City Research Online

City, University of London Institutional Repository

This is the supplemental version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: https://openaccess.city.ac.uk/id/eprint/23383/

Link to published version: https://doi.org/10.1016/j.jet.2017.06.006

Supplement to "Conjugate Information Disclosure in an Auction with Learning"

Arina Nikandrova and Romans Pancs

Birkbeck and ITAM

B Appendix

B.1 A Signal Structure that Rationalizes the Information-Acquisition Technology in Section 2

The model's description, in Section 2, *could* have been specified as follows. The follower exerts effort a. This effort affects the precision of a signal z. This signal's realization induces a conditional probability distribution μ_z of the underlying valuation v. This conditional probability distribution implies the expected conditional valuation $\theta_2 \equiv \mathbb{E}^{\mu_z}[v]$. Before the realization of z has been observed, μ_z and θ_2 are random variables.

The alternative (but equivalent) approach taken in Section 2 makes direct assumptions on how a affects the probability distribution of θ_2. It would have been a mere normalization to identify the set of signal realizations with the set of conditional (on this signal) probability distributions by setting $z = \mu_z$ (Kamenica and Gentzkow, 2011). Because each player is an expected-utility maximizer, however, each cares only about θ_2, and so it is appropriate to identify the set of signal realizations with the set of conditional expectations by setting $z = \theta_2$. The underlying signal structure that induces the probability distribution of θ_2 has been left implicit in the paper's main body, but can be recovered.

For concreteness, this appendix shows how the dependence of θ_2 on a assumed in Condition 1 can be (non-uniquely) rationalized with an appropriate joint probability distribution for v and z. Assume that each c.d.f. F_j in Condition 1 has a p.d.f. f_j, $j = L, H$. Let the follower's underlying valuation be $v \in \{0, 1\}$ with $\Pr\{v = 1\} = p$, where $p \equiv \int_0^1 s \, dF_H(s) = \int_0^1 s \, dF_L(s)$. Then, by construction, $\Pr\{v = 1\} = \mathbb{E}[\theta_2 \mid a]$ for all $a \in A$, meaning that the probability that the follower assigns to $v = 1$ before observing z equals his expectation of the conditional (on z) probability that $v = 1$, which is also his conditional expectation of v, denoted by θ_2. This Bayesian consistency condition is necessary and sufficient for θ_2 to represent the follower's conditional expectation of his underlying valuation (Kamenica and Gentzkow, 2011).

Assume that the signal z can be either more precise, with probability a, or less precise, with probability $1 - a$. The realizations of the more and the less precise signals are governed by the conditional p.d.f.s $\sigma_H(z \mid v)$ and $\sigma_L(z \mid v)$, where

$$\sigma_j(z \mid v) \equiv \frac{z^v (1-z)^{1-v}}{p^v (1-p)^{1-v}} f_j(z), \qquad j \in \{H, L\}, v \in \{0, 1\}, z \in [0, 1]. \tag{B.1}$$

The Law of Total Probability applied to (B.1) implies that, conditional on signal technology j, z is distributed according to the c.d.f. F_j; that is, the probability that the signal realization does not exceed z is

$$\int_{s \leq z} \left[p\sigma_j \left(s \mid 1 \right) + \left(1 - p \right) \sigma_j \left(s \mid 0 \right) \right] \mathrm{d}s = F_j \left(z \right),$$

which immediately implies that unconditionally, for some effort a, z is distributed with the c.d.f. $F \left(\cdot \mid a \right)$.

Bayes' rule implies that z is also the expectation of v conditional on z and on signal technology j:

$$\mathbb{E} \left[v \mid z, j \right] = \Pr \left\{ v = 1 \mid z, j \right\} = \frac{\sigma_j \left(z \mid 1 \right) p}{\sigma_j \left(z \mid 1 \right) p + \sigma_j \left(z \mid 0 \right) \left(1 - p \right)} = z,$$

which immediately implies the expectation that is conditional only on z:

$$\theta_2 \equiv \mathbb{E} \left[v \mid z \right] = a \mathbb{E} \left[v \mid z, j = H \right] + \left(1 - a \right) \mathbb{E} \left[v \mid z, j = L \right] = z.$$

Hence, because z is distributed according to the c.d.f. $F \left(\cdot \mid a \right)$, so is θ_2, as desired.

B.2 What the Information-Acquisition Technology in Section 2 Rules Out

The linear specification (1) and Condition 1 are restrictive. The linearity in (1) rules out information-acquisition technologies that let the follower choose among three or more signals, as Example B.1 clarifies.

Example B.1 (Nonexample). The follower chooses a tuple $\left(a_1, a_2 \right)$ in a two-dimensional probability simplex Δ^2, and then draws θ_2 from the probability distribution with the c.d.f. $F \left(\theta_2 \mid a_1, a_2 \right) = a_2 / 2 + a_1 \theta_2 + \left(1 - a_1 - a_2 \right) \mathbf{1}_{\{\theta_2 \geq 1/2\}}$.

In Example B.1, in addition to allocating probability to a perfectly informative and a somewhat informative signal about the underlying valuation in $\{0, 1\}$, as in Example 1, the follower can also allocate some probability to a completely uninformative signal (with probability $1 - a_1 - a_2$). Ruling out Example B.1 is economically restrictive. If the cost of information acquisition were increasing in a_1 and a_2, one could imagine the follower preferring to set both a_1 and a_2 close to zero if he faced a price close to 0 or 1, and optimally trading off the positive a_1 and a_2 otherwise.

Condition 1 remains restrictive even conditional on the linear specification (1), as Example B.2 illustrates.

Example B.2 (Another Nonexample). $F \left(\theta_2 \mid a \right) = a \left(\frac{1}{4} \mathbf{1}_{\{\theta_2 < 1/2\}} + \frac{1}{2} \mathbf{1}_{\{1/2 \leq \theta_2 < 1\}} + \mathbf{1}_{\{\theta_2 = 1\}} \right) + \left(1 - a \right) \theta_2$.

Example B.2 can be interpreted to say that, with probability a, the follower observes a signal that, with probability $1/2$, reveals his underlying valuation, which is distributed uniformly on $\{0, 1\}$, and, with probability $1/2$, reveals "nothing"; with probability $1 - a$, the follower observes the partially informative signal of Example 1. Even though $F \left(\cdot \mid 1 \right)$ is a mean-preserving spread of $F \left(\cdot \mid 0 \right)$ and, hence, is more informative in some sense (viz. Blackwell's order on the underlying signals), $F \left(\cdot \mid 1 \right)$ and $F \left(\cdot \mid 0 \right)$ are not rotation-ordered.

B.3 An Analytical Equivalent of Condition 2

In applications, Condition 2 can be checked analytically. To do so, let

$$r(\theta_1) \equiv \frac{1 - G(\theta_1)}{g(\theta_1)}, \qquad \theta_1 \in \Theta_1,$$

denote the inverse hazard rate of the leader's c.d.f. As is standard, $r(\theta_1)$ is interpreted as the profit that the seller forgoes—equivalently, the information rent that the leader reaps—when the seller commits to sell to a type-θ_1 leader.[1] In addition, recall that $R(\theta_1)$, defined in (A.2), denotes the planner's return to the follower's information acquisition in the first-best benchmark when the leader's type is θ_1. This return is closely related to the follower's information-acquisition technology (in particular, $R'(\theta_1) = F_H(\theta_1) - F_L(\theta_1)$ and $R''(\theta_1) = f_H(\theta_1) - f_L(\theta_1)$) and so can be treated as a primitive.

