CIRJE-F-112

Search and Knightian Uncertainty

Kiyohiko G. Nishimura
University of Tokyo

Hiroyuki Ozaki
Tohoku University

April 2001

"Search under the Knightian Uncertainty"

K. G. Nishimura and H. Ozaki

Abstract

Suppose that "uncertainty" about labor market conditions has increased. Does this change induce an unemployed worker to search longer, or shorter? This paper shows that the answer is drastically different depending on whether an increase in "uncertainty" is an increase in risk or that in true uncertainty in the sense of Frank Knight. We show in a general framework that, while an increase in risk (the mean-preserving spread of the wage distribution that the worker thinks she faces) increases the reservation wage, an increase in the Knightian uncertainty (a decrease in her confidence about the wage distribution) reduces it.

SEARCH AND KNIGHTIAN UNCERTAINTY*

by

Kiyohiko G. Nishimura

Faculty of Economics
The University of Tokyo

and

Hiroyuki Ozaki

Faculty of Economics
Tohoku University

April 8, 2001

Abstract

Suppose that “uncertainty” about labor market conditions has increased. Does this change induce an unemployed worker to search longer, or shorter? This paper shows that the answer is drastically different depending on whether an increase in “uncertainty” is an increase in risk or that in true uncertainty in the sense of Frank Knight. We show in a general framework that, while an increase in risk (the mean-preserving spread of the wage distribution that the worker thinks she faces) increases the reservation wage, an increase in the Knightian uncertainty (a decrease in her confidence about the wage distribution) reduces it.

*We are grateful to seminar participants at Western Ontario and SUNY-Buffalo for their helpful comments. The first author’s work is partially supported by a grant to the Project of Information Technology and the Market Economy, from the Japan Society for Promotion of Sciences.

1. Introduction

Consider an unemployed worker who is searching for a job. Suppose that "uncertainty" about labor market conditions has increased. Does this change induce her to search longer and more intensively, or shorter and less intensively? The answer to this question has utmost importance both in macroeconomics concerning the aggregate unemployment rate and microeconomics explaining worker behavior. The purpose of this paper is to show that the answer is drastically different depending on what kind of "uncertainty" is involved. If an increase in "uncertainty" is an increase in the variance of the wage offer distribution that the worker thinks she faces, the worker searches longer. If an increase in "uncertainty" is a decrease in her confidence about the wage distribution, the worker searches shorter.

In the tradition of Frank Knight, the former type of "uncertainty," which is reducible to a single distribution with known parameters, is *risk*, while the latter type of "uncertainty," which is irreducible, is true *uncertainty* (see Knight, 1921, and also see Keynes, 1921, 1936). While risk and uncertainty are clearly distinct concepts, they have not been treated separately in economics in an explicit way, at least until recently. This may be because of the celebrated theorem of Savage (Savage, 1954) which shows that if the decision maker's behavior complies to certain axioms, her preference is represented by the expectation of some utility function which is computed by means of some *single* probability measure. Uncertainty that the decision maker faces is thus reduced to risk with some probability measure. However, Ellsberg (1961) presented an example of preference under uncertainty that cannot be justified by Savage's expected utility framework. The decision maker's behavior described in Ellsberg's paradox, which is not at all irrational, clearly violates some of Savage's axioms.

Gilboa (1987), Gilboa and Schmeidler (1989) and Schmeidler (1989) weaken Savage's axioms to settle debates caused by Ellsberg's paradox. They axiomatize the preference which is represented by the minimum among the expected utilities each of which is computed by an element of some *set* of probability measures. This preference is called the *maximin expected utility* or the *Choquet expected utility*. This is a natural extension of preference under uncertainty to the case in which the information is too imprecise to summarize it by a single probability measure. This type of uncertainty is called the *Knightian uncertainty*.

This paper applies the idea of the Knightian uncertainty to the job search model, and compare the effect of its increase on the worker's search behavior with that of an increase in risk. To this end, we extend the stylized model of job search without recall (see, for example, Sargent, 1987, p.66) by assuming that the unemployed worker's preference is represented by the maximin expected utility / Choquet expected utility axiomatized by the authors cited above. Since we focus on the role of the Knightian uncertainty, we assume the time-consistent intertemporal structure in the worker's preference over time. Under these setting, we show that the optimal stopping rule exists, that this optimal stopping rule has a reservation property, and that the reservation wage is characterized by a functional equation.

Then, we exploit the functional equation determining the reservation wage to examine the effect of an increase in the Knightian uncertainty. In the traditional framework where uncertainty is specified by a single probability distribution, an increase in uncertainty, that is, an increase in *risk*, is modeled by a mean-preserving spread of the given distribution. Then, it turns out that the mean-preserving spread, that is, an increase in risk, causes an *in*crease in the reservation wage (see Rothschild and Stiglitz, 1970, 1971, or Section 2.1 of this paper). Thus, the unemployed worker is inclined to keep searching for a job when risk has increased. In contrast, we formulate the Knightian uncertainty in such a way that the worker does not have confidence that a given wage distribution is the true one, and that instead she assumes a set of probability distributions and maximizes the minimum of expected utilities based on each probability distribution. We then show that the reservation wage is *de*creased when the Knightian uncertainty increases, and hence, the worker tends to accept the job offer more quickly. This result conforms our intuition that, when people lose confidence in their forecast about what happens in the future, they generally prefer certainty to uncertainty. An immediate acceptance of the wage offer implies that the uncertainty is turned into certainty.

The organization of the paper is as follows. Section 2 explains the main result of this paper by using a simple example based on the uniform distribution of the wage offer. Technical discussions are kept at the minimum in this section. Section 3 bridges non-technical Section 2 and technical Section 4 in explaining the maximin expected utility, Choquet expected utility, and some continuity problems which must be solved. Section 4 presents the main result in a general framework. This section also explains how the result of Section 2 is derived from the general framework. Proofs of the main theorems of Section 4 are relegated to Section 5. The definitions and some mathematical results about the Choquet capacity, which are extensively utilized in Section 4, are collected in the Appendix. Any lemma in the Appendix will be referred to as Lemma A_.

2. An Example: Risk versus Knightian Uncertainty

Let us first consider a simple job search model (for example, see a stylized example of Sargent, 1987, p.66). In each period, an unemployed worker draws a wage offer, from a wage distribution[1] F_0. *The worker is assumed to know the true distribution* F_0. If the unemployed worker accepts the offer, he earns that wage from this period on. If he decides not to, he gets unemployment compensation, $c > 0$, in this period[2] and will make a draw again in the next period. Let T denote the period that the worker accepts the wage offer. The unemployed worker's objective is, by choosing a suitable stopping rule, to maximize his expected life-time income

$$E_0 \sum_{t=0}^{\infty} \beta^t y_t$$

where

$$y_t = \begin{cases} c & \text{for} \quad t < T \\ w_T & \text{for} \quad t \geq T \end{cases}$$

Under general conditions on F_0, (1) there exists the optimal stopping rule and (2) the stopping rule has a reservation property. That is, the optimal stopping rule is to accept the wage offer if it is no smaller than the reservation wage R and to wait for another offer if otherwise, where the reservation wage R is determined by a choice between accepting this period's offer or waiting for next period's offer:[3]

$$R = c + \frac{\beta}{1-\beta} \int_R^{\infty} [1 - F_0(x)] dx. \tag{1}$$

[1] $F_0(x)$ denotes the probability that the wage offer is no greater than x.

[2] The basic structure of the model is unchanged if instead we assume that the unemployed worker pays a search cost, rather than he gets an unemployment compensation. In the case of the search cost, we have $c < 0$.

[3] Eq (1) is easily derived from Corollary 2 in Section 4.4 as a special case.

If we further specify the wage distribution to be a uniform distribution over $[a,b]$ where $0 < a < b$, we have an explicit solution of the reservation wage. We further assume that $b > c$ because otherwise continuing the search forever would be trivially optimal, and that the parameters of the model satisfy the following conditions:

$$b - a > \beta(2c - a - b) \tag{2}$$

$$\text{and} \quad 2(c - a) > \beta(2c - a - b) \tag{3}$$

to assure that $R \in (a,b)$ holds. Then by (1), we have

$$R = c + \frac{\beta}{1-\beta}\int_R^b \frac{b-x}{b-a}\,dx = c + \frac{\beta(b-R)^2}{2(1-\beta)(b-a)}. \tag{4}$$

By solving this quadratic equation, we get[4]

$$R = \frac{1}{\beta}\left(b - (1-\beta)a - D^{1/2}\right) \tag{5}$$

$$\text{where} \quad D \equiv (1-\beta)(b-a)\left[b - a - \beta(2c - a - b)\right].$$

2.1. Increased Risk: Mean-preserving Spread

Suppose that "uncertainty" over wage offers is slightly increased for the worker. In the above example, the wage distribution may be slightly more dispersed by γ, over $[a-\gamma, b+\gamma]$. See Figure 1, where F_0 (the dotted line) is the probability distribution function of the uniform distribution over $[a,b]$ and the solid line is the one of the new uniform distribution over $[a-\gamma, b+\gamma]$. This is a mean-preserving spread, characterizing *increased risk* (see, Rothschild and Stiglitz, 1970). If this is the case, (4) is modified to

$$R = c + \frac{\beta(b+\gamma-R)^2}{2(1-\beta)(b-a+2\gamma)}. \tag{6}$$

[4] Let us conjecture that $R \in (a,b)$ in order to derive (4) from (1). We then verify that the reservation wage R thus derived in (5) certainly satisfies this condition under (2) and (3).

First, note that $D > 0$ by $b > a$ and by (2). Second, note that the conjugate solution to the quadratic equation (4):

$$\bar{R} \equiv \frac{1}{\beta}\left(b - (1-\beta)a + D^{1/2}\right)$$

violates the condition because $\bar{R} > b$ by $D > 0$ and by the fact that $(1/\beta)(b-(1-\beta)a) > b$, which is equivalent to $(1-\beta)(b-a) > 0$. Third, note that we have $R < b$ since we get $R < b$ if and only if $b > c$, which holds under our assumption. Finally, note that $R > a$ because $R > a$ if and only if $2(c-a) > \beta(2c-a-b)$, which holds since we have (3). The conjecture is thus verified.

Denote the solution R to this equation by $R(\gamma)$ as a function of γ. Then, the implicit function theorem shows that

$$\left.\frac{dR(\gamma)}{d\gamma}\right|_{\gamma=0} = \frac{\beta(b-R)(R-a)}{(b-a)[(1-\beta)(b-a)+\beta(b-R)]},$$

where R in the right-hand side is given by (5). Since $a<R<b$ by (2) and (3),

$$\left.\frac{dR(\gamma)}{d\gamma}\right|_{\gamma=0} > 0.$$

This result shows that increased "uncertainty" in the form of increased risk (a mean-preserving spread) increases the reservation wage.

2.2. Increased Knightian Uncertainty: δ-Approximation of ε-Contamination

In the case of the mean-preserving spread, the worker is still certain of the shape of the wage distribution. It is a uniform distribution $[a-\gamma, b+\gamma]$, spreading out the original distribution by exactly γ. The worker has firm confidence about the new wage distribution.

In reality, however, the worker may not have such firm confidence on the wage distribution when economic conditions are changed. The worker may become uncertain about the shape of the wage distribution itself. The wage distribution may be different from the uniform distribution over $[a,b]$ with a positive (though small) probability. Moreover, the worker may have no idea about the shape of the wage distribution if in fact it is different. It may still be uniform and spreading out by γ ($[a-\gamma, b+\gamma]$), but the worker does not have any confidence about the value of γ. It may be wildly different from uniform distribution. This "uncertainty" that the worker faces clearly cannot be reduced to a change in parameters of known distribution. Thus, the "uncertainty" here is the Knightian uncertainty.

The problem that the worker faces is similar to that of a Bayesian statistician who confronts "uncertainty" in a prior distribution of the Bayesian learning process. One procedure that the Bayesian statistician often follows is to introduce a set of priors obtained by "contaminating" a single hypothetical prior and then to investigate the robustness of the learning process. This procedure is often called as *ε-contamination*.[5] We follow this Bayesian tradition in formulating the Knightian uncertainty by "contaminating" the original wage distribution.

[5] See, for example, Berger (1985) and Wasserman and Kadane (1990). For the use of ε-contamination in economics, see Epstein and Wang (1994, 1995).

We formulate the worker's problem in three steps. Firstly, following the ε-*contamination* literature, we specify the uncertainty that the worker faces by a *set* of distributions, rather than by a single distribution in the traditional framework. Secondly, we postulate an appropriate optimal search problem of the worker facing this multi-distribution uncertainty, using the framework of Gilboa and Schmeidler (1989). Thirdly, we examine whether the optimal strategy has the reservation wage property and if it has, whether increased uncertainty increases the reservation wage.

