Econometrica Supplementary Material

SUPPLEMENT TO "CONTRACT NEGOTIATION AND THE COASE CONJECTURE: A STRATEGIC FOUNDATION FOR RENEGOTIATION-PROOF CONTRACTS"

(*Econometrica*, Vol. 85, No. 2, March 2017, 585–616)

BY BRUNO STRULOVICI

For simplicity, this supplement uses appendix and equation numbering that continue from the main text of the paper.

APPENDIX D: PROOF OF THEOREM 1

IT SUFFICES TO PROVE THE RESULT for R_0 in the H-Rent configuration and $\beta_0 \in (0,1)$: the L-Rent case is shown symmetrically, and existence in other cases (No Rent and degenerate prior) has already been established by Proposition 1. The proof proceeds in two steps:

(i) Step 1—Prove the existence of an equilibrium in an auxiliary game played between P and H.

(ii) Step 2—Construct a strategy profile of the original game based on the equilibrium established in Step 1, and verify that it defines a PBE of the original game.

Step 1: Auxiliary game

The game starts with a contract $R_0 \in \mathcal{H}$ in the H-Rent configuration and a parameter $\beta \in (0,1)$. For this auxiliary game, β is just a parameter affecting the payoff functions and is devoid of its interpretation as a belief.

The auxiliary game is a dynamic game with infinitely many rounds. At each round n, starting in state R_n, P proposes new contracts $R_{n+1} \in \mathcal{H}$ and $C_n \in \mathcal{E}_H$ subject to the constraints

$$u_L(R_{n+1}) \ge u_L(R_n), \tag{30}$$

$$u_L(R_{n+1}) \ge u_L(C_n), \tag{31}$$

$$u_H(C_n) \ge u_H(R_n). \tag{32}$$

H then chooses a number $\mu_n \in [0,1]$. The interpretation of this choice is that H accepts R_{n+1} with probability μ_n and C_n with probability $(1-\mu_n)$. For this auxiliary game, however, μ_n is simply an action deterministically affecting payoffs.

The principal's cost, for strategies $\{R_{n+1}, C_n\}$ and $\{\mu_n\}$, is given by

$$\begin{aligned} &\mathcal{Q}(\{R_{n+1}, C_n\}, \{\mu_n\}) \\ &\quad = \sum_{n\ge 0} Q(C_n)\beta(1-\eta)^n(1-\mu_n)\prod_{k=0}^{n-1}\mu_k \\ &\qquad + \sum_{n\ge 0} Q(R_{n+1})\left(\beta(1-\eta)^n\eta\prod_{k=0}^{n}\mu_j + (1-\beta)(1-\eta)^n\eta\right), \end{aligned} \tag{33}$$

 DOI: 10.3982/ECTA13637

while H's payoff is

$$\mathcal{V}\big(\{R_{n+1}, C_n\}, \{\mu_n\}\big) = \sum_{n\geq 0} u_H(C_n)(1-\eta)^n(1-\mu_n)\prod_{k=0}^{n-1}\mu_k + \sum_{n\geq 0} u_H(R_{n+1})(1-\eta)^n\eta\prod_{k=0}^{n}\mu_n. \tag{34}$$

These payoffs correspond to the expected cost and utility that P and H would obtain in an equilibrium of the *original* game in which P proposes two contracts at each round, the breakdown probability is η, $\{\mu_n\}$ is the mixing strategy of H, L always accepts R_{n+1}, and the initial probability of facing H is equal to β.

LEMMA 17: *For any initial R_0 and $\beta \in (0,1)$, there exists a perfect equilibrium of the auxiliary game.*

PROOF: The result is a direct consequence of Theorem 1 in Harris (1985). We check Assumptions 1–5 of this theorem. The payoff function of the principal is simply the negative of his cost, $\mathcal{Q}$. P's (unconstrained) action set in round n is $S_{P_n} = \mathcal{H} \times \mathcal{E}_H$, while H's action space is $S_{H_n} = [0,1]$, which are both compact and Hausdorff spaces. Hence, Assumptions 1 and 2 are satisfied. P's feasible set at each round n, as defined by the constraints (30) and (32), is closed and depends continuously on the current state. Therefore, the set $\mathcal{S}_f$ of feasible sequences is closed in $\mathcal{S} = \bigtimes_n (S_{P_n} \times S_{H_n})$ endowed with the product topology, and the set of feasible actions in round n depends continuously on past play. Thus, Assumptions 3 and 4 are satisfied. Finally, the payoffs $-\mathcal{Q}$ and $\mathcal{V}$ are clearly continuous on their domain $\mathcal{S}_f$, so Assumption 5 is satisfied as well. The result follows. *Q.E.D.*

REMARK 1: We can similarly define an auxiliary game and equilibrium when instead R_0 is in the L-Rent configuration. This equilibrium yields an expected utility for H, as a passive player of the auxiliary game, given by

$$V_H(\beta) = \sum_{n\geq 0}(1-\eta)^n\eta u_H(R_{n+1}). \tag{35}$$

This equilibrium and payoff are used to define H's strategy, off the equilibrium path, in the PBE construction for the original game.

Step 2: Equilibrium of the original game

Starting from $R_0 \in \mathcal{H}$ and a belief $\beta_0 \in (0,1)$, the equilibrium strategies are defined as follows:

At each round n:

(i) P proposes the sequence of contracts $\{C_n, R_{n+1}\}$ corresponding to the auxiliary game started at (R_0, β_0).

(ii) L accepts R_{n+1} with probability 1, while H accepts R_{n+1} with probability μ_n and C_n with probability $(1-\mu_n)$, where μ_n is H's equilibrium choice in the auxiliary game. If $R_{n+1} \neq R_n$ and the agent accepts R_n, P assigns probability 1 to H, so the continuation play is trivially defined in that case, by Proposition 1 (whose proof is independent of Theorem 1).

(iii) If P proposes, at some round n, a menu M_n that does *not* correspond to the pair of contracts defined by the auxiliary game, let $\bar{R}_{n+1}$ denote the contract of $M_n \cup \{R_n\}$ that maximizes L's utility and $\bar{C}_n$ denote the contract of $M_n \cup \{R_n\}$ that maximizes H's utility.[56] By construction, $\bar{R}_{n+1}$ and $\bar{C}_n$ satisfy (30)–(32). Let $\hat{R}_{n+1}$ denote the L-efficient contract that gives L the same utility as $\bar{R}_{n+1}$ and $\hat{C}_n$ denote the H-efficient contract that gives H the same utility as $\bar{C}_n$. There are three cases to consider: (a) $u_H(\hat{C}_n) \geq u_H(\hat{R}_{n+1})$ and $u_L(\hat{R}_{n+1}) \geq u_L(\hat{C}_n)$, (b) $u_H(\hat{C}_n) < u_H(\hat{R}_{n+1})$ and $u_L(\hat{R}_{n+1}) \geq u_L(\hat{C}_n)$, and (c) $u_H(\hat{C}_n) \geq u_H(\hat{R}_{n+1})$ and $u_L(\hat{R}_{n+1}) < u_L(\hat{C}_n)$. Because of the single-crossing property, the fourth and last logical case cannot occur, as is easily checked.

Continuation play is then defined as follows, according to each case.

(a) L chooses $\bar{R}_{n+1}$ with probability 1, H chooses $\bar{C}_n$ with probability 1. If the agent chooses any other contract R in $M_n \cup \{R_n\}$, then the principal assigns probability 1 to a type θ of the agent such that the other type $\theta' \neq \theta$ cannot benefit from choosing that contract if the principal put probability 1 on θ.[57] There always exists at least one such type, as is easily checked.

(b) Case (b) can occur only if $\bar{R}_{n+1}$ is in the H-Rent configuration. Continuation play is defined by the continuation equilibrium of the auxiliary game in which, following R_n, P proposes $\bar{R}_{n+1}$ and $\hat{C}_n$, but replacing $\hat{C}_n$ by $\bar{C}_n$.[58] In particular, L accepts $\bar{R}_{n+1}$ with probability 1, and H randomizes between the contracts $\bar{C}_n$ and $\bar{R}_{n+1}$ according to the probability μ_n coming from the auxiliary equilibrium if $\bar{C}_n$ is replaced by $\hat{C}_n$. If the agent picks any contract other than $\bar{R}_{n+1}$ and $\bar{C}_n$, P assigns probability 1 to one type according to the same rule as in Case (a). Continuation for rounds $m \geq n+1$ is also determined by the equilibrium of the auxiliary game.

(c) Case (c) is symmetric to Case (b), and can only occur if $\bar{C}_n$ is in the L-Rent configuration. The continuation equilibrium is defined by the continuation equilibrium, from period 1 onwards (see Remark 1), of the auxiliary game starting in period 0 with belief $\tilde{\beta}_0 = 1 - \beta_n$ (since L now plays the role of H and vice versa) and at some fictitious contract $\tilde{R}$ in the L-Rent configuration such that $\hat{R}_{n+1}$ and $\bar{C}_n$ satisfy equations (32) and (30), respectively (the inequalities are reversed, because the equilibrium is in the L-Rent configuration).

This construction defines the continuation strategies after any possible history. We now verify that the strategy profile thus constructed forms an equilibrium. The proof uses the one-shot deviation principle, which applies since the breakdown probability η has the effect of discounting the utility of future rounds at a geometric rate.

Consider first L's strategy, assuming that P follows the prescribed sequence of contracts. From (31), L cannot benefit from picking C_n: doing so would cause β_n to jump to 1 and L to be stuck with utility $u_L(C_n)$, which is (weakly) lower than $u_L(R_{n+1})$ and hence lower than his continuation utility if he chooses R_{n+1}.[59] Similarly, if $R_{n+1} \neq R_n$ and

[56]If there are several maximizers, the equilibrium selects any of them. There must be at least one maximizer, because the menu is finite.

[57]That is, θ' prefers the contract that he is supposed to take with probability 1 in equilibrium (e.g., $\bar{C}_n$ if $\theta' = H$) to the θ-efficient contract that gives θ the same utility as R.

[58]By construction, $\bar{R}_{n+1} \in \mathcal{H}$, $\hat{C}_n$ is H-efficient, and the contracts satisfy conditions (30), (31), and (32), so the contract pair is feasible for the auxiliary games.

[59]That utility is always weakly higher than $u_L(R_{n+1})$, since L can always hold on to R_{n+1}.

L chooses R_n, then β_n jumps to 1, and L's continuation utility is bounded above[60] by $u_L(R_n)$, which is weakly dominated by accepting R_{n+1} by (30) (guaranteeing that L's continuation utility is bounded below by $u_L(R_{n+1})$).

Let us now consider the optimality of H's strategy. From (32), $u_H(C_n) \geq u_H(R_n)$. Therefore, if H holds on to R_n, his continuation utility is equal to $u_H(R_n)$, which is weakly dominated by taking C_n. Moreover, given that H randomizes between C_n and R_{n+1}, his expected payoff is given by (34), and by perfection of the auxiliary equilibrium, the strategy $\{\mu_n\}$ is a best response to the sequence of contracts.

Consider now the agent's strategy after a deviation by P. In Case (a), if L chooses $\bar{C}_n$, his utility is bounded above by $\max\{u_L(\bar{C}_n), u_L(\hat{C}_n)\}$, which is less than $u_L(\bar{R}_{n+1})$, by definition of Case (a). Similarly, if L picks any other contract R, then either P puts probability 1 on L, in which case L gets utility $u_L(R)$, which is less than $u_L(\bar{R}_{n+1})$, by definition of $\bar{R}_{n+1}$, or P puts probability 1 on H, but in this case L cannot benefit from this erroneous belief. The same reasoning applies to H: it is optimal for that type to choose $\bar{C}_n$.