Lemma B.1. *Suppose that Condition 1 holds and $f_L(\theta^*) \neq f_H(\theta^*)$.[2] Then, a prospect set is convex if and only if*

$$r''(\theta_1) + \left(r(\theta_1) \frac{R''(\theta_1)}{R'(\theta_1)} \right)' < 0 \qquad \text{for all } \theta_1 \in (0, \theta^*) \cup (\theta^*, 1). \tag{B.2}$$

Proof. A prospect set, $\Gamma \equiv \{ (\pi(\theta_1), \alpha(\theta_1)) \mid \theta_1 \in \Theta_1 \}$, is a parametrically given plane curve. Its signed curvature at θ_1 is given by[3]

$$\kappa(\theta_1) \equiv \frac{\alpha''(\theta_1) \pi'(\theta_1) - \alpha'(\theta_1) \pi''(\theta_1)}{\left((\pi'(\theta_1))^2 + (\alpha'(\theta_1))^2 \right)^{3/2}}, \tag{B.3}$$

where primes refer to derivatives with respect to θ_1. Because Γ is simple[4] and regular,[5] it is strictly convex if and only if κ is either always positive or always negative. Because the denominator in (B.3) is always positive, requiring that κ does not change the sign is equivalent to requiring that the numerator in (B.3) not change the sign.

When $\theta_1 = \theta^*$, the numerator in (B.3) is negative, or $-r(\theta^*)(f_H(\theta^*) - f_L(\theta^*))^2 / c < 0$, because $f_H(\theta^*) \neq f_L(\theta^*)$ by the lemma's hypothesis and $F_L(\theta^*) = F_H(\theta^*)$ by part (ii) of Condition 1. Thus, the strict convexity of Γ is equivalent to the numerator in (B.3) being always negative:[6]

$$\alpha''(\theta_1) \pi'(\theta_1) - \alpha'(\theta_1) \pi''(\theta_1) < 0.$$

---

[1]When the seller commits to sell to type θ_1 at some price, all types higher than θ_1 may be tempted to imitate type θ_1 and buy at the same price, thereby constraining the seller in how much he can charge these higher types.

[2]Condition $f_L(\theta^*) \neq f_H(\theta^*)$, which can be interpreted to hold "generically," simplifies the analytical characterization in the lemma but is not required for the convexity of the prospect set.

[3]The curvature of Γ at a point is the reciprocal of the radius of the circle osculating Γ at that point; see the Wikipedia entry on curvature: https://en.wikipedia.org/wiki/Curvature.

[4]A curve is simple if it does not intersect itself.

[5]A curve Γ is regular if its derivative $(\alpha', \pi') \neq (0, 0)$ for all $\theta_1 \in \Theta_1$, which holds in our model.

[6]In general, the sign of the curvature κ indicates the direction in which the unit tangent vector rotates as a function of the parameter along the curve. If the unit tangent rotates counterclockwise, then $\kappa > 0$. If it rotates clockwise, then $\kappa < 0$. In our model, as θ_1 increases, the unit tangent vector of Γ rotates clockwise, and, thus, κ must be negative everywhere.

Substituting the definitions of α and π into the above display, dividing by $\left(R'\left(\theta_1\right)\right)^2$, which is positive when $\theta_1 \neq \theta^*$, and rearranging gives the sought inequality (B.2) of Lemma B.1.

This curvature condition captured by (B.2) is local and, alone, does not suffice to conclude that the prospect set is convex (in the sense of Definition 1); a spiral is a counterexample. Condition 1, however, which ensures that $\alpha\left(0\right) = \alpha\left(1\right) = 0$, thereby ruling out a spiral and ensuring that the curvature condition in (B.2) is equivalent to the convexity of Γ. $\square$

B.4   Examples that Illustrate Cases in Definition 2

Case (i) in Definition 2 prevails in examples in which the distribution of the follower's underlying valuations is binary, and the information-acquisition technology grants probabilistic access to a perfectly informative signal, as in Example 1. In this case, the follower's c.d.f. F has mass points at 0 and 1. Example 1, coupled with the assumption of the monotone increasing hazard rate for the leader's c.d.f. G, yields the prospect set in Figure 3c. This prospect set's critical feature is that it slopes upwards near $\theta_1 = 0$ (that is, both α and π are increasing in θ_1 near 0), and so $\underline{\theta} = 0$. That α is increasing near 0 follows from Theorem 1. That π is increasing near 0 follows by taking an arbitrarily small $\varepsilon > 0$ and evaluating

$$\pi'\left(\varepsilon\right) = -r'\left(\varepsilon\right)\left(F_H\left(\varepsilon\right) - F_L\left(\varepsilon\right)\right) + r\left(\varepsilon\right)\left(f_L\left(\varepsilon\right) - f_H\left(\varepsilon\right)\right) > 0,$$

where the inequality follows because $r'\left(\varepsilon\right) < 0$ (the hazard-rate condition on G), $r\left(\varepsilon\right) > 0$, $F_H\left(\varepsilon\right) = 1/2 > F_L\left(\varepsilon\right) = \varepsilon$ (the mass point that corresponds to probabilistically learning that the underlying valuation is 0), and $f_L\left(\varepsilon\right) = 1 > f_H\left(\varepsilon\right) = 0$ (made possible by F_H's mass point at 0).

Figure B.1 illustrates how a downward-sloping segment for Γ near $\theta_1 = 0$ is necessary for case (ii) in Definition 2 not to collapse into case (i). The figure also illustrates the role played by s'.

Figure 3b illustrates an example of case (ii); c.d.f.s F_H and F_L are Beta distributions chosen to satisfy Condition 1. Then, $F_H\left(0\right) = F_L\left(0\right) = 0$. Furthermore, one can (merely to simplify the argument) choose F_H and F_L so that $f_H\left(0\right) > f_L\left(0\right)$. As a result, $\pi'\left(0\right) = r\left(0\right)\left(f_L\left(0\right) - f_H\left(0\right)\right) < 0$; π is decreasing near 0. Because α is increasing near 0, the prospect set is downward-sloping near 0.

B.5   Justifying Equation (A.16) in the Proof of Lemma A.4

To justify (A.16), Lemma B.2 demonstrates that one can approximate any $\nu \in \Delta\left(\Delta\Theta_1\right)$ by a probability measure that puts some mass only on discrete measures in a countable set. The proof proceeds in two steps. First, it shows that by choosing n sufficiently large, any probability measure in $\Delta\Theta_1$ can be approximated by a probability measure that puts some mass on a countable set $\left\{\frac{1}{2^n}, \frac{2}{2^n}, ..., 1\right\}$ in Θ_1. Then, a similar argument is repeated to show that if one chooses n sufficiently large, any measure in $\Delta\left(\Delta\Theta_1\right)$ can be approximated by a measure that puts positive mass only on discrete measures with support $\left\{\frac{1}{2^n}, \frac{2}{2^n}, ..., 1\right\}$. This second half is slightly trickier because it requires finding a countable set of non-overlapping neighborhoods in $\Delta\Theta_1$ that almost cover space
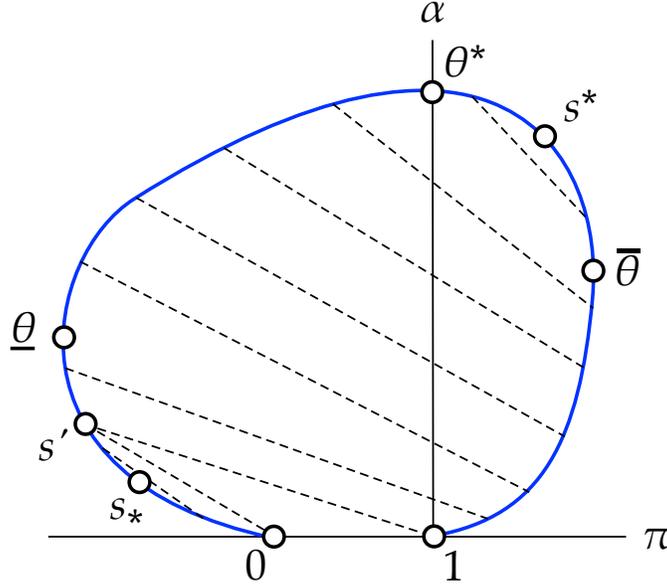
Figure B.1: The convex curve is the prospect set Γ. The circles mark the prospects induced by the leader's types in $0, s_*, s', \underline{\theta}, \theta^*, s^*, \bar{\theta}$, and 1. The dashed links comprise a subset of the links that pool prospects into messages. Type s' demarcates the leader's types that are pooled with types in $[0, s_*)$ and those that are pooled with types in $(s^*, 1]$.

