The effect of learning on membership and welfare in an International Environmental Agreement

Larry Karp
University of California, Berkeley
Department of Agricultural and Resource Economics
karp@are.berkeley.edu

July 29, 2009

Abstract

Better information about the cost-benefit of abatement has an ambiguous effect on both the equilibrium membership and on aggregate welfare of an international environmental agreement. Previous papers claim that (complete) learning increases membership and decreases aggregate welfare. That claim is based on analysis of approximations to the relations between a damage parameter and membership and welfare. Those approximations have characteristics not shared by the functions they are intended to approximate. Conclusions based on the approximations are wrong. The correct result is that better information increases membership and welfare when the damage parameter is "very likely to be high", and the reverse holds when the damage parameter is "very likely to be low".

1 Introduction

A recent paper incorrectly claims that better information increases the equilibrium membership and reduces aggregate welfare in an International Environmental Agreement (IEA): Kolstad and Ulph (2008a), which builds on Ulph (2004), Kolstad (2007) and Kolstad and Ulph (2008b).[1] This note explains and corrects the error, and it is of more general interest because it shows how to analyze the effect of better information on the formation of IEAs, using a standard model.

The above papers use a two-stage game consisting of a participation stage, when nations decide whether to join the IEA, and an abatement stage, when all nations decide whether to reduce emissions. They compare the equilibrium outcome when a damage parameter is uncertain at both stages (no learning), and the outcome when nature reveals the value of this parameter before either the participation or abatement decisions (complete learning).[2]

Ulph and Kolstad show that approximations of the functions relating the damage parameter to the equilibrium number of IEA members and to welfare are, respectively, convex and concave. They then apply Jensen's Inequality to establish their claim. Although the approximations have the curvature that the authors assert, the actual relations are neither convex nor concave. Therefore, Jensen's inequality is inapplicable: better information might increase equilibrium IEA size and decrease welfare, as claimed, or have the opposite effect.

The actual effect of learning on membership and welfare depends on the distribution of the damage parameter. I show that if the probability of low damages is small, then learning increases both expected membership and expected welfare; learning decreases both membership and welfare if the probability of low damages is large. For intermediate probabilities of low damages, learning can have opposite effects on membership and welfare.

[1] These claims appear several places in Kolstad and Ulph (2008a). Their Corollary 1 offers the most precise statement: "Uncertainty with complete learning leads to higher expected membership of IEA but lower expected aggregate world benefits than in the case of uncertainty with no learning."

The abstract to their paper highlights the welfare effect of learning: "Our results are generally pessimistic: the possibility of either complete or partial learning generally reduces the level of global welfare that can be achieved from forming and IEA." On page 130 they write "The general conclusion is that the possibility of learning reduces welfare, so the value of information is negative." On page 139 they write "We can summarize the results as follows: Uncertainty with Complete Learning leads to higher expected membership but lower expected aggregate net benefits than No Learning...".

[2] The papers also consider a more complicated third information structure, "partial learning", in which nature reveals the damage parameter after the participation stage and before the abatement stage. I do not discuss that information structure.

2 Preliminaries: known damage parameter

In the standard model there are N countries, each of which has a binary decision, to emit one unit of pollution or to abate. Emission creates one unit of private benefit for the country, e.g. because higher pollution increases economic output; each country suffers γ units of damage for each unit of pollution emitted by any country, because the pollutant is global. The assumption $\gamma < 1$ means that it is a dominant strategy for a country acting alone to emit. The assumption $\gamma N > 1$ means that all countries are better off if all abate, compared to when all emit. The IEA instructs all members to abate only if there are enough members for the combined benefit of abatement to exceed the combined private costs.

If m countries abate and $N - m$ countries emit, the payoffs of an abater and an emitter are, respectively, $-\gamma(N-m)$, and $1 - \gamma(N-m)$. A nation that abates looses the 1 unit of private benefit associated with its pollution. If m of the N countries join an agreement and all members abate, each member obtains the payoff $-\gamma(N-m)$; if all members pollute, each member obtains the payoff $1 - \gamma N$. Therefore, the IEA that maximizes the joint welfare of its members (as distinct from global, or aggregate welfare) instructs members to abate if and only if

$$-\gamma(N-m) \geq 1 - \gamma N,$$

i.e. if and only if $m \geq \frac{1}{\gamma}$.

There are trivial Nash equilibria to the participation game, those with zero members or with a positive number of members who then decide not to abate. At the unique nontrivial Nash equilibrium to the participation came there are two or more IEA members and these decide to abate. This equilibrium consists of

$$m^* = m(\gamma) \equiv h\left(\frac{1}{\gamma}\right) \tag{1}$$

members, where $h(x)$, the ceiling function, equals the smallest integer not less than x.[3] When there are m abaters, the aggregate benefit of both polluters and abaters is $N - m - \gamma N(N-m) = (1-\gamma N)(N-m)$. Substituting m^* in this expression gives global equilibrium welfare:

$$W(\gamma) = (1-\gamma N)\left(N - h\left(\frac{1}{\gamma}\right)\right). \tag{2}$$

[3]If $m > m^*$ a signatory obtains a higher payoff by leaving the IEA, so $m > m^*$ is not a Nash equilibrium. If $m = m^*$ no signatory wants to defect, because doing so causes the IEA to decide not to abate, leaving the defector worse off. Moreover, no nonsignatory wants to join the IEA with $m = m^*$ members. Therefore $m = m^*$ is a Nash equilibrium. Any level of $m < m^*$ is a trivial equilibrium, because at this level, no abatement occurs in the next stage.