In order to follow the ε-*contamination* literature, we need to be a bit formalistic. Let W be a Borel subset of $\mathbb{R}_+$ and $\mathscr{B}_W$ be the Borel σ-algebra on W. Let P_0 be the probability measure on W corresponding to F_0. In our example, W is $[a,b]$, and P_0 is the uniform distribution over $[a,b]$. Let $\mathcal{M}$ be the set of all probability measures on $\mathscr{B}_W$ and let $\varepsilon > 0$. In our example, $\mathcal{M}$ is the set of all probability measures corresponding to distributions over $[a,b]$. Then, ε-*contamination* of the original distribution is the set of probability measures on W defined by

$$\mathcal{P}_0 \equiv \{ (1-\varepsilon)P_0 + \varepsilon\mu \mid \mu \in \mathcal{M} \}. \tag{7}$$

In fact, if $\varepsilon = 0$, then $\mathcal{P}_0 = \{P_0\}$, and the problem is reduced to the traditional search one. An increase in ε implies that the worker becomes less certain that P_0 is in fact the true distribution. Thus, an increase in ε can be interpreted as an increase in the Knightian uncertainty.[6]

There is, however, one technical problem in the above formulation. This formulation implies that the value of $\inf\{P(A)|P \in \mathcal{P}_0\}$ changes discontinuously at $A = W$ when A approaches W (an illustrative example is given below in Figures 2 and 3), and this discontinuity makes dynamic analysis of this paper much complicated mathematically with no further economic insights (see Section 3.2). To avoid this mathematical problem, we use in this section the following δ-approximation of ε-contamination.

Let δ be a small positive number, and let $\mathcal{M}(P_0,\delta)$ be

$$\mathcal{M}(P_0,\delta) = \{ \mu \in \mathcal{M} \,|\, (\forall A) \;\; \delta\mu(A) \;\leq\; P_0(A) \}.$$

[6] The formal definition of more Knightian uncertainty is given in Section 4.5.

The δ-approximation of ε-contamination, $\mathcal{P}_\delta$, is defined by

$$\mathcal{P}_\delta \equiv \{\,(1-\varepsilon)P_0 + \varepsilon\mu \,|\, \mu \in \mathcal{M}(P_0,\delta)\,\}. \tag{8}$$

Note that $\mathcal{M}(P_0, 0) = \mathcal{M}$. Thus, by appropriately choosing a small δ we can approximate $\mathcal{P}_0$ by $\mathcal{P}_\delta$ as close as we want.

The ε-contamination and its δ-approximation are best explained in our uniform-distribution example by Figures 2 and 3. In our example, we have $W = [a,b]$. In Figure 2, the ε-contamination of P_0, $\mathcal{P}_0$, is given by the set of all probability distribution functions above $(1-\varepsilon)F_0$ and below $(1-\varepsilon)F_0 + \varepsilon$ for all $x \in (a,b)$.[7] It is evident from this figure that $\inf\{P([a,x])|P \in \mathcal{P}_0\}$ is equal to $(1-\varepsilon)F_0(x)$ for all $x \in [a,b)$, and that we have $\inf\{P([a,x])|P \in \mathcal{P}_0\} = 1$ at $x = b$. Thus, $\inf\{P([a,x])|P \in \mathcal{P}_0\}$ becomes discontinuous at $x = b$. In Figure 3, the δ-approximation of ε-contamination of P_0, $\mathcal{P}_\delta$, is given by the set of all probability distribution functions which are (i) above $(1-\varepsilon)F_0$ for $x \in [a,y]$ and above $(1-\varepsilon)F_0 + \varepsilon[(F_0 - 1)/\delta + 1]$ for $x \in [y,b]$, where $F_0(y) = 1-\delta$, and (ii) below $(1-\varepsilon)F_0 + \varepsilon F_0/\delta$ for $x \in [a,z]$ and below $(1-\varepsilon)F_0 + \varepsilon$ for $x \in [z,b]$, where $F_0(z) = \delta$.[8] It is evident by construction that $\inf\{\,P([a,x])\,|\,P \in \mathcal{P}_\delta\}$ is continuous for all $x \in [a,b]$, whereas $\inf\{\,P([a,x])\,|\,P \in \mathcal{P}_0\}$ is discontinuous at $x = b$. The figures also show that the δ-approximation of ε-contamination, $\mathcal{P}_\delta$, expands toward the ε-contamination, $\mathcal{P}_0$, as δ decreases, and that we can "approximate" the latter by the former as close as possible by appropriately choosing a small δ.

To concentrate our attention on the Knightian uncertainty itself, we assume that the unemployed worker faces the same uncertainty characterized by $\mathcal{P}_\delta$ in each period. That is, we assume that the observation of the wage offer does not affect the future uncertainty. Thus, we do not consider explicitly the worker's learning about the uncertainty. This assumption is

[7] From (7), we get the following alternative expression for $\mathcal{P}_0$:

$$\mathcal{P}_0 \;=\; \{P \in \mathcal{M}\,|\,(\forall A)\,P(A) \;\geq\; (1-\varepsilon)P_0(A)\}$$

From this expression, we immediately know that there is a lower bound for $P \in \mathcal{M}$ as described in Figure 2. Moreover, the inequality in the definition must hold for the complement of A. This implies that there is also an upper bound for $P \in \mathcal{M}$, as described in this figure.

[8] From (8), we get the following alternative expressions for $\mathcal{P}_\delta$:

$$\begin{aligned}\mathcal{P}_\delta = \{P \in \mathcal{M}\,|\,(\forall A)\,P(A) \;&\geq\; (1-\varepsilon)P_0(A) \quad \text{if} \quad P_0(A) \leq 1-\delta\,; \quad \text{and} \\ P(A) \;&\geq\; (1-\varepsilon)P_0(A) + \varepsilon[(P_0(A)-1)/\delta + 1] \quad \text{if} \quad P_0(A) > 1-\delta\,\}.\end{aligned}$$

We get the lower bound of $P(A)$ in Figure 3 immediately from this expression. The upper bound of $P(A)$ is obtained by substituting for A its complement in the above formula.

a reasonable one for the Knightian uncertainty. If the uncertainty were specified by a single distribution of some distribution family with an unknown parameter, say, a normal distribution with an unknown mean and a known variance, some updating procedure together with a conjugate prior over the parameter space would be used to detect the true value of the unknown parameter (see DeGroot, 1970 and Rothschild, 1974). In contrast to this case, the worker here does not know even the type of the true distribution, let alone its parameters (recall that a distribution in $\mathcal{P}_\delta$ can be a member of any of uncountably many parametric families). The uncertainty that the worker faces is much broader and deeper. In fact, it is shown that a commonly-used update rule may not resolve the Knightian uncertainty at all in some cases.[9]

Since the Knightian uncertainty is now defined by a set $\mathcal{P}_\delta$ of distributions rather than a single distribution, we must redefine the objective function accordingly. We postulate that the unemployed worker's objective is to maximize the minimum of his expected discounted future income

$$\min\left\{ \int_W I(w)\,dP(w) \,\middle|\, P \in \mathcal{P}_\delta \right\}, \tag{9}$$

where $I(w)$ is the discounted future income which is a bounded measurable function of the observed offer w.[10] The exact formula $I(w)$ is complicated and shown in Section 4.2 so that it is not shown here. Gilboa and Schmeidler (1989) show that if the worker's behavior complies to certain axioms, his objective function is in fact representable by an expression similar to (9). Thus, our formulation is consistent with theirs.[11]

Under these settings we can show (see Section 4) that (1) there exists the optimal stopping rule and (2) the optimal stopping rule has a reservation property. Furthermore, it turns

[9] We could incorporate explicitly some updating rule which is adopted for the Knightian uncertainty. Nishimura and Ozaki (2001) use the Dempster-Shafer rule, which is given some axiomatic foundation by Gilboa and Schmeidler (1993), to show that there are cases in which the adoption of the Dempster-Shafer rule does not resolve at all the uncertainty characterized by $\mathcal{P}_0$.

[10] The minimum is attained since I is assumed to be bounded and measurable, and $\mathcal{P}_\delta$ is weak * compact by the Alaoglu theorem.

[11] They employ a mixture space as a prize space and show that if the decision maker's behavior complies to certain axioms, his objective function is represented by

$$\min\left\{ \int_W u(I(w))\,dP(w) \,\middle|\, P \in \mathcal{P} \right\} \tag{10}$$

for some weak * closed convex set $\mathcal{P}$ of probability charges (for the definition of the probability charge, see the Appendix) and for some utility index u which is unique up to a positive affine transformation.

out that R is characterized as the solution to the following equation for sufficiently small $\varepsilon > 0$ and $\delta > 0$:

$$\begin{aligned} R &= c + \frac{\beta}{1-\beta} \int_R^\infty (1-\varepsilon) P_0(\{w \mid w \geq x\})\, dx \\ &= c + \frac{\beta}{1-\beta}(1-\varepsilon) \int_R^\infty P_0(\{w \mid w \geq x\})\, dx \\ &= c + \frac{\beta(1-\varepsilon)(b-R)^2}{2(1-\beta)(b-a)} \end{aligned} \tag{11}$$

(see right after Corollary 2 in Section 4.4). By solving this equation, we can write R as $R(\varepsilon)$ as a function of ε. The implicit function theorem shows that

$$\left.\frac{dR(\varepsilon)}{d\varepsilon}\right|_{\varepsilon=0} = -\frac{\beta(b-R)^2}{2(1-\beta)(b-a)+2\beta(b-R)},$$

where R in the right-hand side is given by (5). Since $R < b$ by (2) and (3),

$$\left.\frac{dR(\varepsilon)}{d\varepsilon}\right|_{\varepsilon=0} < 0,$$

which shows that an increase in the Knightian uncertainty, specified by an increase in ε, *de*creases the reservation wage.[12] This is exactly the opposite to an increase in risk, specified by an increase in γ. As we already mentioned in the Introduction, this makes sense economically. When they become less confident in what happens in the future, people may prefer "certainty" much more to "uncertainty." The uncertainty is resolved immediately when the worker accepts the wage offer. Hence, an increase in the Knightian uncertainty is likely to persuade the worker to cancel a further search.

Section 4 extends this example to a more general setting and show that qualitatively the same result holds even in the general case.

3. Some Technical Issues

Before proceeding with a formal analysis of Section 4, we deal in this section with two technical issues concerning dynamical analysis. In Section 4, we assume the objective function is intertemporally well-defined. Preferences need to be "continuous" for this property to hold.

[12] A similar result holds globally, not only locally. See right after Theorem 2 in Section 4.5.

The maximin preferences illustrated by the Example in the previous section are not well-suited for the characterization of this continuity requirement. Thus in Section 3.1, we reformulate preferences as those represented by a Choquet integral with a convex probability capacity. In Section 3.2, we explain problems that may arise if the capacity is not continuous in the Choquet-integral-cum-probability-capacity formulation. We thus impose the continuity assumption directly on the probability capacity.

3.1. Representation by Choquet Integral with Convex Capacity

In this section, we use probability capacity, Choquet integral and other related concepts, which are explained in the Appendix. Let P_0 be the probability measure corresponding to the uniform distribution over $[a,b]$ as in Section 2.2. Let $\varepsilon > 0$ and $\delta > 0$. Then, define $\theta_\delta : \mathscr{B}_W \to [0,1]$ by

$$(\forall A) \quad \theta_\delta(A) = \begin{cases} (1-\varepsilon)P_0(A) & \text{if} \quad P_0(A) \leq 1-\delta \\ (1-\varepsilon)P_0(A) + \varepsilon[(P_0(A)-1)/\delta + 1] & \text{if} \quad P_0(A) > 1-\delta\,. \end{cases} \tag{12}$$

Then, Lemma A1 in the Appendix shows that θ_δ is a convex probability capacity. Furthermore, it turns out that the core of θ_δ satisfies $\mathrm{core}(\theta_\delta) = \mathcal{P}_\delta$. Therefore, it follows that

$$\begin{aligned}(\forall I) \qquad \int_W I(w)\,d\theta_\delta(w) &= \min\left\{ \int_W I(w)\,dP(w) \,\middle|\, P \in \mathrm{core}(\theta_\delta) \right\} \\ &= \min\left\{ \int_W I(w)\,dP(w) \,\middle|\, P \in \mathcal{P}_\delta \right\}\end{aligned}$$

where the integral in the left-hand side of the above relations is the Choquet integral. The first equality of the above relations holds by Lemma A8 in the Appendix. This shows that the maximin preferences given by (9) are identical to the preferences which are represented by the Choquet integral with the convex capacity (12).

