A Dutch Book Argument for Belief Consistency -
Online Appendix

Emiliano Catonini∗

February 22, 2022

1 The Dutch book argument for forward consistency

This section formalizes the argument that Bayesian updating, even with the wrong probabilities of contingencies given the states, is necessary and sufficient to reject (some bet in) a diachronic Dutch-book, that is, a system of bets such that the bookmaker makes a positive profit “ex-post”, under each possible realized learning path. The term diachronic is used in the literature for constrast with the classical notion of Dutch book, where all bets are proposed at the same time. Here I use the term “deterministic Dutch book” for contrast with the notion of Dutch book adopted in the main text, where the bookmaker makes a positive profit only “ex-ante”, in objective expected terms.

Given a learning path $l$, let $H(l)$ denote the set of histories that constitute it.

Definition 1 A system of gambles $g = (g(\cdot|h))_{h \in H}$ is a deterministic Dutch book if for every $s \in S$ and $l \in L(s)$,

$$\sum_{h \in H(l)} g(s|h) \leq 0,$$

and for some $s \in S$ and $l \in L(s)$,

$$\sum_{h \in H(l)} g(s|h) < 0.$$

I say that an agent is deterministically Dutch book if she accepts a deterministic Dutch book given her belief system.

As one would expect, a deterministic Dutch book is also a Dutch book.

Remark 2 Every deterministic Dutch book is a Dutch book.

∗New York University in Shanghai, department of economics, emiliano.catonini@nyu.edu
Proof. For each $s \in S$, we have

$$
\sum_{l \in L(s)} \eta^s(l) \sum_{h \in H(l)} g(s|h) = \sum_{h \in H} \left( \sum_{l \in L(s) \cap L(h)} \eta^s(l) \right) g(s|h) = \sum_{h \in H} p(h|s) g(s|h).
$$

Given this, conditions (1) and (2) imply the corresponding conditions in the definition of Dutch book. 

Note that a deterministic Dutch book makes no reference to the objective probabilities of contingencies. This suggests that such probabilities do not play any role in establishing whether the agent can be deterministically Dutch-booked or not. This is indeed the case: if the set of deterministic Dutch-books does not depend on the probabilistic model of how the states trigger different learning paths, this is true also for the set of belief systems that make the agent accept a deterministic Dutch book. Therefore, from now on, we can regard belief systems of the agent that hat expose the agent to Dutch booking should not depend on the probabilistic model she uses.

I say that a belief system $\mu = (\mu(\cdot|h))_{h \in H}$ is forward consistent if for every $h \in H$ and $h' > h$, there exists a profile of distributions over learning paths $\mu(\cdot|h')$ is derived from $\mu(\cdot|h)$ by Bayes rule whenever possible, that is, when $\mu(S(h')|h) > 0$.\footnote{The terminology of forward consistency has been introduced by Battigalli et al. (2021) for the application of the chain rule in the context of a sequential game — see the next section.} Now I state and prove the Dutch book theorem for forward consistency.

Theorem 3 Consider an agent with belief system $\mu = (\mu(\cdot|h))_{h \in H}$. The following are equivalent:

- $\mu$ is forward consistent;
- the agent cannot be deterministically Dutch booked.

Proof. First I show that if $\mu$ is forward consistent, the agent cannot be deterministically Dutch-booked. Fix a system of gambles $g^* = (g^*(\cdot|h))_{h \in H}$. Suppose that $i$ is willing to accept $g^*$. I show that then $g^*$ is not a deterministic Dutch book. Let $g = (g(\cdot|h))_{h \in H}$ be the system of gambles such that, for each $h \in H$ and $s \in S$, if $\sum_{s'} \mu(s'|h) g(s'|h) = 0$ and $\mu(s|h) = 0$, then $g(s|h) = 0 \leq g^*(s|h)$ (the inequality comes from Definition ??), otherwise $g(s|h) = g^*(s|h)$. It is easy to see that $i$ is willing to accept also $g$.

Let $(\eta^*)_{s \in S}$ denote the probability distributions that make $\mu$ forward consistent. Let $H^*$ denote the set of contingencies $h$ such that $\mu(S(h)|h') = 0$ for every $h' < h$. Fix an ordering $(h^1, \ldots, h^{|H^*|})$ of $H^*$ such that if $j < k$, then $h^j \neq h^k$. For each $k = 1, \ldots, |H^*|:

- $H^k$ is the set of contingencies $h$ such that $\mu(S(h)|h^k) > 0$;
- $S^k$ denote the set of states $s$ such that $\mu(s|h^k) > 0$;
- for each \( s \in S^k \), \( \eta^{k,s} \) is the distribution over learning paths conditional on reaching \( h^k \);
- for each \( h \in H^k \), \( p^k(h|s) \) is the probability of reaching \( h \) from \( h^k \) given \( s \) and \( \eta^{k,s} \).

Fix \( k = 1, \ldots, |H^*| \). For each \( h \in H^k \) and \( s \in S^k \), by forward consistency

\[
\mu(s|h) = \frac{p^k(h|s)\mu(s|h^k)}{\sum_{s'} p^k(h|s')\mu(s'|h^k)}.
\]

Thus, since \( i \) is willing to accept \( g(\cdot|h) \), we have

\[
\sum_{s \in S^k} p^k(h|s)\mu(s|h^k) \sum_{l \in L(s) \cap L(h)} \eta^{k,s}(l) \mu(s|h)g(s|h) \geq 0,
\]
which implies

\[
\sum_{s \in S^k} p^k(h|s)\mu(s|h^k)g(s|h) \geq 0.
\] (3)

Since this is true for every \( h \in H^k \), using the definition of \( p(h|s) \), we can write

\[
\sum_{h \in H^k} \sum_{s \in S^k} \left( \sum_{l \in L(s) \cap L(h)} \eta^{k,s}(l) \right) \mu(s|h^k)g(s|h) \geq 0,
\]
and rearrange as

\[
\sum_{s \in S^k} \mu(s|h^k) \sum_{h \in H^k} \sum_{l \in L(s) \cap L(h)} \eta^{k,s}(l)g(s|h) \geq 0,
\]
which can be rewritten as

\[
\sum_{s \in S^k} \sum_{l \in L(s)} \mu(s|h^k)\eta^{k,s}(l) \sum_{h \in H^k \cap H(l)} g(s|h) \geq 0.
\] (4)

Suppose that

\[
\forall s \in S, \forall l \in L(s), \quad \sum_{h \in H(l)} g(s|h) \leq \sum_{h \in H(l)} g^*(s|h) \leq 0,
\] (5)

otherwise \( g^* \) does not satisfy (1). I am going to show with a recursive procedure that, for every \( k = 1, \ldots, |H^*| \),

\[
\sum_{s \in S^k} \sum_{l \in L(s)} \mu(s|h^k)\eta^{k,s}(l) \sum_{h \in H(l)} g(s|h) = 0,
\] (6)

so that, by (5), we get

\[
0 = \sum_{h \in H(l)} g(s|h) \leq \sum_{h \in H(l)} g^*(s|h)
\]
for every \( s \in S^k \) and \( l \in L(s) \), thus \( g^* \) does not satisfy condition (2).

So, fix \( k = 1, \ldots, |H^*| \). Suppose by way of induction that for every \( j \geq k \) and \( s \in S^j \), \( g(s|h) = 0 \) for every \( h \prec h^j \). This is vacuously true for \( k = 1 \) (and possibly more \( k \)'s), because \( h^1 \) is not preceded by any contingency. On the other hand, for every \( h \succ h^k \), if \( h \notin H^k \), then \( s \notin S(h) \), therefore \( g(s|h) = 0 \) as well. Thus, \( g(s|h) = 0 \) for every \( h \notin H^k \). Then, from (3) we get

\[
\sum_{s \in S^k} \sum_{l \in L(s)} \mu(s|h^k) \eta^{h,s}(l) \sum_{h \in H^k \cap H(l)} g(s|h) \leq 0,
\]

and with (4) we get

\[
\sum_{s \in S^k} \sum_{l \in L(s)} \mu(s|h^k) \eta^{h,s}(l) \sum_{h \in H^k \cap H(l)} g(s|h) = 0,
\]

which can be rewritten as (6) because \( g(s|h) = 0 \) for every \( s \in S^k \) and \( h \notin H^k \). There remains to prove the induction hypothesis for \( k + 1 \). Equation (7) can be rewritten as

\[
\sum_{h \in H^k} \sum_{s \in S^k} \left( \sum_{l \in L(s) \cap H(l)} \eta^{h,s}(l) \right) \mu(s|h^k) g(s|h) = 0.
\]

Then, together with (3), for every \( h \in H^k \) we must have

\[
\sum_{s \in S^k} p^k(h|s) \mu(s|h^k) g(s|h) = 0.
\]

Then, the agent’s expected payoff from \( g(\cdot|h) \) is zero, therefore the tie-breaking rule for acceptance requires \( g(s|h) = 0 \) for every \( s \in S \) with \( \mu(s|h) = 0 \), hence for every \( s \in S^j \) for each \( j \) such that \( h^j \succ h \).

