Sequential $\boldsymbol{\varepsilon}$-Contamination with Bayesian Updating and Its Application to Job Search*

Hiroyuki Kato†
Department of Management and Economics, Kaetsu University
Kiyohiko G. Nishimura
National Graduate Institute for Policy Studies (GRIPS)
and CARF, University of Tokyo
and
Hiroyuki Ozaki
Faculty of Economics, Keio University

October 10, 2021

Abstract

The ε-contamination has been studied extensively as a convenient and operational specification of Knightian uncertainty in many economic problems including job search. However, it is formulated in a static environment. When it is applied straightforwardly in a dynamic and sequential framework with Bayesian updating, the Principle of Optimality of dynamic programming may not hold true. We then develop the theory of sequential ε-contamination guaranteeing the applicability of the Principle, which can be represented by a sequence of monotonically increasing ε's that "contaminate" the conditional principal belief, or in mathematical terms, probability charge. The sequential ε-contamination is shown to be a natural extension of static ε-contamination as interpreted in the dynamic and sequential framework, since both have the same initial period "view" (ε) of the world. The sequential ε-contamination is then applied to job search models and it is shown that job searchers become more likely to accept the current job offer in the sequential framework with Bayesian updating than otherwise, which is consistent with empirical evidence.

JEL codes: C61, D81, D83

Keywords: Knightian uncertainty, One-shot ε-contamination, Rectangularity, Sequential $\boldsymbol{\varepsilon}$-contamination, Learning, job search

---

*We appreciate very helpful comments given by an anonymous researcher on the previous version of this paper, as well as the participants of a workshop at Tohoku University and Decision Theory Workshop at Hitotsubashi University. Financial supports to Nishimura from JSPS KAKENHI(S) Grant Number 18H05217 and those to Ozaki from JSPS KAKENHI(C) Grant Number 19K01550 are gratefully acknowledged.

†Corresponding Author.

1 Introduction

The ε-contamination has been studied extensively as a convenient and operational specification of Knightian uncertainty. In the ε-contamination framework, the decision-maker is assumed to be $(1-\varepsilon)\times 100\%$-certain that she faces a particular probability charge, which may be called a "principal" probability charge, but with $\varepsilon \times 100\%$-fear she feels completely ignorant so that she may think she faces the worst case.

This concept is applied to analyze the effect of Knightian uncertainty on the economic agent's behavior such as search (Nishimura and Ozaki, 2004), asset pricing (Epstein and Wang, 1994), voting (Chu and Liu, 2002) and learning (Nishimura and Ozaki, 2017, Chapter 14).[1] Also, it has a simple and intuitive axiomatic foundation (Nishimura and Ozaki, 2006; Kopylov, 2009). The ε-contamination also comes up in the statistics literature on robustness (for example, see Berger, 1985).

The ε-contamination described above, however, is formulated in a static environment. Since many economic problems are dynamic and sequential in nature, we need a dynamic and sequential version of the ε-contamination. However, when it is applied straightforwardly in a dynamic and sequential framework, the Principle of Optimality of dynamic programming may not hold true as demonstrated in an example in Section 2. Thus, we need a more sophisticated version of ε-contamination in a sequential setting, which guarantees applicability of the Principle of Optimality and at the same time has the same initial period "view" (ε) of the world as the static ε-contamination. It is well-known (Epstein and Schneider, 2003) that Knightian uncertainty should exhibit the rectangularity property in order to guarantee time consistency of intertemporal decision-making in the maxmin expected utility model under Knightian uncertainty (Gilboa and Schmeidler, 1989). Since time consistency is necessary for the Principle of Optimality, we formulate a "time-consistent" (or rectangular) dynamic ε-contamination, which we call sequential $\boldsymbol{\varepsilon}$-contamination.[2] In fact, the Principle of Optimality is restored under sequential $\boldsymbol{\varepsilon}$-contamination in the above-mentioned example of Section 2.

There is another issue in the dynamic formulation: *updating*, or equivalently, *conditioning*. The issue can be explained by the analogy to updating under traditional framework of no Knightian uncertainty. Suppose that the

[1] The continuous-time counterpart of the ε-contamination, developed by Chen and Epstein (2002) and known as the κ*-ignorance*, is applied to continuous-time dynamic models with ambiguity, including Chen and Epstein (2002) and Liu (2011) (portfolio choice), and Nishimura and Ozaki (2007), Miao and Wang (2011), Thijssen (2011) and Flor and Hesel (2015) (investment choice) among voluminous literature.

[2] Please note the difference between the boldface "$\boldsymbol{\varepsilon}$" for the sequential $\boldsymbol{\varepsilon}$-contamination and the non-boldface "ε" used everywhere else. This is because $\boldsymbol{\varepsilon}$ appearing in the sequential $\boldsymbol{\varepsilon}$-contamination turns out to be a sequence of (possibly mutually distinct) numbers (*i.e.,* ε's), instead of a mere number.

decision-maker faces a particular probability charge and that one period has elapsed. Then, it is quite natural that the decision-maker likes to update the probability charge by conditioning it on the observation she made in the past period according to Bayes' rule. The situation is the same in the case of Knightian uncertainty, in which the decision-maker faces a set of probability charges. After one period has elapsed, she may want to update all the probability charges in the set by conditioning each of them on the observation she made in the past period, to be left with a new set of updated probability charges which forms new Knightian uncertainty. In this paper, we take this updating procedure which is well-known updating rule for Knightian uncertainty and is often called "generalized Bayes' rule." Interestingly and importantly, we later prove that the sequential $\boldsymbol{\varepsilon}$-contamination updated by the generalized Bayes' rule is again the sequential $\boldsymbol{\varepsilon}'$-contamination with $\boldsymbol{\varepsilon}'$ possibly different from $\boldsymbol{\varepsilon}$.[3]

The results which are proved in the main body of this paper are remarkable. The sequential $\boldsymbol{\varepsilon}$-contamination which is equipped with both generalized Bayesian updating and intertemporal consistency at the same time can be represented by a monotonically increasing sequence of ε's that "contaminate" the conditional "principal" probability charge. Furthermore, each of these subsequent periods' ε's has a simple closed form that is dependent on the initial ε. The degree of contamination, ε, increases in each period after updating, because (1) updating of the set of probability charges in question by generalized Bayes' rule may cause its dilation (that is, an increase in Knightian uncertainty)[4] which in turn may cause time inconsistency of sequential decision if the same ε is kept to be assumed, and thus (2) to take account of this possibility of time inconsistency and to avoid it, the decision-maker may change her view of Knighitian uncertainty and postulate larger uncertainty, that is, she increases her perceived degree of contamination, ε.

Then, in a practical application to job search, we will find two implications of the sequential $\boldsymbol{\varepsilon}$-contamination with Bayesian updating. First, in the multi-period Bayesian updating framework, an increase in the initial Knightian uncertainty decreases the value of continuing search in the subsequent periods monotonically. Thus, the worker is increasingly more likely to accept the current job offer, implying shorter search periods. (See Proposition 11 in Section 4 for the two-period model and Proposition 18 in

[3]There may be another approach, in which we might impose axioms on the primal and all subsequent conditional preferences to derive all of each period's ε's. In an important paper, Kopylov (2016) adopts this approach based on his endogenous characterization of ε in Kopylov (2009). In his approach, however, updating is rather abstract and we do not know the relationship between ε before and after observation (*i.e.,* the one between ε and ε'). Thus, we do not take this approach and instead we start with the single ε in the initial period since we would like to examine the implications of generalized Bayesian updating *explicitly* to the subsequent period's ε's.

[4]See Nishimura and Ozaki (2017, Chapter14).

Appendix B for the multi-period model.)

Second, the presence of multi-period Knightian uncertainty itself is likely to decrease reservation wages over time under Bayesian updating and thus shorten an unemployment spell. (See Proposition 19 of Appendix B) This implication is consistent with empirical evidence which states that the reservation wage declines over the course of an unemployment spell. (See for exmple, Brown, Flinn and Schotter, 2011.)[5]

The organization of the paper is as follows. Section 2 presents a motivating example of a job search model in which possible shortcomings of the traditional static, or one-shot ε-contamination are revealed when it is applied to multi-stage models, which lays the foundation of the sequential $\boldsymbol{\varepsilon}$-contamination. In Section 3, we define formally sequential $\boldsymbol{\varepsilon}$-contamination and show that it is rectangular so as to guarantee time consistency in a two period framework. It also presents basic properties of sequential $\boldsymbol{\varepsilon}$-contamination. Section 4 applies sequential $\boldsymbol{\varepsilon}$-contamination to the job search model introduced in the motivating example of Section 2, and explores its implications in the framework of Bayesian updating. All proofs of propositions are relegated to Appendix A.

Appendix B contains the extension of the two-period models to T-period finite-horizon models with arbitrary T, showing that the same mathematical results as well as the same implications for the job search behavior are also obtained in longer horizon models.

2 A Motivating Example in Job Search: Principle of Optimality under Static ε- and Sequential $\boldsymbol{\varepsilon}$-contamination

This section presents a two-period simple job search model in the presence of uncertainty. In the way of incorporating uncertainty into the model, we adopt mutually distinct two different methods. One is to characterize uncertainty by the *ε-contamination in the initial period* (the period 0), where we interpret the single ε as representing the degree of agent's ignorance about the whole history of the states over the two periods of the world and we assume that ε is given exogenously.[6] This way of specifying ambiguity may be called *static* or *one-shot ε-contamination*. The other method is the one which characterizes uncertainty by the $\boldsymbol{\varepsilon}$-contamination which invokes two (possibly) distinct ε and ε'.

We briefly touched the rough ideas of both in the Introduction, but we states these two kinds of specification of uncertainty more carefully with

[5]See p.948 of their paper, and the literature cited in their footnote 1.

[6]This type of recognition by the agent is axiomatized by Nishimura and Ozaki (2006) and Kopylov (2009).

the ε-contamination in Subsection 2.3 and the sequential $\boldsymbol{\varepsilon}$-contamination in Subsection 2.4.

Furthermore, for each specification of uncertainty, we consider two types of uncertainty-averse agents: the *one-shot minimizer* and the *multi-stage minimizer*.

The first type of the agent, the one-shot minimizer, is supposed to find the single minimum among all the possible expectations of the consequences resultant from her all possible contingent action streams, where the expectations are computed in the period 0, and then she chooses the contingent action stream that attains that minimum. All these are done *once and for all* in the period 0.[7]

An important *restriction* imposed on the one-shot minimizer is that while she may choose any probability charge contained in the time-0 uncertainty (it may be the static ε-contamination or the sequential $\boldsymbol{\varepsilon}$-contamination)[8], she would have to stick to the *identical* probability charge when she needed to invoke the conditional and marginal probability charges during the computational process of reaching the time-0 expectation. In another word, she are allowed to employ only the single probability charge contained in the given time-0 uncertainty, when its unconditional or conditional or marginal form is necessary . It is quite significant to note this difference between the one-shot minimizer and the multi-state minimizer, to the latter of which we will now turn.

The second type of agents, the multi-stage minimizer, is supposed to compute the minimum of the expectations *at each decision node of the period 1* of the consequences resultant from her actions available at that node. She does this after she has observed the period 1's state and by choosing the conditional probability charge as far as the conditional should be based on some probability charge contained in the time-0 uncertainty (it may be the static ε-contamination or the sequential $\boldsymbol{\varepsilon}$-contamination).[9]

Then, after all the minimal "conditional" expectations are thus computed in the period 1, the multi-stage minimizer aggregates these minimal expectations into the overall unconditional minimal expectation *in the period 0* by means of the "marginal" probability charge. In contrast to the one-shot minimizer, the marginal used here may come from any probability charge as far as it is contained in the time-0 uncertainty.

As is apparent from the above procedure, the multi-stage minimizer conducts the act of taking the minimum "backwardly" firstly in the period 1 and secondly in the period 0. Furthermore, during conducting this proce-

[7]The behavior of this type is consistent with the one of the agent axiomatized by Nishimura and Ozaki (2006) and Koylov (2009). See the Footnote 6.

[8]Recall that we are defining the one-shot minimizer for both of two specifications of uncertainty.

[9]Recall that we are defining the multi-shot minimizer for both of two specifications of uncertainty. See the footnote 8.

dure, the conditionals and marginals can be based on the distinct probability charges as far as all of them are contained in the time-0 uncertainty.

In the rest of this subsection, we basically show the following observations:

(1) Under the specification of uncertainty by the static ε-contamination, the behavior of the one-shot minimizer and that of the multi-stage minimizer lead to different consequences in the model.

(2) Under the specification of uncertainty by the sequential $\boldsymbol{\varepsilon}$-contamination, the behavior of the one-shot minimizer and that of the multi-stage minimizer lead to the completely identical consequence in the model.

Before showing these facts, the next subsection formally describes the job search model we consider.

2.1 Description of Simple Job Search Model

We assume that there are two periods and both periods are represented by the identical state space, S, which is composed of the two states, *i.e.*, $S := \{b, s\}$, where we interpret the state b as representing the "boom" and the state s as representing the "slump." Therefore, there are totally four states in this model: (b, b), (b, s), (s, b) and (s, s), where the first of each pair is the state in the first period (period 1) and the second of it is the state in the second period (period 2).

The economic agent of the model is a worker and she is unemployed in the beginning (period 0). This unemployed worker receives the offer of the wage in each period (period 1 and period 2) and the offer in each period is denoted by w_t $(t = 1, 2)$. The wage offer is stochastic and it depends solely on the state when it is offered. The wage offer at a boom is w_b and that at a slump is w_s. That is, for each $t \in \{1, 2\}$, $w_t = w_b$ if the state in period t is b, and $w_t = w_s$ if the state in period t is s. We assume that $w_b > w_s$, which is quite natural to assume given our interpretation of each state.[10]

If the worker accepts the wage offer which is made in the first period, she earns that wage in the first period and she will earn the same wage also in the second period. She is subject to the contract she made in the first period and cannot renege it even if the second-period wage offer is better. On the other hand, if she declines the offer in the first period, she receives unemployed compensation, $c > 0$, in this period and then continues search to face a new wage offer randomly drawn in the second period, which she can either accept or decline again to receive c.

The worker's lifetime income is the discounted sum of her wage/unemployed compensation earned over two periods, where she "discounts" the second pe-

[10] All the assumptions on the parameters of the model, including this one, is summarized in Subsection 2.1.2.

riod income by the discount factor $\beta > 0$.[11] To make the model meaningful we need to and actually do assume that $c < w_b$ because, if otherwise, the worker would always choose to decline the offer to maximize her lifetime income as $c + \beta c$.

As an additional ingredient of our model, we introduce the *grant-in-aid* paid to the worker by the government only in the case when the economy is in the slump. It is a fixed money amount, $\bar{w} > 0$, given to the worker, regardless of whether she is employed or unemployed, without any charge whenever the first-period's state is s. Importantly, the aid is paid only in the first period and will last only for one period. The worker won't get this aid in the second period even if the second-period's state is s. In this regard, the grant-in-aid in this example is quite similar to the one made by the Japanese government in 2020. The government paid 100,000 yen to each Japanese citizen only *once* with the intention of supporting the people whose everyday lives are suffering from COVID-19 while it is prevalent still now for more than one year.

Note that the main motivation of introducing the grant-in-aid is to demonstrate a stark contrast in the consequences resultant from the two different specifications of uncertainty (that is, (1) and (2) mentioned in the preamble of this section). The fundamental source of this discrepancy is the way of specifying uncertainty, not the presence of the grant-in-aid. In fact, we do not assume such an aid in the application of the sequential $\boldsymbol{\varepsilon}$-contamination to the job search model analyzed in the main body of this paper.

2.1.1 The Static ε-Contamination

As we already mentioned, the objective of this section is to illustrate that the two types of time-0 uncertainty will lead to the different implications in the job search model described above.

We now define the first type of time-0 uncertainty over the two periods: the static ε-contamination. We do this firstly because its second type, *i.e.*, the sequential $\boldsymbol{\varepsilon}$-contamination, is defined based on it, and secondly because we like to summarize all the assumptions imposed on the parameters that appear in the example before we start its formal analyses.

Let p^0 be some probability charge[12] on $S \times S$, where $S = \{b, s\}$ as defined above. We call p^0 the *principal* probability charge. Also, let ε be a real number such that $\varepsilon \in (0, 1)$. Then, the *static ε-contamination of* p^0, denoted $\left\{p^0\right\}^\varepsilon$, is defined by

$$\left\{p^0\right\}^\varepsilon := \left\{ (1-\varepsilon)p^0 + \varepsilon q \,\middle|\, q \text{ is any probability charge on } S \times S \right\} . \quad (1)$$

[11] In this finite-horizon model, we need not assume that $\beta < 1$. Thus, the worker may "upcount" the future.

[12] For the precise definition of the probability charge, see Nishimura and Ozaki (2017).

We can interpret the static ε-contamination of p^0 is a form of uncertainty the decision-maker faces at time 0 such that she believe that the risk (probability charge) for the next successive periods will be governed by some specific p^0 with $(1-\varepsilon) \times 100\%$-conviction, and such that if her conviction will turns out to be wrong, any probability charge, which might be wildly different from p^0, could be the case.

The maxmin preference à la Gilboa and Schmeidler (1989) with the set of multi-priors characterized by the the static ε-contamination of p^0, $\left\{p^0\right\}^\varepsilon$, was axiomatized by Nishimura and Ozaki (2006) and Kopylov (2009).

2.1.2 The Assumptions on the Parameters Appearing in the Example

We are now ready to make a whole list of assumptions imposed on the parameters that appear in the job search model whose analyses are to be conducted in this section.

The first group of parametric assumptions are those on the parameters introduced in the preamble of Subsection 2.1:

$$c < w_s < w_b < c + \bar{w} \quad \text{and} \quad w_s + \bar{w} + \beta w_s < w_b + \beta w_b \,. \tag{2}$$

While some are reasonable to assume and we already mention them, some are new.

The second group of parametric assumptions are those on the principal probability charge, p^0:

$$p^0(\{b,b\}) = p^0(\{b,s\}) = p^0(\{s,b\}) = p^0(\{s,s\}) = \frac{1}{4}\,. \tag{3}$$

This is a direct and mainly simplifying assumption without which the computations that follow would be quite messy.

The final group of parametric assumptions are those jointly assumed on the parameters in the first group and the parameter $\varepsilon \in (0,1)$:

$$\begin{aligned}&(1-3\varepsilon)(w_b + \beta w_b) + (1+\varepsilon)(w_s + \bar{w} + \beta w_s) > 0 \\ &\qquad \text{and} \quad w_b + \beta w_b < c + \bar{w} + \beta\left(\frac{1-\varepsilon}{2(1+\varepsilon)} w_b + \frac{1+3\varepsilon}{2(1+\varepsilon)} w_s\right)\,. \end{aligned} \tag{4}$$

Compared with the first two groups of assumptions, this is clearly more complicated. However, thanks to this, we can show the stark contrast between the implications derived from two different ways of specifying the time-0 uncertainty, which we will see in the next two subsections.

There seems to be two mutually distinct approaches to disentangle the complexity embodied in the third group of assumptions. The first approach is to decompose the two inequalities into the more simple ones. But here, we choose the much easier second approach, which simply listed up each of the parameters concretely, as a whole of which both assumptions, (2) and (4), are met.

For instance, setting:

$$c = 0.9\,;\; w_s = 1\,;\; w_b = 1.6\,;\; \bar{w} = 1.1\,;\; \beta = 1\,; \quad \text{with} \;\; \varepsilon < 0.2 \tag{5}$$

meets all the requirements.[13]

In the rest of this section, we maintain the parametric assumptions, (2), (3) and (4).

2.2 The Case of Static ε-contamination: The Principle of Optimality Is Violated

In this subsection, we solve the job search model where the time-0 uncertainty is specified by the static ε-contamination, firstly by the one-shot minimization and secondly by the multi-stage minimization, in turn.

We relegate the formal derivation of the intermediate steps for the claims made in the current and subsequent subsections to the Appendix A.

If we denote the worker's income for each period by y_t $(t = 1, 2)$, the objective of the unemployed worker is to maximize her lifetime income defined by

$$\min_{p \in \{p^0\}^\varepsilon} E^p \left[y_1 + \beta y_2 \right] \tag{6}$$

by choosing her strategy whose nature depends on whether she is the one-shot minimizer or the multi-stage minimizer. Note that the uncertainty the worker faces is characterized by the static ε-contamination in the objective function, (6), where $\left\{p^0\right\}^\varepsilon$ is defined by (1). The "min" operator in the forefront of (6) reflects the agent's uncertainty-aversion.

