beta/8 \rfloor$  out of the last *t* applicants and the best  $n\beta - tb_{t} - \lfloor t\beta/8 \rfloor$  out of the first n - t applicants. Hence, the offline value must satisfy the following when  $b_{t} < \beta/8$ :

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

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

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

The first line above states that your regret is at least as large as the benefit you'd get by replacing the worst  $\lfloor t\beta/8 \rfloor$  applicants hired before period t with the best  $\lfloor t\beta/8 \rfloor$  applicants 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 people exchanged) times the difference between  $h_t^{tb_t+\lfloor t\beta/8 \rfloor}$  (i.e., the value of the worst candidate added) and  $\underline{h}_t^{n\beta-tb_t-\lfloor t\beta/8 \rfloor+1}$  (i.e., the value of the best candidate 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  $t \ge 48 \log(12)/\beta$ ,  $n \ge 336 \log(12)/\beta$ , and  $\sqrt{n} \le t \le n/2$ then

$$\Pr(h_t^{\lfloor t\beta/4 \rfloor} \ge 1 - 3\beta/8) \ge 11/12$$
  
and 
$$\Pr(\underline{h}_t^{n\beta - \lfloor t\beta/4 \rfloor} \le 1 - 7\beta/8) \ge 11/12,$$

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

$$(\log(n+1)+7)/8$$

$$\geq \operatorname{E}(R_n)$$

$$\geq \lfloor t\beta^2/16 \rfloor \operatorname{Pr} \left( b_t < \beta/8 \cap h_t^{\lfloor t\beta/4 \rfloor} \geq 1 - 3\beta/8 \cap \underbrace{h}_t^{\lfloor (n-t)\beta \rfloor} \leq 1 - 7\beta/8 \right)$$

$$\geq \lfloor \sqrt{n\beta^2}/16 \rfloor \left( \operatorname{Pr}(b_t < \beta/8) + 11/12 + 11/12 - 2 \right).$$

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

## 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.<sup>3</sup> See the appendix for a notation guide and the online supplement for the proofs.

As before, I will count backwards from period  $n \in \mathbb{N}$  to period 1, positioning period t-1 after period t. 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 *j*th resource budgeted for the "average" remaining period, where  $e_j$  is the unit vector indicating the *j*th position. If you satisfy the period–ncustomer 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  $b_{n-1} \equiv (nb_n - a_n)/(n-1)$  (N.B., both  $nb_n$  and  $a_n$  can take non-integer values). 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  $\mu$ .

Assumption 2. The utilities and resource requirements are non-negative:  $u_1, a_1 \ge 0$  almost surely. Assumption 3. The utilities have finite expectation:  $E(u_1) < \infty$ . **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:

$$E(v_t^b) \equiv E\left(\max_{x_t \in \{0,1\}} x_t u_t + E(v_{t-1}^{\psi_t^b(x_t a_t)}) \quad \text{s.t.} \quad x_t a_t \le tb\right),\tag{12}$$
$$E(v_0^b) \equiv 0.$$

and 
$$\psi_t^b(a) \equiv \begin{cases} (tb-a)/(t-1) & t > 1, \\ 0 & t = 1. \end{cases}$$
 (13)

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

$$\pi_t^b \equiv \underset{x_t \in \{0,1\}}{\arg\max x_t u_t} + \mathbb{E}(v_{t-1}^{\psi_t^b(x_t a_t)}) \quad \text{s.t.} \quad x_t a_t \le tb.$$
(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.,  $a_t \leq tb$ ) and if the total expected utility conditional on satisfying this customer (i.e.,  $u_t + E(v_{t-1}^{\psi_t^b(1)}))$  exceeds the total expected utility conditional on turning them away (i.e.,  $E(v_{t-1}^{\psi_t^b(0)}))$ .

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

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

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

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

$$R_n \equiv V_n^\beta - v_n^\beta. \tag{17}$$

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

and

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

$$\bar{V}_{n}^{\beta} \geq V_{n}^{\beta},$$
where  $\bar{V}_{t}^{b} \equiv \max_{x \in [0,1]^{t}} \sum_{s=1}^{t} x_{s} u_{s}$  s.t.  $\sum_{s=1}^{t} x_{s} a_{s} \leq tb$ 
(18)

$$= \min_{y \in \mathbb{R}^{m}_{+}, w \in \mathbb{R}^{t}_{+}} tb'y + \sum_{s=1}^{\iota} w_{t} \quad \text{s.t.} \quad a'_{t}y + w_{t} \ge u_{t} \quad \forall \ t,$$
(19)

$$= \min_{y \in \mathbb{R}^m_+} t\Lambda^b_t(y), \tag{20}$$

$$\Lambda_t^b(y) \equiv b'y + \sum_{s=1}^t \Delta_s(y)^+ / t,$$
  
$$\Delta_t(y) \equiv u_t - a'_t y.$$
(21)

Line (18) is the linear programming relaxation of the integer program specified in line (16). Accordingly, whereas  $V_t^b$  is the objective value of the offline optimization problem that gives you resource bundle  $tb \in \mathbb{R}^m$  to allocate over t periods,  $\bar{V}_t^b$  is the objective of the analogous problem that allows you to satisfy a fraction of a customer's demand. Line (19) is the dual of the problem given in line (18), with y corresponding to the  $\sum_{s=1}^t x_s a_s \leq tb$  constraint and w corresponding to the  $x_t \leq 1$ constraints. Finally, we distill this dual linear program to the convex optimization problem given in line (20) by replacing  $w_t$  with its smallest possible value,  $(u_t - a'_t y)^+$ . (To remember that this problem is a dual, it helps to think of  $\Lambda$  as an upside-down V.)

The dual problem in (20) has a not-necessarily-unique shadow price minimizer:

$$y_t^b \in \operatorname*{arg\,min}_{y \in \mathbb{R}^m_+} t\Lambda_t^b(y). \tag{22}$$

Since we initialized  $b_n = \beta$ , the problem in (20) converges, as  $n \to \infty$ , to the following deterministic fluid limit:

$$\min_{y \in \mathbb{R}^m_+} \Lambda^{\beta}_{\infty}(y) \quad \text{where} \quad \Lambda^{b}_{\infty}(y) \equiv b'y + \mathcal{E}(\Delta_1(y)^+).$$
(23)

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

Assumption 5. All resource constraints bind in the fluid approximation: there exists  $y_{\infty}^{\beta} \in$ 

 $\arg\min_{y\in\mathbb{R}^m_+}\Lambda^\beta_\infty(y) \ such \ that \ y^\beta_\infty>0.$ 

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  $u_1 = a'_1 y^{\beta}_{\infty}$  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: there exists a neighborhood of  $y_{\infty}^{\beta}$  such that the Jacobian matrix  $\frac{\partial}{\partial y} \operatorname{E}(\mathbb{1}\{\Delta_{1}(y) > 0\}a_{1})$  exists, is full rank, and is continuous in y in this neighborhood.

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  $E(\mathbb{1}\{\Delta_1(y) > 0\}a_1)$  is the mean resource consumption rate when we satisfy all customers with positive surplus utility, under shadow price vector y. Accordingly, Jacobian matrix  $\frac{\partial}{\partial y} E(\mathbb{1}\{\Delta_1(y) > 0\}a_1)$  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  $(\frac{\partial}{\partial y} E(\mathbb{1}\{\Delta_1(y) > 0\}a_1))^{-1}e_i$  would marginally decrease the consumption of the *i*th 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  $g(u_1 \mid a_1)$ , which almost surely satisfies  $g(a'_1 y^{\beta}_{\infty} \mid a_1) > 0$ . Further, suppose that  $E(a_1 a'_1)$  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  $\ddot{\Lambda}_{\infty}(y) \equiv \frac{\partial^2}{\partial y^2} \Lambda^b_{\infty}(y) = -\frac{\partial}{\partial y} \mathbb{E}(\mathbb{1}\{\Delta_1(y) > 0\}a_1)$  exists, is positive definite (and hence full rank), and its elements are continuous in y in a neighborhood of  $y^{\beta}_{\infty}$ . 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  $\beta$ , then  $\Lambda^b_{\infty}$  has a unique minimizer,  $y^b_{\infty} > 0$ , which is continuously differentiable—and hence Lipschitz continuous—in b, with  $\frac{\partial}{\partial b}y^b_{\infty} = -\ddot{\Lambda}_{\infty}(y^b_{\infty})^{-1}$ .

Together, Lemmas 1 and 2 imply that  $\ddot{\Lambda}_{\infty}(y^b_{\infty})$ —the Hessian matrix of  $\Lambda^b_{\infty}$  at its minimum—is continuous in b in a neighborhood of  $\beta$ . Accordingly,  $\{\omega^b_i\}_{i\in[m]}$  and  $\{\sigma^b_i\}_{i\in[m]}$  are likewise continuous in b, where  $\omega^b_i$  is an eigenvector of  $\ddot{\Lambda}_{\infty}(y^b_{\infty})$  with eigenvalue  $\sigma^b_i$ . Further, since  $\ddot{\Lambda}_{\infty}(y^b_{\infty})$  is positive definite, we can take  $\{\omega^b_i\}_{i\in[m]}$  to be orthonormal and take  $\{\sigma^b_i\}_{i\in[m]}$  to be real numbers that satisfy  $\sigma^b_1 \geq \cdots \geq \sigma^b_m > 0$  (provided that b is sufficiently close to  $\beta$ ).

Lemma 1 also implies that

$$\dot{\Lambda}^{b}_{\infty}(y) \equiv \frac{\partial}{\partial y} \Lambda^{b}_{\infty}(y) = b - \mathcal{E}(\mathbb{1}\{\Delta_{1}(y) > 0\}a_{1})$$
(24)

exists and is continuous in y a neighborhood of  $y_{\infty}^{\beta}$ . Unfortunately, the finite analog,  $\Lambda_t^b$ , is not always differentiable, but when it is, its gradient equals subgradient

$$\dot{\Lambda}_{t}^{b}(y) \equiv b - \sum_{s=1}^{t} \mathbb{1}\{\Delta_{s}(y) > 0\}a_{s}/t.$$
(25)

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 distribution learning yields an expected regret that grows at no more than a log n asymptotic rate:  $E(R_n) = O(\log n)$  as  $n \to \infty$ when distribution  $\mu$  is known to the decision maker.

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

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

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

Since knowing  $\mu$  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 can 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, Example 2 illustrates that adding a second constraint to the multisecretary problem can reduce its expected regret from  $\Theta(\log n)$  to o(1), and Example 3 illustrates that adding a second constraint to two copies of Arlotto and Xie's (2020) stochastic knapsack problem can intertwine the problem instances in a manner that reduces their combined regrets from  $\Theta(\log n)$  to O(1). Hence, some constraints negate the log n regret rate; I must prove that all such negating constraints violate our assumptions.

**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 period-*t* applicant 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 encounter a stream of n identical items. Since they're all the same, a stochastic knapsack problem involving these n items would yield zero regret. Now, imagine that each item consists of two components: A and B, each valued at one dollar. Assume the volume of the A component is a uniform random variable, and the volume of the B component is one minus the volume of the A component, making it also a uniform random variable. Also, suppose you have two backpacks: bag A for storing A components and bag B for storing B components. If both backpacks have a capacity of n/4, then you will face A and B stochastic knapsack problems, both of which are expected to yield  $\Theta(\log n)$  regret, according to the results of Lueker (1998) and Arlotto and Xie (2020). Now, introduce a constraint that prohibits packing only one component

from an item. Compelling you to pack both components or neither effectively reverts the problem to the scenario in which all items have equal value. For example, an item with an attractive A component will have a commensurately unattractive B component. In this case, the regret will be proportional to the unused space in one backpack when the other is filled, a quantity that has O(1) expectation under Algorithm 2, per Corollary 4. Adding the restriction lowers the regret by replacing the component-level selection with item-level selection. While there's plenty of variation in component valuations, there's essentially no variation in item valuations due to the perfectly negative correlation between the components' volumes.

#### 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  $y_{\infty}^b$ . 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  $\Sigma^b \equiv \ddot{\Lambda}_{\infty}(y_{\infty}^b)^{-1} \operatorname{Cov}(\mathbb{1}\{\Delta_1(y_{\infty}^b) > 0\}a_1)\ddot{\Lambda}_{\infty}(y_{\infty}^b)^{-1}$  is full rank and continuous in  $b \in B_{\delta}(\beta)$ .

In the proposition above,  $B_{\delta}(\beta)$  denotes the ball of radius  $\delta$  about  $\beta$ . But don't dwell on these technical  $\delta$  balls; instead, direct your attention to this:  $\sqrt{t}(y_t^b - y_{\infty}^b) \xrightarrow{d} \mathcal{N}(0, \Sigma^b)$ . It's hard to believe, but it seems this basic fact—that the shadow prices of a stochastic linear program converge to a multivariate normal—was previously unknown.

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  $\beta$ . Unfortunately, this 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  $\mathbb{E}(\sup_{b \in B_{\delta}(\beta)} \|y_t^b - y_{\infty}^b\|^2) = O(1/t).$ 

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

**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 distribution  $\mu$  is unknown, in which case shadow prices and inventory vectors become tangled.<sup>4</sup> Proposition 4 is a stronger version of Li and Ye's (2022) first theorem, which states that  $E(\|y_t^b - y_\infty^b\|^2) = O((\log \log t)/t)$ . 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  $y_\infty^b$  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,  $\hat{y}_t^b \equiv (y_t^b + y_\infty^b)/2$ . But  $\hat{y}_t^b$  is difficult to work with, so I then bounded the expected value of  $\|\dot{\Lambda}^b_t(\hat{y}_t^b) - \dot{\Lambda}^b_{\infty}(\hat{y}_t^b)\|^2$  with the expected value of  $\sup_{y \in B_{2\epsilon}(y_\infty^\beta)} \|\dot{\Lambda}^b_t(y) - \dot{\Lambda}^b_{\infty}(y)\|^2$ . 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  $\sqrt{t}(y_t^b - y_t^b)$  is near  $\gamma \in \mathbb{R}^m$ if  $\sqrt{t}(\dot{\Lambda}_t^\beta(y) - \dot{\Lambda}_\infty^\beta(y))$  is near  $\ddot{\Lambda}_\infty(y_\infty^\beta)\gamma$  for all y in a neighborhood of  $y_\infty^\beta$ , and this latter condition holds because the mapping  $(j, y) \mapsto \sqrt{t}e'_j(\dot{\Lambda}_t^b(y) - \dot{\Lambda}_\infty^b(y))$  converges to a sufficiently well-behaved Gaussian process, indexed by y and j.

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

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

**Corollary 3.** There exist  $\delta, C > 0$  such that  $\Pr\left(\sup_{b \in B_{\delta}(\beta)} \|y_t^b - y_{\infty}^b\| > \epsilon\right) \le \exp(-C\epsilon^2 t)$  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  $\dot{\Lambda}_t^b(y_{\infty}^b)$ , Corollary 3 establishes a concentration of measure for random shadow price  $y_t^b$ . This latter result is far harder to prove because  $y_t^b$  is not a sum of *i.i.d.* random variables, unlike  $\dot{\Lambda}_t^b(y_{\infty}^b)$ . 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  $y_t^b$  and  $y_{\infty}^b$ . This box has random faces, so its walls don't meet at 90-degree angles. Still, 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*. More specifically, since  $y_t^b$  is a minimizer, it must satisfy subgradient constraint  $(y_t^b - y_{\infty}^b - \eta k \omega_j^b)' \dot{\Lambda}_t^b(y_{\infty}^b + \eta k \omega_j^b) \leq 0$ for all  $j \in [m]$  and  $k \in \{-1, 1\}$ . These inequalities position  $y_t^b$  in the intersection of 2m halfspaces. Unfortunately, these half-spaces are random, since  $\dot{\Lambda}_t^b$  is stochastic. But  $\dot{\Lambda}_t^b(y_{\infty}^b + \eta k \omega_j^b) \leq 0$ for all  $j \in [m]$  and  $k \in \{-1, 1\}$  resemble those that satisfy  $(y_t^b - y_{\infty}^b - \eta k \omega_j^b)' \dot{\Lambda}_t^b(y_{\infty}^b + \eta k \omega_j^b) \leq 0$ for all  $j \in [m]$  and  $k \in \{-1, 1\}$  resemble those that satisfy  $(y_t^b - y_{\infty}^b - \eta k \omega_j^b)' \dot{\Lambda}_t^b(y_{\infty}^b + \eta k \omega_j^b) \leq 0$   $j \in [m]$  and  $k \in \{-1, 1\}$ . And this latter set of points forms a perfect cube around  $y_{\infty}^{b}$ . Hence, our initial subgradient constraints situate  $y_{t}^{b}$  in a wonky cube about  $y_{\infty}^{b}$ , with off-kilter faces.

#### 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  $\beta$  for which our lemmas apply. Next, I will express the values obtained under Algorithm 2 and those obtained 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.,  $tb_t \ge a_t$ ) 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  $t > 1, b_t \ge \alpha/t$ , and  $b_t$  sufficiently close to  $\beta$ , we have

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

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  $B_{\delta}(\beta)$ :

$$\tau(\delta) \equiv \begin{cases} 0 & \{b_t \mid t \in [n]\} \subset B_{\delta}(\beta), \\ \max\{t \mid b_t \notin B_{\delta}(\beta)\} & \text{otherwise.} \end{cases}$$
(26)

Li and Ye proved that their algorithm yields  $E(\tau(\delta)) = O(\log n \log \log n)$ —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. As the following corollary explains, I managed to tighten it to O(1).