$\Delta\Theta_1$.

Lemma B.2. *Fix an arbitrary measure $\nu \in \Delta(\Delta\Theta_1)$. For every $\epsilon > 0$, there exists N such that for $n \geq N$,*

$$\left| \int_{\Delta\Theta_1} f(P) \, d\nu - \int_{\Delta\Theta_1} f(P) \, d\nu_n \right| < \epsilon,$$

where $f(P)$ is an arbitrary real-valued uniformly continuous, bounded function, and ν_n is a probability measure that puts some mass only on discrete measures in the countable set

$$\mathcal{D}_n \equiv \left\{ \alpha_1 \delta_{1/2^n} + \alpha_2 \delta_{2/2^n} + \ldots + \alpha_{2^n} \delta_1 \; : \; \alpha_1, \ldots, \alpha_{2^n} \in \mathbb{Q} \cap [0,1], \; \sum_{j=1}^{2^n} \alpha_j = 1 \right\} \subset \Delta\Theta_1,$$

where $\mathbb{Q}$ denotes the set of rational numbers, and $\delta_{k/2^n}$ denotes the Dirac measure at $k/2^n \in [0,1]$ (i.e., $\delta_{k/2^n}(B) = \mathbf{1}_{\{k/2^n \in B\}}$, $B \subset \Theta_1$). Set $\mathcal{D}_n$ contains probability measures that put some (rational) mass on a countable set $\left\{ \frac{1}{2^n}, \frac{2}{2^n}, \ldots, 1 \right\}$ in Θ_1.

Proof. The proof proceeds in two steps.

Step 1: It is possible to approximate any measure μ in $\Delta\Theta_1$ with a measure in $\mathcal{D}_n$ by choosing n sufficiently high.

Let $B_j^n \equiv \left[\frac{j-1}{2^n}, \frac{j}{2^n} \right)$ for $j = 1, 2, \ldots, 2^n$, so that the family of disjoint sets $\{B_1^n, \ldots, B_{2^n}^n\}$ completely

5

covers Θ_1. Note that it is possible to approximate a discrete measure

$$\mu\left(B_1^n\right)\delta_{1/2^n} + ... + \mu\left(B_{2^n}^n\right)\delta_1$$

by

$$\mu_n \equiv \alpha_1^n \delta_{1/2^n} + ... + \alpha_{2^n}^n \delta_1,$$

where $\alpha_j^n \in [0,1] \cap \mathbb{Q}$ such that $\sum_{j=1}^{2^n} \alpha_j^n = 1$ and

$$\sum_{j=1}^{2^n} \left|\mu\left(B_j^n\right) - \alpha_j^n\right| < \frac{1}{2^n}.$$

Such a choice of $\left\{\alpha_j^n\right\}$ is possible because rationals are dense in reals. Then, for each n, $\mu_n \in \mathcal{D}_n$. Moreover, as $n \to \infty$, $\mu_n \Rightarrow \mu$, where "$\Rightarrow$" denotes weak convergence and $\mu \in \Delta\Theta_1$.

To show that $\mu_n \Rightarrow \mu$, take a uniformly continuous bounded function g on $\Theta_1 = [0,1]$.[7] Let $\|g\|_\infty \equiv \sup_{x \in \Theta_1} g(x)$ denote the supremum norm. Then,

$$
\begin{aligned}
\left|\int g \mathrm{d}\mu_n - \int g \mathrm{d}\mu\right| &= & \left|\sum_{j=1}^{2^n} \alpha_j^n g\left(\frac{j}{2^n}\right) - \int g \mathrm{d}\mu\right| & \\
&< & \left|\sum_{j=1}^{2^n} \mu\left(B_j^n\right) g\left(\frac{j}{2^n}\right) - \int g \mathrm{d}\mu\right| & + \frac{1}{2^n} \sup_j \left|g\left(\frac{j}{2^n}\right)\right| \\
&\leq & \left|\int \sum_{j=1}^{2^n} g\left(\frac{j}{2^n}\right) \mathbf{1}_{\left\{B_j^n\right\}} \mathrm{d}\mu - \int g \mathrm{d}\mu\right| & + \frac{1}{2^n} \|g\|_\infty \\
&\leq & \left|\sum_{j=1}^{2^n} \int \left(g\left(\frac{j}{2^n}\right) - g\right) \mathbf{1}_{\left\{B_j^n\right\}} \mathrm{d}\mu\right| & + \frac{1}{2^n} \|g\|_\infty \\
&\leq & \left|\sum_{j=1}^{2^n} \sup_{x \in B_j^n} \left|g\left(\frac{j}{2^n}\right) - g(x)\right| \mu\left(B_j^n\right)\right| & + \frac{1}{2^n} \|g\|_\infty .
\end{aligned}
$$

Note that $\left|\frac{j}{2^n} - x\right| < \frac{1}{2^n}$ for each $x \in B_j^n$. Because g is uniformly continuous, for every $\epsilon > 0$, there exists a $\delta > 0$ such that whenever $|x - y| < \delta$, $|g(x) - g(y)| < \epsilon$. Take some $\epsilon > 0$; then, for n such that $\frac{1}{2^n} \leq \delta$, $\left|g\left(\frac{j}{2^n}\right) - g(x)\right| < \epsilon$ for all $x \in B_j^n$ and all j. Then, from previous calculations, it follows that

$$\left|\int g \mathrm{d}\mu_n - \int g \mathrm{d}\mu\right| \leq \epsilon + \frac{1}{2^n} \|g\|_\infty .$$

Because g is bounded, the second term on the right-hand side can be made arbitrarily small by choosing n sufficiently large, whereas ϵ is arbitrary. Hence, $\int g \mathrm{d}\mu_n \to \int g \mathrm{d}\mu$ as $n \to \infty$, which implies that $\mu_n \Rightarrow \mu$.

---

[7]Statements that $\mu_n \Rightarrow \mu$ and that $\lim \int g \mathrm{d}\mu_m = \int g \mathrm{d}\mu$ for all uniformly continuous, bounded functions are equivalent because (i) every Lipschitz function between two metric spaces is uniformly continuous; and (ii) the set of bounded Lipschitz functions on a metric space is dense in the set of continuous bounded functions on that space (Dudley, R. M., *Real Analysis and Probability* 2002, Theorem 11.2.4), which implies that, instead of a wider class of bounded continuous function in the definition of the weak convergence, one may actually consider a smaller class of bounded Lipschitz functions (this fact is sometimes stated as a part of Portmanteau theorem).

Step 2: It is possible to approximate any measure ν in $\Delta\left(\Delta\Theta_1\right)$ with a measure that puts some mass only on measures in $\mathcal{D}_n$.

