Incomplete Information Games with Ambiguity Averse Players: Online Appendix

by

Eran Hanany, Peter Klibanoff and Sujoy Mukerji

American Economic Journal: Microeconomics

ONLINE APPENDIX: PROOFS FOR SECTIONS III-IV AND FURTHER ANALYSIS

B1. Proofs of results in Section III.A

This subsection contains an example and formal results on comparative statics in ambiguity aversion.

EXAMPLE: NEW STRATEGIC BEHAVIOR IN EQUILIBRIUM. — We present a 3-player game, with incomplete information about player 1, in which a path of play can occur as part of an SEA when players 2 and 3 are sufficiently ambiguity averse, but never occurs as part of even an ex-ante equilibrium if we modify the game by making players 2 and 3 ambiguity neutral (expected utility). Furthermore, under the SEA we construct, player 1 achieves a higher expected payoff than under any ex-ante equilibrium of the game with ambiguity neutral players, and even outside the convex hull of such ex-ante equilibrium payoffs. The game is depicted in Figure B1.

There are three players: 1,2 and 3. First, it is determined whether player 1 is of type I or type II and 1 observes her own type. Players 2 and 3 have only one type, so there is complete information about them. The payoff triples in Figure B1 describe vNM utility payoffs given players' actions and player 1's type (i.e., (u_1, u_2, u_3) means that player i receives u_i). Players 2 and 3 have ambiguity about player 1's type and have smooth ambiguity preferences with an associated $\phi_2 = \phi_3 = \phi$ and $\mu_2 = \mu_3 = \mu$. Player 1 also has smooth ambiguity preferences, but nothing in what follows depends on either ϕ_1 or μ_1. Player 1's first and only move in the game is to choose between action P(lay) which leads to players 2 and 3 playing a simultaneous move game in which their payoffs depend on 1's type, and action Q(uit), which ends the game.[24]

PROPOSITION 7: *Suppose players* 2 *and* 3 *are ambiguity neutral and have a common belief* μ*. There is no ex-ante equilibrium such that player* 1 *plays* P *with positive probability.*

PROOF OF PROPOSITION 7:

Observe that player 1 is willing ex-ante to play P with positive probability if and only if, after the play of P, (U, R) will be played with probability at least $\frac{1}{2}$. Suppose there

[24] Note that to eliminate any possible effects of varying players' risk aversion, think of the playoffs being generated using lotteries over two "physical" outcomes, the better of which has utility u normalized to $5/2$ and the worse of which has u normalized to 0. So, for example, the payoff 1 can be thought of as generated by the lottery giving the better outcome with probability $2/5$ and the worse outcome with probability $3/5$.

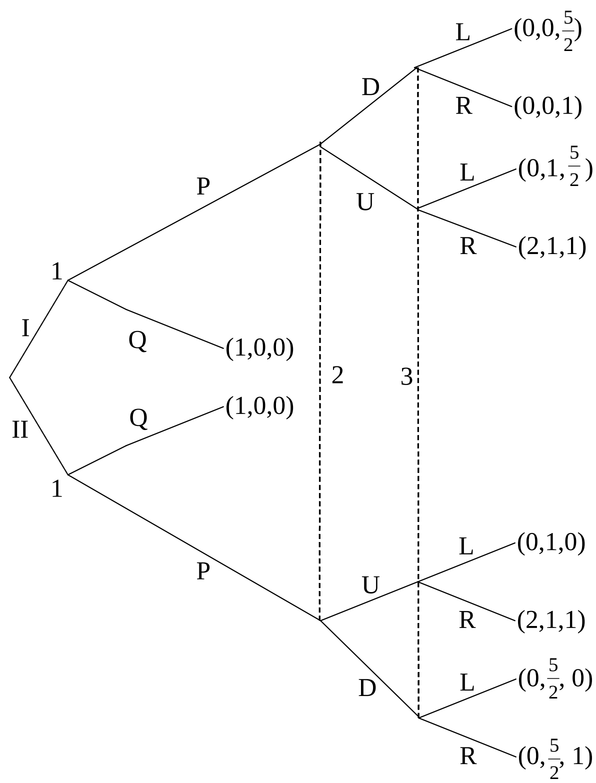

FIGURE B1. NEW EQUILIBRIUM BEHAVIOR WITH AMBIGUITY AVERSION is an ex-ante equilibrium, σ, in which P is played with positive probability. Let p_I and p_{II} denote the probabilities according to σ that types I and II, respectively, of player 1 play P. Then player 2 is finds it optimal to play U with positive probability if and only if

$$p_I \sum_{\pi \in \Delta(\Theta)} \mu(\pi)(\pi(I)) + p_{II} \sum_{\pi \in \Delta(\Theta)} \mu(\pi)(1-\pi(I)) \geq \frac{5}{2} p_{II} \sum_{\pi \in \Delta(\Theta)} \mu(\pi)(1-\pi(I))$$

which is equivalent to

$$\text{(B1)} \qquad p_I \sum_{\pi \in \Delta(\Theta)} \mu(\pi)(\pi(I)) \geq \frac{3}{2} p_{II} \sum_{\pi \in \Delta(\Theta)} \mu(\pi)(1-\pi(I)).$$

Similarly, player 3 finds it optimal to play R with positive probability if and only if

$$p_I \sum_{\pi \in \Delta(\Theta)} \mu(\pi)(\pi(I)) + p_{II} \sum_{\pi \in \Delta(\Theta)} \mu(\pi)(1-\pi(I)) \geq \frac{5}{2} p_I \sum_{\pi \in \Delta(\Theta)} \mu(\pi)(\pi(I))$$

which is equivalent to

$$\text{(B2)} \qquad p_I \sum_{\pi \in \Delta(\Theta)} \mu(\pi)(\pi(I)) \leq \frac{2}{3} p_{II} \sum_{\pi \in \Delta(\Theta)} \mu(\pi)(1-\pi(I)).$$

Since (B1) and (B2) cannot both be satisfied when $p_I + p_{II} > 0$ (i.e., P is played with positive probability), σ must specify that (P, U, R) is never realized as part of a history. This implies that player 1 has an ex-ante profitable deviation to the strategy of always playing Q, contradicting the assumption that σ is an ex-ante equilibrium. ■

Since σ being part of a sequentially optimal (σ, ν) implies σ is an ex-ante equilibrium, Proposition 7 immediately implies that none of the stronger concepts such as SEA, PBE or sequential equilibrium can admit the play of P with positive probability under ambiguity neutrality. The next result shows that the situation changes dramatically under sufficient ambiguity aversion.

PROPOSITION 8: *There exist ϕ and μ (e.g., $\phi(x) \equiv -e^{-x}$ and $\mu(\pi_0) = \mu(\pi_1) = \frac{1}{2}$, where $\pi_0(I) = 1$ and $\pi_1(I) = 0$) such that in an SEA both types of player 1 play P with probability 1, and (U, R) is played with probability greater than $\frac{1}{2}$.*

PROOF OF PROPOSITION 8:

Let μ put probability $\frac{1}{2}$ on π_0 and $\frac{1}{2}$ on π_1, where $\pi_0(I) = 1$ and $\pi_1(I) = 0$.[25] Let $\phi(x) \equiv -e^{-x}$.[26] Let σ be a strategy profile specifying that both types of player 1 play

[25]The degeneracy of the π in the support of μ is not necessary for the argument to go through – it merely shortens some calculations and reduces the ambiguity aversion required.

[26]Any more concave ϕ will also work, as will any ϕ more concave than $-e^{-\alpha x}$ for $\alpha = \frac{-4(\ln(2/3))}{5(2-\sqrt{2})} \approx 0.554$.

P with probability 1, player 2 plays U with probability λ^* if given the move and player 3 plays R with probability λ^* if given the move, where $\lambda^* = 1 - \frac{2}{5}\ln(3/2)$. Notice that according to σ, (P, U, R) occurs with probability $\left(1 - \frac{2}{5}\ln(3/2)\right)^2 > \frac{7}{10}$. Observe that player 1 strictly prefers ex-ante to play P with probability 1 for both types if and only if, after the play of P, (U, R) will be played with probability greater than $\frac{1}{2}$. The same is true for each type of player 1 after her type is realized as well. Player 2 ex-ante chooses the probability, $\lambda \in [0, 1]$, with which to play U if given the move to maximize

$$-\frac{1}{2}e^{-\lambda} - \frac{1}{2}e^{-(\lambda+\frac{5}{2}(1-\lambda))}.$$

One can verify that the maximum is reached at $\lambda = \lambda^*$. Similarly, player 3 ex-ante chooses the probability, $\lambda \in [0, 1]$, with which to play R if given the move to maximize

$$-\frac{1}{2}e^{-(\lambda+\frac{5}{2}(1-\lambda))} - \frac{1}{2}e^{-\lambda}$$

which is again maximized at $\lambda = \lambda^*$.

Now consider the following sequence of completely mixed strategies with limit σ: σ^k has each type of player 1 play P with probability $1 - \frac{1}{k+1}$, and leaves the strategies otherwise the same as in σ. By Lemma 4, Theorem 11 provides a formula (A12) for an interim belief system ν satisfying smooth rule consistency. Recall that player 1 learns the parameter at the beginning of the game. Thus we need only be concerned with the beliefs of players 2 and 3. Therefore (σ, ν) satisfies smooth rule consistency. It remains to show (σ, ν) is sequentially optimal. Since

$$\frac{\nu_{2,\{I,II\}\times\{P\}}(\delta_{(I,P)})}{\nu_{2,\{I,II\}\times\{P\}}(\delta_{(II,P)})} = \frac{\nu_{3,\{I,II\}\times\{P\}\times\{U,D\}}(\delta_{(II,P)})}{\nu_{3,\{I,II\}\times\{P\}\times\{U,D\}}(\delta_{(I,P)})} = \frac{\frac{\phi_i'(\lambda^*)}{\phi_i'(\lambda^*)}\frac{1}{2}}{\frac{\phi_i'\left(\lambda^*+\frac{5}{2}(1-\lambda^*)\right)}{\phi_i'\left(\lambda^*+\frac{5}{2}(1-\lambda^*)\right)}\frac{1}{2}} = 1,$$

σ remains optimal for players 2 and 3 following the play of P given ν. Thus, (σ, ν) is sequentially optimal. It is therefore an SEA. ■

As the proof of Proposition 8 mentions, the example μ is chosen for simplicity, and degeneracy of the measures in its support is not necessary for the result.

FORMAL COMPARATIVE STATICS IN AMBIGUITY AVERSION. —

NOTATION 4: *For a game $\Gamma = (N, H, (\mathcal{I}_i)_{i\in N}, (\mu_i)_{i\in N}, (u_i, \phi_i)_{i\in N})$, let $E_\Gamma((\hat{\mu}_i)_{i\in N}, (\hat{\phi}_i)_{i\in N})$ denote the set of all ex-ante equilibria of the game $\hat{\Gamma} = (N, H, (\mathcal{I}_i)_{i\in N}, (\hat{\mu}_i)_{i\in N}, (u_i, \hat{\phi}_i)_{i\in N})$ differing from Γ only in ambiguity aversions and beliefs. Let $Q_\Gamma((\hat{\mu}_i)_{i\in N}, (\hat{\phi}_i)_{i\in N})$ denote the analogous set of sequentially optimal strategy profiles and $S_\Gamma((\hat{\mu}_i)_{i\in N}, (\hat{\phi}_i)_{i\in N})$ denote the analogous set of SEA strategy profiles.*

NOTATION 5: *Denote the identity function by ι.*

PROOF OF THEOREM 5:

We show that there exists a game Γ and $(\hat{\phi}_i)_{i\in N}$ such that

$$E_\Gamma((\mu_i)_{i\in N},(\hat{\phi}_i)_{i\in N})\cap E_\Gamma((\mu_i)_{i\in N},(\iota)_{i\in N})=\emptyset.$$

Modify the example in Figure B1 by removing the action Q for Player 1. For each player i, let $\mu_i=\mu$ where μ puts probability $\frac{1}{2}$ on π_0 and $\frac{1}{2}$ on π_1, where $\pi_0(I)=1$ and $\pi_1(I)=0$, and let $\hat{\phi}_i(x)=-e^{-x}$ for all i. With these preferences, the unique ex-ante equilibrium has player 2 play U with probability λ^* and player 3 play R with probability λ^*, where $\lambda^*=1-\frac{2}{5}\ln(3/2)$. In contrast, when $\phi_i=\iota$ for all i, using the same μ, then the unique ex-ante equilibrium has player 2 playing D with probability 1 and player 3 play L with probability 1. ■

Examination of the proof shows that, fixing beliefs, not only are the equilibrium strategies distinct under ambiguity aversion compared to ambiguity neutrality, but it can also be that the strategies under ambiguity aversion generate paths of play that do not occur in equilibrium under ambiguity neutrality. An analogue of Theorem 5 is true for sequential optima, SEA and any other refinement of ex-ante equilibria as well, as they are all ex-ante equilibria. Thus, with fixed beliefs, change in ambiguity aversion can impact the set of equilibrium strategies and realized play.