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

Algorithm 2. 1. input n,  $\beta$ ,  $\{u_t\}_{t=1}^n$ ,  $\{a_t\}_{t=1}^n$ ,  $\mu$ 2. initialize  $b_n \coloneqq \beta$ 3. for t from n to 1 do (a) set  $x_t \coloneqq \mathbb{1}\{\Delta_t(y_{\infty}^{b_t}) > 0\}\mathbb{1}\{tb_t \ge a_t\}$ (b) set  $b_{t-1} \coloneqq \psi_t^{b_t}(x_ta_t)$ 4. end for 5. output  $\{x_t\}_{t=1}^n$ 

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

$$\mathcal{E}(\hat{R}_n) \ge \mathcal{E}(R_n),\tag{27}$$

where  $\hat{R}_t \equiv \bar{V}_t^{b_t} - \hat{v}_t$  and  $\hat{v}_t$  is the value collected by Algorithm 2 after period t:

$$\hat{v}_t \equiv \mathbb{1}\{\Delta_t(y_{\infty}^{b_t}) > 0\} \mathbb{1}\{tb_t \ge a_t\} u_t + \hat{v}_{t-1},$$
and
$$\hat{v}_0 \equiv 0.$$
(28)

Accordingly, it will suffice to show that  $E(\hat{R}_n) = O(\log n)$ . As in the multisecretary case, I will bound this benchmark regret by decomposing it into a sum of myopic regrets.

Lemma 4. The benchmark regret under Algorithm 2 can be upper bounded by a sum of approximate

myopic regrets: there exists sufficiently small  $\delta > 0$  such that

$$\begin{split} \hat{R}_n &\leq \sum_{t=1}^n r_t, \\ where \quad r_t &\equiv \mathbb{1}\{b_t \notin B_{\delta/2}(\beta)\}\sum_{s=1}^t u_s \\ &+ \mathbb{1}\{b_t \in B_{\delta/2}(\beta)\}\mathbb{1}\{\Delta_t(y_{\infty}^{b_t}) > 0\}\Delta_t(y_{t-1}^{b_{t-1}})^- \\ &+ \mathbb{1}\{b_t \in B_{\delta/2}(\beta)\}\mathbb{1}\{\Delta_t(y_{\infty}^{b_t}) \leq 0\}\Delta_t(y_{t-1}^{b_{t-1}})^+ \end{split}$$

The indicator variables in the definition of  $r_t$  characterize whether or not Algorithm 2 specifies satisfying the period-t customer, and whether or not  $b_t$  lies in the  $(\delta/2)$ -ball of  $\beta$ . (I use the  $(\delta/2)$ -ball rather than the  $\delta$ -ball, because  $b_t \in B_{\delta/2}(\beta)$  implies  $b_{t-1} \in B_{\delta}(\beta)$  when t is large.)

Finally, combining the preceding lemma with the following lemma yields Theorem 1.

**Lemma 5.** The approximate period-t myopic regret under Algorithm 2 is O(1/t) in expectation: there exists C > 0 such that  $r_t \leq C/t$ .

To control the first term of the myopic regret, I use the fact that  $E(\sum_{s=1}^{t} u_s)$  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  $E(\mathbb{1}\{\Delta_t(y_{\infty}^{b_t}) > 0\}\Delta_t(y_{t-1}^{b_{t-1}})^-)$  in terms of  $E(\sup_{b \in B_{\delta}(\beta)} \|y_{t-1}^b - y_{\infty}^b\|^2)$ ,  $E(\sup_{b \in B_{\delta}(\beta)} \|y_{t-1}^b \notin B_{\epsilon}(y_{\infty}^b)\}\|y_{t-1}^b - y_{\infty}^b\|$ , 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. Finally, I control the third term in a similar manner.

#### 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  $\mu$ , whereas the latter uses *look-back shadow price*  $\underbrace{y}_t^{b_t}$ , which is an estimate of  $y_{\infty}^{b_t}$  given the data observed up until period t + 1. More specifically  $\underbrace{y}_t^{b_t}$  is a minimizer of the backwards-looking problem

$$\underline{A}_{t}^{b}(y) \equiv b'y + \sum_{s=t+1}^{n} \Delta_{s}(y)^{+}/(n-t).$$
<sup>(29)</sup>

Algorithm 3 incorporates learning, as shadow price estimate  $\underbrace{y}_{t}^{b_{t}}$  starts hopelessly crude and ends finely tuned.

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  $E(\mathbb{1}\{b_t \in B_{\delta}(\beta)\} \| \underbrace{y}_t^{b_t} - y_{\infty}^{b_t} \|^2) = O(1/(n - t))$ . (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  $b \mapsto \underbrace{y}_t^{b}$ .)

# Algorithm 3.

1. input n,  $\beta$ ,  $\{u_t\}_{t=1}^n$ ,  $\{a_t\}_{t=1}^n$ 2. initialize  $b_n \coloneqq \beta$ 3. for t from n to 1 do (a) set  $x_t \coloneqq \mathbb{1}\{\Delta_t(\underbrace{y}_t^{b_t}) > 0\}\mathbb{1}\{tb_t \ge a_t\}$ (b) set  $b_{t-1} \coloneqq \psi_t^{b_t}(x_ta_t)$ 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 control its trajectory for all but O(1) periods, as the following results establish.

**Lemma 6.** 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  $\beta$  is asymptotically independent of n under Algorithm 3:  $E(\tau(\delta)) = O(1)$  as  $n \to \infty$ , for all  $\delta > 0$ .

The critical insight underlying Lemma 6 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  $b \mapsto y_t^b$ . 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:  $b_{\tau(\delta/2)+1}$ ,  $\sum_{s=t}^{\tau(\delta/2)} b_s - E(b_s \mid b_{s+1})$ , and  $\sum_{s=t}^{\tau(\delta/2)} E(b_s \mid b_{s+1}) - b_{s+1}$ . By definition, the first part is within  $\delta/2$  of  $\beta$ . The second part follows a martingale and thus concentrates around zero. And the third part is small, provided that  $y_s^{b_s}$  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  $\beta$  for all  $s \in \{t+1, \cdots, \tau(\delta/2)+1\}$ . And thus, I can inductively establish the result:  $b_s$  being near  $\beta$  for  $s \in \{t+1, \cdots, \tau(\delta/2)+1\}$  implies that  $y_s^{b_s}$  is near  $y_{\infty}^{b_s}$  for  $s \in \{t+1, \cdots, \tau(\delta/2)+1\}$ , which implies that  $\sum_{s=t}^{\tau(\delta/2)} E(b_s \mid b_{s+1}) - b_{s+1}$  is small, which implies that  $b_t$  is near  $\beta$ .

Having reigned in our inventory vectors, we are now ready to decompose regret benchmark

$$\hat{R}_t \equiv \bar{V}_t^{b_t} - \hat{v}_t, \tag{30}$$

where  $\hat{v}_t$  now denotes the value collected after period t under Algorithm 3:

$$\hat{v}_t \equiv \mathbb{1}\{\Delta_t(\underbrace{y}_t^{b_t}) > 0\} \mathbb{1}\{tb_t \ge a_t\} u_t + \hat{v}_{t-1},$$
and
$$\hat{v}_0 \equiv 0.$$
(31)

**Lemma 7.** The benchmark regret under Algorithm 3 can be upper bounded by a sum of approximate myopic regrets: there exists sufficiently small  $\delta > 0$  such that

$$\begin{split} \hat{R}_n &\leq \sum_{t=1}^n r_t, \\ where \quad r_t &\equiv \mathbb{1}\{b_t \notin B_{\delta/2}(\beta)\}\sum_{s=1}^t u_s \\ &+ \mathbb{1}\{b_t \in B_{\delta/2}(\beta)\}\mathbb{1}\{\Delta_t(\underbrace{y}_t^{b_t}) > 0\}\Delta_t(y_{t-1}^{b_{t-1}})^- \\ &+ \mathbb{1}\{b_t \in B_{\delta/2}(\beta)\}\mathbb{1}\{\Delta_t(\underbrace{y}_t^{b_t}) \leq 0\}\Delta_t(y_{t-1}^{b_{t-1}})^+ \end{split}$$

Combining the preceding lemma with the following lemma yields Theorem 2.

**Lemma 8.** The approximate period-t myopic regret under Algorithm 3 is O(1/t) + O(1/(n-t)) in expectation: there exists C > 0 such that  $E(r_t) \le C/t + C/(n-t)$ , for all  $n \in \mathbb{N}$  and  $t \le n$ .

This lemma is the same as Lemma 5, 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  $y_t^{b_t}$  and  $y_{\infty}^{b_t}$  contribute to your regret.

#### 4.5 Lower Bound with Known Distribution

We will now prove Theorem 3. To reiterate, the results of Section 3.3 do not make this analysis redundant: whereas we previously established  $\Omega(\log n)$  regret for one specific OLP—the multisecretary problem—we now establish  $\Omega(\log n)$  regret for all OLPs. In other words, Section 3.3 illustrates that the expected regret *can* grow like  $\Omega(\log n)$ , and this section proves that it *must* grow like  $\Omega(\log n)$  (see the discussion at the end of Section 4.1).

We will establish a universal  $\Omega(\log n)$  regret rate by retooling the methodology developed in Section 4.3 for a lower bound. 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  $\beta$  under the optimal algorithm (as opposed to Algorithm 2); and the lower-bounding version of Lemma 5 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  $b_{t-1} = \psi_t^{b_t}(\pi_t^{b_t} a_t).$  (32)

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  $\beta$  for a substantial time interval.

**Lemma 9.** The inventory vector tends to lie near  $\beta$  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  $\beta$  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  $\beta$ . 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:

$$E(\hat{R}_n) = E(R_n) + O(1),$$
  
where  $\hat{R}_t \equiv \bar{V}_t^{b_t} - v_t^{b_t}.$  (33)

The first line above 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  $\beta$  as the linear program does from resource endowment  $\beta - m\alpha/n$ :  $V_n^{\beta} \ge \bar{V}_n^{\beta - m\alpha/n}$ . And since the shadow price decreases in the inventory level, this implies that  $V_n^{\beta} \ge \bar{V}_n^{\beta} - m\alpha' y_n^{\beta - m\alpha/n}$ , and hence that  $R_n \ge \bar{V}_n^{\beta} - v_n^{\beta} - m\alpha' y_n^{\beta - m\alpha/n}$ . Finally, Proposition 4 indicates that  $E(y_n^{\beta - m\alpha/n}) = O(1)$  as  $n \to \infty$ , which establishes the result.

As before, I will now bound the benchmark regret with a sum of myopic regrets.

**Lemma 10.** The benchmark regret under the optimal algorithm can be lower bounded by a sum of approximate myopic regrets: there exists sufficiently small  $\delta > 0$  such that

$$\begin{split} \hat{R}_n &\geq \sum_{t = \lceil 2 \| \alpha \| / \delta \rceil}^n r_t, \\ where \quad r_t \equiv \mathbb{1}\{b_t \in B_{\delta/2}(\beta)\} \left(\pi_t^{b_t} \Delta_t (y_{t-1}^{\psi_t^{b_t}(0)})^- + (1 - \pi_t^{b_t}) \Delta_t (y_{t-1}^{\psi_t^{b_t}(a_t)})^+\right) \end{split}$$

Combining the preceding lemma with the following lemma yields Theorem 3.

**Lemma 11.** The approximate period-t myopic regret under the optimal algorithm is  $\Omega(1/t)$  in expectation for most of the second half of the horizon: There exists C > 0 such that  $E(r_t) \ge C/t$ , for all sufficiently large t that satisfies  $n^{3/4} \le t \le n/2$ .

To establish this last result, I show that if  $b_t \in B_{\delta/2}(\beta)$ —which happens with high probability, by Lemma 9—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 you should have satisfied—and Type II errors—satisfying customers that you should have 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've already said everything I want to, so I'll use this space to recapitulate this work's primary contributions.

- I develop a new methodology for regulating the dual variables of an online linear program: use the convexity of the dual problem to bound the shadow prices in terms of the subgradient of the dual value function and then bound this subgradient by casting it as an empirical process (for which there are many established results). The technique is versatile —I used it five different times:
  - 1. In the proof of Proposition 3, I implicitly use the technique as I leverage an M-estimator result that establishes the asymptotic normality of the dual solution by casting the random dual objective function as an empirical process.
  - 2. In the proof of Proposition 4, I bound  $1\{y_t^b \notin B_{\epsilon}(y_{\infty}^b)\} \|y_t^b y_{\infty}^b\|^2$  with a fixed multiple of  $\|\dot{\Lambda}_t^b((y_t^b + y_{\infty}^b)/2) \dot{\Lambda}_{\infty}^b((y_t^b + y_{\infty}^b)/2)\|^2$ , which I bound with  $\sup_{y \in B_{\epsilon}(y_{\infty}^b)} \|\dot{\Lambda}_t^b(y) \dot{\Lambda}_{\infty}^b(y)\|^2$ , which in turn I bound with an empirical processes result.
  - 3. In the proof of Proposition 6, I bound the probability that  $\sup_{b\in B_{\delta}(\beta)} \|y_t^b y_{\infty}^b\|$  is large with the probability that  $\sup_{b\in B_{\delta}(\beta)} \|\dot{\Lambda}_t^b(y_{\infty}^b + \eta k\omega_j^b) - \dot{\Lambda}_{\infty}^b(y_{\infty}^b + \eta k\omega_j^b)\|$  is large, which I bound with the probability that  $\sup_{y\in B_{\nu}(y_{\infty}^\beta)} \|\dot{\Lambda}_t^\beta(y) - \dot{\Lambda}_{\infty}^\beta(y)\|$  is large, which in turn I bound with an empirical processes result.
  - 4. In the lemma that proves Proposition 5, I lower bound the probability that 
    $$\begin{split} \sup_{b \in B_{\delta}(\beta)} \|\sqrt{t}(y_t^b - y_{\infty}^b) - \gamma\| \text{ is small with the probability that} \\ \sup_{b \in B_{\delta}(\beta)} \|\sqrt{t}(\dot{\Lambda}_t^{\beta}(y_{\infty}^b + (\gamma + \eta k \omega_j^b)/\sqrt{t}) - \dot{\Lambda}_{\infty}^{\beta}(y_{\infty}^b + (\gamma + \eta k \omega_j^b)/\sqrt{t})) + \ddot{\Lambda}_{\infty}(y_{\infty}^{\beta})\gamma\| \text{ is small,} \\ \text{ which I bound with the expected value of } \sup_{y \in B_{\nu}(y_{\infty}^{\beta})} \|\dot{\Lambda}_t^{\beta}(y) - \dot{\Lambda}_{\infty}^{\beta}(y) - (\dot{\Lambda}_t^{\beta}(y_{\infty}^{\beta}) - \dot{\Lambda}_{\infty}^{\beta}(y_{\infty}^{\beta}))\|^2, \\ \text{ which in turn I bound with an empirical processes result.} \end{split}$$
  - 5. In the proof of Lemma 9, I lower bound the cost of an additional restriction that mandates  $b_t = \beta + \xi$ , with  $\left| \Lambda_t^{\beta}(y_n^{\beta}) - \Lambda_{\infty}^{\beta}(y_n^{\beta}) - \Lambda_t^{\beta+\zeta}(y_t^{\beta+\zeta}) + \Lambda_{\infty}^{\beta+\zeta}(y_t^{\beta+\zeta}) \right|$ , which I bound with  $\sup_{y,\bar{y}\in\Omega} |\Lambda_t^b(y) - \Lambda_{\infty}^b(y) - \Lambda_t^{\bar{b}}(\bar{y}) + \Lambda_{\infty}^{\bar{b}}(\bar{y})|$ , which in turn I bound with an empirical processes result.
- I use my new empirical processes methodology to precisely characterize the convergence of dual variable  $y_t^b$  to its deterministic limit,  $y_{\infty}^b$ . Specifically, I show that under weak conditions

1.  $\sqrt{t}(y_t^b - y_{\infty}^b)$  converges to a multivariate normal for all  $b \in B_{\delta}(\beta)$ ,

- 2.  $\mathbb{E}\left(\sup_{b\in B_{\delta}(\beta)} \mathbb{1}\left\{y_{t}^{b} \notin B_{\epsilon}(y_{\infty}^{b})\right\} \|y_{t}^{b} y_{\infty}^{b}\|^{p}\right)$  falls exponentially fast in t for all  $p \geq 0$ ,
- 3.  $\mathrm{E}(\inf_{b \in B_{\delta}(\beta)} \|y_t^b y_{\infty}^b\|^2) = \Omega(1/t)$ , and
- 4.  $\mathbb{E}(\sup_{b \in B_{\delta}(\beta)} \|y_t^b y_{\infty}^b\|^2) = O(1/t).$

I can't find any direct antecedents for the first three results in the literature, but the fourth one is a strengthened version of Li and Ye's (2022) finding that  $E(\|y_t^b - y_\infty^b\|^2) = O((\log \log t)/t)$ . In addition to removing the  $\log \log t$  wiggle room, my bound also holds uniformly across  $b \in B_{\delta}(\beta)$ . Crucially, I position the  $\sup_{b \in B_{\delta}(\beta)}$  inside the expectation, which makes the result especially strong. For example, I would not have been able to accommodate online learning had the E and  $\sup_{b \in B_{\delta}(\beta)}$  operators been commuted. Precisely, if the supremum preceded the expectation, then I could have controlled the variance of look-back shadow price  $\underbrace{y}_t^b$  for any fixed resource vector b, but I could not have done so for the realized  $b_t$ , as this random variable correlates with the random mapping  $b \mapsto y_t^b$ .