In Case (b), L prefers $\bar{R}_{n+1}$ over any other contract in $M_n \cup \{R_n\}$, by an argument similar to Case (a). Now consider H's response to P's deviation. First, H cannot benefit from choosing a contract R other than $\bar{C}_n$ and $\bar{R}_{n+1}$, for the reason explained in Case (a). Moreover, given the continuation play, which is defined by the auxiliary equilibrium, it is optimal to randomize according to the probability μ_n coming from the auxiliary equilibrium in which $\bar{R}_{n+1}$ and $\hat{C}_n$ are proposed.[61]

Case (c) is similar to Case (b).

There remains to show that P's strategy is optimal. By construction of the auxiliary equilibrium, P's strategy is optimal among all strategies that propose contracts (R_{n+1}, C_n) satisfying (30), (31), and (32). As shown by Lemma 2 (whose proof is independent of Theorem 1), P can never benefit from any deviation in which L accepts a contract that is not in the H-Rent configuration. Moreover, any contract R accepted by H with positive probability and that is *not* in the H-Rent configuration immediately results, at the next round, in an H-efficient contract that gives H the same utility as R and is less costly to P than R. We can therefore, without loss of generality, consider deviations in which P proposes one H-efficient contract, $\bar{C}_n$, and a number of contracts in the H-Rent configuration, among which $\bar{R}_{n+1}$ maximizes L's utility, and such that $u_L(\bar{C}_n) \leq u_L(R_{n+1})$. Given the agent's strategy, the menu is equivalent to just proposing $\bar{C}_n$ and $\bar{R}_{n+1}$, which is a feasible strategy in the auxiliary equilibrium and thus has to be weakly dominated by the equilibrium menu, by subgame perfection of that menu in the auxiliary game.

APPENDIX E: PROOF OF INEQUALITIES

PROOF OF LEMMA 13: Consider two contracts C and $\hat{C}$ on $\mathcal{E}_H$ ordered as in the statement of the lemma. The efficiency curve $\mathcal{E}_H$ can be parameterized by a univariate parameter λ such that, letting $C(\lambda) = (x_1(\lambda), x_2(\lambda))$ denote the H-efficient contract corresponding to parameter λ, the map $\lambda \mapsto C(\lambda)$ is continuous, one-to-one, and onto from the parameter set Λ (a compact interval of $\mathbb{R}$) to $\mathcal{E}_H$. We can assume without loss that Λ

[60] Indeed, P then proposes the H-efficient contract R that gives H utility $u_H(R_n)$, and $u_L(R) \leq u_L(R_n)$ by the single-crossing property and the fact that R_n is in the H-Rent configuration.

[61] Because $u_H(\hat{C}_n) = u_H(\bar{C}_n)$, H gets exactly the same utility as in the auxiliary equilibrium, even though the contract $\bar{C}_n$ is not in the H-Rent configuration.

contains $[0,1]$ and that $C(0)=C$ and $C(1)=\hat{C}$. We choose the parameterization to be regular, that is, such that $\lambda \mapsto C(\lambda)$, seen as a function from Λ to $\mathbb{R}^2$, is smooth and does not go "too slow" or "too fast" along $\mathcal{E}_H$.[62] We have

$$Q(\hat{C})-Q(C)=\int_0^1 \frac{dQ\big(x_1(\lambda),x_2(\lambda)\big)}{d\lambda}\cdot dC(\lambda)$$
$$=\int_0^1\left(\frac{\partial Q\big(C(\lambda)\big)}{\partial x_1}\frac{dx_1}{d\lambda}+\frac{\partial Q\big(C(\lambda)\big)}{\partial x_2}\frac{dx_2}{d\lambda}\right)d\lambda$$

and

$$u_H(\hat{C})-u_H(C)=\int_0^1 \frac{du_H\big(x_1(\lambda),x_2(\lambda)\big)}{d\lambda}\cdot dC(\lambda)$$
$$=\int_0^1\left(\frac{\partial u_H\big(C(\lambda)\big)}{\partial x_1}\frac{dx_1}{d\lambda}+\frac{\partial u_H\big(C(\lambda)\big)}{\partial x_2}\frac{dx_2}{d\lambda}\right)d\lambda.$$

By assumption, the partial derivatives of Q and u_H are strictly positive and continuous on the compact domain $\mathcal{C}$, and hence bounded below away from zero as well as bounded above. Therefore, there exist positive constants $\underline{a}<a$ such that $\underline{a}\frac{\partial u_H}{\partial x_i}\leq\frac{\partial Q}{\partial x_i}\leq a\frac{\partial u_H}{\partial x_i}$ for $i=1,2$. Combining these inequalities with the integral representations of $Q(\hat{C})-Q(C)$ and $u_H(\hat{C})-u_H(C)$ then shows (14).

For the second part of the lemma, consider the parameterizations of $\mathcal{E}_H$ and $\mathcal{E}_L$ in which the parameter is the utility that each contract gives to H (i.e., $u_H(C(\lambda))=\lambda$), with generic elements $C(\lambda)=(x_1^H(\lambda),x_2^H(\lambda))$ for $\mathcal{E}_H$ and $E(\lambda)=(x_1^L(\lambda),x_2^L(\lambda))$ for $\mathcal{E}_L$. Since u_H's partial derivatives are strictly positive over the compact domain $\mathcal{C}$ and the curves $\mathcal{E}_\theta$ are nondecreasing, the parameterizations are well defined and regular in the sense of the previous paragraph. Consider two contracts C and $\hat{C}$ of $\mathcal{E}_H$ which provide H with utilities $u_H<\hat{u}_H$ and let E and $\hat{E}$ denote the contracts of $\mathcal{E}_L$ corresponding to utilities u_H and $\hat{u}_H$. Repeating the argument of the previous paragraph, we have

$$Q(\hat{C})-Q(C)=\int_{u_H}^{\hat{u}_H}\left(\frac{\partial Q\big(C(\lambda)\big)}{\partial x_1}\frac{dx_1^H}{d\lambda}+\frac{\partial Q\big(C(\lambda)\big)}{\partial x_2}\frac{dx_2^H}{d\lambda}\right)d\lambda$$

and

$$Q(\hat{E})-Q(E)=\int_{u_H}^{\hat{u}_H}\left(\frac{\partial Q\big(E(\lambda)\big)}{\partial x_1}\frac{dx_1^L}{d\lambda}+\frac{\partial Q\big(E(\lambda)\big)}{\partial x_2}\frac{dx_2^L}{d\lambda}\right)d\lambda.$$

Since the parameterizations are regular and the efficiency curves are nondecreasing, there must exist positive constants $\underline{x}<\bar{x}$ such that $0<\underline{x}\max\{dx_1^H/d\lambda,dx_2^H/d\lambda\}\leq\max\{dx_1^L/d\lambda,dx_2^L/d\lambda\}\leq\bar{x}\max\{dx_1^H/d\lambda,dx_2^H/d\lambda\}$. Moreover, since Q has strictly positive derivatives, bounded below away from zero and bounded above, there also exist positive constants $\underline{q}<\bar{q}$ such that $\underline{q}\partial Q(C(\lambda))/\partial x_i\leq\partial Q(E(\lambda))/\partial x_i\leq\bar{q}\partial Q(C(\lambda))/\partial x_i$ for all

[62]Formally, this means that the norm of the gradient of the function $\lambda\mapsto C(\lambda)$ is uniformly bounded below and above by strictly positive constants.

$\lambda \in [u_H, \hat{u}_H]$ and $i = 1, 2$. Combining these inequalities with the previous integral representations implies the existence of positive constants $\underline{b} < b$ such that

$$\underline{b}\big(Q(\hat{E}) - Q(E)\big) \le Q(\hat{C}) - Q(C) \le b\big(Q(\hat{E}) - Q(E)\big),$$

which concludes the proof. *Q.E.D.*

PROOF OF LEMMA 14: Lemma 1 implies that

$$\beta_n Q_H + (1-\beta_n)Q_L \le \beta_n Q\big(E_H(R_n)\big) + (1-\beta_n)Q\big(E_L(R_n)\big).$$

Moreover, Q_H is bounded below by the cost of the H-efficient contract $C_H(n)$ that provides utility $u_H(n)$ to H, as this contract is the least costly way of providing H with his continuation utility, by convexity of Q. This implies that $Q_L \le Q(E_L(R_n)) + \frac{\beta_n}{1-\beta_n}(Q(E_H(R_n)) - Q(C_H(n))$. The contracts $E_H(R_n)$ and $C_H(n)$ both lie on $\mathcal{E}_H$. Equation (14) implies that $Q(E_H(R_n)) - Q(C_H(n)) \le a(u_H(E_H(R_n)) - u_H(n)) = aw_n$, proving (16).

From (16), R_{n+1} cannot give L a utility greater than the L-efficient contract that costs $Q(E_L(R_n)) + \frac{a\beta_n}{1-\beta_n}w_n$. This implies that $Q(E_L(R_{n+1})) - Q(E_L(R_n))$ is bounded above by $\frac{a\beta_n}{1-\beta_n}w_n$. Combining this with (15) yields[63]

$$Q\big(E_H(R_{n+1})\big) - Q\big(E_H(R_n)\big) \le \frac{ab\beta_n}{1-\beta_n}w_n.$$

This, along with the first part of (14), yields (17). We have

$$\begin{aligned}
w_{n+1} &= u_H\big(E_H(R_{n+1})\big) - u_H(n+1)\\
&= \big[u_H\big(E_H(R_{n+1})\big) - u_H\big(E_H(R_n)\big)\big] + u_H\big(E_H(R_n)\big) - u_H(n+1)\\
&\le \big[u_H\big(E_H(R_{n+1})\big) - u_H\big(E_H(R_n)\big)\big] + u_H\big(E_H(R_n)\big) - u_H(n)\\
&\le w_n\left(\frac{\alpha\beta_n}{1-\beta_n} + 1\right),
\end{aligned}$$

where the first inequality comes from the monotonicity of $u_H(n)$ in n, and the second one comes from (17). This shows (18).

Because L can hold on forever to R_n, his continuation utility $u_L(n)$ is bounded below by $u_L(R_n)$. At round $n+1$, P's expected cost conditional on facing L is bounded above by $Q(E_L(R_{n+1})) + \frac{\beta_{n+1}}{1-\beta_{n+1}}aw_{n+1}$, from (16) applied to round $n+1$. By the same argument that yielded (14), there exists $\alpha_L > 0$ such that $u_L(E) - u_L(E') \le \alpha_L(Q(E) - Q(E'))$ for all $E, E' \in \mathcal{E}_L$. Therefore, the highest utility that may be achieved at this cost is bounded above by $u_L(R_{n+1}) + \hat{a}\beta_{n+1}/(1-\beta_{n+1})w_{n+1}$, for some proportionality constant $\hat{a}$, and

$$u_L(R_n) \le u_L(n) \le u_L(n+1) \le u_L(R_{n+1}) + \hat{a}\beta_{n+1}/(1-\beta_{n+1})w_{n+1},$$

which yields (19).

[63]Equation (15) applies if $Q(E_H(R_{n+1})) - Q(E_H(R_n)) \ge 0$. In the opposite case, the inequality holds trivially since the left-hand side is negative and the right-hand side is positive.

Subtracting $u_H(E_H(R_n))$ from (39) and rearranging (recalling that $w_n = u_H(E_H(R_n)) - u_H(n)$) leads, along any choice sequence, to

$$w_{n+1} = w_n - \eta y_n + \eta w_{n+1} + (1-\eta)\big(u_H\big(E_H(R_{n+1})\big) - u_H\big(E_H(R_n)\big)\big). \tag{36}$$

Combining this with (21) yields $w_{n+1} - w_n \geq \eta w_{n+1} - \eta y_n - b\beta_{n+1} w_{n+1}$, hence (20).