Further examination of the proof shows that ambiguity aversion continues to affect the equilibrium set even if we impose *common beliefs* (i.e., $\mu_i=\mu$ for all players i). The next result addresses the question of whether ambiguity aversion plus the assumption of common beliefs has equilibrium implications that are different from ambiguity neutrality plus the assumption of common beliefs. It shows that, in this case, ambiguity aversion always weakly expands the set of equilibria compared to ambiguity neutrality and may do so strictly:

PROOF OF THEOREM 6:

We show: For all games Γ and $(\hat{\phi}_i)_{i\in N}$, $\bigcup_{\hat{\mu}} E_\Gamma((\hat{\mu})_{i\in N},(\hat{\phi}_i)_{i\in N})\supseteq\bigcup_{\hat{\mu}} E_\Gamma((\hat{\mu})_{i\in N},(\iota)_{i\in N})$, and the same holds when Q or S replaces E; moreover, there exists a game Γ and $(\hat{\phi}_i)_{i\in N}$ such that all these inclusions are strict and some of the new equilibrium strategies induce new paths of play. That $\bigcup_{\hat{\mu}} E_\Gamma((\hat{\mu})_{i\in N},(\hat{\phi}_i)_{i\in N})\supseteq\bigcup_{\hat{\mu}} E_\Gamma((\hat{\mu})_{i\in N},(\iota)_{i\in N})$ follows by considering only degenerate beliefs on the left-hand side and choosing them to have the same reduced measure as the right-hand side beliefs. $\bigcup_{\hat{\mu}} Q_\Gamma((\hat{\mu})_{i\in N},(\hat{\phi}_i)_{i\in N})\supseteq\bigcup_{\hat{\mu}} Q_\Gamma((\hat{\mu})_{i\in N},(\iota)_{i\in N})$ follows using the same construction and additionally taking the left-hand side updated beliefs at each information set to be degenerate with the same reduced measure as the right-hand side updated beliefs at the corresponding information set and noting that this preserves optimality at each information set. $\bigcup_{\hat{\mu}} S_\Gamma((\hat{\mu})_{i\in N},(\hat{\phi}_i)_{i\in N})\supseteq\bigcup_{\hat{\mu}} S_\Gamma((\hat{\mu})_{i\in N},(\iota)_{i\in N})$ follows using the same construction as for sequential optima, observing that the left-hand side degenerate beliefs satisfy smooth rule consistency since the right-hand side beliefs do so. As shown by Propositions 1 and 3, in the running example with sufficient ambiguity aversion and $x\leq 0.5$ the inclusion is strict and the new strategies generate new paths of play. ■

Next, in the constructive proof of Theorem 7, we show that beliefs $\hat{\mu}_i$ and $\hat{\nu}_{i,I_i}$ that support a given equilibrium profile σ are related to the beliefs μ_i and ν_{i,I_i} in the game with the original ambiguity aversion(s) by the formulae in (B3) and (B4) where the ϕ_i are the original and $\hat{\phi}_i$ the new specifications of ambiguity aversions.

PROOF OF THEOREM 7:

Fix a game Γ. We show: For all $(\hat{\phi}_i)_{i\in N}$,

$$\bigcup_{(\hat{\mu}_i)_{i\in N}} E_\Gamma((\hat{\mu}_i)_{i\in N}, (\hat{\phi}_i)_{i\in N}) = \bigcup_{(\hat{\mu}_i)_{i\in N}} E_\Gamma((\hat{\mu}_i)_{i\in N}, (\phi_i)_{i\in N}),$$

and the same holds when Q or S replaces E.

Let $\sigma \in E_\Gamma((\mu_i)_{i\in N}, (\phi_i)_{i\in N})$. Ex-ante equilibrium is equivalent to ex-ante optimality for all players i of σ_i according to i's preferences given σ_{-i}. This ex-ante optimality is equivalent to $\sigma'_i = \sigma_i$ maximizing (11) with respect to σ'_i. Let $\hat{\mu}_i$ be the probability measure such that

$$\hat{\mu}_i(\pi) \propto \frac{\phi'_i\left(\sum_h u_i(h) p_\sigma(h|h^0)\pi(h^0)\right)}{\hat{\phi}'_i\left(\sum_h u_i(h) p_\sigma(h|h^0)\pi(h^0)\right)} \mu_i(\pi). \tag{B3}$$

Observe that replacing ϕ_i and μ_i with $\hat{\phi}_i$ and $\hat{\mu}_i$ leaves the effective beliefs at σ, and so also (11), unchanged up to proportionality. Thus σ_i is ex-ante optimal for player i given $\hat{\phi}_i$, $\hat{\mu}_i$ and σ_{-i}. As this is true for each player i, $\sigma \in E_\Gamma((\hat{\mu}_i)_{i\in N}, (\hat{\phi}_i)_{i\in N})$.

Turn now to sequentially optimal strategy profiles. Suppose $\sigma \in Q_\Gamma((\mu_i)_{i\in N}, (\phi_i)_{i\in N})$ and ν is an interim belief system such that (σ, ν) is sequentially optimal for Γ. Let $\hat{\mu}_i$ be defined as in (B3) and for each I_i, $\hat{\nu}_{i,I_i}$ be the probability measure such that

$$\hat{\nu}_{i,I_i}(\pi) \propto \frac{\phi'_i\left(\sum_{h|h^{s(I_i)}\in I_i} u_i(h) p_\sigma(h|h^{s(I_i)})\pi(h^{s(I_i)})\right)}{\hat{\phi}'_i\left(\sum_{h|h^{s(I_i)}\in I_i} u_i(h) p_\sigma(h|h^{s(I_i)})\pi(h^{s(I_i)})\right)} \nu_{i,I_i}(\pi). \tag{B4}$$

By the argument in the ex-ante equilibrium part of this proof, $\sigma \in E_\Gamma((\hat{\mu}_i)_{i\in N}, (\hat{\phi}_i)_{i\in N})$. Optimality for player i at I_i as a function of σ'_i is equivalent (see (5) and (6)) to $\sigma'_i = \sigma_i$ maximizing

(B5)

$$\sum_{\pi\in\Delta(I_i)} \left(\sum_{h|h^{s(I_i)}\in I_i} u_i(h) p_{(\sigma'_i,\sigma_{-i})}(h|h^{s(I_i)})\pi(h^{s(I_i)})\right) \phi'_i \left(\sum_{h|h^{s(I_i)}\in I_i} u_i(h) p_\sigma(h|h^{s(I_i)})\pi(h^{s(I_i)})\right) \nu_{i,I_i}(\pi)$$

with respect to σ'_i. Observe that replacing ϕ_i and ν_{i,I_i} with $\hat{\phi}_i$ and $\hat{\nu}_{i,I_i}$ leaves (B5) unchanged up to proportionality. This is true for each player i and I_i. Thus, $(\sigma, \hat{\nu})$ is sequentially optimal in $\hat{\Gamma}$.

We now extend the argument to SEA. Suppose $\sigma \in S_\Gamma((\mu_i)_{i\in N}, (\phi_i)_{i\in N})$ and ν is an interim belief system such that (σ, ν) is an SEA for Γ (with corresponding sequence of completely mixed strategy profiles $\{\sigma^k\}_{k=1}^\infty$). By Theorem 11, ν satisfies (A12) using

$\{\sigma^k\}_{k=1}^{\infty}$. As above, let $\hat{\mu}_i$ be as in (B3) and for each I_i, $\hat{\nu}_{i,I_i}$ be defined as in (B4). By our previous arguments, $(\sigma, \hat{\nu})$ is sequentially optimal in $\hat{\Gamma}$. It remains to show that $(\sigma, \hat{\nu})$ satisfies smooth rule consistency in $\hat{\Gamma}$. Observe that replacing ϕ_i, μ_i and ν_{i,I_i} with $\hat{\phi}_i$, $\hat{\mu}_i$ and $\hat{\nu}_{i,I_i}$ in (A12) preserves its validity. Thus, by Theorem 11, $\left(\sigma, \hat{\nu}\right)$ satisfies smooth rule consistency in $\hat{\Gamma}$. Therefore $(\sigma, \hat{\nu})$ is an SEA of $\hat{\Gamma}$.

The above arguments have shown $E_\Gamma((\mu_i)_{i\in N}, (\phi_i)_{i\in N}) \subseteq \bigcup_{(\hat{\mu}_i)_{i\in N}} E_\Gamma((\hat{\mu}_i)_{i\in N}, (\hat{\phi}_i)_{i\in N})$, $Q_\Gamma((\mu_i)_{i\in N}, (\phi_i)_{i\in N}) \subseteq \bigcup_{(\hat{\mu}_i)_{i\in N}} Q_\Gamma((\hat{\mu}_i)_{i\in N}, (\hat{\phi}_i)_{i\in N})$ and $S_\Gamma((\mu_i)_{i\in N}, (\phi_i)_{i\in N}) \subseteq \bigcup_{(\hat{\mu}_i)_{i\in N}} S_\Gamma((\hat{\mu}_i)_{i\in N}, (\hat{\phi}_i)_{i\in N})$. Applying these arguments twice (the second time with the roles of ϕ_i and $\hat{\phi}_i$ interchanged), we obtain that, for any game, the union over all beliefs of the set of equilibrium strategy profiles is independent of ambiguity aversion. ■

Finally, turn to the case of pure strategies and only pure strategy deviations as in Battigalli et al. (2015). Modify the equilibrium set notation to restrict attention to pure strategies:

DEFINITION 12: *For a game $\Gamma = (N, H, (\mathcal{I}_i)_{i\in N}, (\mu_i)_{i\in N}, (u_i, \phi_i)_{i\in N})$, let $\tilde{E}_\Gamma((\hat{\mu}_i)_{i\in N}, (\hat{\phi}_i)_{i\in N})$ be the set of all ex-ante equilibria with respect to pure strategies of a game $\hat{\Gamma} = (N, H, (\mathcal{I}_i)_{i\in N}, (\hat{\mu}_i)_{i\in N}, (u_i, \hat{\phi}_i)_{i\in N})$ differing from Γ only in ambiguity aversions and beliefs. Let $\tilde{Q}_\Gamma((\hat{\mu}_i)_{i\in N}, (\hat{\phi}_i)_{i\in N})$ and $\tilde{S}_\Gamma((\mu_i)_{i\in N}, (\phi_i)_{i\in N})$ be the analogous respective sets of sequentially optimal and SEA strategy profiles with respect to pure strategies.*

PROOF OF THEOREM 8:

Fix a game Γ. We show that: For all $(\hat{\phi}_i)_{i\in N}$ such that, for each i, $\hat{\phi}_i$ is at least as concave as ϕ_i, $\tilde{E}_\Gamma((\mu_i)_{i\in N}, (\phi_i)_{i\in N}) \subseteq \bigcup_{(\hat{\mu}_i)_{i\in N}} \tilde{E}_\Gamma((\hat{\mu}_i)_{i\in N}, (\hat{\phi}_i)_{i\in N})$, and the same holds when $\tilde{Q}$ or $\tilde{S}$ replaces $\tilde{E}$. There exists a game Γ and $(\hat{\phi}_i)_{i\in N}$ such that for each i, $\hat{\phi}_i$ is at least as concave as ϕ_i, all these inclusions are strict and some of the new equilibrium strategies induce new paths of play.