- I broaden the applicability of Vera and Banerjee's (2019) compensated coupling scheme by devising new bounds on the trajectory of the inventory random walk,  $\{b_t\}_{t=n}^1$ . My delicate assumptions hold only in the  $\delta$ -ball around  $\beta$ , so I can't bound the regret without first proving that  $b_t$  resides in  $B_{\delta}(\beta)$ , with high probability. For the upper bound with known demand distribution, I show that following the certainty-equivalent principle—i.e., estimating the unobserved shadow price,  $y_t^{b_t}$ , with its fluid approximation,  $y_{\infty}^{b_t}$ —constrains  $b_t$  to  $B_{\delta}(\beta)$  for all but last O(1) periods. For the upper bound with unknown demand distribution, I inductively prove that the error introduced by replacing deterministic shadow price estimate  $y_{\infty}^{b_t}$  falls fast enough to ensure an orderly  $\{b_t\}_{t=n}^1$  walk. Finally, for the lower bound, I argue that the regret conditional on  $b_t \notin B_{\delta}(\beta)$  must be at least as high as the cost of a  $b_t \notin B_{\delta}(\beta)$  constraint imposed on the offline problem. I then lower bound the cost of this constraint to upper bound the probability that  $b_t \notin B_{\delta}(\beta)$  under the optimal policy.
- I use my new control over shadow prices and inventory levels to extend Lueker's (1998)  $O(\log n)$  and  $\Omega(\log n)$  regret bounds to a multi-resource setting. Rather than use Lueker's approach, I performed this generalization with compensated coupling, as bounding the value function across its entire domain would have been infeasible in higher dimensions (at least for me).
- I tighten Li and Ye's (2022) regret bound for the online linear program (OLP) with online

learning from  $O(\log n \log \log n)$  to  $O(\log n)$ , and I provide a corresponding  $\Omega(\log n)$  lower bound. Hence, I show that incorporating online learning—i.e., the dynamic estimation of  $\mu$ , the joint distribution of utility  $u_t$ , and resource consumption  $a_t$ —does not position a revenue management problem in a new difficulty class.

## Acknowledgments

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

### Notes

<sup>1</sup> Switching from finite to continuous secretary 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  $\sum_{t=1}^{n} (1/\sqrt{t}) \cdot (1/\sqrt{t}) = \Theta(\log n)$ .

<sup>2</sup> 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  $n\beta$  applicants you hire to maximizing one minus the capability of the  $n(1 - \beta)$  applicants you reject. But one minus a uniform is also a uniform, so this mirror-image problem must yield mirror-image regrets.

<sup>3</sup> I make three minor changes to the online linear programming model: I impose additional non-negativity constraints,  $u_1, a_1 \ge 0$ , 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.) However, the third change is 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.

<sup>4</sup> After I posted the result, Jiang et al. (2022) independently developed a slightly weaker version of Proposition 4.

## List of Symbols

[x]	the set $\{1, \cdots, x\}$	
$x \wedge y$	vector with <i>i</i> th element $\min(x_i, y_i)$	
$x \vee y$	vector with <i>i</i> th element is $\max(x_i, y_i)$	
$x^+$	$\max(0,x)$	
$x^-$	$\max(0, -x)$	
$e_j$	unit vector indicating $j$ th position	
ι	vector of ones	
$\mathbb{I}\left\{ \right\}$	indicator function	
$B_{\delta}(b)$	open ball with radius $\delta$ about $b$	
m	number of resources to manage	
n	number of time periods	
t	generic time period	
b	generic inventory vector, defined as total holdings divided by total time	
$b_t$	period- $t$ inventory vector associated with given algorithm	Algorithms $2$ and $3$ and line $(32)$
$\beta$	initial inventory vector	
$ au(\delta)$	first time inventory vector leaves $B_{\delta}(\beta)$	line $(26)$
$u_t$	utility received by satisfying period- $t$ customer	
$a_t$	resources consumed by satisfying period- $t$ customer	
$\Delta_t$	surplus utility function	line $(21)$
$\mu$	joint distribution of $(u_t, a_t)$	Assumption 1
$\alpha$	upper bound on $a_t$	Assumption 4
$x_t$	period- $t$ decision variable	
$\psi_t^b$	function determining period- $(t-1)$ inventory vector	line (13)
$\pi^b_t$	optimal action	line (14)
$v_t^b$	online objective value	line (15)
$\hat{v}_t$	objective value associated with given algorithm	lines $(28)$ and $(31)$
$V_t^b$	offline objective value	line (16)
$\bar{V}_t^b$	offline objective value with linear programming relaxation	line (18)
$R_n$	regret	line (17)
$\hat{R}_t$	benchmark regret associated with given algorithm	lines $(27)$ , $(30)$ , and $(33)$
$r_t$	approximate myopic regret associated with given algorithm	Lemmas $4, 7, and 10$
$\Lambda^b_t$	dual objective	line (20)
$\underline{\Lambda}_t^b$	look-back dual objective	line (29)
$\dot{\Lambda}_t$	dual objective subgradient	line (25)
$\Lambda^b_\infty$	limiting dual objective	line (23)
$\dot{\Lambda}_\infty$	limiting dual gradient	line (24)
$\ddot{\Lambda}_\infty$	limiting dual Hessian	Lemma 1
$\omega_i^b$	ith orthonormal eigenvector of limiting dual Hessian	after Lemma 2
$\sigma^b_i$	ith largest eigenvalue of limiting dual Hessian	after Lemma 2
$y_t^b$	dual optimal solution	line (22)
$\underline{y}_t^b$	look-back dual optimal solution	before line (29)
$y^b_\infty$	limiting dual optimal solution	Assumption 5 and Lemma 2
y	generic dual solution	

## References

- 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.
- Jiang, Jiashuo, Jiawei Zhang. 2020. Online resource allocation with stochastic resource consumption. arXiv (March 2021).
- 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 70(5) 2948–2966.
- 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  $y_{\infty}^{\beta}$ . Hence,  $\Delta_1(y)^+$  is almost surely differentiable in y, which means that

$$\frac{\partial}{\partial y} \Lambda^b_{\infty}(y) = \frac{\partial}{\partial y} (b'y + \mathcal{E}(\Delta_1(y)^+))$$
$$= b + \mathcal{E} \left(\frac{\partial}{\partial y} \Delta_1(y)^+\right)$$
$$= b - \mathcal{E}(\mathbb{1}\{\Delta_1(y) > 0\}a_1).$$

Note we can commute the expectation and differentiation because  $a_1$  is bounded. Combining the derivative above with Assumption 6 and the convexity of  $\Lambda^b_{\infty}$  implies the result.

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

Corollary 1 Proof. This follows immediately from Theorem 3.

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

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

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

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

$$|m_y(u_1, a_1) - m_z(u_1, a_1)| = b'y + \Delta_1(y)^+ - b'z - \Delta_1(z)^+$$
  

$$\leq (||b|| + ||a_1||)||y - z||$$
  

$$= \dot{m}(a_1)||y - z||.$$

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

$$\begin{split} & \mathbb{E}\left(\left(m_{y}(u_{1},a_{1})-m_{y_{\infty}^{b}}(u_{1},a_{1})-\dot{m}_{\infty}(u_{1},a_{1})'(y-y_{\infty}^{b})\right)^{2}\right)\\ &=\mathbb{E}\left(\left(\Delta_{1}(y)^{+}-\Delta_{1}(y_{\infty}^{b})^{+}+\mathbb{I}\left\{\Delta_{1}(y_{\infty}^{b})>0\right\}a_{1}'(y-y_{\infty}^{b})\right)^{2}\right)\\ &=\mathbb{E}\left(\Delta_{1}(y)^{2}|\mathbb{I}\left\{\Delta_{1}(y)>0\right\}-\mathbb{I}\left\{\Delta_{1}(y_{\infty}^{b})>0\right\}|\right)\\ &\leq\mathbb{E}\left((a_{1}'y-a_{1}'y_{\infty}^{b})^{2}|\mathbb{I}\left\{\Delta_{1}(y)>0\right\}-\mathbb{I}\left\{\Delta_{1}(y_{\infty}^{b})>0\right\}|\right)\\ &=\|y-y_{\infty}^{b}\|^{2}\mathbb{E}\left(\|\mathbb{I}\left\{\Delta_{1}(y)>0\right\}a_{1}-\mathbb{I}\left\{\Delta_{1}(y_{\infty}^{b})>0\right\}a_{1}\|^{2}\right)\\ &\leq\|\alpha\|\|y-y_{\infty}^{b}\|^{2}\mathbb{E}\left(\mathbb{I}\left\{\Delta_{1}(y\wedge y_{\infty}^{b})>0\right\}a_{1}-\mathbb{I}\left\{\Delta_{1}(y\vee y_{\infty}^{b})>0\right\}a_{1}\right)\\ &\leq\|\alpha\|\|y-y_{\infty}^{b}\|^{2}O(\|y-y_{\infty}^{b}\|)\\ &=o(\|y-y_{\infty}^{b}\|). \end{split}$$

Finally, Proposition 4 establishes that  $||y_t^b - y_{\infty}^b|| \xrightarrow{p} 0$ .

Proposition 4 Proof. Since Proposition 6 establishes that  $\mathbb{E}(\sup_{b \in B_{\delta}(\beta)} \mathbb{1}\{y_t^b \notin B_{\epsilon}(y_{\infty}^b)\} \|y_t^b - y_{\infty}^b\|^2) = o(1/t)$ , it will suffice to show that  $\mathbb{E}(\sup_{b \in B_{\delta}(\beta)} \mathbb{1}\{y_t^b \in B_{\epsilon}(y_{\infty}^b)\} \|y_t^b - y_{\infty}^b\|^2) = O(1/t)$ , 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  $y_t^b - y_{\infty}^b$  in terms of the magnitude of  $\dot{\Lambda}^b_{\infty}(\hat{y}_t^b) - \dot{\Lambda}^b_t(\hat{y}_t^b)$ , where  $\hat{y}_t^b \equiv (y_t^b + y_{\infty}^b)/2$ . Since  $\hat{y}_t^b$  lies between the mini-

mizers of  $\Lambda_{\infty}^{b}$  and  $\Lambda_{t}^{b}$ , the vector  $\hat{y}_{t}^{b} - y_{\infty}^{b}$  projects positively onto gradient  $\dot{\Lambda}_{\infty}^{b}(\hat{y}_{t}^{b})$  and projects negatively onto subgradient  $\dot{\Lambda}_{t}^{b}(\hat{y}_{t}^{b})$ . I use this fact to show that  $(\hat{y}_{t}^{b} - y_{\infty}^{b})'(\dot{\Lambda}_{\infty}^{b}(\hat{y}_{t}^{b}) - \dot{\Lambda}_{t}^{b}(\hat{y}_{t}^{b}))$  is larger than some fixed multiple of  $\|\hat{y}_{t}^{b} - y_{\infty}^{b}\|^{2}$ , which indicates that  $\|\dot{\Lambda}_{\infty}^{b}(\hat{y}_{t}^{b}) - \dot{\Lambda}_{t}^{b}(\hat{y}_{t}^{b})\|$  is larger than some fixed multiple of  $\|y_{t}^{b} - y_{\infty}^{b}\|$ . This, in turn, implies that the expectation of the maximum of  $\|\dot{\Lambda}_{t}^{b}(y) - \dot{\Lambda}_{\infty}^{b}(y)\|^{2}$ , across y in some small ball of  $y_{\infty}^{\beta}$ , is larger than some fixed multiple of the expectation of  $\mathbbm{1}\{\|y_{t}^{b} - y_{\infty}^{b}\| \le \epsilon\}\|y_{t}^{b} - y_{\infty}^{b}\|^{2}$ . And bounding the expectation of the maximum of  $\|\dot{\Lambda}_{t}^{b}(y) - \dot{\Lambda}_{\infty}^{b}(y)\|^{2}$  is a classic empirical processes problem.

Now, let's get to the proof. First, Lemma 2 establishes that we can choose  $\delta$  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  $\sigma_m^b$  denote the smallest singular value of  $\ddot{\Lambda}_{\infty}(y_{\infty}^b)$ . Lemmas 1 and 2 imply that we can set  $\delta$  small enough so that for all  $b \in B_{\delta}(\beta)$  we have  $\sigma_m^b \ge \sigma_m^{\beta}/2$ , and hence

$$(\hat{y}_t^b - y_\infty^b)' \ddot{\Lambda}_\infty(y_\infty^b) (\hat{y}_t^b - y_\infty^b) \ge \sigma_m^\beta \|\hat{y}_t^b - y_\infty^b\|^2 / 2.$$

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

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

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

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

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

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

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 12, yields the following, for  $y_t^b \in B_{2\epsilon}(y_\infty^\beta)$ :

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

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

$$\|\dot{\Lambda}^b_{\infty}(\hat{y}^b_t) - \dot{\Lambda}^b_t(\hat{y}^b_t)\| \ge \sigma^{\beta}_m \|\hat{y}^b_t - y^b_{\infty}\|/4 = \sigma^{\beta}_m \|y^b_t - y^b_{\infty}\|/8$$

And thus, we have

$$\begin{split} & \operatorname{E}\Big(\sup_{b\in B_{\delta}(\beta)}\mathbbm{1}\{y_{t}^{b}\in B_{\epsilon}(y_{\infty}^{b})\}\|y_{t}^{b}-y_{\infty}^{b}\|^{2}\Big) \\ & \leq \operatorname{E}\Big(\sup_{b\in B_{\delta}(\beta)}\mathbbm{1}\{y_{t}^{b}\in B_{2\epsilon}(y_{\infty}^{\beta})\}\|y_{t}^{b}-y_{\infty}^{b}\|^{2}\Big) \\ & \leq (8/\sigma_{m}^{\beta})^{2}\operatorname{E}\Big(\sup_{b\in B_{\delta}(\beta)}\mathbbm{1}\{y_{t}^{b}\in B_{2\epsilon}(y_{\infty}^{\beta})\}\|\dot{\Lambda}_{t}^{b}(\hat{y}_{t}^{b})-\dot{\Lambda}_{\infty}^{b}(\hat{y}_{t}^{b})\|^{2}\Big) \\ & \leq (8/\sigma_{m}^{\beta})^{2}\operatorname{E}\Big(\sup_{b\in B_{\delta}(\beta)}\sup_{y\in B_{2\epsilon}(y_{\infty}^{\beta})}\|\dot{\Lambda}_{t}^{b}(y)-\dot{\Lambda}_{\infty}^{b}(y)\|^{2}\Big) \\ & = (8/\sigma_{m}^{\beta})^{2}\operatorname{E}\Big(\sup_{y\in B_{2\epsilon}(y_{\infty}^{\beta})}\|\dot{\Lambda}_{t}^{\beta}(y)-\dot{\Lambda}_{\infty}^{\beta}(y)\|^{2}\Big), \end{split}$$

where the last line holds because  $\dot{\Lambda}_t^b - \dot{\Lambda}_\infty^b$  is independent of b. Finally, Lemma 16 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 13.  $\hfill \Box$ 

Corollary 2 Proof. This follows from Proposition 4 and Lemma 13.

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

$$\Pr\left(\sup_{b\in B_{\delta}(\beta)}\|y_t^b - y_{\infty}^b\| > \epsilon\right) \le 4m^2 \exp(-C\epsilon^2 t),\tag{34}$$

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

$$\mathbb{E}\left(\sup_{b\in B_{\delta}(\beta)}\mathbb{1}\left\{y_{t}^{b}\notin B_{\gamma^{1/p}}(0)\right\}\left\|y_{t}^{b}\right\|^{p}\right)\leq \exp(-Ct),\tag{35}$$

for all sufficiently large t.

The p = 0 case follows immediately from the line (34). Deriving the p > 0 case from lines (34) and (35) will take a bit more work. To that end, choose  $\gamma$  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 (34) and (35) imply that we can choose C > 0 so that we have the following for all sufficiently small  $\epsilon$  and large t:

$$\begin{split} & \mathcal{E} \left( \sup_{b \in B_{\delta}(\beta)} \mathbbm{1}\{y_{t}^{b} \notin B_{\epsilon}(y_{\infty}^{b})\} \| y_{t}^{b} - y_{\infty}^{b} \|^{p} \right) \\ & \leq \mathcal{E} \left( \sup_{b \in B_{\delta}(\beta)} \mathbbm{1}\{y_{t}^{b} \notin B_{\epsilon}(y_{\infty}^{b})\} \mathbbm{1}\{y_{t}^{b} \notin B_{\gamma^{1/p}}(0)\} (\|y_{t}^{b}\| + \gamma^{1/p})^{p} \right) \\ & + \mathcal{E} \left( \sup_{b \in B_{\delta}(\beta)} \mathbbm{1}\{y_{t}^{b} \notin B_{\epsilon}(y_{\infty}^{b})\} \mathbbm{1}\{y_{t}^{b} \in B_{\gamma^{1/p}}(0)\} (\|y_{t}^{b}\| + \gamma^{1/p})^{p} \right) \\ & \leq \mathcal{E} \left( \sup_{b \in B_{\delta}(\beta)} \mathbbm{1}\{y_{t}^{b} \notin B_{\epsilon}(y_{\infty}^{b})\} \mathbbm{1}\{y_{t}^{b} \notin B_{\gamma^{1/p}}(0)\} 2^{p} \| y_{t}^{b} \|^{p} \right) \\ & + \mathcal{E} \left( \sup_{b \in B_{\delta}(\beta)} \mathbbm{1}\{y_{t}^{b} \notin B_{\epsilon}(y_{\infty}^{b})\} \mathbbm{1}\{y_{t}^{b} \in B_{\gamma^{1/p}}(0)\} 2^{p} \gamma \right) \\ & \leq 2^{p} \mathcal{E} \left( \sup_{b \in B_{\delta}(\beta)} \mathbbm{1}\{y_{t}^{b} \notin B_{\gamma^{1/p}}(0)\} \| y_{t}^{b} \|^{p} \right) + 2^{p} \gamma \mathcal{Pr} \left( \sup_{b \in B_{\delta}(\beta)} \| y_{t}^{b} - y_{\infty}^{b} \| > \epsilon \right) \\ & \leq 2^{p} \exp(-Ct) + 2^{p+2} \gamma m^{2} \exp(-C\epsilon^{2} t). \end{split}$$

The inequality above establishes the p > 0 case. Hence, proving lines (34) and (35) will complete the argument.