2.2.1 The One-Shot Minimization

By definition, the one-shot minimizer employs the plans that are contingent upon the whole history of the states over the two periods. That is, her strategy depends upon both period 1 and period 2's states together.

Let us first consider such strategies the agent may choose. There are four possible histories of states, and there are three possible actions: "stop," "continue and then stop," and "continue and then continue again," for each of the four histories. We thus conclude that there is a total of twelve possible (contingent) strategies available to the agent.

However, some simple reasoning eliminates some irrelevant strategies to a large extent. First, the first half of the assumption (2) kills the third action mentioned above. Second, if the state b is realized in the first period, the best

[13]Because the assumptions, (2) and (4), are stated in the form of the inequalities, the list of parameters, (5), is "robust" in the sense that the assumptions keep to be satisfied if we slightly shake the values listed in (5).

action for the agent is to stop immediately by the first half of the assumption (2) again, because the grant-in-aid will be never paid in the second period by definition. In sum, we are left with only the two (contingent) strategies worth considering: $\{b \to \text{stop} \; ; \; s \to \text{stop}\}$ and $\{b \to \text{stop} \; ; \; s \to \text{continue}\}$, in which the state denotes the one in the first period. (Note that the state realized in the second period does not matter at all because the agent had already accepted the offer in the first period or if otherwise, she definitely accepts the second period's offer regardless of the realized state then.)

Our next task is to compare the agent's minimal expected lifetime incomes, (6), each of which is computed by each of the two strategies mentioned above.

$\{b \to \text{stop} \; ; \; s \to \text{stop}\}$. If the state in the first period were b, her lifetime income would be

$$w_b + \beta w_b \, . \tag{7}$$

On the other hand, if the state in the first period were s, her lifetime income would be

$$w_s + \bar{w} + \beta w_s \, . \tag{8}$$

Because (7) $>$ (8) by the second half of the assumption (2), her minimal expected lifetime income, (6), turns out to be

$$\frac{1-\varepsilon}{2}\left(w_b + \beta w_b\right) + \frac{1+\varepsilon}{2}\left(w_s + \bar{w} + \beta w_s\right) \, , \tag{9}$$

where the "worst" weight on (7) and the "best" weight on (8), which should be chosen from $\{p^0\}^\varepsilon$ so as to minimize the expected lifetime income, are found by setting, say, $q := (0, 0, 1/2, 1/2)$ in (1).[14]

$\{b \to \text{stop} \; ; \; s \to \text{continue}\}$. If the state in the first period were b, her lifetime income would be determined by (7) as above. If otherwise, it depends also on the state which will be realized in the second period. While her lifetime income will be $c + \bar{w} + \beta w_b$ when the second-period's state is b, it will be $c + \bar{w} + \beta w_s$ when the second-period's state is s. Thus, her minimal expected lifetime income turns out to be the weighted sum of three levels of the lifetime incomes:

$$\begin{aligned} & \frac{1-\varepsilon}{2}\left(w_b + \beta w_b\right) + \frac{1-\varepsilon}{4}\left(c + \bar{w} + \beta w_b\right) + \frac{1+3\varepsilon}{4}\left(c + \bar{w} + \beta w_s\right) \\ = \; & \frac{1-\varepsilon}{2}\left(w_b + \beta w_b\right) + \frac{1+\varepsilon}{2}\left(c + \bar{w}\right) + \beta\left(\frac{1-\varepsilon}{4} w_b + \frac{1+3\varepsilon}{4} w_s\right) . \end{aligned} \tag{10}$$

[14] The first co-ordinate of q, *i.e.*, 0, denotes the wight on the history, (b, b), because the wight by q has no restriction except that the sum of weights must be unity. See (1). The co-ordinate of q is listed by the same order as that of the histories indicated in (3). The same convention applies also elsewhere.

Here, in order to minimize the expected lifetime income, the "worst" weight in $\{p^0\}^\varepsilon$, $(1-\varepsilon)/4$, should be put on the best lifetime income; the "best" weight in $\{p^0\}^\varepsilon$, $(1+3\varepsilon)/4$, should be put on the worst lifetime income; and the remaining weight, $(1-\varepsilon)/2$, which is determine by the fact that the weights sum up to unity, should be put on the middle lifetime income, where the rank of the lifetime incomes follows from the second half of the assumption (2). The weights above are corresponding to setting $q := (0, 0, 0, 1)$ in (1).

Finally, the agent decides which strategy to choose by comparing (9) with (10), and it turns out to be (9) < (10) holds. (See A.1 in the Appendix A.) Thus, the one-shot minimizing worker chooses the second strategy described by $\{b \to \text{stop} \ ; \ s \to \text{continue}\}$.

We call the maximized minimal (that is, maxmin) expected lifetime income for the one-shot-minimizing worker, where the maximum is achieved by her appropriate state-contingent strategy, as the *value* of the job search and denote it by V_0. We state the result obtained so far in the form of the proposition below:

Proposition 1 (Static ε-Cont. + One-Shot Min.) *The value of the job search, V_0, is given by* (10):

$$V_0 = \frac{1-\varepsilon}{2}\left(w_b + \beta w_b\right) + \frac{1+\varepsilon}{2}\left(c + \bar{w}\right) + \beta\left(\frac{1-\varepsilon}{4}w_b + \frac{1+3\varepsilon}{4}w_s\right),$$

which is attained by the worker's choosing "to stop" if the first period's state is b and by choosing "to continue" if it is s (and then choosing "to stop" in the second period).

2.2.2 The Multi-Stage Minimization

This subsection illustrates how to solve the job search model if the worker is a multi-stage minimizer introduced in the preamble of this section.

According to the definition of the multi-stage minimization, we assume both the agent's belief-updating by Bayes' rule and her stage-by-stage minimization, where the second-stage minimization is conducted among her conditional beliefs.

Assuming this, we will solve the problem in a backward induction method. The worker can thus select the "worst" conditionals independently at each decision node, each of which is derived from *some* (not necessarily the same) probability charge as far as such a charge is contained in the given static ε-contamination of p^0, $\{p^0\}^\varepsilon$.

The worker of this type likes to maximize:

$$\min_{\alpha}\left[\alpha \max\left\{w_b + \beta w_b\, ,\, c + \beta \min_{\alpha'}\left[\alpha' w_b + (1-\alpha')w_s\right]\right\}\right.$$

$$+ (1-\alpha)\max\left\{w_s + \bar{w} + \beta w_s\,,\ c + \bar{w} + \beta \min_{\alpha''}\left[\alpha'' w_b + (1-\alpha'')w_s\right]\right\}\Bigg]. \tag{11}$$

Here, each term inside the first braces denotes the agent's "pessimistic" lifetime income when she chooses to "stop" or to "continue," respectively, after knowing that the first-period's state is b, and each term inside the second braces denotes the agent's "pessimistic" lifetime income when she does so after knowing that the first-period's state is s.[15] The "pessimism" is represented by the "min" operators there which are indicating the minimum second-period (current) incomes computed by means of the conditional beliefs.

In (11), we assume that, quite differently from the one-shot minimizer in the previous subsection, α (the marginal probability charge for the first period), and α' and α'' (the conditional probability charges for the second period when the first-period's state is b or s, respectively) may be based upon mutually distinct three probability charges as far as all of three charges are contained in the given static ε-contamination of p^0, $\left\{p^0\right\}^\varepsilon$.

Finally, each "max" operator in (11) indicates the maximum lifetime income computed in the first period given the first-period's state, b or s, and the "min" operator in the head of (11) indicates the overall lifetime income computed by the agent "pessimistically" before the first period begins.

We now tackle with the problem by backward induction. Suppose that the worker is in the end of the first period just after she observed the state b. Because $w_b > w_s$ by (2), the "worst" value of α' turns out to be $(1-\varepsilon)/2(1+\varepsilon)$, which is derived by setting $q = q' := (0,1,0,0)$ in (1).[16]

Similarly, suppose that the worker is in the end of the first period just after she observed the state s. Then, by the same logic as above, the "worst" value of α'' turns out to be $(1-\varepsilon)/2(1+\varepsilon)$, which is derived by setting $q = q'' := (0,0,0,1)$ in (1). (See the Footnote 16.)

By substituting these values of α' and α'', (11) is now reduced to

$$\min_{\alpha}\Bigg[\alpha\max\left\{w_b + \beta w_b\,,\ c + \beta\left(\frac{1-\varepsilon}{2(1+\varepsilon)}w_b + \frac{1+3\varepsilon}{2(1+\varepsilon)}w_s\right)\right\}$$

$$+(1-\alpha)\max\left\{w_s + \bar{w} + \beta w_s\,,\ c + \bar{w} + \beta\left(\frac{1-\varepsilon}{2(1+\varepsilon)}w_b + \frac{1+3\varepsilon}{2(1+\varepsilon)}w_s\right)\right\}\Bigg]. \tag{12}$$

[15] Recall that the agent always stops in the second period by the maintained assumption, (2).

[16] In order to find q', seek for $p = (p_1, p_2, p_3, p_4) \in \left\{p^0\right\}^\varepsilon$ such that $p_1/(p_1+p_2)$ (the conditional charge of b given b in the first period) will be minimized, where $p_1 := p(\{b,b\})$, and so on, according to the order of the state's histories listed in (3). The minimum must be achieved by some $p \in \left\{p^0\right\}^\varepsilon$ and such a minimizing p turns out to be $p = ((1-\varepsilon)/4, (1-\varepsilon)/4+\varepsilon, (1-\varepsilon)/4, (1-\varepsilon)/4)$, which corresponds to q defined by $q' := (0,1,0,0)$, and the weight on w_b appearing in the first line of (12) must be $(1-\varepsilon)/2(1+\varepsilon)$. The similar reasoning applies to the derivation of q'' in the next paragraph.

By solving the two maximization problems contained in (12) under the maintained assumptions, we can show that the formula (12) is equal to

$$\min_{\alpha}\Big[\alpha\left(w_b+\beta w_b\right)\ ,\qquad (1-\alpha)\left(c+\bar{w}+\beta\left(\frac{1-\varepsilon}{2(1+\varepsilon)}w_b+\frac{1+3\varepsilon}{2(1+\varepsilon)}w_s\right)\right)\Big] \tag{13}$$

(see A.2 in the Appendix A).

Finally, by the second inequality of the assumption (4), we know that the minimum in (13) can be achieved when α is maximized. If we recall that α is the first-period marginal of any probability charge contained in $\{p^0\}^{\varepsilon}$, such an α turns out to be $(1+\varepsilon)/2$, which is derived by setting, say, $q=(1/2,1/2,0,0)$ (and thus $1-\alpha$ should be $(1-\varepsilon)/2$, which is derived by setting, say, $q=(0,0,1/2,1/2)$).

If we denote by V_0' the value of the job search for the multi-stage minimizer, what we have established so far is summarized by the next proposition.

Proposition 2 (Static ε-Cont. + Multi-Stage Min.) *The value of the job search, V_0', is given by:*

$$\begin{aligned}
V_0' &= \frac{1+\varepsilon}{2}\left(w_b+\beta w_b\right)+\frac{1-\varepsilon}{2}\left[c+\bar{w}+\beta\left(\frac{1-\varepsilon}{2(1+\varepsilon)}w_b+\frac{1+3\varepsilon}{2(1+\varepsilon)}w_s\right)\right]\\
&= \frac{1+\varepsilon}{2}\left(w_b+\beta w_b\right)\\
&\qquad +\frac{1-\varepsilon}{2}\left(c+\bar{w}\right)+\beta\left(\frac{(1-\varepsilon)^2}{4(1+\varepsilon)}w_b+\frac{(1-\varepsilon)(1+3\varepsilon)}{4(1+\varepsilon)}w_s\right),
\end{aligned} \tag{14}$$

which is attained by the worker's choosing "to stop" if the first period's state is b and by choosing "to continue" if it is s (and then choosing "to stop" in the second period).

Comparing Propositions 1 and 2 shows that, when the time-0 uncertainty is specified by the static ε-contamination, the value of the job search varies depending on whether the worker takes the one-shot minimization or the multi-stage minimization, while the worker's best strategy is the same between the two types of her behaviors. Thus, the principle of optimality is violated for the two-period job search model with the static ε-contamination considered in this subsection.

This happens because of two reasons: (1) there exists uncertainty, that is, $\varepsilon>0$; (2) the static ε-contamination is *not* the way of specifying the time-0 uncertainty that guarantees the time-consistent behavior of the worker. This is one of our chief concerns which motivate us to write this paper, and we discuss this issue in more detail later in this section.

2.3 The Case of Sequential $\boldsymbol{\varepsilon}$-contamination: The Principle of Optimality Is Restored

In this subsection, we solve the job search model where the time-0 uncertainty is specified by the sequential $\boldsymbol{\varepsilon}$-contamination, firstly by the one-shot minimization and secondly by the multi-stage minimization, in turn, according to the manner we employed in the previous subsection.

First of all, however, we need to define the sequential $\boldsymbol{\varepsilon}$-contamination. Let $\varepsilon \in (0,1)$ be the one which is used in the definition of the static ε-contamination, (1). We then define a series of real numbers by:

$$\underline{\varepsilon} := -\frac{\varepsilon}{2}\ ; \quad \bar{\varepsilon} := \frac{\varepsilon}{2}\ ; \quad \underline{\varepsilon}' := -\frac{\varepsilon}{1+\varepsilon} \quad \text{and} \quad \bar{\varepsilon}' := \frac{\varepsilon}{1+\varepsilon}\,. \tag{15}$$

We use these "bounds" to define the *sequential $\boldsymbol{\varepsilon}$-contamination of* p^0, denoted $\left\{p^0\right\}^{seq\varepsilon}$, by

$$\left\{p^0\right\}^{seq\varepsilon} := \left\{ \left(\frac{1}{2}+\varepsilon_i\right)\cdot\left(\frac{1}{2}+\varepsilon'_{ij}\right)_{\substack{i=b,s\\ j=b,s}} \;\middle|\; \begin{array}{l} (\forall i=b,s)\ \varepsilon_i \in [\underline{\varepsilon},\bar{\varepsilon}];\ \ \varepsilon_b+\varepsilon_s=0; \\ (\forall i=b,s)(\forall j=b,s)\ \varepsilon'_{ij} \in [\underline{\varepsilon}',\bar{\varepsilon}'] \ \text{ and } \ (\forall i=b,s)\ \varepsilon'_{ib}+\varepsilon'_{is}=0 \end{array} \right\}. \tag{16}$$

Here, the dot ("$\cdot$") in the first line simply denotes the product of two numbers, and it should be understood as that of the first-period's "marginal" and the second-period's "conditional."

Also, for each i and j, ε_i and ε'_{ij} may vary within the range defined by the "bounds," (15), which makes $\left\{p^0\right\}^{seq\varepsilon}$ a set of probability charges on $S \times S$, rather than a single probability charge.

Note that the definition of the sequential $\boldsymbol{\varepsilon}$-contamination given by (16) is the one adapted for the example of this section. Presenting both its general definition and an intuition behind it will be deferred until the next section.

2.3.1 The One-Shot Minimization

This subsection illustrates how to solve the job search model if the worker is a one-shot minimizer when the time-0 uncertainty is specified by the sequential $\boldsymbol{\varepsilon}$-contamination introduced above.

By the same logic as Subsection 2.2.1, it suffices to consider only two strategies: $\{b \to \text{stop}\ ;\ s \to \text{stop}\}$ and $\{b \to \text{stop}\ ;\ s \to \text{continue}\}$.

$\{b \to \text{stop}\ ;\ s \to \text{stop}\}$. Under this strategy, the lifetime income would be $w_b + \beta w_b$ whenever the first-period's state were b; and it would be $w_s + \bar{w} + \beta w_s$ whenever it is s. In order to minimize the weighted sum of these two lifetime incomes by choosing the wights from among the sequential $\boldsymbol{\varepsilon}$-contamination, we need to find the minimal first-period marginals for b, and the maximal first-period's marginal for s, subject to the constraints imposed by the definition of the sequential $\boldsymbol{\varepsilon}$-contamination, because $w_b + \beta w_b > w_s + \bar{w} + \beta w_s$ by the second half of the assumption (2).

Such marginals can be found immediately and thus the worker's minimal expected lifetime income given this strategy turns out to be

$$\frac{1-\varepsilon}{2}\left(w_b + \beta w_b\right) + \frac{1+\varepsilon}{2}\left(w_s + \bar{w} + \beta w_s\right) . \tag{17}$$

$\{b \to \text{stop} \ ; \ s \to \text{continue}\}$. Given this strategy, the worker's lifetime income takes three possible values depending on the states' whole history: (i) $w_b + \beta w_b$ (taking place both at (b, b) and (b, s) and being the second largest among three); (ii) $c + \bar{w} + \beta w_b$ (taking place at (s, b) and being the largest among three); and (iii) $c + \bar{w} + \beta w_s$ (taking place at (s, s) and being the smallest among three). Here, the relative magnitude among the three lifetime incomes is derived form the assumption (2).

Our task is to find the wights on (i), (ii) and (iii) so as to minimize the expected lifetime income, subject to the constraints that the wights are consistent with the sequential $\boldsymbol{\varepsilon}$-contamination, $\left\{p^0\right\}^{seq\varepsilon}$. We do this task in the Appendix (see A.3 in the Appendix A), and as a result, we obtain the minimal expected lifetime income given this strategy below:

$$\begin{aligned}
&\frac{1+\varepsilon}{2}\left(w_b + \beta w_b\right) + \frac{(1-\varepsilon)^2}{4(1+\varepsilon)}\left(c + \bar{w} + \beta w_b\right) \\
&\qquad\qquad\qquad\qquad + \frac{(1-\varepsilon)(1+3\varepsilon)}{4(1+\varepsilon)}\left(c + \bar{w} + \beta w_s\right) \\
= \;& \frac{1+\varepsilon}{2}\left(w_b + \beta w_b\right) \\
&\quad + \frac{1-\varepsilon}{2}\left(c + \bar{w}\right) + \beta\left(\frac{(1-\varepsilon)^2}{4(1+\varepsilon)} w_b + \frac{(1-\varepsilon)(1+3\varepsilon)}{4(1+\varepsilon)} w_s\right) .
\end{aligned} \tag{18}$$

Finally, the agent decides which strategy to apply by comparing (17) with (18). The Appendix A shows that the latter is larger than the former (see A.4 in the Appendix A), and thus we have proved the next proposition, where V_0'' denotes the value of the job search in the current context.

Proposition 3 (Seq. $\boldsymbol{\varepsilon}$-Cont. + One-Shot Min.) *The value of the job search by the* one-shot minimizer *is given by* (18):

$$V_0'' = \frac{1+\varepsilon}{2}\left(w_b + \beta w_b\right)$$

$$+\frac{1-\varepsilon}{2}(c+\bar{w})+\beta\left(\frac{(1-\varepsilon)^2}{4(1+\varepsilon)}w_b+\frac{(1-\varepsilon)(1+3\varepsilon)}{4(1+\varepsilon)}w_s\right),$$

which is attained by the worker's choosing "to stop" if the first period's state is b and by choosing "to continue" if it is s (and then choosing "to stop" in the second period).

We close the sub-subsection with some remarks. The "worst" weight in (18), $(1-\varepsilon)^2 \big/ (4(1+\varepsilon))$, attached to the largest lifetime income, $c+\bar{w}+\beta w_b$, is *not* contained in the static ε-contamination, $\left\{p^0\right\}^{\varepsilon}$, because it is always true that

$$\frac{(1-\varepsilon)^2}{4(1+\varepsilon)}<\frac{1-\varepsilon}{4}$$

as far as $\varepsilon \in (0,1)$, where the right-hand side is the "worst" weight available in $\left\{p^0\right\}^{\varepsilon}$.

On the other hand, the "best" weight available in $\left\{p^0\right\}^{seq\varepsilon}$ for the smallest lifetime income, $c+\bar{w}+\beta w_s$, would be[17]

$$\left(\frac{1}{2}+\frac{\varepsilon}{2}\right)\cdot\left(\frac{1}{2}+\frac{\varepsilon}{1+\varepsilon}\right)=\frac{1+3\varepsilon}{4},$$

which is also in $\left\{p^0\right\}^{\varepsilon}$.

All of these suggest that while $\left\{p^0\right\}^{seq\varepsilon}$ dilates $\left\{p^0\right\}^{\varepsilon}$ properly (which we will prove formally in the later section), the dilation may take place only *downward*.