Suppose $\varsigma \in \tilde{E}_\Gamma((\mu_i)_{i\in N}, (\phi_i)_{i\in N})$, for each i, $\hat{\phi}_i = \chi_i(\phi_i)$ for some increasing, differentiable and concave χ_i (note that differentiability of χ_i is implied by the continuous differentiability of $\hat{\phi}_i$ in the class of games considered in this paper) and $\hat{\mu}_i$ is the probability measure such that

$$\hat{\mu}_i(\pi) \propto \frac{\mu_i(\pi)}{\chi_i'\left(\phi_i\left(\sum_h u_i(h) p_\varsigma(h|h^0)\pi(h^0)\right)\right)}. \tag{B6}$$

By definition of $\tilde{E}_\Gamma((\mu_i)_{i\in N}, (\phi_i)_{i\in N})$, for each i and each ς_i',

(B7)

$$\sum_\pi \phi_i\left(\sum_h u_i(h) p_\varsigma(h|h^0)\pi(h^0)\right)\mu_i(\pi) \geq \sum_\pi \phi_i\left(\sum_h u_i(h) p_{(\varsigma_i', \varsigma_{-i})}(h|h^0)\pi(h^0)\right)\mu_i(\pi).$$

Since χ_i is increasing, differentiable and concave, for each π,

$$\begin{aligned}
& \chi_i\left(\phi_i\left(\sum_h u_i(h)p_\varsigma(h|h^0)\pi(h^0)\right)\right)-\chi_i\left(\phi_i\left(\sum_h u_i(h)p_{(\varsigma_i',\varsigma_{-i})}(h|h^0)\pi(h^0)\right)\right)\\
\geq\ & \chi_i'\left(\phi_i\left(\sum_h u_i(h)p_\varsigma(h|h^0)\pi(h^0)\right)\right)\\
& \cdot\left[\phi_i\left(\sum_h u_i(h)p_\varsigma(h|h^0)\pi(h^0)\right)-\phi_i\left(\sum_h u_i(h)p_{(\varsigma_i',\varsigma_{-i})}(h|h^0)\pi(h^0)\right)\right].
\end{aligned}$$

Thus, dividing both sides by $\chi_i'\left(\phi_i\left(\sum_h u_i(h)p_\varsigma(h|h^0)\pi(h^0)\right)\right)$ and taking the expectation with respect to μ_i yields

$$\begin{aligned}
& \sum_\pi \chi_i\left(\phi_i\left(\sum_h u_i(h)p_\varsigma(h|h^0)\pi(h^0)\right)\right)-\chi_i\left(\phi_i\left(\sum_h u_i(h)p_{(\varsigma_i',\varsigma_{-i})}(h|h^0)\pi(h^0)\right)\right)\hat{\mu}_i(\pi)\\
\geq\ & \sum_\pi\left[\phi_i\left(\sum_h u_i(h)p_\varsigma(h|h^0)\pi(h^0)\right)-\phi_i\left(\sum_h u_i(h)p_{(\varsigma_i',\varsigma_{-i})}(h|h^0)\pi(h^0)\right)\right]\mu_i\geq 0,
\end{aligned}$$

where the last inequality follows from B7. Since this is true for each i and each ς_i', $\varsigma\in\tilde{E}_\Gamma((\hat{\mu}_i)_{i\in N},(\hat{\phi}_i)_{i\in N})$. This shows $\tilde{E}_\Gamma((\mu_i)_{i\in N},(\phi_i)_{i\in N})\subseteq\bigcup_{(\hat{\mu}_i)_{i\in N}}\tilde{E}_\Gamma((\hat{\mu}_i)_{i\in N},(\hat{\phi}_i)_{i\in N})$.

Turn now to the part of the theorem about sequentially optimal strategy profiles. Suppose $\varsigma\in\tilde{Q}_\Gamma((\mu_i)_{i\in N},(\phi_i)_{i\in N})$ and ν is an interim belief system such that (ς,ν) is sequentially optimal for Γ with respect to pure strategies. Further suppose that for each i, $\hat{\phi}_i=\chi_i(\phi_i)$ for some increasing, differentiable and concave χ_i, $\hat{\mu}_i$ is defined as in (B6), and for each I_i, $\hat{\nu}_{i,I_i}$ is the probability measure such that

$$\hat{\nu}_{i,I_i}(\pi)\propto\frac{\nu_{i,I_i}(\pi)}{\chi_i'\left(\phi_i\left(\sum_{h|h^{s(I_i)}\in I_i}u_i(h)p_\varsigma(h|h^{s(I_i)})\pi(h^{s(I_i)})\right)\right)}.$$

By the argument in the ex-ante equilibrium part of this proof, $\varsigma\in\tilde{E}_\Gamma((\hat{\mu}_i)_{i\in N},(\hat{\phi}_i)_{i\in N})$. By definition of $\tilde{Q}_\Gamma((\mu_i)_{i\in N},(\phi_i)_{i\in N})$, for each i, each I_i and each ς_i',

$$\begin{aligned}
\text{(B8)}\qquad & \sum_\pi\phi_i\left(\sum_{h|h^{s(I_i)}\in I_i}u_i(h)p_\varsigma(h|h^{s(I_i)})\pi(h^{s(I_i)})\right)\nu_{i,I_i}(\pi)\\
\geq\ & \sum_\pi\phi_i\left(\sum_{h|h^{s(I_i)}\in I_i}u_i(h)p_{(\varsigma_i',\varsigma_{-i})}(h|h^{s(I_i)})\pi(h^{s(I_i)})\right)\nu_{i,I_i}(\pi).
\end{aligned}$$

Since χ_i is increasing, differentiable and concave, for each π we repeat the argument

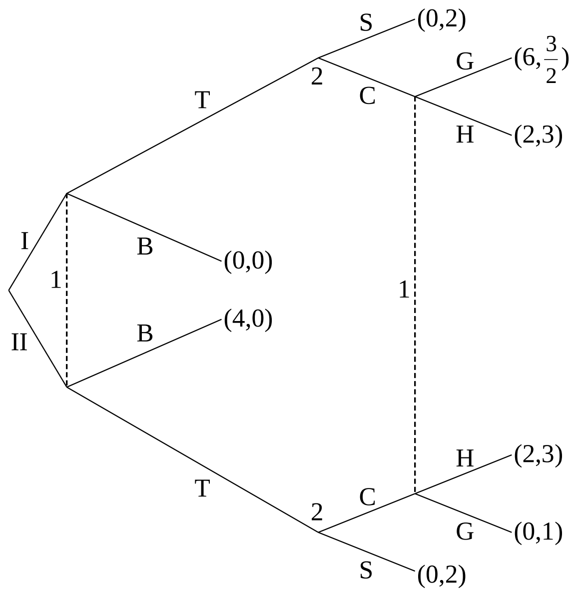

FIGURE B2. AMBIGUITY AVERSION GENERATES NEW EQUILIBRIA WITH RESPECT TO PURE STRATEGIES

in the ex-ante equilibrium part of this proof to conclude that $\varsigma \in \tilde{Q}_{\Gamma}((\hat{\mu}_i)_{i\in N}, (\hat{\phi}_i)_{i\in N})$. This shows $\tilde{Q}_{\Gamma}((\mu_i)_{i\in N}, (\phi_i)_{i\in N}) \subseteq \bigcup_{(\hat{\mu}_i)_{i\in N}} \tilde{Q}_{\Gamma}((\hat{\mu}_i)_{i\in N}, (\hat{\phi}_i)_{i\in N})$.

Finally, turn to the part of the theorem about SEA strategy profiles. Suppose $\varsigma \in \tilde{S}_{\Gamma}((\mu_i)_{i\in N}, (\phi_i)_{i\in N})$ and ν is an interim belief system such that (ς, ν) is an SEA for Γ with respect to pure strategies, where the sequence used in satisfying smooth rule consistency is $\{\sigma^k\}_{k=1}^{\infty}$. By Theorem 11, ν satisfies (A12) using $\{\sigma^k\}_{k=1}^{\infty}$. Further suppose that for each i, $\hat{\phi}_i = \chi_i(\phi_i)$ for some increasing, differentiable and concave χ_i, $\hat{\mu}_i$ is defined as in (B6), and, for each I_i, $\hat{\nu}_{i,I_i}$ is defined as in (B8). By the arguments in the sequentially optimal part of the proof, $\varsigma \in \tilde{Q}_{\Gamma}((\hat{\mu}_i)_{i\in N}, (\hat{\phi}_i)_{i\in N})$. Observe that replacing ϕ_i, μ_i and ν_{i,I_i} with $\hat{\phi}_i$, $\hat{\mu}_i$ and $\hat{\nu}_{i,I_i}$ in (A12) preserves its validity. Thus, by Theorem 11, $(\sigma, \hat{\nu})$ satisfies smooth rule consistency in $\hat{\Gamma}$. Thus $\varsigma \in \tilde{S}_{\Gamma}((\hat{\mu}_i)_{i\in N}, (\hat{\phi}_i)_{i\in N})$. This shows $\tilde{S}_{\Gamma}((\mu_i)_{i\in N}, (\phi_i)_{i\in N}) \subseteq \bigcup_{(\hat{\mu}_i)_{i\in N}} \tilde{S}_{\Gamma}((\hat{\mu}_i)_{i\in N}, (\hat{\phi}_i)_{i\in N})$.

To prove that strict inclusions may happen, consider the game depicted in Figure B2. There are two players, 1 and 2. First, it is determined whether player 2 is of type I or type II and 2 observes the type. Player 1 does not observe the type. The payoff pairs in Figure B2 describe vNM utility payoffs given players' actions and type (i.e., (u_1, u_2) means that player i receives u_i). Player 1's first move in the game is to choose between action T(wo) which gives the move to player 2 and action B(et) (i.e., betting that player 2 is of type II) which reveals the type and ends the game. If T, then player 2's move is a choice between C(ontinue) which leads to player 1 again having a non-trivial move, and S(top) which reveals the type and ends the game. If C, then player 1 has a choice between G(amble) and H(edge) after which the game ends.

Under ambiguity neutrality for both players, $\bigcup_{(\mu_i)_{i\in N}} \tilde{E}_\Gamma((\mu_i)_{i\in N},(\iota)_{i\in N}) = \{(B,(S,S),H),$ $(B,(S,S),G), (B,(C,S),H), (B,(C,S),G), (B,(S,C),H), (B,(S,C),G), (B,(C,C),H),$ $(B,(C,C),G)\}$. To see this, first note that if $\sum_\pi \pi(I)\mu_1(\pi) \in (0,\frac{2}{5})$, then all the pure profiles where 1 plays B are ex-ante equilibria under ambiguity neutrality. Second, any pure profile where 1 plays T cannot be an ex-ante equilibrium under ambiguity neutrality. Observe that 2 plays C following T (under either type) only if 1 plays H, 1 can play H rather than G on-path if and only if $2 \geq 6\sum_\pi \pi(I)\mu_1(\pi)$, and 1 can play T followed by H rather than B if and only if $p(C)2 \geq 4(1-\sum_\pi \pi(I)\mu_1(\pi))$ where $0 \leq p(C) \leq 1$ is 1's reduced probability that the type is such that 2 plays C. Since $\sum_\pi \pi(I)\mu_1(\pi)$ cannot be simultaneously $\leq \frac{1}{3}$ and $\geq \frac{1}{2}$, 1 cannot play T in pure strategy equilibrium under ambiguity neutrality.