Before getting into the math, let me roughly sketch the proof of line (34). The key tool will be Lemma 14, which is our only means for positioning  $y_t^b$ . The lemma corresponds to a set of inequalities that describe a small box, which aligns roughly 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  $y_{\infty}^b$  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 (34), 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  $y_{\infty}^{b}$  and  $y_{\infty}^{b} + \eta k \omega_{j}^{b}$ , and small enough to ensure that  $\sigma_{1}^{b} \leq 2\sigma_{1}^{\beta}$ ,  $\sigma_{m}^{b} \geq \sigma_{m}^{\beta}/2 > 0$ , and  $y_{\infty}^{b} + \eta k \omega_{j}^{b} \geq 0$ , for all  $j \in [m]$ ,  $k \in \{-1,1\}, b \in B_{\delta}(\beta)$ , and  $\eta \equiv \epsilon/(1 + 8\sqrt{m}\sigma_{1}^{\beta}/\sigma_{m}^{\beta})$ . Now, with these conditions, we can use Lemma 14 to bound the left-hand side of (34) in terms of more amenable subgradients:

$$\Pr\left(\sup_{b\in B_{\delta}(\beta)} \|y_{t}^{b} - y_{\infty}^{b}\| > \epsilon\right) \\
\leq \Pr\left(\sup_{b\in B_{\delta}(\beta)} \|y_{t}^{b} - y_{\infty}^{b}\| > \eta(1 + 2\sqrt{m}\sigma_{1}^{b}/\sigma_{m}^{b})\right) \\
\leq \Pr\left(\sup_{b\in B_{\delta}(\beta)} \max_{j\in[m]} \max_{k\in\{-1,1\}} \|\dot{\Lambda}_{t}^{b}(y_{\infty}^{b} + \eta k\omega_{j}^{b}) - \eta k\sigma_{j}^{b}\omega_{j}^{b}\| - \eta\sigma_{m}^{b}/(2\sqrt{m}) > 0\right) \\
\leq \Pr\left(\sup_{b\in B_{\delta}(\beta)} \max_{j\in[m]} \max_{k\in\{-1,1\}} \|\dot{\Lambda}_{t}^{b}(y_{\infty}^{b} + \eta k\omega_{j}^{b}) - \eta k\sigma_{j}^{b}\omega_{j}^{b}\| - \eta\sigma_{m}^{\beta}/(4\sqrt{m}) > 0\right). \quad (36)$$

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

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

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

$$\begin{split} \|\dot{\Lambda}_{t}^{b}(y_{\infty}^{b}+\eta k\omega_{j}^{b})-\eta k\sigma_{j}^{b}\omega_{j}^{b}\|\\ &\leq \|\dot{\Lambda}_{t}^{b}(y_{\infty}^{b}+\eta k\omega_{j}^{b})-\dot{\Lambda}_{\infty}^{b}(y_{\infty}^{b}+\eta k\omega_{j}^{b})\|+\|\dot{\Lambda}_{\infty}^{b}(y_{\infty}^{b}+\eta k\omega_{j}^{b})-\eta k\sigma_{j}^{b}\omega_{j}^{b}\|\\ &\leq \|\dot{\Lambda}_{t}^{b}(y_{\infty}^{b}+\eta k\omega_{j}^{b})-\dot{\Lambda}_{\infty}^{b}(y_{\infty}^{b}+\eta k\omega_{j}^{b})\|+\eta \sigma_{m}^{\beta}/(8\sqrt{m}). \end{split}$$

And, finally, combining this with line (36) and the fact that  $\dot{\Lambda}_t^b - \dot{\Lambda}_t^b = \dot{\Lambda}_t^\beta - \dot{\Lambda}_t^\beta$  yields the following:

$$\begin{split} &\Pr\left(\sup_{b\in B_{\delta}(\beta)}\|y_{t}^{b}-y_{\infty}^{b}\|>\epsilon\right) \\ &\leq \Pr\left(\sup_{b\in B_{\delta}(\beta)}\max_{j\in[m]}\max_{k\in\{-1,1\}}\|\dot{\Lambda}_{t}^{b}(y_{\infty}^{b}+\eta k\omega_{j}^{b})-\eta k\sigma_{j}^{b}\omega_{j}^{b}\|>\eta\sigma_{m}^{\beta}/(4\sqrt{m})\right) \\ &\leq \Pr\left(\sup_{b\in B_{\delta}(\beta)}\max_{j\in[m]}\max_{k\in\{-1,1\}}\|\dot{\Lambda}_{t}^{b}(y_{\infty}^{b}+\eta k\omega_{j}^{b})-\dot{\Lambda}_{\infty}^{b}(y_{\infty}^{b}+\eta k\omega_{j}^{b})\|>\eta\sigma_{m}^{\beta}/(8\sqrt{m})\right) \\ &\leq \sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{k\in\{-1,1\}}\Pr\left(\sup_{b\in B_{\delta}(\beta)}|e_{i}'\dot{\Lambda}_{t}^{b}(y_{\infty}^{b}+\eta k\omega_{j}^{b})-e_{i}'\dot{\Lambda}_{\infty}^{b}(y_{\infty}^{b}+\eta k\omega_{j}^{b})|\geq \eta\sigma_{m}^{\beta}/(8m)\right) \\ &\leq \sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{k\in\{-1,1\}}\Pr\left(\sup_{y\in B_{\nu}(y_{\infty}^{\beta})}|e_{i}'\dot{\Lambda}_{t}^{\beta}(y)-e_{i}'\dot{\Lambda}_{\infty}^{\beta}(y)|\geq \eta\sigma_{m}^{\beta}/(8m)\right), \end{split}$$

where  $\nu > 0$  is a constant that's large enough to ensure that  $y^b_{\infty} + \eta k \omega^b_j \in B_{\nu}(y^{\beta}_{\infty})$  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 16 for confirmation of this theorem's hypothesis). Note that the  $\sup_{b \in B_{\delta}(\beta)}$  prevents us from bounding the probability in the penultimate line with a more standard concentration of measure result.

This establishes line (34), which establishes the p = 0 case. I will now prove line (35), 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  $e'_j \dot{\Lambda}^b_t(e_j \gamma^{1/p} / \sqrt{m})$  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  $e'_j \dot{\Lambda}^b_t(e_j \gamma^{1/p} / \sqrt{m})$  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  $e'_j y^b_t > \omega$  implies  $e'_j \dot{\Lambda}^b_t (\omega e_j) \leq 0$ , for  $\omega \in \mathbb{R}$ . To see this,

take  $e_j' y_t^b \ge \omega$  and  $\hat{y} \equiv y_t^b - e_j (e_j' y_t^b - \omega)/2$ , and apply Lemma 12:

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

Since  $e'_j y^b_t - \omega$  is positive, by assumption, it follows that  $e'_j \dot{\Lambda}^b_t(\omega e_j)$  must be non-positive.

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

$$\Pr\left(\sup_{b\in B_{\delta}(\beta)} \left\|y_{t}^{b}\right\|^{p} > \gamma\right) \leq \sum_{j=1}^{m} \Pr\left(\sup_{b\in B_{\delta}(\beta)} e_{j}' y_{t}^{b} > \gamma^{1/p} / \sqrt{m}\right)$$
$$\leq \sum_{j=1}^{m} \Pr\left(\sup_{b\in B_{\delta}(\beta)} e_{j}' \dot{\Lambda}_{t}^{b}(e_{j} \gamma^{1/p} / \sqrt{m}) \leq 0\right).$$

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:

$$\begin{split} \sup_{b \in B_{\delta}(\beta)} e'_{j} \dot{\Lambda}^{b}_{t}(e_{j} \gamma^{1/p} / \sqrt{m}) \\ &= \sup_{b \in B_{\delta}(\beta)} e'_{j} b - \sum_{s=1}^{t} \mathbb{1}\{u_{s} > a'_{s} e_{j} \gamma^{1/p} / \sqrt{m})\} e'_{j} a_{s} / t \\ &\geq \sup_{b \in B_{\delta}(\beta)} e'_{j} b - \sum_{s=1}^{t} \left(\mathbb{1}\{e'_{j} a_{s} \le e'_{j} b / 2\}(e'_{j} b / 2) / t + \mathbb{1}\{e'_{j} a_{s} > e'_{j} b / 2\}\mathbb{1}\{u_{s} > a'_{s} e_{j} \gamma^{1/p} / \sqrt{m})\} \alpha / t\right) \\ &\geq \sup_{b \in B_{\delta}(\beta)} e'_{j} b / 2 - \sum_{s=1}^{t} \mathbb{1}\{u_{s} > e'_{j} b \gamma^{1/p} / (2\sqrt{m})\} \alpha / t \\ &\geq \eta / 2 - \xi_{t} \alpha / t, \end{split}$$

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

$$\Pr\left(\sup_{b\in B_{\delta}(\beta)} \left\|y_{t}^{b}\right\|^{p} > \gamma\right) \leq \sum_{j=1}^{m} \Pr(\eta/2 - \xi_{t}\alpha/t \leq 0)$$
$$= m \Pr(\xi_{t} \geq t\eta/(2\alpha))$$
$$\leq m \exp\left(-\frac{t\eta}{2\alpha} \left(\log\frac{\eta}{2\alpha\rho_{\gamma}} - 1\right)\right)$$
$$\leq m \exp\left(-\frac{t\eta}{4\alpha}\log\frac{\eta\gamma^{1/p}}{2\alpha}\right)$$
$$= m(\eta/(2\alpha))^{\frac{-t\eta}{4\alpha}} \gamma^{\frac{-t\eta}{4p\alpha}},$$

where the penultimate line supposes that  $\gamma$  is large enough to satisfy  $\log(\frac{\eta\gamma^{1/p}}{2\alpha})/2 \geq 1$ . Now choosing  $\gamma$  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:

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

The last expression above falls exponentially fast in t, by line (34), so this establishes line (35).  $\Box$ Corollary 3 Proof. This is the p = 0 case of Proposition 6.  $\Box$ 

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  $\|\hat{b}_t - \hat{b}_{t+1}\| \leq (\|\beta\| + \delta + \|\alpha\|)/t$ :

$$\Pr(b_t \notin B_{\delta}(\beta))$$

$$\leq \sup_{N \geq t} \Pr(\hat{b}_t \notin B_{\delta}(\beta)))$$

$$\leq \sup_{N \geq t} \sum_{j=1}^m \Pr\left(|e'_j \hat{b}_t - e'_j b_n| \geq \delta/\sqrt{m}\right)$$

$$\leq \sup_{N \geq t} 2m \exp\left(-\frac{\delta^2/m}{2\sum_{s=t}^{N-1} (\|\beta\| + \delta + \|\alpha\|)^2/s^2}\right)$$

$$< 2m \exp\left(-\frac{\delta^2}{2m(\|\beta\| + \delta + \|\alpha\|)^2 \int_{s=t-1}^{\infty} dss^2}\right)$$

$$< 2m \exp\left(-\frac{\delta^2(t-1)}{2m(\|\beta\| + \delta + \|\alpha\|)^2}\right). \tag{37}$$

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 (37) implies the result:

$$E(\tau(\delta)) = \sum_{s=1}^{n} \Pr(\tau(\delta) \ge s)$$
  
=  $\sum_{t=1}^{n} \Pr(\hat{b}_t \notin B_{\delta}(\beta))$   
<  $\sum_{t=1}^{\infty} 2m \exp\left(-\frac{\delta^2(t-1)}{2m(\|\beta\| + \delta + \|\alpha\|)^2}\right)$   
=  $O(1).$ 

Lemma 4 Proof. First, we can express the value function recursively:

$$\bar{V}_{t}^{b} \equiv \begin{cases} \max_{x_{t} \in [0,1]} x_{t} u_{t} + \bar{V}_{t-1}^{\psi_{t}^{b}(x_{t} a_{t})} & tb \geq a_{t}, \\ \bar{V}_{t-1}^{\psi_{t}^{b}(0)} & tb < a_{t}. \end{cases}$$
(38)

Second, since the shadow price weakly decreases with the inventory level, we have the following for

 $x \in [0,1]$  and  $tb \ge a_t$ :

$$(1-x)a_t'y_{t-1}^{\psi_t^b(0)} \le \bar{V}_{t-1}^{\psi_t^b(xa_t)} - \bar{V}_{t-1}^{\psi_t^b(a_t)} \le (1-x)a_t'y_{t-1}^{\psi_t^b(a_t)}$$
(39)

and 
$$xa'_{t}y^{\psi^{b}_{t}(0)}_{t-1} \leq \bar{V}^{\psi^{b}_{t}(0)}_{t-1} - \bar{V}^{\psi^{b}_{t}(xa_{t})}_{t-1} \leq xa'_{t}y^{\psi^{b}_{t}(a_{t})}_{t-1}.$$
 (40)

Third,  $\Delta_t(y_{\infty}^{b_t}) > 0$  and  $tb_t \ge a_t$  imply  $b_{t-1} = \psi_t^{b_t}(a_t)$  and hence  $\hat{v}_{t-1} = \bar{V}_{t-1}^{\psi_t^{b_t}(a_t)} - \hat{R}_{t-1}$ . Accordingly, lines (38)–(39) yield the following, when  $\Delta_t(y_{\infty}^{b_t}) > 0$  and  $tb_t \ge a_t$ :

$$\hat{R}_{t} = \bar{V}_{t}^{b_{t}} - \hat{v}_{t}$$

$$= \max_{x \in [0,1]} u_{t}x + \bar{V}_{t-1}^{\psi_{t}^{b_{t}}(xa_{t})} - u_{t} - \hat{v}_{t-1}$$

$$= \max_{x \in [0,1]} u_{t}(x-1) + \bar{V}_{t-1}^{\psi_{t}^{b_{t}}(xa_{t})} - \bar{V}_{t-1}^{\psi_{t}^{b_{t}}(a_{t})} + \hat{R}_{t-1}$$

$$\leq \max_{x \in [0,1]} u_{t}(x-1) + (1-x)a_{t}'y_{t-1}^{\psi_{t}^{b_{t}}(a_{t})} + \hat{R}_{t-1}$$

$$= (a_{t}'y_{t-1}^{\psi_{t}^{b_{t}}(a_{t})} - u_{t})^{+} + \hat{R}_{t-1}$$

$$= \Delta_{t}(y_{t-1}^{\psi_{t}^{b_{t}}(a_{t})})^{-} + \hat{R}_{t-1}.$$
(41)

Likewise, if  $\Delta_t(y^{b_t}_{\infty}) \leq 0$  and  $tb_t \geq a_t$  then (28), (38), and (40) yield

$$\hat{R}_t \leq \Delta_t (y_{t-1}^{\psi_t^{b_t}(0)})^+ + \hat{R}_{t-1}, \tag{42}$$

And if  $tb_t < a_t$  then (28) and (38) yield

$$\hat{R}_t = \hat{R}_{t-1}.\tag{43}$$

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

$$\hat{R}_t \le \sum_{s=1}^t u_s + \hat{R}_{t-1}.$$
(44)

Combining inequalities (41)–(44) inductively yields the result.  $\hfill \Box$ 

Lemma 5 Proof. I will begin by bounding the expectation of the myopic regret's first term. 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

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

I will next bound the expectation of the myopic regret's second term. This second term is zero unless  $b_t \in B_{\delta/2}(\beta)$ . To streamline the math, I will henceforth suppress all  $\mathbb{1}\{b_t \in B_{\delta/2}(\beta)\}$  indicator variables and implicitly suppose that the subsequent results condition on the event  $b_t \in B_{\delta/2}(\beta)$ .

To begin, note that  $||b_{t-1} - b_t|| \le ||2\beta + \alpha||/(t-1)$  when  $b_t \in B_{\delta/2}(\beta)$  and  $\delta$  is sufficiently small. With this, Lemma 2 implies that we can choose  $\delta$  small enough and t large enough to ensure that

$$\|y_{\infty}^{b_{t-1}} - y_{\infty}^{b_t}\| \le \|4\beta + 2\alpha\|/(\sigma_m^{\beta}(t-1)),$$

when  $b_t \in B_{\delta/2}(\beta)$ . Further, if t is sufficiently large then  $b_t \in B_{\delta/2}(\beta)$  implies  $b_{t-1} \in B_{\delta}(\beta)$ , and hence

$$\|y_{t-1}^{b_{t-1}} - y_{\infty}^{b_{t-1}}\| \le \sup_{b \in B_{\delta}(\beta)} \|y_{t-1}^b - y_{\infty}^b\|.$$

Combining the previous two results yields the following, for sufficiently large t:

$$\Delta_{t}(y_{t-1}^{b_{t-1}})^{-} \leq \Delta_{t}(y_{\infty}^{b_{t}} + \|y_{\infty}^{b_{t-1}} - y_{\infty}^{b_{t}}\|\iota + \|y_{t-1}^{b_{t-1}} - y_{\infty}^{b_{t-1}}\|\iota)^{-}$$

$$\leq \Delta_{t}(y_{\infty}^{b_{t}} + \xi_{t-1}\iota)^{-}, \qquad (45)$$
where  $\xi_{t-1} \equiv \|4\beta + 2\alpha\|/(\sigma_{m}^{\beta}(t-1)) + \sup_{b \in B_{\delta}(\beta)}\|y_{t-1}^{b} - y_{\infty}^{b}\|.$ 

Note, it's easier to work with  $\Delta_t (y_{\infty}^{b_t} + \xi_{t-1})^-$  than  $\Delta_t (y_{t-1}^{b_{t-1}})^-$ , because  $y_{\infty}^{b_t} + \xi_{t-1}$  is independent of the random function  $\Delta_t$ , whereas  $y_{t-1}^{b_{t-1}}$  is not.