2.3.2 The Multi-Stage Minimization

With uncertainty specified by means of $\left\{p^0\right\}^{seq\varepsilon}$, instead of $\left\{p^0\right\}^{\varepsilon}$, the multi-stage minimizer will maximize:

$$\min_{\alpha\in\left[\frac{1}{2}+\underline{\varepsilon},\frac{1}{2}+\bar{\varepsilon}\right]}\Bigg[\alpha\max\left\{w_b+\beta w_b\,,\,c+\beta\min_{\alpha'\in\left[\frac{1}{2}+\underline{\varepsilon}',\frac{1}{2}+\bar{\varepsilon}'\right]}\left[\alpha' w_b+(1-\alpha')w_s\right]\right\}$$
$$+(1-\alpha)\max\Bigg\{w_s+\bar{w}+\beta w_s\,,$$
$$c+\bar{w}+\beta\min_{\alpha''\in\left[\frac{1}{2}+\underline{\varepsilon}',\frac{1}{2}+\bar{\varepsilon}'\right]}\left[\alpha'' w_b+(1-\alpha'')w_s\right]\Bigg\}\Bigg], \quad (19)$$

Here, we explicitly invoked both the definition (16) of $\left\{p^0\right\}^{seq\varepsilon}$ and the definition (15) for its bounds. Importantly, the formula (19) shows that the multi-stage minimizer solves the involved minimization problems "backwardly."

[17] This weight, however, is not employed in (18), in purpose of achieving the overall minimum.

Solving the first maximization problem and the two minimization problems in the brackets and substituting the bounds for the conditionals with the help of the first half of the assumption (2) simplifies the inside of the brackets to show that the formula (19) is equivalent to

$$\min_{\alpha \in \left[\frac{1}{2}+\varepsilon, \frac{1}{2}+\bar{\varepsilon}\right]} \left[\alpha \left(w_b + \beta w_b\right) + (1-\alpha) \max \left\{ w_s + \bar{w} + \beta w_s \, , \right.\right.$$
$$\left.\left. c + \bar{w} + \beta \left(\frac{1-\varepsilon}{2(1+\varepsilon)} w_b + \frac{1+3\varepsilon}{2(1+\varepsilon)} w_s \right) \right\} \right] . \quad (20)$$

Furthermore, solving the remaining optimization problems, substituting the bounds for the marginals, and the second halves of both the assumptions (2) and (4) show that (20) is equivalent to

$$\begin{aligned} & \frac{1+\varepsilon}{2} \left(w_b + \beta w_b\right) + \frac{1-\varepsilon}{2} \left(c + \bar{w} + \beta \left(\frac{1-\varepsilon}{2(1+\varepsilon)} w_b + \frac{1+3\varepsilon}{2(1+\varepsilon)} w_s \right) \right) \\ = \; & \frac{1+\varepsilon}{2} \left(w_b + \beta w_b\right) \\ & \quad + \frac{1-\varepsilon}{2} \left(c + \bar{w}\right) + \beta \left(\frac{(1-\varepsilon)^2}{4(1+\varepsilon)} w_b + \frac{(1-\varepsilon)(1+3\varepsilon)}{4(1+\varepsilon)} w_s \right) . \end{aligned} \quad (21)$$

In sum, we have proved

Proposition 4 (Seq. ε-Cont. + Multi-Stage Min.) *The value of the job search by the* multi-stage minimizer*, V_0''', is given by* (21)*:*

$$V_0''' = \frac{1+\varepsilon}{2} \left(w_b + \beta w_b\right) + \frac{1-\varepsilon}{2} \left(c + \bar{w}\right) + \beta \left(\frac{(1-\varepsilon)^2}{4(1+\varepsilon)} w_b + \frac{(1-\varepsilon)(1+3\varepsilon)}{4(1+\varepsilon)} w_s \right) ,$$

which is attained by the worker's choosing "to stop" if the first period's state is b and by choosing "to continue" if it is s (and then choosing "to stop" in the second period).

2.4 Discussion: Updating and Increased Uncertainty

In this section, we presented the two ways of specifying the time-0 uncertainty: the static ε-contamination and the sequential $\boldsymbol{\varepsilon}$-contamination, and then we applied these concepts to the simple job search model. The result is summarized by the four propositions we have shown so far (Propositions 1, 2, 3, and 4).

Both Propositions 1 and 2 assume the *static* ε-contamination, but each considers the different type of agent: the one-shot minimizer and the multi-stage minimizer. Here, the former type of agent takes the minimization once and for all at time 0 and the latter takes it piece by piece backwardly, though both types can choose appropriate beliefs from among the *common* set of the joint probability charges over the two periods. Then, the two propositions indicate clear distinction between the consequences implied by the behaviors of these two types of agent: the value of the job search are different between the two types.

Next, Propositions 3 and 4 assume the *sequential* ε-contamination and repeat the same exercise. In particular, both types can choose appropriate beliefs from among the *common* set of the joint probability charges. However, these propositions highlight the stark contrast between the results of these propositions and those of Propositions 1 and 2. That is, the values of the job search are completely the same between the two types of agent in Propositions 3, and 4!

In the terminology for dynamic models where sequential decision-making may be permitted, the situation where both the one-shot minimization and the multi-stage minimization lead to the same consequence (the same value and the same optimal strategy) is named *time-consistent*, or *satisfying Bellman's principle of optimality.*[18] Given this terminology, we can say that Propositions 3 and 4 show that the job search model with the time-0 uncertainty specified by the *sequential* ε-contamination considered in this section *is* time-consistent, while Propositions 3 and 4 show that the one with the time-0 uncertainty specified by the *static* ε-contamination is *not*, that is, it is time-*in*consistent.

In the dynamic context, the time-consistency is quite often regarded as a desired property the model should possess basically by the following two reasons: rationality and operational convenience.

Rationality. The time-*in*consistency means that the value found by the backward induction disagrees with the value that can be achieved by applying some contingent plan once and for all, which is the plan made at time 0 by taking all possible contingencies into consideration. Furthermore, any value attained by such contingent plan is always at least weakly dominated by the backwardly-found value because any contingent plan can be available or mimicked also by the backward induction. All of these indicate that the one-shot minimizer should have a strong incentive to revise the plan shaped by herself at time 0 as time goes by. The one-shot minimizer thus deserves to be named *not* to be rational, or to be *ir*rational because she employs the contingent plan regardless of her knowing she will revise that plan in the next period.[19] The cause of this irrationality rests with the

[18]Mathematically speaking, Bellman's principle of optimality holds if both so-called *Bellman's equation* derived from the model under investigation has a unique fixed point and this fixed point coincides with the value of the model. In most cases, however, both "time-consistency" and ''Bellman's principle of optimality" are used interchangeably. We, too, do so here.

[19]Recall the way of us solving the job search models by a backward induction in the

time-inconsistency with which the model with the static ε-contamination is endowed.[20]

Operational convenience. Thanks to the coincidence between the one-shot and multi-stage minimization in the time-consistent model, we can choose whichever is easier for solving the problem. In the one-shot minimization, the agent has to seek for the best contingent plan at time 0. In other words, this search must be conducted in the set of stochastic processes, which set can be easily quite complicated, in particular, when the state space is large and the time-horizon is lengthy. On the other hand, in the multi-stage minimization, the whole problem can be decomposed into some sub-problems, each of which has a much lower dimension of the set of available plans and is much easier to solve because of this lower dimension of the sub-problem. It takes the form of backward induction in the context of a dynamic model, which is widely known as the *dynamic programming* method. To see how the backward induction makes it easy to solve the problem, it should be enough to recall the proofs of Propositions 2 and 4, and compare them to those of Propositions 1 and 3.

With that being said, for the dynamic model with uncertainty like the job search model considered in this section, *we conclude that it is appropriate to specify the time-0 uncertainty by the sequential* $\boldsymbol{\varepsilon}$*-contamination because of its time-consistent property.*

It is well-known that if the uncertainty is reduced to the *risk*, that is, when the set of probability charges representing the worker's belief is a singleton, and if the objective function is time-separable like the lifetime income of the job search model in this section, then the model is *automatically* time-consistent.

On the other hand, if the uncertainty is given by not the risk but the *ambiguity*, which is characterized by the set of probability charges, it is not always the case as the static ε-contamination typically shows. (See Propositions 1 and 2.)

Epstein and Schneider proved, in their important article (2003), that dynamic models with the ambiguity (as well as the time-separable objective function) combined with the maxmin preference à la Gilboa and Schmeidler (1989) exhibit the time-consistency if the ambiguity satisfies some property which they call *rectangularity*. In their terminology, we have shown so far that the sequential $\boldsymbol{\varepsilon}$-contamination *is* rectangular, while the static ε-contamination is *not*.

In the rest of this subsection, we scrutinize the reason why the sequential $\boldsymbol{\varepsilon}$-contamination is rectangular in view of the updating of beliefs as well as the dilation of uncertainty.

former subsections.

[20]For the time-inconsistent model, it is customary to solve it as the one of finding Nash equilibria with multiple players, each of which is the same individual indexed by a different date. Such model is known as a *game with multiple-selves*.

2.4.1 Why Is the Sequential ε-Contamination Rectangular?

Let us first focus on the worker's updating behavior on her beliefs. Note that the worker's each belief over the two periods represented by the joint probability capacity in the relevant set *is* updated in a Bayesian manner *both* in the one-shot and multi-stage minimization. That is, her each belief is updated by Bayes' rule upon observing the first-period's state, for the purpose of deriving the conditional probability charge to calculate the second-period's conditional income in both minimization schemes. In this sense, the updating plays an essential role for the truly dynamic model we consider in this paper.

The significant difference between the one-shot and multi-stage minimization, then, exists in the fact that each available conditional probability charge used for the calculating the second-period's conditional income must be derived by the *same* joint probability charge as that whose first-period's marginal probability charge served to calculate the first-period's marginal income along some two-period's state history in the one-shot minimization, while the multi-stage minimization is free from such restriction.

In the model considered in this section, the optimal strategy for the one-shot-minimizing agent and for the multi-stage-minimizing agent are common: Stop the search upon observing b in the first period, but continue it upon observing s in the same period. However, the value function for the former agent, V_0, is given by (10) and that for the latter agent, V_0', is given by (14), and they are obviously different from each other.

This fact starkly exhibits that the behavior of each type of agents does not generate the same consequences in a given dynamic economic model. In another word, the method by finding the best stochastic process among many such processes and the method of backward induction do not coincide in general.

More generally, the source of this distinction between the two types is the presence of *ambiguity*, that is, the multiplicity of the agent's beliefs represented by the set of probability charges. This is easy to see if we realize that the minimization in each problem (in particular, the one after observing the first-period's state in the multi-stage-minimization problem) is vacuous as well as the law of iterated expectations holds true under risk, that is, when the set of agent's beliefs is a singleton.

Therefore, if ambiguity is reduced to the risk, the distinction between the two types of the behavior vanishes. Such a case is called as a *time-consistent* situation, which is quite common in the dynamic economic models because of its mathematical tractability, in particular, because the backward induction and the dynamic programming methods can be invoked.

It is well-known that even in the presence of ambiguity, a version of the law of iterated expectations holds true under the assumption of *rectangularity*, which was developed by Epstein and Schneider (2003). That is, the rectangular ambiguity guarantees that the one-shot minimization and the multi-stage minimization should coincide. (The more detailed discussion will be deferred until later. See, in particular, Proposition 5 in Subsection 3.2.) The one-shot- or multi-stage-minimizing behavior with rectangular ambiguity thus corresponds exactly to a recursive version of the maxmin preference by Gilboa and Schmeidler (1989) when it is extended to a temporal context.

The disparity of the implications derived from the two types of the behavior shown in the previous two subsections tersely reveals that the one-shot ε-contamination defined by (1) is *not* rectangular. One of the main motivations of the current paper is to develop and study a rectangular version of the one-shot ε-contamination, which we will call *sequential* $\boldsymbol{\varepsilon}$*-contamination*.

The two propositions presented in the previous two sub-subsections show that, with the *sequential* ε-contamination, apparently mutually distinct methods lead to the same solution, we know which must be the case because the *sequential* ε-contamination is rectangular (from the results formally proven in the later section).

On the contrary, in Subsections 2.2 and 2.2.2, the "value functions" which are contradicting each other neatly show that the backward induction method is solving some problem which is irrelevant to the original problem. Intuitively, this happens because the backward induction method fail to pick up the correct "worst" probability charge in evaluating the best lifetime income (so as to minimize the over-all lifetime income). In fact, the backward induction method picks up "too" worst probability charge, which is *not* included by the one-shot ε-contamination. This is the source of the disparity observed in the case of the one-shot ε-contamination and it disappears if we replace the one-shot ε-contamination by the sequential ε-contamination, which dilates the former and makes it rectangular.

Such a dilation takes place typically in the region containing small probability charges. (See the end of sub-subsection 2.3.1.) We thus add the grant-in-aid into the model by making a "bad" income into a "good" income. By this trick, the worst probability charges would become quite relevant and the dilation do matter in order to show clear disparity between one-shot and multi-stage minimization. Without the grant-in-aid as it is in the model, even in the scheme of the one-shot ε-contamination the disparity between the two solution methods would vanish.

3 Rectangularity and Sequential $\boldsymbol{\varepsilon}$-Contamination over Two Periods

The rectangularity is a concept developed by Epstein and Schneider (2003). After giving an overview of this concept and its basic properties, this section proposes a version of the one-shot ε-contamination that is rectangular, which we call *sequential* $\boldsymbol{\varepsilon}$*-contamination*. We then derive its important properties in the rest of this subsection.

3.1 Some Notations and Definitions

The following notations draw on Chapter 14 of Nishimura and Ozaki (2017) at the outset, and then, they will be further simplified for the later use in this paper.

Let S be a state space for each single period and let $\Omega := S \times S$ be the whole state space. A generic element of Ω is denoted by (s_1, s_2). In the main text of the paper, we exclusively consider the two-period model. Appendix B extends it to an arbitrarily finite horizon models.

The information structure, which represents the basis of the decision-maker's view of the world, is exogenously given by a filtration $\mathcal{F} := \langle \mathcal{F}_t \rangle_{t=0,1,2}$. Let $m, n \geq 2$ and let $\langle E_i \rangle_{i=1}^m$ and $\langle F_j \rangle_{j=1}^n$ be two finite partitions of S. Throughout this section, we fix these two partitions. We assume that $\mathcal{F}_1$ is represented by a finite partition of Ω of the form: $\langle E_i \times S \rangle_i$, and that $\mathcal{F}_2$ is represented by a finite partition of Ω of the form: $\langle E_i \times F_j \rangle_{i,j}$.

We abuse a notation by using a partition also to denote the algebra generated by that partition on S and Ω. By this convention, $\mathcal{F}_1$ and $\mathcal{F}_2$ are the algebras on Ω and it holds that $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2$, where $\mathcal{F}_0 := \{\phi, \Omega\}$. Thus, information increases as time goes by.

We now turn to notations for probability charges. Let $\mathscr{M}(\Omega, \mathcal{F}_i)$ be the space of all probability charges on the measurable space $(\Omega, \mathcal{F}_i)$ $(i = 1, 2)$.

Given $p \in \mathscr{M}(\Omega, \mathcal{F}_2)$, we denote by $p|_1$ its restriction on $(\Omega, \mathcal{F}_1)$. Although $p|_1$ is formally a charge on Ω, it can be naturally regarded as the one on the measurable space, $(S, \langle E_i \rangle_i)$, and in that case, $p|_1(\cdot) = p(\cdot \times S)$. Thus viewed, $p|_1$ can be considered as the <u>first-period marginal probability charge</u> of p. We henceforth write $p|_1(E_i)$ simply as p_i for all $i \leq m$.

Given $p \in \mathscr{M}(\Omega, \mathcal{F}_2)$, $i \leq m$ and E_i satisfying $p(E_i \times S) > 0$, we denote by $p|_{E_i}(\cdot)$ the "posterior" probability charge on $(S, \langle F_j \rangle_j)$ conditional on the occurrence of $E_i \times S$. Here, the adjective "posterior" signifies the fact that this is a probability charge the decision-maker obtains after she made an observation, E_i, in the first period (and when she updates based on it according to Bayes' rule). That is, $(\forall i, j)$ $p|_{E_i}(F_j) := p(E_i \times F_j)/p(E_i \times S)$. With the conventional wording, we henceforth write $p|_{E_i}(F_j)$ simply as p_{ij} for all $i \leq m$ and $j \leq n$. In a plain English, p_{ij} is the *conditional probability charge* of F_j *given* E_i.

Finally, given $p \in \mathscr{M}(\Omega, \mathcal{F}_2)$, we henceforth write the *joint* probability charge $p(E_i \times F_j)$ simply as $p_{i,j}$ for all $i \leq m$ and $j \leq n$. Note that regardless of similar notations between p_{ij} and $p_{i,j}$, their meanings are quite different.

3.1.1 The Decomposition of a Probability Charge

So far, we have defined three real numbers: p_i, p_{ij}, and $p_{i,j}$, for each $p \in \mathscr{M}(\Omega, \mathcal{F}_2)$ and each $i \leq m$ and $j \leq n$.

By means of all notations introduced thus far, an important result called the "decomposition of a probability charge into its marginal and conditional" can be stated as follows: Given any probability charge $p \in \mathscr{M}(\Omega, \mathcal{F}_2)$, p can be written as

$$(\forall i, j) \quad p_{i,j} = p_i \cdot p_{ij} \tag{22}$$

as far as $p(E_i \times S) > 0$.

Conversely, given any list (vector) of first-period marginals of $p \in \mathscr{M}(\Omega, \mathcal{F}_2)$ (where each marginal is identified with an element of $\mathscr{M}(S, \langle E_i \rangle_i)$), $(p_i)_i := (p_1, p_2, \ldots, p_m)$, as well as any set of lists (vectors) of well-defined conditionals (where any element of the set is similarly identified with an elements of $\mathscr{M}(S, \langle F_j \rangle_j)$),

$$\{(p_{ij})_j\}_i := \{(p_{11}, p_{12}, \ldots, p_{1n}), (p_{21}, p_{22}, \ldots, p_{2n}), \ldots, (p_{m1}, p_{m2}, \ldots, p_{mn})\},$$

the right-hand side of Equation (22) "defines" a probability charge $p \in \mathscr{M}(\Omega, \mathcal{F}_2)$ with i and j varying.

This decomposition (that is, Equation (22)) will be used repeatedly in what follows. In particular, one of significant implications of Equation (22) is the *law of iterated expectations*:

$$E^p [u_1 + u_2] = E^{p|_1} \left[u_1 + E^{p|_E} [u_2] \right], \tag{23}$$

where u_1 is a real-valued $\mathcal{F}_1$-measurable function, u_2 is a real-valued $\mathcal{F}_2$-measurable function, E is an arbitrary element of the partition, $\langle E_i \rangle_i$, and E^p is the mathematical expectation with respect to a relevant probability charge of p.

3.1.2 The Formal Definition of the One-Shot ε-Contamination

We now formally define the *one-shot ε-contamination* by means of the notations introduced thus far. (Recall that it was already defined "loosely" by Equation (1)).

Let p^0 be a probability charge on $(\Omega, \mathcal{F}_2)$ such that $(\forall i)\ p_i^0 > 0$, and let $\varepsilon \in (0, 1)$. As we already suggested, the probability charge p^0 may be called as a "principal" probability charge, which the agent believes to be the true probability charge with the $((1 - \varepsilon) \times 100)\%$ conviction. When the agent's conviction were wrong, she would have completely no idea about the true probability charge that governs the world.

Then, the *one-shot ε-contamination of* p^0, denoted by $\{p^0\}^\varepsilon$, is defined by

$$\{p^0\}^\varepsilon := \left\{ (1 - \varepsilon) p^0 + \varepsilon q \,\middle|\, q \in \mathscr{M}(\Omega, \mathcal{F}_2) \right\}. \tag{24}$$

We now move to the maxmin preference with the one-shot ε-contamination. The maxmin preference with a general set of probability charges was axiomatized by Gilboa and Schmeidler (1989) and the one with the set specified by the one-shot ε-contamination was axiomatized by Nishimura and Ozaki (2006). Both maxmin preferences represent the agent's pessimistic attitude toward ambiguity. i

Let u_i denote an arbitrary real-valued $\mathcal{F}_i$-measurable function on Ω ($i = 1, 2$). Here, we interpret u_1 as an agent's felicity function in period 1 and u_2 as her (present-valued) felicity function in period 2, and thus, her lifetime (stochastic) utility is given by $u_1 + u_2$. Then, the *maxmin expected utility with* $\left\{p^0\right\}^\varepsilon$ is defined by

$$\min_{p \in \left\{p^0\right\}^\varepsilon} E^p \left[u_1 + u_2\right] , \tag{25}$$

where E^p denotes the standard mathematical expectation with respect to a probability charge p. ((25) would be basically the same as (6) if we understand that the current-valued felicity function here would happen to be a common affine function of money with a positive coefficient.)