Now I show that if \( \mu \) is not forward consistent, the agent can be deterministically Dutch-booked. If \( \mu \) is not forward consistent, for all distributions over learning paths, it violates Bayes rule between some contingencies, so in particular under the distributions for which Bayes rule boils down to conditioning. So, there exist \( h, h' \in H \) such that \( h \prec h' \), but \( \mu(s|h) \neq \mu(s|h') \mu(S(h')|h) \) for some \( s \in S \). Thus, there are \( s, s' \in S(h') \) such that \( x := \mu(s|h)/\mu(s'|h) > \mu(s|h')/\mu(s'|h') =: y \). Consider a system of gambles \( g = (g(\cdot|h))_{h \in H} \) such that, for some \( \varepsilon \in (0, x - y) \), \( g(s|h) = 1 \), \( g(s'|h) = -x + \varepsilon/3 \), \( g(s|h') = -1 - y \varepsilon/4 \), \( g(s'|h') = y + \varepsilon/3 \), and \( g(s|h) = 0 \) otherwise. It is easy to see that \( g \) is a deterministic Dutch book that the agent is willing to accept. \( \blacksquare \)
2 Sequential games: connection with Battigalli et al. (2021) and Siniscalchi (2020)

In the context of a sequential game with complete information and perfect recall, Battigalli et al. (2021) introduce a classification of the belief systems of a player over the opponents’ strategies. I study their classification of belief systems as a special case of the analysis of the main body. In a game, the moves of a player are independent of the moves of the co-players that are yet to be observed, and the two jointly determine the path of the game, thus the learning path. Therefore, the assumption that every information state is reached with the same probability under every consistent state is satisfied in that context. For concreteness, I stick to the sequential games application, but the analysis extends seamlessly to situations where this assumption is satisfied, for instance in situations where the an agent does not choose what to observe but the factors that determine the observations are independent of the state itself, as in the running example of the paper.

Let $S_{-i}$ be the set of strategy profiles of the opponents of player $i$. Let $H_i$ be the arborescence of information sets of player $i$, endowed with the precedence relation $\prec$. Each information set $h_i$ corresponds to the observation of some past moves of the opponents, summarized by the event $S_{-i}(h_i) \subseteq S_{-i}$. A belief system is an array $\mu_i = (\mu_i(\cdot|h_i))_{h_i \in H_i}$ of probability measures over $S_{-i}$ such that $\mu_i(S_{-i}(h_i)|h_i) = 1$ for every $h_i \in H_i$.

Taking as primitive a collection of conditioning events $C \subseteq 2^{S_{-i}} \setminus \{\emptyset\}$, a conditional probability system over $S_{-i}$ is an array of probability measures $\bar{\mu}_i = (\bar{\mu}_i(\cdot|C))_{C \in C}$ such that $\bar{\mu}_i(C) = 1$ for every $C \in C$, and for every $D \in C$ with $D \subset C$, for each $E \subseteq D$,

$$\bar{\mu}_i(E) = \bar{\mu}_i(E) \bar{\mu}_i(D|C).$$

A conditional probability system is complete when it is defined with respect to the collection $C = 2^{S_{-i}} \setminus \{\emptyset\}$

**Definition 4 (Battigalli et al., 2021)** A belief system $\mu_i = (\mu_i(\cdot|h_i))_{h_i \in H_i}$ over the opponents’ strategies is forward consistent if, for all $h_i, h_i' \in H_i$, if $h_i \prec h_i'$, then

$$\mu_i(E|h_i) = \mu_i(S_{-i}(h_i')|h_i) \mu_i(E|h_i').$$

for every $E \subseteq S_{-i}(h_i')$.

**Definition 5** A belief system $\mu_i = (\mu_i(\cdot|h_i))_{h_i \in H_i}$ over the opponents’ strategies is completely consistent if there exists a complete CPS $\bar{\mu}_i = (\bar{\mu}_i(\cdot|C))_{C \in 2^{S_{-i}} \setminus \{\emptyset\}}$ such that, for each $h_i \in H_i$,

$$\mu_i(\cdot|h_i) = \bar{\mu}_i(\cdot|S_{-i}(h_i)).$$

Forward consistency corresponds to Bayesian updating of conjectures along the path of play. Differently from the definition of this paper, the definition
of complete consistency of Battigalli et al. uses complete CPS’s to discipline beliefs at counterfactual contingencies. An information set $h$ does not carry any additional information about the strategies of the opponents besides ruling out all the strategies that do not belong to $S_{-i}(h)$; for this reason, it is appropriate to derive each belief by mere conditioning in place of Bayes rule, from the most general theory that explains the contingency.

**Proposition 6** For every information set $h_i$, for every pair $s_{-i}, s'_{-i} \in S_{-i}(h_i)$, and for every probability distribution over player $i$’s strategies, the probability of reaching the information set is identical under $s_{-i}$ and $s'_{-i}$: $p(h_i|s_{-i}) = p(h_i|s'_{-i})$.

**Proof.** Under each $s_{-i}$, the probability of reaching $h_i$ is the probability of the strategies of $i$ that, with $s_{-i}$, induce a path that goes through $h_i$. So all we need to show is that such set of strategies of $i$ is the same under all $s_{-i} \in S_{-i}(h_i)$. By perfect recall, all these strategies must prescribe the same moves before reaching $h_i$. Suppose by contradiction that there exists a strategy of $i$ that induces a path that goes through $h_i$ with some $s_{-i} \in S_{-i}(h_i)$ but not with some other $s'_{-i} \in S_{-i}(h_i)$. But then, all other strategies of $i$ that prescribe the same moves before $h_i$, will have the same effect: $h_i$ is not reached under $s'_{-i}$. This contradicts that $s'_{-i}$ is consistent with $h_i$. \(\blacksquare\)

Proposition 6 Although the use of complete CPSs is appropriate in this context, one might as well model the theories from which a player derives just as a LCPS, and derive each belief by conditioning the first theory in the list that is consistent with the contingency; the two definitions of complete consistency are equivalent.\(^2\)

**Proposition 7** Fix a belief system $\mu_i = (\mu_i(\cdot|h_i))_{h_i \in H_i}$. The following are equivalent:

1. there exists a complete CPS $\tilde{\mu}_i = (\tilde{\mu}_i(\cdot|C))_{C \in 2^{S_{-i}\setminus \{\emptyset\}}}$ such that for each $h_i \in H_i$,
   $$\mu_i(\cdot|h_i) = \tilde{\mu}_i(\cdot|S_{-i}(h_i));$$
2. there exists a LCPS $\hat{\mu}_i = (\mu_1^i, ..., \mu_m^i)$ such that for each $h_i \in H_i$ and $s_{-i} \in S_{-i}(h_i)$,
   $$\mu_i(s_{-i}|h_i) = \frac{\mu_k^i(s_{-i})}{\mu_k^i(S_{-i}(h_i))},$$
   where $k$ is the smallest $m$ such that $\mu_k^i(S_{-i}(h_i)) > 0$.

**Proof.**

2⇒1) For each $C \in 2^{S_{-i}\setminus \{\emptyset\}}$, derive $\tilde{\mu}_i(\cdot|C)$ by conditioning the first measure in the LCPS that assigns positive probability to $C$, so that if $C = S_{-i}(h_i)$ for some $h_i \in H_i$, then $\tilde{\mu}_i(\cdot|C) = \mu_i(\cdot|h_i)$. There remains to show that $\tilde{\mu}_i = $
Corollary 8
Fix a belief system following is a corollary of the main theorem and of the last two propositions. With the coherence rule among discounted odds ratios becomes particularly significant for sequential games. In particular, the characterization of complete consistency for every $j < k$. Therefore, also $\overline{\mu}_i(\cdot|D)$ is derived from $\mu^k_i$ and the chain rule is satisfied.