Now set  $\delta$  small enough so that  $b_t \in B_{\delta/2}(\beta)$  implies  $y_{\infty}^{b_t} \in B_{\epsilon}(y_{\infty}^{\beta})$ . In this case,  $y_{\infty}^{b_t} + \xi_{t-1}\iota \in B_{\delta/2}(\beta)$ 

 $B_{\epsilon}(y_{\infty}^{\beta})$  implies the following conditional expectation bound, by Line (45) and Lemma 15:

$$\begin{split} & E\left(\mathbb{1}\{\Delta_{t}(y_{\infty}^{b_{t}})>0\}\Delta_{t}(y_{t-1}^{b_{t-1}})^{-} \mid b_{t},\xi_{t-1}\right) \\ & \leq E\left(\mathbb{1}\{\Delta_{t}(y_{\infty}^{b_{t}})>0\}\Delta_{t}(y_{\infty}^{b_{t}}+\xi_{t-1}\iota)^{-} \mid b_{t},\xi_{t-1}\right) \\ & \leq 2\sigma_{1}^{\beta}\|\xi_{t-1}\iota\|^{2} \\ & = 2m\sigma_{1}^{\beta}\xi_{t-1}^{2}. \end{split}$$

Conversely, if  $y_{\infty}^{b_t} + \xi_{t-1} \iota \notin B_{\epsilon}(y_{\infty}^{\beta})$  then line (45) yields the following conditional expectation bound:

$$\begin{split} & E\left(\mathbb{1}\{\Delta_{t}(y_{\infty}^{b_{t}})>0\}\Delta_{t}(y_{t-1}^{b_{t-1}})^{-} \mid b_{t}, \ \xi_{t-1}\right) \\ & \leq E\left(\Delta_{t}(y_{\infty}^{b_{t}}+\xi_{t-1}\iota)^{-} \mid b_{t}, \ \xi_{t-1}\right) \\ & \leq E\left(a_{t}'(y_{\infty}^{b_{t}}+\xi_{t-1}\iota) \mid b_{t}, \ \xi_{t-1}\right) \\ & \leq \|\alpha\|\|y_{\infty}^{b_{t}}+\xi_{t-1}\iota\| \\ & \leq \|\alpha\|(2\|y_{\infty}^{\beta}\|+\xi_{t-1}\sqrt{m}). \end{split}$$

For the final line, I suppose  $\delta$  is small enough to ensure that  $b_t \in B_{\delta/2}(\beta)$  implies  $\|y_{\infty}^{b_t}\| \leq 2\|y_{\infty}^{\beta}\|$ . Now, combining the previous two results yields the following, for sufficiently large t and small  $\epsilon$  and  $\delta$ :

$$\begin{split} & E\left(\mathbb{1}\{\Delta_{t}(y_{\infty}^{b_{t}})>0\}\Delta_{t}(y_{t-1}^{b_{t-1}})^{-}\right)\\ &\leq E\left(\mathbb{1}\{y_{\infty}^{b_{t}}+\xi_{t-1}\iota\in B_{\epsilon}(y_{\infty}^{\beta})\}2m\sigma_{1}^{\beta}\xi_{t-1}^{2}\right)\\ &\quad + E\left(\mathbb{1}\{y_{\infty}^{b_{t}}+\xi_{t-1}\iota\notin B_{\epsilon}(y_{\infty}^{\beta})\}\|\alpha\|(2\|y_{\infty}^{\beta}\|+\xi_{t-1}\sqrt{m})\right)\\ &\leq 2m\sigma_{1}^{\beta} E(\xi_{t-1}^{2})+2\|\alpha\|\|y_{\infty}^{\beta}\|\Pr(y_{\infty}^{b_{t}}+\xi_{t-1}\iota\notin B_{\epsilon}(y_{\infty}^{\beta})))\\ &\quad + \sqrt{m}\|\alpha\| E\left(\mathbb{1}\{y_{\infty}^{b_{t}}+\xi_{t-1}\iota\notin B_{\epsilon}(y_{\infty}^{\beta})\}\xi_{t-1}\right)\\ &\leq 2m\sigma_{1}^{\beta} E(\xi_{t-1}^{2})+2\|\alpha\|\|y_{\infty}^{\beta}\|\Pr(\xi_{t-1}\geq\epsilon/(2\sqrt{m}))\\ &\quad + \sqrt{m}\|\alpha\| E\left(\mathbb{1}\{\xi_{t-1}\geq\epsilon/(2\sqrt{m})\}\xi_{t-1}\right). \end{split}$$

The last line holds because we can make  $\delta$  small enough so that  $b_t \in B_{\delta/2}(\beta)$  implies  $y_{\infty}^{b_t} \in B_{\epsilon/2}(y_{\infty}^{\beta})$ , in which case  $b_t \in B_{\delta/2}(\beta)$  and  $y_{\infty}^{b_t} + \xi_{t-1}\iota \notin B_{\epsilon}(y_{\infty}^{\beta})$  imply that  $\|\xi_{t-1}\iota\| \ge \epsilon/2$ , and hence that  $\xi_{t-1} \ge \epsilon/(2\sqrt{m})$ . Finally, Proposition 4, Corollary 3, and Proposition 6 respectively establish that the first, second, and third terms of the last expression above are O(1/t).

Finally, the same argument yields an analogous bound for the expectation of the myopic regret's

third term.

Lemma 6 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  $\beta$  with the following inequality:

$$\|\hat{b}_{t} - \beta\| \leq \|\hat{b}_{\tau(\delta/2)+1} - \beta\| + \|\xi_{t}\| + \sum_{s=t}^{\tau(\delta/2)} \|\mathbf{E}(\hat{b}_{s} \mid \hat{b}_{s+1}) - \hat{b}_{s+1}\|,$$
(46)  
where  $\xi_{t} \equiv \sum_{s=t}^{\tau(\delta/2)} \hat{b}_{s} - \mathbf{E}(\hat{b}_{s} \mid \hat{b}_{s+1}).$ 

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 lookback 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

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

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

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

where the lower bound is within  $\delta/2$  of  $\beta$  unless  $t \leq \frac{n}{1+\delta/(2\|\alpha-\beta\|)}$ , and the upper bound is within  $\delta/2$  of  $\beta$  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

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

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  $t \ge \tau(\delta/2) + 1$ , 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  $t \le \tau(\delta/2)$  will require showing that  $\|E(\hat{b}_s \mid \hat{b}_{s+1}) - \hat{b}_{s+1}\|$  is small for all  $s \in \{t, \dots, \tau(\delta/2)\}$ , which in turn will require showing that  $\|\hat{y}_s^{\hat{b}_s} - y_{\infty}^{\hat{b}_s}\|$  is small for all  $s \in \{t+1, \dots, \tau(\delta/2)+1\}$ , which in turn will require showing that  $\|\hat{b}_s - \beta\|$  is small for all  $s \in \{t+1, \dots, \tau(\delta/2)+1\}$ .

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

$$\|\hat{b}_{\tau(\delta/2)+1} - \beta\| \le \delta/2, \\ \|\xi_t\| \le \delta/4,$$
and
$$\sum_{s=t}^{\tau(\delta/2)} \| \mathbf{E}(\hat{b}_s \mid \hat{b}_{s+1}) - \hat{b}_{s+1} \| \le \delta/4,$$
(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  $s \in \{t + 1, \dots, \tau(\delta/2) + 1\}$ . Since  $\{\xi_t\}_{t=\tau(\delta/2)}^1$  is a martingale that satisfies  $\|\xi_t - \xi_{t+1}\| = \|\hat{b}_t - \mathbb{E}(\hat{b}_t \mid \hat{b}_{t+1})\| \le (\|\beta\| + \delta + \|\alpha\|)/t$ , by design, the argument underlying line (37) analogously implies that there exists C > 0 such that  $\Pr(\|\xi_t\| > \delta/4) \le \exp(-Ct)$ , for all sufficiently large t. And since  $\Pr(A|B) = \Pr(A \cap B)/\Pr(B) \le \Pr(A)/\Pr(B)$ , it follows that

$$\Pr\left(\|\xi_t\| > \delta/4 \mid \frac{\tau^{(\delta/2)+1}}{\max} \|\hat{b}_s - \beta\| \le \delta\right)$$

$$\le \frac{\Pr(\|\xi_t\| > \delta/4)}{\Pr\left(\max_{s=t+1}^{\tau(\delta/2)+1} \|\hat{b}_s - \beta\| \le \delta\right)}$$

$$\le 2\exp(-Ct). \tag{50}$$

Note, the last line holds because  $\Pr(\max_{s=t+1}^{\tau(\delta/2)+1} \|\hat{b}_s - \beta\| \le \delta) \ge 1/2$ , 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  $s \in \{t + 1, \dots, \tau(\delta/2) + 1\}$ . 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

τ

$$\begin{split} \mathbf{E}(\hat{b}_{s} \mid \hat{b}_{s+1}) &- \hat{b}_{s+1} \\ &= \mathbf{E}(b_{s} \mid b_{s+1}) - b_{s+1} \\ &= ((s+1)b_{s+1} - \mathbf{E}(\mathbbm{1}\{\Delta_{s+1}(\underbrace{y}_{s+1}^{b_{s+1}}) > 0\}a_{s+1} \mid b_{s+1}))/s - b_{s+1} \\ &= ((s+1)b_{s+1} - b_{s+1} + \dot{\Lambda}_{\infty}^{b_{s+1}}(\underbrace{y}_{s+1}^{b_{s+1}}))/s - b_{s+1} \\ &= \dot{\Lambda}_{\infty}^{b_{s+1}}(\underbrace{y}_{s+1}^{b_{s+1}})/s \\ &= \ddot{\Lambda}_{\infty}(y_{\infty}^{b_{s+1}})(\underbrace{y}_{s+1}^{b_{s+1}} - y_{\infty}^{b_{s+1}})/s + o(\|\underbrace{y}_{s+1}^{b_{s+1}} - y_{\infty}^{b_{s+1}}\|)/s \\ &= \ddot{\Lambda}_{\infty}(y_{\infty}^{\hat{b}_{s+1}})(\underbrace{y}_{s+1}^{\hat{b}_{s+1}} - y_{\infty}^{\hat{b}_{s+1}})/s + o(\|\underbrace{y}_{s+1}^{\hat{b}_{s+1}} - y_{\infty}^{\hat{b}_{s+1}}\|)/s, \end{split}$$

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  $\delta$  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  $\max_{s=t+1}^{\tau(\delta/2)+1} \|\underbrace{y}_s^{\hat{b}_s} - y_{\infty}^{\hat{b}_s}\| \leq n^{-1/4}$  imply

$$\sum_{s=t}^{(\delta/2)} \| \mathbf{E}(\hat{b}_{s} \mid \hat{b}_{s+1}) - \hat{b}_{s+1} \| \\
\leq \sum_{s=t}^{\tau(\delta/2)} \| \ddot{\Lambda}_{\infty}(y_{\infty}^{\hat{b}_{s+1}}) (\underbrace{y}_{s+1}^{\hat{b}_{s+1}} - y_{\infty}^{\hat{b}_{s+1}}) \| / s + o(\| \underbrace{y}_{s+1}^{\hat{b}_{s+1}} - y_{\infty}^{\hat{b}_{s+1}} \|) / s \\
\leq \sum_{s=t}^{\tau(\delta/2)} 2\sigma_{1}^{\beta} \| \underbrace{y}_{s+1}^{\hat{b}_{s+1}} - y_{\infty}^{\hat{b}_{s+1}} \| / s + o(\| \underbrace{y}_{s+1}^{\hat{b}_{s+1}} - y_{\infty}^{\hat{b}_{s+1}} \|) / s \\
\leq \sum_{s=t}^{\tau(\delta/2)} 3\sigma_{1}^{\beta} n^{-1/4} / s \\
\leq \delta/4.$$
(51)

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  $\delta$  is small, and the fourth line holds because the little-o term is less than  $\sigma_1^{\beta} n^{-1/4}$  when  $\|\underbrace{y}_{s+1}^{\hat{b}_{s+1}} - y_{\infty}^{\hat{b}_{s+1}}\| \leq n^{-1/4}$  and n is large.

Further, we can use Corollary 3 upper bound the probability that  $\max_{s=t+1}^{\tau(\delta/2)+1} \|\underline{y}_{s}^{\hat{b}_{s}} - y_{\infty}^{\hat{b}_{s}}\| > n^{-1/4}$ , 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:

$$\Pr\left(\max_{s=t+1}^{\tau(\delta/2)+1} \| \underbrace{y}_{s}^{\hat{b}_{s}} - y_{\infty}^{\hat{b}_{s}} \| > n^{-1/4} | \max_{s=t+1}^{\tau(\delta/2)+1} \| \hat{b}_{s} - \beta \| \leq \delta\right)$$

$$\leq \sum_{s=t+1}^{\tau(\delta/2)+1} \Pr\left(\sup_{b\in B_{\delta}(\beta)} \| \underbrace{y}_{s}^{b} - y_{\infty}^{b} \| > n^{-1/4} | \max_{s=t+1}^{\tau(\delta/2)+1} \| \hat{b}_{s} - \beta \| \leq \delta\right)$$

$$\leq \sum_{s=t+1}^{\tau(\delta/2)+1} \frac{\Pr\left(\sup_{b\in B_{\delta}(\beta)} \| \underbrace{y}_{s}^{b} - y_{\infty}^{b} \| > n^{-1/4}\right)}{\Pr\left(\max_{s=t+1}^{\tau(\delta/2)} \| \hat{b}_{s} - \beta \| \leq \delta\right)}$$

$$\leq \sum_{s=t+1}^{\tau(\delta/2)+1} \frac{\exp(-Cn^{-1/2}(n-s))}{1/2}$$

$$\leq 2n \exp(-Cn^{-1/2}(n-\tau(\delta/2)-1))$$

$$\leq 2n \exp(-C(1-\gamma)\sqrt{n}).$$
(52)

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

$$\begin{aligned} \Pr(\hat{b}_{t} \notin B_{\delta}(\beta) \mid \max_{s=t+1}^{n} \|\hat{b}_{s} - \beta\| &\leq \delta) \\ &\leq \Pr\left(\|\xi_{t}\| > \delta/4 \mid \max_{s=t+1}^{n} \|\hat{b}_{s} - \beta\| \leq \delta\right) \\ &+ \Pr\left(\sum_{s=t}^{\tau(\delta/2)} \|\operatorname{E}(\hat{b}_{s} \mid \hat{b}_{s+1}) - \hat{b}_{s+1}\| > \delta/4 \mid \max_{s=t+1}^{n} \|\hat{b}_{s} - \beta\| \leq \delta\right) \\ &\leq 2\exp(-Ct) + \Pr\left(\max_{s=t+1}^{\tau(\delta/2)+1} \|\underbrace{y}_{s}^{\hat{b}_{s}} - y_{\infty}^{\hat{b}_{s}}\| > n^{-1/4} \mid \max_{s=t+1}^{n} \|\hat{b}_{s} - \beta\| \leq \delta\right) \\ &\leq 2\exp(-Ct) + 2n\exp(-C(1-\gamma)\sqrt{n}). \end{aligned}$$

And with our inductive hypothesis, this implies that

$$\Pr\left(\max_{s=t}^{n} \|\hat{b}_{s} - \beta\| > \delta\right)$$
  
= 
$$\Pr\left(\max_{s=t+1}^{n} \|\hat{b}_{s} - \beta\| > \delta\right) + \Pr(\hat{b}_{t} \notin B_{\delta}(\beta) \mid \max_{s=t+1}^{n} \|\hat{b}_{s} - \beta\| \le \delta)$$
  
$$\le (\tau(\delta/2) + 1 - t - 1)(2\exp(-C(t+1)) + 2n\exp(-C(1-\gamma)\sqrt{n}))$$
  
$$+ 2\exp(-Ct) + 2n\exp(-C(1-\gamma)\sqrt{n})$$
  
$$\le (\tau(\delta/2) + 1 - t)(2\exp(-Ct) + 2n\exp(-C(1-\gamma)\sqrt{n})).$$

Lemma 8 Proof. This proof will closely follow the proof of Lemma 5. For example,  $E(\mathbb{1}\{b_t \notin B_{\delta/2}(\beta)\}\sum_{s=1}^t u_s) = o(1/t)$  trivially follows from Lemma 6, as it previously followed from Lemma 3.

Bounding the expectation of the myopic regret's second term will prove more difficult. As in the proof of Lemma 5, I will henceforth suppress all  $1\{b_t \in B_{\delta/2}(\beta)\}$  indicator variables and implicitly suppose that the subsequent results condition on the event  $b_t \in B_{\delta/2}(\beta)$ .