Finally, consider any two rounds $n < n'$, and let Q_L denote P's expected cost conditional on facing L at round n'. L's continuation utility in round n' is bounded above by the utility $u_L(E)$ of the L-efficient contract E that costs Q_L. From (16), Q_L is bounded above by $Q(E_L(R_{n'})) + a\frac{w_{n'}\beta_{n'}}{1-\beta_{n'}}$. This implies, by the analogue of (14) for $\mathcal{E}_L$, that $u_L(n')$ is bounded above by $u_L(R_{n'}) + \hat{u}\frac{w_{n'}\beta_{n'}}{1-\beta_{n'}}$ for some constant $\hat{u} > 0$. Since L's continuation utility is weakly increasing in n (see Lemma 5, which also holds for L), we have $u_L(n') \geq u_L(n) \geq u_L(R_n)$. This yields $u_L(R_{n'}) \geq u_L(R_n) - \hat{u}\frac{w_{n'}\beta_{n'}}{1-\beta_{n'}}$. From (14) applied to $\mathcal{E}_L$ instead of $\mathcal{E}_H$, this implies that $Q(E_L(R_{n'})) \geq Q(E_L(R_n)) - \hat{q}_L\frac{w_{n'}\beta_{n'}}{1-\beta_{n'}}$ for some $\hat{q}_L > 0$. This, together with (15), yields $Q(E_H(R_{n'})) \geq Q(E_H(R_n)) - \hat{q}_H\frac{w_{n'}\beta_{n'}}{1-\beta_{n'}}$, where $\hat{q}_H = b\hat{q}_L$. Combining this result with (14) implies that $u_H(E_H(R_{n'})) - u_H(E_H(R_n)) \geq -\hat{b}\beta_{n'}w_{n'}/(1-\beta_{n'})$, where $\hat{b} = \hat{q}_H/\underline{a}$, which proves (21). *Q.E.D.*

PROOF OF LEMMA 15: Fix some $C \in \mathcal{E}_L$ and consider the referential centered at C whose x-axis is the common tangent of u_L and Q at C, oriented towards $\mathcal{H}$, and whose y-axis is the normal vector pointing northeast in $\mathcal{C}$. The components of a contract, in this referential, are denoted x_t and x_n. We parameterize the iso-utility curve of L going through C, $\mathcal{U}_L(C) = \{R \in \mathcal{C} : u_L(\tilde{C}) = u_L(C)\}$, in terms of x_t: $\{C(x_t) = (x_t, x_n(x_t))\}$. With this parameterization, $C(0) = C$ and $C(x_t) \in \mathcal{H}$ if and only if $x_t \geq 0$.[64]

Let $Q(x_t) = Q(C(x_t))$ and $u_H(x_t) = u_H(C(x_t))$. By L-efficiency of C, $Q'(0) = 0$.[65] Since $-u_L$ and Q are convex, the iso-utility curve of u_L going through C is convex and corresponds to positive values of x_n, while the iso-cost curve going through C is concave and corresponds to negative values of x_n. Moreover, by assumption, at least one of these curves has a nonzero curvature at C. In the (x_t, x_n) space, this means that either $d^2u_L/dx_t^2 > 0$ or $d^2Q/dx_t^2 < 0$. We wish to show the existence of a constant $\hat{q} > 0$ such that $Q(x_t) - Q(0) \geq \hat{q}x_t^2$ for x_t in a right neighborhood of 0. Suppose first that $d^2u_L/dx_t^2 > 0$. This implies that $x_n(x_t) \geq q_x x_t^2$ for some $q_x > 0$ and x_t in a neighborhood of zero. Therefore, $Q(x_t) \geq Q(0) + q_x\|\nabla Q(C)\|x_t^2$ for that neighborhood. Now suppose that $d^2Q/dx_t^2 < 0$. In this case, let $D(x_t)$ denote the contract of the iso-cost curve with x-value x_t in the new referential (hence, just below $C(x_t)$ in the new referential), so that $Q(D(x_t)) = Q(C)$ for all x_t. By tangency of the curves, we have $\|C(x_t) - D(x_t)\| = o(x_t)$. Moreover, $Q(x_t) = Q(C(x_t)) = Q(D(x_t) + \nabla Q(D(x_t)) \cdot (C(x_t) - D(x_t)) + O(\|C(x_t) - D(x_t)\|^2)$ by a standard Taylor expansion. Finally, for x_t in a neighborhood of 0, $\nabla Q(D(x_t)) = \nabla Q(C) + O(\|D(x_t) - C\|) = \nabla Q(C) + o(x_t)$. Combining this, we get $Q(x_t) = Q(D(x_t)) + \nabla Q(C) \cdot (C(x_t) - D(x_t)) + o(x_t)(\|C(x_t) - D(x_t)\| + \|D(x_t) - C\|)$. Since $d^2Q/dx_t^2 < 0$, the

[64]The parameterization is well defined because $\mathcal{U}_L(C)$ can have only one point for each x_t, by strict monotonicity of u_L in the original coordinates (x_1, x_2) and the fact that increasing x_n corresponds to increasing both x_1 and x_2 with at least one of these increases being strict, since the normal vector defining x_n points northeastwards.

[65]Formally, $Q'(x_t) = \frac{\partial Q}{\partial x_1}\frac{dx_1}{dx_t} + \frac{\partial Q}{\partial x_2}\frac{dx_2}{dx_t}$. Since C is L-efficient, Q and u_L are tangent at C. This implies that the tangent vector $(dx_1/dx_t, dx_2/dx_t)$ is orthogonal to the normal vector $(\partial Q/\partial x_1, \partial Q/\partial x_2)$ at C.

y-value of $D(x_t)$ in the new referential satisfies $x_n^D(x_t) \leq -\hat{q}_x x_t^2$ for some $\hat{q}_x > 0$. Hence, $\nabla Q(C)\cdot(C(x_t)-D(x_t)) \geq \tilde{q}_x(x_n(x_t)-x_n^D(x_t)) \geq \check{q}_x x_t^2$ for some positive constants $\tilde{q}_x, \check{q}_x$. These observations imply that

$$Q(x_t) \geq Q(C) + \hat{q}x_t^2 + o(x_t^2), \tag{37}$$

proving the result for the second case.

By compactness of $\mathcal{C}$ and convexity of $\mathcal{U}_L(C)$, $\hat{q}$ may be chosen small enough so that the inequality

$$Q(x_t) - Q(0) \geq \hat{q}x_t^2$$

holds for all nonnegative x_t. Since u_H has bounded derivatives, there must exist $\bar{u} > 0$ such that $|u_H(x_t) - u_H(0)| \leq \bar{u}x_t$ (the single-crossing property between u_H and u_L imply that $u_H(x_t) \leq u_H(0)$ for all $x_t \geq 0$ and that $\nabla u_H(C)\cdot(C(x_t) - C) \neq 0$ for x_t in a neighborhood of 0). Combining these inequalities, there exists $\underline{q}(C) > 0$ such that

$$Q(C_\lambda) - Q(C) \geq \underline{q}(C)\big(u_H(C) - u_H(C_\lambda)\big)^2.$$

Moreover, $\underline{q}(C)$ can be chosen to vary continuously in $C \in \mathcal{E}_L$.[66] By compactness of $\mathcal{E}_L$, $\underline{q} = \min_{C\in\mathcal{E}_L} \underline{q}(C)$ is strictly positive and yields the desired inequality. *Q.E.D.*

PROOF OF LEMMA 16: We have

$$\begin{aligned}
y_n^2 &= \big[\big(u_H\big(E_H(R_n)\big) - u_H\big(E_H(R_{n+1})\big)\big) + \big(u_H\big(E_H(R_{n+1})\big) - u_H(R_{n+1})\big)\big]^2 \\
&\leq 2\big[u_H\big(E_H(R_n)\big) - u_H\big(E_H(R_{n+1})\big)\big]^2 + 2\big[u_H\big(E_H(R_{n+1})\big) - u_H(R_{n+1})\big]^2 \\
&\leq k_1\big(\max\{\beta_n w_n/(1-\beta_n), \beta_{n+1}w_{n+1}\}\big)^2 + 2\big[u_H\big(E_H(R_{n+1})\big) - u_H(R_{n+1})\big]^2 \\
&\leq k_1 \max\big\{\big(\beta_n w_n/(1-\beta_n)\big)^2, (\beta_{n+1}w_{n+1})^2\big\} + k_2\big[Q(R_{n+1}) - Q\big(E_L(R_{n+1})\big)\big] \\
&= k_1 \max\big\{\big(\beta_n w_n/(1-\beta_n)\big)^2, (\beta_{n+1}w_{n+1})^2\big\} \\
&\quad + k_2\big[Q\big(E_L(R_n)\big) - Q\big(E_L(R_{n+1})\big) + Q(R_{n+1}) - Q\big(E_L(R_n)\big)\big].
\end{aligned}$$

The first inequality is standard ($(a+b)^2 \leq 2a^2 + 2b^2$). The second inequality comes from (17) and (21), which taken together imply an upper bound on $|u_H(E_H(R_n)) - u_H(E_H(R_{n+1}))|$. The third inequality comes from the equality $u_H(E_H(R_{n+1})) = u_H(E_L(R_{n+1}))$ and Lemma 15 applied to the contracts $C = E_L(R_{n+1})$ and $R = R_{n+1}$. The difference $Q(E_L(R_n)) - Q(E_L(R_{n+1}))$ is bounded above in proportion to $u_L(R_n) - u_L(R_{n+1})$ (by a simple transposition to $\mathcal{E}_L$ of the proof of (14)), itself bounded by $\gamma\beta_{n+1}w_{n+1}$, from (19). Therefore,

$$y_n^2 \leq k_2\big[Q(R_{n+1}) - Q\big(E_L(R_n)\big)\big] + k_3\big(\max\big\{\big(\beta_n w_n/(1-\beta_n)\big)^2, (\beta_{n+1}w_{n+1})^2\big\} + \beta_{n+1}w_{n+1}\big),$$

where $k_3 = \max\{k_1, \gamma k_2\}$, which yields the result. *Q.E.D.*

[66]Indeed, all the constants involved in the previous steps are based on the curvature of the iso-utility and iso-cost curves at C, which only involve the second derivative of the utility and cost functions at C. These functions were assumed to be C^2 over $\mathcal{C}$.

APPENDIX F: PROOFS FOR PART I

PROOF OF LEMMA 4: Let $\underline{u}_H = u_H(R_0)$ and, for any $\tilde{\epsilon} \geq 0$,

$$D(\tilde{\epsilon}) = \inf\{Q(C) - Q(E) : C \in \mathcal{H}, E \in \mathcal{E}_H : \underline{u}_H \leq u_H(E) \leq u_H(E_H(C)) + \tilde{\epsilon}\}. \quad (38)$$

$D(\tilde{\epsilon})$ is nonincreasing in $\tilde{\epsilon}$, as a higher $\tilde{\epsilon}$ only increases the set of (C, E) pairs over which the objective is minimized. Since R_0 is regular, the contracts C arising in (38) are bounded away from $\mathcal{E}_H$ for $\tilde{\epsilon}$ small enough, which implies that $D(\tilde{\epsilon})$ is strictly positive.[67] We choose any ε such that $D(\varepsilon) > 0$ and set $D = D(\varepsilon)$.

From Lemma 5 (whose proof, below, is independent of the present lemma), $u_H(n) \geq u_H(m)$ for any round $n \geq m$, and Proposition 1, part (iv), yields $u_H(n) \leq u_H(E_H(R_n))$. Let C denote the L-efficient contract that gives H utility $u_H(m)$. We have $u_H(C) = u_H(m) \leq u_H(n) \leq u_H(E_H(R_n)) = u_H(E_L(R_n))$, which implies that $u_L(C) \leq u_L(E_L(R_n))$.[68] Therefore, R_n must cost weakly more than C. By assumption, $u_H(E_H(R_m)) - u_H(C) = u_H(E_H(R_m)) - u_H(m) = w_m \leq \varepsilon$. Since also $u_H(R_m) \leq u_H(E_H(R_m))$, (38) implies that $Q(C) \geq Q(E_H(R_m)) + D$. Combining this with $Q(R_n) \geq Q(C)$ proves the lemma. *Q.E.D.*

PROOF OF LEMMA 5: Let R_{n+1} denote any contract chosen by H with positive probability among $R_n \cup \{M_n\}$. H's utility satisfies the dynamic equation[69]

$$u_H(n) = \eta u_H(R_{n+1}) + (1 - \eta) u_H(n + 1). \quad (39)$$

Therefore, $u_H(n)$ is a convex combination of $u_H(R_{n+1})$ and $u_H(n+1)$. Because H can hold on to R_{n+1} in all rounds $m \geq n + 1$, $u_H(n + 1)$ is bounded below by $u_H(R_{n+1})$. Combining these observations yields $u_H(n) \leq u_H(n + 1)$. Since $u_H(n + 1) - u_H(n) = \eta(u_H(n + 1) - u_H(R_{n+1}))$, the second claim follows. *Q.E.D.*

LEMMA 18: *For any round* n_0, $\check{\varepsilon} > 0$, *and choice sequence*, *there exists a round* $n > n_0$ *such that* $u_H(R_n) \geq \max\{u_H(E_H(R_m)) : m \leq n_0\} - \check{\varepsilon}$.