1$\Rightarrow$2 Create a decreasing chain of conditioning events $(C^1, \ldots, C^n)$ as follows. Let $C^1 = S_{-i}$. For each $k > 1$, let $C^k$ be the union of all the $s_{-i} \in S_{-i}$ such that $\overline{\mu}_i(s_{-i}|C^j) = 0$ for every $j < k$. Let $n$ be the smallest $k$ such that $C^{k+1} = \emptyset$. Consider the list of measures $\overline{\mu}_i = (\overline{\mu}_i(\cdot|C^1), \ldots, \overline{\mu}_i(\cdot|C^n))$. For each $h_i \in H_i$, let $k$ be the smallest $j \leq n$ such that $\overline{\mu}_i(S_{-i}(h_i)|C^j) > 0$ — it exists by $C^{n+1} = \emptyset$. By $\overline{\mu}_i(S_{-i}(h_i)|C^j) = 0$ for every $j < k$, it follows that $S_{-i}(h_i) \subseteq C^j$. Hence, $\mu_i(\cdot|C^k)$ can be derived from $\overline{\mu}_i(\cdot|C^k)$ by updating, and $\overline{\mu}_i(\cdot|C^k)$ is also the first measure in $\overline{\mu}_i$ that gives positive probability to $S_{-i}(h_i)$. There remains to show that $\overline{\mu}_i$ is a LCPS. For each $s_{-i} \in S_{-i}$, by $C^{n+1} = \emptyset$, $\overline{\mu}_i(s_{-i}|C^k) > 0$ for some $k \leq n$, so $\overline{\mu}_i$ has full joint support. Moreover, by $s_{-i} \in C^k$, $\overline{\mu}_i(s_{-i}|C^j) = 0$ for every $j < k$, and by $\overline{\mu}_i(s_{-i}|C^k) > 0$, $s_{-i} \notin C^j$ for every $j > k$, which implies $\overline{\mu}_i(s_{-i}|C^j) = 0$. Hence, $\overline{\mu}_i$ has disjoint supports and is therefore a LCPS.

Proposition [7] implies that the main result of this paper applies to the case of sequential games. In particular, the characterization of complete consistency with the coherence rule among discounted odds ratios becomes particularly simple, because by Proposition [6] discounted odds ratios becomes odds ratios. The following is a corollary of the main theorem and of the last two propositions.

**Corollary 8** Fix a belief system $\mu_i = (\mu_i(\cdot|h_i))_{h_i \in H_i}$. The following are equivalent:

- The belief system is completely consistent;
- For every $s_{-i}, s'_{-i} \in S_{-i}$, for every $h_i, h'_i \in H_i$ such that $s_{-i}, s'_{-i} \in S_{-i}(h_i) \cap S_{-i}(h'_i)$, $\mu_i(s_{-i}|h_i) > 0$ and

$$\frac{\mu_i(s_{-i}|h_i)}{\mu_i(s'_{-i}|h_i)} = \frac{\mu_i(s_{-i}|h'_i)}{\mu_i(s'_{-i}|h'_i)}.$$

Thus, complete consistency in games is characterized by keeping constant the odds ratios between any two strategy profiles of the opponents, as long as neither of them is ruled out.

It is interesting to note that Siniscalchi (2020) comes to a similar characterization of complete consistency starting from a different idea. Consider a belief system that is isomorphic to a conditional probability system defined over the collection of observable events. Such belief system thus satisfies the chain rule between contingencies that correspond to nested observations, but the chain rule may have little bite when the observable events are not nested, and the corresponding beliefs are not disciplined by the preliminary observation of a larger event. Therefore, Siniscalchi (2020) generalizes the chain rule to apply also between overlapping, but non-nested observable events. I conclude this section by showing the convergence between Siniscalchi’s approach and mine.
Call “generalized self-odds ratio” a generalized odds ratio between \( s_{-i} \) and itself.

**Remark 9** All the generalized odds ratios of each pair are identical if and only if all generalized self-odds ratios are 1.

**Proof.** Only if. Fix a generalized self-odds ratios. Break it into two and invert one of the two halves. We have two concatenations of odds ratios that start and end with the same element. Thus, the two corresponding generalized odds ratios are identical by assumption. Therefore, the generalized self-odds ratio is 1.

If. Fix two concatenations of odds ratios where either no odds ratio is infinite or there is no zero. If both contain a zero, the corresponding generalized odds ratios are both zero. If they both contain an infinite, they are both infinite. If one contains an infinite and the other contains a zero, the generalized self-odds ratio obtained by inverting one of the two is zero or infinite, which violates the assumption. Otherwise, the generalized self-odds ratio (in one direction) is finite. By assumption, it is 1. Therefore the two generalized odds ratios are identical. \( \blacksquare \)

**Definition 10 (Siniscalchi 2020)** A belief system \((\mu^i(\cdot|h_i))_{h_i \in H_i}\) is consistent if for every \((h_1^n, ..., h_i^n) \in H_i\) and \(E \subseteq S_{-i}(h_1^n) \cap S_{-i}(h_i^n)\), \[
\mu^i(E|h_i^n) \prod_{m=1}^{n-1} \mu^i(S_{-i}(h_i^m) \cap S_{-i}(h_i^m+1)|h_i^n) = \frac{\mu^i(E|h_i^n)}{\prod_{m=1}^{n-1} \mu^i(S_{-i}(h_i^m) \cap S_{-i}(h_i^m+1)|h_i^n)}.
\]

(8)

**Theorem 11** A belief system \((\mu^i(\cdot|h_i))_{h_i \in H_i}\) is consistent if and only if it is completely consistent.

**Proof.** We will use the characterization of compete consistency of Theorem ??.

If part. Fix \((h_1^n, ..., h_i^n) \in H_i\) and \(E \subseteq S_{-i}(h_1^n) \cap S_{-i}(h_i^n)\). If both sides of equation \(8\) are zero, consistency holds. Otherwise, suppose without loss of generality that the right-hand side is not zero, so that equation \(8\) can be rewritten as

\[
\frac{\mu^i(E|h_i^n)}{\mu^i(S_{-i}(h_1^n) \cap S_{-i}(h_2^n)|h_1^n)} \cdot \frac{\mu^i(S_{-i}(h_i^n) \cap S_{-i}(h_i^n)|h_1^n)}{\mu^i(E|h_i^n)} = 1.
\]

(9)

Note that equation \(8\) holds if for every \( s_{-i} \in E \) such that the ratio \( \mu^i(s_{-i}|h_1^n)/\mu^i(s_{-i}|h_i^n) \) is not indeterminate,

\[
\frac{\mu^i(s_{-i}|h_1^n)}{\mu^i(s_{-i}(h_1^n) \cap S_{-i}(h_2^n)|h_1^n)} \cdot \frac{\mu^i(s_{-i}(h_i^n) \cap S_{-i}(h_i^n)|h_1^n)}{\mu^i(s_{-i}|h_i^n)} = 1,
\]

(10)

because this means that all the non-indeterminate ratios \( \mu^i(s_{-i}|h_1^n)/\mu^i(s_{-i}|h_i^n) \) are identical, which implies that the ratio \( \mu^i(E|h_1^n)/\mu^i(E|h_i^n) \) takes the same
value too. We are going to show that for every \( m = 2, ..., n \), there is \( s_{m_i} \in S_{-i}(h_i^{m-1}) \cap S_{-i}(h_i^m) \) such that

\[
\frac{\mu^i(S_{-i}(h_i^{m-1}) \cap S_{-i}(h_i^m)|h_i^m)}{\mu^i(S_{-i}(h_i^{m-1}) \cap S_{-i}(h_i^m)|h_i^{m-1})} = \frac{\mu^i(s_{m_i}|h_i^m)}{\mu^i(s_{m_i}|h_i^{m-1})},
\]

then, we can rewrite equation 10 as

\[
\frac{\mu^i(s_{-i}|h_i^1)}{\mu^i(s_{-i}|h_i^1)} \cdot ... \cdot \frac{\mu^i(s_{-i}|h_i^n)}{\mu^i(s_{-i}|h_i^n)} = 1,
\]

which is true by complete consistency (cf. Remark 9). Let \( s_{m_i} \) be any \( s_{-i} \in S_{-i}(h_i^{m-1}) \cap S_{-i}(h_i^m) \) such that \( \mu^i(s_{m_i}|h_i^{m-1}) > 0 \); one exists by the fact that the right-hand side of 8 is not zero. By complete consistency, for each \( s'_{m_i} \in S_{-i}(h_i^{m-1}) \cap S_{-i}(h_i^m) \), we have

\[
\frac{\mu^i(s_{m_i}|h_i^m)}{\mu^i(s'_{m_i}|h_i^m)} = \frac{\mu^i(s_{m_i}|h_i^{m-1})}{\mu^i(s'_{m_i}|h_i^{m-1})}
\]

With some algebra, it is easy to show that then

\[
\frac{\mu^i(s_{m_i}|h_i^m)}{\mu^i(S_{-i}(h_i^{m-1}) \cap S_{-i}(h_i^m)|h_i^m)} = \frac{\mu^i(s_{m_i}|h_i^{m-1})}{\mu^i(S_{-i}(h_i^{m-1}) \cap S_{-i}(h_i^m)|h_i^{m-1})},
\]

as we wanted to show.