Gilboa (1987) and Schmeidler (1989) show that if the worker's behavior complies to certain axioms, his objective function is in fact representable by a Choquet integral with some convex probability capacity. Thus, our formulation is consistent with theirs.[13]

For mathematical tractability (see Section 3.2), we formulate a general search model under the Knightian uncertainty in the next section with preferences represented by a Choquet

[13]Gilboa (1987) employs a general prize space and Schmeidler (1989) employs a mixture space as a prize space, and they show that if the decision maker's behavior complies to certain axioms, his objective function is represented

integral with a convex probability capacity. It should, however, be kept in mind that we lose some mathematical generality by this procedure, though not much in economics. In particular, although the maximin representation and the Choquet representation of the preference coincide exactly for the case of the δ-approximation of ε-contamination, they are not always so. In fact, while the preferences represented by a Choquet integral with a convex capacity is a proper subset of the maximin preferences, the converse is not necessarily true.[14]

3.2. Problems with Non-Continuous Capacity

The convex capacity which is *not* continuous poses technical difficulty in a dynamic context. To see this point, let us consider the capacity θ_0 corresponding to the original ε-contamination.[15] It can be shown that θ_0 is not continuous.[16] Let $I(w_-, w)$ denote the discounted future income when the wage offer w_- has been observed today and the wage offer w will be observed tomorrow. Then, it turns out that

$$\int_W I(w_-, w)\, d\theta_0(w) = (1-\varepsilon) \int_W I(w_-, w)\, dP_0(w) + \varepsilon \inf_{w \in W} I(w_-, w). \tag{14}$$

---

by

$$\int_W u(I(w))\, d\theta(w) \tag{13}$$

for some convex probability capacity θ and for some utility index u which is unique up to a positive affine transformation.

[14]By Lemma A8, (13) is always reduced to (10). This footnote discusses that the converse is not necessarily true. Let $\mathcal{P}$ be an arbitrary weak * closed convex set of probability charges. If there exists a convex capacity θ with which the Choquet integral is identical to (10) with this $\mathcal{P}$, it must be that $(\forall A)\ \theta(A) = \inf \mathcal{P} \equiv \inf\{P(A) \mid P \in \mathcal{P}\}$ (let I be the indicator function of a set A). Therefore, we need to have that $\inf \mathcal{P}$ be convex and that $\mathrm{core}(\inf \mathcal{P}) = \mathcal{P}$. However, the latter equality may not hold and $\inf \mathcal{P}$ may not even be convex. (See Huber and Strassen, 1973, for both counter-examples.)

[15]The convex capacity θ_0 is defined by

$$(\forall A) \quad \theta_0(A) = \begin{cases} (1-\varepsilon)P_0(A) & \text{if} \quad A \neq W \\ 1 & \text{if} \quad A = W, \end{cases}$$

and it holds that $\mathrm{core}(\theta_0) = \mathcal{P}_0$ ($\mathcal{M}$ in the definition of $\mathcal{P}_0$ is now understood to be the set of all probability *charges*, rather than probability *measures*).

[16]To see this, consider the increasing sequence of sets, $\langle W_n \rangle_n$, each of which is not equal to W and such that $\cup_n W_n = W$. Such a sequence exists, for example, when W is an open interval.

It should be noted that θ_0 satisfies some weaker continuity requirement on the capacity defined on the Borel σ-algebra. If this weaker continuity requirement is satisfied, (15) below turns out to be *analytic* in w_-. Analyticity is a weaker property than measurability. To handle analytic functions, much more mathematical sophistication is required than to handle measurable functions. Epstein and Wang (1995) refer to Dellacherie and Meyer (1988) to cope with analytic functions in a dynamic asset pricing model.

For the unemployed worker's objective function is well-defined over time, (14) must be measurable in w_{i} today because *yesterday* the worker had computed the expectation of (14) over w_{i}. However, it is well-known that $\inf_{w2W} I(w_{\mathrm{i}}, w)$ is not necessarily measurable in w_{i} even if I is measurable jointly in (w_{i}, w). More generally, the Choquet integral

$$\int_W I(w_{\mathrm{i}}, w)\, d\theta(w) \tag{15}$$

is not necessarily measurable in w_{i} unless θ is continuous.

Although θ_0 is not continuous, Lemma A1 shows that θ_δ defined by (12) is continuous. Note that (15) is always measurable in w_{i} when the capacity is continuous by the Fubini property (Lemma A13). Thus, the worker's objective is well-defined in this case. This is why we use the δ-approximation of the ε-contamination in the previous section, instead of the ε-contamination.

One mathematical advantage to formulate preferences with a Choquet integral with a convex capacity, rather than maximin preferences, is that we can impose the continuity assumption directly on the primitive of the model by assuming that the capacity is continuous. This procedure greatly simplifies formal dynamic analysis without losing any economic insights.

4. The Formal Model

4.1. Stochastic Environment

Let $(W, \mathscr{B}_W)$ be a measurable space, where W is a Borel subset of $\mathbb{R}_+$ and $\mathscr{B}_W$ is the Borel σ-algebra on W. We regard each element $w \in W$ as an offer of wage in each single-period. For any $t \geq 0$, we construct the t-dimensional product measurable space $(W^t, \mathscr{B}_{W^t})$ (we let $\mathscr{B}_{W^0} \equiv \{\phi, W^\infty\}$) and embed it in the infinite-dimensional product measurable space $(W^\infty, \mathscr{B}_{W^\infty})$ in a usual manner.[17] We write a history of realized offers as ${}_1w_t \equiv (w_1, w_2, \ldots, w_t) \in W^t$, ${}_1w \equiv (w_1, w_2, \ldots) \in W^\infty$, and so on.

[17]More precisely, the construction is as follows: First, let $W^\infty = W \times W \times \cdots$ be the countably-infinite-dimensional Cartesian product of W, and let $W^t = W \times \cdots \times W$ be the t-dimensional Cartesian product of W. That is, W^∞ is the set of infinite sequences $(w_1, w_2, \ldots)$, and W^t is the set of finite sequences $(w_1, \ldots, w_t)$, where $(\forall i)\, w_i \in W$. Second, let $\mathscr{B}_{W^\infty}$ be the σ-algebra on W^∞ generated by the family of sets of the form $E_1 \times E_2 \times \cdots$, and let $\mathscr{B}_{(W^t)}$ be the σ-algebra on W^t generated by the family of sets of the form $E_1 \times \cdots \times E_t$, where for each i, $E_i \in \mathscr{B}_W$, that is, E_i is a Borel set. Since W is a separable metric space, $\mathscr{B}_{(W^t)}$ is identical to $(\mathscr{B}_W)^t \equiv \mathscr{B}_W \otimes \cdots \otimes \mathscr{B}_W$, the t-dimensional

Let θ be a *capacitary kernel*, that is, let $\theta : W \times \mathscr{B}_W \to [0,1]$ be a function such that

$$(\forall w \in W) \quad \theta_w \text{ is a probability capacity on } (W, \mathscr{B}_W) \quad \text{and}$$
$$(\forall B \in \mathscr{B}_W) \quad \theta_\cdot(B) \text{ is } \mathscr{B}_W\text{-measurable.}$$

Throughout the paper, we assume that $(\forall w)\ \theta_w$ *is convex and continuous.* We specify the *uncertainty* about the offer of the next period when the current wage offer is w by the core of θ_w. That is, we assume that the offer in each period is 'distributed' in a Markovian manner according to core(θ_w). While we now allow that the uncertainty is Markovian, we still retain the assumption of no learning as in Section 2.2 by restricting θ to be time-homogeneous. To incorporate a reasonable learning process into the case of the Markovian-Knightian uncertainty is an important agenda of future research.

4.2. Objective Function

An *income process* ${}_0y = (y_0, y_1, y_2, \dots)$ is a $\mathbb{R}_+$-valued stochastic process which is $\langle \mathscr{B}_{W^t} \rangle$-adapted, that is, which satisfies $(\forall t \geq 0)\ y_t$ is $\mathscr{B}_{W^t}$-measurable. Given an income process ${}_0y$, we denote the *continuation* of ${}_0y$ after the realization of a history ${}_1w_t$ by ${}_ty|_{{}_1w_t}$:

$${}_ty|_{{}_1w_t} = (y_t({}_1w_t), y_t({}_1w_t, \cdot), y_t({}_1w_t, \cdot, \cdot), \dots).$$

Obviously, the continuation ${}_ty|_{{}_1w_t}$ is $\langle \mathscr{B}_{W^t} \rangle$-adapted given ${}_1w_t$.

Given any adapted income process ${}_0y$ and an initial wage offer $w_0 \in W$, we define the

product measurable space of $\mathscr{B}_W$. Third and finally, define the σ-algebra $\hat{\mathscr{B}}_{W^t}$ on W^∞ (*not* on W^t) as the σ-algebra generated by the family of *cylinder sets* $E_1 \times \cdots \times E_t \times W \times W \times \cdots$, where $(\forall i)\ E_i$ is a Borel set. In particular, $\hat{\mathscr{B}}_{W^0} = \{\phi, W^\infty\}$ represents *no* information. Then, any function defined on W^∞ which is $\hat{\mathscr{B}}_{W^t}$-measurable takes on the same value given the realization of $(w_1, \dots, w_t)$ regardless of the realization of $(w_{t+1}, w_{t+2}, \dots)$, and hence, it can be identified with the function defined on $\mathscr{B}_{(W^t)}$. Exactly in this manner, we can *embed* $\mathscr{B}_{(W^t)}$ in $\hat{\mathscr{B}}_{W^t}$. Therefore, we do not distinguish these two objects and use the notation $\mathscr{B}_{W^t}$ to represent both. This convention is convenient when we consider stopping rules which is defined on W^∞.

lifetime expected income $I_{w_0}({}_0y)$ by[18]

$$I_{w_0}({}_0y) = \lim_{T\to\infty} y_0 + \beta \int_W \left(y_1 + \beta \int_W \left(y_2 + \cdots \beta \int_W y_T \,\theta(dw_T) \cdots \right) \theta(dw_2) \right) \theta_{w_0}(dw_1) \tag{16}$$

where $\beta \in (0,1)$ is the discount factor and $\int_W \cdot \, d\theta$ is the *Choquet integral* with respect to a capacitary kernel θ. Note that each element of the sequence defining I is well-defined by the continuity of θ and by the Fubini property (Lemma A13), and that the limit exists (allowing $+\infty$) since the sequence is non-decreasing by the nonnegativity of y_t's and by Lemma A4. Furthermore, θ's continuity and the monotone convergence theorem (Lemma A12) imply that

$$(\forall {}_0y)(\forall w_0) \quad I_{w_0}({}_0y) = y_0 + \beta \int_W I_{w_1}({}_1y|_{w_1}) \theta_{w_0}(dw_1) \,,$$

which is called *Koopmans' equation.*

When it happens to be the case that θ is a stochastic kernel, Eq (16) is reduced to

$$\begin{aligned} I_{w_0}({}_0y) &= \lim_{T\to\infty} y_0 + \beta \int_W \left(y_1 + \beta \int_W \left(y_2 + \cdots \beta \int_W y_T \, d\theta \cdots \right) d\theta \right) d\theta_{w_0} \\ &= \lim_{T\to\infty} y_0 + \beta \int_W y_1 \, d\theta_{w_0} + \beta^2 \int_W \int_W y_2 \, d\theta \, d\theta_{w_0} \\ &\qquad + \cdots + \beta^T \int_W \int_W \cdots \int_W y_T \, d\theta \cdots d\theta \, d\theta_{w_0} \\ &= E_{w_0} \sum_{t=0}^{\infty} \beta^t y_t \,, \end{aligned}$$

where the expectation operator E in the last line is taken with respect to the infinite-dimensional product probability measure constructed from θ. When θ is not a stochastic kernel, the second equality may not hold since y_t and $\int_W y_{t+1} \theta(dw_{t+1})$ may not be co-monotonic (see Lemmas A7 and A9), and the third equality may not hold since the 'product capacity' is not well-defined uniquely.[19]

---

[18]In Eq (16), we suppressed the arguments of the integrand. If we did not, it would be

$$\begin{aligned} &I_{w_0}({}_0y) \\ &= \lim_{T\to\infty} y_0 + \beta \int_W \left(y_1(w_1) + \beta \int_W \left(y_2(w_1, w_2) + \cdots \beta \int_W y_T({}_1w_T) \, \theta_{w_{T-1}}(dw_T) \cdots \right) \theta_{w_1}(dw_2) \right) \theta_{w_0}(dw_1) \,. \end{aligned}$$

Here, the t-th integral from the most inside is a function of $(w_1, \ldots, w_{T-t})$. Then, the whole integral is a real number since w_0 is given. This definition implies that the randomness will be aggregated backward from the future to the current period one by one in each period.

[19]Let $(X, \mathscr{X})$ and $(Y, \mathscr{Y})$ be two measurable spaces, let $(Z, \mathscr{Z}) = (X \times Y, \mathscr{X} \otimes \mathscr{Y})$ be the product measurable space, and let μ and ν be a capacity on $\mathscr{X}$ and $\mathscr{Y}$, respectively. Consider a capacity σ on $\mathscr{Z}$ which satisfies

$$(\forall S \in \mathscr{X})(\forall T \in \mathscr{Y}) \quad \sigma(S \times T) = \mu(S)\nu(T) \,.$$

4.3. Stopping Rule and Optimization Problem

Each period the prospective worker is given an offer w. Upon observing the value of w, she has two alternatives, to accept it or to reject it. If she accepts the offer, she will obtain w each period from that period on; if she rejects the offer, she will get the unemployment compensation $c > 0$ that period and will be given a random offer again in the next period.