PROOF: Fix $\check{\varepsilon} > 0$. Proposition 2 guarantees that, along any choice sequence, R_n converges to some L-efficient contract $\bar{C}_L$. Continuity of $u_H(\cdot)$ implies that there exists a round $\check{n}$ such that $u_H(R_n) \geq u_H(\bar{C}_L) - \check{\varepsilon}$ for all $n \geq \check{n}$. Therefore, it suffices to show that $u_H(\bar{C}_L) \geq \max\{u_H(E_H(R_m)) : m \leq n_0\}$ for all n_0. Equivalently, we must show that $u_H(\bar{C}_L) \geq \max\{u_H(E_L(R_m)) : m \leq n_0\}$ for all n_0 since, by construction, $E_H(R)$ and $E_L(R)$ give the same utility to H for any $R \in \mathcal{H}$. For contracts C, C' on the L-efficiency curve $\mathcal{E}_L$, $u_H(C) \leq u_H(C')$ if and only if $u_L(C) \leq u_L(C')$. Therefore, it suffices to show that $u_L(\bar{C}_L) \geq \max_{m \in \mathbb{N}}\{u_L(E_L(R_m))\}$. By construction, $u_L(E_L(R)) = u_L(R)$ for all $R \in \mathcal{H}$. It thus suffices to show that

$$u_L(\bar{C}_L) \geq \max_{m \in \mathbb{N}}\{u_L(R_m)\}.$$

[67]For any C entering the definition of $D(\tilde{\epsilon})$, $u_H(E_L(C)) = u_H(E_H(C)) \geq u_H(R_0) - \tilde{\epsilon}$. Regularity of R_0 implies that any contract for which this inequality holds for $\tilde{\epsilon} = 0$ is strictly H-inefficient. By compactness of the contract space and continuity of u_H and Q, this inefficiency must be bounded below by a positive constant for $\tilde{\epsilon} = 0$, and thus also for all $\tilde{\epsilon}$ small enough. Since C gives H a utility within $\tilde{\epsilon}$ of the H-efficient contract E, the cost $Q(C)$ must exceed $Q(E)$ by a strictly positive amount provided that $\tilde{\epsilon}$ is small enough. Again by compactness, this amount is uniformly bounded below a strictly positive constant.

[68]For contracts C, C' on the L-efficiency curve $\mathcal{E}_L$, $u_H(C) \leq u_H(C')$ if and only if $u_L(C) \leq u_L(C')$.

[69]More generally, H's utility satisfies the Bellman equation $u_H(n) = \max_{R \in \{R_n\} \cup M_n}\{\eta u_H(R) + (1 - \eta) u_H(n + 1)\}$. Equation (39) then follows for all contracts that are optimal for H in round n.

We recall that for all n, $u_L(n) \geq u_L(R_n)$ since holding on to R_n is always a feasible strategy for L, and that $u_L(n)$ is nondecreasing in n for all choice sequences (see Lemma 5; the argument there also applies to L). Since R_n converges to $\bar{C}_L$, $u_L(n)$ must converge to $u_L(\bar{C}_L)$. Finally, because $u_L(n)$ is nondecreasing, we have

$$u_L(R_n) \leq u_L(n) \leq u_L(\bar{C}_L),$$

for all n, which concludes the proof. *Q.E.D.*

PROOF OF LEMMA 6: P's gain from reducing H's rent between rounds n_0 and n_1 is bounded above by $\hat{\beta}_0(1-\hat{\mu}_0)a(\hat{e}_0-\hat{u}_0)$ for some Lipschitz constant $a > 0$. To see this, first note that $\hat{\beta}_0(1-\hat{\mu}_0)$ is the probability that the agent is of type H *and* that accepts some H-efficient contract during the first block. Moreover, because H accepts only H-efficient contracts that give him at least his continuation utility,[70] and because this continuation utility is nondecreasing (Lemma 5), the smallest utility that H receives when choosing an H-efficient contract within this first block is $\hat{u}_0$. Finally, $\hat{e}_0 = u_H(E_H(R_{n_0}))$ is the utility that P provides to H if he chooses the immediate jump. Therefore, the maximum rent that P can take away from H is $\hat{e}_0 - \hat{u}_0$. The constant a is a Lipschitz constant that "translates" utility differences for H along $\mathcal{E}_H$ into cost differences for P, derived in Lemma 13. Similarly, the expected net gain made after round n_1, but seen from round n_0, is bounded above by $\hat{\beta}_0\hat{\mu}_0 a(\hat{e}_0 - \hat{u}_1)$, because $\hat{\beta}_0\hat{\mu}_0$ is the probability of facing H *and* reaching round n_1, and $\hat{u}_1$ is the smallest utility that P must provide to H at any round following n_1.

To compute a lower bound on the inefficiency loss, we note that as long as H accepts contracts in $\mathcal{H}$, Lemma 4 implies that any breakdown causes an inefficiency cost greater than $D > 0$.[71] The number of rounds between n_0 and n_1 is bounded below by $\underline{n}(1) = \lfloor(\hat{u}_1 - \hat{u}_0)/\eta\Delta_H)\rfloor$, as explained in the main text, below Lemma 5. The probability of a breakdown before the end of the block is thus bounded below by[72]

$$\begin{aligned} 1-(1-\eta)^{\underline{n}(1)} &= 1-\exp\bigl(\underline{n}(1)\ln(1-\eta)\bigr) \\ &\geq -\underline{n}(1)\ln(1-\eta) - \frac{1}{2}\underline{n}(1)^2\bigl(\ln(1-\eta)\bigr)^2. \end{aligned} \tag{40}$$

Because the gain is of order ε, which is small, while the loss conditional on a breakdown is of order D, the probability of a breakdown must be (at most) of order ε, which means that $\underline{n}(1)\ln(1-\eta)$ must also be small (of order ε, from the second expression in (40)). The quadratic term of (40) is therefore of order ε^2 and will be neglected (alternatively, the breakdown probability could be slightly scaled down to account for this term without changing the analysis). Moreover, the analysis is concerned with the limit as η goes to zero, $\ln(1-\eta)$ can be approximated by $-\eta$. Combining these bounds on gains and losses proves the lemma.[73] *Q.E.D.*

[70]By accepting such a contract, H reveals his type. His continuation utility is thus equal to the utility provided by the last accepted contract (Proposition 1, Part (i)).

[71]The lower bound D is valid if the rent reduction index at the beginning of a block is less than ε, which holds without loss of generality as explained in Remark 2 below.

[72]The inequality comes from the standard inequality $1-\exp(x) \geq -x - x^2/2$, valid for all $x \leq 0$.

[73]For expositional simplicity, the "floor" operator is dropped. This change is negligible because $\underline{n}(1)$ is large since $\hat{u}_1 - \hat{u}_0 = \frac{1}{t}(\hat{e}_0 - \hat{u}_0) \gg \eta\Delta_H$, for η small. This observation applies to each block k as explained in Footnote 24.

LEMMA 19: *There exists a pushdown sequence for Block* 1.

PROOF: Let $\mu^\theta(\{\tilde{R}_n\})$ denote the probability, conditional on facing type θ, of observing a choice sequence $\{\tilde{R}_n\}$ until $\hat{u}_1$ is reached. By definition, summing over all choice sequences with elements in $\mathcal{H}$ truncated at the first round at which H's continuation utility reaches $\hat{u}_1$, we have $\sum_{\{\tilde{R}_n\}} \mu^H(\{\tilde{R}_n\}) = \mu_0$. Because L always chooses contracts in $\mathcal{H}$, we also have $\sum_{\{\tilde{R}_n\}} \mu^L(\{\tilde{R}_n\}) = 1$. These two equations imply that there exists a choice sequence $\{R_n^0\}$ such that $\mu^H(\{R_n^0\})/\mu^L(\{R_n^0\}) \le \mu_0$. Conditional on observing this sequence, the posterior is given by Bayesian updating

$$\hat{\beta}_1 = \frac{\mu^H(\{R_n^0\})\hat{\beta}_0}{\mu^H(\{R_n^0\})\hat{\beta}_0 + \mu^L(\{R_n^0\})(1-\hat{\beta}_0)}.$$

Dividing by $\mu^L(\{R_n^0\})$ and using that $\mu^H(\{R_n^0\})/\mu^L(\{R_n^0\}) \le \mu_0$ yields the result. *Q.E.D.*

PROOF OF LEMMA 7: Let $e^* = \max\{\hat{e}_k : k \le K\}$. For each block $k \le K$, we have

$$e^* - \hat{u}_k = (e^* - \hat{u}_{k+1}) + (\hat{u}_{k+1} - \hat{u}_k) \le (e^* - \hat{u}_{k+1}) + \frac{1}{t-1}(\hat{e}_k - \hat{u}_{k+1}) \le \frac{t}{t-1}(e^* - \hat{u}_{k+1}).$$

By induction, $e^* - \hat{u}_0 \le [t/(t-1)]^K (e^* - \hat{u}_K)$. Since $\hat{e}_0 \le e^*$, this implies that

$$w_0 = \hat{e}_0 - \hat{u}_0 \le \left(\frac{t}{t-1}\right)^K (e^* - \hat{u}_K). \tag{41}$$

Applying (21) to rounds n_k and n_K and observing that $w_{n_K} \le \hat{w}_K$ (since $u_H(n_K)$ exceeds $\hat{u}_K$, by definition of n_K), we have

$$\hat{e}_K \ge \hat{e}_k - \frac{\hat{b}\beta_K}{1-\beta_K}\hat{w}_K \quad \text{for each } k \le K. \tag{42}$$

Since $\hat{w}_K \le \bar{W}\eta$ and $\beta_K \le \beta_0$, this yields $\hat{e}_K \ge e^* - \hat{e}\eta$, where $\hat{e} = \hat{b}\bar{W}\beta_0/(1-\beta_0)$. Therefore, $e^* - \hat{u}_K = e^* - \hat{e}_K + \hat{w}_K \le \hat{e}\eta + \bar{W}\eta \le c_w\eta$, where $c_w = \hat{e} + \bar{W}$. Combining this with (41) shows the first part of the lemma.

For the second part, we have

$$\hat{w}_K = \hat{e}_K - \hat{u}_K \ge \hat{e}_{K-1} - \frac{\hat{b}\hat{\beta}_K}{1-\hat{\beta}_K}\hat{w}_K - \hat{u}_K = (\hat{e}_{K-1} - \hat{u}_{K-1}) - (\hat{u}_K - \hat{u}_{K-1}) - \frac{\hat{b}\hat{\beta}_K}{1-\hat{\beta}_K}\hat{w}_K,$$

where the inequality come from (42) applied to block $K-1$. Since $\hat{e}_{K-1} - \hat{u}_{K-1} = t(\hat{u}_K - \hat{u}_{K-1})$ and $\hat{\beta}_K \le \beta_0$, this yields

$$\hat{w}_K\left(1 + \frac{\hat{b}\beta_0}{1-\beta_0}\right) \ge \frac{t-1}{t}(\hat{e}_{K-1} - \hat{u}_{K-1}).$$

By definition of K, $\hat{e}_{K-1} - \hat{u}_{K-1} \ge \bar{W}\eta$. This, together with $\bar{W} \ge \frac{t}{t-1}(1 + \hat{b}\beta_0/(1-\beta_0))(1 + \Delta_H)$, yields

$$\hat{w}_K \ge (1 + \Delta_H)\eta.$$

Finally, since $w_{n_K} = \hat{w}_K - (u_H(n_K) - \hat{u}_K)$ and $u_H(n_K) - \hat{u}_K \leq \Delta_H \eta$, by Lemma 5,[74] we get

$$w_{n_K} = \hat{e}_K - u_H(n_K) \geq \eta. \tag{43}$$

Q.E.D.