Only if part. By Remark 9 we can show that for each \( s_{-i} \in S_{-i} \), for every generalized self-odds ratio, equation 11 holds. Without loss of generality, we can assume that the generalized self-odds ratio is not infinite (because an infinite one can be inverted and becomes zero), so that no denominator in equation 11 is zero. For each \( m = 2, ..., n \), by consistency, equation 8 for \( E = \{ s_{m_i} \} \) and \( (h_1^1, ..., h_n^1) = (h_i^{m-1}, h_i^m) \) yields

\[
\frac{\mu^i(s_{m_i}|h_i^m)}{\mu^i(s_{m_i}|h_i^{m-1})} = \frac{\mu^i(S_{-i}(h_i^{m-1}) \cap S_{-i}(h_i^m)|h_i^m)}{\mu^i(S_{-i}(h_i^{m-1}) \cap S_{-i}(h_i^m)|h_i^{m-1})}.
\]

Then, substituting in 11 we obtain equation 10 which coincides with equation 9 for \( E = \{ s_{-i} \} \), and thus holds by consistency.
A Dutch Book Argument for Belief Consistency

Emiliano Catonini†

February 22, 2022

Abstract: An agent progressively learns about a state of the world. A bookmaker is ready to offer one bet after every new discovery. I say that the agent is Dutch-booked when she is willing to accept every single bet, but her expected payoff is negative under each state, where the expected payoff is computed with the objective probabilities of different discoveries conditional on the state. I introduce a rule of coherence among beliefs after counterfactual discoveries that is necessary and sufficient to avoid being Dutch-booked. This rule characterizes an agent who derives all her beliefs with Bayes rule from the same Lexicographic Conditional Probability System (Blume, Brandenburger and Dekel, 1991).

Keywords: Dutch book, lexicographic beliefs, Bayesian learning

1 Introduction

The classical Dutch book argument (Ramsey 1931, De Finetti 1937) is a cornerstone of probability theory. It says that an agent may accept a set of bets that return an aggregate loss under every state (i.e., a Dutch book) if her assessments of the relative likelihoods of states cannot be summarized by a probabilistic belief. Later (Lehman 1955, Kemeny 1955), also the converse was proven: an agent cannot be “Dutch-booked” if she evaluates all bets with the same, well-defined probability measure over the possible states.

In the classical Dutch book argument there is no time dimension. But an agent may progressively learn about the state and the bookmaker could offer the different bets in a sequence — a diachronic Dutch book. In this case, the agent can be Dutch-booked even if she holds a well-defined probabilistic belief at every point in time. If the agent does not deem any state impossible, Teller

*I would like to thank Pierpaolo Battigalli, Julien Manili, and Brad Weslake for the conversations, as well as all the attendants of the FLSS seminar at NYU-Shanghai.
†New York University in Shanghai, emiliano.catonini@nyu.edu.
‡Differently from the disease and the auction, there is no known historical reason why such a system of bets was ascribed to the Dutch as well.
(1973) and Lewis (1999) have shown that she can be Dutch-booked if and only if her beliefs are not updated with the rules of conditioning.

This paper deals with mutually exclusive discoveries. Constructing a Dutch book across counterfactual contingencies may sound like a pointless exercise: if some contingency is not reached, some bet is not sold, so the bookmaker cannot secure the profit. Yet, there are at least two possible objections against this view.

- There do exist meaningful ways to aggregate bets across counterfactual contingencies. While the probabilities of states in different contingencies may be subjective, the probabilities of the contingencies given the states can be objective, and the agent may be Dutch-booked in objective expected terms. Then, the agent, or a population of like-minded agents subject to identical experiments, will suffer a certain loss in the long run, if the situation is repeated, or in aggregate.

- The currently accepted interpretation of Dutch book arguments is not the literal one, whereby the agent incurs in a certain loss in a particular problem, rather that the agent is exposed to making detrimental choices, because of inconsistencies in the process of belief formation that the Dutch book argument brings to light. A population of agents that is exposed to Dutch-booking may therefore suffer an evolutionary disadvantage.

I illustrate these ideas with an example.

**Example 1**  
Larry wants to go out and watch the new year’s fireworks, but the municipality does not disclose anymore their location to avoid crowding. There are three possible locations: the main square, the marina, and the central park. Larry is late, so his only chance to see the fireworks is to take the first bus that arrives and hope it goes to the right location. There are three bus lines on his street: the red line goes to the square, the blue line to the marina, and the green line to the park. The busses always arrive in this order, with constant time intervals. Suppose that Larry gets on the blue line. When he arrives to the marina, he discovers whether the fireworks are there or not. In this second scenario, Larry is disappointed and regrets not arriving to the bus stop a little earlier, which meant catching the red line: “I bet the fireworks are at the square!” Overhearing this, a bookie approaches Larry and proposes him a bet on which of the other two locations is hosting the fireworks. Because of his regret, he accepts a bet with a win of 9 if the fireworks are at the square and a loss of 10 if the fireworks are in the park. By the same token, in the park he would accept a bet with a win of 9 for marina and a loss of 10 for square, and at the square he would

2If the agent assigns probability 0 to a state, she might be trivially Dutch-booked by being offered a loss for that state. I will rule out this uninteresting possibility with a simple tie-breaking rule. Rescorla (2021) extends the diachronic Dutch book theorem of Teller (1973) and Lewis (1999) to the case of subjectively impossible states with a more sophisticated rule of belief revision.

3See Rescorla (2021) for a recent, critical review.
accept a bet with a gain of 9 for park and a loss of 10 for marina. Note that, if Larry was to find himself in all of these three mutually exclusive contingencies, he would accept bets that, in aggregate, induce a loss of 1 no matter the location of the fireworks.

In the example, the bus schedule determines the objective probabilities of the three contingencies of interest, and is independent of the state (a special case that will be analysed in detail). If this situation repeats every year, Larry will find himself with the same frequency in the two contingencies that are consistent with each state, therefore, in the long run, his balance will approach −1 under each state. Or, a large population of agents who form beliefs like Larry would split themselves on the three different busses, and those who do not take the right bus would accept bets that in aggregate constitute a Dutch book. This happens because Larry exhibits a form of incoherence between his beliefs in different contingencies: the location that he deems more likely changes every time. In this simple example, if Larry had a prior, he could update the prior in at least two of the contingencies, and the induced coherence between beliefs would shelter him from Dutch-booking. In more complex environments, having a prior belief is not enough.

The main result of this paper determines the coherence rule among beliefs that is necessary and sufficient to avoid Dutch-booking in objective expected terms. The coherence rule is easy to interpret; it generalizes the idea that differences in the subjective odds ratio of two states across contingencies shall only reflect differences in the relative likelihoods of the contingencies under the states. The generalization is about comparing the likelihood of two states indirectly, using the odds ratios with other states: if Larry deems square twice more likely than park when he’s at the marina and marina twice more likely than square when he’s at the park, he should deem marina four times more likely than park when he’s at the square. In turn, I show that this coherence rule characterizes a Completely Consistent Belief System (henceforth, CCBS): an array of beliefs, one for each contingency, that can be derived with Bayes rule from a Lexicographic Conditional Probability System (Blume et al., 1991; henceforth, LCPS), a list of theories in a priority order. This means that each belief is derived by updating with Bayes rule the first theory in the list that is consistent with the contingency. It is worth nothing that a system of beliefs that satisfies Bayesian updating may not satisfy the coherence rule, thus be susceptible of Dutch-booking, because Bayesian updating does not discipline beliefs across counterfactual surprises.

In the context of sequential games, Siniscalchi (2020) and Battigalli, Catonini and Manili (2021) define a CCBS over the opponents’ strategies as an array of

---

4To keep the definition of Dutch book simple, sufficiency relies on two light auxiliary assumptions: (i) the agent can recognize when a discovery objectively rules out a state (e.g., see that there are no fireworks at the marina) and thus gives probability 0 to it, otherwise she could be easily Dutch-booked with a bet on that state; (ii) when the agent deems a state subjectively impossible, she turns down bets with a loss under that state and no win otherwise, an obvious Dutch book.
beliefs, one for each information set, that is consistent with a complete Conditional Probability System (henceforth, CPS), one belief conditional on each subset of opponents’ strategies, disciplined by the chain rule of probability. For a player of a sequential game, the observation of the opponents’ moves depends on the own moves, and the two are independent. Therefore, given a distribution over our player’s strategies, each information set is reached with the same probability under all the strategies of the opponents that are consistent with it. Under the condition that the probability of each contingency is the same under all the consistent states, Bayes rule boils down to the chain rule; moreover, a complete CPS can be derived from a LCPS by conditioning; the two things combined imply that the definition of CCBS in Siniscalchi (2020) and Battigalli et al. (2021) is a special case of the definition of this paper. Thus, the Dutch-book theorem of this paper implies that the standard notion of CPS defined on the collection of observable events falls short of the coherence that prevents Dutch-booking. Moreover, the coherence rule provides an alternative foundation of CCBSs, which can be interpreted as a property of introspection or as a wired-in property of belief formation, in place of the less plausible formulation of conditional beliefs on virtual, unobserved events (see also Battigalli et. al, 2021).