Let

$$\mathcal{V} \equiv \left\{ \sum_{n=0}^{k}\sum_{j=1}^{\infty} \beta_{nj}\mu_n^j : \mu_n^j \in \mathcal{D}_n, \ \beta_{0j}, ..., \beta_{kj} \in \mathbb{Q}\cap[0,1], \ \sum_{n=0}^{k}\sum_{j=1}^{\infty}\beta_{nj} = 1, k = 0, 1, 2, ... \right\}$$

be a countable subset of probability measures in $\Delta\left(\Delta\Theta_1\right)$. It contains measures that put positive mass only on measures in a countable set $\mathcal{D} \equiv \cup_{n=0}^{\infty}\mathcal{D}_n$. It will be demonstrated that $\mathcal{V}$ is dense in $\Delta\left(\Delta\Theta_1\right)$, and, thus, an arbitrary measure in $\Delta\left(\Delta\Theta_1\right)$ can be approximated by some measure in $\mathcal{V}$.

Let $\nu \in \Delta\left(\Delta\Theta_1\right)$ and

$$B\left(\mu_n^j, 1/m\right) \equiv \left\{ \mu \in \Delta\Theta_1 : d_p\left(\mu_n^j, \mu\right) < 1/m \right\}$$

be an open ball in $\Delta\left(\Delta\Theta_1\right)$ with radius $1/m$ centered around measure $\mu_n^j \in \mathcal{D}_n$, where $d_p\left(\mu_n^j, \mu\right)$ denotes Prohorov distance between measures μ_n^j and μ. For each $m \geq 1$,

$$\cup_{j=1}^{\infty}B\left(\mu_0^j, \frac{1}{m}\right) \subset \cup_{j=1}^{\infty}B\left(\mu_1^j, \frac{1}{m}\right) \subset ... \text{ and } \lim_{n\to\infty}\cup_{j=1}^{\infty}B\left(\mu_n^j, \frac{1}{m}\right) = \Delta\left(\Delta\Theta_1\right).$$

Take N and J such that

$$\nu\left(\cup_{j=1}^{J}B\left(\mu_N^j, \frac{1}{m}\right)\right) \geq 1 - 1/m.$$

Modify the balls $B\left(\mu_N^j, \frac{1}{m}\right)$ into disjoint sets by taking

$$B_1^m \equiv B\left(\mu_N^1, \frac{1}{m}\right), \ B_k^m \equiv B\left(\mu_N^k, \frac{1}{m}\right) \setminus \left[\cup_{j=1}^{k-1}B\left(\mu_N^j, \frac{1}{m}\right)\right], \ k = 2, ..., J.$$

Then, $B_1^m, ..., B_J^m$ are disjoint and $\cup_{k=1}^{j}B_k^m = \cup_{k=1}^{j}B\left(\mu_N^k, \frac{1}{m}\right)$ for all j. Consequently,

$$\nu\left(\cup_{k=1}^{J}B_k^m\right) = \nu\left(\cup_{k=1}^{J}B\left(\mu_N^k, \frac{1}{m}\right)\right) \geq 1 - 1/m. \tag{B.4}$$

It is possible to approximate

$$\nu\left(B_1^m\right)\delta_{\mu_N^1} + ... + \nu\left(B_J^m\right)\delta_{\mu_N^J}$$

by

$$\nu_m \equiv \beta_1^m\delta_{\mu_N^1} + ... + \beta_J^m\delta_{\mu_N^J},$$

where $\beta_j^m \in [0,1]\cap\mathbb{Q}$ is such that $\sum_{j=1}^{J}\beta_j^m = 1$ and

$$\sum_{j=1}^{J}\left|\nu\left(B_j^m\right) - \beta_j^m\right| < \frac{2}{m}.$$

7

Since rationals are dense in reals, such a choice of $\left\{\beta_j^m\right\}$ is always possible through an appropriate rescaling. Then, for each m, $\nu_m \in \mathcal{D}$.

To show that $\nu_m \Rightarrow \nu$, take a uniformly continuous bounded function f on $\Delta\Theta_1$. Then,

$$
\begin{aligned}
\left|\int f \mathrm{d}\nu_m - \int f \mathrm{d}\nu\right| &= && \left|\Sigma_{j=1}^J \beta_j^m f\left(\mu_N^j\right) - \int f \mathrm{d}\nu\right| \\
&\leq && \left|\Sigma_{j=1}^J \nu\left(B_j^m\right) f\left(\mu_N^j\right) - \int f \mathrm{d}\nu\right| && +\frac{2}{m}\sup_j \left|f\left(\mu_N^j\right)\right| \\
&\leq && \left|\int \Sigma_{j=1}^J f\left(\mu_N^j\right)\mathbf{1}_{\left\{B_j^m\right\}}\mathrm{d}\nu - \int f \mathrm{d}\nu\right| && +\frac{2}{m}\|f\|_\infty \\
&\leq && \left|\Sigma_{j=1}^J \int \left(f\left(\mu_N^j\right)-f\right)\mathbf{1}_{\left\{B_j^m\right\}}\mathrm{d}\nu\right| + \int f\mathbf{1}_{\left\{\left(\cup_{j=1}^J B_j^m\right)^c\right\}}\mathrm{d}\nu\Bigg| && +\frac{2}{m}\|f\|_\infty \\
&\leq && \left|\Sigma_{j=1}^J \sup_{\mu\in B_j^m}\left|f\left(\mu_N^j\right)-f\left(\mu\right)\right|\nu\left(B_j^m\right)\right. && +\|f\|_\infty\,\nu\left(\left(\cup_{j=1}^J B_j^m\right)^c\right)\Bigg| +\frac{2}{m}\|f\|_\infty.
\end{aligned}
$$

Each B_j^m is contained in a ball with radius $1/m$ around μ_N^j, and, thus, $d_p\left(\mu_N^j,\mu\right) < \frac{1}{m}$ for each $\mu \in B_j^m$. Because f is uniformly continuous, for every $\epsilon > 0$, there is a $\delta > 0$ such that whenever $d_p\left(\mu,\upsilon\right) < \delta$, $|f\left(\mu\right)-f\left(\upsilon\right)| < \epsilon$. Take some $\epsilon > 0$; then, for $m \geq 1/\delta$, $\left|f\left(\mu_N^j\right)-f\left(\mu\right)\right| < \epsilon$ for all $\mu \in B_j^m$ and all j. Then, from previous calculations

$$
\left|\int f \mathrm{d}\nu_m - \int f \mathrm{d}\nu\right| \leq \epsilon + \frac{1}{m}\|f\|_\infty + \frac{2}{m}\|f\|_\infty.
$$

Because f is bounded, the last two terms on the right-hand side can be made arbitrarily small by choosing m sufficiently large, whereas ϵ is arbitrary. Hence, $\int f \mathrm{d}\nu_m \to \int f \mathrm{d}\nu$ as $m \to \infty$, which implies that $\nu_m \Rightarrow \nu$. □

B.6 Preliminary Results for the Proof of Theorem 4

The proof of Theorem 4 demonstrates that the boundedness of $a^{*\prime}$ on (θ^*, s^*) follows from Π being connected and from Π and Γ being disjoint on $(s^*, 1]$. This section proves the two required intermediate results in Lemma B.3 and Lemma B.6, respectively.

Lemma B.3. Π *is connected.*

Proof. By contradiction, suppose that Π is not connected. Any non-connectedness in Π must be caused by some prospects being revealed, as in Figure B.2 . The revealed prospects, which lie in Γ, belong to Π. As a result, Π fails to be nondecreasing, thereby contradicting optimality. □

It remains to show that the optimal-prospect path Π is disjoint from the prospect set Γ on $(s^*, 1]$. Lemma B.6 accomplishes this task. The lemma's proof relies on a number of preliminary results.