By the weak inclusions already shown, and since SEA implies sequentially optimal, which in turn implies ex-ante equilibrium, it is enough to show that for some strictly concave $\hat{\phi}_1$ there is an SEA strategy profile with respect to pure strategies not contained in $\bigcup_{(\mu_i)_{i\in N}} \tilde{E}_\Gamma((\mu_i)_{i\in N},(\iota)_{i\in N})$. To this end, suppose $\hat{\phi}_1(x) \equiv -e^{-2x}$, $\hat{\phi}_2 \equiv \iota$ and $\mu_1(\pi_1) = \mu_1(\pi_2) = \frac{1}{2}$, where $\pi_1(I) = \frac{1}{4}$ and $\pi_2(I) = \frac{3}{4}$. Consider the pure strategy profile $(T,(C,C),H)$ and a sequence of completely mixed strategy profiles approaching it where the k^{th} element of the sequence has player 1 and each type of player 2 playing the action not assigned by $(T,(C,C),H)$ with probability $\frac{1}{k+1}$ at any point they are given the move. By Lemma 4 and Theorem 11, ν calculated using (A12) with $\hat{\phi}_1$ and μ_1 satisfies smooth rule consistency. By Theorem 3, for sequential optimality, it is sufficient to check against one-stage deviations, and therefore only at information sets where the player has a non-trivial move. For $I_1 = \{I, II\}$ and $I_1 = \{I, II\}\times\{T\}\times\{C\}$, $\nu_{1,I_1}((\pi_1)_{I_1}) = \nu_{1,I_1}((\pi_2)_{I_1}) = \frac{1}{2}$. Since $\hat{\phi}_1(2) > \frac{1}{2}(\hat{\phi}_1(3)+\hat{\phi}_1(1))$ and $\hat{\phi}_1(2) > \frac{1}{2}(\hat{\phi}_1(\frac{3}{2})+\hat{\phi}_1(\frac{9}{2}))$, 1 is best responding, and since C is a best response for player 2 given any beliefs, $(T,(C,C),H)$ is sequentially optimal with respect to pure strategies given ν. Therefore $(T,(C,C),H) \in \tilde{S}_\Gamma((\mu_i)_{i\in N},(\hat{\phi}_i)_{i\in N})$ and the proof is complete. ■

B2. *Proofs of results in Section III.B*

Parts of the next proof (of Theorem 9) make use of the following particularly convenient set of $\hat{\phi}_i$ at least as concave as ϕ_i, parametrized by $l \geq 1$ and $b \geq 1$, with $\hat{\phi}_i$ strictly more concave than ϕ_i when $b > 1$ and equal to ϕ_i when $b = 1$:

Let e_i^l denote the l^{th} lowest distinct value of $\sum_{h\in H} u_i(h)p_\sigma(h|h^0)\pi(h^0)$ generated by π in the support, Π_i, of μ_i.

DEFINITION 13: *For any $l \geq 1$ such that e_i^{l+1} exists, for $b \geq 1$ let $\hat{\phi}_i^l \equiv \psi_i^l \circ \phi_i$, where ψ_i^l is defined by*

$$\psi_i^l(y) = \begin{cases} y + \frac{1}{2}(b-1)[\phi_i\left(e_i^l\right) + \phi_i\left(e_i^{l+1}\right)] & , \quad y \geq \phi_i\left(e_i^{l+1}\right) \\ \frac{-(b-1)y^2 + 2[b\phi_i\left(e_i^{l+1}\right) - \phi_i\left(e_i^l\right)]y - (b-1)[\phi_i\left(e_i^l\right)]^2}{2[\phi_i\left(e_i^{l+1}\right) - \phi_i\left(e_i^l\right)]} & , \quad \phi_i\left(e_i^l\right) < y < \phi_i\left(e_i^{l+1}\right) \\ b\cdot y & , \quad y \leq \phi_i\left(e_i^l\right). \end{cases}$$

When $b > 1$, it may be verified that any ψ_i^l is continuously differentiable, concave, strictly increasing and not affine. Notice that for all $x \leq e_i^l$, $\hat{\phi}_i^{l\prime}(x) = b\phi_i'(x)$ and for all $x \geq e_i^{l+1}$, $\hat{\phi}_i^{l\prime}(x) = \phi_i'(x)$.

PROOF OF THEOREM 9:

This proof makes use of the E_Γ, Q_Γ and S_Γ notations for sets of equilibria given in Notation 4 (see p. 59). Fixing an ex-ante equilibrium $\sigma \in E_\Gamma((\mu_i)_{i\in N}, (\phi_i)_{i\in N})$ (resp. Q_Γ with associated interim belief system ν or S_Γ with associated ν and sequence of completely mixed strategy profiles $\{\sigma^k\}_{k=1}^\infty$) and a player i, say that ambiguity aversion makes σ_i *ex-ante (resp. sequentially optimal or SEA) belief robust* if for each $\varepsilon_i \in (0, \frac{1}{|\operatorname{supp}\mu_i|})$, there exists $\bar{\phi}_i^{\varepsilon_i}$ at least as concave as ϕ_i so that σ_i is an ex-ante best response to σ_{-i} given each $\hat{\mu}_i$ and $\hat{\phi}_i$ such that $\min_{\pi\in\operatorname{supp}\mu_i} \hat{\mu}_i(\pi) > \varepsilon_i$ and $\hat{\mu}_i$ has the same support, Π_i, as μ_i, and such that $\hat{\phi}_i$ at least as concave as $\bar{\phi}_i^{\varepsilon_i}$ (resp. that plus also a best response to σ_{-i} at each information set I_i given $\hat{\phi}_i$ and ν_{i,I_i} or (for SEA) all of the previous plus satisfying the part for player i of smooth rule consistency using $\{\sigma^k\}_{k=1}^\infty$). To prove that ambiguity aversion makes σ ex-ante (resp. sequentially optimal or SEA) belief robust, it is sufficient to show, for each player i, that ambiguity aversion makes σ_i ex-ante (resp. sequentially optimal or SEA) belief robust. The argument is the same for each player, so for the remainder of the proof fix a player i. Also assume for the remainder of the argument that $|\Pi_i| > 1$, as otherwise the result follows immediately because there is only one possible belief with that support.

We begin by proving that ambiguity aversion makes σ_i ex-ante belief robust. Recall that σ_i is an ex-ante best response to σ_{-i} for player i given $\hat{\mu}_i$ and $\hat{\phi}_i$ if and only if σ_i maximizes, among all σ_i', (11) with $\hat{\mu}_i$ replacing μ_i and $\hat{\phi}_i$ replacing ϕ_i. Observe that any strategies σ_i' that are weakly worse than σ_i (in terms of ex-ante expected payoff, $\sum_{h\in H} u_i(h) p_{(\sigma_i',\sigma_{-i})}(h|h^0)\pi(h^0)$) for all $\pi \in \Pi_i$ can never interfere with optimality of σ_i and will thus, without loss of generality, be ignored whenever making statements about strategies other than σ_i in what follows. For each l, denote by π_i^l the unique $\pi \in \Pi_i$ under which σ_i gives e_i^l, the l^{th} lowest distinct ex-ante expected payoff generated by Π_i. For each strategy σ_i' and $1 \leq l \leq |\Pi_i|$, denote

$$\begin{aligned} d_i^l(\sigma_i') \quad \equiv \quad & \left(\sum_{h\in H} u_i(h) p_\sigma(h|h^0)\pi_i^l(h^0) - \sum_{h\in H} u_i(h) p_{(\sigma_i',\sigma_{-i})}(h|h^0)\pi_i^l(h^0)\right) \\ & \cdot\phi_i'\left(\sum_{h\in H} u_i(h) p_\sigma(h|h^0)\pi_i^l(h^0)\right). \end{aligned}$$

The conclusion of the theorem in the ex-ante case is immediate when all strategies σ_i' are weakly worse than σ_i for all π_i^l (i.e., $d_i^l(\sigma_i') \geq 0$ for all l), so assume that there exists a strategy σ_i' with $d_i^l(\sigma_i') < 0$ for some $1 \leq l \leq |\Pi_i|$.

We next show that all strategies σ_i' must have $d_i^1(\sigma_i') \geq 0$. To see this, suppose, to the contrary, there exists a strategy $\hat{\sigma}_i$ with $d_i^1(\sigma_i') < 0$. Since σ is ex-ante robust to increased ambiguity aversion, σ_i is an ex-ante best response to σ_{-i} for player i given μ_i

and $\hat{\phi}_i^1$ (from Definition 13), and, in particular, is at least as good as $\hat{\sigma}_i$. Using (11) with $\hat{\phi}_i^1$ replacing ϕ_i, this implies

$$\sum_{l=1}^{|\Pi_i|} d_i^l(\hat{\sigma}_i)\mu_i(\pi) + (b-1)d_i^1(\hat{\sigma}_i) \geq 0.$$

Since the value of the first term is bounded and $d_i^1(\hat{\sigma}_i) < 0$, taking b large enough generates a contradiction.

For each pure strategy ς_i', let $m(\varsigma_i') < |\Pi_i|$ be the smallest number l for which $d_i^{l+1}(\varsigma_i') < 0$. By the previous paragraph, $m(\varsigma_i') \geq 1$. By the definition of $m(\varsigma_i')$, $d_i^l(\varsigma_i') \geq 0$ for all $1 \leq l \leq m(\varsigma_i')$. Furthermore, $d_i^l(\varsigma_i') > 0$ for some $1 \leq l \leq m(\varsigma_i')$, because otherwise ς_i' could be used together with μ_i and $\hat{\phi}_i^{m(\varsigma_i')+1}$ to generate a contradiction to σ_i being ex-ante robust to increased ambiguity aversion. Thus $\sum_{l=1}^{m(\varsigma_i')} d_i^l(\varsigma_i') > 0$ and $\min_{m(\varsigma_i')+1\leq l\leq|\Pi_i|} d_i^l(\varsigma_i') < 0$. For each $1 \leq m < |\Pi_i|$, if there exists no pure strategy ς_i' with $m(\varsigma_i') = m$, then let $B(m) = 1$, otherwise let

$$B(m) \equiv \max\{1, \max_{\varsigma_i'|m(\varsigma_i')=m} \frac{-\min_{m+1\leq l\leq|\Pi_i|} d_i^l(\varsigma_i')}{\varepsilon_i \sum_{l=1}^{m} d_i^l(\varsigma_i')}\},$$

which is well defined because the set of pure strategies is finite. Define $\bar{\phi}_i^{\varepsilon_i} = \psi_i^1 \circ ... \circ \psi_i^{|\Pi_i|-1} \circ \phi_i$, for ψ_i^m with $b = B(m)$ for each $1 \leq m < |\Pi_i|$. Consider $\hat{\phi}_i$ at least as concave as $\bar{\phi}_i^{\varepsilon_i}$, i.e., $\hat{\phi}_i = \psi_i \circ \bar{\phi}_i^{\varepsilon_i}$ for some ψ_i continuously differentiable, concave and strictly increasing. For any $\hat{\mu}_i$ such that $\min_{\pi\in\operatorname{supp}\mu_i} \hat{\mu}_i(\pi) > \varepsilon_i$ and $\hat{\mu}_i$ has the same support as μ_i, and any pure strategy ς_i',

$$\begin{aligned}
& \sum_{\pi\in\Pi_i}\left(\sum_{h\in H} u_i(h)p_\sigma(h|h^0)\pi(h^0) - \sum_{h\in H} u_i(h)p_{(\varsigma_i',\sigma_{-i})}(h|h^0)\pi(h^0)\right) \\
& \cdot\hat{\phi}_i'\left(\sum_{h\in H} u_i(h)p_\sigma(h|h^0)\pi(h^0)\right)\hat{\mu}_i(\pi) \\
= \quad & \sum_{l=1}^{|\Pi_i|} d_i^l(\varsigma_i')\psi_i'\left(\bar{\phi}_i^{\varepsilon_i}\left(\sum_{h\in H} u_i(h)p_\sigma(h|h^0)\pi_i^l(h^0)\right)\right)\left(\prod_{m=l}^{|\Pi_i|-1} B(m)\right)\hat{\mu}_i(\pi_i^l) \\
\geq \quad & \psi_i'\left(\bar{\phi}_i^{\varepsilon_i}\left(\sum_{h\in H} u_i(h)p_\sigma(h|h^0)\pi_i^{m(\varsigma_i')}(h^0)\right)\right)\left(\prod_{m=m(\varsigma_i')+1}^{|\Pi_i|-1} B(m)\right) \\
& \cdot\left(B[m(\varsigma_i')]\varepsilon_i \sum_{l=1}^{m(\varsigma_i')} d_i^l(\varsigma_i') + \min_{m(\varsigma_i')+1\leq l\leq|\Pi_i|} d_i^l(\varsigma_i')\right) \\
\geq \quad & 0,
\end{aligned}$$

where the last inequality follows by applying the definition of $B[m(\varsigma_i')]$. Therefore, σ_i does at least as well as any pure strategy ς_i' given σ_{-i} according to (11) with $\hat{\mu}_i$ and $\hat{\phi}_i$ replacing μ_i and ϕ_i. Since (11) is linear in the mixing weights in σ_i', this is sufficient to conclude that σ_i is a best response to σ_{-i} given $\hat{\mu}_i$ and $\hat{\phi}_i$. Therefore ambiguity aversion makes σ_i ex-ante belief robust.