In the problems considered so far, and more generally in this paper, the state does not determine entirely the agent’s discoveries; this is why there are different, counterfactual contingencies that are consistent with the same state. In principle, one could enrich the state space with the possible learning paths of the agent; then, counterfactual discoveries would never be consistent with the same state. But the Dutch book argument hinges on a clear separation between objective and subjective uncertainty, hence the choice of the former model. Interestingly, given a diachronic Dutch book, the agent may have in mind wrong objective probabilities of the discoveries given the states, but as long as she does Bayesian updating with them, she will reject some bet. I prove this in the online appendix, and I obtain a diachronic Dutch book argument for “forward consistency”. Since Bayes rule boils down to mere conditioning when the probability of each contingency is taken to be the same under every consistent state, this result reconciles the perspective of this paper with the Dutch book theorems in the literature, and gives them a foundation for when the state space (i.e. the subject of betting) cannot fully explain the observable events.

5CPSs were first introduced by Renyi (1995). Complete CPSs were introduced in the analysis of games by Myerson (1986).
6Siniscalchi (2020) characterizes CCBSs with a generalization of the chain rule that applies also between non-nested conditional events. As my generalization of the constant odds ratios property to comparisons between different odds ratios, Siniscalchi’s generalization of the chain rule is meant to strengthen a coherence rule that loses bite in the passage from the rich collection of conditioning events used by complete CPSs to the poorer one induced by the actual contingencies. In the Online Appendix, I show directly the convergence between my approach and Siniscalchi’s approach.
2 Framework

The space of uncertainty of the agent is represented by a finite set $S$ of possible states. I describe the possible learning processes through an arborescence of contingencies $H$, endowed with the (chronological) precedence relation $\prec$. A learning path is a sequence $(h_1, ..., h_n)$ of contingencies where $h_1$ does not follow any $h \in H$, each $h_{k+1}$ immediately follows $h_k$, and $h_n$ is not followed by any $h \in H$. Let $L$ denote the set of possible learning paths and $L(h)$ the set of paths that go through a contingency $h$. Each state $s \in S$ induces an objective distribution $\eta^s \in \Delta(L)$ over learning paths. Let $L(s)$ denote the set of paths $l$ that are consistent with $s$, that is, $\eta^s(l) > 0$. For each $h \in H$, let $p(h|s) = \sum_{l \in L(h)} \eta^s(l)$ denote the probability of reaching discovery $h$ under state $s$. Let $S(h)$ denote the set of states $s$ that are consistent with contingency $h$, that is, $p(h|s) > 0$. Without loss of generality, I assume that every contingency is consistent with some state. As anticipated, an important special case is the one in which the probability of a contingency $h$ is the same across all the consistent states $s$, that is, $p(h|s) = p(h|s') =: p_h$ for all $h \in H$ and $s, s' \in S(h)$.

I illustrate the framework by formalizing the example of the introduction.

Example 2 There are three possible states, corresponding to the location of the fireworks: $S = \{sq, ma, pa\}$. There are six possible contingencies, each with a different set of consistent states, so I am going to label each contingency $h$ with the (initials of) the states in $S(h)$:

$$H = \{sq, ma, pa, sm, mp, ps\}.$$  

For instance, Larry ends up in contingency $sq$ if he takes the bus that goes to the square and find the fireworks there; in contingency $mp$ if he takes the bus to the square and the fireworks are not there. Note that all the contingencies are unordered, therefore contingencies learning paths coincide. Under each state $s \in S$, the bus schedule induces a uniform distribution over the three contingencies $h$ such that $s \in S(h)$. For instance, in state $sq$, Larry may end up in contingencies $sq, sm$, and $ps$. So, this is a situation where, for each $h \in H$, $p(h|s)$ is the same ($1/3$) for every $s \in S$. $\Delta$
3 Belief systems

I assume that after each discovery $h$ the agent forms a belief $\mu(\cdot|h) \in \Delta(S)$. I call $(\mu(\cdot|h))_{h \in H}$ belief system. To avoid uninteresting cases, throughout the paper I assume $\mu(S(h)|h) = 1$ for every $h \in H$.

I will define a Completely Consistent Belief System (CCBS) as a system of beliefs that can be derived from a Lexicographic Conditional Probability System (LCPS).

**Definition 3 (Blume et al., 1991)** A list of probability measures $\bar{\mu} = (\mu^1, ..., \mu^n)$ over $S$ is a Lexicographic Conditional Probability System if for every $s \in S$, there exists exactly one $m \in \{1, ..., n\}$ such that $\mu^m(s) > 0$.

Definition 3 embodies a “full-support” requirement of the LCPS: every state is given positive probability at some level. In this way, from every LCPS, one can derive a belief system as follows.

**Definition 4** A belief system $(\mu(\cdot|h))_{h \in H}$ is completely consistent if there exists a LCPS $\bar{\mu} = (\mu^1, ..., \mu^n)$ such that, for each $h \in H$,

$$\mu(s|h) = \frac{p(h|s)\mu^m(s)}{\sum_{s'} p(h|s')\mu^m(s')}$$

where $m$ is the smallest $k$ such that $\mu^k(S(h)) > 0$.

In a CCBS, given each contingency $h$, the belief $\mu(\cdot|h)$ is derived with Bayes rule from the first theory in the LCPS that is able to explain $h$, i.e., assigns positive probability to some state that can induce $h$. Note that at two contingencies $h, h'$ with the same set of consistent states $(S(h) = S(h'))$, the beliefs derived from the LCPS with Bayes rule may still differ, although they are derived from the same measure in the LCPS. This is the reason to prefer LCPSs over complete CPSs when the contingencies are not reached with the same probability under every state; technically one could derive both $\mu(\cdot|h)$ and $\mu(\cdot|h')$ with Bayes rule from the belief conditional on $S(h) = S(h')$, but the fact that all these beliefs would be different clashes with the existing notion of complete consistency (Siniscalchi 2020, Battigalli et al. 2021), which requires the contingent beliefs to coincide with the corresponding conditional belief in the complete CPS.

The key notion of coherence among beliefs will compare the relative probabilities of states at different contingencies. However, given that the objective probabilities of different contingencies may differ across states, these comparisons are meaningful only after discounting the subjective probabilities of states by the likelihood that the state induces the contingency.

---

7The related, deeper reason is that it is hard to conceive an agent who thinks “what would I believe if I were to observe just $S(h)$”, in a context where the actual contingencies contain more information than ruling out some states. The formulation of alternative theories of the world as in a LCPS, just conditional on the previous theories being wrong, seems more plausible.
Definition 5 Given a belief system $(\mu(h))_{h \in H}$, a contingency $h$, and a pair of states $s, s' \in S(h)$, the discounted odds ratio between $s$ and $s'$ at $h$ is

$$\frac{\mu(s|h) \cdot p(h|s')}{p(h|s) \cdot \mu(s'|h)}$$

provided that it is not indeterminate. A generalized odds ratio between $s$ and $s'$ is the product of a finite concatenation of discounted odds ratios

$$\frac{\mu(s|1) \cdot p(1|s)}{p(1|s) \cdot \mu(s')} \cdot \frac{\mu(s'|2) \cdot p(2|s)}{p(2|s) \cdot \mu(s'')} \cdots \cdot \frac{\mu(s'|n) \cdot p(n|s)}{p(n|s) \cdot \mu(s''')}$$

provided that it is not indeterminate.

A discounted odds ratio compares the probabilities attributed to two states at the same contingency. When the probabilities $p(h|s)$ do not depend on the state, discounted odds ratios boil down to (standard) odds ratios and will be referred to as such. A generalized odds ratio compares the probabilities of two states at two different contingencies, via a chain of comparisons with other states.

Now I illustrate the notions of belief system and odds ratios through the running example.

Example 6 For each state $s$, Larry’s belief system $\mu = (\mu(h))_{h \in H}$ must assign probability 1 to $s$ at the contingency labelled with $s$.

