Online Appendix to "Endogenous Beliefs in Models of Politics"*

William Minozzi†

November 26, 2013

The Endogenous Beliefs Models. The Endogenous Beliefs Models (EBM) applies to any finite extensive form game with perfect information and chance moves. The process (in brief) is (1) for any game G, identify the associated family of games Γ with heterogeneous beliefs, (2) embed Γ as the set of subgames in a new endogenous beliefs game $ebg(G)$, (3) find the subgame perfect equilibria of $ebg(G)$. What follows is a formalization of this process.

Let G be a *finite extensive form game with perfect information and chance moves*, referred to here as a *game*. Fix a set of players $N = \{1, 2, ..., n\}$ and a set of histories H, such that the root history $\varnothing \in H$ and, for any $a = (a_1, ..., a_{t-1}, a_t) \in H$, $(a_1, ..., a_{t-1}) \in H$. The set of terminal histories is $Z \subset H$, and the *player assignment function* identifies a subset of players (or Nature) at each nonterminal history, $P : H \setminus Z \to 2^N \cup \{\text{Nature}\}$. The set of i's moves is $H_i = h \in H : i \in P(h)$. For any nonterminal history $h \in H \setminus Z$, the *action set* for i at h is $A_i(h) = \{a : (h, a) \in H\}$. Each player $i \in N$ has preferences over Z represented by a utility function $u_i : Z \to \mathbb{R}$. At Nature's moves, each player $i \in N$ has the beliefs given by conditional probability distribution $F(a|h) \in \mathcal{F}$, where $\mathcal{F}$ is the family of distributions with support $A(h)$. These beliefs are common to all players, as is standard. Then $G = (N, H, P, (u_i)_{i \in N}, F)$.

The EBM relies on the *subgame perfect equilibrium* solution concept, referred to here as *equilibrium*. A *strategy* for $i \in N$ is a function $s_i : H \to H$ such that $s_i(h) \in A(h), \forall h \in H^i$.[1]

*I thank Andrew Coe for alerting me to mistakes in the proofs section. These mistakes did not alter the findings presented in the article.

†Assistant Professor, Department of Political Science, 2189 Derby Hall, Ohio State University, Columbus, OH 43210 Phone: 614-247-7017, Email: minozzi.1@osu.edu

[1] For simplicity of presentation, I focus on pure strategies, but allowing mixed strategies does not alter the

A *strategy profile* is $s = (s_i)_{i \in N}$, the set of strategies for i is S^i, and the set of all strategy profiles is S. For any $s = (s_i)_{i \in N}$, the *outcome distribution* for G when each $i \in N$ uses s_i is a probability distribution $\zeta(z; s)$ with support Z. For any $h \in H \setminus Z$, the *subgame* that follows h is $G(h) = (N, H|_h, P|_h, (u_i|_h)_{i \in N}, F)$, where $H|_h = \{\tilde{h} \in H : \tilde{h} = (h, h')\}$, $P|_h(h') = P(h, h')$, and $u_i|_h(h') = u_i(h, h')$. Extending definitions to $G(h)$, a subgame strategy is $s_i|_h(h') = s_i(h, h')$, the subgame outcome function is $\zeta|_h(z; s|_h) = \zeta(z; h, s)$, and the set of subgame strategies $S_i|_h$. An *equilibrium* of G is a strategy profile s^* such that for all $h \in H \setminus Z$ and $i \in N$, if $P(h) = i$, then

$$\sum_{z \in Z} u_i(s^*|_h)\zeta|_h(z; s^*|_h) \geq \sum_{z \in Z} u_i(s_i|_h, s^*_{-i}|_h)\zeta|_h(z; s_i|_h, s^*_{-i}), \ \forall s_i|_h \in S_i|_h. \tag{1}$$

The one-step deviation principle and Kuhn's theorem provide for the existence of equilibrium for G and its identification via backwards induction.