Let  $\xi_{t-1}$  be as defined in the proof of Lemma 5, and analogously define  $\underline{\xi}_{t+1} \equiv \sup_{b \in B_{\delta/2}(\beta)} \| \underline{y}_t^b - y_\infty^b \|$ . I subscript this latter variable with t + 1 because  $\underline{y}_t^b$  is determined by that time. Note, that  $b_t \in B_{\delta/2}(\beta)$  implies  $\underline{y}_t^{b_t} \geq y_\infty^{b_t} - \underline{\xi}_{t+1}\iota$ , and hence that  $\Delta_t(\underline{y}_t^{b_t}) \leq \Delta_t(y_\infty^{b_t} - \underline{\xi}_{t+1}\iota)$ . With this, Line (45) implies the following:

$$\mathbb{E}\left(\mathbb{1}\left\{\Delta_{t}(\underbrace{y}_{t}^{b_{t}}) > 0\right\}\Delta_{t}(y_{t-1}^{b_{t-1}})^{-}\right) \\
 \leq \mathbb{E}\left(\mathbb{1}\left\{\Delta_{t}(y_{\infty}^{b_{t}} - \underbrace{\xi}_{t+1}\iota) > 0\right\}\Delta_{t}(y_{\infty}^{b_{t}} + \xi_{t-1}\iota)^{-}\right).$$
(53)

Now suppose  $y_{\infty}^{b_t} + \xi_{t-1}\iota \in B_{\epsilon}(y_{\infty}^{\beta})$  and  $y_{\infty}^{b_t} - \xi_t \iota \in B_{\epsilon}(y_{\infty}^{\beta})$ . In this case, Line (53) and Lemma 15 yield the following conditional expectation bound:

$$\begin{split} & E\left(\mathbb{1}\{\Delta_{t}(\underbrace{y}_{t}^{b_{t}}) > 0\}\Delta_{t}(y_{t-1}^{b_{t-1}})^{-} \mid b_{t}, \xi_{t-1}, \underbrace{\xi}_{t+1}\right) \\ & \leq E\left(\mathbb{1}\{\Delta_{t}(y_{\infty}^{b_{t}} - \underbrace{\xi}_{t+1}\iota) > 0\}\Delta_{t}(y_{\infty}^{b_{t}} + \xi_{t-1}\iota)^{-} \mid b_{t}, \xi_{t-1}, \underbrace{\xi}_{t+1}\right) \\ & \leq 2\sigma_{1}^{\beta} \|\underbrace{\xi}_{t+1}\iota - \xi_{t-1}\iota\|^{2} \\ & \leq 4m\sigma_{1}^{\beta}(\underbrace{\xi}_{t+1}^{2} + \xi_{t-1}^{2}). \end{split}$$

Conversely, if  $y_{\infty}^{b_t} + \xi_{t-1}\iota \notin B_{\epsilon}(y_{\infty}^{\beta})$  or  $y_{\infty}^{b_t} - \xi_t \iota \notin B_{\epsilon}(y_{\infty}^{\beta})$  then Line (53) yields the following conditional expectation bound:

$$E\left(\mathbb{1}\left\{\Delta_{t}\left(\underbrace{y}_{t}^{b_{t}}\right)>0\right\}\Delta_{t}\left(y_{t-1}^{b_{t-1}}\right)^{-}\mid b_{t},\xi_{t-1},\underbrace{\xi}_{t+1}\right)$$

$$\leq E\left(\Delta_{t}\left(y_{\infty}^{b_{t}}+\xi_{t-1}\iota\right)^{-}\mid b_{t},\xi_{t-1},\underbrace{\xi}_{t+1}\right)$$

$$\leq E\left(a_{t}'\left(y_{\infty}^{b_{t}}+\xi_{t-1}\iota\right)\mid b_{t},\xi_{t-1}\right)$$

$$\leq \|\alpha\|\|y_{\infty}^{b_{t}}+\xi_{t-1}\iota\|$$

$$\leq \|\alpha\|(2\|y_{\infty}^{\beta}\|+\xi_{t-1}m).$$

For the last line, I suppose  $\delta$  is small enough to ensure that  $b_t \in B_{\delta/2}(\beta)$  implies  $\|y_{\infty}^{b_t}\| \leq 2\|y_{\infty}^{\beta}\|$ .

Now, define  $\mathcal{E} \equiv \mathbb{1}\{y_{\infty}^{b_t} + \xi_{t-1}\iota \in B_{\epsilon}(y_{\infty}^{\beta})\}\mathbb{1}\{y_{\infty}^{b_t} - \xi_t\iota \in B_{\epsilon}(y_{\infty}^{\beta})\}$ . With this, the previous two results yield the following, for sufficiently large t and small  $\epsilon$  and  $\delta$ :

$$\begin{split} & \operatorname{E}\left(\mathbb{1}\left\{\Delta_{t}\left(\underbrace{y}_{t}^{b_{t}}\right)>0\right\}\Delta_{t}\left(y_{t-1}^{b_{t-1}}\right)^{-}\right)\\ &\leq \operatorname{E}\left(\mathcal{E}4m\sigma_{1}^{\beta}\left(\underbrace{\xi}_{t+1}^{2}+\xi_{t-1}^{2}\right)\right)+\operatorname{E}\left((1-\mathcal{E})\|\alpha\|(2\|y_{\infty}^{\beta}\|+\xi_{t-1}m)\right)\\ &\leq 4m\sigma_{1}^{\beta}\operatorname{E}\left(\underbrace{\xi}_{t+1}^{2}+\xi_{t-1}^{2}\right)\\ &\quad +2\|\alpha\|\|y_{\infty}^{\beta}\|\operatorname{Pr}\left(y_{\infty}^{b_{t}}+\xi_{t-1}\iota\notin B_{\epsilon}(y_{\infty}^{\beta})\right)\\ &\quad +2\|\alpha\|\|y_{\infty}^{\beta}\|\operatorname{Pr}\left(y_{\infty}^{b_{t}}-\underbrace{\xi}_{t}\iota\notin B_{\epsilon}(y_{\infty}^{\beta})\right)\\ &\quad +m\operatorname{E}\left(\mathbb{1}\left\{y_{\infty}^{b_{t}}+\xi_{t-1}\iota\notin B_{\epsilon}(y_{\infty}^{\beta})\right\}\xi_{t-1}\right)\\ &\quad +m\operatorname{Pr}\left(y_{\infty}^{b_{t}}-\underbrace{\xi}_{t}\iota\notin B_{\epsilon}(y_{\infty}^{\beta})\right)\operatorname{E}(\xi_{t-1}). \end{split}$$

Note, I can separate  $\xi_t$  and  $\xi_{t-1}$  in the last term of the final expression because these variables are independent of one another. And each of the terms in this final expression is either O(1/t) or O(1/(n-t)), by the argument used at the end of Lemma 5.

Finally, the argument above yields an analogous bound for the expectation of the myopic regret's third term.  $\hfill \Box$ 

Corollary 5 Proof. Copy the proof of Corollary 4.

Lemma 9 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

$$v_n^{\beta} \leq \bar{V}_t^{\beta+\xi} + \underline{\bar{V}}_t^{\beta-\frac{t}{n-t}\xi}$$
$$= t\Lambda_t^{\beta+\xi}(y_t^{\beta+\xi}) + (n-t)\underline{\Lambda}_t^{\beta-\frac{t}{n-t}\xi}(\underline{y}_t^{\beta-\frac{t}{n-t}\xi})$$
$$\equiv \bar{\Lambda}_t^{\xi}$$
(54)

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

$$\hat{R}_n \ge \bar{V}_n^\beta - \bar{\Lambda}_t^{b_t - \beta} = n\Lambda_n^\beta(y_n^\beta) - \bar{\Lambda}_t^{b_t - \beta}.$$
(55)

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

$$n\Lambda_{\infty}^{\beta}(y_{\infty}^{\beta}) - \bar{\Lambda}_{\infty}^{\xi} \ge Ct\min(\|\xi\|^2, 1), \tag{56}$$

for some C > 0, where<sup>5</sup>

$$\bar{\Lambda}^{\xi}_{\infty} \equiv t\Lambda^{\beta+\xi}_{\infty}(y^{\beta+\xi}_{\infty}) + (n-t)\Lambda^{\beta-\frac{t}{n-t}\xi}_{\infty}(y^{\beta-\frac{t}{n-t}\xi}_{\infty}).$$
(57)

Line (56) suggests that reserving  $t(\beta + \xi)$  units of inventory for the first t periods should sacrifice  $\Omega(t)$  units of value when  $\xi$  non-negligible. I will leverage this fact to show that  $\hat{R}_n$  is almost always large when  $\|b_t - \beta\|$  is non-negligible. And since  $E(\hat{R}_n)$  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 (56), our only tool for establishing the cost of  $b_t$ diverging from  $\beta$ . Now with (56), it's relatively easy to lower bound the regret when  $b_t = \beta + \xi$ , for some specific  $\xi$  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  $\xi$  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  $\underline{y}_t^{\beta-\frac{t}{n-t}\zeta}$ , 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  $\epsilon$  is a small positive number and  $\{\omega_i^{\beta}\}_{i\in[m]}$  are the orthonormal eigenvectors of  $\bar{\Lambda}_{\infty}(y_{\infty}^{\beta})^{-1}$ . 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  $\{k\epsilon\omega_j^{\beta} \mid k \in \{-1,1\} \times j \in [m]\}$ . 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  $\Lambda^b_{\infty}(y^b_{\infty})$  is concave in b, since  $\frac{\partial^2}{\partial b^2}\Lambda^b_{\infty}(y^b_{\infty}) = \frac{\partial}{\partial b}y^b_{\infty} = -\ddot{\Lambda}_{\infty}(y^b_{\infty})^{-1}$  is negative definite. This concavity implies that we can restrict attention to small  $\xi$  vectors since  $\bar{\Lambda}^{\xi}_{\infty}$  decreases in the magnitude of  $\xi$ . But, more importantly, the concavity implies line (56), as I will now show.

Let  $i \in [m]$  denote the index of the largest element of  $\xi$ , 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

$$\bar{\Lambda}_{\infty}^{\xi} \le \sup_{\{\zeta \in \mathbb{R}^m \mid e_i' \zeta = \gamma\}} \bar{\Lambda}_{\infty}^{\zeta}.$$
(58)

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

$$0 = \frac{\partial}{\partial \zeta} \left( t \Lambda_{\infty}^{\beta+\zeta}(y_{\infty}^{\beta+\zeta}) + (n-t) \Lambda_{\infty}^{\beta-\frac{t}{n-t}\zeta}(y_{\infty}^{\beta-\frac{t}{n-t}\zeta}) - \lambda(e_i'\zeta-\gamma) \right)$$
$$= t(y_{\infty}^{\beta+\zeta} - y_{\infty}^{\beta-\frac{t}{n-t}\zeta}) - \lambda e_i.$$

Now we'll use Lemma 2 to differentiate this with respect to  $\gamma$ :

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

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

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

which implies that

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

Further, since  $\lambda = 0$  when  $\gamma = 0$ , by the concavity of  $\Lambda^b_{\infty}(y^b_{\infty})$  in b, it follows that for sufficiently small  $\gamma$  we have

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

By definition, our Lagrange multiplier also satisfies  $\frac{\partial}{\partial \gamma} \sup_{\{\zeta \in \mathbb{R}^m \mid e'_i \zeta = \gamma\}} \bar{\Lambda}^{\zeta}_{\infty} = \lambda$ , which with the

result above yields the following, for sufficiently small  $\gamma$ :

$$\begin{split} n\Lambda_{\infty}^{\beta}(y_{\infty}^{\beta}) &- \sup_{\{\zeta \in \mathbb{R}^{m} \mid e_{i}^{\prime}\zeta = \gamma\}} \bar{\Lambda}_{\infty}^{\zeta} \\ &= - \big( \sup_{\{\zeta \in \mathbb{R}^{m} \mid e_{i}^{\prime}\zeta = \gamma\}} \bar{\Lambda}_{\infty}^{\zeta} - \sup_{\{\zeta \in \mathbb{R}^{m} \mid e_{i}^{\prime}\zeta = 0\}} \bar{\Lambda}_{\infty}^{\zeta} \big) \\ &= - \int_{g=0}^{\gamma} \frac{\partial}{\partial g} \sup_{\{\zeta \in \mathbb{R}^{m} \mid e_{i}^{\prime}\zeta = g\}} \bar{\Lambda}_{\infty}^{\zeta} dg \\ &\geq - \int_{g=0}^{\gamma} \frac{-ntg}{2(n-t)e_{i}^{\prime}\bar{\Lambda}_{\infty}(y_{\infty}^{\beta})e_{i}} dg \\ &= \frac{nt\gamma^{2}}{4(n-t)e_{i}^{\prime}\bar{\Lambda}_{\infty}(y_{\infty}^{\beta})e_{i}} \\ &\geq \frac{t(\|\xi\|/\sqrt{m})^{2}}{4\max_{j \in [m]} e_{j}^{\prime}\bar{\Lambda}_{\infty}(y_{\infty}^{\beta})e_{j}}. \end{split}$$

Note, the first line above holds because the concavity of  $\Lambda^b_{\infty}(y^b_{\infty})$  in *b* implies that  $n\Lambda^{\beta}_{\infty}(y^{\beta}_{\infty}) = \sup_{\{\zeta \in \mathbb{R}^m \mid e'_i \zeta = 0\}} \bar{\Lambda}^{\zeta}_{\infty}$ , and the last line holds because  $\gamma = \|\xi\|_{\infty} \ge \|\xi\|/\sqrt{m}$ . Finally, combining the result above with line (58) yields line (56).

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  $\underbrace{y}_t^{\beta-\frac{t}{n-t}\zeta}$ , which will enable us to invoke our shadow price convergence results. To this end, first note that

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

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

$$\bar{\Lambda}_t^{\xi} - \bar{\Lambda}_t^{\zeta} \le t(\xi - \zeta)' \left( y_t^{\beta + \zeta} - \underbrace{y_t^{\beta - \frac{t}{n - t}\zeta}}_{t} \right).$$
(59)

Third, I will use the preceding inequality to establish that

$$\Pr\left(\sup_{\xi \notin B_{2\sqrt{m}\epsilon}(0)} \bar{\Lambda}_t^{\xi} \le \max_{k \in \{-1,1\}} \max_{j \in [m]} \bar{\Lambda}_t^{k\epsilon\omega_j^{\beta}}\right) \ge 3/4,\tag{60}$$

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 (59) yields the following, for  $t \le n/2$ :

$$\bar{\Lambda}_{t}^{\xi} - \bar{\Lambda}_{t}^{k\epsilon\omega_{i}^{\beta}} \leq t(\xi - k\epsilon\omega_{i}^{\beta})'(y_{t}^{\beta + k\epsilon\omega_{i}^{\beta}} - \underbrace{y_{t}^{\beta - \frac{t}{n-t}k\epsilon\omega_{i}^{\beta}}}_{i}) \\
= t(\xi - k\epsilon\omega_{i}^{\beta})'(y_{\infty}^{\beta + k\epsilon\omega_{i}^{\beta}} - y_{\infty}^{\beta} - y_{\infty}^{\beta - \frac{t}{n-t}k\epsilon\omega_{i}^{\beta}} + y_{\infty}^{\beta}) \\
+ t(\xi - k\epsilon\omega_{i}^{\beta})'(y_{t}^{\beta + k\epsilon\omega_{i}^{\beta}} - y_{\infty}^{\beta + k\epsilon\omega_{i}^{\beta}}) - t(\xi - k\epsilon\omega_{i}^{\beta})'(\underbrace{y_{t}^{\beta - \frac{t}{n-t}k\epsilon\omega_{i}}}_{i} - y_{\infty}^{\beta - \frac{t}{n-t}k\epsilon\omega_{i}}) \\
= \frac{-ntk\epsilon}{n-t}(\xi - k\epsilon\omega_{i}^{\beta})'\omega_{i}^{\beta}/\sigma_{i}^{\beta} + t\|\xi - k\epsilon\omega_{i}^{\beta}\|(o(\epsilon) + O_{p}(t^{-1/2})),$$
(61)

where the last line holds because  $\|y_t^{\beta+\epsilon\omega_i^{\beta}} - y_{\infty}^{\beta+\epsilon\omega_i^{\beta}}\|$  and  $\|\underbrace{y}_t^{\beta-\frac{t}{n-t}\epsilon\omega_i} - y_{\infty}^{\beta-\frac{t}{n-t}\epsilon\omega_i}\|$  are  $O_p(t^{-1/2})$  when  $t \leq n/2$ , by Proposition 4.

Now, to derive (60) from (61), 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  $k = \operatorname{sign}(\gamma_j)$ , 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

$$k(\xi - k\epsilon\omega_j^{\beta})'\omega_j^{\beta} = k\gamma_j - \epsilon \ge \|\xi\|/(2\sqrt{m})$$
  
and  $\|\xi - k\epsilon\omega_i^{\beta}\| \le 2\|\xi\|.$ 

Finally, set  $\epsilon$  small enough so that the  $o(\epsilon)$  term in (61) 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

(61) yield the following, for all  $\xi \notin B_{2\sqrt{m}\epsilon}(0)$ :

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

This establishes line (60).

Fourth, I will use (56) to show that

$$\Pr\left(n\Lambda_n^\beta(y_n^\beta) - \max_{k \in \{-1,1\}} \max_{j \in [m]} \bar{\Lambda}_t^{k\epsilon\omega_j^\beta} \ge n^{2/3}\right) \ge 3/4.$$
(62)

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 (60), 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  $\zeta \equiv k \epsilon \omega_j^\beta$ :

$$\begin{split} n\Lambda_{n}^{\beta}(y_{n}^{\beta}) &- \bar{\Lambda}_{t}^{\zeta} \geq n\Lambda_{\infty}^{\beta}(y_{\infty}^{\beta}) - t\Lambda_{\infty}^{\beta+\zeta}(y_{\infty}^{\beta+\zeta}) + (n-t)\Lambda_{\infty}^{\beta-\frac{t}{n-t}\zeta}(y_{\infty}^{\beta-\frac{t}{n-t}\zeta}) \\ &- n|\Lambda_{\infty}^{\beta}(y_{n}^{\beta}) - \Lambda_{\infty}^{\beta}(y_{\infty}^{\beta})| \\ &- t|\Lambda_{\infty}^{\beta+\zeta}(y_{t}^{\beta+\zeta}) - \Lambda_{\infty}^{\beta+\zeta}(y_{\infty}^{\beta+\zeta})| \\ &- (n-t)|\Lambda_{\infty}^{\beta-\frac{t}{n-t}\zeta}(\underbrace{y}_{t}^{\beta-\frac{t}{n-t}\zeta}) - \Lambda_{\infty}^{\beta-\frac{t}{n-t}\zeta}(y_{\infty}^{\beta-\frac{t}{n-t}\zeta})| \\ &- t\Big|\Lambda_{t}^{\beta}(y_{n}^{\beta}) - \Lambda_{\infty}^{\beta}(y_{n}^{\beta}) - \Lambda_{t}^{\beta+\zeta}(y_{t}^{\beta+\zeta}) + \Lambda_{\infty}^{\beta+\zeta}(y_{t}^{\beta+\zeta})\Big| \\ &- (n-t)\Big|\underbrace{\underline{\lambda}_{t}^{\beta}(y_{n}^{\beta}) - \Lambda_{\infty}^{\beta}(y_{n}^{\beta}) - \underbrace{\underline{\lambda}_{t}^{\beta-\frac{t}{n-t}\zeta}(\underbrace{y}_{t}^{\beta-\frac{t}{n-t}\zeta}) + \Lambda_{\infty}^{\beta-\frac{t}{n-t}\zeta}(\underbrace{y}_{t}^{\beta-\frac{t}{n-t}\zeta})\Big|. \end{split}$$

Each term on the right is bounded in probability: First, line (56) 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  $\Lambda_{\infty}^{\beta}$  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  $(n-t)|\Lambda_{\infty}^{\beta-\frac{t}{n-t}\zeta}(y_{t}^{\beta-\frac{t}{n-t}\zeta}) - \Lambda_{\infty}^{\beta-\frac{t}{n-t}\zeta}(y_{\infty}^{\beta-\frac{t}{n-t}\zeta})|$  are  $O_{p}(t^{1/2})$  and  $O_{p}((n-t)^{1/2})$ , respectfully. Finally, Proposition 4 and Lemma 18 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 (62).