LEMMA 20: *Suppose there exists $K_w > 0$ such that $w_0 \leq K_w \eta$ whenever $w_0 \leq \varepsilon$. Then, for all η small enough and corresponding equilibrium, $w_0 \leq K_w \eta$.*

PROOF: It suffices to show that w_0 cannot exceed ε when η is small enough. Suppose by contradiction that there exists an equilibrium, for η arbitrarily small, such that $w_0 > \varepsilon$. Following any choice sequence along which β_n is non-increasing,[75] let n_0 denote the first round such that $w_{n_0} \leq \varepsilon$. The lemma's hypothesis implies that $w_{n_0} \leq K_w \eta$. By construction, $w_{n_0-1} > \varepsilon$. Recalling that $w_n = u_H(E_H(R_n)) - u_H(n)$, we have

$$w_{n_0} = w_{n_0-1} + \big(u_H(n_0 - 1) - u_H(n_0)\big) + \big(u_H\big(E_H(R_{n_0})\big) - u_H\big(E_H(R_{n_0-1})\big)\big).$$

The middle term is of order η by (39) and hence negligible compared to ε for η small enough. Moreover, (21) implies that $u_H(E_H(R_{n_0})) - u_H(E_H(R_{n_0-1})) \geq -\hat{b}\beta_{n_0} w_{n_0}$. Combining this,

$$w_{n_0} \geq \frac{w_{n_0-1} - O(\eta)}{1 + \hat{b}\beta_{n_0}}.$$

Setting $\underline{\varepsilon} = \frac{\varepsilon}{2(1+\hat{b})}$, we thus have $w_{n_0} \geq \underline{\varepsilon}$. For η small enough, this contradicts the lemma's condition that $w_{n_0} \leq K_w \eta$. *Q.E.D.*

REMARK 2: One could reach a block k at the end of which $\hat{w}_k$ exceeds ε, invalidating the inefficiency lower bound D for the following block. In this case, the argument just used to prove Lemma 20 implies that there is a continuation choice sequence and later round n for which $w_n \in (\underline{\varepsilon}, \varepsilon)$ and $\beta_n \leq \hat{\beta}_{k-1}$. After reaching this round, we resume the block construction. One may encounter a new block for which the problem arises again, in which case we repeat the previous step. Since w_n converges to zero along any choice sequence as n goes to infinity, there can only be finitely many such iterations: one must reach a round n_0 such that $w_{n_0} \in (\underline{\varepsilon}, \varepsilon)$ and $\hat{w}_k$ remains below ε for all blocks constructed from n_0. Moreover, by construction we have $\beta_{n_0} \leq \beta_0$, because the posterior decreases across each of previous steps. The analysis applied from this round, instead of round 0, shows that w_{n_0} is of order η.[76]

F.1. *Complement to Proposition 4*

This appendix shows that the limitations of Proposition 4 are robust to changes in the parameter t used to determine the size of each block. Let $\bar{t} = 1 + \frac{D}{a\Delta_H}$. To guarantee

[74] The reason for using $u_H(n_K)$ instead of $\hat{u}_K$ is that H's continuation utility at round n_K may slightly overshoot $\hat{u}_K$: it is only guaranteed to lie between $\hat{u}_K$ and $\hat{u}_K + \Delta_H \eta$.

[75] Since L puts probability 1 on contracts in $\mathcal{H}$, there must be in each round n a contract R_{n+1} chosen by L with a weakly higher probability than by H, implying that $\beta_{n+1} \leq \beta_n$.

[76] This, for η small enough, shows that $\hat{w}_k$ could not have exceeded ε and thus rules out, ex post, the possibility considered by this remark.

that $\hat{u}_k \geq \hat{u}_{k-1}$ and $\bar{\mu}_k \leq 1$ for all blocks k of Part I, t must lie in $(1, \bar{t})$, the value used in Section 5 corresponding to $t = \sqrt{\bar{t}}$. For other values of t, the relation $w_0 \leq c_w(t/(t-1))^K \eta$ is the same as before, as is easily checked. What changes is the value of g, which is now equal to $1/(\beta_0 + (1-\beta_0)(\bar{t}/t))$ and implies a new value for ρ_0 of

$$\rho_0(t) = \frac{\ln\big(t/(t-1)\big)}{\ln(\bar{t}/t)}.$$

As $\bar{t}$ gets close to 1, the numerator goes to $+\infty$ and the denominator goes to 0, both uniformly over $t \in (1, \bar{t})$. In particular, $\inf_{t\in(1,\bar{t})} \rho_0(t) > 1/3$. This shows that Proposition 4 cannot be modified to cover all primitives of the model.

APPENDIX G: PROOFS FOR PART II

PROOF OF LEMMA 8: Equation (19) together with (14) applied to $\mathcal{E}_L$ instead of $\mathcal{E}_H$ implies the existence of $k_Q > 0$ such that

$$Q(R_{n+1}) - Q\big(E_L(R_n)\big) \geq Q\big(E_L(R_{n+1})\big) - Q\big(E_L(R_n)\big) \geq -k_Q \beta_{n+1} w_{n+1}. \tag{44}$$

Letting $\mu_n = \mu_n(R_{n+1})$, this equation, together with (7), yields

$$\beta_n w_n a \geq \beta_n \mu_n \eta D - \eta k_Q \beta_{n+1} w_{n+1}$$

or

$$\mu_n \leq \frac{w_n a}{\eta D} + k_Q \frac{\beta_{n+1}}{\beta_n D} w_{n+1}. \tag{45}$$

Since $w_N \leq \frac{\eta D}{2a}$, (18) implies that $w_{N+1} \leq \frac{\eta D}{2a}(1 + \alpha\beta_N/(1-\beta_N))$. Combined with (45), this shows that $\mu_N \leq \frac{1}{2} + O(\eta) \leq \frac{3}{5}$.

From Bayesian updating, we have $\beta_{n+1} = \frac{\mu_n \beta_n}{\mu_n\beta_n + (1-\beta_n)}$. Since $\beta_N \leq \hat{\beta}$, we can choose $\hat{\beta}$ small enough to guarantee that the denominator exceeds $1-\epsilon$ for $n = N$. More generally, we have

$$\beta_{n+1} \leq \mu_n \beta_n (1+\epsilon), \tag{46}$$

where ϵ is a small positive constant, as long as $\beta_n \leq \hat{\beta}$.

Consider the first round $M > N$ for which $\mu_M \geq 3/4$. The sequence β_n is decreasing[77] from $n = N$ until at least round M. Proceeding by induction from round N to round M and using (18), the previous inequalities imply that

$$w_{N+m} \leq w_N \prod_{i=0}^{m-1} \big(1 + \alpha(1-\hat{\beta})^{-1} \beta_{N+i}\big) \tag{47}$$

and

$$\beta_{N+i} \leq \beta_N \prod_{j=0}^{i-1} \big(\mu_{N+j}(1+\epsilon)\big). \tag{48}$$

[77]From Bayes's rule, $\beta_{n+1} = \frac{\mu_n \beta_n}{\mu_n\beta_n + (1-\beta_n)}$, which is nondecreasing in μ_n. Taking $\mu_n = 1$ shows that $\beta_{n+1} \leq \beta_n$ as long as $\mu_n \leq 1$.

For $j \leq N - M$, $\mu_{N+j} \leq 3/4$, so

$$\beta_{N+i} \leq \left(\frac{3(1+\epsilon)}{4}\right)^i \beta_N.$$

Equation (47) then implies that

$$w_M \leq w_N \prod_{i=1}^{M-N} \left(1 + \alpha(1-\hat{\beta})^{-1}\hat{\beta}\left(\frac{3(1+\epsilon)}{4}\right)^i\right).$$

The product

$$\prod_{i=1}^{\infty} \left(1 + \alpha(1-\hat{\beta})^{-1}\hat{\beta}\left(\frac{3(1+\epsilon)}{4}\right)^i\right) \tag{49}$$

is finite for $\epsilon < 1/3$, and converges to 1 as $\hat{\beta}$ goes to zero.[78] Therefore, for $\hat{\beta}$ small enough, w_M is bounded above by $\frac{5}{4} w_N \leq \frac{5\eta D}{8a}$. From (45), this implies that μ_M is bounded above by $5/8 + O(\eta) < 3/4$, so M cannot be finite. This shows that for β_N below some threshold $\hat{\beta}$, μ_n is bounded above by 3/4 for all $n \geq N$ and, from (46), that β_n is decreasing. Since w_n is bounded above by $\frac{3}{2} w_N$ and $w_N \leq \frac{\eta D}{2a}$, Part (iii) follows easily.

From Part (ii) and the fact that $\beta_{n+1} \sim \beta_n \mu_n$ (which comes from Bayes's rule), the second term in the right-hand side of (45) is of order w_{n+1} and thus negligible compared to the first one, which is of order $\frac{w_n}{\eta}$. Therefore, by slightly increasing a, whose specific value does not matter for the proof, we get (8) and (9). *Q.E.D.*

LEMMA 21: *There exists a positive constant* $\bar{A}$ *such that*

$$y_n^2 \leq \frac{\bar{A}\beta_{n+1}}{1-\beta_0}. \tag{50}$$

PROOF: Equation (9) together with $\beta_n \leq \beta_0$ implies that $Q(R_{n+1}) - Q(E_L(R_n)) \leq \frac{\beta_n w_n a}{\eta(1-\beta_0)}$. This, along with (8), yields[79]

$$Q(R_{n+1}) - Q\big(E_L(R_n)\big) \leq \frac{D\beta_{n+1}}{1-\beta_0}.$$

Combining this inequality with Lemma 16 yields

$$y_n^2 \leq k_2 \frac{D\beta_{n+1}}{1-\beta_0} + k_3\Big(\max\big\{\big(\beta_n w_n/(1-\beta_n)\big)^2, (\beta_{n+1}w_{n+1})^2\big\} + \beta_{n+1}w_{n+1}\Big).$$

[78] Indeed, taking the logarithm of that product, we obtain a sequence that is approximately geometric with geometric factor $3(1+\epsilon)/4$ and, hence, converges, uniformly in $\hat{\beta}$. Moreover, each term of the sequence converges to 0 as $\hat{\beta}$ goes to zero. This implies that all partial sums converge to zero and, by uniform convergence, that the sequence converges to zero as well. By continuity of the exponential function, the product itself thus converges to 1 as $\hat{\beta}$ goes to zero.

[79] We are using $\beta_{n+1} \geq \mu_n \beta_n$, which is implied by the Bayesian updating equation $\beta_{n+1} = \frac{\mu_n \beta_n}{\mu_n \beta_n + (1-\beta_n)}$ and the fact that $\mu_n \leq 1$, from Lemma 8.