It is convenient to cast the analysis in terms of a decreasing function $\lambda \equiv \tau^{-1}$, the inverse of the matching function τ, whenever this inverse exists. Function λ is defined on the interval
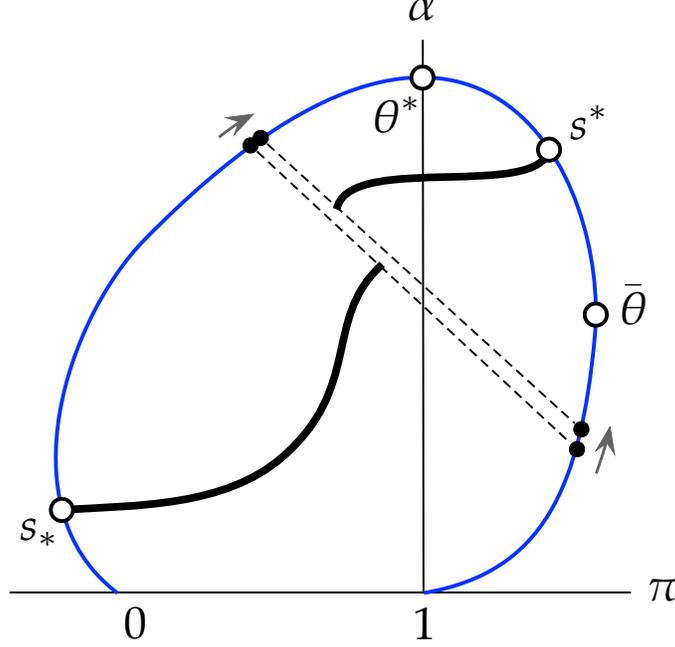
Figure B.2: The broken solid thick curve is a counterfactually non-connected Π. Prospects that lie on Γ's chords that are demarcated by the two pooling segments are revealed and belong to Π; Π fails to be nondecreasing, thereby contradicting optimality.

$[s^*, 1]$. Operating under the assumption that λ exists is justified because the goal is to rule out the situation in which λ is flat (and so, by implication, exists). Define $\gamma \equiv -\lambda'$.

For expositional purposes only, assume that the leader's type is distributed uniformly: $G(\theta_1) = \theta_1$.[8] After a change of variables, the seller's objective function (21) restricted to $[s^*, 1]$ (which is the interval of particular interest in our analysis) can be written in terms of γ and λ as

$$\int_{s^*}^{1} \Lambda(s, \lambda, \gamma)\, ds, \tag{B.5}$$

where

$$\Lambda(s, \lambda, \gamma) \equiv \frac{(\pi(s) + \gamma(s)\, \pi(\lambda(s)))\,(\alpha(s) + \gamma(s)\, \alpha(\lambda(s)))}{1 + \gamma(s)}.$$

Towards optimality, take some λ and perturb it towards $\lambda + \epsilon\eta$ for some $\epsilon \in \mathbb{R}$ and for some $\eta : [s^*, 1] \to \mathbb{R}$ with $\eta(s^*) = \eta(1) = 0$. The value of the perturbed objective function is denoted by

$$\Phi \equiv \int_{s^*}^{1} \Lambda(s, \lambda + \epsilon\eta, \gamma - \epsilon\eta')\, ds,$$

where the dependence of λ, γ, and η on s is implicit and will remain so as long as no ambiguity

---

[8]The Euler equation, in (B.7), and the subsequent arguments all hold for a general G. The unwieldy derivation for this general case is available upon request.

arises. The marginal benefit from perturbing λ in the direction of η is denoted by

$$J \equiv \frac{d\Phi}{d\epsilon}\Big|_{\epsilon=0} = \int_{s^*}^{1} \eta \left(\frac{\partial\Lambda}{\partial\lambda} - \frac{d}{ds}\frac{\partial\Lambda}{\partial\gamma} \right) ds.$$

If λ is optimal, any (feasible) perturbation η requires $J \leq 0$. If, in addition, an optimal λ is interior, the parenthetical term in the display above must be identically zero. The parenthetical term equated to zero becomes the Euler equation.

Computing $\partial\Lambda/\partial\lambda$, $\partial\Lambda/\partial\gamma$, and $(d/ds)(\partial\Lambda/\partial\gamma)$ and substituting the results into J in the display above gives

$$J = \int_{s^*}^{1} \eta \left[\frac{(\pi(\lambda) - \pi)(\alpha' - \gamma^2\alpha'(\lambda)) - (\pi' - \gamma^2\pi'(\lambda))(\alpha - \alpha(\lambda))}{(1+\gamma)^2} - \frac{2\gamma'(\pi(\lambda) - \pi)(\alpha - \alpha(\lambda))}{(1+\gamma)^3} \right] ds.$$

To interpret the expression for J graphically, in terms of the slope of Π, with some abuse of notation, write $\Pi \equiv \{(\pi^*(s), a^*(s)) \mid s \in [s^*, 1]\}$, where[9]

$$\pi^* = \frac{\pi + \gamma\pi(\lambda)}{1+\gamma} \qquad \text{and} \qquad a^* = \frac{\alpha + \gamma\alpha(\lambda)}{1+\gamma}.$$

Combining

$$\frac{d\pi^*}{ds} = \frac{(\pi' - \gamma^2\pi'(\lambda))(1+\gamma) - \gamma'(\pi - \pi(\lambda))}{(1+\gamma)^2}$$

and

$$\frac{da^*}{ds} = \frac{(\alpha' - \gamma^2\alpha'(\lambda))(1+\gamma) - \gamma'(\alpha - \alpha(\lambda))}{(1+\gamma)^2},$$

one obtains the slope of Π:

$$\frac{da^*}{d\pi^*} = \frac{da^*/ds}{d\pi^*/ds} = \frac{(\alpha' - \gamma^2\alpha'(\lambda))(1+\gamma) - \gamma'(\alpha - \alpha(\lambda))}{(\pi' - \gamma^2\pi'(\lambda))(1+\gamma) - \gamma'(\pi - \pi(\lambda))}.$$

Rearranging the expression for J and substituting $d\pi^*/ds$ and $da^*/d\pi^*$ gives

$$J = \int_{s^*}^{1} \eta\sigma \left(\frac{da^*}{d\pi^*} - \frac{\alpha(\lambda) - \alpha}{\pi - \pi(\lambda)} \right) ds, \qquad \text{where} \qquad \sigma \equiv \frac{\pi - \pi(\lambda)}{1+\gamma} \left(-\frac{d\pi^*}{ds} \right). \tag{B.6}$$

Now, from (B.6), we can extract the Euler equation. On $(s^*, 1)$, $\pi - \pi(\lambda) > 0$ and $\gamma \geq 0$ (and so $1 + \gamma > 0$). Because Π cannot be vertical, $d\pi^*/ds < 0$. As a result, $\sigma > 0$. Therefore, whenever Π is interior, it must satisfy the Euler equation:

$$\frac{da^*}{d\pi^*} = \frac{\alpha(\lambda) - \alpha}{\pi - \pi(\lambda)}. \tag{B.7}$$

---

[9]The abuse of notation here consists in reinterpreting $\pi^*(s)$ and $a^*(s)$ to be indexed by s in $[s^*, 1]$ (not $[s_*, s^*]$), which is consistent with the dummy variable in the integrand of the rewritten objective function (B.5) being defined on $[s^*, 1]$ (not $[s_*, s^*]$).

To summarize,

Lemma B.4. *When Π is interior, its slope, $da^*/d\pi^*$, obeys the Euler equation (B.7).*

Proof. The argument for the uniform G precedes the lemma's statement. The argument for a general, non-uniform, G is available upon request. □

Lemma B.5 is an auxiliary result that roughly says that, if Π were to touch Γ at a point in $[\bar{\theta}, 1]$, it would do so at an angle, instead of pasting smoothly at that point.