To obtain a CCBS, let us postulate a LCPS $\tilde{\mu} = (\mu^1, \mu^2, \mu^3)$ ($n \in \{1, 2, 3\}$) from which Larry derives his beliefs at the other three contingencies. Suppose that $\mu^1$ assigns positive probability to at least 2 states. Then, $\mu^1(S(h)) > 0$ for each $h = sm, mp, ps$. Because a contingency has the same probability under every state, Bayes rule boils down to the chain rule (and discounted odds ratios boil down to odds ratios). So, for each $h = sm, mp, ps$, writing $S(h) = \{s, s'\}$, we have

$$\mu(s|h) = \frac{\mu^1(s)}{\mu^1(s) + \mu^1(s')}$$

If instead $\mu^1$ assigns probability 1 to a state $\tilde{s}$, we obtain $\mu(\tilde{s}|h) = 1$ whenever $\tilde{s} \in S(h)$, and the belief at the contingency $h$ with $S(h) = S \setminus \{s\} = \{s, s'\}$ is

$$\mu(s|h) = \frac{\mu^2(s)}{\mu^2(s) + \mu^2(s')}$$

Given any two states $s, s'$, one can compute an odds ratio between $s$ and $s'$ at the contingency $h$ with $S(h) = \{s, s'\}$; call it $o'(s, s')$. Then, one can also compute a generalized odds ratio between $s$ and $s'$ as $o'(s, s'') \cdot o(s''', s')$ ($s'' \neq s, s'$), unless one is zero and the other infinite. For instance, when $\mu^1(ma) \in (0, 1)$, we can compute

$$o(sq, ma) \cdot o(ma, pa) = \frac{\mu(sq|sm)}{\mu^{ma|sm}} \cdot \frac{\mu^{ma|mp}}{\mu^{pa|mp}} = \frac{\mu^1(sq)}{\mu^1(ma)} \cdot \frac{\mu^1(ma)}{\mu^1(pa)} = \frac{\mu^1(sq)}{\mu^1(pa)} = \frac{\mu^1(sq)}{\mu^1(pa)} = o(sq, pa),$$
which can be infinite if $\mu_1(pa) = 0$ and thus $\mu_1(sq) > 0$, or zero, vice versa. Note the equality between the odds ratio and the generalized odds ratio.

An alternative belief system, which rationalizes Larry’s betting behavior, is the following:

$$\bar{\mu}(sq|ps) = \bar{\mu}(ma|sm) = \bar{\mu}(pa|mp) = \frac{3}{4}.$$  

In this case, the odds ratio and the generalized odds ratio between any two states differ: the odds ratio between $sq$ and $pa$, or $ma$ and $sq$, or $pa$ and $ma$ is 3; however, the generalized odds ratios are all $\frac{1}{9}$. $\triangle$

It is also worth noting that an agent does not need to know the probability of each contingency, when it is constant across states, to form completely consistent beliefs; Larry could also believe that the green line is suppressed if he does not take it, but as long as he does not perceive inexistent correlations between the location of the firework and the arrivals of the different lines, he has no reason to introduce that bias in his beliefs.

4 Dutch books and Dutch-booking

In this section, I introduce the notion of Dutch Book, and a novel sense in which an agent can be “Dutch-booked”. Traditionally, a Dutch book is a set of bets such that, no matter the realization of the uncertainty, the aggregate payoff from the bets is never positive and sometimes negative for the gambler (thus never negative and sometimes positive for the bookmaker). In general, an agent is Dutch-booked if she accepts all the bets in a Dutch book. Our agent, however, would never accept a classical Dutch book if all the bets are proposed at the same time, as she has a well-defined probabilistic belief at any point in time — this follows from the classical Dutch book theorem. In particular, all the bets that are offered to our agent at the same contingency can be coalesced into an aggregate bet, which she will accept whenever she is willing to accept every single bet. Thus, for the sake of Dutch-booking, we can assume without loss of generality that our agent is proposed just one bet over the possible states at every information state.

Definition 7 A gamble is a map $\gamma : S \rightarrow \mathbb{R}$ that specifies, for each $s \in S$, the gain or loss $\gamma(s)$ for player $i$ in case of $s$.

We say that $\gamma$ is a gamble on $\tilde{S} \subseteq S$ if $\gamma(s) = 0$ for every $s \in S \setminus \tilde{S}$.

A system of gambles is an array $g = (g(\cdot|h))_{h \in \mathcal{H}}$ where, for each $h \in \mathcal{H}$, $g(\cdot|h)$ is a gamble on $S(h)$.

What is a meaningful notion of Dutch book for a system of gambles? Since only the gambles along the realized learning path will actually be proposed, a strong notion of Dutch book requires that, under every state, the agent makes no cumulative gain, and sometimes make a cumulative loss, along each possible learning path. This is the notion of diachronic Dutch book employed in the literature, which I formalize and analyze in the Online Appendix. But given
that there are objective probabilities of contingencies, it also makes sense to introduce a weak notion of Dutch book, whereby the agent makes no gain or a loss in objective expected terms.

**Definition 8** A system of gambles \( g = (g(h))_{h \in \mathcal{H}} \) is **Dutch book** if for every \( s \in \mathcal{S} \),

\[
\sum_{h \in \mathcal{H}} p(h|s)g(s|h) \leq 0,
\]

and for some \( s \in \mathcal{S} \),

\[
\sum_{h \in \mathcal{H}} p(h|s)g(s|h) < 0.
\]

As anticipated, we are taking the view that the agent is proposed each gamble in the system at the corresponding contingency. The acceptance of each gamble will then depend on the belief at that contingency.

**Definition 9** Fix a probability measure \( \nu \in \Delta(\mathcal{S}) \) and a gamble \( \gamma \). We say that player \( i \) is **willing to accept** \( \gamma \) given \( \nu \), if either

\[
\sum_{s \in \mathcal{S}} \nu(s)\gamma(s) > 0,
\]

or

\[
\sum_{s \in \mathcal{S}} \nu(s)\gamma(s) = 0
\]

and \( \gamma(s) \geq 0 \) for every \( s \in \mathcal{S} \) such that \( \nu(s) = 0 \).

I say that \( i \) with belief system \( (\mu(h))_{h \in \mathcal{H}} \) is willing to accept a system of gambles \( g = (g(h))_{h \in \mathcal{H}} \) if, for every \( h \in \mathcal{H} \), she is willing to accept \( g(h) \) given \( \mu(h) \).

I say that \( i \) is **Dutch-booked** if she is willing to accept a Dutch book.

Definition 9 features a tie-breaking rule for the gambler: when the subjective expected payoff of a gamble is zero, the gambler will turn it down if there is a loss for some subjectively impossible state. The only aim of the tie-breaking rule is focusing on the Dutch book argument of interest for this paper, by ruling out in a simple way that the agent can be Dutch-booked just because she accepts bets with zero expected payoff but a negative payoff for the subjectively impossible states — note that this would still constitute Dutch-booking because the definition of Dutch book is agnostic regarding the probabilities of states.

To construct a Dutch book, a bookmaker needs to know the objective probabilities \( p(h|s) \). To construct a Dutch book that will certainly be accepted by the agent, i.e. to Dutch-book the agent, the bookmaker needs to know the agent’s beliefs; but even without knowing \( (\mu(h))_{h \in \mathcal{H}} \), the bookmaker can still try to

---

8By doing so, the tie-breaking rule substitutes the rule of belief revision of Rescorla (2021). The tie-breaking rule could be relaxed in favour of acceptance of negative returns under a null state in exchange for positive returns under another null state, provided that the positive returns more than compensate negative returns for that same state already accepted in the past. The advantage of the present, simpler rule is that it does not require to look at the system of gambles as a whole.
Dutch-book the agent. Be as it may, as argued in the introduction, the currently preferred interpretation of Dutch book arguments is not that an agent can be certainly tricked into a loss by an omniscient opponent, rather that the agent is exposed to making irrational, detrimental choices.

To conclude the section, I illustrate Dutch-booking in the running example.

**Example 10** Suppose that Larry has the belief system $\mu$ constructed before. The bookmaker construct the following system of bets:

\[
g(sq|ps) = g(ma|sm) = g(pa|mp) = 9; \\
g(pa|ps) = g(sq|sm) = g(ma|mp) = -10; \\
g(s|h) = 0 \quad \forall h \in \{sq, ma, pa\}, \forall s \in S(h).
\]

Under every state $s$, we have

\[
\sum_{h \in H} p(h|s)g(s|h) = \frac{1}{3}(9 - 10 + 0) = -\frac{1}{3} < 0,
\]

i.e., Larry has an objective expected payoff of $-1/3$. Therefore, the system is a Dutch book. There remains to show that Larry is willing to accept it. For each $h = sm, mp, ps$, we have

\[
\sum_{s \in S} \mu(s|h)g(s|h) = \frac{3}{4} \cdot 9 + \frac{1}{4}(-10) + 0 \cdot 0 > 0.
\]

Therefore, Larry is Dutch-booked. $\triangle$

## 5 The Dutch book argument for complete consistency

The main theorem generalizes the Dutch-book argument and the relationship between CCPSs and generalized odds ratios of the running example: incongruities among generalized odds ratios can be exploited for Dutch-booking and can only be displayed by a belief system that is not completely consistent. But the main theorem also shows the opposite implications: when there is no incongruity between generalized odds ratios, the agent’s belief system is completely consistent and she cannot be Dutch-booked. Thus, consistency among generalized odds ratios is the coherence requirement that characterizes complete consistency and impossibility of Dutch-booking.