Given a game G as above, the family of *heterogeneous beliefs games* associated with G is $\Gamma(G) = \{G' : G' = (N, H, P, (u_i)_{i \in N}, (F_i)_{i \in N})\}$, where each F_i is a (potentially) different probability distribution. The definition and existence of equilibrium extends to this family of games with only minor amendments; neither the one-step deviation principle nor Kuhn's Theorem makes use of the commonality of F. The only changes come in the statements of outcome distributions $\zeta(z; s)$, which now depend on players' (potentially) different beliefs about Nature's moves. Define i's outcome distribution as $\zeta_i(z; s)$. An equilibrium of G' is a strategy profile s^* such that for all $h \in H \setminus Z$ and for all $i \in N$, if $P(h) = i$, then

$$\sum_{z \in Z} u_i(s^*|_h)\zeta_i|_h(z; s^*|_h) \geq \sum_{z \in Z} u_i(s_i|_h, s^*_{-i}|_h)\zeta_i|_h(z; s_i|_h, s^*_{-i}), \ \forall s_i|_h \in S_i|_h. \tag{2}$$

As stated above, both the one-step deviation principle and Kuhn's theorem apply without amendment to each $G' \in \Gamma$. Thus, for each G', there exists an equilibrium.

To complete the interpretation of the EBM advanced in the paper, I utilize two more

arguments substantively.

game-theoretic tools. Each begins with a given $G = (N, H, P, (u_i)_{i \in N}, F)$ that maps to *ebg*(G), the *endogenous beliefs game* associated with G. First, I use Selten's (1975) interpretation of an extensive form game as an agent-normal form. According to this interpretation, the set of players in *ebg*(G) is $\hat{N}$, where $\hat{N} = \{\hat{1}, ..., \hat{n}\}$, the player assignment function is $\hat{P}(\varnothing) = N$, and each $\hat{i}$ has action set $\hat{A}_i(\varnothing) = F_i \in \mathcal{F}_i$, where $\mathcal{F}_i \subset \mathcal{F}$.[2] Nature begins the game by choosing a distribution $F \in \mathcal{F}$. Players, histories, and player assignment are defined without significant amendment of G. Each player is associated with a belief-forming agent who has access to the information in F at $t = 0$. Thus, the terminal histories of $\hat{H}$ are the pairs $(\varnothing, G')$, where $G' \in \Gamma$. For each $G' \in \Gamma$, let $s^*(F_i, F_{-i}; G')$ be an equilibrium of G'. With a slight abuse of notation, let $s^*(F_i, F_{-i})$ select a single equilibrium of G' via some selection criteria. Utility functions in *ebg*(G) are

$$v_i(F_i, F_{-i}) = \sum_{z \in Z} u_i(s^*(F_i, F_{-i}))\zeta(z; s^*(F_i, F_{-i})) + \delta \sum_{z \in Z} u_i(s^*(F_i, F_{-i}))\zeta_i(z; s^*(F_i, F_{-i})). \quad (3)$$

Thus, *ebg*$(G) = (\hat{N}, \hat{H}, \hat{P}, (v_{\hat{i}})_{\hat{i} \in \hat{N}}, F)$, and any equilibrium of *ebg*(G) constitutes a set of endogenous beliefs.

The second game-theoretic tool I use in the EBM is the assumption that players have *imperfect recall* of belief formation (Piccione and Rubinstein 1997).[3] In every post-belief formation subgame in *ebg*(G), information sets are singletons. However, players do not have access to F for $t > 0$. Preferences are given as in (3), and beliefs depend on information sets. If at $t = 0$ players form beliefs $(F_i)_{i \in N}$, the $t > 0$ information set for i is $\{(F, G'_\varnothing) : F \in \mathcal{F}\}$, where $G'_\varnothing$ is the root history of the G' identified by $(F_i)_{i \in N}$. The outcome distribution corresponding to this information set depends only on F_i because the remainder of the elements in the information set are uninformative. An equilibrium of this game identifies

[2] In all the applications in the paper, $\mathcal{F}_i = \mathcal{F}$, but this need not be the case. The EBM can easily accommodate limits on allowable beliefs.