A $\{0,1,2,\dots\} \cup \{+\infty\}$-valued random variable δ on $(W^\infty, \mathscr{B}_{W^\infty})$ is called a *stopping rule* if it satisfies

$$(\forall t = 0,1,2,\dots) \quad \{\delta = t\} \in \mathscr{B}_{W^t}$$

where $\{\delta = t\}$ abbreviates $\{{}_1w \,|\, \delta({}_1w) = t\}$. We allow δ to be $+\infty$ for some history. We denote the set of all stopping rules by Δ. Given any stopping rule $\delta \in \Delta$, define a process ${}_0y^\delta = (y_0^\delta, y_1^\delta, y_2^\delta, \dots)$ by

$$(\forall t \geq 0) \quad y_t^\delta = \begin{cases} c & \text{if} \quad \delta > t \\ w_T & \text{if} \quad \delta = T \quad (T = 0,1,\dots,t) \end{cases}$$

Lemma 1 (Section 5) shows that ${}_0y^\delta$ is $\langle \mathscr{B}_{W^t} \rangle$-adapted, and hence it is actually an income process. Given an initial wage offer $w_0 \in W$, we denote the lifetime expected income under a stopping rule $\delta \in \Delta$ by the symbol I for notational simplicity (there should be no confusion about this):

$$I_{w_0}(\delta) = I_{w_0}({}_0y^\delta).$$

Similarly, given any $t \geq 1$ and any history ${}_1w_t \in W^t$, we denote the income under $\delta \in \Delta$ after the realization of ${}_1w_t$ by $I_{w_t}(\delta|_{{}_1w_t})$, that is,

$$I_{w_t}(\delta|_{{}_1w_t}) = I_{w_t}({}_ty^\delta|_{{}_1w_t}),$$

where ${}_ty^\delta|_{{}_1w_t}$ is the continuation of y^δ after the realization of ${}_1w_t$ as is defined in Section 4.2. Lemma 2 (Section 5) proves that

$$I_{w_0}(\delta) = \frac{w_0}{1-\beta}\chi_{\{\delta=0\}} + \left(c + \beta \int I_{w_1}(\delta|_{w_1})\, d\theta_{w_0}\right)\chi_{\{\delta>0\}}. \tag{17}$$

If both of μ and ν are probability charges, such a product capacity σ is uniquely determined. However, if at least one of them is not additive, there could be many capacities which satisfy the above relation. This implies that the 'product capacity' cannot be determined uniquely from the 'marginal capacity.' For more details, see Ghirardato (1997).

Here, χ denotes the indicator function on W^∞.[20] Eq (17) is Koopmans' equation for a stopping rule δ.

A stopping rule $\delta \in \Delta$ is *optimal* from w_0 if

$$\delta \in \arg\max\{ I_{w_0}(\delta) \mid \delta \in \Delta \}.$$

A stopping rule δ is *admissible* if it dictates more search as long as the observed offer is strictly less than c. Any stopping rule which is not admissible is suboptimal since it is dominated by the stopping rule which never stops, and hence, it can be safely ignored. When an optimal stopping rule exists, we define the *value function* $V^* : W \to \bar{\mathbb{R}}_+$ by

$$(\forall w \in W) \quad V^*(w) = I_w(\delta^*_w),$$

where we denote an optimal stopping rule from w by δ^*_w.

4.4. Existence and Characterization of Optimal Stopping Rule

Given random variables w, w_0, $\ldots$, w_t, let $c \vee w$ and $\vee_{i=0}^t w_i$ denote the random variables defined by $\max\{c, w\}$ and $\max\{w_1, \ldots, w_t\}$. Throughout the rest of this paper, we assume that the primitives of the model satisfy the following two conditions:

$$\textbf{E1.} \quad (\forall w_0)(\forall t > 0) \quad \bar{W}^t(w_0) \equiv \int \cdots \int \left(c \vee \vee_{i=0}^t w_i\right) \theta'(dw_t) \cdots \theta'_{w_0}(dw_1) < +\infty$$

$$\textbf{E2.} \quad (\forall w_0) \quad \overline{\lim}_{t \to \infty} \left(\bar{W}^t(w_0)\right)^{1/t} < \beta^{-1}$$

where θ' is the conjugate of θ. The integrand in E1, $c \vee \vee_{i=0}^t w_i$, is the overly optimistic income the worker expects in time t. This is overly optimistic because it is the highest offer up to time t (our model is on search *without* recall). The integral in E1 is its overly optimistic 'expectation' evaluated at time 0. This is overly optimistic because it is evaluated by the *conjugate* of θ rather

[20] For example,

$$\chi_{\{\delta=t\}} = \begin{cases} 1 & \text{if} \quad \omega \in \{\delta = t\}, \quad \textit{i.e.}, \quad \delta(\omega) = t \\ 0 & \text{if} \quad \omega \notin \{\delta = t\}, \quad \textit{i.e.}, \quad \delta(\omega) \neq t. \end{cases}$$

Since δ is a stopping rule, $\chi_{\{\delta=t\}}$ and $\chi_{\{\delta>t\}}$ are $\mathscr{B}_{W^t}$-measurable.

than θ itself.[21] E1 assumes that this is finite for any t. This optimistic 'expected' income, $\bar{W}^t$, grows as t increases since it takes the maximum offer up to time t. The left-hand-side of E2 defines the time-average of the rate of growth in $\bar{W}^t$. Hence, E2 as a whole assumes that this time-average is lower than the worker's impatience. When E2 holds, the effect caused by the high income in a far future can be safely ignored since the worker's impatience dominates the income growth along any optimistic path. This is an analogue to the condition for the dynamic programming technique introduced in Ozaki and Streufert (1996). If θ is simply a probability measure, the left-hand side of E2 is 1 as long as the expectation of w is finite (Chung, 1974, p.49), and hence, E2 is automatically satisfied.

Define a (constant) function $V^- : W \to \mathbb{R}_+$ by $(\forall w)\, V^-(w) = c/(1-\beta)$ and a function $V^+ : W \to \bar{\mathbb{R}}_+$ by $(\forall w_0)$

$$V^+(w_0) = \lim_{T \to \infty} (c \vee w_0) + \beta \int \left((c \vee \vee_{i=0}^{1} w_i) + \cdots \beta \int (c \vee \vee_{i=0}^{T} w_i)\, \theta(dw_T) \cdots \right) \theta_{w_0}(dw_1),$$

which is a well-defined $\mathscr{B}_W$-measurable function (let $y_T = c \vee \vee_{i=0}^{T} w_i$ in Eq (16)). Clearly, $V^- \le V^+$, and Lemma 3 (Section 5) shows that $(\forall w_0)\, V^+(w_0) < +\infty$.

A $\mathscr{B}_W$-measurable function $V : W \to \mathbb{R}_+$ is *admissible* if it satisfies $V^- \le V \le V^+$. Note that for any admissible stopping rule δ, $I_{\cdot}(\delta)$ is admissible. Let $\mathcal{V}$ be the space of all admissible functions from W into $\mathbb{R}_+$, and let B be the operator from $\mathcal{V}$ into itself defined by

$$(\forall V \in \mathcal{V})(\forall w \in W) \quad BV(w) = \max \left\{ \frac{w}{1-\beta},\ c + \beta \int_W V(w')\theta_w(dw') \right\}. \tag{18}$$

Lemma 4 (Section 5) shows that BV is admissible for any admissible function V, and hence, that B is well-defined.

A function $V \in \mathcal{V}$ is said to *solve Bellman's equation* if $BV = V$. We then have the main result of this paper, summarized in the following theorem. The proof of this theorem and those of other theorems and corollaries are relegated to Section 5.

Theorem 1. *The value function V^* exists and is the unique admissible solution to Bellman's equation. Furthermore, V^* is attained by the stopping rule δ^* such that for all $t \ge 0$, $\delta^* = t$*

[21] When we prove that the solution to Bellman's equation (specified later) is the value function, we apply the method of "squeezing" (see Lemma 6 in Section 5). In order to "squeeze", we need to bound the increment by the Choquet integral with respect to the conjugate (Lemma A11). This is why we need to define the 'expected' income via the conjugate capacity in E1.

as soon as

$$\frac{w_t}{1-\beta} \geq c + \beta \int_W V^*(w_{t+1})\theta_{w_t}(dw_{t+1})$$

holds; and $\delta^* > t$ *otherwise.*

Let $R : W \to \mathbb{R}_+$ be a $\mathscr{B}_W$-measurable function defined by

$$(\forall w) \quad R(w) = (1-\beta)\left(c + \beta \int_W V^*(w')\theta_w(dw')\right).$$

We call $R(w)$ *reservation wage* at a state w, that is, when w is observed. We say that the capacitary kernel θ is *monotonic* if for any weakly increasing function $y : W \to \mathbb{R}_+$ and for any $x \geq 0$,

$$w' \geq w \Rightarrow \theta_{w'}(\{y \geq x\}) \geq \theta_w(\{y \geq x\}).$$

The next result is an extension of Lippman and MaCall (1976, Theorem 1).

Corollary 1. *If* θ *is monotonic, then* R *is weakly increasing in* w.

The next result further characterizes the reservation wage when the capacitary kernel θ is *i.i.d.*, that is, when θ is independent of the current wage offer w.

Corollary 2. *If the capacitary kernel* θ *is independent of the current wage offer, then the reservation wage* $R(w)$ *will be constant and is given by the solution* R *to the next equation:*

$$\begin{aligned} R &= c + \frac{\beta}{1-\beta}\int_R^\infty \theta(\{w \mid w \geq x\})\,dx \\ &= c + \frac{\beta}{1-\beta}\int_R^\infty \left[1 - F_{\theta'}(x)\right]dx, \end{aligned}$$

where $F_{\theta'}$ *is the "distribution" of* θ' *defined by* $F_{\theta'}(x) = \theta'(\{w \mid w \leq x\})$.

As an application of Corollary 2, we show that (11) characterizes the reservation wage for sufficiently small ε and δ in the Example of the δ-approximation of ε-contamination provided in Section 2.2. First, let $\varepsilon > 0$ be small enough to be such that $R(\varepsilon) > a$ where $R(\varepsilon)$ is the solution to (11). This is possible since $R(\cdot)$ is continuous and $R(0) > a$ by (2) and (3). Second, let $\delta > 0$ be small enough to be such that $a + (b-a)\delta < R(\varepsilon)$. This is possible since $R(\varepsilon) > a$. Third, note that for any $R > a + (b-a)\delta$, it holds that

$$c + \frac{\beta}{1-\beta}\int_R^\infty \theta_\delta(\{w \mid w \geq x\})\,dx = c + \frac{\beta}{1-\beta}\int_R^\infty (1-\varepsilon)P_0(\{w \mid w \geq x\})\,dx$$

by the definition of θ_δ. Finally, note that

$$R(\varepsilon) = c + \frac{\beta}{1-\beta}\int_{R(\varepsilon)}^\infty \theta_\delta(\{w \mid w \geq x\})\,dx. \tag{19}$$

This holds because $R(\varepsilon)$ solves (11) and becasue $R(\varepsilon) > a + (b-a)\delta$. Therefore, the continuity of θ_δ, Corollary 2 and (19) imply that $R(\varepsilon)$ in ((11)) is certainly the reservation wage.

4.5. Increase in Uncertainty

Let θ^1 and θ^2 be two capacitary kernels. According to Epstein and Zhang (1999), we say that θ^2 *represents more Knightian uncertainty than* θ^1 if there exists a weakly increasing, surjective and convex function $g : [0,1] \to [0,1]$ such that

$$(\forall w)(\forall B) \quad \theta_w^2(B) = g(\theta_w^1(B)).$$

It immediately follows that if θ_w^1 is convex and continuous for each w, which we henceforth assume, and if θ^2 represents more Knightian uncertainty than θ^1, then θ_w^2 is also convex and continuous for each w. It also follows that if θ^2 represents more Knightian uncertainty than θ^1, then

$$(\forall w) \quad \mathrm{core}(\theta_w^2) \supseteq \mathrm{core}(\theta_w^1), \tag{20}$$

which justifies our definition of more Knightian uncertainty[22]. Let R^1 and R^2 be a reservation wage of an unemployed worker with θ^1 and θ^2, respectively. The next result shows that the reservation wage decreases if the Knightian uncertainty increases.

Theorem 2. *If θ^2 represents more Knightian uncertainty than θ^1, then* $(\forall w)\, R^2(w) \leq R^1(w)$.

[22] For further motivation of this definition of more Knightian uncertainty, see Epstein and Zhang (1999, Theorem 3.1).