Finally, combining (55), (60), and (62) yields the following, for sufficiently small  $\epsilon$ , sufficiently large n, and  $n^{3/4} \le t \le n/2$ :

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

Finally, since  $E(\hat{R}_n) = O(\log n)$ , by Theorem 1 and line (33), the result above implies

$$O(\log n) = \mathcal{E}(\hat{R}_n)$$
  

$$\geq n^{2/3} \Pr(b_t \notin B_{2\sqrt{m\epsilon}}(\beta)) \Pr(\hat{R}_n \geq n^{2/3} \mid b_t \notin B_{2\sqrt{m\epsilon}}(\beta))$$
  

$$\geq n^{2/3} \Pr(b_t \notin B_{2\sqrt{m\epsilon}}(\beta))/2.$$

And this implies the result.

*Lemma 10 Proof.* First, note that  $\pi_t^{b_t} = 1$  implies  $b_{t-1} = \psi_t^{b_t}(a_t)$ , and hence  $v_{t-1}^{\psi_t^{b_t}(a_t)} = \bar{V}_{t-1}^{\psi_t^{b_t}(a_t)} - \bar{V}_{t-1}^{\psi_t^{b_t}(a_t)}$ 

 $\hat{R}_{t-1}$ . Also,  $\pi_t^{b_t} = 1$  implies  $tb_t \ge a_t$ , which with lines (38) and (39) yield

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

Analogously, if  $\pi_t^{b_t} = 0$  and  $tb_t \ge a_t$  then lines (38) and (40) yield

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

Further, we always trivially have

$$\hat{R}_t \ge \hat{R}_{t-1}.$$

Now choose  $\delta > 0$  small enough to ensure that  $\delta \iota \leq \beta$ , where  $\iota$  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 yields the result.

Lemma 11 Proof. I will show that there exists C > 0 that satisfies

$$\inf_{b \in B_{\delta/2}(\beta)} \mathbb{E} \left( \pi_t^b \Delta_t (y_{t-1}^{\psi_t^b(0)})^- + (1 - \pi_t^b) \Delta_t (y_{t-1}^{\psi_t^b(\alpha)})^+ \right) \ge Cm\sigma_m^\beta/(2t),$$
(63)

for all sufficiently large t. Combining this result with Lemma 9 yields the desired result:

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

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 t satisfies  $n^{3/4} \le t \le n/2$ .

Let me briefly outline how we will establish line (63). First, Lemma 13 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:

$$E\left((1-\pi_t^b)\Delta_t(y_{t-1}^{\psi_t^b(a_t)})^+\right) \ge C E\left((1-\pi_t^b)\left(\mathbb{1}\{\Delta_t(y_{\infty}^b-\iota/\sqrt{t})>0\}-\mathbb{1}\{\Delta_t(y_{\infty}^b+\iota/\sqrt{t})>0\}\right)a_t'\iota/\sqrt{t}\right).$$

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,  $E\left(\pi_t^b \Delta_t(y_{t-1}^{\psi_t^b(0)})^-\right)$  honors the same bound, except with  $\pi_t^b$  replacing  $1 - \pi_t^b$ , which means that  $E\left(\pi_t^b \Delta_t(y_{t-1}^{\psi_t^b(0)})^- + (1 - \pi_t^b) \Delta_t(y_{t-1}^{\psi_t^b(\alpha)})^+\right)$  has a corresponding  $\pi_t^b$ -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 13 establishes that there exists C > 0 that satisfies the following, for all  $b \in B_{\delta/2}(\beta)$ :

$$\Pr\left(\left\|\sqrt{t}(y_{t-1}^{\psi_{t}^{b}(a_{t})} - y_{\infty}^{\psi_{t}^{b}(a_{t})}) + 4\iota\right\| \leq 1\right)$$
$$\geq \Pr\left(\sup_{b\in B_{\delta}(\beta)}\left\|\sqrt{t}(y_{t-1}^{b} - y_{\infty}^{b}) + 4\iota\right\| \leq 1\right)$$
$$= \Pr\left(\sup_{b\in B_{\delta}(\beta)}\left\|\sqrt{t}(y_{t-1}^{b} - y_{\infty}^{b}) + 4\iota\right\| \leq 1\right)$$
$$> C.$$

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

$$y_{t-1}^{\psi_{t}^{b}(a_{t})} - y_{\infty}^{b} = (y_{t-1}^{\psi_{t}^{b}(a_{t})} - y_{\infty}^{\psi_{t}^{b}(a_{t})}) + (y_{\infty}^{\psi_{t}^{b}(a_{t})} - y_{\infty}^{b})$$
$$\leq -3\iota/\sqrt{t} + \iota/\sqrt{t}$$
$$= -2\iota/\sqrt{t},$$

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:

$$\begin{split} & E\left((1-\pi_{t}^{b})\Delta_{t}(y_{t-1}^{\psi_{t}^{b}(a_{t})})^{+}\right) \\ & \geq E\left(\mathbb{1}\{\|\sqrt{t}(y_{t-1}^{\psi_{t}^{b}(a_{t})}-y_{\infty}^{\psi_{t}^{b}(a_{t})})+4\iota\|\leq 1\}(1-\pi_{t}^{b})\Delta_{t}(y_{\infty}^{b}-2\iota/\sqrt{t})^{+}\right) \\ & \geq C E\left((1-\pi_{t}^{b})\Delta_{t}(y_{\infty}^{b}-2\iota/\sqrt{t})^{+}\right) \\ & \geq C E\left((1-\pi_{t}^{b})\mathbb{1}\{\Delta_{t}(y_{\infty}^{b}-\iota/\sqrt{t})\geq 0\}\Delta_{t}(y_{\infty}^{b}-2\iota/\sqrt{t})^{+}\right) \\ & \geq C E\left((1-\pi_{t}^{b})\mathbb{1}\{\Delta_{t}(y_{\infty}^{b}-\iota/\sqrt{t})\geq 0\}a_{t}'\iota/\sqrt{t}\right) \\ & \geq C E\left((1-\pi_{t}^{b})\mathbb{1}\{\Delta_{t}(y_{\infty}^{b}-\iota/\sqrt{t})\geq 0\}-\mathbb{1}\{\Delta_{t}(y_{\infty}^{b}+\iota/\sqrt{t})>0\}\big)a_{t}'\iota/\sqrt{t}\big). \end{split}$$

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

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:

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

Finally, adding our two bounds establishes line (63):

$$\begin{split} & \operatorname{E}\left(\pi_{t}^{b}\Delta_{t}(y_{t-1}^{\psi_{t}^{b}(0)})^{-} + (1-\pi_{t}^{b})\Delta_{t}(y_{t-1}^{\psi_{t}^{b}(\alpha)})^{+} \mid b \in B_{\delta/2}(\beta)\right) \\ &= C\operatorname{E}\left(\left(\mathbbm{1}\{\Delta_{t}(y_{\infty}^{b}-\iota/(2\sqrt{t}))>0\} - \mathbbm{1}\{\Delta_{t}(y_{\infty}^{b}+\iota/(2\sqrt{t}))>0\}\right)a_{t}'\iota/\sqrt{t}\right) \\ &= C/\sqrt{t}\int_{\gamma=-1}^{1}\iota'\frac{\partial}{\partial\gamma}\operatorname{E}\left(\mathbbm{1}\{\Delta_{1}(y_{\infty}^{b}-\gamma\iota/(2\sqrt{t}))>0\}a_{1}\right)d\gamma \\ &= C/\sqrt{t}\int_{\gamma=-1}^{1}\iota'\ddot{\Lambda}_{\infty}(y_{\infty}^{b}-\gamma\iota/\sqrt{t})\iota/(2\sqrt{t})d\gamma \\ &\geq C/\sqrt{t}\int_{\gamma=-1}^{1}\iota'\iota\sigma_{m}^{\beta}/(4\sqrt{t})d\gamma \\ &\geq Cm\sigma_{m}^{\beta}/(2t). \end{split}$$

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  $\beta$  and t is large.

**Lemma 12.**  $(y_t^b - y)'\dot{\Lambda}_t^b(y) \leq 0$  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  $\Lambda_t^b(y_t^b) \leq 0$ . **Lemma 13.** For all  $\gamma \in \mathbb{R}^m$  and  $\epsilon > 0$  there exist  $\delta, C > 0$  such that  $\Pr(\sup_{b \in B_{\delta}(\beta)} \|\sqrt{t}(y_t^b - y_{\infty}^b) - \gamma\| \le \epsilon) \ge C$  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 14, 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  $\nu$ -ball of  $y_{\infty}^\beta$ , for some  $\nu > 0$ . Finally, I use Lemma 19 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  $\ddot{\Lambda}_{\infty}(y_{\infty}^\beta)\gamma$ .

Now, I will begin the proof in earnest. Lemmas 1 and 2 imply that we can choose  $\delta$  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 14 indicates that  $\sup_{b \in B_{\delta}(\beta)} \|\sqrt{t}(y_t^b - y_{\infty}^b) - \gamma\| \leq \epsilon$ if

$$\sup_{b\in B_{\delta}(\beta)} \max_{j\in[m]} \max_{k\in\{-1,1\}} \|\sqrt{t}\dot{\Lambda}_{t}^{b}(y_{\infty}^{b} + (\gamma + \eta k\omega_{j}^{b})/\sqrt{t}) - \eta k\sigma_{j}^{b}\omega_{j}^{b}\| \le \kappa,$$
(64)

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  $k \in \{-1, 1\}$ :

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

$$\left|\sqrt{t}\dot{\Lambda}_{t}^{b}(y_{\infty}^{b}+(\gamma+\eta k\omega_{j}^{b})/\sqrt{t})-\dot{\Lambda}_{\infty}(y_{\infty}^{b})\gamma-\eta k\sigma_{j}^{b}\omega_{j}^{b}\right|\leq\kappa/3,\tag{66}$$

and 
$$\|\ddot{\Lambda}_{\infty}(y_{\infty}^{b})\gamma - \ddot{\Lambda}_{\infty}(y_{\infty}^{\beta})\gamma\| \le \kappa/3.$$
 (67)

Lines (65)–(67) imply line (64) 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 (65) from b to  $\beta$ . I will now show that there's a non-negligible chance that these inequalities hold universally across b, j, and kin 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  $\delta$  small enough to make inequality (67) hold universally.

Second, since  $\dot{\Lambda}^b_{\infty}(y^b_{\infty}) = 0$  and  $\ddot{\Lambda}_{\infty}$  is locally continuous near  $y^{\beta}_{\infty}$ , the mean value theorem indicates that there exists  $\xi^b_{tjk} \in (0,1)$  for which

$$\begin{split} \sqrt{t}\dot{\Lambda}^{b}_{\infty}(y^{b}_{\infty}+(\gamma+\eta k\omega^{b}_{j})/\sqrt{t}) \\ &=\sqrt{t}(\dot{\Lambda}^{b}_{\infty}(y^{b}_{\infty}+(\gamma+\eta k\omega^{b}_{j})/\sqrt{t})-\dot{\Lambda}^{b}_{\infty}(y^{b}_{\infty})) \\ &=\ddot{\Lambda}_{\infty}(y^{b}_{\infty}+\xi^{b}_{tjk}(\gamma+\eta k\omega^{b}_{j})/\sqrt{t})(\gamma+\eta k\omega^{b}_{j}) \\ &=\ddot{\Lambda}_{\infty}(y^{b}_{\infty})\gamma+\eta k\sigma^{b}_{j}\omega^{b}_{j}+\zeta^{b}_{tjk}(\gamma+\eta k\omega^{b}_{j}), \end{split}$$
where  $\zeta^{b}_{tjk}\equiv\ddot{\Lambda}_{\infty}(y^{b}_{\infty}+\xi^{b}_{tjk}(\gamma+\eta k\omega^{b}_{j})/\sqrt{t})-\ddot{\Lambda}_{\infty}(y^{b}_{\infty}).$ 

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

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

for all sufficiently large t. Hence, inequality (66) will hold universally for all sufficiently large t and small  $\delta$ .

Finally, I will show that for all sufficiently large t the probability that inequality (65) holds universally across  $b \in B_{\delta}(\beta)$ ,  $j \in [m]$ , and  $k \in \{-1, 1\}$  exceeds some C > 0. Since  $y_{\infty}^{b}$  and  $\omega_{j}^{b}$  are continuous in b, by Lemmas 1 and 2, it will suffice to show that there exist  $\nu > 0$  such that

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

I will prove this inequality with Lemma 19, which with proposition 3.13 of Eaton (1983) implies that conditional on  $\zeta_t \equiv \sqrt{t}(\dot{\Lambda}_t^{\beta}(y_{\infty}^{\beta}) - \dot{\Lambda}_{\infty}^{\beta}(y_{\infty}^{\beta}))$ , the random map  $(j, y) \mapsto \sqrt{t}e'_j(\dot{\Lambda}_t^{\beta}(y) - \dot{\Lambda}_{\infty}^{\beta}(y))$  weakly converges to a Gaussian process with domain  $[m] \times B_{\nu}(y_{\infty}^{\beta})$ , mean function  $\rho^{\zeta_t}$ , and covariance function  $\Xi$ , where

$$\begin{split} \rho_j^{\zeta_t}(y) &\equiv e_j' \Omega(y, y_\infty^\beta) \Omega(y_\infty^\beta, y_\infty^\beta)^{-1} \zeta_t, \\ \Xi_{j\bar{j}}(y, \bar{y}) &\equiv e_j' \Omega(y, \bar{y}) e_{\bar{j}} - e_j' \Omega(y, y_\infty^\beta) \Omega(y_\infty^\beta, y_\infty^\beta)^{-1} \Omega(y_\infty^\beta, \bar{y}) e_{\bar{j}}, \\ \text{and} \quad \Omega(y, \bar{y}) &\equiv \mathrm{E}(\mathbbm{1}\{\Delta_1(y) > 0\} \mathbbm{1}\{\Delta_1(\bar{y}) > 0\} a_1 a_1') - \mathrm{E}(\mathbbm{1}\{\Delta_1(y) > 0\} a_1) \,\mathrm{E}(\mathbbm{1}\{\Delta_1(\bar{y}) > 0\} a_1'). \end{split}$$

Since  $\Xi$  is independent of  $\zeta_t$ , the random map  $(j, y) \mapsto \sqrt{t} e'_j (\dot{\Lambda}^{\beta}_t(y) - \dot{\Lambda}^{\beta}_{\infty}(y)) - e'_j \rho_j^{\zeta_t}(y)$  is assymptotically independent of the random map  $y \mapsto \rho^{\zeta_t}(y)$ .

Accordingly, for sufficiently large t, we have

$$\begin{split} \Pr & \left( \sup_{y \in B_{\nu}(y_{\infty}^{\beta})} \| \sqrt{t} (\dot{\Lambda}_{t}^{\beta}(y) - \dot{\Lambda}_{\infty}^{\beta}(y)) + \ddot{\Lambda}_{\infty}(y_{\infty}^{\beta}) \gamma \| \leq \kappa/3 \right) \\ \geq \Pr & \left( \sup_{y \in B_{\nu}(y_{\infty}^{\beta})} \| \sqrt{t} (\dot{\Lambda}_{t}^{\beta}(y) - \dot{\Lambda}_{\infty}^{\beta}(y)) - \rho^{\zeta_{t}}(y) \| \leq \kappa/6 \right) \\ & \cap \sup_{y \in B_{\nu}(y_{\infty}^{\beta})} \| \rho^{\zeta_{t}}(y) - \ddot{\Lambda}_{\infty}(y_{\infty}^{\beta}) \gamma \| \leq \kappa/6 \right) \\ \geq p_{t}^{1} p_{t}^{2}/2, \\ \text{where} \quad p_{t}^{1} \equiv \Pr & \left( \sup_{y \in B_{\nu}(y_{\infty}^{\beta})} \| \sqrt{t} (\dot{\Lambda}_{t}^{\beta}(y) - \dot{\Lambda}_{\infty}^{\beta}(y)) - \rho^{\zeta_{t}}(y) \| \leq \kappa/6 \right) \\ \text{and} \quad p_{t}^{2} \equiv \Pr & \left( \sup_{y \in B_{\nu}(y_{\infty}^{\beta})} \| \rho^{\zeta_{t}}(y) - \ddot{\Lambda}_{\infty}(y_{\infty}^{\beta}) \gamma \| \leq \kappa/6 \right). \end{split}$$

I will now lower bound probability  $p_t^2$ . It is straightforward to confirm that  $\|\zeta_t - \rho^{\zeta_t}(y)\| = O(\|y - y_\infty^\beta\|)O(\|\zeta_t\|)$ , which implies that we can choose  $\nu$  small enough so that  $\|\zeta_t - \ddot{\Lambda}_\infty(y_\infty^\beta)\gamma\| \leq \kappa/12$  implies  $\|\zeta_t - \rho^{\zeta_t}(y)\| \leq \kappa/12$  for all  $y \in B_\nu(y_\infty^\beta)$ . This, in turn, implies that

$$p_t^2 \ge \Pr\left(\|\zeta_t - \ddot{\Lambda}_{\infty}(y_{\infty}^{\beta})\gamma\| \le \kappa/12 \ \cap \ \sup_{y \in B_{\nu}(y_{\infty}^{\beta})} \|\zeta_t - \rho^{\zeta_t}(y)\| \le \kappa/12\right)$$
$$= \Pr\left(\|\zeta_t - \ddot{\Lambda}_{\infty}(y_{\infty}^{\beta})\gamma\| \le \kappa/12\right).$$

Finally, the limit inferior of this last probability is strictly positive, as  $t \to \infty$ , because  $\zeta_t$  converges to a multivariate normal with a full-rank covariance matrix, by Lemma 19.