3.2 Knightian Uncertainty and Its Rectangularity

Any nonempty subset $\mathscr{P}$ of $\mathscr{M}(\Omega, \mathcal{F}_2)$ is called *Knightian uncertainty* or *ambiguity*. The one-shot ε-contamination introduced above is an example of Knightian uncertainty.

Given Knightian uncertainty $\mathscr{P}$, its *first-period marginal Knightian uncertainty*, denoted by $\mathscr{P}|_1$, is the nonempty subset of $\mathscr{M}(S, \langle E_i \rangle_i)$ that is defined by

$$\mathscr{P}|_1 := \{\, p|_1 \,|\, p \in \mathscr{P} \,\} ,$$

where $p|_1$ is the first-period marginal probability charge of p defined in the previous subsection and it is written in the conventional wording so that $p|_1 \in \mathscr{M}(S, \langle E_i \rangle_i)$.

Next, let $\mathscr{P}$ be Knightian uncertainty, suppose that $E \in \langle E_i \rangle_i$ was observed in the first period, and suppose that *every* probability charge in $\mathscr{P}$ is updated given E by Bayes' rule. As a result of this procedure, we obtain $\mathscr{P}|_E \subseteq \mathscr{M}(S, \langle F_j \rangle_j)$ which is defined by

$$\mathscr{P}|_E := \{\, p|_E \,|\, p \in \mathscr{P} \,\} ,$$

where $p|_E$ is the "posterior" or conditional probability charge defined in the previous subsection. Note that $\mathscr{P}|_E = \phi_{GB}(\mathscr{P}, E)$ by the notation of Nishimura and Ozaki (2017, Chapter 14), where "GB" abbreviates "generalized Bayes." The set $\mathscr{P}|_E$ may be thought of as the state of uncertainty in the second period after the observation E was made in the first period. We may call $\mathscr{P}|_E$ the *conditional Knightian uncertainty given* E in the first period.

Knightian uncertainty $\mathscr{P}$ is *rectangular* by definition if for any $p', p'' \in \mathscr{P}$, it holds that $(p'_i \cdot p''_{ij})_{i,j} \in \mathscr{P}$, where p' is decomposed into $(\forall i, j)$ $p'_{i,j} = p'_i \cdot p'_{ij}$; p'' is decomposed into $(\forall i, j)$ $p''_{i,j} = p''_i \cdot p''_{ij}$; and $(p'_i \cdot p''_{ij})_{i,j}$ defines a joint probability charge in $\mathscr{M}(\Omega, \mathcal{F}_2)$ by Equation (22). The concept of rectangularity was introduced by Epstein and Schneider (2003).

One of the novelties of the rectangularity is that a version of the "law of iterated expectations" (see (23)) for the minimum expectations with Knightian uncertainty holds when it is rectangular. In the terminology of Section 2, the objective function of the one-shot minimizer and that of multi-stage minimizer coincides, and hence, the problem for the one-shot minimizer can be solved by the method of the backward induction. The problem of the time-inconsistency does not arise when the rectangularity is satisfied, which brings about the extreme operational convenience for economic applications.

Proposition 5 *Let $\mathscr{P}$ be a weak *compact subset of $\mathscr{M}(\Omega, \mathcal{F}_2)$. Then, both $\mathscr{P}|_1$ and $\mathscr{P}|_E$ are also weak *compact. Thus, all the minima in* (26) *below exist. Furthermore, if $\mathscr{P}$ is rectangular, we have*

$$\min_{p \in \mathscr{P}} E^p \left[u_1 + u_2\right] = \min_{p' \in \mathscr{P}|_1} E^{p'} \left[u_1 + \min_{p'' \in \mathscr{P}|_E} E^{p''} \left[u_2\right]\right] , \tag{26}$$

where u_i is an arbitrary real-valued $\mathcal{F}_i$-measurable function on Ω $(i = 1, 2)$ and E is an arbitrary element of the partition, $\langle E_i \rangle_i$. (Proof B.1 in Appendix A)

If you look at the proof, you will find that the inequality "$\geq$" holds in general even without the rectangularity, which seems to be natural to see. On the other hand, note that in the previous section, it was quite likely that $V_0 > V'_0$, where V_0 and V'_0 are the values of the job search for the one-shot minimizer and for the multi-stage minimizer, respectively. As we will see soon, the one-shot ε-contamination is not rectangular. So, the same direction of the two inequalities seems to be (partly) a consequence of Proposition 5. But, this is misleading because Proposition 5 only compares the objective functions themselves, not the values of the problem, the latter of which incorporate the best strategies of the agents.

3.3 The Sequential ε-Contamination

This subsection defines a variant of the one-shot ε-contamination which IS rectangular, which we call *sequential ε-contamination*. Before doing it formally, we sketch its definition intuitively .

First, note that any element $p \in \mathscr{M}(\Omega, \mathcal{F}_2)$ can be alternatively expressed as an $(m \times n)$-dimensional vector (or list of $m \times n$ joint probability charges) as

$$p = (p_{1,1}, p_{1,2}, \ldots, p_{1,n}; p_{2,1}, p_{2,2}, \ldots, p_{2,n}; \ldots; p_{m,1}, p_{m,2}, \ldots, p_{m,n})$$

which satisfies $p \in [0,1]^{m\times n}$ and $\sum_{i=1}^{m}\sum_{j=1}^{n} p_{i,j} = 1$. For simplicity, we write this as $p = (p_{i,j})_{i,j}$ (while we already used this notation in some occasions). As is apparent in the above formulation, p includes all possible "evolutions" of probability charges over two periods.

Second, by means of this notation, the one-shot ε-contamination of p^0, defined by (24), will be clearly rewritten as

$$\left\{p^0\right\}^{\varepsilon} = \left\{ (1-\varepsilon)\left(p^0_{i,j}\right)_{i,j} + \varepsilon\left(q_{i,j}\right)_{i,j} \;\middle|\; (q_{i,j})_{i,j} \in [0,1]^{m\times n} \text{ and } \textstyle\sum_{i,j} q_{i,j} = 1 \right\}. \tag{27}$$

Third, for any $q \in \mathscr{M}(\Omega, \mathcal{F}_2)$, define $(\forall i,j)\ \delta_{i,j} := \varepsilon(-p^0_{i,j} + q_{i,j})$. Then, it can be easily verified that the requirement that $(\forall i,j)\ q_{i,j} \in [0,1]$ and $\sum_{i,j} q_{i,j} = 1$ can be turned into the requirement on $(\delta_{i,j})_{i,j}$ that $(\forall i,j)\ \delta_{i,j} \in [-\varepsilon p^0_{i,j}\,,\ \varepsilon(1-p^0_{i,j})]$ and $\sum_{i,j}\delta_{i,j} = 0$. Therefore, (27) is further rewritten as

$$\left\{p^0\right\}^{\varepsilon} = \left\{ \left(p^0_i \cdot p^0_{ij} + \delta_{i,j}\right)_{i,j} \;\middle|\; (\forall i,j)\ \delta_{i,j} \in [\underline{\delta}_{i,j}, \bar{\delta}_{i,j}] \text{ and } \textstyle\sum_{i,j}\delta_{i,j} = 0 \right\} \tag{28}$$

with

$$\underline{\delta}_{i,j} := -\varepsilon p^0_{i,j} \quad \text{and} \quad \bar{\delta}_{i,j} := \varepsilon(1-p^0_{i,j})\,, \tag{29}$$

where we used $p^0_{i,j} = p^0_i \cdot p^0_{ij}$, which is justified by (22).

Fourth and finally, suppose that we could find two vectors $(\varepsilon_i)_{i=1}^{m}$ and $(\varepsilon_{ij})_{i=1}^{m}{}_{j=1}^{n}$, and their ranges they can move around within, such that it holds that

$$(\forall i,j) \quad \delta_{i,j} = \varepsilon_{ij} p^0_i + \varepsilon_i p^0_{ij} + \varepsilon_i \varepsilon_{ij}$$

and $(\delta_{i,j})_{i,j}$ thus specified should satisfy the constraints in (28) and (29). Then, we would have

$$\left(p^0_i \cdot p^0_{ij} + \delta_{i,j}\right)_{i,j} = \left((p^0_i + \varepsilon_i)(p^0_{ij} + \varepsilon_{ij})\right)_{i,j}\,,$$

and thus, we had shown that the one-shot ε-contamination could be rewritten so as to become rectangular, concluding that the one-shot ε-contamination is rectangular.

Unfortunately, it is impossible for us to execute this procedure, and thus, one-shot ε-contamination is *not* rectangular in general.

3.3.1 Counter-Example

To substantiate the statement we gave in the end of the last paragraph, the next simple example constructs the one-shot ε-contamination which is *not* rectangular. In fact, it is exactly the same as the one we used for presenting a job search model with the one-shot ε-contamination.

Example 1 Let $S := \{b, s\}$. Then, we have $\Omega = \{(b,b),(b,s),(s,b),(s,s)\}$. Let p^0 be a principal probability charge on $(\Omega, 2^\Omega)$ defined by $p^0_{b,b} = p^0_{b,s} = p^0_{s,b} = p^0_{s,s} = 1/4$ and we consider $\{p^0\}^\varepsilon$ for an arbitrary $\varepsilon \in (0,1)$.

We see that $p := \left(\frac{1}{4}+\frac{\varepsilon}{4}, \frac{1}{4}+\frac{\varepsilon}{4}, \frac{1}{4}-\frac{\varepsilon}{4}, \frac{1}{4}-\frac{\varepsilon}{4}\right) \in \{p^0\}^\varepsilon$ (let $q := \left(\frac{1}{2}, \frac{1}{2}, 0, 0\right)$), and that $p' := \left(\frac{1}{4}+\frac{\varepsilon}{4}, \frac{1}{4}-\frac{\varepsilon}{4}, \frac{1}{4}+\frac{\varepsilon}{4}, \frac{1}{4}-\frac{\varepsilon}{4}\right) \in \{p^0\}^\varepsilon$ (let $q := \left(\frac{1}{2}, 0, \frac{1}{2}, 0\right)$).

From p, we can compute the first-period marginal of p as $p|_1 = (p_b, p_s) = \left(\frac{1}{2}+\frac{\varepsilon}{2}, \frac{1}{2}-\frac{\varepsilon}{2}\right)$, and the conditionals of p, p_{bb} and so on, as $p_{bb} = p_{bs} = p_{sb} = p_{ss} = 1/2$. Similarly, from p', we can compute $p'|_1 = (p'_b, p'_s) = \left(\frac{1}{2}, \frac{1}{2}\right)$, and conditionals: $p'_{bb} = \frac{1}{2}+\frac{\varepsilon}{2}$, $p'_{bs} = \frac{1}{2}-\frac{\varepsilon}{2}$, $p'_{sb} = \frac{1}{2}+\frac{\varepsilon}{2}$ and $p'_{ss} = \frac{1}{2}-\frac{\varepsilon}{2}$.

These computations show that

$$\left(p_b p'_{bb}, p_b p'_{bs}, p_s p'_{sb}, p_s p'_{ss}\right) = \left(\frac{1}{4}+\frac{\varepsilon}{2}+\frac{\varepsilon^2}{4}, \frac{1}{4}-\frac{\varepsilon^2}{4}, \frac{1}{4}-\frac{\varepsilon^2}{4}, \frac{1}{4}-\frac{\varepsilon}{2}+\frac{\varepsilon^2}{4}\right)$$

would be an element of $\{p^0\}^\varepsilon$, if it were rectangular. But it would *not*, in fact. Because $p_s p'_{ss} = \frac{1}{4}-\frac{\varepsilon}{2}+\frac{\varepsilon^2}{4} < \frac{1}{4}-\frac{\varepsilon}{4} = (1-\varepsilon)p^0_{s,s}$, which is the minimum value $p_{s,s}$ can take on as long as p belongs to $\{p^0\}^\varepsilon$, we must conclude that $p_s p'_{ss} \notin \{p^0\}^\varepsilon$. □

3.3.2 The Formal Definition of the Sequential ε-Contamination

We are now ready to formally define the sequential $\boldsymbol{\varepsilon}$-contamination along the line developed in the beginning of this subsection.

Let $p^0 \in \mathscr{M}(\Omega, \mathcal{F}_2)$ be a "principal" probability charge such that $(\forall i)\ p^0_i > 0$. Now, let $\boldsymbol{\varepsilon}$ be a lengthy real vector, $\boldsymbol{\varepsilon} = (\underline{\varepsilon}_i; \bar{\varepsilon}_i; \underline{\varepsilon}_{ij}; \bar{\varepsilon}_{ij})_{i,j}$, which is defined by $(\forall i, j)$

$$\underline{\varepsilon}_i := -\varepsilon p^0_i\ ;\ \bar{\varepsilon}_i := \varepsilon(1-p^0_i)\ ;\ \underline{\varepsilon}_{ij} := \frac{-\varepsilon p^0_{ij}}{(1-\varepsilon)p^0_i + \varepsilon}\ ;\ \bar{\varepsilon}_{ij} := \frac{\varepsilon(1-p^0_{ij})}{(1-\varepsilon)p^0_i + \varepsilon}\ , \tag{30}$$

where ε is the one with which the one-shot ε-contamination is defined in (24), or in (28) and (29).[21]

Then, we use $\boldsymbol{\varepsilon}$ to define the *sequential $\boldsymbol{\varepsilon}$-contamination of* p^0, $\{p^0\}^{seq\boldsymbol{\varepsilon}}$, by[22]

$$\{p^0\}^{seq\boldsymbol{\varepsilon}} := \left\{ \left((p^0_i+\varepsilon_i)(p^0_{ij}+\varepsilon_{ij})\right)_{i,j} \,\middle|\, (\forall i)\ \varepsilon_i \in [\underline{\varepsilon}_i, \bar{\varepsilon}_i];\ \textstyle\sum_i \varepsilon_i = 0; \right.$$
$$\left. (\forall i,j)\ \varepsilon_{ij} \in [\underline{\varepsilon}_{ij}, \bar{\varepsilon}_{ij}] \ \text{ and } \ (\forall i)\ \textstyle\sum_j \varepsilon_{ij} = 0 \right\}. \tag{31}$$

Note that the restrictions imposed by (30) on the ranges within which $(\varepsilon_i)'s$ and $(\varepsilon_{ij})'s$ may move around are sufficient for $(p^0_i + \varepsilon_i)_i$ and $(\forall i)\ (p^0_{ij} + \varepsilon_{ij})_j$ to be probability charges as well as for p^0 to be included in $\{p^0\}^{seq\boldsymbol{\varepsilon}}$.

[21] Recall that (28) and (29) are still correct as reformulations of the one-shot ε-contamination.

[22] We use a boldface letter for the epsilon in the term of "sequential $\boldsymbol{\varepsilon}$-contamination" because the epsilon there is a vector, not a single number.

In contrast with the one-shot ε-contamination, the sequential $\boldsymbol{\varepsilon}$-contamination allows the agent to "renew" Knightian uncertainty after making an observation in the first period because ε's may depend on i. This "renewness" is quite important for the sequential $\boldsymbol{\varepsilon}$-contamination: the agent may discard or change the probability charge she has initially in mind when she updates her belief, while she must keep the same probability charge and update it even after making a new observation.

3.4 The Basic Properties of Sequential $\boldsymbol{\varepsilon}$-Contamination

Because we have finished its definition, we now examine some properties of the sequential $\boldsymbol{\varepsilon}$-contamination.

Above all, as we notified beforehand, the sequential $\boldsymbol{\varepsilon}$-contamination is in fact rectangular as the next proposition shows.

Proposition 6 *The sequential $\boldsymbol{\varepsilon}$-contamination defined by* (31) *is rectangular.* (Proof B.2 in Appendix A)

The next proposition is also necessary to justify the sequential $\boldsymbol{\varepsilon}$-contamination, which states that the first-period marginals of the two kinds of ε-contamination: the one-shot ε-contamination and the sequential ε-contamination, are the same. If this is not the case, the agent may face either larger or smaller uncertainty at the very outset depending on which type of the ε-contamination represents her belief, regardless of what she will obtain as a new piece of information and how to update uncertainty as time goes by.

Proposition 7 (First-Period Marginals) *The first-period marginal Knightian uncertainty of the sequential $\boldsymbol{\varepsilon}$-contamination* ((31) and (30)) *and that of the one-shot ε-contamination* ((28) and (29)) *coincide. That is,*

$$\left\{p^0\right\}^{seq\varepsilon}\Big|_1 = \left\{p^0\right\}^{\varepsilon}\Big|_1 .$$

(Proof B.3 in Appendix A)

The proof for Proposition 7 suggests that the "bounds," $\underline{\varepsilon}_i$ and $\bar{\varepsilon}_i$, that appear in the sequential $\boldsymbol{\varepsilon}$-contamination are "tight."

Here, some may wonder if Knightian uncertainty which is both rectangular and having the same first-period marginal as the given one-shot ε-contamination should be unique. In other words, someone may think that such Knightian uncertainty must be nothing but the sequential $\boldsymbol{\varepsilon}$-contamination itself. Unfortunately, however, this is not the case.

To see this, recall the definition of the sequential $\boldsymbol{\varepsilon}$-contamination, (31), and note that a *necessary* and sufficient condition for $(p_i^0+\varepsilon_i)_i$ and $(\forall i)$ $(p_{ij}^0+\varepsilon_{ij})_j$ to be probability charges as well as for p^0 to be included in $\left\{p^0\right\}^{seq\varepsilon}$ is that $(\forall i,j)$ $-p_i^0 \leq \underline{\varepsilon}_i \leq 0 \leq \bar{\varepsilon}_i \leq 1-p_i^0$ and $-p_{ij}^0 \leq \underline{\varepsilon}_{ij} \leq 0 \leq \bar{\varepsilon}_{ij} \leq 1-p_{ij}^0$.

The bounds (30) that define the sequential $\boldsymbol{\varepsilon}$-contamination obviously satisfy these conditions.

The proof for the rectangularity (Proof B.2) goes through only with this necessary and sufficient conditions. The proof that the first-period marginals coincide for both types of the ε-contamination (Proof B.3) invoked only the bounds, $\underline{\varepsilon}_i$ and $\bar{\varepsilon}_i$, that are defined by (30).

Therefore, there are some bounds, $\underline{\varepsilon}_{ij}$ and $\bar{\varepsilon}_{ij}$, that are (possibly quite complicated and) different from the ones in (30) but still keeping satisfying the necessary and sufficient conditions above. Thus, there are some degree of freedom in bounds defining the set of *conditional* probability charges. All this reveals the non-uniqueness we promised to show.

In sum, the rectangularity and the coincidence of the first-period marginals are not strong enough to pin down Knightian uncertainty given the one-shot ε-contamination.

3.5 Comparison between the One-Shot and Sequential $\boldsymbol{\varepsilon}$-Contamination

Thus far, we have found that the sequential $\boldsymbol{\varepsilon}$-contamination is conveniently rectangular and has the identical first-period marginal with the one-shot ε-contamination. Although these are certainly good properties we should expect for it, we have also shown that a way of formulating Knightian uncertainty possessing these properties is not unique, given a one-shot ε-contamination. (See the end of the previous section.)

The main objective of this section is to mathematically scrutinize the relationship between the two versions of ε-contamination: the one-shot ε-contamination and the sequential $\boldsymbol{\varepsilon}$-contamination we proposed in the previous subsection, in order to further motivate the latter specification of Knightian uncertainty as a suitable specification of the ε-contamination in a truly dynamic context.

The first result in this section is that the sequential $\boldsymbol{\varepsilon}$-contamination is at least as large as the one-shot ε-contamination.

Proposition 8 *It holds that* $\left\{p^0\right\}^{\varepsilon} \subseteq \left\{p^0\right\}^{seq\boldsymbol{\varepsilon}}$. (Proof B.4 in Appendix A)

However, it does not hold that $\left\{p^0\right\}^{\varepsilon} \supseteq \left\{p^0\right\}^{seq\boldsymbol{\varepsilon}}$, and thus, the inclusion established in Proposition 8 is strict in general. See Example 1 in Subsection 3.3.1.