As an application of Theorem 2, we show that an increase in ε decreases the reservation wage for any $\delta > 0$ in the Example of the δ-approximation of ε-contamination provided in Section 2.2. (Recall that we made only a local analysis there.) Let $\delta > 0$ and let $\varepsilon_2 > \varepsilon_1 > 0$. Write $\mathcal{P}_\delta$ explicitly as $\mathcal{P}_\delta^\varepsilon$ to denote its dependence on ε. Finally, suppose that θ_δ^i is the probability capacity corresponding to $\mathcal{P}_\delta^{\varepsilon_i}$. Then, it turns out that

$$(\forall A) \quad \theta_\delta^2(A) = \begin{cases} \dfrac{1-\varepsilon_2}{1-\varepsilon_1}\theta_\delta^1(A) & \text{if} \quad \theta_\delta^1(A) \leq (1-\varepsilon_1)(1-\delta) \\ \dfrac{\delta-\varepsilon_2\delta+\varepsilon_2}{\delta-\varepsilon_1\delta+\varepsilon_1}\theta_\delta^1(A) - \dfrac{(\varepsilon_2-\varepsilon_1)(1-\delta)}{\delta-\varepsilon_1\delta+\varepsilon_1} & \text{if} \quad \theta_\delta^1(A) > (1-\varepsilon_1)(1-\delta), \end{cases}$$

which shows that θ_δ^2 is a convex transformation of θ_δ^1, and hence, θ_δ^2 represents more Knightian uncertainty than θ_δ^1. Then, Theorem 2 shows that an increase in ε decreases the reservation wage.

5. Lemmas and Proofs

Lemma 1. ${}_0y^\delta$ *is* $\langle \mathscr{B}_{W^t} \rangle$*-adapted.*

Proof. This is immediate since for any $t \geq 0$, $y_t^\delta = \sum_{T=0}^t w_T \chi_{\{\delta=T\}} + c\chi_{\{\delta>t\}}$ and the components in the right-hand side are all $\mathscr{B}_{W^t}$-measurable. □

Lemma 2. *For any* $\delta \in \Delta$, *it holds that*

$$I_{w_0}(\delta) = \frac{w_0}{1-\beta}\chi_{\{\delta=0\}} + \left(c + \beta \int I_{w_1}(\delta|_{w_1})\, d\theta_{w_0}\right)\chi_{\{\delta>0\}}.$$

Proof. This holds because

$$\begin{aligned}
& I_{w_0}(\delta) \\
= \; & \lim_{T\to\infty} y_0^\delta + \beta \int \left(y_1^\delta + \cdots \beta \int y_T^\delta\, d\theta \cdots\right) d\theta_{w_0} \\
= \; & \lim_{T\to\infty} \left[\left(y_0^\delta + \beta \int \left(y_1^\delta + \cdots \beta \int y_T^\delta\, d\theta \cdots\right) d\theta_{w_0}\right)\chi_{\{\delta=0\}} \right. \\
& \left. + \left(y_0^\delta + \beta \int \left(y_1^\delta + \cdots \beta \int y_T^\delta\, d\theta \cdots\right) d\theta_{w_0}\right)\chi_{\{\delta>0\}}\right] \\
= \; & \lim_{T\to\infty} \left[\left(w_0 + \beta \int \left(w_0 + \cdots \beta \int w_0\, d\theta \cdots\right) d\theta_{w_0}\right)\chi_{\{\delta=0\}} \right.
\end{aligned}$$

$$
\begin{aligned}
&\quad + \left(c+\beta \int \left(y_1^\delta + \cdots \beta \int y_T^\delta \, d\theta \cdots \right) d\theta_{w_0}\right) \chi_{\{\delta>0\}} \Bigg] \\
&= \lim_{T\to\infty} \left[\left(\frac{1-\beta^{T+1}}{1-\beta} w_0\right) \chi_{\{\delta=0\}} + \left(c+\beta \int \left(y_1^\delta + \cdots \beta \int y_T^\delta \, d\theta \cdots \right) d\theta_{w_0}\right) \chi_{\{\delta>0\}}\right] \\
&= \frac{w_0}{1-\beta} \chi_{\{\delta=0\}} + \left(c+\beta \lim_{T\to\infty} \int \left(y_1^\delta + \cdots \beta \int y_T^\delta \, d\theta \cdots \right) d\theta_{w_0}\right) \chi_{\{\delta>0\}} \\
&= \frac{w_0}{1-\beta} \chi_{\{\delta=0\}} + \left(c+\beta \int \lim_{T\to\infty} \left(y_1^\delta + \cdots \beta \int y_T^\delta \, d\theta \cdots \right) d\theta_{w_0}\right) \chi_{\{\delta>0\}} \\
&= \frac{w_0}{1-\beta} \chi_{\{\delta=0\}} + \left(c+\beta \int I_{w_1}(\delta|_{w_1}) d\theta_{w_0}\right) \chi_{\{\delta>0\}},
\end{aligned}
$$

where the last equality but one holds by the monotone convergence theorem (Lemma A12). □

Lemma 3. $(\forall w_0 \in W)\ V^+(w_0) < +\infty$.

Proof. Let $w_0 \in W$, and let δ and $\hat{t}$ be such that

$$
\delta > \beta \quad \text{and} \quad (\forall t \geq \hat{t}) \ \left(\bar{W}^t(w_0)\right)^{1/t} < \delta^{-1}.
$$

Such δ and $\hat{t}$ exist by E2. Then,

$$
\begin{aligned}
& V^+(w_0) \\
\leq\ & \lim_{T\to\infty} (c \vee w_0) + \beta \int (c \vee \vee_{i=0}^1 w_i) d\theta'_{w_0} + \cdots + \beta^T \int \cdots \int (c \vee \vee_{i=0}^T w_i) d\theta'_{w_{T-1}} \cdots d\theta'_{w_0} \\
=\ & \lim_{T\to\infty} \bar{W}^0(w_0) + \beta \bar{W}^1(w_0) + \cdots + \beta^T \bar{W}^T(w_0) \\
\leq\ & \lim_{T\to\infty} \bar{W}^0(w_0) + \beta \bar{W}^1(w_0) + \cdots + \beta^{\hat{t}-1} \bar{W}^{\hat{t}-1}(w_0) + \beta^{\hat{t}} \delta^{-\hat{t}} + \cdots + \beta^T \delta^{-T} \\
=\ & \bar{W}^0(w_0) + \beta \bar{W}^1(w_0) + \cdots + \beta^{\hat{t}-1} \bar{W}^{\hat{t}-1}(w_0) + \frac{(\beta/\delta)^{\hat{t}}}{1-\beta/\delta},
\end{aligned}
$$

where the first inequality holds by Lemmas A2, A9 and A10. The last line of the whole inequality is finite by E1 and the fact that $\delta > \beta$. □

Lemma 4. *$BV^- \geq V^-$, $BV^+ \leq V^+$, and for any admissible function $V \in \mathcal{V}$, BV is admissible.*

Proof. The first claim holds because $(\forall w_0)$

$$
\begin{aligned}
BV^-(w_0) &= \max\left\{\frac{w_0}{1-\beta},\ c+\beta \int V^-(w_1)\theta_{w_0}(dw_1)\right\} \\
&= \max\left\{\frac{w_0}{1-\beta},\ \frac{c}{1-\beta}\right\}
\end{aligned}
$$

$$\geq \quad V^{-}(w_0)\,.$$

The secnd claim holds because $(\forall w_0)$

$$
\begin{aligned}
& BV^{+}(w_0) \\
= \;& \max\left\{ \frac{w_0}{1-\beta},\, c+\beta \int V^{+}(w_1)\theta_{w_0}(dw_1) \right\} \\
= \;& \max\left\{ \frac{w_0}{1-\beta},\, c+\beta \int \lim_{T \to \infty} \left((c \vee w_1) + \cdots \beta \int (c \vee \vee_{i=1}^{T} w_i) d\theta \cdots \right) d\theta_{w_0} \right\} \\
= \;& \max\left\{ \frac{w_0}{1-\beta},\, \lim_{T \to \infty} c+\beta \int \left((c \vee w_1) + \cdots \beta \int (c \vee \vee_{i=1}^{T} w_i) d\theta \cdots \right) d\theta_{w_0} \right\} \\
\leq \;& \max\left\{ \frac{w_0}{1-\beta},\, \lim_{T \to \infty} (c \vee w_0) + \beta \int \left((c \vee \vee_{i=0}^{1} w_i) + \cdots \beta \int (c \vee \vee_{i=0}^{T} w_i) d\theta \cdots \right) d\theta_{w_0} \right\} \\
\leq \;& \max\left\{ \frac{w_0}{1-\beta},\, V^{+}(w_0) \right\} \\
= \;& V^{+}(w_0)\,,
\end{aligned}
$$

where the third equality holds by the monotone convergence theorem (Lemma A12). The final claim follows from the first two claims and the fact that $BV \leq BV'$ whenever $V \leq V'$. $\square$

Lemma 5. *For any* $w_0 \in W$,

$$\lim_{t \to \infty} \beta^{t} \int \int \cdots \int V^{+}(w_t) d\theta'_{w_{t-1}} \cdots d\theta'_{w_1} d\theta'_{w_0} = 0\,.$$

Proof. Let $w_0 \in W$, and let δ and $\hat{t}$ be such that

$$\delta > \beta \quad \text{and} \quad (\forall t \geq \hat{t}) \;\; \left(\bar{W}^{t}(w_0)\right)^{1/t} < \delta^{-1}\,.$$

Such δ and $\hat{t}$ exist by E2. Then, for any $t \geq \hat{t}$,

$$
\begin{aligned}
& \beta^{t} \int \cdots \int V^{+}(w_t) d\theta'_{w_{t-1}} \cdots d\theta'_{w_0} \\
\leq \;& \beta^{t} \int \cdots \int \lim_{T \to \infty} \Bigg[(c \vee w_t) + \beta \int (c \vee \vee_{i=0}^{1} w_{t+i}) d\theta'_{w_t} \\
& + \cdots + \beta^{T} \int \cdots \int (c \vee \vee_{i=0}^{T} w_{t+i}) d\theta'_{w_{t+T-1}} \cdots d\theta'_{w_t} \Bigg] d\theta'_{w_{t-1}} \cdots d\theta'_{w_0} \\
= \;& \lim_{T \to \infty} \beta^{t} \int \cdots \int \Bigg[(c \vee w_t) + \beta \int (c \vee \vee_{i=0}^{1} w_{t+i}) d\theta'_{w_t} \\
& + \cdots + \beta^{T} \int \cdots \int (c \vee \vee_{i=0}^{T} w_{t+i}) d\theta'_{w_{t+T-1}} \cdots d\theta'_{w_t} \Bigg] d\theta'_{w_{t-1}} \cdots d\theta'_{w_0}
\end{aligned}
$$

$$
\begin{aligned}
&\le \lim_{T\to\infty} \beta^t \int\cdots\int \Big[(c \vee \vee_{i=0}^{t} w_i) + \beta \int (c \vee \vee_{i=0}^{t+1} w_i)\, d\theta'_{w_t} \\
&\qquad + \cdots + \beta^T \int \cdots \int (c \vee \vee_{i=0}^{t+T} w_i)\, d\theta'_{w_{t+T-1}} \cdots d\theta'_{w_t} \Big]\, d\theta'_{w_{t-1}} \cdots d\theta'_{w_0} \\
&\le \lim_{T\to\infty} \beta^t \int\cdots\int (c \vee \vee_{i=0}^{t} w_i)\, d\theta'_{w_{t-1}} \cdots d\theta'_{w_0} + \beta^{t+1} \int\cdots\int (c \vee \vee_{i=0}^{t+1} w_i)\, d\theta'_{w_t} \cdots d\theta'_{w_0} \\
&\qquad + \cdots + \beta^{t+T} \int\cdots\int (c \vee \vee_{i=0}^{t+T} w_i)\, d\theta'_{w_{t+T-1}} \cdots d\theta'_{w_0} \\
&= \lim_{T\to\infty} \beta^t \bar{W}^t(w_0) + \beta^{t+1} \bar{W}^{t+1}(w_0) + \cdots + \beta^{t+T} \bar{W}^{t+T}(w_0) \\
&\le \lim_{T\to\infty} \beta^t \delta^{-t} + \beta^{t+1} \delta^{-(t+1)} + \cdots + \beta^{t+T} \delta^{-(t+T)} \\
&= \frac{(\beta/\delta)^t}{1-\beta/\delta},
\end{aligned}
$$

where the first inequality holds by Lemmas A2, A9 and A10; the first equality holds by the monotone convergence theorem (Lemma A12); and the third inequality holds by Lemma A9. The last line of the whole inequality converges to 0 as $t \to \infty$ by the fact that $\delta > \beta$. □

Lemma 6. *For any $\delta \in \Delta$ and for any admissible function V,*

$$
\begin{aligned}
\lim_{t\to\infty} \Bigg| I_{w_0}(\delta) - \Bigg[\frac{w_0}{1-\beta} \chi_{\{\delta=0\}} + \Bigg(c + \beta \int \Bigg(\frac{w_1}{1-\beta} \chi_{\{\delta=1\}} + \Bigg(c + \cdots \beta \int \Bigg(\frac{w_{t-1}}{1-\beta} \chi_{\{\delta=t-1\}} \\
+ \Big(c + \beta \int V(w_t)\, d\theta_{w_{t-1}} \Big) \chi_{\{\delta > t-1\}} \Bigg) d\theta_{w_{t-2}} \cdots \Bigg) \chi_{\{\delta>1\}} \Bigg) d\theta_{w_0} \Bigg) \chi_{\{\delta>0\}} \Bigg] \Bigg| = 0
\end{aligned}
$$