Lemma B.5. *Suppose that Π coincides with Γ at a single point $s' \in [\bar{\theta}, 1]$ and is disjoint from Γ on some subset $S \subset [\bar{\theta}, 1]$ adjacent to s' (i.e., $s' \in cl\,(S)$). Then, Π cannot be tangent to Γ at s'.*

Proof. By contradiction, suppose that Π coincides with, and is tangent to, Γ at a point $s' \in [\bar{\theta}, 1]$. Define a subset $S \subset (\bar{\theta}, 1)$ so that (i) for some $\varepsilon > 0$, S is either $(s', s' + \varepsilon)$ or $(s' - \varepsilon, s')$; (ii) on S, Π and Γ are disjoint; and (iii) ε is "sufficiently small" in the sense that will be made precise. That is, S is chosen so that, on S, Π is close to, but disjoint from, Γ. The disjointness of Π and Γ implies that, on S, Π satisfies the Euler equation (B.7). Differentiating the Euler equation on S gives

$$\frac{d}{ds}\left(\frac{da^*}{d\pi^*}\right) \;=\; \frac{d}{ds}\frac{\alpha(\lambda) - \alpha}{\pi - \pi(\lambda)} = -\frac{\alpha' + \gamma\alpha'(\lambda) + (da^*/d\pi^*)(\pi' + \gamma\pi'(\lambda))}{\pi - \pi(\lambda)}.$$

By requirement (iii) in the construction of S, Π is "close" to Γ on S, and so $a^* \approx \alpha$, which requires $\gamma \approx 0$ (by Bayes' rule, $a^* = (\alpha + \gamma\alpha(\lambda))/(1 + \gamma)$).[10] With this justification, we neglect the terms multiplied by γ. Then, $\pi - \pi(\lambda) > 0$ combined with $da^*/d\pi^* > 0$ (Π is increasing, by optimality), $\alpha' < 0$, and $\pi' < 0$ (by $s \in (\bar{\theta}, 1)$), imply

$$\frac{d}{ds}\left(\frac{da^*}{d\pi^*}\right) > 0,$$

which, in turn, implies that Π is concave.[11]

Thus, we have reached a contradiction because, at s', Π and Γ coincide and are tangent to each other, and yet, on S, Π is concave, whereas Γ is convex (by Condition 2). This situation is a geometric impossibility; it requires Π to exit the convex hull of Γ. Hence, Π and Γ cannot be tangent at s'. □

We are now ready to formulate and prove

Lemma B.6. *Suppose that Condition 3 holds. Then, the intersection of Π and Γ on $(s^*, 1]$ is empty.*

---

[10]The role of tangency in the contradiction hypothesis is to ensure that Π is "close" to Γ on S, and so $\gamma \approx 0$ indeed holds.

[11]As s rises, the induced prospect is moving southwest along Γ—or from right to left if one projects this movement on the horizontal axis. Hence, the sign in the criterion for concavity is positive, flipped from the customary. Formally, Π, a parametric curve, is concave if its second derivative is negative; that is, if $d^2 a^*/(d\pi^*)^2 = d\,(da^*/d\pi^*)/d\pi^* = (d\,(da^*/d\pi^*)/ds)/(d\pi^*/ds) < 0$. Because, on $[\bar{\theta}, 1] \subseteq [s^*, 1]$, $d\pi^*/ds < 0$, Π is concave if $d\,(da^*/d\pi^*)/ds > 0$.
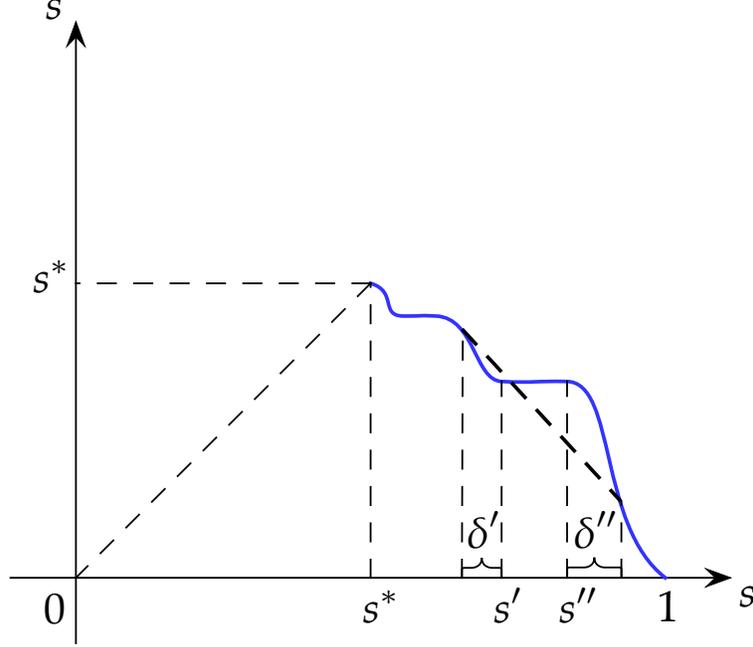
Figure B.3: The thick solid curve is an inverse matching function λ with a (counterfactually) flat segment on $[s', s'']$. The dashed segment is the perturbed (by η) inverse matching function $\hat{\lambda} \equiv \lambda + \eta$.

Proof. By contradiction, first suppose that Π and Γ have a nonempty intersection on $(s^*, \bar{\theta})$. Then, because Π and Γ coincide at s^*, and because Γ is decreasing on $(s^*, \bar{\theta})$, Π must have a decreasing segment, thereby contradicting optimality, which requires that Π be nondecreasing.

The remainder of the proof is concerned with showing that Π cannot intersect Γ on $[\bar{\theta}, 1]$. By way of contradiction, let $[s', s'']$ with $s'' \geq s'$ be the largest (lengthwise, $s'' - s'$) interval on which Π and Γ coincide in $[\bar{\theta}, 1]$. The argument goes through a list of cases with all the possible values for s' and s'' relative to $\bar{\theta}$ and 1.

$\underline{\text{Case 1}}$: $\bar{\theta} < s' < s'' < 1$.

Given λ, define an additive perturbation η parametrized by positive scalars δ' and δ''. This perturbation is illustrated in Figure B.3 . The scalars δ' and δ'', and the perturbation η are chosen so that (i) on $[s^*, s' - \delta'] \cup [s'' + \delta'', 1]$, $\eta = 0$; (ii) on $s \in (s' - \delta', s'' + \delta'') \subset (s^*, 1)$, η induces a $\hat{\lambda}$ that linearly interpolates between $\lambda(s' - \delta')$ and $\lambda(s'' + \delta'')$:

$$\hat{\lambda}(s) \equiv \eta(s) + \lambda(s) = \lambda(s'' + \delta'') + \frac{s'' + \delta'' - s}{s'' + \delta'' - (s' - \delta')} \left(\lambda(s' - \delta') - \lambda(s'' + \delta'')\right); \quad \text{(B.8)}$$

and (iii) δ' and δ'' are such that $s^* < s' - \delta'$, $s'' + \delta'' < 1$, Π and Γ are disjoint on $(s^*, s' - \delta') \cup (s'' + \delta'', 1)$,

$$\int_{s' - \delta'}^{s'' + \delta''} \eta \sigma ds = 0, \quad \text{(B.9)}$$

12

and η crosses zero only once (at a single point), from above.[12] Because σ, defined in (B.6), is positive, the single-crossing property of η implies that $\eta\sigma$, too, crosses zero only once (at a single point), from above.