Proposition 8 suggests that the sequential $\boldsymbol{\varepsilon}$-contamination may be constructed from the one-shot ε-contamination by adding to it all products of a marginal and a conditional both of which is derived from any (possibly distinct) probability charge included in the one-shot ε-contamination.

In order to make this idea mathematically more rigorous, we develop a new concept of what we call the "rectangular-hull." To this end, let $\mathscr{P} \subseteq$

$\mathscr{M}(\Omega, \mathcal{F}_2)$ be any (not necessarily rectangular) Knightian uncertainty. Then, consider *rectangular* Knightian uncertainty, say $\mathscr{P}'$, which is containing $\mathscr{P}$ and is *minimal* in the sense that, if $\mathscr{P}''$ is another Knightian uncertainty which is rectangular and containing $\mathscr{P}$, then $\mathscr{P}' \subseteq \mathscr{P}''$. We call such $\mathscr{P}'$ the *rectangular-hull* of $\mathscr{P}$, if any, and denote it by rect$\mathscr{P}$.

Proposition 9 (Rectangular-Hull) *For any $p^0 \in \mathscr{M}(\Omega, \mathcal{F}_2)$ and any $\varepsilon \in (0,1)$, $\mathrm{rect}\big(\{p^0\}^{\varepsilon}\big)$ exists and it equals $\{p^0\}^{seq\varepsilon}$.* (Proof B.5 in Appendix A)

Proposition 9 is important: it claims that the sequential $\boldsymbol{\varepsilon}$-contamination is very near to the original one-shot ε-contamination in the sense that it is "minimal" among the rectangular Knightian uncertainty that contains that ε-contamination.

Our next result is quite remarkable. It exhibits another equivalent expression of the sequential $\boldsymbol{\varepsilon}$-contamination which is intuitive and convenient for its applications. (For instance, see the application in the next section.) In particular, it expresses the sequential $\boldsymbol{\varepsilon}$-contamination via successive ε's together with the conditional principal probability charges.

In the two-period framework we are now working with, we call it *ε-$\boldsymbol{\varepsilon}'$ contamination*, where $\boldsymbol{\varepsilon}'$ can be described by the simple formula that depends on the original ε with which the one-shot ε-contamination was defined. The proposition's beauty will be emphasized if we look at an arbitrarily finite horizon case, which we will conduct in Appendix B.

Proposition 10 (ε-$\boldsymbol{\varepsilon}'$ Contamination) *For any $p^0 \in \mathscr{M}(\Omega, \mathcal{F}_2)$ and any $\varepsilon \in (0,1)$, it holds that*

$$\{p^0\}^{seq\varepsilon} = \Big\{ \big(((1-\varepsilon)p_i^0 + \varepsilon q_i\big) \cdot \big((1-\varepsilon_i')p_{ij}^0 + \varepsilon_i' q_{ij})\big)_{i,j} \,\Big|\, (\forall i)\, q_i \in [0,1]; \\ \textstyle\sum_i q_i = 1;\; (\forall i,j)\, q_{ij} \in [0,1] \text{ and } (\forall i)\, \sum_j q_{ij} = 1 \Big\}, \quad (32)$$

where $(\forall i)$ ε_i' is defined by

$$\varepsilon_i' := \frac{\varepsilon}{(1-\varepsilon)p_i^0 + \varepsilon}\,.$$

Furthermore, $(\forall i)$ $\varepsilon_i' > \varepsilon$. (Proof B.6 in Appendix A)

We conclude this section by comparing the conditional Knightian uncertainty given the first-period's observation when Knightian uncertainty is characterized by the one-shot ε-contamination with when it is characterized by the *sequential* $\boldsymbol{\varepsilon}$-contamination.

Nishimura and Ozaki (2017, Chapter 14) extensively studies the conditional Knightian uncertainty when it is characterized by the one-shot ε-contamination over two periods. In particular, they prove that

$$(\forall i) \qquad \{p^0\}^{\varepsilon}\big|_{E_i} = \{p_i^0\}^{\varepsilon_i'}$$

where ε_i' is exactly the same as the one defined above in Proposition 10 (Nishimura and Ozaki, 2017, Theorem 14.5.1, p.242).

If we paraphrase their result in plain English, it says that the set of conditional probability charges in the one-shot ε-contamination of p^0 given E_i is again the "(one-shot)[23] form of ε-contamination" with "new"$\varepsilon = \varepsilon_i'$ and "new"p^0 is the conditional probability charge of p^0 given E_i. Furthermore, ε_i' is larger than ε, and hence, uncertainty "dilates."

Now, let $\{p^0\}^{seq\boldsymbol{\varepsilon}}$ be the sequential $\boldsymbol{\varepsilon}$-contamination defined by (31) and let $E_i \in \langle E_k \rangle_k$ for some i. Then, the definition of the conditional Knightian uncertainty and Proposition 10 immediately imply

$$\{p^0\}^{seq\boldsymbol{\varepsilon}}\big|_{E_i} = \left\{ \left((1-\varepsilon_i')p_{ij}^0 + \varepsilon_i' q_{ij}\right)_j \,\middle|\, (\forall j)\, q_{ij} \in [0,1] \text{ and } \textstyle\sum_j q_{ij} = 1 \right\}, \tag{33}$$

where ε_i' is as defined in Proposition 10.

We thus obtain the next proposition.

Proposition 11 ("Posteriors") *For any i and for any $E_i \in \langle E_k \rangle_k$,*

$$\{p^0\}^{seq\boldsymbol{\varepsilon}}\big|_{E_i} = \{p^0\}^{\varepsilon}\big|_{E_i} .$$

Proposition 11 shows that "posterior" Knightian uncertainty of the sequential $\boldsymbol{\varepsilon}$-contamination after making an observation is the same as the "posterior" of the one-shot ε-contamination after making the same observation. This fact further strengthens the "nearness" of the sequential $\boldsymbol{\varepsilon}$-contamination to the one-shot ε-contamination.

The sequential $\boldsymbol{\varepsilon}$-contamination is a reasonable modification of, or even better than the one-shot ε-contamination *in dynamic setups*. We may argue so because the former has the same first-period marginal as the latter, because the former has the same posterior as the latter ("nearness" from both sides) , as well as because the former is rectangular above all.

From all of this, we see that the sequential $\boldsymbol{\varepsilon}$-contamination shares exactly the same important property as the one-shot ε-contamination: by updating according to Bayes's rule after making an observation, *uncertainty dilates*, which may be regarded as a remarkable and robust property of the ε-contamination-type Knightian uncertainty.[24]

4 An Application to Job Search Model with Bayesian Updating

This section applies the sequential $\boldsymbol{\varepsilon}$-contamination developed thus far in this paper to a job search model. The model specifies the one by Nishimura

[23] It is the "one-shot" because there is only one period left.

[24] Shishkin Ⓡ Ortoleva (2019) try to measure the degree of the dilation of ambiguity upon Bayesian updating experimentally in a more general framework.

and Ozaki (2004) by assuming that the ambiguity is now specified by the sequential ε-contamination in the framework of the two-period model.[25] On the other hand, it generalizes their model by incorporating Bayesian updating behavior by the worker.

We then conduct very similar sensitivity analyses to Nishimura and Ozaki (2004) and show that the basically the same result (an increase in uncertainty shortens the search period) holds also in the current framework.

4.1 The Model

We employ almost the same model as we used in Section 2 except that the assumption (2) is now weakened to

$$c \vee w_s < w_b \,, \tag{34}$$

where $c \vee w_s$ abbreviates $\max\{c\,, w_s\}$; the assumption (**??**) is not assumed any more; and the assumption (3) is replaced by

Assumption 1 The principal probability charge p^0 satisfies

$$p^0_{bb} < \frac{1}{2} + \frac{\varepsilon}{2(1-\varepsilon)p^0_b} \quad \text{and} \quad p^0_{sb} < \frac{1}{2} + \frac{\varepsilon}{2(1-\varepsilon)(1-p^0_b)} \,.$$

And most importantly, the objective function of the worker is now given by

$$\min_{p \in \{p^0\}^{seq\varepsilon}} E^p \left[y_1 + \beta y_2 \right] \,, \tag{35}$$

where $\{p^0\}^{seq\varepsilon}$ denotes the *sequential ε-contamination of* p^0.

We rely on Assumption 1 when we compute the worker's maxmin expected utility later in this section. Note that it will be easier to be satisfied if the value of ε increases because the inequalities becomes slacker then. Lemma B.7 in Appendix A provides a sufficient condition for Assumption 1 to be satisfied.

4.2 The Second-Period Value Function

An important novelty of the sequential ε-contamination is that it is rectangular (Proposition 6). Therefore, we can invoke Proposition 5 to rewrite (35) as

$$\min_{p \in \{p^0\}^{seq\varepsilon}} E^p \left[y_1 + \beta y_2 \right] = \min_{p' \in \{p^0\}^{seq\varepsilon}|_1} E^{p'} \left[y_1 + \beta \min_{p'' \in \{p^0\}^{seq\varepsilon}|_E} E^{p''} \left[y_2 \right] \right] \,, \tag{36}$$

where $\{p^0\}^{seq\varepsilon}|_1$ and $\{p^0\}^{seq\varepsilon}|_E$ be the marginal and conditional sequential ε-contamination of p^0. We thus apply the backward induction in order to maximize Equation (35), or equivalently, Equation (36).

[25] An extension to an arbitrarily-finite-horizon model is executed in Appendix B.

According to the backward induction technique, we first concentrate on the maximization in the second period *after* the states both in periods 1 and 2 were realized and observed by the worker and *after* her period 1's action, which is denoted by a_1 and may be dependent on period 1's state, has been already chosen in period 1.

Note here that the worker's available actions in period 1, a_1, is either "stop the search" or "continue it," and that while her available actions in period 2, denoted by a_2, is also either "stop" or "continue," an opportunity for her to take action arises only when $a_1 =$ "continue."

The *value function in the second period*, denoted by V_2, is defined as the maximized value of the worker's (current-value) second-period's income, y_2, when the second-period's action, a_2, is optimally chosen by her, given that the first- and second-period states is (s_1, s_2) and that the worker actually took her action a_1 in the first period. Then, it follows that

$$V_2(s_1, s_2)|_{a_1=\text{“stop”}} = w_{s_1} \tag{37}$$

and

$$V_2(s_1, s_2)|_{a_1=\text{“continue”}} = \begin{cases} w_b & \text{if} \quad s_2 = b \\ w_s \vee c & \text{if} \quad s_2 = s\,. \end{cases} \tag{38}$$

In Equation (37), w_{s_1} denotes the wage offer when the state in period 1 is $s_1 \in \{b, s\}$ and it means that the worker accepted the wage offer which was predetermined in period 1 because she stopped the search in the first period (and she will have no opportunity to take a new action in period 2).

For Equation (38), she needs to contemplate which action she will take in the second period and she should conclude to accept the offer when $s_2 = b$ by the assumption (34), while, if otherwise, she decides whether to accept the offer or to decline it depending on whether $w_s > c$ (receiving w_s by acceptance) or not (receiving c by declination).

So far, we solved the second-period maximization problem, and hence, the right-hand side of Equation (36) may be modified to

$$\min_{p' \in \{p^0\}^{seq\varepsilon}|_1} E^{p'} \left[y_1(s_1) + \beta \min_{p'' \in \{p^0\}^{seq\varepsilon}|_E} E^{p''} \left[V_2(s_1, \cdot)|_{a_1} \right] \right] . \tag{39}$$

Our next task is to compute the "minimum" of the (many) expected values of V_2. However, if $a_1 =$ "stop", V_2 is constant with respect to s_2 (see Equation (37)), and hence, in such a case, Equation (39) is immediately simplified to

$$\min_{p' \in \{p^0\}^{seq\varepsilon}|_1} E^{p'} \left[y_1(s_1) + \beta w_{s_1} \right] . \tag{40}$$

Thus, for a while, we consider the case where $a_1 =$ "continue". We then obtain the next result.

Proposition 12 *Suppose that Assumption 1 is satisfied. Then, it holds that*

$$\min_{p'' \in \{p^0\}^{seq\varepsilon}|_E} E^{p''}\left[V_2(b,\cdot)|_{a_1=\text{“continue”}}\right] = \frac{(1-\varepsilon)p^0_{b,b}}{(1-\varepsilon)p^0_b+\varepsilon}w_b + \frac{(1-\varepsilon)p^0_{b,s}+\varepsilon}{(1-\varepsilon)p^0_b+\varepsilon}(w_s \vee c) \quad \textit{and} \tag{41}$$

$$\min_{p'' \in \{p^0\}^{seq\varepsilon}|_E} E^{p''}\left[V_2(s,\cdot)|_{a_1=\text{“continue”}}\right] = \frac{(1-\varepsilon)p^0_{s,b}}{(1-\varepsilon)p^0_s+\varepsilon}w_b + \frac{(1-\varepsilon)p^0_{s,s}+\varepsilon}{(1-\varepsilon)p^0_s+\varepsilon}(w_s \vee c)\,. \tag{42}$$

(Proof B.8 in Appendix A)

4.3 The First-Period Value Function

We are now ready to derive the *value function in the first period*, V_1, which is defined as the maximized value of the worker's lifetime income given that the first-period state is realized and observed by the worker and that the worker has optimally chosen her first-period action based on her observation. In other words, it is the maximized value of the formula inside the outermost brackets in Equation (39) given the first-period state.

Importantly, the backward induction guarantees that it is taken for granted that the worker's action in the *second* period has been chosen optimally.

We first consider the case where $s_1 = b$. In this case, the *pre*-maximized value of the worker's lifetime income which depends on the worker's first-period action, a_1, is summarized as follows by Proposition 12 and by the analyses we conducted so far under Assumption 1:

$$\begin{cases} w_b + \beta w_b & \text{if} \quad a_1 = \text{“stop”} \\ c + \beta\frac{(1-\varepsilon)p^0_{b,b}}{(1-\varepsilon)p^0_b+\varepsilon}w_b + \beta\frac{(1-\varepsilon)p^0_{b,s}+\varepsilon}{(1-\varepsilon)p^0_b+\varepsilon}(w_s \vee c) & \text{if} \quad a_1 = \text{“continue”} \end{cases} \tag{43}$$

Here, the first line is obtained from Equation (40) because $y_1(b) = w_b$ when the offer is accepted. Also, the second line is obtained from Equation (41) by multiplying it by β and by adding c to it.

It is easy to see that the first line is greater than the second line by the assumption (34). In particular, note that the sum of the second and third terms is less than βw_b because it is β times a mean of w_b and a smaller number than w_b.

We thus conclude that

$$V_1(b) = w_b + \beta w_b \tag{44}$$

because the worker's best action in period 1 is an acceptance of the wage offer when $s_1 = b$ as we verified in the previous paragraph.[26]

We now turn to the case where $s_1 = s$. Very similarly to the above case, under Assumption 1, the *pre*-maximized value of the worker's lifetime income is given by

$$\begin{cases} w_s + \beta w_s & \text{if} \quad a_1 = \text{“stop”} \\ c + \beta \dfrac{(1-\varepsilon)p^0_{s,b}}{(1-\varepsilon)p^0_s + \varepsilon} w_b + \beta \dfrac{(1-\varepsilon)p^0_{s,s} + \varepsilon}{(1-\varepsilon)p^0_s + \varepsilon} (w_s \vee c) & \text{if} \quad a_1 = \text{“continue”} \end{cases} \tag{45}$$

Here, the first line is obtained from Equation (40) because $y_1(s) = w_s$ when the wage offer is accepted. Also, the second line is obtained from Equation (42) as before.

In contrast to the case where $s_1 = b$, the relative relation between the magnitude of the first line and that of the second line of Equations (45) is now *indeterminate*. If it is the case that $c > w_s$, the value of the second line always dominates that of the first line because, ignoring β, the sum of the second and third terms in the second line there is a mean of two numbers both of which are greater than w_s. If otherwise (*i.e.*, if $c < w_s$), however, the value of the first line may dominate that of the second line, say, when w_s is much larger than c, $p^0_{s,b}$ is very close to zero, and/or ε is very close to unity. Thus, the relative size between the two lines hinges upon the configuration of the parameters.

In sum, the value function in the first period when $s_1 = s$ is given by

$$V_1(s) = \max \left\{ w_s + \beta w_s, c + \beta \frac{(1-\varepsilon)p^0_{s,b}}{(1-\varepsilon)p^0_s + \varepsilon} w_b + \beta \frac{(1-\varepsilon)p^0_{s,s} + \varepsilon}{(1-\varepsilon)p^0_s + \varepsilon} (w_s \vee c) \right\} \tag{46}$$

under Assumption 1. Here, the optimal action in the first period is determined so as to choose the maximum between the two terms: accept the wage offer if the former term is the larger and decline it if otherwise.

Finally, the "minimum" of the (many) expected values of the lifetime income, which we call the *value* of the job search as we did in Section 2 and which we denote by V_0, is interpreted as a potential gain the worker can exploit in the framework of our job search model *before* the first-period state is revealed and if we suppose that the worker never makes mistakes in their future decision-making.

In other words, V_0 is expressed by

$$V_0 = \min_{p' \in \{p^0\}^{seq\varepsilon}|_1} E^{p'} \left[V_1(s_1) \right] ,$$

where $V_1(s_1)$ are defined by Equations (44) and (46).

[26] In fact, in order to derive Equation (44), we do not need Assumption 1.

Furthermore, if we assume that

$$p_b^0 < \frac{1}{2(1-\varepsilon)}, \tag{47}$$

V_0 can be explicitly computed as[27]

$$V_0 = (1-\varepsilon)p_b^0 V_1(b) + \left((1-\varepsilon)p_s^0 + \varepsilon\right) V_1(s), \tag{48}$$

because it can be easily seen that $V_1(b)$ always dominates $V_1(s)$ by (34) and because Inequality (47) is equivalent to $(1-\varepsilon)p_b^0 < (1-\varepsilon)p_s^0 + \varepsilon$. For example, Inequality (47) always holds if $\varepsilon > 1/2$.

At last, we have "almost" solved the job search model with the sequential $\boldsymbol{\varepsilon}$-contamination. we say "almost" because we have not completely specified the worker's optimal strategy when the first-period state is s. Whether the worker stops the search or not hinges on the configuration of the model's parameters (see Equation (46)).

4.4 Comparative Statics

We close this section by analyzing the effect on the timing when the uncertainty-averse worker stops the search that is caused by an increase in ε (*i.e.* an increase in uncertainty).

When the state observed in the first period is b, the worker always stops the search regardless of the value of ε (see Equation (44)). Therefore, an increase in ε does not cause any effect on the worker's behavior.

In contrast, however, if period 1's state happens to be s, an increase in ε may affect the worker's search behavior. To see this, define $\pi(\varepsilon)$ as the coefficient of w_b in the second element of the right-hand side of Equation (46), ignoring β. That is, let

$$\pi(\varepsilon) := \frac{(1-\varepsilon)p_{s,b}^0}{(1-\varepsilon)p_s^0 + \varepsilon}.$$

It is easy to see that $\pi(\varepsilon)$ is in effect the probability of b occurring in the second period conditional on s occurring in the first period.

Then, it immediately follows that

$$\frac{d\pi(\varepsilon)}{d\varepsilon} = -\frac{p_{s,b}^0}{((1-\varepsilon)p_s^0 + \varepsilon)^2} < 0,$$

which shows that an increase in ε always *decreases* the value of choosing "continue" in the first period after observing s. Also, recall that Assumption

[27] The following argument does not require Assumption 1, and hence, if we would be satisfied with not knowing the exact formula of $V_1(s)$, the formula (48) would be correct as it is.

1 keeps being satisfied after an increase in ε (see the remark right after the statement of Assumption 1). All this shows that an increase in ε may urge the unemployed worker to stop the job search and to make her income stream determinate by lowering the reward she would get when she continues the search. In particular, it is never the case that the worker who has decided to stop search upon observing s changes her initial decision into continuing the search when she becomes more pessimistic in the sense that ε increases.