Proof. By the iterative applications of Eq (17), we have for any $t > 0$,

$$
\begin{aligned}
I_{w_0}(\delta) &= \frac{w_0}{1-\beta} \chi_{\{\delta=0\}} + \Big(c + \beta \int I_{w_1}(\delta|_{w_1})\, d\theta_{w_0} \Big) \chi_{\{\delta>0\}} \\
&= \frac{w_0}{1-\beta} \chi_{\{\delta=0\}} + \Bigg(c + \beta \int \Bigg(\frac{w_1}{1-\beta} \chi_{\{\delta=1\}} \\
&\quad + \Big(c + \beta \int I_{w_2}(\delta|_{1 w_2})\, d\theta_{w_1} \Big) \chi_{\{\delta>1\}} \Bigg) d\theta_{w_0} \Bigg) \chi_{\{\delta>0\}} \\
&= \cdots\cdots\cdots \\
&= \frac{w_0}{1-\beta} \chi_{\{\delta=0\}} + \Bigg(c + \beta \int \Bigg(\frac{w_1}{1-\beta} \chi_{\{\delta=1\}} + \Bigg(c + \cdots \beta \int \Bigg(\frac{w_{t-1}}{1-\beta} \chi_{\{\delta=t-1\}} \\
&\quad + \Big(c + \beta \int I_{w_t}(\delta|_{1 w_t})\, d\theta_{w_{t-1}} \Big) \chi_{\{\delta>t-1\}} \Bigg) d\theta_{w_{t-2}} \cdots \Bigg) \chi_{\{\delta>1\}} \Bigg) d\theta_{w_0} \Bigg) \chi_{\{\delta>0\}}.
\end{aligned}
$$

Therefore, for any $t > 0$,

$$
\Bigg| I_{w_0}(\delta) - \Bigg[\frac{w_0}{1-\beta} \chi_{\{\delta=0\}} + \Bigg(c + \beta \int \Bigg(\frac{w_1}{1-\beta} \chi_{\{\delta=1\}} + \Bigg(c + \cdots \beta \int \Bigg(\frac{w_{t-1}}{1-\beta} \chi_{\{\delta=t-1\}}
$$

$$
\begin{aligned}
&\qquad + \Bigg(c+\beta\int V(w_t)d\theta_{w_{t-1}}\Bigg)\chi_{\{\delta>t-1\}}\Bigg)d\theta_{w_{t-2}}\cdots\Bigg)\chi_{\{\delta>1\}}\Bigg)d\theta_{w_0}\Bigg)\chi_{\{\delta>0\}}\Bigg]\Bigg| \\
= \;& \beta\Bigg|\int\Bigg(\frac{w_1}{1-\beta}\chi_{\{\delta=1\}}+\Bigg(c+\cdots\beta\int\Bigg(\frac{w_{t-1}}{1-\beta}\chi_{\{\delta=t-1\}} \\
&\qquad + \Bigg(c+\beta\int I_{w_t}(\delta|_{1w_t})d\theta_{w_{t-1}}\Bigg)\chi_{\{\delta>t-1\}}\Bigg)d\theta_{w_{t-2}}\cdots\Bigg)\chi_{\{\delta>1\}}\Bigg)d\theta_{w_0} \\
&- \int\Bigg(\frac{w_1}{1-\beta}\chi_{\{\delta=1\}}+\Bigg(c+\cdots\beta\int\Bigg(\frac{w_{t-1}}{1-\beta}\chi_{\{\delta=t-1\}} \\
&\qquad + \Bigg(c+\beta\int V(w_t)d\theta_{w_{t-1}}\Bigg)\chi_{\{\delta>t-1\}}\Bigg)d\theta_{w_{t-2}}\cdots\Bigg)\chi_{\{\delta>1\}}\Bigg)d\theta_{w_0}\Bigg|\chi_{\{\delta>0\}} \\
\leq \;& \beta\int\beta\Bigg|\int\Bigg(\frac{w_2}{1-\beta}\chi_{\{\delta=2\}}+\cdots\beta\int\Bigg(\frac{w_{t-1}}{1-\beta}\chi_{\{\delta=t-1\}} \\
&\qquad + \Bigg(c+\beta\int I_{w_t}\left(\delta|_{1w_t}\right)d\theta_{w_{t-1}}\Bigg)\chi_{\{\delta>t-1\}}\Bigg)\cdots\Bigg)d\theta_{w_1} \\
&- \int\Bigg(\frac{w_2}{1-\beta}\chi_{\{\delta=2\}}+\cdots\beta\int\Bigg(\frac{w_{t-1}}{1-\beta}\chi_{\{\delta=t-1\}} \\
&\qquad + \Bigg(c+\beta\int V(w_t)d\theta_{w_{t-1}}\Bigg)\chi_{\{\delta>t-1\}}\Bigg)\cdots\Bigg)d\theta_{w_1}\Bigg|\chi_{\{\delta>1\}}d\theta'_{w_0}\chi_{\{\delta>0\}} \\
\leq \;& \cdots\cdots\cdots \\
\leq \;& \beta\int\beta\int\cdots\beta\int\left|I_{w_t}\left(\delta|_{1w_t}\right)-V(w_t)\right|d\theta'_{w_{t-1}}\chi_{\{\delta>t-1\}}\cdots d\theta'_{w_1}\chi_{\{\delta>1\}}d\theta'_{w_0}\chi_{\{\delta>0\}} \\
= \;& \beta^t\int\int\cdots\int\left|I_{w_t}\left(\delta|_{1w_t}\right)-V(w_t)\right|\chi_{\{\delta>t-1\}}d\theta'_{w_{t-1}}\cdots d\theta'_{w_1}d\theta'_{w_0} \\
\leq \;& \beta^t\int\int\cdots\int\max\{I_{w_t}\left(\delta|_{1w_t}\right)\chi_{\{\delta>t-1\}},V(w_t)\}d\theta'_{w_{t-1}}\cdots d\theta'_{w_1}d\theta'_{w_0} \\
\leq \;& \beta^t\int\int\cdots\int V^+(w_t)\,d\theta'_{w_{t-1}}\cdots d\theta'_{w_1}d\theta'_{w_0},
\end{aligned}
$$

where a series of inequalities in the middle holds by successive applications of Lemma A11, and the last inequality holds by the admissibility of V. Finally, Lemma 5 completes the proof. $\square$

Lemma 7. *Any admissible solution to Bellman's equation is the value function.*

Proof. Let V be any admissible solution to Bellman's equation, and let $w_0 \in W$. This paragraph shows that $V(w_0) \geq I_{w_0}(\delta)$ for any $\delta \in \Delta$. Let $\delta \in \Delta$ be any stopping rule. Then, that V solves Bellman's equation implies that for any $t > 0$,

$$
\begin{aligned}
& V(w_0) \\
\geq \;& \frac{w_0}{1-\beta}\chi_{\{\delta=0\}}+\left(c+\beta\int V(w_1)d\theta_{w_0}\right)\chi_{\{\delta>0\}}
\end{aligned}
$$

$$
\begin{aligned}
&\geq \frac{w_0}{1-\beta}\chi_{\{\delta=0\}} + \Big(c+\beta\int\Big(\frac{w_1}{1-\beta}\chi_{\{\delta=1\}} + \Big(c+\beta\int V(w_2)d\theta_{w_1}\Big)\chi_{\{\delta>1\}}\Big)d\theta_{w_0}\Big)\chi_{\{\delta>0\}} \\
&\geq \cdots\cdots\cdots \\
&\geq \frac{w_0}{1-\beta}\chi_{\{\delta=0\}} + \Big(c+\beta\int\Big(\frac{w_1}{1-\beta}\chi_{\{\delta=1\}} + \Big(c+\cdots\beta\int\Big(\frac{w_{t-1}}{1-\beta}\chi_{\{\delta=t-1\}} \\
&\quad + \Big(c+\beta\int V(w_t)d\theta_{w_{t-1}}\Big)\chi_{\{\delta>t-1\}}\Big)d\theta_{w_{t-2}}\cdots\Big)\chi_{\{\delta>1\}}\Big)d\theta_{w_0}\Big)\chi_{\{\delta>0\}}.
\end{aligned}
$$

Hence, Lemma 6 proves the claim by the admissibility of V. This paragraph shows that there exists a stopping rule $\delta\in\Delta$ such that $V(w_0)=I_{w_0}(\delta)$. Let δ be the stopping rule such that for all $t\geq 0$, $\delta=t$ as soon as

$$\frac{w_t}{1-\beta}\geq c+\beta\int_W V(w_{t+1})\theta_{w_t}(dw_{t+1})$$

holds; and $\delta>t$ otherwise. Then, that V solves Bellman's equation implies that for any $t>0$,

$$
\begin{aligned}
&V(w_0) \\
&= \frac{w_0}{1-\beta}\chi_{\{\delta=0\}} + \Big(c+\beta\int V(w_1)d\theta_{w_0}\Big)\chi_{\{\delta>0\}} \\
&= \frac{w_0}{1-\beta}\chi_{\{\delta=0\}} + \Big(c+\beta\int\Big(\frac{w_1}{1-\beta}\chi_{\{\delta=1\}} + \Big(c+\beta\int V(w_2)d\theta_{w_1}\Big)\chi_{\{\delta>1\}}\Big)d\theta_{w_0}\Big)\chi_{\{\delta>0\}} \\
&= \cdots\cdots\cdots \\
&= \frac{w_0}{1-\beta}\chi_{\{\delta=0\}} + \Big(c+\beta\int\Big(\frac{w_1}{1-\beta}\chi_{\{\delta=1\}} + \Big(c+\cdots\beta\int\Big(\frac{w_{t-1}}{1-\beta}\chi_{\{\delta=t-1\}} \\
&\quad + \Big(c+\beta\int V(w_t)d\theta_{w_{t-1}}\Big)\chi_{\{\delta>t-1\}}\Big)d\theta_{w_{t-2}}\cdots\Big)\chi_{\{\delta>1\}}\Big)d\theta_{w_0}\Big)\chi_{\{\delta>0\}}.
\end{aligned}
$$

Again, Lemma 6 proves the claim by the admissibility of V. □

Lemma 8. *Both of* $\lim_{n\to\infty}B^nV^-$ *and* $\lim_{n\to\infty}B^nV^+$ *are admissible solutions to Bellman's equation.*

Proof. $\langle B^nV^-\rangle$ is weakly increasing and $\langle B^nV^+\rangle$ is weakly decreasing by Lemma 4, and hence, the limits exist. Lemma 4 also shows that these limits are admissible. Finally, the limits solve Bellman's equation by the monotone convergence theorem (Lemma A12). □

Proof of Theorem 1. The first half of the claim follows immediately from Lemmas 7 and 8. The second half of the claim also follows immediately from the proof of Lemma 7. □

Proof of Corollary 1. It suffices to show that V^* is weakly increasing in w by the monotonicity of θ and the definition of R. This, however, follows immediately from the facts that BV is weakly increasing whenever so is V by the monotonicity of θ, that V^- is weakly increasing (actually it's constant), and that $V^* = \lim_{n\to\infty} B^n V^-$. □

Proof of Corollary 2. By the definition of the reservation wage, R is constant. Furhermore, R satisfies the first equation because

$$\begin{aligned}
\frac{R}{1-\beta} &= c+\beta \int_W V^*(w)\theta(dw) \\
&= c+\beta \int_0^\infty \theta(\{w|V^*(w)\geq x\})\,dx \\
&= c+\beta V^*(R)+\beta \int_{V^*(R)}^\infty \theta(\{w|V^*(w)\geq x\})\,dx \\
&= c+\beta V^*(R)+\beta \int_{V^*(R)}^\infty \theta(\{w\,|\,w/(1-\beta)\geq x\})\,dx \\
&= c+\frac{\beta}{1-\beta}R+\beta \int_{R/(1-\beta)}^\infty \theta(\{w\,|\,w/(1-\beta)\geq x\})\,dx \\
&= c+\frac{\beta}{1-\beta}R+\frac{\beta}{1-\beta} \int_R^\infty \theta(\{w\,|\,w\geq x\})\,dx,
\end{aligned}$$

where the first equality holds by the definition of R; the second equality holds by the definition of Choquet integral; the third, fourth, and fifth equalities hold since V^* solves Bellman's equation; and the final equality holds by the change of variable. To show the second equality, note that for almost all x, it holds that $\theta(\{w\,|\,w\geq x\}) = \theta(\{w\,|\,w>x\})$ (see the proof of Lemma A12). Then, it follows that

$$\begin{aligned}
\int_R^\infty \theta(\{w\,|\,w\geq x\})\,dx &= \int_R^\infty \theta(\{w\,|\,w>x\})\,dx \\
&= \int_R^\infty \theta(\{w\,|\,w\leq x\}^c)\,dx \\
&= \int_R^\infty [1-(1-\theta(\{w\,|\,w\leq x\}^c))]\,dx \\
&= \int_R^\infty \left[1-\theta'(\{w\,|\,w\leq x\})\right]dx.
\end{aligned}$$

□

Proof of Theorem 2. For each $i = 1, 2$, let B^i and V^{i*} be the operator defined by (18) and the value function corresponding to θ^i. Then, $(\forall V)(\forall w)$

$$\begin{aligned}
B^2 V(w) &= \max\left\{ \frac{w}{1-\beta}\,,\, c+\beta \int_W V(w')\theta_w^2(dw') \right\} \\
&\leq \max\left\{ \frac{w}{1-\beta}\,,\, c+\beta \int_W V(w')\theta_w^1(dw') \right\} \\
&= B^1 V(w)\,,
\end{aligned}$$

where the inequality holds by Lemma A8 and (20). Therefore, it follows that $V^{2*} \leq V^{1*}$ by this, the fact that $B^i V' \geq B^i V$ whenever $V' \geq V$, and that $V^{i*} = \lim_{n \to \infty} (B^i)^n V$ by Lemmas 7 and 8. Finally, we conclude that $(\forall w)$

$$\begin{aligned}
R^2(w)/(1-\beta) &= c+\beta \int_W V^{2*}(w')\theta_w^2(dw') \\
&\leq c+\beta \int_W V^{1*}(w')\theta_w^2(dw') \\
&\leq c+\beta \int_W V^{1*}(w')\theta_w^1(dw') \\
&= R^1(w)/(1-\beta)\,,
\end{aligned}$$

where the first inequality holds by the remark made right before, and the second inequality holds by Lemma A8 and (20). $\square$

APPENDIX

This appendix provides some mathematics for the theory of Choquet capacity, which we rely upon in the text. We omit the proof whenever it is easily available somewhere in the literature (see, for example, Dellacherie (1970), Shapley (1971) and Schmeidler (1972, 1986) among others).