[3] Games with imperfect recall constitute something of a "road not taken" since Kuhn (1953) formalized extensive form games with *perfect recall*.

endogenous beliefs.[4]

Proofs

Voter Turnout with Endogenous Beliefs. Using backwards induction, at $t = 1$, the citizen only has access subjective beliefs, and not to objective probabilities. Thus, she regards anticipatory and outcome utility as being identical. If the citizen chooses to vote, she has subjective expected utility $\hat{p}(1) - c$, and if she chooses not to vote, she expects utility 0. Thus the citizen votes if and only if $\hat{p} > c$. At $t = 0$, if the citizen forms beliefs $\hat{p} > c$, she expects to vote and thus total utility $\delta(\hat{p} - c) + p - c$, and if she forms beliefs $\hat{p} \leq c$, she expects not to vote and total utility 0. Because total utility is (weakly) increasing in $\hat{p}$, all beliefs $\hat{p} \in (c, 1)$ are dominated by $\hat{p} = 1$. Thus, the citizen forms beliefs $\hat{p}^* = 1$ if and only if $\delta(1 - c) + p - c > 0$, and forms beliefs $\hat{p} \leq c$ otherwise. Thus, she votes if and only if $p > c - \delta(1 - c)$. $\square$

Conditions for Endogenous Prospects of Upward Mobility. Consider a candidate for equilibrium in which (1) all the likely rich citizens expect income $\hat{y}^* = y_r$, (2) $\frac{n+1}{2}$ of the likely poor citizens expect income $\hat{y}^* = \bar{y}$, and (3) the remaining likely poor citizens expect $\hat{y}^* = y_r$. By the median voter theorem, the tax rate will be $\tau^* = 1$ since $\frac{n+1}{2}$ are a majority of citizens, and tax preferences are single-peaked with respect to expected income. Each citizen i with expected income $\hat{y}^*$ expects total utility $(\delta + 1)\bar{y}$. If i increases $\hat{y}_i$ above $\bar{y}$ the tax rate goes to $\tau^* = 0$, and i expects total utility $\delta y_r + y_p$. Thus, $\tau^* = 1$ if and only if $(\delta + 1)\bar{y} \geq \delta y_r + y_p$, which is true if and only if $\frac{\bar{y} - y_p}{y_r - \bar{y}} \geq \delta$. Substituting in $\bar{y} = \frac{1}{n}(my_p + (n - m)y_r)$ and rearranging yields the condition $\frac{(n-m)(y_r - y_p)}{m(y_r - y_p)} \geq \delta$. Since $y_r > y_p$, this conditions holds if and only if $\delta \leq \frac{(n-m)}{m}$. $\square$

Subgame Perfect Equilibrium of the Crisis Bargaining Game. First, consider the case that $\hat{p}_A \leq c$. In the last period, if B rejects any demand that A has made, A must

[4] Because this solution concept does not involve *forward induction*, players are not assumed to reason about the identity of F given their choices of F_i.

choose whether to initiate conflict, $a \in \{0, 1\}$. The subjective expected payoff for initiating conflict ($a = 1$) is $\hat{p}_A - c$ and that for not initiating conflict ($a = 0$) is 0. Since $\hat{p}_A \leq c$, A prefers $a = 0$. Working backwards, B observes that A will not initiate conflict. Paying any amount x yields payoff $1 - x$, while rejecting demands yields payoff 1. Thus B will not pay any amount demanded, and $\bar{x}^* = 0$.