This is summarized by the next proposition.[28]

Proposition 13 *Suppose that Assumption 1 is satisfied and also suppose that the unemployed worker observed state s in the first period. Then, an increase in ε may discourage the worker's behavior of continuing the job search and of drawing a new wage offer in the second period, while that increase in ε never encourages more search.*

In a similar framework to this paper, Nishimura and Ozaki (2004) studied a job search model with ambiguity where an uncertainty-averse unemployed worker seeks to maximize her lifetime income in an infinite-horizon framework, while we consider a two-period finite-horizon model.[29]

Another important difference between theirs and ours is that they assume that a worker faces a stationary set of priors that does not change over time, while a worker in this paper updates a set of priors according to Bayes's rule after making observations of states.[30]

In spite of such a slight difference, this paper's comparative statics result is the same as that of Nishimura and Ozaki (2004). That is, we showed an increase of ambiguity (dilation of a set of probability charges or an increase in ε) motivates the worker to stop the search earlier. In other words, this paper shows that the result by Nishimura and Ozaki (2004) is robust in the sense that basically the same result holds even if the worker updates her belief by Bayes' rule, though the analyses here are confined within the class of the ε-contamination with updating suitably adapted so as to retain time consistency according to the sequential $\boldsymbol{\varepsilon}$-contamination developed in this paper.

Intuitively, this is plausible because Bayesian updating dilates ambiguity (see the last inequality in Proposition 10) and the uncertainty-averse worker hates successively lower wage offer she might draw by waiting longer.

APPENDICES

[28]In order to derive only the comparative statics result without knowing the exact form of the value function, Assumption 1 is not necessary. See Footnote 39.

[29]See, however, Footnote 25.

[30]As another important difference, Nishimura and Ozaki (2004) assume that the worker's preference is defined by any stationary (thus, constant) set of priors which can be also represented by the core of a convex capacity, which includes as a special case the atemporal ε-contamination adopted to the dynamic recursive framework.

A Derivations of the Claims Made in Section 2

A.1 Proof of Proposition 1

A few lines of computations will reveal that (9) < (10) holds if and only if the following inequality holds:

$$w_s + \bar{w} + \beta w_s < c + \bar{w} + \beta \left(\frac{1-\varepsilon}{2(1+\varepsilon)} w_b + \frac{1+3\varepsilon}{2(1+\varepsilon)} w_s \right) . \tag{49}$$

However, the inequality (49) right above is always true because of both the second half of the assumption (2) and the assumption (4). □

A.2 Derivation of Formula (13)

For the first maximization in (12), the first out of the two terms is always larger than the second by the assumption (2).

For the second maximization there, it is easy to see that the second out of the two terms is larger than the first if and only if

$$w_s + \bar{w} + \beta w_s < c + \bar{w} + \beta \left(\frac{1-\varepsilon}{2(1+\varepsilon)} w_b + \frac{1+3\varepsilon}{2(1+\varepsilon)} w_s \right) .$$

However, this inequality is identical with the inequality (49) that appeared in A.1 right above, and the inequality (49) always holds true under the maintained assumptions as is said there. Thus, the claim follows. □

A.3 Derivation of Formula (18)

To find the weights appearing in (18), we apply a linear programming method.

Note that the second largest lifetime income, $w_b + \beta w_b$, takes place when the first-period state is b. Let the first-period marginal probability charge of b be $1/2 + x$, where x is a free variable and must be in the range, $[-\varepsilon/2, \varepsilon/2]$, by the definition of the bounds, (15). Given x, the (joint) probability charge of (s, b), at which the largest lifetime income, $c + \bar{w} + \beta w_b$, is realized, will be minimized at

$$\left(\frac{1}{2} - x\right) \cdot \left(\frac{1}{2} + \underline{\varepsilon}'\right) = \left(\frac{1}{2} - x\right) \cdot \left(\frac{1}{2} - \frac{\varepsilon}{1+\varepsilon}\right) = \left(\frac{1}{2} - x\right) \frac{1-\varepsilon}{2(1+\varepsilon)} ,$$

and the (joint) probability charge of (s, s), at which the smallest lifetime income, $c + \bar{w} + \beta w_s$, is realized, will be maximized at

$$\left(\frac{1}{2} - x\right) \cdot \left(\frac{1}{2} + \bar{\varepsilon}'\right) = \left(\frac{1}{2} - x\right) \cdot \left(\frac{1}{2} + \frac{\varepsilon}{1+\varepsilon}\right) = \left(\frac{1}{2} - x\right) \frac{1+3\varepsilon}{2(1+\varepsilon)} ,$$

where the first term in each product in each line is given as such because it is the first-period's marginal of s, and the second terms of the products are given as such by the definition of the sequential $\boldsymbol{\varepsilon}$-contamination, (16).

We then define a linear programming problem by

$$\min_{x\in[-\varepsilon/2,\varepsilon/2]} \left[\left(\frac{1}{2}+x\right)(w_b+\beta w_b) + \left(\frac{1}{2}-x\right)\frac{1-\varepsilon}{2(1+\varepsilon)}(c+\bar{w}+\beta w_b) \right. \\ \left. + \left(\frac{1}{2}-x\right)\frac{1+3\varepsilon}{2(1+\varepsilon)}(c+\bar{w}+\beta w_s) \right]. \quad (50)$$

By the ranking of the value of each lifetime income and the preliminary consideration made in the previous paragraph, the solution to (50), which is denoted x^* and takes place inevitably at a "corner," achieves the minimal expected lifetime income.

Also, we know that if the derivative of the objective function of (50) with respect to x is positive, then $x^* = -\varepsilon/2$; and if otherwise, then $x^* = \varepsilon/2$, because it is a minimization problem.

However, it is immediately follows that the derivative turns out to be negative by the assumption (4), and thus, we conclude that $x^* = \varepsilon/2$. Substituting this back into (50), the worker's minimal expected lifetime income given this strategy is determined by

$$\frac{1+\varepsilon}{2}(w_b+\beta w_b) + \frac{(1-\varepsilon)^2}{4(1+\varepsilon)}(c+\bar{w}+\beta w_b) \\ + \frac{(1-\varepsilon)(1+3\varepsilon)}{4(1+\varepsilon)}(c+\bar{w}+\beta w_s) ,$$

which is equal to (18), and thus, the claim is verified. □

A.4 Proof of Proposition 3

Some tedious but straightforward computations show that (17) < (18) holds if and only if

$$\frac{2\varepsilon}{1-\varepsilon}(w_b+\beta w_b) + \frac{1+\varepsilon}{1-\varepsilon}(w_s+\bar{w}+\beta w_s) \\ < c+\bar{w}+\beta\left(\frac{(1-\varepsilon)}{2(1+\varepsilon)}w_b + \frac{1+3\varepsilon}{2(1+\varepsilon)}\right). \quad (51)$$

But, by the second halves of both the assumptions (2) and (4), we can show that (51) holds true if the following inequality is satisfied:

$$\frac{2\varepsilon}{1-\varepsilon}(w_b+\beta w_b) - \frac{1+\varepsilon}{1-\varepsilon}(w_s+\bar{w}+\beta w_s) < w_b+\beta w_b\,. \quad (52)$$

It follows immediately that the inequality (52) is equivalent to

$$(1-3\varepsilon)(w_b+\beta w_b) + (1+\varepsilon)(w_s+\bar{w}+\beta w_s) > 0\,,$$

which is the very assumption we imposed in the first half of the assumption (4). We, thus, conclude that the worker's maxmin-expected lifetime income is given by (18). □

B A Lemma and Proofs

B.1 Proof of Proposition 5

First, because a set of probability charges with the common finite support can be identified as a subset of the finite-dimensional Euclidean space, the weak * compactness of $\mathscr{P}$ implies that so are $\mathscr{P}|_1$ and $\mathscr{P}|_E$.

Second, given any $p = (p_{i,j})_{i,j} \in \mathscr{P}$, note that (22) implies that p can be written as $p = (p'_i \cdot p''_{ij})_{i,j}$, where $(p'_i)_i \in \mathscr{P}|_1$ is the list of the first-period marginals of p and $(p''_{ij})_j \in \mathscr{P}|_E$ is the list of well-defined conditionals when $E = E_i$ is observed in the first period.

Third, by the law of iterated expectations, and by the remark made right before the statement of the proposition, we obtain

$$E^p [u_2] = E^{p'} \left[E^{p''} [u_2] \right] = E^{p'} \left[E^{p''} [u_2 \,|\langle E_i \times S \rangle_i] \,(s_1, s_2) \right] ,$$

where the outer expectations of the middle and right terms aggregate with respect to s_1. Therefore, we obtain

$$E^p [u_1 + u_2] = E^{p'} \left[u_1 + E^{p''} [u_2] \right] .$$

Fourth, we show that "$\geq$" holds in (26). The equality in the previous paragraph immediately implies that for any $p \in \mathscr{P}$,

$$E^p [u_1 + u_2] \geq \min_{p' \in \mathscr{P}|_1} E^{p'} \left[u_1 + \min_{p'' \in \mathscr{P}|_E} E^{p''} [u_2] \right] ,$$

which proves the claim. We remark that we did not use the rectangularity of $\mathscr{P}$. Thus, this direction of the equality always holds.

Fifth and Finally, we show "$\leq$" holds in (26). To this end, on the contrary, assume that $>$ holds there. By the compactness of the relevant sets, there exist $p'^* = (p'^*_i)_i \in \mathscr{P}|_1$ and $p''^* = (p''^*_{ij})_j \in \mathscr{P}|_E$ that attain the minima in the right-hand side of (26). *By the rectangularity of* $\mathscr{P}$, $p^* := (p^*_{i,j})_{i,j} := (p'^*_i \times p''^*_{ij})_{i,j}$ must be contained by $\mathscr{P}$. Thus,

$$\begin{aligned} \min_{p \in \mathscr{P}} E^p [u_1 + u_2] &> \min_{p' \in \mathscr{P}|_1} E^{p'} \left[u_1 + \min_{p'' \in \mathscr{P}|_E} E^{p''} [u_2] \right] \\ &= E^{p'^*} \left[u_1 + E^{p''^*} [u_2] \right] \\ &= E^{p^*} [u_1 + u_2] \\ &\geq \min_{p \in \mathscr{P}} E^p [u_1 + u_2] , \end{aligned}$$

where we invoked the law of iterated expectations again. This is a contradiction we desire. □

B.2 Proof of Proposition 6

First, observe that, for any $p \in \{p^0\}^{rec\varepsilon}$ and for any i, j, $p_i = p_i^0 + \varepsilon_i$ and $p_{ij} = p_{ij}^0 + \varepsilon_{ij}$ because $(\forall i)\ \sum_j \varepsilon_{ij} = 0$ by assumption.

Second, to complete the proof, let $p', p'' \in \{p^0\}^{rec\varepsilon}$. Then, p' can be written as $\Big((p_i^0 + \varepsilon_i')(p_{ij}^0 + \varepsilon_{ij}')\Big)_{i,j}$ for some $(\varepsilon_i')_i$ and $(\varepsilon_{ij}')_{ij}$ satisfying (31), and p'' can be so as $\Big((p_i^0 + \varepsilon_i'')(p_{ij}^0 + \varepsilon_{ij}'')\Big)_{i,j}$ for some $(\varepsilon_i'')_i$ and $(\varepsilon_{ij}'')_{ij}$ satisfying (31). Then, from the first paragraph, we conclude that

$$(p_i' \cdot p_{ij}'')_{i,j} = \big((p_i^0 + \varepsilon_i')(p_{ij}^0 + \varepsilon_{ij}'')\big)_{i,j}$$

holds. Because $(\varepsilon_i')_i$ and $(\varepsilon_{ij}'')_{ij}$ satisfy all the requirements in (31), $(p_i' \cdot p_{ij}'')_{i,j} \in \mathscr{P}$ and the proof is complete. □

B.3 Proof of Proposition 7

First note that

$$\{p^0\}^{rec\varepsilon}\big|_1 = \left\{ \left(p_i^0 + \varepsilon_i\right)_i \,\middle|\, (\forall i)\, \varepsilon_i \in [\underline{\varepsilon}_i, \bar{\varepsilon}_i] \text{ and } \textstyle\sum_i \varepsilon_i = 0 \right\}.$$

Furthermore, it is easy to see that

$$\{p^0\}^{\varepsilon}\big|_1 = \{p^0\big|_1\}^{\varepsilon}.$$

By (28), we know that $p^0\big|_1 = \Big(p_i^0 + \sum_j \delta_{i,j}\Big)_i$. But, from (29) and the first two definitions in (30), it follows that $(\forall i)\ \sum_j \underline{\delta}_{i,j} = -\varepsilon p_i^0 = \underline{\varepsilon}_i$ and that $(\forall i)\ \sum_j \bar{\delta}_{i,j} = \varepsilon(1 - p_i^0) = \bar{\varepsilon}_i$, which completes the proof. □

B.4 Proof of Proposition 8

Let $p \in \{p^0\}^{\varepsilon}$ and write it as $p = (p_{i,j}^0 + \delta_{i,j})_{i,j}$ with some $(\delta_{i,j})_{i,j}$ that satisfies the requirements stated in (28).

By definition, it follows that $p|_1 \in \{p^0\}^{\varepsilon}\big|_1$. This, (28), and (29) mean that we can write $p|_1$ as $p|_1 = \big(p_i^0 + \delta_i'\big)_i$ with some $(\delta_i')_i$ such that $(\forall i)\ -\varepsilon p_i^0 \le \delta_i' \le \varepsilon(1 - p_i^0)$ and $\sum_i \delta_i' = 0$. That is, let $(\forall i)\ \delta_i' := \sum_j \delta_{i,j}$.

For each i, define ε_i by $\varepsilon_i := \delta_i'$. Then, by definition and the previous paragraph, it holds that $\sum_i \varepsilon_i = 0$, that $\underline{\varepsilon}_i = -\varepsilon p_i^0 \le \delta_i' = \varepsilon_i$ and that $\bar{\varepsilon}_i = \varepsilon(1 - p_i^0) \ge \delta_i' = \varepsilon_i$, where $(\forall i)\ \underline{\varepsilon}_i$ and $\bar{\varepsilon}_i$ are defined by (30). Thus, all the requirements for ε_i in (31) are now met.

Next, for each i, j, define ε_{ij} by $\varepsilon_{ij} := (\delta_{i,j} - \delta_i' p_{ij}^0)/(p_i^0 + \delta_i')$. Then, $(\forall i)\ \sum_j \varepsilon_{ij} = (\sum_j \delta_{i,j} - \delta_i' \sum_j p_{ij}^0)/(p_i^0 + \delta_i') = (\delta_i' - \delta_i')/(p_i^0 + \delta_i') = 0$, where we used the definition of δ'_i and the fact that $(\forall i)$ p^0_{ij} is a (conditional) charge.

Let $\underline{\varepsilon}_{ij}$ and $\bar{\varepsilon}_{ij}$ be as defined in (30). First, we show that $(\forall i,j)$ $\varepsilon_{ij} \geq \underline{\varepsilon}_{ij}$. To this end, note that $\partial\varepsilon_{ij}/\partial\delta'_i = (-p^0_{i,j} - \delta_{i,j})/(p^0_i + \delta'_i)^2 < 0$, where the numerator must be negative because $-p^0_{i,j} - \delta_{i,j} \leq -p^0_{i,j} - (-\varepsilon p^0_{i,j}) = (\varepsilon - 1)p^0_{i,j} < 0$ since $\delta_{i,j} \geq -\varepsilon p^0_{i,j}$ and $\varepsilon < 1$. Therefore, ε_{ij} attains its lower bound when δ'_i attains its upper bound. We thus obtain $(\forall i,j)$

$$\begin{aligned}
\varepsilon_{ij} &\geq \frac{-\varepsilon p^0_{i,j} - \varepsilon(1-p^0_i)p^0_{ij}}{p^0_i + \varepsilon(1-p^0_i)} \\
&= \frac{-\varepsilon p^0_{i,j} - \varepsilon p^0_{ij} + \varepsilon p^0_{i,j}}{p^0_i + \varepsilon(1-p^0_i)} \\
&= \frac{-\varepsilon p^0_{ij}}{(1-\varepsilon)p^0_i + \varepsilon} = \underline{\varepsilon}_{ij}\,.
\end{aligned}$$

Second, we show that $(\forall i,j)$ $\varepsilon_{ij} \leq \bar{\varepsilon}_{ij}$, note that $\varepsilon_{ij} = (\delta_{i,j} - \sum_\ell \delta_{i,\ell}\, p^0_{ij})/(p^0_i + \sum_\ell \delta_{i,\ell})$, and hence that $\partial\varepsilon_{ij}/\partial\delta_{i,j} = \sum_{\ell\neq j}(p^0_{i,\ell} + \delta_{i,\ell})/(p^0_i + \sum_\ell \delta_{i,\ell})^2 > 0$.[31] Therefore, ε_{ij} attains its maximum when $\delta_{i,j}$ is maximal, that is, when $\delta_{i,j} = \varepsilon(1-p^0_{i,j})$. However, this occurs precisely only when $\delta_{i,\ell} = -\varepsilon p^0_{i,\ell}$ for $\ell \neq j$ because if otherwise, the two requirements that $\sum_{i,j}\delta_{i,j} = 0$ and that $(\forall i,j)$ $\delta_{i,j} \geq -\varepsilon p^0_{i,j}$ cannot be satisfied simultaneously. Therefore, at this time, it holds that $\sum_\ell \delta_{i,\ell} = \varepsilon(1-p^0_i)$, and we obtain

$$\begin{aligned}
\varepsilon_{ij} &\leq \frac{\varepsilon(1-p^0_{i,j}) - \varepsilon(1-p^0_i)p^0_{ij}}{p^0_i + \varepsilon(1-p^0_i)} \\
&= \frac{\varepsilon(1-p^0_{ij})}{(1-\varepsilon)p^0_i + \varepsilon} = \bar{\varepsilon}_{ij}\,.
\end{aligned}$$

So far, we have found $(\varepsilon_i)_i$ and $(\varepsilon_{ij})_{i,j}$ that satisfies (31) and (30).

Finally, it suffices to verify that for these $(\varepsilon_i)_i$ and $(\varepsilon_{ij})_{i,j}$, it holds that $p^0_{i,j} + \delta_{i,j} = (p^0_i + \varepsilon_i)(p^0_{ij} + \varepsilon_{ij})$ for each i and j. But, this is immediate from the definitions of ε_i and ε_{ij}: $(\forall i,j)$

$$\begin{aligned}
& (p^0_i + \varepsilon_i)(p^0_{ij} + \varepsilon_{ij}) \\
=\ & p^0_{i,j} + \varepsilon_i p^0_{ij} + (p^0_i + \varepsilon_i)\varepsilon_{ij} \\
=\ & p^0_{i,j} + \delta'_i p^0_{ij} + (p^0_i + \delta'_i)\frac{\delta_{i,j} - \delta'_i p^0_{ij}}{p^0_i + \delta'_i} \\
=\ & p^0_{i,j} + \delta'_i p^0_{ij} + \delta_{i,j} - \delta'_i p^0_{ij}
\end{aligned}$$

[31] We may assume the strict positivity here because the numerator being zero will take place only when $p_{i,j} = 1$ for some i and j, which implies, together with the assumption that $p \in \{p^0\}^\varepsilon$, that $p^0 = (0,\ldots,0,1,0,\ldots,0)$, meaning that p^0 represents no risk, which we excluded throughout this paper.

$$= \quad p^0_{i,j} + \delta_{i,j}\,,$$

which completes the proof. □

B.5 Proof of Proposition 9

Because $\{p^0\}^{seq\varepsilon}$ is rectangular (Proposition 6), it suffices to prove that any rectangular set including $\{p^0\}^{\varepsilon}$ contains $\{p^0\}^{seq\varepsilon}$.

Let $p^0 \in \mathscr{M}(\Omega, \mathcal{F}_2)$, let $\varepsilon \in (0,1)$, and let i and j be arbitrarily fixed below.

First, note that there exists $p' \in \{p^0\}^{\varepsilon}$ such that $p'_i = (1-\varepsilon)p^0_i + \varepsilon$, which is the maximum value the first-period marginal, p_i, can assume subject to $p \in \{p^0\}^{\varepsilon}$. To do this, we can let $(q_{i,j})_{i,j}$ be such that $\sum_\ell q_{i,\ell} = 1$.

Second, note that there exists $p'' \in \{p^0\}^{\varepsilon}$ such that

$$p''_{ij} = \frac{(1-\varepsilon)p^0_{i,j} + \varepsilon}{(1-\varepsilon)p^0_i + \varepsilon}\,,$$

which is the maximum value the conditional, p_{ij}, can assume subject to $p \in \{p^0\}^{\varepsilon}$. To do this, we can let $(q_{i,j})_{i,j}$ be such that $q_{i,j} = 1$.