Probability Capacity and Probability Charge Let $(S, \mathscr{F})$ be a measurable space, where $\mathscr{F}$ is a σ-algebra on S. A *probability capacity* on $(S, \mathscr{F})$ is a function $\theta : \mathscr{F} \to [0,1]$ which satisfies

$$\theta(\phi) = 0$$

$$\theta(S) = 1$$

$$\text{and} \quad (\forall A, B \in \mathscr{F}) \;\; A \subseteq B \;\Rightarrow\; \theta(A) \leq \theta(B)\,.$$

A probability capacity is *convex* if

$$(\forall A, B \in \mathscr{F}) \quad \theta(A \cup B) + \theta(A \cap B) \geq \; \theta(A) + \theta(B) \tag{21}$$

while it is *concave* if the inequality in (21) is reversed. A probability capacity is a *probability charge* if the inequality in (21) holds with an equality.

A capacity θ is *continuous from below* if

$$(\forall \langle A_i \rangle_i \subseteq \mathscr{F}) \quad A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots \;\Rightarrow\; \theta(\cup_i A_i) = \lim_{i \to \infty} \theta(A_i)\,.$$

A capacity θ is *continuous from above* if

$$(\forall \langle A_i \rangle_i \subseteq \mathscr{F}) \quad A_1 \supseteq A_2 \supseteq A_3 \supseteq \cdots \;\Rightarrow\; \theta(\cap_i A_i) = \lim_{i \to \infty} \theta(A_i)\,.$$

A capacity θ is *continuous* if it is continuous both from below and from above. Note that any finite *measure* is continuous, and that continuity and finite additivity (that is, (21) with the inequality replaced by the equality) together imply countable additivity.

The *conjugate* of a probability capacity θ is the function $\theta^{0} : \mathscr{F} \to [0,1]$ defined by

$$(\forall A \in \mathscr{F}) \quad \theta^{0}(A) = 1 - \theta(A^c)$$

where A^c denotes the complement of A in S. The *core* of a probability capacity θ, $\mathrm{core}(\theta)$, is defined by

$$\mathrm{core}(\theta) = \{\, P \in \mathcal{M} \,|\, (\forall A \in \mathscr{F})\; P(A) \geq \theta(A) \,\}\,,$$

where $\mathcal{M}$ is the set of all probability charges on $(S, \mathscr{F})$. Note that a probability charge in the core of a continuous capacity is countably additive and hence a probability measure.

Lemma A 1. *Given a probability measure P on $(S, \mathscr{F})$ and a weakly increasing function f : $[0,1] \to [0,1]$ such that $f(0) = 0$ and $f(1) = 1$, define the function $f \circ P : \mathscr{F} \to [0,1]$ by*

$$(\forall A \in \mathscr{F}) \quad f \circ P(A) = f\,(P(A))\,.$$

Then $f \circ P$ is a probability capacity. Furthermore, $f \circ P$ is concave (resp. convex, continuous) when f is a concave (resp. convex, continuous) function.

Lemma A 2. *Suppose θ is a probability capacity. Then θ is concave (resp. convex) if and only if θ' is convex (resp. concave).*

Lemma A 3. *If θ is a convex probability capacity, then $core(\theta)$ is non-empty.*

Choquet Integral Let $L(S,\bar{\mathbb{R}})$ be the space of $\mathscr{F}$-measurable functions from S into $\bar{\mathbb{R}}$, and let $B(S,\mathbb{R})$ be the subspace of $L(S,\bar{\mathbb{R}})$ which consists of the bounded functions. Then, the *Choquet integral* of $u \in L(S,\bar{\mathbb{R}})$ with respect to a probability capacity θ is defined by

$$\int u\,d\theta \equiv \int_S u(s)\theta(ds) \equiv \int_{-\infty}^{0} (\theta(\{s|u(s) \geq x\}) - 1)dx + \int_0^{+\infty} \theta(\{s|u(s) \geq x\})dx$$

unless the expression is $-\infty + \infty$.

Two functions $u, v \in L(S,\bar{\mathbb{R}})$ are said to be *co-monotonic* if $(\forall s,t \in S)\ (u(s)-u(t))(v(s)-v(t)) \geq 0$.

Lemma A 4 (Monotonicity). *Let θ be a probability capacity. Then,*

$$(\forall u, v \in B(S,\mathbb{R}))\quad u \leq v \Rightarrow \int u\,d\theta \leq \int v\,d\theta\,.$$

Lemma A 5 (Positive Homogeneity). *Let θ be a probability capacity. Then*

$$(\forall u \in B(S,\mathbb{R}))(\forall a \in \mathbb{R})(\forall b \in \mathbb{R}_+)\quad \int (a+bu)d\theta = a + b\int u\,d\theta$$

where a in the left-hand side is understood to be a constant function.

Lemma A 6. *Let θ be a probability capacity. Then*

$$(\forall u \in B(S,\mathbb{R}))\quad \int u\,d\theta' = -\int -u\,d\theta\,.$$

Lemma A 7 (Co-monotonic Additivity). *Let θ be a probability capacity. If $u, v \in B(S,\mathbb{R})$ are co-monotonic, then*

$$\int (u+v)d\theta = \int u\,d\theta + \int v\,d\theta\,.$$

Lemma A 8. *If θ is a convex probability capacity, then*

$$(\forall u \in B(S,\mathbb{R})) \qquad \int u\, d\theta = \min\left\{ \int u\, dP \,\middle|\, P \in core(\theta) \right\}.$$

Lemma A 9. *If θ is a convex (resp. concave) probability capacity, then*

$$(\forall u, v \in B(S,\mathbb{R})) \qquad \int (u+v)\, d\theta \;\geq (resp. \leq) \int u\, d\theta + \int v\, d\theta\,.$$

Lemma A 10. *If θ is a convex probability capacity, then*

$$(\forall u \in B(S,\mathbb{R})) \qquad \int u\, d\theta \;\leq \int u\, d\theta'.$$

Lemma A 11. *If θ is a convex probability capacity, then*

$$(\forall u \in B(S,\mathbb{R})) \qquad \left| \int u\, d\theta - \int v\, d\theta \right| \leq \int |u-v|\, d\theta'$$

Proof. This holds because

$$\begin{aligned}
& \int u\, d\theta - \int v\, d\theta \\
= & \int u\, d\theta - \int (v-u+u)\, d\theta \\
\leq & \int u\, d\theta - \left[\int (v-u)\, d\theta + \int u\, d\theta \right] \\
= & -\int (v-u)\, d\theta \\
= & \int (u-v)\, d\theta' \\
\leq & \int |u-v|\, d\theta'
\end{aligned}$$

where the first inequality holds by Lemma A9, the third equality holds by Lemma A6, and the last inequality holds by Lemma A4. □

Lemma A 12 (Monotone Convergence Theorem). *(a) Let θ be a probability capacity which is continuous from below and let $\langle u_n \rangle_{n=0}^{\infty}$ be a sequence of $\mathscr{F}$-measurable functions such that $u_0 \leq u_1 \leq u_2 \leq u_3 \leq \cdots$ and $\int u_0\, d\theta > -\infty$. Then,*

$$\lim_{n\to\infty} \int u_n\, d\theta = \int \lim_{n\to\infty} u_n\, d\theta\,.$$

(b) Let θ be a probability capacity which is continuous from above and let $\langle u_n\rangle_{n=0}^{\infty}$ be a sequence of $\mathscr{F}$-measurable functions such that $u_0 \geq u_1 \geq u_2 \geq u_3 \geq \cdots$ and $\int u_0\, d\theta < +\infty$. Then,

$$\lim_{n\to\infty} \int u_n\, d\theta = \int \lim_{n\to\infty} u_n\, d\theta\,.$$

Proof. (b) follows from (a) if we let $\langle u_n\rangle$ and θ be $\langle -u_n\rangle$ and θ' in (a). We thus prove only (a). We first note that it holds that

$$\theta(\{u \geq x\}) = \theta(\{u > x\})$$

for almost all x. Obviously, $\theta(\{u \geq x\}) \geq \theta(\{u > x\})$ holds for all x. On the other hand, $x \mapsto \theta(\{u > x\})$ is weakly decreasing in x, and hence, continuous in x except on at most countably many points. Let x be a point at which $x \mapsto \theta(\{u > x\})$ is continuous. Then,

$$\theta(\{u > x\}) = \lim_{n\to\infty} \theta(\{u > x - 1/n\}) \geq \theta(\{u \geq x\})\,.$$

Therefore, (a) holds because

$$\begin{aligned}
&\lim_{n\to\infty} \int u_n d\theta \\
= &\lim_{n\to\infty} \left[\int_{-\infty}^{0} (\theta(\{u_n \geq x\}) - 1)dx + \int_{0}^{\infty} \theta(\{u_n \geq x\})dx\right] \\
= &\lim_{n\to\infty} \left[\int_{-\infty}^{0} (\theta(\{u_n > x\}) - 1)dx + \int_{0}^{\infty} \theta(\{u_n > x\})dx\right] \\
= &\int_{-\infty}^{0} \lim_{n\to\infty} (\theta(\{u_n > x\}) - 1)dx + \int_{0}^{\infty} \lim_{n\to\infty} \theta(\{u_n > x\})dx \\
= &\int_{-\infty}^{0} (\theta(\cup_n\{u_n > x\}) - 1)dx + \int_{0}^{\infty} \theta(\cup_n\{u_n > x\})dx \\
= &\int_{-\infty}^{0} (\theta(\{\lim_{n\to\infty} u_n > x\}) - 1)dx + \int_{0}^{\infty} \theta(\{\lim_{n\to\infty} u_n > x\})dx \\
= &\int_{-\infty}^{0} (\theta(\{\lim_{n\to\infty} u_n \geq x\}) - 1)dx + \int_{0}^{\infty} \theta(\{\lim_{n\to\infty} u_n \geq x\})dx \\
= &\int \lim_{n\to\infty} u_n d\theta
\end{aligned}$$

where the first and last equalities hold by the definition of the Choquet integral; the second and sixth hold by the remark we have just made; the third holds by the Lebesgue monotone convergence theorem; the fourth holds by θ's continuity from below; and the fifth holds by the strict inequality defining the set. $\square$

Note that by the monotone convergence theorem (Lemma A12), all of the above lemmas concerning the Choquet integral hold true for any continuous capacity θ and for any function $u \in L(S, \bar{\mathbb{R}})$ whenever the integral is well-defined.

Capacitary Kernel and Fubini Property A mapping $\theta : S \times \mathscr{F} \to [0,1]$ is a *capacitary kernel* (from S to S) if it satisfies

$$(\forall s \in S) \quad \theta_s \text{ is a probability capacity on } (S, \mathscr{F}) \quad \text{and}$$
$$(\forall B \in \mathscr{F}) \quad \theta_{(\cdot)}(B) \text{ is } \mathscr{F}\text{-measurable.}$$

In particular, if θ_s is a probability measure for all s, θ is called *stochastic kernel.*

Lemma A 13 (Fubini Property). *Let θ be a capacitary kernel such that $(\forall s)\ \theta_s$ is continuous. Then for any $(\mathscr{F} \otimes \mathscr{F})$-measurable function u, the mapping*

$$s \mapsto \int u(s, s_+)\, \theta_s(ds_+)$$

is $\mathscr{F}$-measurable.