Since $w_{n+1} \leq \frac{\eta D}{2a} \ll 1$, the last term is negligible compared to β_{n+1}, as is the maximum since β_{n+1} is of order $\beta_n \mu_n \gg \beta_n w_n$. Choosing $\bar{A}$ slightly greater than $k_2 D$ yields the lemma. *Q.E.D.*

PROOF OF LEMMA 9: From (36), we have

$$w_{n+1} = w_n - \eta y_n + \eta w_{n+1} + (1-\eta)\big(u_H\big(E_H(R_{n+1})\big) - u_H\big(E_H(R_n)\big)\big).$$

Combining this with (21) yields

$$(1-\eta) w_{n+1} \geq w_n - \eta y_n - \hat{b}\beta_{n+1} w_{n+1}.$$

Using the upper bound for y_n^2 provided by Lemma 21, this implies that

$$w_{n+1} \geq w_n - \eta \sqrt{\frac{\bar{A}}{1-\beta_0}} \sqrt{\beta_{n+1}} - \hat{b}\beta_{n+1} w_{n+1}. \tag{51}$$

Since $w_n \leq \hat{w}\eta$ for $n \geq N$, by Part (iii) of Lemma 8, the last term is negligible compared to the penultimate one. Slightly increasing the value of $\bar{A}$, whose precise value does not affect the proof, and letting $A_w = \sqrt{\frac{\bar{A}}{1-\beta_0}}$, this yields

$$w_{n+1} \geq w_n - \eta A_w \sqrt{\beta_{n+1}}.$$ *Q.E.D.*

PROOF OF LEMMA 10: Let $q_n = \frac{a w_n}{\eta D}$. Equation (11) may be re-expressed as

$$q_{n+1} \geq q_n - c\sqrt{\beta_{n+1}} \tag{52}$$

with $c = \frac{a}{D} A_w$. The Bayesian updating equation

$$\beta_{n+1} = \frac{\beta_n \mu_n}{\beta_n \mu_n + (1-\beta_n)}$$

implies that[80]

$$\frac{\beta_{n+1}}{\beta_n} \leq \mu_n + \beta_{n+1} + O(\beta_n \beta_{n+1}). \tag{53}$$

From (8), this implies that $\frac{\beta_{n+1}(1-O(\beta_n))}{\beta_n} \leq \frac{a w_n}{\eta D} = q_n$. By slightly increasing a, whose precise value does not matter anyway, we can get rid of the term $O(\beta_n)$, which shows that

$$\frac{\beta_{n+1}}{\beta_n} \leq q_n \tag{54}$$

for all $n \geq N$. Since $\beta_{n+1} = \beta_N \times \prod_{k=N}^{n} \frac{\beta_k+1}{\beta_k}$, (52) and (54) yield

$$q_{n+1} \geq q_n - c' \sqrt{\Pi_N^n q_k}, \tag{55}$$

[80]We have $\frac{\beta_{n+1}}{\beta_n} = \mu_n \frac{1}{1-\beta_n(1-\mu_n)} = \mu_n(1+\beta_n(1-\mu_n)) + \mu_n O(\beta_n^2)$. Rearranging the expression yields $\mu_n \geq \frac{\beta_{n+1}}{\beta_n} - \mu_n\beta_n + \mu_n O(\beta_n^2)$. Since the Bayesian updating equation also implies that $\mu_n \leq \frac{\beta_{n+1}}{\beta_n}$, the last two terms are respectively bounded in absolute value by β_{n+1} and $O(\beta_{n+1}\beta_n)$.

where $c' = \sqrt{\beta_N}c$. The first hypothesis of Lemma 10 implies that

$$q_N \geq 4c^2\beta_N. \tag{56}$$

Taking the square root on both sides of (56) and multiplying the result by $\sqrt{\frac{\beta_{N+1}}{\beta_N}}$,

$$\sqrt{\frac{q_N\beta_{N+1}}{\beta_N}} \geq 2c\sqrt{\beta_{N+1}}.$$

From (54), this implies that $q_N \geq 2c\sqrt{\beta_{N+1}}$. Combining this with (52) then yields

$$q_{N+1} \geq c\sqrt{\beta_{N+1}}.$$

Taking the square root of this expression and dividing both sides by $\sqrt{\beta_{N+1}}$, we get

$$\sqrt{\frac{q_{N+1}}{\beta_{N+1}}} \geq \frac{\sqrt{c}}{\beta_{N+1}^{1/4}}. \tag{57}$$

The second hypothesis of the lemma may be re-expressed as

$$\beta_N^{1/4} \leq \frac{1}{2\sqrt{c}}. \tag{58}$$

Since $\beta_{N+1} \leq \beta_N$, (58) implies that the RHS of (57) is greater than $2c$ and, therefore, that (56) holds at round $N+1$. Since β_n is non-increasing in n for $n \geq N$ and hence satisfies (58) for all $n \geq N$, we can apply the previous argument by induction to conclude that (56) and (57) hold for all $n \geq N$. Dividing (52) by q_n yields

$$\frac{q_{n+1}}{q_n} \geq 1 - c\frac{\sqrt{\beta_{n+1}}}{q_n} = 1 - c\sqrt{\frac{\beta_{n+1}}{\beta_n q_n}}\sqrt{\frac{\beta_n}{q_n}}.$$

From (54), the first factor of the last term is bounded above by 1, and from (57) applied to round n, the second factor is bounded above by $\beta_n^{1/4}/\sqrt{c}$, which converges to zero as n goes to infinity. *Q.E.D.*

PROOF OF LEMMA 11: Suppose by contradiction that $\{q_n\}$ converges to zero. This together with the second assumption of the lemma implies the existence, for any fixed $\varepsilon > 0$, of an integer $\bar{N} \geq N$ such that (i) $\frac{q_{n+1}}{q_n} \geq 1 - \varepsilon$ and (ii) $q_n \leq q_{\bar{N}} \leq \varepsilon$ for all $n \geq \bar{N}$.[81] Let $\Pi^* = \Pi_N^{\bar{N}} q_k$ and $\tilde{\varepsilon} = \sqrt{q_{\bar{N}}}$. We have $\Pi_{\bar{N}+1}^{\bar{N}+k} q_k \leq \tilde{\varepsilon}^{2k}$ for all $k \geq 1$. Therefore, for any $K \geq 1$,

$$q_{\bar{N}+K} = q_{\bar{N}+K} - q_\infty = \sum_{n \geq \bar{N}+K} (q_n - q_{n+1}) \leq \tilde{c}\tilde{\varepsilon}^K \sum_{k \geq 0} \tilde{\varepsilon}^k,$$

[81] Indeed, there exist N_1 such that (i) holds for all $n \geq N_1$ and N_2 such that $q_n \leq \varepsilon$ for all $n \geq N_2$. Letting $N_3 = \max\{N_1, N_2, N\}$, any $\bar{N} \in \arg\max_{n \geq N_3}\{q_n\}$ satisfies conditions (i) and (ii).

where $\tilde{c} = c'\sqrt{\Pi^*}$ and the last inequality comes from the first hypothesis of the lemma. Taking $K = 3$ and using $\sum_{k\geq 0} \tilde{\varepsilon}^k = 1/(1-\tilde{\varepsilon})$, this yields

$$q_{\bar{N}+3} \leq \frac{c'}{1-\tilde{\varepsilon}} q_{\bar{N}}^{3/2} \leq 2c' q_{\bar{N}}^{3/2}. \tag{59}$$

Applying inequality (i) above to $n = \bar{N}$, $\bar{N}+1$, and $\bar{N}+2$ yields

$$q_{\bar{N}+3} \geq q_{\bar{N}}(1-\varepsilon)^3. \tag{60}$$

Combining (59) and (60), we get $(1-\varepsilon)^3 \leq 2c' q_{\bar{N}}^{1/2} \leq 2c'\varepsilon^{1/2}$, which is impossible for ε small enough and yields the desired contradiction. *Q.E.D.*

PROOF OF PROPOSITION 6: Suppose by contradiction that $w_N > \hat{c}\eta\beta_N$. Since $\beta_N \leq \tilde{\beta} \leq c^{-2}/16$, the hypotheses of Lemma 10 are satisfied and $\liminf_{n\to+\infty} \frac{q_{n+1}}{q_n} \geq 1$. Lemma 11 then implies that w_n does not converge to zero and contradicts Proposition 2. *Q.E.D.*

APPENDIX H: PROOFS FOR PART III

PROOF OF LEMMA 12: Consider any round n and contract R_{n+1} in $(M_n \cup \{R_n\}) \cap \mathcal{H}$. If $\mu_n^L(R_{n+1}) \geq \bar{\varepsilon}\mu_n^H(R_{n+1})$, then $\mu_n(R_{n+1}) \leq \frac{1}{\bar{\varepsilon}}$ and, by Bayesian updating, $\beta_{n+1} \leq \frac{\beta_n}{\bar{\varepsilon}}$.[82] Contracts R_{n+1} for which $\mu_n^L(R_{n+1}) \leq \bar{\varepsilon}\mu_n^H(R_n+1)$ arise with probability at most $\bar{\varepsilon} + \beta_n$, because L chooses among these contracts with probability at most $\bar{\varepsilon}$ and H has probability β_n.[83] If $\beta_n \leq \bar{\varepsilon}$, this implies that with probability at least $1-2\bar{\varepsilon}$,

$$\beta_{n+1} \leq \frac{\beta_n}{\bar{\varepsilon}}.$$

At round $\bar{n}$, we have $\beta_{\bar{n}} \leq \bar{\varepsilon}^{\bar{N}}$, which implies that $\beta_{\bar{n}+1} \leq \bar{\varepsilon}^{\bar{N}-1} \leq \bar{\varepsilon}$ with probability at least $1-2\bar{\varepsilon}$. By induction, with probability at least $(1-2\bar{\varepsilon})^{\bar{N}}$ we have

$$\beta_n \leq \bar{\varepsilon}^{\bar{N}-(n-\bar{n})} \leq \bar{\varepsilon}$$

for all $n \in \{\bar{n}, \ldots, \bar{n}+\bar{N}\}$. Since $w_{\bar{n}} \leq \bar{W}\eta$, (18) implies that $w_{\bar{n}+1} \leq k_1\eta$ for some constant k_1. Applying this reasoning by induction to rounds $n = \bar{n}, \ldots, \bar{n}+\bar{N}-1$ shows that with probability at least $(1-2\bar{\varepsilon})^{\bar{N}}$,

$$\beta_n \leq \bar{\varepsilon}^{-\bar{N}}\beta_{\bar{N}} \leq \bar{\varepsilon} \quad \text{and} \quad w_n \leq k_w\eta$$

for all $n \in \{\bar{n}, \ldots, \bar{n}+\bar{N}-1\}$, where k_w is independent of $\bar{\varepsilon}$ and η.[84]

Consider any choice sequence such that β_n and w_n satisfy the above inequalities throughout the block. From (5), we have

$$\sum_{R_{n+1}\in(\mathcal{M}_n\cup\{R_n\})\cap\mathcal{H}} \mu_n^L(R_{n+1})\big(Q(R_{n+1}) - Q(E_L(R_n))\big) \leq \frac{aw_n\beta_n}{\eta(1-\beta_n)}.$$

[82]Equation (53) implies that $\beta_{n+1} \sim \mu_n(R_{n+1})\beta_n$, the term $\beta_{n+1}\beta_n$ being negligible.

[83]Formally, let $\mathcal{A}_n$ denote the set of such contracts and $\mu^\theta(\mathcal{A}_n)$ denote the probability that θ chooses a contract in $\mathcal{A}_n$. By assumption, we have $\mu^L(\mathcal{A}_n) \leq \bar{\varepsilon}\mu^H(\mathcal{A}_n) \leq \bar{\varepsilon}$. Therefore, the unconditional probability that the agent chooses a contract in $\mathcal{A}_n$ is bounded above by $(1-\beta_n)\mu^L(\mathcal{A}_n) + \beta_n\mu^H(\mathcal{A}_n) \leq \bar{\varepsilon} + \beta_n$.

[84]k_w depends on $\bar{N}$, which is set independently of this lemma so as to satisfy the condition $a\bar{W}/D\bar{N} < 1/4$.