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

$$\begin{aligned} &\Pr\left(\sup_{\substack{y\in B_{\nu}(y_{\infty}^{\beta})}}\|\sqrt{t}(\dot{\Lambda}_{t}^{\beta}(y)-\dot{\Lambda}_{\infty}^{\beta}(y))-\rho^{\zeta_{t}}(y)\|\leq\kappa/6\right)\\ &\geq &\Pr\left(\sup_{\substack{y\in B_{\nu}(y_{\infty}^{\beta})}}\|\sqrt{t}(\dot{\Lambda}_{t}^{\beta}(y)-\dot{\Lambda}_{\infty}^{\beta}(y))-\zeta_{t}\|\leq\kappa/12\ \cap\ \sup_{\substack{y\in B_{\nu}(y_{\infty}^{\beta})}}\|\zeta_{t}-\rho^{\zeta_{t}}(y)\|\leq\kappa/12\right)\\ &\geq &\Pr\left(\sup_{\substack{y\in B_{\nu}(y_{\infty}^{\beta})}}\|\sqrt{t}(\dot{\Lambda}_{t}^{\beta}(y)-\dot{\Lambda}_{\infty}^{\beta}(y))-\zeta_{t}\|\leq\kappa/12\ \cap\ \|\zeta_{t}\|< M\right)\\ &\geq &\Pr\left(\sup_{\substack{y\in B_{\nu}(y_{\infty}^{\beta})}}\|\sqrt{t}(\dot{\Lambda}_{t}^{\beta}(y)-\dot{\Lambda}_{\infty}^{\beta}(y))-\zeta_{t}\|\leq\kappa/12\right)/2.\end{aligned}$$

Further, Lemma 17 implies that we can set  $\nu$  small enough so that the last inequality in the expression below holds:

$$\Pr\left(\sup_{\substack{y\in B_{\nu}(y_{\infty}^{\beta})}} \|\sqrt{t}(\dot{\Lambda}_{t}^{\beta}(y) - \dot{\Lambda}_{\infty}^{\beta}(y)) - \zeta_{t}\| > \kappa/12\right)$$
$$= \Pr\left(\sup_{\substack{y\in B_{\nu}(y_{\infty}^{\beta})}} \|\sqrt{t}(\dot{\Lambda}_{t}^{\beta}(y) - \dot{\Lambda}_{\infty}^{\beta}(y)) - \zeta_{t}\|^{2} > \kappa^{2}/144\right)$$
$$\leq \operatorname{E}\left(\sup_{\substack{y\in B_{\nu}(y_{\infty}^{\beta})}} \|\sqrt{t}(\dot{\Lambda}_{t}^{\beta}(y) - \dot{\Lambda}_{\infty}^{\beta}(y)) - \zeta_{t}\|^{2}\right)/(\kappa^{2}/144)$$
$$\leq 1/2.$$

Accordingly, we can set  $\nu$  small enough so that  $p_t^1 \ge 1/4$  for all sufficiently large t.

**Lemma 14.** If b is close enough to  $\beta$  to ensure that  $\{\omega_i^b\}_{i\in[m]}$  and  $\{\sigma_i^b\}_{i\in[m]}$  exist, and if  $y \in \mathbb{R}_{>0}^m$ and  $\eta > 0$  satisfy  $y + \eta k \omega_j^b \ge 0$  and  $\|\dot{\Lambda}_t^b(y + \eta k \omega_j^b) - \eta k \sigma_j^b \omega_j^b\| \le \eta \sigma_m^b/(2\sqrt{m})$  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 12 with the hypotheses of the current lemma implies the following:

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

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

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

Finally, rearranging the terms yields the result.

**Lemma 15.** There exists  $\epsilon > 0$  such that if  $y, \bar{y} \in B_{\epsilon}(y_{\infty}^{\beta})$  then  $\mathbb{E}(\mathbb{1}\{\Delta_{1}(y) > 0\}\Delta_{1}(\bar{y})^{-}) \leq 2\sigma_{1}^{\beta}\|\bar{y}-y\|^{2}$ .

Proof. I will first consider the case in which  $\bar{y} \ge y$ . 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  $y_{\infty}^{\beta}$ :

$$\begin{split} & E\left(\left(\mathbb{1}\{\Delta_{1}(y+dy)>0\}-\mathbb{1}\{\Delta_{1}(y)>0\}\right)\Delta_{1}(\bar{y})^{-}\right)\\ &= E\left(\left(\mathbb{1}\{\Delta_{1}(y+dy)>0\}-\mathbb{1}\{\Delta_{1}(y)>0\}\right)a_{1}\right)'(\bar{y}-y+O(dy))\\ &= \left(\frac{\partial}{\partial y}E(\mathbb{1}\{\Delta_{1}(y)>0\}a_{1})dy+o(dy)\right)'(\bar{y}-y+O(dy))\\ &= (\bar{y}-y)'\frac{\partial}{\partial y}E(\mathbb{1}\{\Delta_{1}(y)>0\}a_{1})dy+o(dy)\\ &= (y-\bar{y})'\ddot{\Lambda}(y)dy+o(dy). \end{split}$$

Accordingly, for y near  $y_{\infty}^{\beta}$  we have  $\frac{\partial}{\partial y} \operatorname{E}(\mathbb{1}\{\Delta_1(y) > 0\}\Delta_1(\bar{y})^-) = (y - \bar{y})'\ddot{\Lambda}(y)$ . And thus, for y and  $\bar{y}$  sufficiently close to  $y_{\infty}^{\beta}$  we have

$$\begin{split} & \mathrm{E}(\mathbbm{1}\{\Delta_{1}(y) > 0\}\Delta_{1}(\bar{y})^{-}) \\ &= \mathrm{E}(\mathbbm{1}\{\Delta_{1}(y) > 0\}\Delta_{1}(\bar{y})^{-}) - \mathrm{E}(\mathbbm{1}\{\Delta_{1}(\bar{y}) > 0\}\Delta_{1}(\bar{y})^{-}) \\ &= \int_{\gamma=0}^{1} \frac{\partial}{\partial \gamma} \mathrm{E}(\mathbbm{1}\{\Delta_{1}(\bar{y} + \gamma(y - \bar{y})) > 0\}\Delta_{1}(\bar{y})^{-})d\gamma \\ &= \int_{\gamma=0}^{1} (y - \bar{y})'\ddot{\Lambda}(\bar{y} + \gamma(y - \bar{y}))(y - \bar{y})d\gamma \\ &\leq \int_{\gamma=0}^{1} 2\sigma_{1}^{\beta} \|y - \bar{y}\|^{2}d\gamma \\ &= 2\sigma_{1}^{\beta} \|y - \bar{y}\|^{2}, \end{split}$$

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  $y_{\infty}^{\beta}$ , when y and  $\bar{y}$  are sufficiently close to  $y_{\infty}^{\beta}$ , by Lemma 1.

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

$$E(\mathbb{1}\{\Delta_{1}(y) > 0\}\Delta_{1}(\bar{y})^{-})$$
  

$$\leq E(\mathbb{1}\{\Delta_{1}(y) > 0\}\Delta_{1}(\bar{y} \lor y)^{-})$$
  

$$\leq 2\sigma_{1}^{\beta} \|y - \bar{y} \lor y\|^{2}$$
  

$$\leq 2\sigma_{1}^{\beta} \|y - \bar{y}\|^{2}.$$

**Lemma 16.** There exists C > 0 such that  $\mathbb{E}\left(\sup_{b \in \mathbb{R}^m_+} \sup_{y \in \mathbb{R}^m_+} \|\dot{\Lambda}^b_t(y) - \dot{\Lambda}^b_\infty(y)\|^2\right) \leq C/t$  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

$$\begin{split} e_j'(\dot{\Lambda}_t^b(y) - \dot{\Lambda}_\infty^b(y)) &= \sum_{s=1}^t \lambda_j^y(u_s, a_s)/t - \mathcal{E}(\lambda_j^y(u_1, a_1)), \\ \text{where} \quad \lambda_j^y(u_1, a_1) \equiv \mathbbm{1}\{\Delta_1(y) > 0\} e_j'a_1. \end{split}$$

The  $\lambda_j^y$  functions lie under an upper envelope—since  $|\lambda_j^y(u_1, a_1)| \leq ||\alpha||$ —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  $\ell(y) \equiv \max\{g \in G \mid g \leq y\}$  represent the largest gridpoint that's weakly less than  $y \in \mathbb{R}^m_+$  and let  $h(y) \equiv \min\{g \in G \mid g > \ell(y)\}$  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  $\nu$ -bracket that contains  $\lambda_j^y$ . First, if  $\|y\|_{\infty} \leq Y$  then

$$E\left(\left(\lambda_{j}^{\ell(y\wedge Y)}(u_{1},a_{1})-\mathbb{1}\{\|y\|_{\infty}\leq Y\}\lambda_{j}^{h(\bar{y})}(u_{1},a_{1})\right)^{2}\right)^{1/2} \\ \leq \|\alpha\|\Pr\left(u_{1}\in(\ell(y)'a_{1},\ell(y)'a_{1}+\gamma\iota'a_{1}]\right) \\ \leq \|\alpha\|E\left(\Pr\left(u_{1}\in(\ell(y)'a_{1},\ell(y)'a_{1}+\gamma\|\alpha\|_{1}]\mid a_{1}\right)\right) \\ \leq \gamma M\|\alpha\|_{1}^{2} \\ = \nu.$$

Next, if  $e'_i y > Y$  then

$$\mathbb{E}\left(\left(\lambda_{j}^{\ell(y\wedge Y)}(u_{1},a_{1})-\mathbb{1}\left\{\|y\|_{\infty}\leq Y\right\}\lambda_{j}^{h(\bar{y})}(u_{1},a_{1})\right)^{2}\right)^{1/2} \\
 \leq \|\alpha\|\Pr\left(u_{1}>a_{1}'(y\wedge Y)\right) \\
 \leq \|\alpha\|\Pr(u_{1}>Ya_{1}'e_{i}) \\
 \leq \|\alpha\|\left(\Pr(a_{1}'e_{i}\leq A)+\Pr(u_{1}>YA)\right) \\
 \leq \nu.$$

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  $\nu$ . Hence, for small enough  $\nu$  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  $\int_{\nu=0}^{1} \sqrt{\log N_{\nu}} d\nu < \infty$ .

**Lemma 17.** For all  $\delta > 0$  there exist  $\epsilon > 0$  such that  $\mathbb{E}\left(\sup_{y \in B_{\epsilon}(y_{\infty}^{\beta})} \|\dot{\Lambda}_{t}^{b}(y) - \dot{\Lambda}_{\infty}^{b}(y) - \dot{\Lambda}_{t}^{b}(y_{\infty}^{\beta}) + \dot{\Lambda}_{\infty}^{b}(y_{\infty}^{\beta})\|^{2}\right) \leq \delta/t$  for all  $t \in \mathbb{N}$ .

*Proof.* The proof is similar to the proof of Lemma 16, except with

$$\lambda_j^y(u_1, a_1) \equiv (\mathbb{1}\{\Delta_1(y) > 0\} - \mathbb{1}\{\Delta_1(y_{\infty}^{\beta}) > 0\})e_j'a_1.$$

Modifying the proof of Lemma 16 establishes that the bracketing integral of  $\{\lambda_j^y\}_{y \in B_{\epsilon}(y_{\infty}^{\beta}), j \in [m]}$  is uniformly bounded in  $\epsilon \in [0, 1]$ . Further,  $\lambda_j^y$  is bounded by envelope function  $\bar{\lambda}_j^y$ , where

$$\begin{split} \bar{\lambda}_{j}^{y}(u_{1},a_{1}) &\equiv (\mathbb{1}\{\Delta_{1}(\underline{y}) > 0\} - \mathbb{1}\{\Delta_{1}(\bar{y}) > 0\})e_{j}'a_{1},\\ &\underline{y} \equiv y \wedge y_{\infty}^{\beta},\\ &\text{and} \quad \bar{y} \equiv y \vee y_{\infty}^{\beta}. \end{split}$$

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:

$$\begin{split} \mathbf{E}(\bar{\lambda}_{j}^{y}(u_{1},a_{1})^{2}) &= \mathbf{E}\left((\mathbb{1}\{\Delta_{1}(\underline{y})>0\} - \mathbb{1}\{\Delta_{1}(\bar{y})>0\})(e_{j}'a_{1})^{2}\right)\\ &\leq \|\alpha\|e_{j}' \mathbf{E}\left((\mathbb{1}\{\Delta_{1}(\underline{y})>0\} - \mathbb{1}\{\Delta_{1}(\bar{y})>0\})a_{1}\right)\\ &= \|\alpha\|e_{j}' \int_{\gamma=0}^{1} \frac{\partial}{\partial\gamma} \mathbf{E}\left(\mathbb{1}\{\Delta_{1}(\bar{y}+\gamma(\underline{y}-\bar{y}))>0\}a_{1}\right)d\gamma\\ &= -\|\alpha\|e_{j}' \int_{\gamma=0}^{1} \ddot{\Lambda}_{\infty}(\bar{y}+\gamma(\underline{y}-\bar{y}))(\underline{y}-\bar{y})d\gamma\\ &= -\|\alpha\|e_{j}'\ddot{\Lambda}_{\infty}(\bar{y}+\bar{\gamma}(\underline{y}-\bar{y}))(\underline{y}-\bar{y}), \end{split}$$

for some  $\bar{\gamma} \in [0, 1]$ . And since we constrain  $y \in B_{\epsilon}(y_{\infty}^{\beta})$ , it follows that  $\ddot{\Lambda}_{\infty}(\bar{y} + \bar{\gamma}(\underline{y} - \bar{y})) \to \ddot{\Lambda}_{\infty}(y_{\infty}^{\beta})$ and  $(\underline{y} - \bar{y}) \to 0$  as  $\epsilon \to 0$ . Accordingly,  $\mathrm{E}(\bar{\lambda}_{j}^{y}(u_{1}, a_{1})^{2}) \to 0$  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 18.** For any compact  $\Omega \subset \mathbb{R}^m_+$ , there exists C > 0 such that  $\operatorname{E}\left(\sup_{y,\bar{y}\in\Omega} |\Lambda^b_t(y) - \Lambda^b_\infty(y) - \Lambda^{\bar{b}}_t(\bar{y}) + \Lambda^{\bar{b}}_\infty(\bar{y})|^2\right) \leq C/t$  for all  $t \in \mathbb{N}$ ,  $b, \bar{b} \in \mathbb{R}^m_+$ .

*Proof.* Like in the proofs of Lemmas 16 and 17, 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:

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

Note that  $|\lambda_y^{\bar{y}}(u_1, a_1)| \leq ||\alpha|| \operatorname{diam}(\Omega) < \infty$ . 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  $\ell(y) \equiv \max\{g \in G \mid g \leq y\}$  represent the largest gridpoint that's weakly less than  $y \in \mathbb{R}^m_+$ and let  $h(y) \equiv \min\{g \in G \mid g > \ell(y)\}$  represent the smallest gridpoint that's strictly larger than  $\ell(y)$ . By design, we have

$$\lambda_{\ell(y)}^{h(\bar{y})} - \lambda_{h(y)}^{\ell(\bar{y})} \le (u_1 - a_1'(y - \gamma \iota))^+ - (u_1 - a_1'(\bar{y} + \gamma \iota))^+ - (u_1 - a_1'(y + \gamma \iota))^+ + (u_1 - a_1'(\bar{y} - \gamma \iota))^+ \le 4\gamma \|\alpha\| \|\iota\| = \nu.$$

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

**Lemma 19.** For all sufficiently small  $\epsilon > 0$ , the random mapping  $(j, y) \mapsto \sqrt{t}e'_j(\dot{\Lambda}^b_t(y) - \dot{\Lambda}^b_\infty(y))$ , 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  $\sum_{j\bar{j}}(y, \bar{y}) \equiv e'_j \operatorname{E}(\mathbb{1}\{\Delta_1(y) > 0\}\mathbb{1}\{\Delta_1(\bar{y}) > 0\}a_1a'_1) - \operatorname{E}(\mathbb{1}\{\Delta_1(y) > 0\}a_1) \operatorname{E}(\mathbb{1}\{\Delta_1(\bar{y}) > 0\}a'_1)e_{\bar{j}}$ .

*Proof.* This is a direct application of theorem 2.3 of Kosorok, a classical empirical processes result. The proof of Lemma 16 establishes that the corresponding bracketing integral is finite.  $\Box$