Proof. Given $E \in \mathscr{F} \otimes \mathscr{F}$ and $s \in S$, we denote by $E(s)$ the s-section of E: $E(s) = \{ s' \in S \,|\, (s, s') \in E \}$. We first prove that the mapping $s \mapsto \theta_s(E(s))$ is $\mathscr{F}$-measurable for any $E \in \mathscr{F} \otimes \mathscr{F}$. Define $\mathscr{E}$ by

$$\mathscr{E} = \{ E \in \mathscr{F} \otimes \mathscr{F} \,|\, s \mapsto \theta_s(E(s)) \text{ is } \mathscr{F}\text{-measurable } \}.$$

Then the collection of finite disjoint unions of rectangles is a subfamily of $\mathscr{E}$ because, if $E = \cup_{i=1}^n (A_i \times B_i)$ where $A_i, B_i \in \mathscr{F}$ and $(A_i \times B_i) \cap (A_j \times B_j) = \phi$ for $i \neq j$, then

$$\theta_s(E(s)) = \max_{N' \subseteq \{1,2,\ldots,n\}} \theta_s(\cup_{i \in N'} B_i) I_{\cap_{i \in N'} A_i}(s)$$

and the right-hand side is $\mathscr{F}$-measurable. It remains to show that $\mathscr{E}$ is a monotone class. To this end, let $\langle E_n \rangle_{n=1}^\infty \subseteq \mathscr{E}$ and $E_n \uparrow E$. Then $E_n(s) \uparrow E(s)$ for any $s \in S$ and $\lim_{n \to \infty} \theta_s(E_n(s)) = \theta_s(E(s))$ by the continuity of $\theta_s(\cdot)$, which implies $E \in \mathscr{E}$. The similar argument applies to decreasing sequences. We now prove the theorem for the simple functions which is sufficient thanks to the monotone convergence theorem (Lemma A12). Let u be a simple function on $S \times S$. Then, we can write $u(s,s_+) = \sum_{i=1}^{n} a_i \chi_{E_i}(s,s_+)$, where $0 < a_1 < \cdots < a_n$, χ is the indicator function, and $\langle E_i \rangle$ is a partition of $S \times S$. It follows that

$$
\begin{aligned}
(\forall s \in S) \qquad \int u(s,s_+)\,\theta_s(ds_+) &= \int \sum_{i=1}^{n} a_i \chi_{E_i}(s,s_+)\,\theta_s(ds_+) \\
&= \int \sum_{i=1}^{n} a_i \chi_{E_i(s)}(s_+)\,\theta_s(ds_+) \\
&= \sum_{i=1}^{n} (a_i - a_{i-1})\,\theta_s\big(\cup_{k=i}^{n} (E_k(s))\big) \\
&= \sum_{i=1}^{n} (a_i - a_{i-1})\,\theta_s\big((\cup_{k=i}^{n} E_k)(s)\big),
\end{aligned}
$$

where $a_0 \equiv 0$ and the third equality holds by the definition of Choquet integral. Then the claim follows since the last expression is $\mathscr{F}$-measurable by the first paragraph. $\square$

REFERENCES

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

Chung, Kai Lai (1974): *A Course in Probability Theory* (Second Edition), Academic Press.

DeGroot, Morris H. (1970): *Optimal Statistical Decisions*, McGraw-Hill.

Dellacherie, C. (1970): *Quelques commentaires sur les prolongements de capacités*, Seminaire Probabilités V, Strasbourg, Lecture Notes in Math., vol. 191, Springer-Verlag, Berlin and New York.

Dellacherie, C. and P. A. Meyer (1988): *Probabilities and Potential C*, North-Holland, New York.

Ellsberg, D. (1961): “Risk, Ambiguity, and the Savage Axioms,” *Quarterly Journal of Economics* 75, 643-669.

Epstein, Larry G. and Tan Wang (1994): “Intertemporal Asset Pricing under Knightian Uncertainty,” *Econometrica* 62, 283-322.

Epstein, Larry G. and Tan Wang (1995): “Uncertainty, Risk-Neutral Measures and Security Price Booms and Crashes,” *Journal of Economic Theory* 67, 40-82.

Epstein, Larry G. and Jiankang Zhang (1999): “Least Convex Capacities,” *Economic Theory* 13, 263-286.

Ghirardato, Paolo (1997): “On Independence for Non-Additive Measures, with a Fubini Theorem,” *Journal of Economic Theory* 73, 261-291.

Gilboa, Itzhak (1987): “Expected Utility with Purely Subjective Non-Additive Probabilities,” *Journal of Mathematical Economics* 16, 65-88.

Gilboa, Itzhak and David Schmeidler (1989): “Maxmin Expected Utility with Non-Unique Prior,” *Journal of Mathematical Economics* 18, 141-153.

Gilboa, Itzhak and David Schmeidler (1993): “Updating Ambiguous Beliefs,” *Journal of Economic Theory* 59, 33-49.

Huber, Peter J. and Volker Strassen (1973): “Minimax Tests and the Neyman-Pearson Lemma for Capacities,” *The Annals of Statistics* 1, 251-263.

Keynes, John Maynard (1921): *A Treatise on Probability*, London: Macmillan.

Keynes, John Maynard (1936): *The General Theory of Employment, Interest and Money*, London: Macmillan.

Knight, Frank (1921): *Risk, Uncertainty and Profit*, Boston: Houghton Mifflin.

Lippman, Steven A. and John J. McCall (1976): "The Economics of Job Search: A Survey," *Economic Inquiry* 14, 347-368.

Nishimura, Kiyohiko G. and Hiroyuki Ozaki (2001): "A Note on Learning under Knightian Uncertainty," the University of Tokyo, mimeo.

Ozaki, Hiroyuki and Peter A. Streufert (1996): "Dynamic Programming for Non-additive Stochastic Objectives," *Journal of Mathematical Economics* 25, 391-442.

Rothschild, Michael (1974): "Searching for the Lowest Price When the Distribution of Prices is Unknown," *Journal of Political Economy* 82, 689-711.

Rothschild, Michael and Joseph Stiglitz (1970): "Increasing Risk I: A Definition," *Journal of Economic Theory* 2, 225-243.

Rothschild, Michael and Joseph Stiglitz (1971): "Increasing Risk II: Its Economic Consequences," *Journal of Economic Theory* 3, 66-84.

Sargent, Thomas J. (1987) *Dynamic Macroeconomic Theory*, Cambridge: Harvard University Press.

Savage, L. (1954): *The Foundations of Statistics.* John Wiley, New York.

Schmeidler, David (1972): "Cores of Exact Games I," *Journal of Mathematical Analysis and Applications* 40, 214-225.

Schmeidler, David (1986): "Integral Representation without Additivity," *Proceedings of the American Mathematical Society* 97, 255-261.

Schmeidler, David (1989): "Subjective Probability and Expected Utility without Additivity," *Econometrica* 57, 571-587.

Shapley, Lloyd S. (1971): "Cores of Convex Games," *International Journal of Game Theory* 1, 11-26.

Wasserman, Larry A. and Joseph B. Kadane (1990): "Bayes' Theorem for Choquet Capacities," *The Annals of Statistics* 18, 1328-1339.

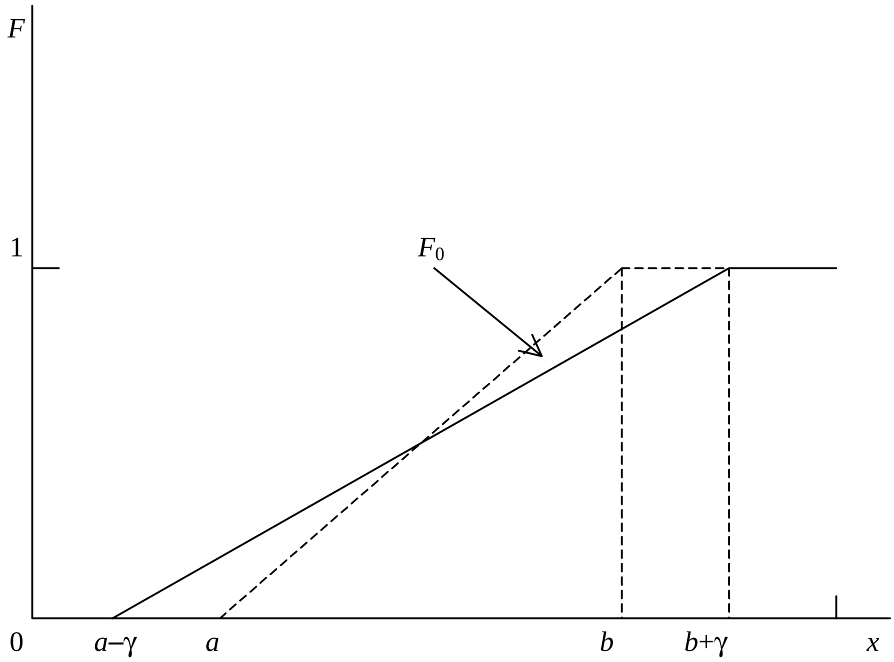

FIGURE 1 : Mean–Preserving Spread

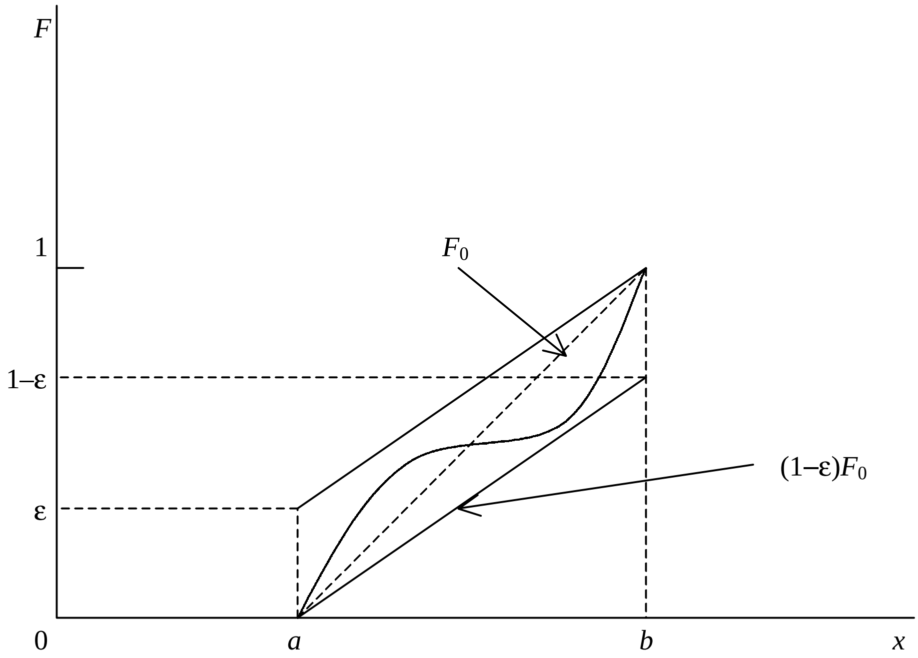

FIGURE 2 : ε–contamination

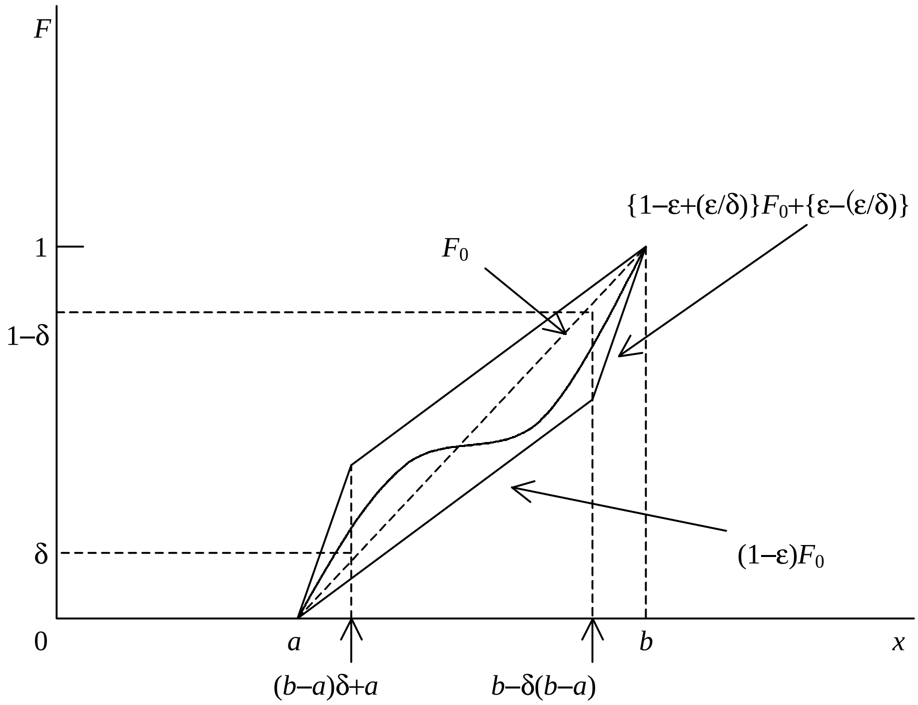

FIGURE 3 : δ–approximation of ε–contamination