Next, consider the case that $\hat{p}_A > c$ and $\hat{p}_A + \hat{p}_B \leq 1 + 2c$. Unlike in the last case, A's threat is now credible, and $a^* = 1$. Working backwards, B believes that rejecting A's demand x will lead to conflict and payoff $\hat{p}_B - c$, and paying the demand x yields payoff $1 - x$. Therefore, B will pay any amount $x \leq 1 - \hat{p}_B + c = \bar{x}^*$. When A chooses the amount to demand x, A believes that choosing $x > \bar{x}^*$ will lead to conflict and payoff $\hat{p}_A - c$, and that B will accept any demand $x \leq \bar{x}^*$, yielding payoff x. Thus, A will demand $\bar{x}^*$ if $\hat{p}_A - c \leq \bar{x}^* = 1 - \hat{p}_B + c$. By assumption, $\hat{p}_A + \hat{p}_B \leq 1 + 2c$. Hence, the equilibrium demand will be $x^* = \bar{x}^*$, and the outcome will be settlement.

Finally, consider the case that $\hat{p}_A + \hat{p}_B > 1 + 2c$. Because $\hat{p}_B \leq 1$, it must be that $\hat{p}_A > c$, and therefore $a^* = 1$ by the argument from the previous case. Furthermore, B's maximum acceptable demand is the same as in the previous case. In fact, the only difference between these two cases is that A now prefers the subjective expected payoff of conflict to that from settlement because of the case assumption. Hence, A will demand any unacceptable amount $x^* > \bar{x}^*$, and the outcome will be conflict. $\square$

Conditions for Endogenous Mutual Optimism. First, observe that A will not form beliefs $\hat{p}_A \leq c$ because such beliefs are weakly dominated by $\hat{p}_A = 2c$. If $\hat{p}_A = 2c$, then the equilibrium outcome of the crisis bargaining game is always settlement, since $\hat{p}_B \leq 1$. Thus, total utility for any $\hat{p}_A \leq c$ is 0, while total utility for $\hat{p}_A = 2c$ is $x^* = 1 - \hat{p}_B + c > 0$.

Next consider the conditions on beliefs for settlement. Suppose first that $\hat{p}_B > 2c$, so that conflict is possible for some beliefs A might form. In this case, A might form beliefs that lead to settlement or to conflict. If A forms beliefs $\hat{p}_A \in [2c, 1 - \hat{p}_B + 2c]$, settlement ensues and A expects total utility $(\delta_A + 1)x^* = (\delta_A + 1)(1 - \hat{p}_B + c)$. If A forms beliefs $\hat{p}_A > 1 - \hat{p}_B + 2c$, conflict ensues and A expects total utility $\delta_A(\hat{p}_A - c) + (p_A - c)$. In the latter case, total utility increases in $\hat{p}_A$, so $\hat{p}_A = 1$ is the best of these beliefs for A. Thus, A will form beliefs that lead to settlement if and only if $\delta_A(1-c) + (p_A - c) \leq (\delta_A + 1)(1 - \hat{p}_B + c)$, or

$$\hat{p}_B \leq \frac{p_B}{\delta_A + 1} + 2c, \tag{4}$$

where p_B has been substituted for $1 - p_A$. Thus, A will form beliefs that lead to settlement only if $\hat{p}_B$ is not too optimistic.

A similar line of reasoning shows that B will form beliefs that lead to settlement if and only if A is not too optimistic. First, suppose $\hat{p}_A > 2c$, so that conflict is possible for some beliefs that B might form. If B forms beliefs $\hat{p}_B \leq 1 - \hat{p}_A + 2c$, settlement ensues and B expects total utility $(\delta_B + 1)(1 - x^*) = (\delta_B + 1)(\hat{p}_B - c)$. Total utility in this case is increasing in $\hat{p}_B$ and B will therefore form beliefs

$$\hat{p}_B = 1 - \hat{p}_A + 2c. \tag{5}$$

If instead B forms beliefs $\hat{p}_B > 1 - \hat{p}_A + 2c$, conflict ensues and B expects total utility $\delta(\hat{p}_B - c) + p_B - c$. As with A, total utility increases in $\hat{p}_B$, so $\hat{p}_B = 1$ is the best of these beliefs for B. Thus, B forms beliefs that lead to settlement if and only if $\delta_B(1-c)+(p_B-c) \leq (\delta_B + 1)(1 - \hat{p}_A + c)$, or

$$\hat{p}_A \leq \frac{p_A}{\delta_B + 1} + 2c. \tag{6}$$

So B only forms beliefs that lead to conflict when $\hat{p}_A$ is not too optimistic.