While some terms in this sum may be negative, (44) permits to bound these terms below:

$$Q(R_{n+1}) - Q(E_L(R_n)) \geq Q(E_L(R_{n+1})) - Q(E_L(R_n)) \geq -k_Q \beta_{n+1} w_{n+1}.$$

This implies that

$$\begin{aligned} &\sum_{R_{n+1} \in (\mathcal{M}_n \cup \{R_n\}) \cap \mathcal{H}} \mu_n^L(R_{n+1}) \left| Q(R_{n+1}) - Q(E_L(R_n)) \right| \\ &\quad \leq \frac{a w_n \beta_n}{\eta(1-\beta_n)} + k_Q \sum_{R_{n+1} \in (\mathcal{M}_n \cup \{R_n\}) \cap \mathcal{H}} \mu_n^L(R_{n+1}) \beta_{n+1} w_{n+1}. \end{aligned} \tag{61}$$

Since $w_{n+1} \leq w_n(1 + \frac{\alpha \beta_n}{1-\beta_n})$ from (18), and $\mu_n^L(R_{n+1})\beta_{n+1}$ is of order β_n by Bayesian updating, the sum in the RHS of (61) is of order $w_n \beta_n$ and thus negligible compared to the first term of the RHS. It can be dropped from the computation by slightly increasing the value of a.

In order to satisfy (61), the set of contracts R_{n+1} for which $Q(R_{n+1}) - Q(E_L(R_n)) \geq \sqrt{\beta_n}$ must be chosen by L with probability at most

$$\frac{a w_n \sqrt{\beta_n}}{\eta(1-\beta_n)},$$

which is less than $\frac{a\bar{W}}{1-\beta_0}\sqrt{\beta_n}$ since $w_n \leq \bar{W}\eta$ and $\beta_n \leq \beta_0$.

This, combined with Lemma 16, implies the existence of a constant $k_y > 0$ such that y_n is less than $k_y \beta_n^{1/4}$ with probability at least $1 - \frac{a\bar{W}}{1-\beta_0}\sqrt{\beta_n}$. Indeed, y_n^2 is bounded above by terms proportional to $Q(R_{n+1}) - Q(E_L(R_n))$ and a term proportional to $\max\{(\beta_n w_n/(1-\beta_n))^2, (\beta_{n+1} w_{n+1})^2\} + \beta_{n+1} w_{n+1}$. The first term is of order $\sqrt{\beta_n}$ with probability at least $1 - \frac{a\bar{W}}{1-\beta_0}\sqrt{\beta_n}$ as explained in the previous paragraph, while the second term is of order $\beta_n \eta$ and thus negligible compared to the first one. For the sequences considered, we have $\beta_n \leq \bar{\varepsilon}$ for all n in the block, so $\sqrt{\beta_n} \leq \sqrt{\bar{\varepsilon}}$. Combining these observations, we conclude that with probability at least $(1 - \frac{a\bar{W}}{1-\beta_0}\sqrt{\bar{\varepsilon}})^{\bar{N}}(1-2\bar{\varepsilon})^{\bar{N}}$, we have $\beta_n \leq \bar{\varepsilon}$, $w_n \leq k_w \eta$ and $y_n \leq k_y \bar{\varepsilon}^{1/4}$ for all rounds of the block. Choosing $k_\varepsilon = \max\{2, a\bar{W}/(1-\beta_0)\}$ yields the result. *Q.E.D.*

PROOF OF PROPOSITION 7: Consider any block of the kind studied by Lemma 12. Let $\mu_{\bar{n}}^{\mathcal{S}}$ (resp. $\mu_{\bar{n}}^{\mathcal{B}}$) denote the probability that H rejects all H-efficient contracts conditional on event $\mathcal{S}$ (resp. conditional on its complement, $\mathcal{B}$), and let $p_{\mathcal{S}}$ (resp. $p_{\mathcal{B}}$) be the probability that $\mathcal{S}$ ($\mathcal{B}$) occurs. We have $\mu_{\bar{n}} = p_{\mathcal{S}}\mu_{\bar{n}}^{\mathcal{S}} + p_{\mathcal{B}}\mu_{\bar{n}}^{\mathcal{B}}$. Since $p_{\mathcal{S}} \geq (1 - k_\varepsilon\sqrt{\bar{\varepsilon}})^{2\bar{N}}$, (13) implies that

$$\mu_{\bar{n}}^{\mathcal{S}} \leq \mu_n / p_{\mathcal{S}} \leq \frac{a\bar{W}}{D\bar{N}}(1 - k_\varepsilon\sqrt{\bar{\varepsilon}})^{-2\bar{N}}. \tag{62}$$

Choose $\bar{N}$ first so that $\frac{a\bar{W}}{D\bar{N}} < \frac{1}{4}$ and $\bar{\varepsilon}$ second, small enough to satisfy $(1 - k_\varepsilon\sqrt{\bar{\varepsilon}})^{-2\bar{N}} < 2$ and $k_y \bar{\varepsilon}^{1/4} < \frac{D}{4a}$ (the last condition is used in the next paragraph). This yields $\mu_{\bar{n}}^{\mathcal{S}} \leq \frac{1}{2}$. As with the blocks of Part I (Lemma 19), there exists a pushdown choice sequence in $\mathcal{S}$ such that the ex post probability that H has not chosen an H-efficient contract is weakly less than $\mu_{\bar{n}}^{\mathcal{S}}$. Along such a sequence, (i) $y_n \leq k_y \bar{\varepsilon}^{1/4}$, (ii) $\beta_n \leq \bar{\varepsilon}$, and (iii) $\beta_{\bar{n}+\bar{N}} \leq \frac{\beta_{\bar{n}}}{2}$. In particular β_n ends up *smaller* at the end of the block than at its beginning.

Let $\check{\beta} = \bar{\varepsilon}^{\bar{N}}$. Starting from round N, we build a sequence of $\bar{N}$-sized blocks as described above. Because w_n converges to zero (Proposition 2), it eventually drops below $\frac{D\eta}{2a}$. Let M denote the first round at which this happens. From (20), we have

$$w_{n+1} - w_n \geq -b\beta_{n+1}w_{n+1} - \eta y_n. \tag{63}$$

The sequence was constructed so that $y_n \leq k_y \bar{\varepsilon}^{1/4} \leq \frac{D}{4a}$. Combining this with (63) applied to round $M-1$ yields

$$w_M - w_{M-1} \geq -\frac{\eta D}{4a}.$$

Since $w_{M-1} \geq \frac{\eta D}{2a}$ by definition of M, this shows that $w_M \geq \frac{\eta D}{4a}$ and proves Proposition 7. *Q.E.D.*

REMARK 3: Some block may end with $w_n > \bar{W}\eta$. One then restarts the blocks of Part I from round n until reaching $\bar{W}\eta$ again. Since w_n converges to zero along any sequence, this back and forth between blocks of Parts I and III must end in finite time, leading to a round M for which $w_M \in (\frac{\eta D}{4a}, \frac{\eta D}{2a})$ and $\beta_M \geq \beta^*$. Since β_n decreases along blocks of both types, we have $\beta_M \leq \beta_N$; this yields the desired bound for β_N.

PROPOSITION 7 IMPLIES PROPOSITION 5: Let $\bar{\beta} = \min\{\frac{\tilde{\beta}}{k_\beta}, \check{\beta}\}$ and suppose that $w_N \in (\frac{\eta D}{2a}, \bar{W}\eta)$. If $\beta_N \leq \bar{\beta}$, we have $\beta_M \leq \tilde{\beta}$, from Part 2 of Proposition 7. From Part 1, moreover, w_M lies in $[\frac{1}{2}\frac{\eta D}{2a}, \frac{\eta D}{2a}]$. Proposition 6 applied to round M then implies that $\beta_M \geq \frac{1}{2\hat{c}}$ and, hence, $\beta_N \geq \frac{1}{2\hat{c}k_\beta}$. Therefore, either $\beta_N \geq \bar{\beta}$ or $\beta_N \geq \frac{1}{2\hat{c}k_\beta}$. Let $\beta' = \min\{\bar{\beta}, \frac{1}{2\hat{c}k_\beta}\}$. To prove Proposition 5, consider any round N for which $w_N \geq \eta$. If $w_N \geq \frac{\eta D}{2a}$, then we just showed that $\beta_N \geq \beta'$. And if $w_N \leq \frac{\eta D}{2a}$, Proposition 6 applies, showing that $\beta_N \geq \frac{2a}{D\hat{c}}$. Setting $\beta^* = \min\{\beta', \frac{1}{\hat{c}}\}$ yields the desired lower bound.

APPENDIX I: PROOF OF THEOREM 2, STATEMENT B

Fix an initial belief $\beta_0 \in (0, 1)$ and suppose without loss that $R_0 \in \mathcal{H}$. Let $\hat{Q}(u, p)$ denote the minimal expected cost of providing an expected utility u to H with a contract distribution that puts probability at least p on contracts lying in $\mathcal{H}$. We have $\hat{Q}(u_H(E_H(R_0)), 0) = Q(E_H(R_0))$, and $\hat{Q}(u_H(E_H(R_0)), p)$ is strictly increasing for p in a neighborhood of zero because contracts in $\mathcal{H}$ are inefficient for H. Statement A of Theorem 2 guarantees that H must get a utility arbitrarily close to $u_H(E_H(R_0))$ and that the cost to P conditional on facing H must be arbitrarily close to $Q(E_H(R_0))$ as η goes to zero.[85] For any $\varepsilon > 0$, this implies that there exists a threshold $\tilde{\eta}(\varepsilon)$ such that the probability $p_{\mathcal{H}}$ that H ends up with a contract in $\mathcal{H}$ satisfies $p_{\mathcal{H}} < \varepsilon$ for all PBEs corresponding to any $\eta < \tilde{\eta}(\varepsilon)$. For the remainder of the proof, we consider some small[86] $\varepsilon > 0$ and focus on η's below the threshold $\tilde{\eta} = \tilde{\eta}(\varepsilon^4)$, so that $p_{\mathcal{H}} \leq \varepsilon^4$.

[85]Indeed, each type θ gets an expected utility arbitrarily close to $u_\theta(E_\theta(R_0))$ while P's expected cost is bounded above by $\beta_0 Q(E_H(R_0)) + (1-\beta_0)Q(E_L(R_0))$, from Lemma 1. Since the contracts $E_\theta(R_0)$'s are efficient, the claim follows.

[86]It suffices to show the claim for all ε small enough, as it immediately implies that claim for higher values of ε.

From Statement A, there exists a threshold $\hat{\eta}$ below which θ's expected utility at the beginning of the game is bounded below by $v_\theta(\varepsilon) = u_\theta(E_\theta(R_0)) - \varepsilon^4$. The least costly contract $E_\theta(v_\theta(\varepsilon))$ that provides this utility costs $Q(E_\theta(R_0)) - O(\varepsilon^4)$ and lies within $O(\varepsilon^4)$ of $E_\theta(R_0)$. We also recall from Lemma 1 that P's expected cost is bounded above by $\beta_0 Q(E_H(R_0)) + (1-\beta_0)Q(E_L(R_0))$. Fix a PBE associated with some $\eta \le \min\{\hat{\eta}, \tilde{\eta}\}$ and let Q_θ denote P's expected cost conditional on facing θ and u_θ denote θ's expected utility. The previous observations imply that $|Q_\theta - Q(E_\theta(R_0))| = O(\varepsilon^4)$.

Let $\mathcal{T}_\varepsilon^H$ denote the set of contracts lying within a distance ε of $\mathcal{E}_H$ (for the Euclidean distance, say). For any contract R' outside of $\mathcal{T}_\varepsilon^H$, there exists $\hat{q}_H > 0$ such that $Q(R') \ge Q(E_H(R')) + \hat{q}_H \varepsilon^2$, by an argument similar to the one made in the proof of Lemma 15. Let p_H^{out} denote the probability that H ends up with a contract outside of $\mathcal{T}_\varepsilon^H$. Repeating the proof of Lemma 3 that led to (24) (page 29), we have, letting $\bar{Q}_H = Q(E_H(R_0))$,

$$\bar{Q}_H + O(\varepsilon^4) \ge Q_H \ge \bar{Q}_H + p_H^{\text{out}} \hat{q}_H \varepsilon^4,$$

and hence $p_H^{\text{out}} = O(\varepsilon^2)$. With probability $1 - O(\varepsilon^2)$, H thus gets contracts in the thin tube $\mathcal{T}_\varepsilon^H$. There remains to show that, within this tube, almost all contracts are clustered within ε of $E_H(R_0)$.