**Theorem 11** Fix a belief system $\mu$. The following are equivalent.

1. for every pair of states $s, s'$, all generalized odds ratios are identical;
2. $\mu$ is completely consistent;

3. player $i$ cannot be Dutch-booked given $\mu$.

**Proof.** We prove that 1 implies 2, that 2 implies 3, and that 3 implies 1. The acronyms DOR and GOR will stand for discounted/generalized odds ratio.

1$\Rightarrow$2)

Note preliminarily that, by Condition 1, if there is a DOR between two states $s, s'$, all the DORs between $s$ and $s'$ coincide with the same value; call it $o(s, s')$.

Let $P^1$ be the set of all $s \in S$ that have no zero GOR with any $s'$.

Let $P^2$ be the set of all $s \in S \setminus P^1$ that have a zero GOR only with some $s' \in P^1$.

Let $P^3$ be the set of all $s \in S \setminus (P^1 \cup P^2)$ that have a zero GOR only with some $s' \in P^1 \cup P^2$.

And so on. Call $P^{n+1}$ the first empty set obtained in this way. (Of course, it can already be $P^2$.)

**Claim 1:** $(P^1, ..., P^n)$ partitions $S$.

Proof: The sets are disjoint by construction. There remains to show that $\bigcup_{m=1}^n P^m = S$. Suppose not. Let $\tilde{S} = S \setminus (\bigcup_{m=1}^n P^m)$. Emptiness of $P^{n+1}$ implies that every $s \in \tilde{S}$ has a zero GOR with some other $s' \in \tilde{S}$. Consider a directed graph with nodes $\tilde{S}$ and arrows that represent a zero GOR between the start and the end node. Since every node is the start of an arrow, the graph has a cycle $(s^1, s^2, ..., s^k = s^1)$. So, the product of all the (zero) GORs in the cycle is a zero GOR between $s^1$ and itself. Inverting all the DORs in this GOR, we obtain an infinite GOR between $s^1$ and itself. This violates the assumption that all the GORs between $s^1$ and itself are identical (Condition 1).

**Claim 2:** for each $m = 1, ..., n$, there exists a probability measure $\nu^m$ with support $P^m$ such that, for every pair $(s, s') \in P^m \times P^m$ that has a DOR given $\mu$, $\nu^m(s)/\nu^m(s') = o(s, s')$.

Proof: Create an ordering $(s^1, ..., s^{|P^m|})$ of $P^m$ and a map $\tilde{\nu} : P^m \to \mathbb{R}^+$ with the following recursive procedure. Pick $s^1$ arbitrarily and let $\tilde{\nu}(s^1) = 1$. For each $q = 2, ..., k$, if possible, pick $s^q$ that has a DOR with $s^p$ for some $p < q$, let $p$ be the largest such $p'$, and let $\tilde{\nu}(s^q) = \tilde{\nu}(s^p) \cdot o(s^q, s^p)$; otherwise, pick $s^q$ arbitrarily and let $\tilde{\nu}(s^q) = 1$. By definition of $P^m$, no $s \in P^m$ can have a zero or infinite DOR with any other $s' \in P^m$, so $\tilde{\nu}(s) \in \mathbb{R}^+$ for every $s \in P^m$. Thus, we can get a probability measure $\nu^m$ with support $P^m$ by rescaling $\tilde{\nu}$.

To show that $\nu^m$ satisfies the desiderata, fix $p$ and $q$ with $1 \leq p < q \leq |P^m|$. By construction of $\tilde{\nu}$, if there is a DOR between $s^q$ and $s^p$, then there is a finite
sequence \((\bar{s}^1, ..., \bar{s}^T)\) with \(s^1 = s^p\) and \(s^T = s^q\) such that, for each \(t = 2, ..., T\), we have \(\nu(\bar{s}^t) = \nu(\bar{s}^{t-1}) \cdot o(\bar{s}^t, \bar{s}^{t-1})\). Therefore
\[
\frac{\nu_m(s^q)}{\nu_m(s^p)} = \frac{\bar{\nu}_i(s^q)}{\bar{\nu}_i(s^p)} = o(\bar{s}^T, \bar{s}^{T-1}) \cdot \frac{\bar{\nu}(\bar{s}^{T-1})}{\bar{\nu}(s^{T-1})} = \prod_{t=2}^{T} o(\bar{s}^t, \bar{s}^{t-1}),
\]
and the latter is a GOR between \(s^q\) and \(s^p\), which is equal to \(o(s^q, s^p)\) by Condition 1.

Now let \(\bar{\mu} = (\nu_1, ..., \nu_n)\). By Claim 1, it is a LCPS. Now fix \(h \in \mathcal{H}\). Let \(m\) be the smallest \(k\) such that \(\nu^k(S(h)) > 0\). I show that \(\mu(\cdot|h)\) can be derived from \(\nu^m\) with Bayes rule. Fix \(s\) in the support of \(\mu(\cdot|h)\). Thus, \(s\) has non-zero DOR with every \(s' \in S(h)\), so in particular with some \(s' \in P^m\) such that \(\nu^m(s') > 0\). Since \(s' \in P^m\), it has non-zero DOR with every \(s'' \in P^m \cup ... \cup P^n\), so, in turn, by condition 1, \(s\) has non-zero DOR with \(s''\) as well. Hence \(s \in P^1 \cup ... \cup P^m\). Since \(\nu^k(S(h)) = 0\) for every \(k < m\), we conclude that \(s \in P^m\). By contrast, every \(s' \in S(h)\) with \(\mu(s'|h) = 0\) has zero DOR with \(s\), thus cannot be in \(P^m\). We conclude that the support of \(\mu(\cdot|h)\) is \(P^m \cap S(h)\). Given this, we only need to check that the (non-discounted) odds ratio of any pair \(s, s' \in P^m \cap S(h)\) obtained from \(\nu^m\) with Bayes rule coincides with \(\mu(s|h)/\mu(s'|h)\). We have
\[
\frac{p(h|s)\mu^m(s)}{p(h|s')\mu^m(s')} = \frac{\mu(s|h)}{\mu(s'|h)} = \frac{\mu(s|h)}{\mu(s'|h)},
\]
where the first equality is by Claim 2.

2⇒3)

Let \(\bar{\mu} = (\mu_1, ..., \mu_n)\) a LCPS from which the belief system is derived. For each \(m = 1, ..., n\), let \(S^m\) denote the support of \(\mu^m\), and let \(H^m\) denote the set of all \(h \in \mathcal{H}\) such that \(S(h) \cap S^m \neq \emptyset\) and \(S(h) \cap S^k = \emptyset\) for every \(k < m\), so that the belief at \(h\) is derived from \(\mu^m\).

Fix a system of gambles \(g^* = (g^*(\cdot|h))_{h \in \mathcal{H}}\). Suppose that \(i\) is willing to accept \(g^*\). I show that then \(g^*\) is not a Dutch book. Let \(g = (g(\cdot|h))_{h \in \mathcal{H}}\) be the system of gambles such that, for each \(h \in \mathcal{H}\) and \(s \in S\), if \(\sum_{s'} \mu(s'|h)g(s'|h) = 0\) and \(\mu(s|h) = 0\), then \(g(s|h) = 0 \leq g^*(s|h)\) (the inequality comes from Definition 9), otherwise \(g(s|h) = g^*(s|h)\). It is easy to see that \(i\) is willing to accept also \(g\).

Fix \(m = 1, ..., n\). Since \(i\) is willing to accept \(g\), for every \(h \in H^m\), we have
\[
\sum_{s \in S} \frac{p(h|s)\mu^m(s)}{\sum_{s'} p(h|s')\mu^m(s')} g(s|h) \geq 0
\]
which implies
\[
\sum_{s \in S^m} p(h|s)\mu^m(s) g(s|h) \geq 0.
\]
So we can write
\[
\sum_{h \in H^m} \sum_{s \in S^m} p(h|s)\mu^m(s) g(s|h) \geq 0,
\]

12
and rewrite it as
\[ \sum_{s \in S^m} \mu^m(s) \sum_{h \in H^m} p(h|s)g(s|h) \geq 0. \] (5)

Suppose that
\[ \forall s \in S, \quad \sum_{h \in H} p(h|s)g(s|h) \leq \sum_{h \in H} p(h|s)g^*(s|h) \leq 0, \] (6)
otherwise, obviously, \( g^* \) is not a Dutch book. I am going to show with a recursive procedure that, for every \( m = 1, \ldots, n \),
\[ \sum_{s \in S^m} \mu^m(s) \sum_{h \in H^m} p(h|s)g(s|h) = 0, \] (7)
so that, by (6), we get
\[ 0 = \sum_{h \in H} p(h|s)g(s|h) \leq \sum_{h \in H} p(h|s)g^*(s|h) \]
for every \( s \in S \), thus \( g^* \) does not satisfy condition (2) of Dutch book.