The construction of η implies the inequality

$$J = \int_{s^*}^1 \eta\sigma \left(\frac{\mathrm{d}a^*}{\mathrm{d}\pi^*} - \frac{\alpha(\lambda) - \alpha}{\pi - \pi(\lambda)} \right) \mathrm{d}s$$
$$= \int_{s'}^{s''} \eta\sigma \left(\frac{\mathrm{d}a^*}{\mathrm{d}\pi^*} - \frac{\alpha(\lambda) - \alpha}{\pi - \pi(\lambda)} \right) \mathrm{d}s > 0. \tag{B.10}$$

Indeed, in the display above, the equality in the second line follows because, by construction, $\eta \equiv 0$ outside the interval $(s' - \delta', s'' + \delta'')$, and because the Euler equation holds on the set $(s' - \delta', s') \cup (s'', s'' + \delta'')$, so that the parenthetical term is zero on this set. To conclude the inequality in (B.10), note that, on (s', s''), the parenthetical term is decreasing in s because $\mathrm{d}a^*/\mathrm{d}\pi^*$ is decreasing in s (because, by hypothesis, Π coincides with Γ on $[s', s'']$ and, by Condition 2, Γ is convex), and because the term $(\alpha(\lambda) - \alpha) / (\pi - \pi(\lambda))$ is increasing in s (because, by hypothesis, λ does not vary with s; α and π both decrease in s; and, since pooling links are non-increasing, $\pi - \pi(\lambda)$ and $\alpha(\lambda) - \alpha$ are both positive). Then, because, by construction of η, $\eta\sigma$ crosses zero once and from above, (B.9) implies the positive sign in (B.10); indeed, the left-hand side of (B.10) puts larger weights on positive $\eta\sigma$'s, and smaller weights on negative $\eta\sigma$'s, than (B.9) does. Thus, when $s'' > s'$, the constructed η induces a profitable perturbation away from λ, and so the coincidence of Π and Γ on $[s', s'']$ is suboptimal.

Case 2: $\bar{\theta} < s' = s'' < 1$.

The Euler equation (B.7) describes the slope of Π arbitrarily close to s'. This slope is the same from whichever direction s' is approached. Thus, Π must be tangent to Γ at s'. Lemma B.5, however, rules out such a tangency; a contradiction is reached. Thus, the coincidence of Π and Γ at a single point s' is suboptimal.

Case 3: $\bar{\theta} \le s' < s'' = 1$.

Figure B.4 illustrates that if Π coincides with Γ on $[s', 1]$, then Π coincides with Γ also on a nonempty interval $[0, \theta_1]$, for some $\theta_1 > 0$. Such a Π fails to be nondecreasing at θ_1, thereby contradicting optimality. Therefore, the coincidence of Π and Γ on $[s', 1]$ is suboptimal.

Case 4: $\bar{\theta} < s' = s'' = 1$.

The Euler equation applies in a sufficiently small neighborhood of s'. Because Π lies in the convex hull of Γ, near s', Π must be at least as steep as Γ; that is, the slope of Π must be at least $\lim_{s\to 1} \alpha'(s) / \pi'(s)$. If this slope is nonzero, the pooling pattern implied by the Euler equation requires Π to coincide with Γ on $[0, \theta_1]$ for some $\theta_1 > 0$, thereby leading to a Π that fails to be nondecreasing and, thus, contradicting optimality. Figure B.5 illustrates the contradiction.

---

[12]Because $\sigma > 0$, by the Second Mean Value Theorem for Integrals, $\int_{s'-\delta'}^{s''+\delta''} \eta\sigma \mathrm{d}s = \eta(x) \int_{s'-\delta'}^{s''+\delta''} \sigma \mathrm{d}s$ for some $x \in (s' - \delta', s'' + \delta'')$. Because η is positive at first and then negative, it must cross zero at some point. Consequently, it is possible to find δ' and δ'' to satisfy $\eta(x) = 0$ and, hence, also (B.9).
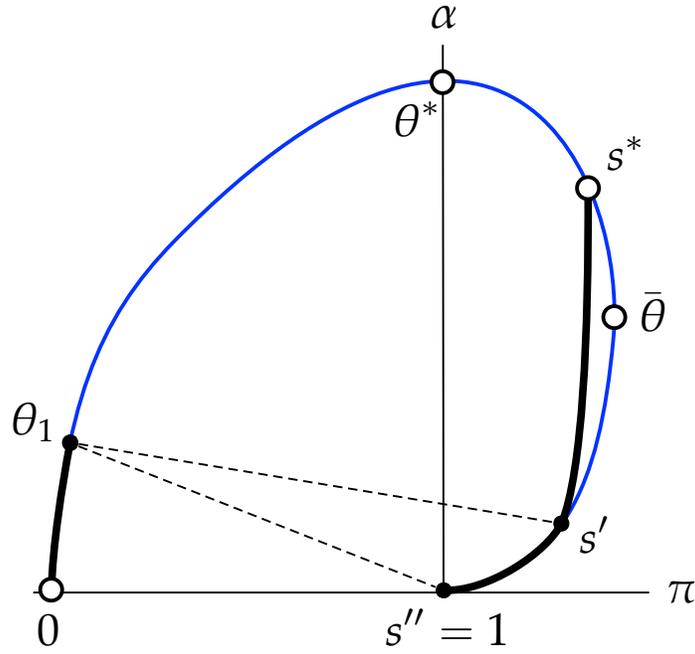
Figure B.4: If Π lies on Γ for $s \in [s', 1]$, then the pooling links must be nonincreasing as indicated, thereby implying that Π lies on Γ for $s \in [0, \theta_1]$, for some $\theta_1 > 0$. As a result, the implied Π fails to be nondecreasing, which contradicts optimality.
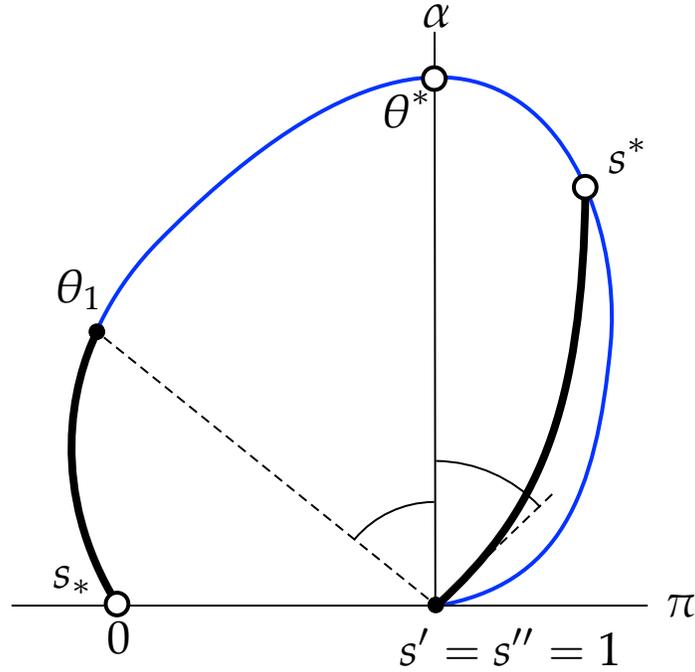

Figure B.5: The thick curves constitute Π. The two angles emanating from s' and marked by arcs are equal by the Euler equation. The dashed pooling link that connects θ_1 and s' suggests two possibilities: (i) $s_* = 0$ and Π coinciding with Γ on $[0, \theta_1]$ (shown); and (ii) $s_* > 0$ (not shown). In either case, Π fails to be nondecreasing, thereby contradicting optimality.
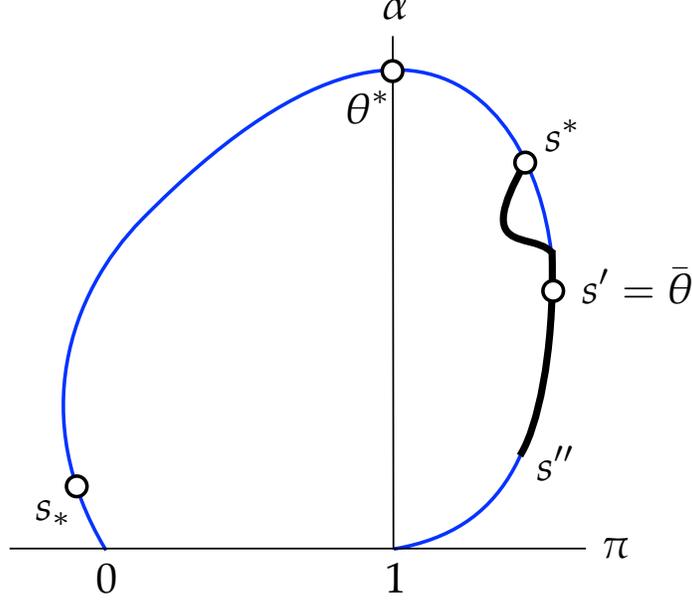
Figure B.6: The thick curve denotes the segment of Π invoked in the argument. If $s^* < s' = \bar{\theta}$, then Π must bend backwards to reach s^*.