Both (4) and (5) are true when $1 - \hat{p}_A + 2c \leq \frac{p_B}{\delta_A+1} + 2c$, which is the case if and only if

$$1 - \frac{p_B}{\delta_A + 1} \leq \hat{p}_A. \tag{7}$$

Combining (6) and (7) yields

$$1 - \frac{p_B}{\delta_A + 1} \leq \hat{p}_A \leq \frac{p_A}{\delta_B + 1} + 2c, \tag{8}$$

which can only be satisfied if

$$1 - 2c \leq \frac{p_A}{\delta_B + 1} + \frac{p_B}{\delta_A + 1}. \tag{9}$$

Condition (9) defines the endogenous beliefs of the crisis bargaining game. If (9) is satisfied, both A and B prefer settlement to conflict, A forms beliefs given by (7), and B forms beliefs given by (5). Otherwise, both A and B form beliefs $\hat{p}_A^* = \hat{p}_B^* = 1$ and there is conflict in equilibrium. $\square$

Additional Examples

Dominance Solvability. In the text, I state that no strategy that is strictly dominated in the standard model can be included in an equilibrium in the EBM.[5] But does this identity between the equilibria of the standard model and the EBM extend to all dominance solvable games? In this section, I provide a counterexample to this conjecture.

For a player i, an action a_i is weakly dominated by another action a_i' if a_i' yields at least as much utility as a_i regardless of the profile of other players' actions (and chance outcomes), and in at least one such profile a_i' yields more strictly utility than a_i. Consider the iterative procedure in which all the weakly dominated actions of each player are eliminated at each stage of the procedure. A game is *dominance solvable* if all players are indifferent between the outcomes that survive this procedure. Although dominance solvability is not often applied in games with chance moves, the definition is easily extended to include moves available to Nature (although she is indifferent between all outcomes, of course).

Now consider the strategic form game described by Figure 1. In that game there are two players, Row and Column, each of which chooses one of two actions. Nature is the matrix player, and she chooses between matrix A and matrix B, each with equal probability. Consider Row's reasoning. There are four action profiles to consider (A, L), (A, R), (B, L), and (B, R). In the first three cases, the action U offers strictly more utility than the action D. In the last case, the actions offer Row an equal amount of utility. Thus, U weakly dominates D. By symmetry, L also weakly dominates R. Thus, the game is dominance solvable and has a unique equilibrium at (U, L).

But yet there is another equilibrium in the EBM. If Row believes that Nature chooses matrix B with probability 1 and Column believes that Nature chooses matrix A with probability 1, then (D, R) is also an equilibrium. Consider whether there are any profitable deviations, first in the subgame after beliefs have been formed. Row believes he is playing matrix 2. If he conjectures that Column will play R, then T and B offer the same utility, so

[5] I gratefully acknowledge an anonymous reviewer for pointing out this result.

there is no profitable deviation. By symmetry, Column has no profitable deviations, and so (B, R) is an equilibrium in the subgame. Now consider the belief-formation stage. The total utility attendant to these beliefs is $2\delta_i + 1$. If either player instead forms *any* other beliefs, then the subgame that follows has only the dominance solvable equilibrium at (T, L). As such, the total utility for those beliefs is $\delta_i + 1$, which is strictly less than $2\delta_i + 1$ for any $\delta_i > 0$. Thus, Figure 1 provides a counterexample to the conjecture.