The above two paragraphs show that $p'_i \cdot p''_{ij} = (1-\varepsilon)p^0_{i,j} + \varepsilon$ is a value $p_{i,j}$ can assume as long as p is an element of any rectangular set that contains $\{p^0\}^{\varepsilon}$. Here, some computations exhibit

$$\begin{aligned} p'_i \cdot p''_{ij} &= (1-\varepsilon)p^0_{i,j} + \varepsilon = \left(p^0_i + \varepsilon(1-p^0_i)\right)\left(p^0_{ij} + \frac{\varepsilon(1-p^0_{ij})}{(1-\varepsilon)p^0_i + \varepsilon}\right) \\ &= (p^0_i + \bar{\varepsilon}_i)(p^0_{ij} + \bar{\varepsilon}_{ij})\,, \end{aligned}$$

where $\bar{\varepsilon}_i$ and $\bar{\varepsilon}_{ij}$ are defined by (30).

Similarly, note that there exists $p''' \in \{p^0\}^{\varepsilon}$ such that $p'''_i = (1-\varepsilon)p^0_i$, which is the minimum value the first-period marginal, p_i, can assume subject to $p \in \{p^0\}^{\varepsilon}$. (Let $(q_{i,j})_{i,j}$ be such that $(\forall \ell)\, q_{i,\ell} = 0$.) Also note that there exists $p'''' \in \{p^0\}^{\varepsilon}$ such that

$$p''''_{ij} = \frac{(1-\varepsilon)p^0_{i,j}}{(1-\varepsilon)p^0_i + \varepsilon}\,,$$

which is the minimum value the conditional, p_{ij}, can assume subject to $p \in \{p^0\}^{\varepsilon}$. (Let $(q_{i,j})_{i,j}$ be such that $\sum_{\ell \neq j} q_{i,\ell} = 1$.)

Therefore, by a similar reasoning as above, $p'''_i \cdot p''''_{ij}$ is a value $p_{i,j}$ can assume as long as p is an element of any rectangular set that contains $\{p^0\}^{\varepsilon}$. Here, some computations exhibit

$$p'''_i \cdot p''''_{ij} \quad = \quad (1-\varepsilon)p^0_i \cdot \frac{(1-\varepsilon)p^0_{i,j}}{(1-\varepsilon)p^0_i + \varepsilon} = (1-\varepsilon)p^0_i \cdot \left(p^0_{ij} - \frac{\varepsilon p^0_{ij}}{(1-\varepsilon)p^0_i + \varepsilon}\right)$$

$$= \quad (p_i^0 + \underline{\varepsilon}_i)(p_{ij}^0 + \underline{\varepsilon}_{ij})\,,$$

where $\underline{\varepsilon}_i$ and $\underline{\varepsilon}_{ij}$ are defined by (30).

By the first paragraph, we know that $\mathrm{rect}\big(\{p^0\}^\varepsilon\big) \subseteq \{p^0\}^{seq\varepsilon}$. Furthermore, the arguments so far show that we can always find a probability charge in any rectangular set containing $\{p^0\}^\varepsilon$ that achieves the "upper rim" of $\{p^0\}^{seq\varepsilon}$ for arbitrary i and j and (possibly) another probability charge in such a set that achieves the "lower rim" of $\{p^0\}^{seq\varepsilon}$ for arbitrary i and j. This fact proves that both sets are identical. □

B.6 Proof of Proposition 10

First, note that we have $(\forall i)\ (1-\varepsilon)p_i^0 + \varepsilon q_i = p_i^0 + \varepsilon(q_i - p_i^0)$. Then, it is immediate that $\varepsilon_i := \varepsilon(q_i - p_i^0)$ satisfies all the requirements in (31) and (30) by (32).

Second, note that $(\forall i,j)\ (1-\varepsilon_i')p_{ij}^0 + \varepsilon_i' q_{ij} = p_{ij}^0 + \varepsilon_i'(q_{ij} - p_{ij}^0)$. Then, it is immediate that $\varepsilon_{ij} := \varepsilon_i'(q_{ij} - p_{ij}^0)$ satisfies all the requirements in (31) and (30) by (32) and the definition of ε_i'.

Finally, the last claim follows from the fact that $(1-\varepsilon)p_i^0 + \varepsilon < 1$ for each i, which holds because $(\forall i)\ p_i^0 < 1$ since we assume that $(\forall i)\ p_i^0 > 0$ (see the start of Subsection 3.1.2). □

B.7 Lemma

Assume that $p^0 = p^1 \otimes p^1$ with some probability charge p^1 on S. Then, the two inequalities in Assumption 1 is satisfied whenever $\varepsilon > 1/2$.[32]

Proof First, note that all of p_{bb}^0, p_{sb}^0, p_b^0 are now equal to p_b^1, where p_b^1 is simply a probability charge of b because p^0 is now the product of p^1.

Then, the first inequality in Assumption 1 turns out to be

$$\big(p_b^1\big)^2 - \frac{1}{2}p_b^1 - \frac{\varepsilon}{2(1-\varepsilon)} < 0\,.$$

If we solve this by replacing the inequality by the equality, we obtain

$$p_b^1 = \frac{1}{4} \pm \frac{1}{2}\sqrt{\frac{1}{4} + \frac{2\varepsilon}{1-\varepsilon}}\,.$$

Thus, the desired inequality holds true if p_b^1 lies between the smaller and larger solutions of this quadratic equation. Firstly, because $\varepsilon > 0$, its smaller solution is always negative and p_b^1 is obviously larger than this. Secondly, if

[32] The learning procedure does not take place when the probability charge is defined as a product one and our main concern in this section is consequences of an agent's updating behavior. In this sense, the lemma should be understood as providing purely theoretical information about the content of Assumption 1.

$\varepsilon = 1/2$, the larger solution is unity and the larger solution increases when ε increases, which shows that p_b^1 is less than the larger solution of the above quadratic equation when $\varepsilon > 1/2$. We are done about the first inequality in Assumption 1.

Next, we do the same thing for the second inequality in Assumption 1 and get the quadratic equation:

$$\left(p_b^1\right)^2 - \frac{3}{2}p_b^1 - \frac{1}{2(1-\varepsilon)} > 0$$

and its smaller and larger solutions as

$$p_b^1 = \frac{3}{4} \pm \frac{1}{2}\sqrt{\frac{9}{4} - \frac{2}{1-\varepsilon}}\,.$$

It turn out that the inside of the square-root is negative when $\varepsilon > 1/9$, verifying that the quadratic equation is always above zero. Because $\varepsilon > 1/2$ implies this situation, we are done also about the other inequality. □

B.8 Proof of Proposition 12

We prove only Equation (41). A symmetric reasoning also applies to show Equation (42).

By Assumption 1, we have

$$p_{bb}^0 < \frac{1}{2} + \frac{\varepsilon}{2(1-\varepsilon)p_b^0}\,.$$

By noting the decomposition of a joint probability charge into its marginal and conditional, $p_{s,s}^0 = p_s^0 \cdot p_{ss}^0$ (see Equation (22)), it immediately follows that the above inequality is equivalent to

$$(1-\varepsilon)p_{b,b}^0 < (1-\varepsilon)p_{b,s}^0 + \varepsilon\,,$$

which, in turn, is clearly equivalent to

$$\frac{(1-\varepsilon)p_{b,b}^0}{(1-\varepsilon)p_b^0 + \varepsilon} < \frac{(1-\varepsilon)p_{b,s}^0 + \varepsilon}{(1-\varepsilon)p_b^0 + \varepsilon}\,.$$

It then turns out that the left-hand side of the last inequality is equal to $(1-\varepsilon_b')\,p_{bb}^0$ and that its right-hand side is equal to $(1-\varepsilon_b')\,p_{bs}^0 + \varepsilon_b'$, where ε_b' is defined in Proposition 10. The former corresponds to the smallest conditional in (32) (by letting $q_{bb} = 0$ there) and the latter corresponds to the largest conditional in (32) (by letting $q_{bs} = 1$ there).

Finally, an application of Proposition 10 with the assumption (34) proves that Equation (41) is the smallest conditional expectation given $s_1 = b$ and given the sequential $\boldsymbol{\varepsilon}$-contamination of p^0, (32). □

C Arbitrarily-Finite-Horizon Models

The purpose of Appendix B is to show that the results obtained for the two-period models in the main text go through as it is even in arbitrarily-finite-horizon models. We decided to relegate it to the appendix because we are afraid that its heavy notations conceal the main message of this paper, which we hope can be conveyed to the reader even only in a simple two-period setting.

C.1 (Slightly Heavy) Notations and Definitions

Let $T \in \mathbb{N}\backslash\{0,1\}$ be a length of a finite horizon, and for each $t \in \{1,2,\ldots,T\}$, $n_t\ (\geq 2)$ be a number of elements of each finite partition of S; that is, let $\langle E_{1,i_1}\rangle_{i_1=1}^{n_1}$, $\langle E_{2,i_2}\rangle_{i_2=1}^{n_2}$, … , $\langle E_{T,i_T}\rangle_{i_T=1}^{n_T}$ be finite partitions of S, which are fixed throughout the rest of the paper. We identify each of $(S, \langle E_{1,i_1}\rangle_{i_1=1}^{n_1})$, $(S^2, \langle E_{1,i_1}\times E_{2,i_2}\rangle_{i_1=1\ i_2=1}^{n_1\ \ n_2})$, … , $(S^T, \langle E_{1,i_1}\times E_{2,i_2}\times\cdots\times E_{T,i_T}\rangle_{i_1=1\ i_2=1\ \cdots\ i_T=1}^{n_1\ \ n_2\ \ \cdots\ \ n_T})$ with each of the measurable spaces, $(\Omega,\mathcal{F}_1)$, $(\Omega,\mathcal{F}_2)$, … , $(\Omega,\mathcal{F}_T)$, exactly as we did in the main text. (Note that $\Omega := S^T$.)

Let $\mathscr{M}(\Omega,\mathcal{F}_T)$ be the space of all probability charges on $(\Omega,\mathcal{F}_T)$. Given $p \in \mathscr{M}(\Omega,\mathcal{F}_T)$, we write the *joint probability charge*, $p(E_{1,i_1}\times E_{2,i_2}\times\cdots\times E_{T,i_T})$, simply as $p_{i_1,i_2,\ldots,i_T}$, where for each $t \leq T$, $E_{t,i_t} \in \langle E_{t,i_t}\rangle_{i_t=1}^{n_t}$. For each $p \in \mathscr{M}(\Omega,\mathcal{F}_T)$, the *first-t-period marginal* of p is denoted and defined by[33]

$$(\forall t)(\forall i_1,i_2,\ldots,i_t)\ \ p_{i_1\,i_2\,\ldots\,i_t} := p(E_{1,i_1}\times E_{2,i_2}\times\cdots\times E_{t,i_t}\times S\times\cdots\times S)\,.$$

Obviously, the first-T-period marginal is identical to the joint probability charge.

Next, the *one-period-ahead conditional* of p *given* $E_{1,i_1}\times E_{2,i_2}\times\cdots\times E_{t,i_t}$ is denoted and defined by

$$(\forall t \leq T-1)(\forall i_1,i_2,\ldots,i_t)\ \ p^+_{i_{t+1}|i_1\,i_2\,\ldots\,i_t} := \frac{p_{i_1\,i_2\,\ldots\,i_t\,i_{t+1}}}{p_{i_1\,i_2\,\ldots\,i_t}}\,,$$

where the probability charges in the numerator and denominator are the first-some-appropriate-period marginals defined above.[34]

For any $p \in \mathscr{M}(\Omega,\mathcal{F}_T)$, we denote its first-$t$-period marginal *charge* in $\mathscr{M}(\Omega,\mathcal{F}_t)$ (*not* a single number) by p_t,[35] as well as its one-period-ahead conditional *charge* in $\mathscr{M}(S,\langle E_{t+1,i_{t+1}}\rangle_{i_{t+1}})$ (*not* a single number) given $E_{1,i_1}\times E_{2,i_2}\times\cdots\times E_{t,i_t}$ by $p^+(\cdot|E_{1,i_1}\times E_{2,i_2}\times\cdots\times E_{t,i_t})$.

[33]When $T=2$, $p_{i_1} = p|_1(E_{1,i_1})$, where the right-hand side was introduced in the main text.

[34]When $T=2$, p^+ is well-defined only when $t=1$ and $p^+_{i_2|i_1} = p_{i_1 i_2}$, where the right-hand side was the notation we used in the main text.

[35]Note that it is only when $t=1$ that $p|_t = p_t$ holds, where the former notation appeared in the main text.

C.1.1 The Decomposition of a Probability Charge

For each $p \in \mathscr{M}(\Omega, \mathcal{F}_T)$, its decomposition into its marginal and (one-period-ahead) conditional is now represented as follows:

$$(\forall t \leq T-1) \quad p_{i_1\, i_2\, \ldots\, i_t\, i_{t+1}} = p_{i_1\, i_2\, \ldots\, i_t} \cdot p^{+}_{i_{t+1}|i_1\, i_2\, \ldots\, i_t}\,, \tag{53}$$

where the left-hand side in (53) is p's fisrt-$(t+1)$-period marginal, while its right-hand side is the product of p's fisrt-t-period marginal and its one-period-ahead marginal given the first t observations.

These decompositions (that is, Equations (53)) will be used repeatedly in what follows.

C.2 Knightian Uncertainty and Its Rectangularity

A nonempty subset $\mathscr{P}$ of $\mathscr{M}(\Omega, \mathcal{F}_T)$ is called *Knightian uncertainty*.

Given any Knightian uncertainty, $\mathscr{P}$, and any $t \leq T-1$, its *first-t-period marginal Knightian uncertainty*, denoted by $\mathscr{P}_t$, is the nonempty subset of $\mathscr{M}(\Omega, \mathcal{F}_t)$ that is defined by

$$\mathscr{P}_t := \{\, p_t \,|\, p \in \mathscr{P} \,\}\ ,$$

where p_t is the first-t-period marginal probability charge defined in C.1.[36]

Let $\mathscr{P}$ be Knightian uncertainty, $t \leq T-1$, suppose that $E_1 \times \cdots \times E_t \in \langle E_{1,i_1} \times \cdots \times E_{t,i_t} \rangle_{i_1,\ldots,i_t}$ has been observed in the first t periods, and suppose that *every* probability charge in $\mathscr{P}_{t+1}$ is updated by Bayes' rule. As a result of this procedure, we obtain the *one-period-ahead conditional* Knightian uncertanty, denoted $\mathscr{P}^{+}|_{E_1 \times \cdots \times E_t}$, which is a subset of $\mathscr{M}(S, \langle E_{t+1,i_{t+1}} \rangle_{i_{t+1}})$. That is,

$$\mathscr{P}^{+}|_{E_1 \times \cdots \times E_t} := \left\{\, p^{+}(\cdot | E_1 \times \cdots \times E_t) \,|\, p \in \mathscr{P} \,\right\}\ , \tag{54}$$

where p^{+} is the one-period-ahead probability charge defined in C.1.

Knightian uncertainty $\mathscr{P}$ is *rectangular* by definition if for any $p', p'' \in \mathscr{P}$, it holds that

$$(\forall t \leq T-1) \quad \left(p'_{i_1\, i_2\, \ldots\, i_t} \cdot p''^{+}_{i_{t+1}|i_1\, i_2\, \ldots\, i_t} \right)_{i_1, i_2, \ldots, i_t, i_{t+1}} \in \mathscr{P}_{t+1}\, .$$

Note that if Knighian uncertainty is given by a singleton set (that is, if it is a risk), then it is clearly rectangular in view of (53).

In order to state an important result, let $t \leq T$ and denote by u_t an arbitrary real-valued function on Ω that is $\mathcal{F}_t$-measurable. As we did in the main text, we denote by $E^p[u]$ the standard mathematical expectation of such a function with respect to a probability charge $p \in \mathscr{M}(\Omega, \mathcal{F}_T)$.

The next proposition states that with rectangular Knightian uncertainty, *the iterated or recursive maxmin preference, which is a dynamic extension*

[36] Note that it is only when $t = 1$ that $\mathscr{P}|_t = \mathscr{P}_t$ holds. See Footnote 35.

of the atemporal preference à la Gilboa and Schmeidler (1989), is identified with the "one-shot" maxmin preference.

Proposition 14 (Epstein-Schneider, 2003) *Let $\mathscr{P}$ be rectangular Knightian uncertainty that is weak *compact. Then,*

$$
\begin{aligned}
& \min_{p\in\mathscr{P}} E^{p}\left[\textstyle\sum_{i=1}^{T} u_i\right] \\
= \; & \min_{p'\in\mathscr{P}_1} E^{p'}\left[u_1 + \min_{p''\in\mathscr{P}^+|_{E_1}} E^{p''}\left[\textstyle\sum_{i=2}^{T} u_i\right]\right] \\
= \; & \cdots\cdots\cdots \\
= \; & \min_{p'\in\mathscr{P}_1} E^{p'}\left[u_1 + \min_{p''\in\mathscr{P}^+|_{E_1}} E^{p''}\left[u_2 + \min_{p^{(3)}\in\mathscr{P}^+|_{E_1\times E_2}} E^{p^{(3)}}\left[u_3 + \cdots \right.\right.\right. \\
& \left.\left.\left. \min_{p^{(T)}\in\mathscr{P}^+|_{E_1\times\cdots\times E_{T-1}}} E^{p^{(T)}}[u_T]\cdots\right]\right]\right],
\end{aligned}
$$

where $p^{(t)}$ abbreviates $p^{''\cdots'}$ (t primes), which is a generic probability charge relevant there.

Proof We only prove the equation:

$$
\begin{aligned}
& \min_{p\in\mathscr{P}} E^{p}\left[\textstyle\sum_{i=1}^{T} u_i\right] \\
= \; & \min_{p'\in\mathscr{P}_1} E^{p'}\left[u_1 + \min_{p''\in\mathscr{P}^+|_{E_1}} E^{p''}\left[u_2 + \min_{p^{(3)}\in\mathscr{P}^+|_{E_1\times E_2}} E^{p^{(3)}}\left[u_3 + \cdots \right.\right.\right. \\
& \left.\left.\left. \min_{p^{(T)}\in\mathscr{P}^+|_{E_1\times\cdots\times E_{T-1}}} E^{p^{(T)}}[u_T]\cdots\right]\right]\right].
\end{aligned}
$$

The other equations can be proved in a very similar manner. Also, we only prove that "$\leq$" holds there because the other direction of the inequality can be proved almost by the same way as Proposition 5 without invoking the rectangularity.

To this end, assume that $>$ holds there on the contrary. By the compactness of the relevant Knightian Uncertainty, which is guaranteed by the assumed weak *compactness of $\mathscr{P}$, there exists a sequence of probability charges such that $p'^* \in \mathscr{P}_1; p''^* \in \mathscr{P}^+|_{E_1}; \ldots; p^{(T)*} \in \mathscr{P}^+|_{E_1\times\cdots\times E_{T-1}}$, each of which attains the corresponding minimum.

Here, note that by the decomposition of a risk, for any sequence of observations, $E_1, \ldots, E_T$, it holds that

$$
\begin{aligned}
p^*_{i_1,i_2,\ldots,i_T} = p^*_{i_1 i_2\ldots i_T} \; & := \; p^{(T-1)*}_{i_1 i_2\ldots i_{T-1}} \cdot p^{(T)*}_{i_T|i_1 i_2\ldots i_{T-1}} \\
& = \; p^{(T-2)*}_{i_1 i_2\ldots i_{T-2}} \cdot p^{(T-1)*}_{i_{T-1}|i_1 i_2\ldots i_{T-2}} \cdot p^{(T)*}_{i_T|i_1 i_2\ldots i_{T-1}} \\
& = \; \ldots\;\ldots\;\ldots
\end{aligned}
$$

$$= \quad p'^{*}_{i_1} \cdot p''^{*}_{i_2|i_1} \cdot \cdots \cdot p^{(T)*}_{i_T|i_1 i_2 \ldots i_{T-1}}$$

(see Equation (53)).