First, H never accepts a contract near $\mathcal{E}_H$ that gives a utility less than $u_H(E_H(R_0)) - \varepsilon^4$, as this would be his final utility (his type being revealed) but less than his expected utility in the game. Therefore, all contracts accepted in $\mathcal{T}_\varepsilon^H$ are located in the upper half of the tube starting $O(\varepsilon^4)$ below $E_H(R_0)$. Second, H's expected utility being less than $u_H(E_H(R_0))$, the probability of H getting a contract R' such that $u_H(R') \ge u_H(R_0) + \varepsilon^2$ is $O(\varepsilon^2)$. This shows that H gets within ε of $E_H(R_0)$ with probability $1 - O(\varepsilon^2)$.

The analysis for L is similar: the probability that L gets a contract outside of the set $\mathcal{T}_\varepsilon^L$ of contracts lying within a distance ε of $\mathcal{E}_L$ is of order $O(\varepsilon^2)$. Moreover, H never accepts contracts $R' \in \mathcal{H}$ such that $u_H(E_L(R')) \le u_H(E_H(R_0)) - \varepsilon^4$, by Proposition 1, (iv), and the remaining contracts all give L a utility of at least $u_L(R_0) - O(\varepsilon^4)$. Since L's expected utility is within ε^4 of $u_L(R_0)$, this implies that the probability of L getting contracts R' such that $u_L(R') \ge u_L(R_0) + \varepsilon^2$ is of order $O(\varepsilon^2)$.

For ε small enough, the terms $O(\varepsilon^2)$ are all less than ε. Therefore, each type θ gets a contract within ε of $E_\theta(R_0)$ with probability exceeding $1 - \varepsilon$. The threshold $\min\{\tilde{\eta}, \hat{\eta}\}$ thus delivers the conclusions of Statement B.

APPENDIX J: SMOOTHNESS AND MONOTONICITY OF EFFICIENCY CURVES

It suffices to show the result for $\mathcal{E}_H$, the analysis for $\mathcal{E}_L$ being identical. The set $\mathcal{E}_H$ of interior, H-efficient contracts is characterized by the tangency condition (letting $u = u_H$ for notational simplicity and letting f_i denote f's ith partial derivative for $f \in \{u, Q\}$)

$$\frac{u_2(x_1, x_2)}{u_1(x_1, x_2)} = \frac{Q_2(x_1, x_2)}{Q_1(x_1, x_2)},$$

or

$$F(x_1, x_2) = u_2(x_1, x_2)Q_1(x_1, x_2) - u_2(x_1, x_2)Q_1(x_1, x_2) = 0,$$

where we recall that the first-order derivatives of Q and u are strictly positive by assumption. We have $F_2(x_1, x_2) = u_{22}Q_1 + u_2Q_{12} - Q_{22}u_1 - Q_2u_{12} < 0$, since (i) all first-order derivatives are positive, (ii) u_{12} and $-Q_{12}$ are nonnegative by supermodularity and $-u_{11}$, $-u_{22}$, Q_{11}, and Q_{22} are positive by concavity of u and convexity of Q, and (iii) $u_{22}Q_{22} \neq 0$

by assumption. Therefore, we can apply the implicit function theorem: $\mathcal{E}_H$ is smooth, and its slope at any H-efficient contract (x_1, x_2) is given by

$$\frac{dx_2}{dx_1}\bigg|_{(x_1,x_2)} = \frac{-F_1}{F_2}\bigg|_{(x_1,x_2)} = -\frac{u_{12}Q_1 + u_2Q_{11} - Q_{12}u_1 - Q_2u_{11}}{u_{22}Q_1 + u_2Q_{12} - Q_{22}u_1 - Q_2u_{12}}\bigg|_{(x_1,x_2)}.$$

Under the assumed monotonicity, convexity, and supermodularity conditions, the numerator of the right-hand side is nonnegative, which shows that the slope is nonnegative.

The reason for requiring the efficiency curves to be upward sloping is to preserve the one-to-one mapping between $E_H(R)$ and $E_L(R)$. If, say, $\mathcal{E}_L$ was downward sloping over some range, then there could be multiple contracts of $\mathcal{E}_L$ giving the same utility to H, which would destroy Lipschitz bounds used in some of the proofs.

APPENDIX K: THE IMPOSSIBILITY OF FINITELY MANY ACTIVE ROUNDS

Consider any equilibrium and history with current contract R in $\mathcal{H}$ and non-degenerate belief β. We will show that the principal never jumps immediately to $E_H(R), E_L(R)$. This will imply that no L-efficient contract $\tilde{R}$ is ever accepted in equilibrium, since this would be equivalent to jumping.[87]

If P jumps immediately, he gets $\beta Q_H(E_H(R)) + (1-\beta)Q_L(E_L(R))$. Suppose that he proposes instead a contract $R' \in \mathcal{H}$ giving L the same utility as R and lying at a small distance x from $E_L(R)$, and the H-efficient contract C that gives H utility

$$u_H(C) = \eta u_H\big(R'\big) + (1-\eta)u_H\big(E_H(R)\big), \tag{64}$$

and jumps one period later. Also suppose that H accepts C with probability 1 and L accepts R' with probability 1—we will see shortly that P can guarantee this outcome by modifying the contracts by an infinitesimal amount which does not affect the strict benefit of this deviation over the immediate jump. Since $u_L(R') = u_L(R)$, we have $E_L(R') = E_L(R)$ and $E_H(R') = E_H(R)$, so the only payoff difference between the two strategies (jumping now or one period later) concerns what happens if a breakdown occurs at the end of the current period.

By efficiency of $E_L(R)$, P's iso-cost curve and L's iso-utility curve are tangent at $E_L(R)$, so the cost of R' is of second order, x^2, above $E_L(R)$'s cost.

From (64), we have $u_H(E_H(R)) - u_H(C) = \eta(u_H(E_H(R)) - u_H(R'))$. The difference on the right-hand side is of order x, because H's iso-utility curve is not tangent to L's iso-utility curve at $E_L(R)$, from the strict single-crossing property. Therefore, P's gain from the deviation, relative to an immediate jump, is bounded below by $\eta x \underline{a}\beta$, where $\underline{a}$ is a Lipschitz constant translating bounds on utility decrements for H into bounds on cost savings (cf. Lemma 13), while the cost is of order $(1-\beta)\eta x^2$. The deviation is thus strictly beneficial, no matter how small the current belief β is, provided that x is below some fixed threshold.[88]

[87]If an agent of type θ accepts $\tilde{R} \in \mathcal{E}_L$, his continuation utility jumps to $u_\theta(\tilde{R})$ by Proposition 1. Therefore, L accepts such a contract only if $u_L(\tilde{R}) \geq u_L(R) = u_L(E_L(R))$. But this implies that $u_H(\tilde{R}) \geq u_H(E_L(R)) = u_H(E_H(R))$. From Lemma 1, this can only occur if $\tilde{R} = E_L(R)$, as proposing $\tilde{R}$ would otherwise be too costly for P.

[88]We also choose x small enough to guarantee that R' is closer to $\mathcal{E}_L$ than R, that is, $Q(R') < Q(R)$ or, equivalently, $u_H(R') > u_H(R)$.

There remains to check that P can implement the desired deviation by forcing H to take C with probability 1, so that P can reap the full benefit from the deviation, and forcing L to take R' with probability 1, so that the loss on L is indeed of order x^2. Suppose that H rejects C. In this case, his utility if there is a breakdown at the end of the current period is bounded above by $u_H(R')$. Moreover, from Part (iv) of Proposition 1, his continuation utility if no breakdown occurs is bounded above by $u_H(E_H(R))$, regardless of whether he takes R' or stays at R, since $E_H(R) = E_H(R')$. From (64), it is therefore strictly beneficial for H to accept C provided that P moves C higher by an arbitrarily small amount ϵ, negligible compared to η. Similarly, P can guarantee that L takes R' instead of holding on to R by moving R' by ϵ^2 above L's current iso-utility curve (L never wants to take C for reasons explained earlier). Provided that ϵ is small enough (i.e., of order less than $\eta^2\beta^2$), the cost increase is negligible compared to the gain from the deviation, which is of order $\eta\beta^2$.

APPENDIX L: EXTENSION TO MORE AGENT TYPES

While the analysis has focused on two types, it offers a path to analyze more general type structures. To illustrate how the ideas presented here may be used to carry such an extension, suppose that the agent may have a third, intermediate type M, with utility functions still ordered by single-crossing. The efficiency curves $\mathcal{E}_L$, $\mathcal{E}_M$, and $\mathcal{E}_H$ are distinct. The setting is otherwise identical to the binary case.

The extension of Theorem 2 is straightforward to conjecture: The contract space is now divided into four regions, separated by the efficiency curves. If the initial contract R lies below $\mathcal{E}_L$, $E_L(R)$ is defined as before, $E_M(R)$ is the M-efficient contract giving M the same utility as $E_L(R)$ does, and $E_H(R)$ is the H-efficient contract giving H the same utility as $E_M(R)$ does. If R lies between $\mathcal{E}_L$ and $\mathcal{E}_M$, then define $E_L(R)$ and $E_M(R)$ as the L-efficient and M-efficient contracts giving L and M the same utility as R does and $E_H(R)$ be the H-efficient contract giving H the same utility as $E_M(R)$ does. These contracts are defined analogously when R lies in the other two regions. The conjecture is that if P initially assigns strictly positive probability to each type, each type θ gets with arbitrarily high probability an outcome arbitrarily close to $E_\theta(R)$ as η goes to zero.

To show this, the suggested strategy is as follows. First, notice that after any history at which P assigns probability 0 to any of the three types, we are back to the binary case, and Theorem 2 predicts a unique (up to η) outcome. So suppose instead that the belief distribution puts positive weight on all types and, for example, that R lies in the first configuration, that is, below $\mathcal{E}_L$. In this case, one can extend Proposition 1 and its corollaries to show that L will never accept a contract in another configuration, derive upper bounds on H's and M's continuation utilities, and an upper bound on P's expected cost.

To analyze equilibrium behavior, one must keep track of the probabilities β_H and β_M of facing H and M, respectively. However, it turns out that many steps of earlier proofs can still be performed using the probability $\beta = \beta_H + \beta_M$ of facing a type other than L. In particular, the loss bound D conditional on facing a type other than L in case of a breakdown is still valid. Lipschitz bounds are still valid, too. As noted, given Theorem 2, it is (almost) without loss to assume that P proposes only contracts in the first configuration and a pair of H and M efficient contracts that make H indifferent between these contracts. Indeed, any contract above $\mathcal{E}_L$ reveals that the agent type is not L, bringing the analysis back to the binary type case with H and M. One difference is that Theorem 5 implies only that the agent's utility is within $O(\eta)$ of the outcome, a difference which should be minor for a full-blown analysis.

With three or more types, the principal is more likely to benefit from proposing at least three contracts in each period, and the belief assigned to any given type is likely to be non-monotonic. Fortunately, the analysis conducted in this paper already allows for an arbitrary number of contracts proposed in each period and already handles arbitrary belief non-monotonicity. This should facilitate any formal extension of the results to more than two types.

REFERENCES

HARRIS, C. (1985): "Existence and Characterization of Perfect Equilibrium in Games of Perfect Information," *Econometrica*, 53 (3), 613–628. [2]

Northwestern University, Evanston, 60208, U.S.A.; b-strulovici@northwestern.edu.

Co-editor Joel Sobel handled this manuscript.

Manuscript received 14 July, 2015; final version accepted 12 December, 2016; available online 12 December, 2016.