Consider now sequential optimality. Consider $\bar{\phi}_i^{\varepsilon_i}$ and $\hat{\phi}_i$ as defined above. Since σ is sequentially optimal robust to increased ambiguity aversion, σ_i is an ex-ante best response to σ_{-i} given μ_i and $\hat{\phi}_i$, and for each information set I_i there exists a belief ν_{i,I_i} such that σ_i is a best response at I_i to σ_{-i} given ν_{i,I_i} and $\hat{\phi}_i$. Consider any $\hat{\mu}_i$ such that $\min_{\pi\in\text{supp}\,\mu_i}\hat{\mu}_i(\pi) > \varepsilon_i$ and $\hat{\mu}_i$ has the same support as μ_i. By the ex-ante equilibrium argument above, σ_i is an ex-ante best response to σ_{-i} given $\hat{\mu}_i$ and $\hat{\phi}_i$. Given $\hat{\mu}_i$ and $\hat{\phi}_i$, derive $\hat{\nu}_{i,I_i}$ from the smooth rule using σ for those information sets I_i for which that rule implies that $\hat{\nu}_{i,I_i}$ must vary with ex-ante beliefs. By Lemma 1, σ_i an ex-ante best response to σ_{-i} given $\hat{\mu}_i$ and $\hat{\phi}_i$ implies σ_i is a best response to σ_{-i} at these information sets given $\hat{\phi}_i$ and $\hat{\nu}_{i,I_i}$. Extend $\hat{\nu}$ by setting $\hat{\nu}_{i,I_i} = \nu_{i,I_i}$ elsewhere. Thus σ_i is a best response to σ_{-i} also at these remaining information sets given $\hat{\nu}_{i,I_i}$ and $\hat{\phi}_i$, as this fact is not affected by the shift from μ_i to $\hat{\mu}_i$. This shows that ambiguity aversion makes σ_i sequentially optimal belief robust.

Finally turn to SEA. We establish the existence of beliefs at each I_i so that i's part of both sequential optimality and smooth rule consistency are satisfied. Since (σ,ν) satisfies smooth rule consistency using $\{\sigma^k\}_{k=1}^{\infty}$, Theorem 11 yields that $\bar{p}_{-i,\sigma_{-i}}(h^t|h^0)$ exists for each player i and each $h^t \in I_i \in \mathcal{F}_{i,\sigma}$. Given any $\hat{\phi}_i$, for each I_i, construct a belief $\hat{\nu}_{i,I_i}$ as defined in (A12) using $\{\sigma^k\}_{k=1}^{\infty}$ with $\hat{\mu}_i$ and $\hat{\phi}_i$ replacing μ_i and ϕ_i. Theorem 11 applied with $\hat{\mu}_i$ and $\hat{\phi}_i$ replacing μ_i and ϕ_i (and noting that $\bar{p}_{-i,\sigma_{-i}}(h^t|h^0)$ is independent of the choice of $\hat{\mu}_i$ and $\hat{\phi}_i$) implies that σ together with $\hat{\nu}$ satisfies player i's part of smooth rule consistency using $\{\sigma^k\}_{k=1}^{\infty}$ given $\hat{\mu}_i$ and $\hat{\phi}_i$.

Showing that σ_i is a best response to σ_{-i} for player i at each I_i given $\hat{\nu}_{i,I_i}$ and $\hat{\phi}_i$ is equivalent to showing that

$$\sigma_i \in \arg\max_{\sigma_i'}\sum_{\hat{\pi}\in\Pi_i}\left(\sum_{h|h^{s(I_i)}\in I_i}u_i(h)\,p_{(\sigma_i',\sigma_{-i})}(h|h^{s(I_i)})\,\bar{p}_{-i,\sigma_{-i}}(h^{s(I_i)}|h^0)\hat{\pi}(h^0)\right)$$
$$\cdot\hat{\phi}_i'\left(\sum_{h\in H}u_i(h)\,p_\sigma(h|h^0)\hat{\pi}(h^0)\right)\hat{\mu}_i(\hat{\pi}), \tag{B9}$$

as can be seen by considering (B5) with $\hat{\nu}_{i,I_i}$ replacing ν_{i,I_i} and $\hat{\phi}_i$ replacing ϕ_i, substituting for $\hat{\nu}_{i,I_i}$ and $\overline{\hat{\pi}}_{I_i}$ using (A12), replacing the summation over $\pi \in \Delta(I_i)$ and $\hat{\pi} \in \Delta(\Theta)$ such that $\overline{\hat{\pi}}_{I_i} = \pi$ with summation over $\hat{\pi} \in \Pi_i$ since each element in the support of $\hat{\nu}_{i,I_i}$ is $\overline{\hat{\pi}}_{I_i}$ for some $\hat{\pi}$ in the support of $\hat{\mu}_i$, and simplifying, including, since π no longer appears in the expression, replacing the notation $\hat{\pi} \in \Delta(\Theta)$ with $\pi \in \Delta(\Theta)$.

Since σ is SEA robust to increased ambiguity aversion, σ_i is an ex-ante best response to σ_{-i} given μ_i and $\hat{\phi}_i^1$, and for any I_i there exists a belief ν_{i,I_i}, constructed as was $\hat{\nu}_{i,I_i}$ at the beginning of the SEA part of the proof except now using μ_i and $\hat{\phi}_i^1$, such that σ_i is a best response to σ_{-i} given ν_{i,I_i} and $\hat{\phi}_i^1$. From the definition of $\hat{\phi}_i^1$, the assumption that π_i^1 is well-defined and the assumption that the same sequence $\{\sigma^k\}_{k=1}^{\infty}$ can be used in smooth rule consistency for each $\hat{\phi}_i$ (which ensures use of the same $\bar{p}_{-i,\sigma_{-i}}(h^{s(I_i)}|h^0)$), (B9) with $\hat{\mu}_i = \mu_i$ and $\hat{\phi}_i = \hat{\phi}_i^1$ with b large enough (i.e., $\hat{\phi}_i^1$ sufficiently concave) implies that, for each I_i, $\sigma_i' = \sigma_i$ must maximize the following expected payoff under π_i^1,

$$\sum_{h|h^{s(I_i)}\in I_i} u_i(h)p_{(\sigma_i',\sigma_{-i})}(h|h^{s(I_i)})\bar{p}_{-i,\sigma_{-i}}(h^{s(I_i)}|h^0)\pi_i^1(h^0),$$

and, by the corresponding argument for ex-ante equilibrium, also maximizes the ex-ante expected payoff under π_i^1.

For each I_i, σ_i' and $1 \le l \le |\Pi_i|$, denote

$$d_{i,I_i}^l(\sigma_i') \equiv \left(\sum_{h|h^{s(I_i)}\in I_i} u_i(h)p_{\sigma}(h|h^{s(I_i)})\bar{p}_{-i,\sigma_{-i}}(h^{s(I_i)}|h^0)\pi_i^l(h^0)\right.$$
$$\left. - \sum_{h|h^{s(I_i)}\in I_i} u_i(h)p_{(\sigma_i',\sigma_{-i})}(h|h^{s(I_i)})\bar{p}_{-i,\sigma_{-i}}(h^{s(I_i)}|h^0)\pi_i^l(h^0)\right)\cdot\phi_i'\left(\sum_{h\in H}u_i(h)p_{\sigma}(h|h^0)\pi_i^l(h^0)\right).$$

That σ_i is a best response to σ_{-i} for player i at I_i is immediate when $d_{i,I_i}^l(\varsigma_i') \ge 0$ for all pure strategies ς_i' and l, so assume that $d_{i,I_i}^l(\varsigma_i') < 0$ for some ς_i' and $1 \le l \le |\Pi_i|$. For each I_i and any such ς_i' (as any other strategy can never interfere with optimality of σ_i at I_i and thus, without loss of generality, may be ignored), let $m_{i,I_i}(\varsigma_i') < |\Pi_i|$ be the smallest number l for which $d_{i,I_i}^{l+1}(\varsigma_i') < 0$. By the previous paragraph, $m_{i,I_i}(\varsigma_i') \ge 1$. By the definition of $m_{i,I_i}(\varsigma_i')$, $d_{i,I_i}^l(\varsigma_i') \ge 0$ for all $1 \le l \le m(\varsigma_i')$. Furthermore, $d_{i,I_i}^l(\varsigma_i') > 0$ for some $1 \le l \le m_{i,I_i}(\varsigma_i')$, because otherwise ς_i' could be used together with μ_i and $\hat{\phi}_i^{m_{i,I_i}(\varsigma_i')+1}$ to generate a contradiction to σ_i being SEA robust to increased ambiguity aversion. Thus $\sum_{l=1}^{m_{i,I_i}(\varsigma_i')} d_{i,I_i}^l(\varsigma_i') > 0$ and $\min_{m_{i,I_i}(\varsigma_i')+1\le l\le|\Pi_i|} d_{i,I_i}^l(\varsigma_i') < 0$. For each $1 \le m < |\Pi_i|$, if there exists no pure strategy ς_i' with $m_{i,I_i}(\varsigma_i') = m$, then let $B_{i,I_i}(m) = 1$, otherwise let

$$B_{i,I_i}(m) \equiv \max\{1, \max_{\varsigma_i'|m_{i,I_i}(\varsigma_i')=m} \frac{-\min_{m+1\le l\le|\Pi_i|} d_{i,I_i}^l(\varsigma_i')}{\varepsilon_i\sum_{l=1}^m d_{i,I_i}^l(\varsigma_i')}\},$$

which is well defined because the set of pure strategies is finite. Define $\bar{\phi}_i^{\varepsilon_i} = \psi_i^1 \circ \ldots \circ \psi_i^{|\Pi_i|-1} \circ \phi_i$, for ψ_i^m with $b = \bar{B}(m) \equiv \max\{B(m), \max_{I_i\in\mathcal{I}_i} B_{i,I_i}(m)\}$ for each $1 \le m < |\Pi_i|$, where $B(m)$ is as defined in the ex-ante part of the proof. Consider $\hat{\phi}_i$

at least as concave as $\bar{\phi}_i^{\varepsilon_i}$, i.e., $\hat{\phi}_i = \psi_i \circ \bar{\phi}_i^{\varepsilon_i}$ for some ψ_i continuously differentiable, concave and strictly increasing. Consider any $\hat{\mu}_i$ such that $\min_{\pi \in \operatorname{supp} \mu_i} \hat{\mu}_i(\pi) > \varepsilon_i$ and $\hat{\mu}_i$ has the same support as μ_i. By the argument above for ex-ante equilibrium, σ_i is an ex-ante best response to σ_{-i} given $\hat{\mu}_i$ and $\hat{\phi}_i$. For any I_i and pure strategy ς'_i,