Nonnestedness of the Set of Standard Model Equilibria. In each of the examples included in the text, the equilibria of the example in the standard model are also equilibria in the EBM, at least for some values of taste for anticipation. Therefore, it is reasonable to conjecture that any equilibrium in the standard model will be an equilibrium in the EBM for some tastes for anticipation. However, this conjecture is false.

Consider the game depicted in Figure 2. When $p < \frac{1}{2}$, the unique equilibrium of this game is (L_1, R_2). To see this, observe that $p < \frac{1}{2}$ implies by backwards induction that player 2 will choose R_2. But R_2 ensures player 2 of getting utility 0, and therefore 1 will choose L_1. If the conjecture is correct, then this equilibrium in the standard model should be an equilibrium in the EBM, but it is not. It suffices to consider 2's subjective beliefs $\hat{p}$ about what Nature chooses after R_2. The equilibrium in the game given subjective beliefs is straightforward: 2 chooses L_2 if and only if $\hat{p} > \frac{1}{2}$ (excluding the non-generic case $\hat{p} = \frac{1}{2}$). Otherwise, 2 chooses R_2. But if player 2 chooses L_2, then player 1 receives utility 2 if she chooses R_1 as opposed to utility 1 if she chooses L_1. Thus, 1 will choose R_1 if and only if $\hat{p} > \frac{1}{2}$.

Finally, consider the belief-forming stage. Again, it is sufficient to focus on player 2's beliefs. If player 2 forms beliefs $\hat{p} < \frac{1}{2}$, the equilibrium will be (L_1, R_2), and her total utility will be 0. If instead she forms beliefs $\hat{p} > \frac{1}{2}$, the equilibrium will be (L_2, R_1), and her total utility will be 1. Therefore, the total utility for $\hat{p} > \frac{1}{2}$ exceeds that for $\hat{p} < \frac{1}{2}$, and the unique equilibrium of the EBM is (L_2, R_1). Thus, Figure 2 constitutes a counterexample to the conjecture.

	L	R
U	$1,1$	$2,0$
D	$0,2$	$0,2$

A

	L	R
U	$1,1$	$2,0$
D	$0,2$	$2,0$

B

Figure 1: Nature is the matrix player, and she randomizes between A and B with equal probabilities. The game is dominance solvable in the standard model, and the unique equilibrium is (U, L). In the EBM, there is a second equilibrium at (D, R) in which Row believes that Nature chooses B with probability 1 and Column believes Nature chooses A with probability 1.

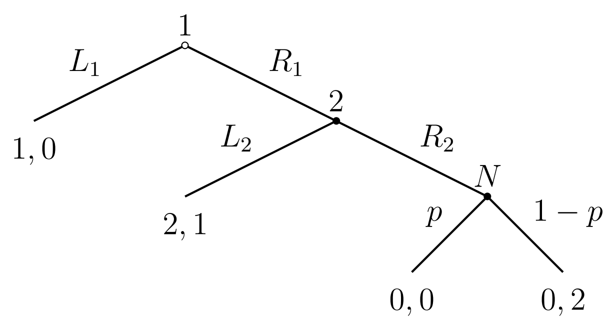

Figure 2: If $p < \frac{1}{2}$, then the unique equilibrium in the standard model is (L_1, R_2). However, regardless of the tastes for anticipation, the unique equilibrium of the game in the EBM is (R_1, L_2) with beliefs $\hat{p} > \frac{1}{2}$.

References

Kuhn, H. W. 1953. "Extensive Games and the Problem of Information." *Contributions to the Theory of Gamess II* pp. 307–317.

Selten, R. 1975. "Reexamination of the perfectness concept for equilibrium points in extensive games." *International journal of game theory* 4(1): 25–55.

Piccione, Michele, and Ariel Rubinstein. 1997. "On the Interpretation of Decision Problems with Imperfect Recall* 1." *Games and Economic Behavior* 20(1): 3–24.