The slope is nonzero if

$$\lim_{s \to 1} \frac{\alpha'(s)}{\pi'(s)} > 0,$$

which holds by Condition 3, in the lemma's hypothesis. Thus, the coincidence of Π and Γ at 1 is suboptimal.

Case 5: $\bar{\theta} = s' < s'' \leq 1$.

Note that $s' = \bar{\theta}$ implies that $s^* = \bar{\theta}$. Indeed, if $s' = \bar{\theta}$ and, by contradiction, $s^* < \bar{\theta}$, then Π must necessarily be decreasing somewhere on $[s^*, \bar{\theta}]$ (see Figure B.6), thereby contradicting optimality.

However, it will be shown that $s^* = \bar{\theta}$ is not possible either. By contradiction, suppose that $s^* = \bar{\theta}$, as in Figure B.7 . If $s^* = \bar{\theta} = s'$ and Π coincides with Γ on $[s', s'']$, then Π must also coincide with Γ on an interval $[\theta_1, s^*]$ for some $\theta_1 < s^*$. This geometric arrangement contradicts Π being nondecreasing because Γ is decreasing on $[\theta^*, \bar{\theta}]$. Thus, the coincidence of Π and Γ on $[s', s'']$ is suboptimal.

Case 6: $\bar{\theta} = s' = s'' < 1$.

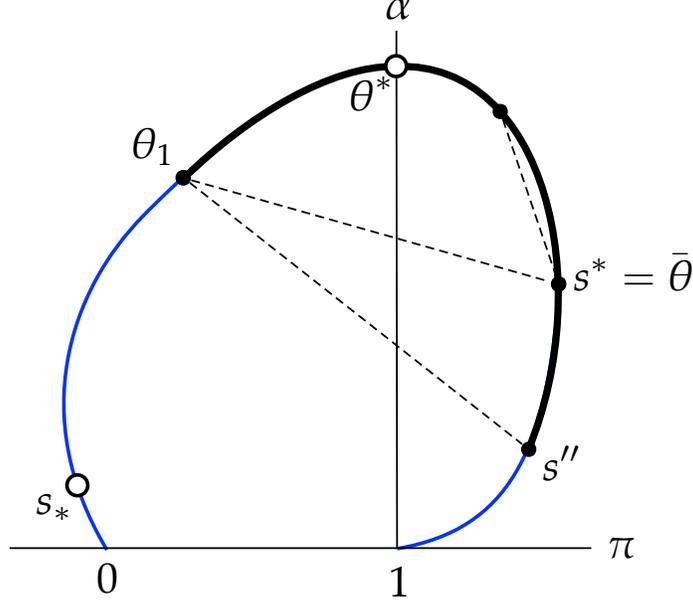As in the preceding case, $s' = \bar{\theta}$ requires $s^* = \bar{\theta}$.

15

Figure B.7: The thick curve denotes the segment of Π invoked in the argument. If $s' = s^* = \bar\theta < s''$, then all prospects in $[s^*, s'']$ must be connected to the same prospect θ_1 by nonincreasing pooling segments. Then, by $s^* = \bar\theta$, Π must coincide with Γ on $(\theta_1, \bar\theta)$. In particular, a segment of Π, on $(\theta^*, \bar\theta)$, must be downward-sloping, which contradicts optimality.

By the Euler equation, (B.7),

$$
\begin{aligned}
\lim_{s \downarrow s^*} \frac{\mathrm{d}a^*}{\mathrm{d}\pi^*} &= \lim_{s \downarrow s^*} \frac{\alpha\left(\lambda\left(s\right)\right) - \alpha\left(s\right)}{\pi\left(s\right) - \pi\left(\lambda\left(s\right)\right)} \\
&= -\lim_{s \downarrow s^*} \frac{\gamma\left(s\right)\alpha'\left(\lambda\left(s\right)\right) + \alpha'\left(s\right)}{\pi'\left(s\right) + \gamma\left(s\right)\pi'\left(\lambda\left(s\right)\right)} \\
&= -\lim_{s \downarrow s^*} \frac{\alpha'\left(s\right)\left(1 + \gamma\left(s\right)\right)}{\pi'\left(s\right)\left(1 + \gamma\left(s\right)\right)} \\
&= -\lim_{s \downarrow s^*} \frac{\alpha'\left(s\right)}{\pi'\left(s\right)},
\end{aligned}
\tag{B.11}
$$

where the second equality is by L'Hôpital's rule, the third one uses the smoothness of the prospect set (i.e., $\lim_{s \downarrow s^*}\left(\alpha'\left(s\right) - \alpha'\left(\lambda\left(s\right)\right)\right) = 0$) and the continuity of λ, and the fourth one uses $1 + \gamma\left(s^*\right) > 0$. When $s^* = \bar\theta$, equation (B.11) implies that

$$
\lim_{s \downarrow s^* = \bar\theta} \frac{\mathrm{d}a^*}{\mathrm{d}\pi^*} = -\lim_{s \downarrow s^* = \bar\theta} \frac{\alpha'}{\pi'} = \infty;
$$

that is, at $s^* = \bar\theta$, Π is tangent to Γ. Lemma B.5 rules out such a tangency; a contradiction is reached. Figure B.8 illustrates the contradiction. Thus, the coincidence of Π and Γ at $\bar\theta$ is suboptimal. $\square$
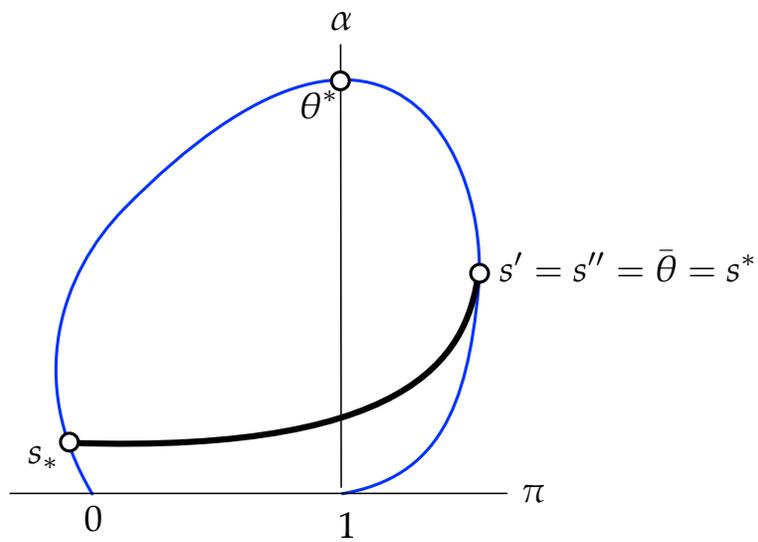
16

Figure B.8: The thick curve is Π. If $s' = s'' = \bar{\theta}$, it must be that $s^* = \bar{\theta}$, and that Π has an infinite slope at s^*, which contradicts optimality.