$$\begin{aligned}
&\sum_{\pi \in \Pi_i} \Bigg(\sum_{h|h^{s(I_i)} \in I_i} u_i(h) p_\sigma(h|h^{s(I_i)}) \bar{p}_{-i,\sigma_{-i}}(h^{s(I_i)}|h^0)\pi(h^0) \\
&- \sum_{h|h^{s(I_i)} \in I_i} u_i(h) p_{(\sigma'_i,\sigma_{-i})}(h|h^{s(I_i)}) \bar{p}_{-i,\sigma_{-i}}(h^{s(I_i)}|h^0)\pi(h^0) \Bigg) \cdot \hat{\phi}'_i \left(\sum_{h \in H} u_i(h) p_\sigma(h|h^0)\pi(h^0) \right) \hat{\mu}_i(\pi) \\
= \;& \sum_{l=1}^{|\Pi_i|} d^l_{i,I_i}(\varsigma'_i) \psi'_i \left(\bar{\phi}_i^{\varepsilon_i} \left(\sum_{h \in H} u_i(h) p_\sigma(h|h^0) \pi^l_i(h^0) \right) \right) \left(\prod_{m=l}^{|\Pi_i|-1} \bar{B}(m) \right) \hat{\mu}_i(\pi^l_i) \\
\geq \;& \psi'_i \left(\bar{\phi}_i^{\varepsilon_i} \left(\sum_{h \in H} u_i(h) p_\sigma(h|h^0) \pi_i^{m_{i,I_i}(\varsigma'_i)}(h^0) \right) \right) \left(\prod_{m=m_{i,I_i}(\varsigma'_i)+1}^{|\Pi_i|-1} \bar{B}(m) \right) \\
&\cdot \left(\bar{B}[m_{i,I_i}(\varsigma'_i)] \varepsilon_i \sum_{l=1}^{m_{i,I_i}(\varsigma'_i)} d^l_i(\varsigma'_i) + \min_{m_{i,I_i}(\varsigma'_i)+1 \leq l \leq |\Pi_i|} d^l_i(\varsigma'_i) \right) \\
\geq \;& 0,
\end{aligned}$$

where the last inequality follows by applying the definition of $\bar{B}[m(\varsigma'_i)]$. Therefore, σ_i does at least as well as any pure strategy ς'_i given σ_{-i} according to (B9) with $\hat{\mu}_i$ and $\hat{\phi}_i$ replacing μ_i and ϕ_i. Since (B9) is linear in the mixing weights in σ'_i, this is sufficient to conclude that σ_i is a best response to σ_{-i} for player i at each I_i given $\hat{\nu}_{i,I_i}$ and $\hat{\phi}_i$. Furthermore, by construction, σ together with beliefs $\hat{\nu}_{i,I_i}$ for player i satisfy player i's part of smooth rule consistency using $\{\sigma^k\}_{k=1}^\infty$ given $\hat{\mu}_i$ and $\hat{\phi}_i$. Therefore ambiguity aversion makes σ_i SEA belief robust. ■

PROOF OF REMARK 2:

Assume ϕ_i is twice continuously differentiable with strictly positive first derivative and recall that, all along, it was assumed to be strictly increasing and concave. In the proof of Theorem 9, $\bar{\phi}_i^{\varepsilon_i}$ was taken to be $\psi_i^1 \circ ... \circ \psi_i^{|\Pi_i|-1} \circ \phi_i$. From the definition of the ψ_i^m, it follows that $\bar{\phi}_i^{\varepsilon_i}$ is twice differentiable, strictly increasing, and concave and has bounded second derivative. Take this $\bar{\phi}_i^{\varepsilon_i}$ and let $(\bar{\phi}_i^{\varepsilon_i})^{-1}$ be its inverse. We want to show that there exists an $\alpha(\varepsilon_i) > 0$ and an increasing, concave transformation ζ such that $-e^{-\alpha(\varepsilon_i)x} = \zeta[\bar{\phi}_i^{\varepsilon_i}(x)]$ for all $x \in \operatorname{co}(u_i(H))$. For any $y \in \bar{\phi}_i^{\varepsilon_i}[\operatorname{co}(u_i(H))]$, $-e^{-\alpha(\varepsilon_i)x} = \zeta[\bar{\phi}_i^{\varepsilon_i}(x)]$ implies $\zeta(y) = -e^{-\alpha(\varepsilon_i)(\bar{\phi}_i^{\varepsilon_i})^{-1}(y)}$, which is increasing. Thus $\zeta'(y) = \frac{\alpha(\varepsilon_i)e^{-\alpha(\varepsilon_i)(\bar{\phi}_i^{\varepsilon_i})^{-1}(y)}}{(\bar{\phi}_i^{\varepsilon_i})'[(\bar{\phi}_i^{\varepsilon_i})^{-1}(y)]}$, and the sign of $\zeta''(y)$ is the sign of $-\alpha(\varepsilon_i)e^{-\alpha(\varepsilon_i)(\bar{\phi}_i^{\varepsilon_i})^{-1}(y)} \left(\alpha(\varepsilon_i) - (-\frac{(\bar{\phi}_i^{\varepsilon_i})''[(\bar{\phi}_i^{\varepsilon_i})^{-1}(y)]}{(\bar{\phi}_i^{\varepsilon_i})'[(\bar{\phi}_i^{\varepsilon_i})^{-1}(y)]}) \right)$, so ζ is concave for all sufficiently large $\alpha(\varepsilon_i)$, since $-\frac{(\bar{\phi}_i^{\varepsilon_i})''(x)}{(\bar{\phi}_i^{\varepsilon_i})'(x)}$, the coefficient of ambiguity aversion at $x \in \text{co}(u_i(H))$ is non-negative and bounded above. Note that $(\bar{\phi}_i^{\varepsilon_i})''(x)$ is bounded because the composition of any functions f and g that have bounded second derivatives and continuous, strictly positive first derivatives has bounded second derivatives and continuous, strictly positive first derivatives as follows from the formula $f[g(x)]'' = [f'[g(x)]g'(x)]' = f''[g(x)][g'(x)]^2 + f'[g(x)]g''(x)$ and the fact that since ϕ_i has bounded derivatives, as do the ψ_i^m. Observe that the $\alpha(\varepsilon_i)$ may need to be much higher than some $-\frac{(\bar{\phi}_i^{\varepsilon_i})''(x)}{(\bar{\phi}_i^{\varepsilon_i})'(x)}$ since it must be at least the supremum of this over x. ■

B3. Proofs of results in Section IV

The next result relates to analysis of the ambiguous cheap talk example.

PROOF OF PROPOSITION 4:

Since all information sets are on-path under the given strategies, by Theorems 10 and 12 it is sufficient to establish that the given strategies form an ex-ante equilibrium. P's strategy is an ex-ante best response because it leads to payoff 2 for all parameters, which is the highest feasible payoff for this player. Let γ_m be the probability with which agent r plays w after message $m \in \{\alpha, \beta\}$, and similarly let δ_m be the corresponding probabilities for agent c. The proposed strategies correspond to $\gamma_\alpha = \gamma_\beta = \delta_\beta = 1$ and $\delta_\alpha = 0$. We now verify that these are ex-ante best responses. Denoting $\pi_k(IIU) + \pi_k(IID)$ by $\pi_k(II)$, given the strategies of the others, r maximizes

$$\frac{1}{2}\sum_{k=1}^{2}\phi_r\left(\pi_k(IU)\gamma_\alpha + 2\pi_k(ID)\gamma_\beta + \pi_k(II)[2\gamma_\beta + 5(1-\gamma_\beta)]\right).$$

Since this function is strictly increasing in γ_α, it is clearly maximized at $\gamma_\alpha = 1$. The first derivative with respect to γ_β evaluated at $\gamma_\alpha = \gamma_\beta = 1$ is

$$\begin{aligned}&\frac{1}{2}\sum_{k=1}^{2}[2\pi_k(ID) - 3\pi_k(II)]\,\phi_r'\,(2-\pi_k(IU))\\ =\;&\frac{11}{8}e^{-11\cdot\frac{39}{20}}\left(e^{-11(\frac{7}{4}-\frac{39}{20})} - \frac{42}{5}\right) > 0,\end{aligned}$$

where the last line uses $\phi_r(x) = -e^{-11x}$ and the values of the π_k. Thus, by concavity in γ_β, the maximum is attained at $\gamma_\alpha = \gamma_\beta = 1$. Similarly, given the strategies of the others, c maximizes

$$\frac{1}{2}\sum_{k=1}^{2}\phi_c\left(\pi_k(IU)[2\delta_\alpha + 5(1-\delta_\alpha)] + \pi_k(ID)[2\delta_\beta + 5(1-\delta_\beta)] + 2\pi_k(II)\delta_\beta\right).$$

Since this function is strictly decreasing in δ_α, it is clearly maximized at $\delta_\alpha = 0$. The first derivative with respect to δ_β evaluated at $\delta_\alpha = 0$ and $\delta_\beta = 1$ is

$$\begin{aligned} &\frac{1}{2}\sum_{k=1}^{2}[-3\pi_k(ID) + 2\pi_k(II)]\phi_c'(3\pi_k(IU)+2) \\ = \quad &-\frac{1}{2}\phi_c'\left(\frac{11}{4}\right) + \frac{23}{40}\phi_c'\left(\frac{43}{20}\right) \geq \frac{3}{40}\phi_c'\left(\frac{11}{4}\right) > 0, \end{aligned}$$

where the last line uses the values of the π_k. Since ϕ_c is weakly concave, the problem is weakly concave in δ_β, thus the maximum is attained at $\delta_\alpha = 0$ and $\delta_\beta = 1$. ■

PROOF OF PROPOSITION 5:

Limit attention to strategies for P conditioning only on the payoff relevant component of the parameter, I and II. Denote P's probability of playing α conditional on the payoff relevant component by ρ_I and ρ_{II}, respectively. Let γ_m be the probability with which r plays w after message $m \in \{\alpha, \beta\}$, and similarly let δ_m be the corresponding probabilities for c. Given ρ_I and ρ_{II}, r chooses $\gamma_\alpha, \gamma_\beta$ to maximize

$$\text{(B10)} \qquad \frac{1}{2}\sum_{k=1}^{2}\phi_r\left(\begin{array}{c} \pi_k(I)[\rho_I(1+\delta_\alpha)\gamma_\alpha + (1-\rho_I)(1+\delta_\beta)\gamma_\beta] \\ +\pi_k(II)[\rho_{II}((1+\delta_\alpha)\gamma_\alpha + 5(1-\gamma_\alpha)) \\ +(1-\rho_{II})((1+\delta_\beta)\gamma_\beta + 5(1-\gamma_\beta))] \end{array}\right)$$

and c chooses $\delta_\alpha, \delta_\beta$ to maximize

$$\text{(B11)} \qquad \frac{1}{2}\sum_{k=1}^{2}\phi_c\left(\begin{array}{c} \pi_k(I)[\rho_I((1+\gamma_\alpha)\delta_\alpha + 5(1-\delta_\alpha)) \\ +(1-\rho_I)((1+\gamma_\beta)\delta_\beta + 5(1-\delta_\beta))] \\ +\pi_k(II)[\rho_{II}(1+\gamma_\alpha)\delta_\alpha + (1-\rho_{II})(1+\gamma_\beta)\delta_\beta] \end{array}\right).$$

The proof proceeds by considering four cases, which together are exhaustive:

Case 1: When $\rho_I = \rho_{II} = 1$ (resp. $\rho_I = \rho_{II} = 0$) so that only one message is sent, for P to always receive the maximal payoff of 2 it is necessary that the agents play w, w with probability 1 after this message, i.e. $\gamma_\alpha = \delta_\alpha = 1$ (resp. $\gamma_\beta = \delta_\beta = 1$). But w is not a best response for c, as can be seen by the fact that the partial derivative of (B11) with respect to δ_α (resp. δ_β) evaluated at those strategies is

$$\frac{1}{2}(4 - 5\sum_{k=1}^{2}\pi_k(I))\phi_c'(2) = -\frac{3}{8}\phi_c'(2) < 0.$$

Similarly, one can show that w is not a best response for r.

Case 2: When $0 < \rho_{II} < 1$, since under II, P sends both messages with positive probability, it is necessary that w, w is played with probability 1 after both messages in order that the principal always receive the maximal payoff of 2. A necessary condition for this to be a best response for c is that the partial derivatives of (B11) with respect to $\delta_\alpha, \delta_\beta$ are non-negative at $\gamma_\alpha = \gamma_\beta = \delta_\alpha = \delta_\beta = 1$. This is, respectively, equivalent to

$14\rho_{II} \geq 19\rho_I$ and $14(1-\rho_{II}) \geq 19(1-\rho_I)$, which implies $14 \geq 19$, a contradiction.