So, fix \( m = 1, \ldots, n \). Suppose by way of induction that for every \( k < m \) and \( h \in H^k \), \( g(s|h) = 0 \) for every \( s \in S^m \ldots \cup S^n \). This is vacuously true for \( m = 1 \). Note that, for each \( s \in S^m \), for each \( k > m \) and \( h \in H^k, s \notin S(h) \), therefore \( g(s|h) = 0 \) as well. Thus, \( g(s|h) = 0 \) for every \( h \notin H^m \). Then, from (6) we get
\[ \sum_{s \in S^m} \mu^m(s) \sum_{h \in H^m} p(h|s)g(s|h) \leq 0, \]
and with (5) we get
\[ \sum_{s \in S^m} \mu^m(s) \sum_{h \in H^m} p(h|s)g(s|h) = 0, \]
which can be rewritten as (7) because \( g(s|h) = 0 \) for every \( s \in S^m \) and \( h \notin H^m \).

There remains to prove the induction hypothesis for \( m + 1 \). From (7) and (4), we must have
\[ \sum_{s \in S} p(h|s)\mu^m(s)g(s|h) = 0 \]
for every \( h \in H^m \). Then the agent’s expected payoff from \( g(\cdot|h) \) is zero, therefore acceptance requires \( g(s|h) = 0 \) for every \( s \in S^{m+1} \cup \ldots S^n \).

3⇒1)

We prove the counterpositive. Fix a belief system \( \mu = (\mu(\cdot|h))_{h \in H} \) and suppose that it is not completely consistent. We are going to show that the agent can be Dutch-booked given \( \mu \).

We have already proven (1⇒2, counterpositively) that if a belief system is not completely consistent, there exist two different GORs between some \( s \) and
s'. Then, inverting the second GOR and multiplying by the first GOR\footnote{Since the two GORs are different, they cannot be both infinite or zero, therefore after inverting one of the two their product is not indeterminate.} we obtain a new GOR between \( s \) and itself, which is not 1; write it as

\[
\begin{align*}
    g(s|h^1) &= -1, \\
    g(s^1|h^1) &= -\frac{\mu(s|h^1)}{\mu(s^1|h^1)} + \varepsilon,
\end{align*}
\]

where each \( o(s, s^1|h) \) is a DOR between \( s' \) and \( s \) at \( h \) between \( s_{-i} \) and itself which is not 1. We can suppose without loss of generality that \( r < 1 \): if a GOR is larger than 1, inverting all the odds ratios we obtain a GOR smaller than 1 (still between \( s \) and itself). With this, we construct a Dutch Book that the agent is willing to accept. For every \( s \in S \) is larger than 1, inverting one of the two their product is not indeterminate. Finally, we need to show that, for some \( \varepsilon > 0 \),\footnote{Since the two GORs are different, they cannot be both infinite or zero, therefore after inverting one of the two their product is not indeterminate.} we also have

\[
\begin{align*}
    g(s^m|\tilde{s}^m) &= 0 \quad \text{for every } \tilde{s} \in S \setminus \{ s^1, \ldots, s^n \} \text{ trivially satisfies } (\ref{eq:1}).
\end{align*}
\]
is a Dutch book. Note that, for $\varepsilon = 0$,
\[
\sum_{h \in H} p(h|s)g(s|h) = p(h^1|s)g(s|h^1) + p(h^n|s)g(s|h^n) = -p(h^1|s) - p(h^n|s)g(s^{n-1}|h^{n-1})\frac{\mu(s^{n-1}|h^n)}{\mu(s|h^n)}
\]
\[
= -p(h^1|s) + p(h^{n-1}|s^{n-1})g(s^{n-1}|h^{n-1})\frac{p(h^n|s)}{p(h^n|s^{n-1})} \cdot \frac{\mu(s^{n-1}|h^{n-1})}{\mu(s|h^{n-1})} \cdot o(s^{n-1}, s|h)
\]
\[
= -p(h^1|s) + p(h^{n-2}|s^{n-2})g(s^{n-2}|h^{n-2})\frac{p(h^{n-1}|s^{n-1})}{p(h^{n-1}|s^{n-2})} \cdot \frac{\mu(s^{n-2}|h^{n-1})}{\mu(s^{n-2}|h^{n-2})} \cdot o(s^{n-2}, s^{n-1}|h^{n-1}) \cdot o(s^{n-1}, s|h)
\]
\[
= \ldots
\]
\[
= -p(h^1|s) - p(h^1|s)\frac{\mu(s|h^1)}{\mu(s|h^1)} \cdot o(s^1, s^2|h^2) \cdot \ldots \cdot o(s^{n-2}, s^{n-1}|h^{n-1}) \cdot o(s^{n-1}, s|h)
\]
\[
= -p(h^1|s) + p(h^1|s)r < 0.
\]

By continuity, the strict inequality is preserved for sufficiently small $\varepsilon > 0$. $\blacksquare$

6 Conclusion

Dutch-book arguments have been used as a benchmark for the rationality of beliefs in different circumstances, but they have never been used to check the coherence of beliefs across mutually exclusive contingencies. The reason is intuitive: if the bets of a Dutch book are spread across counterfactual contingencies, some bets will not be proposed, therefore the bookmaker will not realize the aggregate win that the sum of all bets assigns to each state. Yet, while the uncertainty about the state that is subject of betting cannot be quantified with objective probabilities, the additional factors that determine one contingency or the other may follow objective statistics. Therefore, it is possible to calculate the objective expected payoff from the Dutch book under each state. If this expected payoff is positive for the bookmaker under every state, I say that the agent is Dutch-booked (in objective expected terms). A Dutch-booked agent, or a population of like-minded agents subject to identical experiments, may experience a sure loss in the long run, or in aggregate. More generally, Dutch-bookable populations may suffer an evolutionary disadvantage because of incoherent decision-making.

From this novel Dutch-book argument, I derive a rule of coherence among beliefs at counterfactual contingencies. To avoid Dutch-booking, the way the subjective odds ratios between states change across contingencies must reflect
the objective probabilities of the contingencies given each state. While having a prior and doing Bayesian updating may seem sufficient, at some point the prior may be falsified in mutually exclusive ways, and then Bayesian updating as commonly intended does not impose any coherence among the revised beliefs after the counterfactual discoveries. I show that coherence is re-established when every belief can be derived, by Bayesian updating, from the same Lexicographic Conditional Probability System over the states. Instead, to avoid deterministic (diachronic) Dutch-booking, an agent can simply update her beliefs by conditioning, even when conditioning does not correspond to a correct use of Bayes rule. This shows that the diachronic Dutch book arguments for conditioning (i.e., the Dutch book theorem of Teller, 1973, and Lewis, 1999) are robust to the presence of unmodeled factors that are not in the state space but influence the agent’s observations.

References

[1] Battigalli, P., Catonini, E., Manili, J. (2021): “Belief change, rationality, and strategic reasoning in sequential games,” Iigier Working Paper n. 679, Universita Bocconi.

[2] de Finetti, B. (1937): “Foresight. Its Logical Laws, Its Subjective Sources,” report in Studies in Subjective Probability, Huntington: Robert E. Krieger.

[3] Blume, L., A. Brandenburger, and E. Dekel (1991): “Lexicographic Probabilities and Choice Under Uncertainty,” Econometrica, 59, 61-79.

[4] Kemeny, J. (1955): “Fair Bets and Inductive Probabilities,” Journal of Symbolic Logic, 20 (3): 263–273.

[5] Lehman, R. (1955): “On Confirmation and Rational Betting,” Journal of Symbolic Logic, 20 (3): 251–262.

[6] Lewis, D. (1999): “Why Conditionalize?” in Papers in Metaphysics and Epistemology, Cambridge: Cambridge University Press.

[7] Myerson, R. 1986. “Multistage Games with Communication.” Econometrica 54 (2): 323-358.

[8] Ramsey, F. P. (1931): “Truth and Probability,” in The Foundations of Mathematics and Other Logical Essays, London: Routledge and Kegan.

[9] Renyi, A. 1955. “On a New Axiomatic Theory of Probability.” Acta Mathematica Academiae Scientiarum Hungaricae 6: 285-335.

[10] Rescorla, M. (2021): “An improved Dutch book theorem for conditioning,” Erkenntnis, forthcoming.
[11] Siniscalchi, M. (2020): “Structural Rationality in Dynamic Games,” working paper.

[12] Teller, P. (1973): “Conditionalization and Observation,” Synthese, 26: 218–258.