By exactly the same reasoning as Proposition 5, we reached the contradiction because $\left(p^{*}_{i_1,i_2,\ldots,i_T}\right)_{i_1,i_2,\ldots,i_T} \in \mathscr{P}$ *by its rectangularity.* □

C.3 The Formal Definition of the Sequential $\boldsymbol{\varepsilon}$-Contamination

Let $p^0 \in \mathscr{M}(\Omega, \mathcal{F}_T)$ and let $\varepsilon \in (0,1)$. Then, the *one-shot ε-contamination of* p^0 is denoted and defined by

$$\left\{p^0\right\}^{\varepsilon} := \left\{ (1-\varepsilon)p^0 + \varepsilon q \,\middle|\, q \in \mathscr{M}(\Omega, \mathcal{F}_T) \right\} . \tag{55}$$

Use the same p^0 and ε to define $\underline{\varepsilon}_{i_1} := -\varepsilon p^0_{i_1}$, $\bar{\varepsilon}_{i_1} := \varepsilon(1 - p^0_{i_1})$, $(\forall t \in \{2, \ldots, T\})$

$$\underline{\varepsilon}_{i_1 \ldots i_t} := \frac{-\varepsilon p^{0+}_{i_t|i_1 \ldots i_{t-1}}}{(1-\varepsilon)p^0_{i_1 \ldots i_{t-1}} + \varepsilon} \quad \text{and} \quad \bar{\varepsilon}_{i_1 \ldots i_t} := \frac{\varepsilon(1 - p^{0+}_{i_t|i_1 \ldots i_{t-1}})}{(1-\varepsilon)p^0_{i_1 \ldots i_{t-1}} + \varepsilon}, \tag{56}$$

where p^{0+} is the one-period-ahead conditional of p^0 defined in C.1. Then, the *sequential $\boldsymbol{\varepsilon}$-contamination of* p^0 is defined by

$$\begin{aligned} \left\{p^0\right\}^{seq\varepsilon} := \Bigg\{ \Big((p^0_{i_1} + \varepsilon_{i_1})(p^{0+}_{i_2|i_1} + \varepsilon_{i_1 i_2}) \cdots (p^{0+}_{i_T|i_1 \ldots i_{T-1}} + \varepsilon_{i_1 \ldots i_T}) \Big)_{i_1,\ldots,i_T} \,\Bigg|\, \\ (\forall i_1)\ \varepsilon_{i_1} \in [\underline{\varepsilon}_{i_1}, \bar{\varepsilon}_{i_1}];\ \textstyle\sum_{i_1} \varepsilon_{i_1} = 0;\ \ldots\ ; \\ (\forall i_1, \ldots, i_T)\ \varepsilon_{i_1 \ldots i_T} \in [\underline{\varepsilon}_{i_1 \ldots i_T}, \bar{\varepsilon}_{i_1 \ldots i_T}] \\ \text{and}\ \ (\forall i_1, \ldots, i_{T-1})\ \textstyle\sum_{i_T} \varepsilon_{i_1 \ldots i_{T-1} i_T} = 0 \Bigg\}. \end{aligned} \tag{57}$$

C.4 Properties of the Sequential $\boldsymbol{\varepsilon}$-Contamination and Its Comparison to the One-Shot ε-Contamination

Here, we show that a series of results established for the two-period setting in the main text can be extended to an arbitrarily-finite-horizon setting.

Proposition 15 *The sequential $\boldsymbol{\varepsilon}$-contamination is rectangular.*

Proof The proof can be conducted very closely following the proof for the case where $T = 2$ (Proposition 6). Therefore, it is omitted. □

Proposition 16 *It holds that* $\left\{p^0\right\}^{\varepsilon} \subseteq \left\{p^0\right\}^{seq\varepsilon}$.

Proof When $T = 2$, the claim holds true (Proposition 8).

Now, let $\{p^0\}^\varepsilon$ be the one-shot ε-contamination with $T = 3$ and use the same p^0 and ε to define $p_2 := \left((p^0_{i_1} + \varepsilon_{i_1})(p^{0+}_{i_2|i_1} + \varepsilon_{i_1 i_2})\right)_{i_1 i_2} \in \mathscr{M}(S^2, \langle E_{1,i_1} \times E_{2,i_2}\rangle_{i_1,i_2})$. By the way of the construction of $\{p^0\}^{seq\varepsilon}$ with $T = 3$ from p_2, the claim for $T = 3$ can be proved by very closely following the proof for the case where $T = 2$ by letting $\varepsilon_{i_1 i_2 i_3} := (\delta_{i_1,i_2,i_3} - \delta_{i_1,i_2} p^{0+}_{i_3|i_1 i_2})/(p^0_{i_1 i_2} + \delta_{i_1,i_2})$, where $\delta_{i_1,i_2} := \sum_{i_3} \delta_{i_1,i_2,i_3}$ and $p^0_{i_1 i_2} := \sum_{i_3} p^0_{i_1 i_2 i_3}$ for some δ_{i_1,i_2,i_3}. Thus, we omit the details of the proof.

For $T \geq 4$, repeat the procedure briefly described in the previous paragraph. □

Proposition 17 (ε^1-ε^2-$\cdots$-ε^T Contamination) *It holds that*

$$\{p^0\}^{seq\varepsilon} = \Big\{ \Big(\big((1-\varepsilon^1)p^0_{i_1} + \varepsilon^1 q_{i_1}\big) \big((1-\varepsilon^2_{i_1})p^{0+}_{i_2|i_1} + \varepsilon^2_{i_1} q_{i_1 i_2}\big) \cdots \big((1-\varepsilon^T_{i_1\dots i_{T-1}})p^{0+}_{i_T|i_1\dots i_{T-1}} + \varepsilon^T_{i_1\dots i_{T-1}} q_{i_1 i_2 \dots i_T}\big)\Big)_{i_1,i_2,\dots,i_T} \Big| \; (\forall i_1)\ q_{i_1} \in [0,1];\ \textstyle\sum_{i_1} q_{i_1} = 1; \dots;\ (\forall i_1,\dots,i_T)\ q_{i_1\dots i_T} \in [0,1] \text{ and } (\forall i_1,\dots,i_{T-1})\ \textstyle\sum_{i_T} q_{i_1\dots i_{T-1} i_T} = 1 \Big\}, \tag{58}$$

where $\varepsilon^1 := \varepsilon$, which is defining the one-shot ε-contamination, and $(\forall t \in \{2,\dots,T\})(\forall (i_1,\dots,i_{t-1}))$ $\varepsilon^t_{i_1\dots i_{t-1}}$ is defined by

$$\varepsilon^t_{i_1\dots i_{t-1}} := \frac{\varepsilon}{(1-\varepsilon)p^0_{i_1\dots i_{t-1}} + \varepsilon}.$$

Furthermore, $(\forall i_1,\dots,i_{T-1})$ $\varepsilon^1 < \varepsilon^2_{i_1} < \cdots < \varepsilon^T_{i_1\dots i_{T-1}}$, unless $(\exists t)$ $E_t = S$.

Proof The first half of the claim can be proved by very closely following the proof for the case when $T = 2$, and hence, the details of the proof is omitted.

To show that it holds that, along any sequence of observations, $\varepsilon^1 < \varepsilon^2_{i_1} < \cdots < \varepsilon^T_{i_1\dots i_{T-1}}$, simply note that $p^0_{i_1\dots i_{t-1}} > p^0_{i_1\dots i_{t-1} i_t}$ unless $E_t = S$ by the definition of the marginal. □

The one-period-ahead conditional Knightian uncertainty of any sequential $\boldsymbol{\varepsilon}$-contamination has a very convenient form.

To be precise, let $p^0 \in \mathscr{M}(\Omega, \mathcal{F}_T)$ and let $\varepsilon \in (0,1)$. Then, note that for any $t \leq T$, Proposition 17 and the definition of the marginal imply that

$$\{p^0\}^{seq\varepsilon}_t = \Big(\big((1-\varepsilon^1)p^0_{i_1} + \varepsilon^1 q_{i_1}\big) \big((1-\varepsilon^2_{i_1})p^{0+}_{i_2|i_1} + \varepsilon^2_{i_1} q_{i_1 i_2}\big) \cdots \big((1-\varepsilon^t_{i_1\dots i_{t-1}})p^{0+}_{i_t|i_1\dots i_{t-1}} + \varepsilon^t_{i_1\dots i_{t-1}} q_{i_1 i_2\dots i_t}\big)\Big)_{i_1,i_2,\dots,i_t},$$

where q's satisfy the conditions imposed in (58) and $\varepsilon^t_{i_1\dots i_{t-1}}$ is defined in Proposition 17.

Furthermore, this together with the definition of the one-period-ahead conditional Knightian uncertainty, (54), adapted to the sequential $\boldsymbol{\varepsilon}$-contamination shows that for any $t \leq T$ and any $E_1 \times \cdots \times E_{t-1}$,

$$\begin{aligned}
&\left(\{p^0\}^{seq\varepsilon}\right)^+ |_{E_1\times\cdots\times E_{t-1}} \\
&= \left\{ (1-\varepsilon^t_{i_1\dots i_{t-1}})p^{0+}(\cdot|E_1\times\cdots\times E_{t-1}) + \varepsilon^t_{i_1\dots i_{t-1}} q \,\middle|\, q \in \mathscr{M}(S, \langle E_{t,i_t}\rangle_{i_t}) \right\}.
\end{aligned} \tag{59}$$

Here, note that the resemblance between (33) and (59), which strongly suggests that the next proposition holds true.

Proposition 18 ("Posteriors") *For any $t \leq T$ and any $E_1 \times \cdots \times E_{t-1}$, it holds that*

$$\left(\{p^0\}^{rec\varepsilon}\right)^+ |_{E_1\times\cdots\times E_{t-1}} = \left(\{p^0\}^{\varepsilon}\right)^+ |_{E_1\times\cdots\times E_{t-1}}.$$

Proof The way of the construction of the one-period-ahead conditional allows us to mimic the proof for the case where $T = 2$ (Proposition 11), and hence we omit the proof. □

In order to state our final result on the updating behavior with the sequential $\boldsymbol{\varepsilon}$-contamination, we largely simplify the dynamic structure underlying our model.[37]

Let $(\Omega, \mathcal{F}_T) := (S^T, \otimes_{t=1}^T \langle E_i\rangle_{i=1}^n)$, where $\otimes_{t=1}^T \langle E_i\rangle_{i=1}^n$ is the T-time self-direct-product of the identical finite partition of S, $\langle E_i\rangle_{i=1}^n$. Also, assume that $p^0 := p^{00} \otimes p^{00} \otimes \cdots \otimes p^{00}$, which is the T-time self-direct-product of some $p^{00} \in \mathscr{M}(S, \langle E_i\rangle_{i=1}^n)$.

Then, we can prove the following proposition.

Proposition 19 *Assume that the stochastic structure is as described in the previous paragraph. Then, for any $t \leq T$ and any $E_{i_1} \times \ldots \times E_{i_{t-1}} \times E_{i_t}$ such that $(\forall i_t)\ E_{i_t} \neq S$, it holds that*

$$\left(\{p^0\}^{rec\varepsilon}\right)^+ |_{E_{i_1}\times\cdots\times E_{i_{t-1}}} \subsetneq \left(\{p^0\}^{rec\varepsilon}\right)^+ |_{E_{i_1}\times\cdots\times E_{i_{t-1}}\times E_{i_t}}.$$

[37] We can dispense with this simplifying assumption if we incur a cost that the conclusion of the next proposition holds only when some condition are met. For such a condition, see Nishimura and Ozaki (2017, Theorem 14.5.2, p.243)

Proof In view of the Equation (59) and the current underlying stochastic structure, the left-hand side of the inclusion in the proposition turns out to be

$$\left\{ \left(1 - \varepsilon^t_{i_1 \dots i_{t-1}}\right) p^{00} + \varepsilon^t_{i_1 \dots i_{t-1}} q \,\middle|\, q \in \mathscr{M}(S, \langle E_i \rangle_i) \right\},$$

while its right-hand side turns out to be

$$\left\{ \left(1 - \varepsilon^{t+1}_{i_1 \dots i_t}\right) p^{00} + \varepsilon^{t+1}_{i_1 \dots i_t} q \,\middle|\, q \in \mathscr{M}(S, \langle E_i \rangle_i) \right\}.$$

Then, the strict inclusion we desire follows because $\varepsilon^t_{i_1 \dots i_{t-1}} < \varepsilon^{t+1}_{i_1 \dots i_t}$ by Proposition 17. □

The last proposition exhibits that an active updating behavior upon observing the occurrence of an event always dilates the degree of uncertainty whenever Knightian uncertainty is specified by the sequential $\boldsymbol{\varepsilon}$-contamination and underlying stochastic structure is as specified in the proposition.[38]

As we already claimed with respect to the two-period models, we repeat to emphasize that Propositions 6, 10 (in particular, its very last inequality), 15 as well as 17 (in particular, its last successive inequalities) all together show that time-consistency (rectangularity) and dilation of ambiguity upon Bayesian updating are completely consistent.

C.5 Job Search with Bayesian Updating in an Arbitrarily-Finite-Horizon

Let $T \in \mathbb{N} \backslash \{0, 1\}$. We now consider the T-period job search model with the sequential $\boldsymbol{\varepsilon}$-contamination. Except for the number of the periods, the model is exactly the same as the one we considered in Section 2.

For any $t \geq 1$, let $V_t : S^t \to \mathbb{R}$ be the value function at period t. We assume that the value function is measured by the current value, that is, V_t is measured in terms of the value at period t.

First, note that it holds that

$$(\forall t \geq 1) \quad V_t(s_1, s_2, \dots, s_t) = w_b + \beta w_b + \cdots + \beta^{T-t} w_b$$

if $t = \min\{\, i \in \{1, 2, \dots, t\} \,|\, s_i = b\}$, because the worker accepts the wage offer as soon as the state turns out to be b by the assumption (34).

Therefore, it suffices to concentrate on the value functions only when the state s keeps occurring in order to completely characterize the whole system of the value functions. By denoting the states' history $(s, s, \dots, s)$, where the state s is taking place successively t-times, by s^t, we will invoke the backward induction method to obtain

$$V_T(s^T) = w_s \vee c\,,$$

[38] Note that in Proposition 19, the principal probability charge in each period, p^{00}, is independent over time, but uncertainty itself is *not* because of the presence of non-independent term $\varepsilon \cdot q$. Thus, Bayesian updating does matter.

$$V_{T-1}(s^{T-1}) = \max\left\{w_s + \beta w_s\,,\; c + \beta\left(p^+_{b|s^{T-1}} w_b + p^+_{s|s^{T-1}}(w_s \vee c)\right)\right\}\,,$$

$$V_{T-2}(s^{T-2}) = \max\left\{w_s + \beta w_s + \beta^2 w_s\,,\right.$$
$$\left. c + \beta\left(p^+_{b|s^{T-2}} V_{T-1}(s^{T-2}, b) + p^+_{s|s^{T-2}} V_{T-1}(s^{T-1})\right)\right\}\,,$$

$$\cdots\cdots\cdots$$

$$V_1(s) = \max\left\{w_s + \beta w_s + \beta^2 w_s + \cdots + \beta^{T-1} w_s\,,\right.$$
$$\left. c + \beta\left(p^+_{b|s} V_2(s, b) + p^+_{s|s} V_2(s, s)\right)\right\}\,,$$

where

$$p^+_{b|s^t} = \frac{(1-\varepsilon)p^0_{s^t b}}{(1-\varepsilon)p^0_{s^t} + \varepsilon} \quad \text{and} \quad p^+_{s|s^t} = \frac{(1-\varepsilon)p^0_{s^t s} + \varepsilon}{(1-\varepsilon)p^0_{s^t} + \varepsilon}\,, \tag{60}$$

for $t = 1, 2, \ldots, T-1$.

Here, note that the former term in (60) is the (conditional) probability weight on the "better" value function in the next period and that

$$(\forall t = 1, 2, \ldots, T-1) \quad \frac{\partial}{\partial \varepsilon}\left(\frac{(1-\varepsilon)p^0_{s^t b}}{(1-\varepsilon)p^0_{s^t} + \varepsilon}\right) < 0\,,$$

which shows that in the sequence of the value functions listed above, an increase in ε tends to make the second element (out of two) in the braces smaller.

We thus established the following result.[39]

Proposition 20 *An increase in ε (i.e., an increase in uncertainty) may discourage the worker's behavior of continuing the job search beyond the current period at period t if the state s keeps occurring up to period t. If otherwise, the worker should have already stopped the search regardless of the states' history up to period t. In particular, it is not the case that the worker who has decided to stop the search will change her mind to continue to search because of an increase in ε in any event.*

This proposition clearly states that Proposition 13 holds true regardless of the length of the periods as far as it is finite.

We conclude the paper by presenting a comparative statics result in a special situation, that is, when the principal probability charge p^0 is given by the T-time self-direct-product as $p^0 := p^{00} \otimes p^{00} \otimes \cdots \otimes p^{00}$.

[39] In this appendix, we do not invoke Assumption 1 because we do not derive the explicit closed-form of the value function here. See Footnote 28.

If this is the case, (60) will become

$$p^{+}_{b|s^t} = \frac{(1-\varepsilon)(p^{00}_s)^t p^{00}_b}{(1-\varepsilon)(p^{00}_s)^t + \varepsilon} \quad \text{and} \quad p^{+}_{s|s^t} = \frac{(1-\varepsilon)(p^{00}_s)^t p^{00}_s + \varepsilon}{(1-\varepsilon)(p^{00}_s)^t + \varepsilon},$$

for $t = 1, 2, \cdots T-1$. Hence, because $p^{00}_b, p^{00}_s \in (0,1)$, it follows that

$$\lim_{t\to\infty} p^{+}_{b|s^t} = 0 \quad \text{and} \quad \lim_{t\to\infty} p^{+}_{s|s^t} = 1 .$$

All this shows the next proposition.

Proposition 21 *If the principal probability charge p^0 is given by the T-time self-direct-product as $p^0 = p^{00} \otimes p^{00} \otimes \cdots \otimes p^{00}$, then there exists a (possibly large, but finite) T such that the worker stops the search and accepts the wage offer by period T. (T may depend only on ε and p^{00}.)*

This proposition may be understood as insisting that not only an increase in uncertainty but also even *its presence itself* tends to urge the agent to make uncertain prospects determinate at some point over the course of an employment spell because of the gradual decrease of the reservation wage. This is consistent with accumulated empirical evidence showing that the reservation wage declines over such periods. (See Brown, Flinn and Schotter (2011, p.948) and literature cited in the footnote 1 there.)

References

[1] Berger, J. O. (1985): *Statistical Decision Theory and Bayesian Analysis* (Second Edition), Springer-Verlag.

[2] Brown, M., C. J. Flinn and A. Schotter (2011): "Real-time search in the laboratory and the market," *American Economic Review* 101, 948-974.

[3] Chu, C. Y. C. and W.-F. Liu (2002): "Knightian uncertainty and the voting paradox," mimeo., University of Washington.

[4] Epstein, L. G. and M. Schneider (2003): "Recursive multiple-priors," *Journal of Economic Theory* 113, 1-31.

[5] Epstein, L. and T. Wang (1994): "Intertemporal asset pricing and Knightian uncertainty," *Econometrica* 62, 283-382, 1994.

[6] Flor, C. R. and S. Hesel (2015): "Uncertain dynamics, correlation effects, and robust investment decisions," *Journal of Economic Dynamic & Control* 51, 278-298.

[7] Gilboa, I., and D. Schmeidler (1989): "Maxmin expected utility with non-unique prior," *Journal of Mathematical Economics* 18, 141-153.

[8] Kopylov, I. (2009): "Choice deferral and ambiguity aversion", *Theoretical Economics* 4, 199-225.

[9] Kopylov, I. (2016): "Subjective probability, confidence, and Bayesian uadating," *Economic Theory* 62, 635-658.

[10] Liu, H. (2011): "Dynamic portfolio choice under ambiguity and regime switching mean returns," *Journal of Economic Dynamic & Control* 35, 623-640.

[11] Miao, J. and N. Wang (2011): "Risk, uncertainty, and option exercise," *Journal of Economic Dynamic & Control* 35, 442-461.

[12] Nishimura, K. G., and H. Ozaki (2004): "Search and Knightian uncertainty," *Journal of Economic Theory* 119, 299-333.

[13] Nishimura, K. G., and H. Ozaki (2006): "An axiomatic approach to ε-contamination," *Economic Theory* 27, 333-340.

[14] Nishimura, K. G., and H. Ozaki (2007): "Irreversible investment and Knightian uncertainty," *Journal of Economic Theory* 136(1), 668-694.

[15] Nishimura, K. G., and H. Ozaki (2017): *Economics of Pessimism and Optimism: Theory of Knightian Uncertaonty and Its Applications*, Springer.

[16] Shishkin, D. ® Ortoleva, P. (2019): "Ambiguous information and dilation: an experiment," mimeo. http://pietroortoleva.com/

[17] Thijssen, J. J. (2011): "Incomplete markets, ambiguity, and irreversible investment," *Journal of Economic Dynamics and Control* 35, 909-921.