Case 3: When $\rho_{II} = 0$ and $0 < \rho_I \leq 1$, (B11) is strictly decreasing in δ_α, thus the maximum is attained at $\delta_\alpha = 0$. For P to always receive the maximal payoff of 2, it is necessary that $\gamma_\alpha = \gamma_\beta = \delta_\beta = 1$. However, this is not a best response for r because the partial derivative of (B10) with respect to γ_β evaluated at these strategies using the values for the π_k is,

$$\frac{3}{4}(\frac{1}{2}-\rho_I)\phi_r'(2-\frac{3}{4}\rho_I)+(-\frac{1}{5}\rho_I-1)\phi_r'(2-\frac{\rho_I}{5}) < 0.$$

To see this, note that the second term is always negative, the first term is non-positive for $\frac{1}{2} \leq \rho_I \leq 1$, and, when $0 < \rho_I < \frac{1}{2}$, substituting $\phi_r(x) = -e^{-11x}$ yields that the left-hand side is negative.

Case 4: When $\rho_{II} = 1$ and $0 \leq \rho_I < 1$, the argument is identical to Case 3 except the roles of the messages α and β are swapped. ■

The next result relates to analysis of the limit pricing example. Denote the entrant's Cournot profit net of entry costs when facing an incumbent of type θ by $\omega_\theta \equiv b(\frac{a+c_\theta-2c_E}{3b})^2 - K$.

LEMMA 5: *Under Assumption 3, σ^{LP} is an ex-ante equilibrium if and only if (ICH for I), (ICM for I), $\omega_H \geq 0$ and*

$$\text{(ICL for E)} \qquad \sum_\pi \mu(\pi)(\pi(L)\omega_L + \pi(M)\omega_M)\phi'(\pi(H)\omega_H) \leq 0.$$

The conditions above correspond to the following incentives in the game: (ICH for I), (ICM for I) were described in the main text, $\omega_H \geq 0$ ensures that the entrant is willing to enter when it is sure the incumbent is type H, and ICL for E ensures the entrant does not want to enter after observing the monopoly quantity for type L.

PROOF OF LEMMA 5:

Since there is complete information in the final stage, the Cournot or monopoly quantities respectively are ex-ante optimal there. Taking the incumbent's point of view, consider its action in the first stage. Since the incumbent learns its cost before taking any action and there is no other uncertainty, checking ex-ante optimality for the incumbent is equivalent to checking optimality for each incumbent type separately given the entrant's strategy. This is true no matter what the incumbent's ambiguity aversion or beliefs.

When does type H not prefer to pool with M,L at the monopoly quantity for L and thereby deter entry? Profits for H in the conjectured equilibrium are $b(\frac{a-c_H}{2b})^2+b(\frac{a+c_E-2c_H}{3b})^2$. Profits if it instead pools with M,L at monopoly quantity for L and deters entry are $\frac{a-c_L}{2b}(a-\frac{a-c_L}{2}-c_H)+b(\frac{a-c_H}{2b})^2$. H at least as well off not pooling if and only if

$$b(\frac{a+c_E-2c_H}{3b})^2 \geq \frac{a-c_L}{2b}(a-\frac{a-c_L}{2}-c_H).$$

This is equivalent to (ICH for I).

When does type M not prefer to produce the monopoly quantity for M and fail to deter entry? Profits for M in the conjectured equilibrium are $\frac{a-c_L}{2b}(a - \frac{a-c_L}{2} - c_M) + b(\frac{a-c_M}{2b})^2$. If it instead produced at the monopoly quantity for M and fails to deter entry, profits are $b(\frac{a-c_M}{2b})^2 + b(\frac{a+c_E-2c_M}{3b})^2$. M is at least as well off pooling with L if and only if

$$\frac{a-c_L}{2b}(a - \frac{a-c_L}{2} - c_M) \geq b(\frac{a+c_E-2c_M}{3b})^2.$$

This is equivalent to (ICM for I).

Type L is playing optimally since its monopoly quantity also deters entry.

It remains to examine the entry decision of the entrant. As a best-response to the incumbent's strategy, ex-ante the entrant wants to maximize

$$\text{(B12)} \qquad \sum_{\pi} \mu(\pi)\phi\left[\lambda_L(\pi(L)w_L + \pi(M)w_M) + \lambda_H \pi(H)w_H\right]$$

with respect to $\lambda_H, \lambda_L \in [0, 1]$, where λ_H and λ_L are the mixed-strategy probabilities of entering contingent on seeing the monopoly quantity for H and the monopoly quantity for L, respectively. When is this maximized at $\lambda_H = 1$ and $\lambda_L = 0$? Notice, by monotonicity, some maximum involves $\lambda_H = 1$ if and only if $w_H \geq 0$, and $w_H > 0$ is equivalent to $\lambda_H = 1$ being part of every maximum. This says that entering against a known high cost incumbent is profitable. Assuming this is satisfied, so that $\lambda_H = 1$ is optimal, then $\lambda_L = 0$ is optimal if and only if the derivative of (B12) with respect to λ_L evaluated at $\lambda_L = 0$ and $\lambda_H = 1$ is non-positive, which yields (ICL for E). ■

Before turning to the proof of Proposition 6, we remark that we actually prove a slightly stronger result, allowing for the possibility that $\mu(\{\pi \mid \pi(L)w_L + \pi(M)w_M = 0\}) = 1$ (i.e., that the entrant unambiguously believes that it will exactly break even if it enters conditional on the incumbent's type being in $\{L, M\}$). This appears in the proof only in the proof of Lemma 6.

PROOF OF PROPOSITION 6:

Consider the limit pricing strategy profile σ^{LP}.

By Lemma 6, under the assumptions of the proposition there exists a $\hat{\phi}$ such that if the entrant's ϕ is at least as concave as $\hat{\phi}$, then (ICL for E) is satisfied. By Lemma 5, the assumptions of the proposition together with (ICL for E) are sufficient for σ^{LP} to be an ex-ante equilibrium.

Next, we construct an interim belief system that, together with σ^{LP}, satisfies smooth rule consistency. Consider a sequence of completely mixed strategy profiles, σ^k, where $\gamma^k_{\theta,q} > 0$ is the probability that type θ of the incumbent chooses first period quantity q, $\lambda^k_q > 0$ is the probability that the entrant enters after observing quantity q, $\delta^k_{\theta,(q,enter,r)} > 0$ and $\delta^k_{(q,enter,r)} > 0$ are the probabilities of second period quantity r being chosen by, respectively, type θ of the incumbent and the entrant, after observing first period quantity q followed by entry and revelation of θ, and $\delta^k_{\theta,(q,no\ entry,r)} > 0$ is the probability of second period quantity r being chosen by type θ of the incumbent after observing first period quantity q followed by no entry. Specifically, let $\gamma^{k}_{\theta,q} \equiv \frac{\beta^k_{\theta,q}}{\sum_{\hat{q}\in\mathcal{Q}} \beta^k_{\theta,\hat{q}}}$ for $k = 1, 2, \ldots$, where $\beta^k_{\theta,q}$ is defined according to Figure B3.

θ	$q \in \mathcal{Q}$			
	$q = q_H$	$q = q_L$	$q_H \neq q < q_L$	$q > q_L$
L	1	k^2	1	k
M	1	k^2	1	1
H	k^2	1	k	1

FIGURE B3. DEFINITION OF $\beta^k_{\theta,q}$

Additionally, let λ^k_q converge to 1 as $k \to \infty$ when $q < q_L$ and converge to 0 otherwise, $\delta^k_{\theta,(q,enter,r)}$ converge to 1 as $k \to \infty$ when r is the Cournot quantity for type θ and converge to 0 otherwise, $\delta^k_{(q,enter,r)}$ converge to 1 as $k \to \infty$ when r is the Cournot quantity for the entrant and converge to 0 otherwise, and $\delta^k_{\theta,(q,no\ entry,r)}$ converge to 1 as $k \to \infty$ when r is the monopoly quantity for type θ and converge to 0 otherwise. Note that σ^k converges to σ^{LP}. By Lemma 4, Theorem 11 delivers an interim belief system ν such that $\left(\sigma^{\mathrm{LP}}, \nu\right)$ satisfies smooth rule consistency.

The final step in the proof is to verify that $\left(\sigma^{\mathrm{LP}}, \nu\right)$ satisfies the optimality conditions (8) at all information sets. By Theorem 3, for optimality, it is sufficient to check against one-stage deviations, and therefore only at information sets where the player has a non-trivial move. The Cournot strategies in the last stage given entry are optimal because all distributions over type become degenerate when conditioned on the entrant learning the incumbent's type. The fact that $\omega_L < 0$ plus $\omega_H \geq 0$ implies that it is optimal for the entrant to stay out if its objective function after observing q places all weight on type L and to enter if that objective function places all weight on type H. We now verify that when $q \neq q_L$ this objective function does exactly that when entry/no entry are supposed to occur according to σ^{LP}. Entry is supposed to occur if and only if $q < q_L$. When $q = q_H$, since π_{I_i} is the degenerate distribution on type H for all π that may be so conditioned, it is optimal to enter. When $q_H \neq q < q_L$, since $\bar{p}_{-i,\sigma^{\mathrm{LP}}_{-i}}(\theta, q|\theta)$ places all weight on (H, q), (A12) implies that all π in the support of $\nu_{E,\Theta\times\{q\}}$ puts weight only on (H, q), and so it is again optimal to enter. Similarly, when $q > q_L$, since $\bar{p}_{-i,\sigma^{\mathrm{LP}}_{-i}}(\theta, q|\theta)$ places all weight on (L, q), all π in the support of $\nu_{E,\Theta\times\{q\}}$ puts weight only on (L, q), and so it is optimal not to enter.

Not entering being optimal after observing $q = q_L$ is equivalent (see 11) to the following:

$$\text{(B13)} \qquad \sum_{\pi\in\Delta(\Theta\times\{q_L\})} (\pi(L, q_L)\omega_L + \pi(M, q_L)\omega_M)\phi'(0)\nu_{E,\{L,M\}\times\{q_L\}}(\pi) \leq 0.$$

Using the formula (A12) to substitute for $\nu_{E,q_L}(\pi)$ in (B13) yields that not entering remaining optimal is equivalent to (ICL for E). Therefore $\left(\sigma^{\mathrm{LP}}, \nu\right)$ satisfies the optimality conditions (8) at all information sets as long as the entrant's ϕ is at least as concave as the $\hat{\phi}$ identified from Lemma 6. For such sufficiently concave ϕ, having shown $\left(\sigma^{\mathrm{LP}}, \nu\right)$ is sequentially optimal and satisfies smooth rule consistency, it is therefore an SEA.

Since the only assumption on ϕ made in the above argument that σ^{LP} is part of an SEA was that it was sufficiently concave for the entrant, the argument goes through in its entirety for all $\tilde{\phi}$ at least as concave as ϕ. Furthermore, the same sequence $\{\sigma^k\}_{k=1}^{\infty}$ may be used for all $\tilde{\phi}$. Thus, σ^{LP} is SEA robust to increased ambiguity aversion.

We next verify that the other conditions in the antecedents of Theorem 9 are satisfied. We begin by showing that, for each player, $\sum_{h\in H} u_i(h)p_{\sigma^{\mathrm{LP}}}(h|h^0)\pi(h^0)$ can be strictly ordered across the π in the support of μ. For the entrant,

$$\sum_{h\in H} u_i(h)p_{\sigma^{\mathrm{LP}}}(h|h^0)\pi(h^0) = \pi(H)\omega_H.$$

Thus, strict ordering corresponds to strict ordering by $\pi(H)$. The assumption that the support of μ can be ordered in the likelihood-ratio ordering ensures the latter, as it implies that for any two distinct $\pi, \pi' \in \operatorname{supp}\mu$, $\pi(H) \neq \pi'(H)$. To see this, suppose to the contrary that $\pi(H) = \pi'(H)$. By distinctness and that weights must sum to one, $\pi(M) \neq \pi'(M)$, $\pi(L) \neq \pi'(L)$ and $\pi(M) > \pi'(M)$ if and only if $\pi(L) < \pi'(L)$, a violation of likelihood-ratio ordering. For the incumbent,

$$\begin{aligned}\sum_{h\in H} u_i(h)p_{\sigma^{\mathrm{LP}}}(h|h^0)\pi(h^0) \quad = \quad & \pi(L)2b(\frac{a-c_L}{2b})^2 \\ & +\pi(M)[\frac{a-c_L}{2b}(a-\frac{a-c_L}{2}-c_M)+b(\frac{a-c_M}{2b})^2] \\ & +\pi(H)[b(\frac{a-c_H}{2b})^2+b(\frac{a+c_E-2c_H}{3b})^2].\end{aligned}$$

By Assumption 3 and (ICM for I), the expression multiplied by $\pi(L)$ is strictly larger than the one multiplied by $\pi(M)$, which is, in turn, strictly larger than the one multiplied by $\pi(H)$. Thus, likelihood-ratio ordering of the support of μ implies strict ordering of $\sum_{h\in H} u_i(h)p_{\sigma^{\mathrm{LP}}}(h|h^0)\pi(h^0)$. By Theorem 9, ambiguity aversion makes σ^{LP} SEA belief robust. ■

LEMMA 6: *Under the assumptions of Proposition 6 there exists an $\alpha > 0$ such that if ϕ is at least as concave as $-e^{-\alpha x}$ then (ICL for E) is satisfied.*

PROOF OF LEMMA 6:

Assume the conditions of the proposition. We show that (ICL for E) is satisfied for concave enough ϕ. The assumption in the proposition that some $\pi \in \operatorname{supp}\mu$ makes entry conditional on $\{L, M\}$ strictly unprofitable means $\mu(\{\pi \mid \pi(L)\omega_L + \pi(M)\omega_M < 0\}) > 0$. If $\mu(\{\pi \mid \pi(L)\omega_L + \pi(M)\omega_M \leq 0\}) = 1$ then (ICL for E) is trivially satisfied for any ϕ. For the remainder of the proof, therefore, suppose that $\mu(\{\pi \mid \pi(L)\omega_L + \pi(M)\omega_M > 0\}) >$

0. Let $\Pi^- \equiv \{\pi \mid \pi(L)w_L + \pi(M)w_M < 0\}$, $\Pi^+ \equiv \{\pi \mid \pi(L)w_L + \pi(M)w_M > 0\}$, $N \equiv \sum_{\pi\in\Pi^-} \mu(\pi)(\pi(L)w_L + \pi(M)w_M)$, and $P \equiv \sum_{\pi\in\Pi^+} \mu(\pi)(\pi(L)w_L + \pi(M)w_M)$. Let $\pi^- \in \arg\max_{\pi\in\Pi^-} \pi(H)$ and $\pi^+ \in \arg\min_{\pi\in\Pi^+} \pi(H)$. The left-hand side of (ICL for E) can be bounded from above as follows:

$$\begin{aligned}
& \sum_{\pi\in\Pi^-} \mu(\pi)(\pi(L)w_L + \pi(M)w_M)\phi'(\pi(H)w_H) + \sum_{\pi\in\Pi^+} \mu(\pi)(\pi(L)w_L + \pi(M)w_M)\phi'(\pi(H)w_H) \\
\leq\ & \sum_{\pi\in\Pi^-} \mu(\pi)(\pi(L)w_L + \pi(M)w_M)\phi'\left(\pi^-(H)w_H\right) + \sum_{\pi\in\Pi^+} \mu(\pi)(\pi(L)w_L + \pi(M)w_M)\phi'\left(\pi^+(H)w_H\right) \\
=\ & N\phi'\left(\pi^-(H)w_H\right) + P\phi'\left(\pi^+(H)w_H\right).
\end{aligned}$$

Consider $\phi(x) = -e^{-\alpha x}$, $\alpha > 0$. The upper bound above becomes

$$\alpha N e^{-\alpha\pi^-(H)w_H} + \alpha P e^{-\alpha\pi^+(H)w_H}.$$

We show that this upper bound is non-positive for sufficiently large α, implying (ICL for E). The upper bound is non-positive if and only if $Pe^{-\alpha\pi^+(H)w_H} \leq -Ne^{-\alpha\pi^-(H)w_H}$ if and only if $e^{\alpha(\pi^-(H)-\pi^+(H))w_H} \leq -\frac{N}{P}$ if and only if $\alpha\left(\pi^-(H) - \pi^+(H)\right)w_H \leq \ln(-\frac{N}{P})$. Since $\pi^-(L)w_L + \pi^-(M)w_M < 0 < \pi^+(L)w_L + \pi^+(M)w_M$ and $c_L < c_M$, we have $w_L < 0 < w_M$. Thus, $\frac{\pi^-(L)}{\pi^-(M)} > -\frac{w_M}{w_L} > \frac{\pi^+(L)}{\pi^+(M)}$. By our assumption on the support of μ and Lemma 7, $\frac{\pi^-(L)}{\pi^-(M)} > \frac{\pi^+(L)}{\pi^+(M)}$ implies $\pi^-(H) < \pi^+(H)$. Therefore, $\alpha\left(\pi^-(H) - \pi^+(H)\right)w_H \leq \ln(-\frac{N}{P})$ if and only if $\alpha \geq \frac{\ln(-\frac{N}{P})}{\left(\pi^-(H)-\pi^+(H)\right)w_H}$.

To complete the proof, fix α satisfying this inequality and consider ϕ such that $\phi(x) = h(-e^{-\alpha x})$ for all x with h concave and strictly increasing on $(-\infty, 0)$. We show that (ICL for E) holds. Observe that $\phi'(x) = h'(-e^{-\alpha x})\alpha e^{-\alpha x}$. Since $\pi^-(H) - \pi^+(H) < 0$ and $w_H > 0$, we have

$$-e^{-\alpha\pi^-(H)w_H} \leq -e^{-\alpha\pi^+(H)w_H}$$

and, by concavity of h,

$$h'(-e^{-\alpha\pi^-(H)w_H}) \geq h'(-e^{-\alpha\pi^+(H)w_H}).$$

Therefore the upper bound derived above satisfies

$$\begin{aligned}
& N\phi'\left(\pi^-(H)w_H\right) + P\phi'\left(\pi^+(H)w_H\right) \\
=\ & \alpha N e^{-\alpha\pi^-(H)w_H} h'(-e^{-\alpha\pi^-(H)w_H}) + \alpha P e^{-\alpha\pi^+(H)w_H} h'(-e^{-\alpha\pi^+(H)w_H}) \\
\leq\ & (\alpha N e^{-\alpha\pi^-(H)w_H} + \alpha P e^{-\alpha\pi^+(H)w_H}) h'(-e^{-\alpha\pi^-(H)w_H}) \leq 0
\end{aligned}$$

by the first part of the proof and the assumption on α. This implies (ICL for E). ■

LEMMA 7: *If the support of μ can be ordered in the likelihood-ratio ordering, then, for any $\pi, \pi' \in \operatorname{supp}\mu$, $\frac{\pi(L)}{\pi(M)} > \frac{\pi'(L)}{\pi'(M)}$ implies $\pi(H) < \pi'(H)$.*

PROOF OF LEMMA 7:

Suppose the support of μ can be so ordered. Fix any $\pi, \pi' \in \text{supp}\,\mu$. Suppose $\frac{\pi(L)}{\pi(M)} > \frac{\pi'(L)}{\pi'(M)}$. Then $\frac{\pi'(L)}{\pi(L)} < \frac{\pi'(M)}{\pi(M)}$, and thus, by likelihood-ratio ordering, $\frac{\pi'(L)}{\pi(L)} < \frac{\pi'(M)}{\pi(M)} \leq \frac{\pi'(H)}{\pi(H)}$. This implies $\pi'(H) > \pi(H)$ since the last two ratios cannot be less than or equal to 1 without violating the total probability summing to 1. ■

B4. *Details on the analysis of the game in Figure 2 and the comparison with no profitable one-stage deviations and consistent planning*

A strengthening of no profitable one-stage deviations used in some of the existing literature investigating games with ambiguity is the following condition, describing a consistent planning requirement in the spirit of Strotz (1955-56) (for a formal decision theoretic treatment see Siniscalchi 2011):

DEFINITION 14: *Fix a game Γ and a pair (σ, ν) consisting of a strategy profile and interim belief system. Specify V_i and V_{i,I_i} as in (1) and (4). For each player i and information set $I_i \in \mathcal{I}_i^T$, let*

$$CP_{i,I_i} \equiv \operatorname*{argmax}_{\hat{\sigma}_i \in \Sigma_i} V_{i,I_i}(\hat{\sigma}_i, \sigma_{-i}).$$

Then, inductively, for $0 \leq t \leq T-1$, and $I_i \in \mathcal{I}_i^t$ let

$$CP_{i,I_i} \equiv \operatorname*{argmax}_{\hat{\sigma}_i \in \bigcap_{\hat{I}_i \in \mathcal{I}_i^{t+1} | \hat{I}_i^{-1} = I_i} CP_{i,\hat{I}_i}} V_{i,I_i}(\hat{\sigma}_i, \sigma_{-i}).$$

Finally, let

$$CP_i \equiv \operatorname*{argmax}_{\hat{\sigma}_i \in \bigcap_{\hat{I}_i \in \mathcal{I}_i^0} CP_{i,\hat{I}_i}} V_i(\hat{\sigma}_i, \sigma_{-i}).$$

(σ, ν) *is* optimal under consistent planning *if, for all players i,*

$$\sigma_i \in CP_i.$$

Equivalently, (σ, ν) is such that for all players i,

$$V_i(\sigma) \geq V_i(\hat{\sigma}_i, \sigma_{-i}) \text{ for all } \hat{\sigma}_i \in \bigcap_{\hat{I}_i \in \mathcal{I}_i^0} CP_{i,\hat{I}_i}$$

and, for all information sets $I_i \in \mathcal{I}_i^t$, $0 \leq t \leq T-1$,

$$V_{i,I_i}(\sigma) \geq V_{i,I_i}(\hat{\sigma}_i, \sigma_{-i}) \text{ for all } \hat{\sigma}_i \in \bigcap_{\hat{I}_i \in \mathcal{I}_i^{t+1} | \hat{I}_i^{-1} = I_i} CP_{i,\hat{I}_i}$$

and, for all information sets $I_i \in \mathcal{I}_i^T$,

$$V_{i,I_i}(\sigma) \geq V_{i,I_i}(\hat{\sigma}_i, \sigma_{-i}) \text{ for all } \hat{\sigma}_i \in \Sigma_i.$$

If (σ, ν) is sequentially optimal then it is also optimal under consistent planning. However, if (σ, ν) is optimal under consistent planning it may fail to be sequentially optimal (even when limiting attention to ambiguity neutrality). For such a failure to occur, the optimal strategy from player i's point of view at some earlier stage must have a continuation that fails to be optimal from the viewpoint of some later reachable stage. This is what makes the extra constraints imposed in the optimization inequalities under consistent planning bind. Just as with no profitable one-stage deviations, when updating is according to the smooth rule, (σ, ν) optimal under consistent planning implies (σ, ν) is sequentially optimal, making the three equivalent under smooth rule updating.

Recall that the example in Figure 2 in Section II.B showed how the no profitable one-stage deviation criterion under Bayesian updating allowed strategy profiles that are not ex-ante equilibria of a game (and thus clearly not sequentially optimal). Replacing no profitable one-stage deviations by consistent planning does not change this fact. The main text used the following specification of preferences for the example: $\phi_1(x) = -e^{-10x}$, μ is $1/2$ on $(1/3, 1/9, 5/9)$ and $1/2$ on $(1/3, 5/9, 1/9)$, and 1's beliefs after seeing U are given by Bayes' rule applied to μ: $1/3$ on $(3/4, 1/4, 0)$ and $2/3$ on $(3/8, 5/8, 0)$. With these parameters and beliefs, the following strategy profile satisfies no profitable one-stage deviations and consistent planning: player 1 plays o with probability $1 - \frac{9}{20}\ln(\frac{29}{11}) \approx 0.564$ and mixes evenly between u and d if U, while player 2 plays her strictly dominant strategy if given the move. Notice, if we consider any more concave ϕ_1, playing o with even higher probability will be consistent with consistent planning or no profitable one-stage deviations given these beliefs. In the limit where the decision maker is Maxmin EU with set of priors equal to the convex combinations of $(1/3, 1/9, 5/9)$ and $(1/3, 5/9, 1/9)$ and applies Bayes' rule to each measure in the set, playing o with probability 1 is consistent with consistent planning and no profitable one-stage deviations.