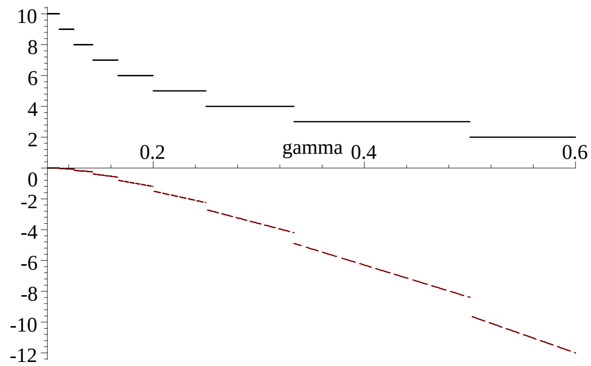

Figure 1: Equlibrium membership (the solid step function in the top part of the figure) and welfare (the dashed graph in the lower part of the figure) as functions of γ

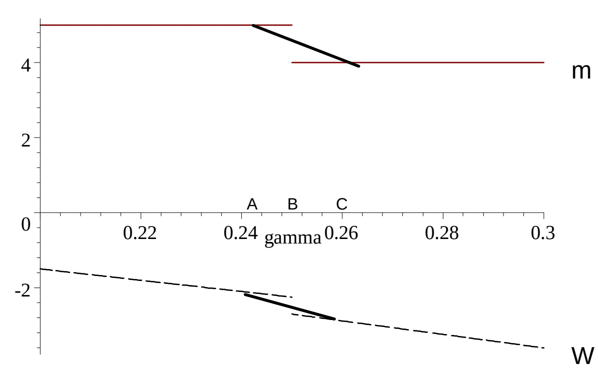

Figure 2: Enlargement of Figure 1: Illustration of non-cacavity and non-convexity and basis for intuition

Figure 1 graphs membership $m(\gamma)$, the solid step function, and scaled welfare $\frac{3}{N}W(\gamma)$, the dashed negatively sloped discontinuous curve, for $\gamma \in [0.1, 0.6]$ and $N = 10$.[4] The functions $m(\gamma)$ and $W(\gamma)$ are neither convex nor concave, although the former might appear "approximately convex" and the latter might appear "approximately concave". To demonstrate this non-convexity and non-concavity and to provide intuition for the propositions below, Figure 2 provides a close-up of Figure 1 and adds to each graph a (heavy, solid) line segment between points A and C. These two points lie on either side of point B, the value $\gamma = 0.25$, a point of discontinuity of the graphs.

I draw attention to the values of $m(\gamma)$ and $W(\gamma)$ for $\gamma \in [A, C]$, and especially for $\gamma \in \{A, B, C\}$. For $\gamma \in (A, B)$, the graph of $m(\gamma)$ and of $W(\gamma)$ lie above the heavy solid line segments between (respectively) $m(A)$ and $m(C)$, and between $W(A)$ and $W(C)$. The ordering is reversed for $\gamma \in (B, C)$. Therefore, neither graph is concave and neither is convex. I return to Figure 2 after stating the propositions.

3 Uncertainty about γ and the effect of learning

To introduce uncertainty, I use the assumption common to many papers, that the damage parameter has a two-point distribution: γ equals γ_L with probability p and γ_H with probability $1 - p$, with $\gamma_L < \gamma_H$ If countries learn the true value before they make their participation decision ("complete learning"), the ex ante (before learning) expected equilibrium membership and expected aggregate welfare are

$$\begin{aligned} m^{\text{learn}} &= ph\left(\frac{1}{\gamma_L}\right) + (1-p)\,h\left(\frac{1}{\gamma_H}\right) \\ W^{\text{learn}} &= pW(\gamma_L) + (1-p)\,W(\gamma_H)\,. \end{aligned}$$

[4] Over this range, the assumptions $\gamma < 1$ and $\gamma N \geq 1$ are satisfied. I use the scaling factor $\frac{3}{N}$ in the graph of W so that the two graphs are on approximately the same scale. In the text I drop further reference to this scaling factor, since it plays no role except to improve the visuals.

If they make their participation and abatement decisions before learning the value of γ, the equilibrium membership and expected welfare are[5]

$$\begin{aligned} m^{\text{no learn}} &= h\left(\frac{1}{p\gamma_L + (1-p)\gamma_H}\right) \\ W^{\text{no learn}} &= W\left(p\gamma_L + (1-p)\gamma_H\right). \end{aligned}$$

The reduction in the expected number of members due to learning, denoted δ, and the reduction in aggregate expected welfare due to learning, denoted Δ, are

$$\begin{aligned} \delta &= m^{\text{no learn}} - m^{\text{learn}} \\ \Delta &= W^{\text{no learn}} - W^{\text{learn}}. \end{aligned}$$

In order to reduce the number of special cases, I adopt:

Assumption 1 *Neither* $\frac{1}{\gamma_L}$ *nor* $\frac{1}{\gamma_H}$ *is an integer.*

This assumption excludes a set of measure 0 from parameter space. Therefore, the results described in the following propositions are generic. The papers cited above claim that $\Delta > 0 > \delta$, i.e., learning increases membership and reduces welfare. Proposition 1 shows that the claim $0 > \delta$ is true for small p but false for large p. Proposition 2 shows that the claim $\Delta > 0$ is false for small p but true for large p. For completeness, parts (i) of both propositions consider the case where learning has no effect on the outcome. Parts (ii) of the propositions contain the significant results.

Proposition 1 *Adopt Assumption 1. (i) If* $h\left(\frac{1}{\gamma_H}\right) = h\left(\frac{1}{\gamma_L}\right)$ *then learning has no effect on equilibrium expected membership. (ii) If* $h\left(\frac{1}{\gamma_H}\right) < h\left(\frac{1}{\gamma_L}\right)$ *then learning increases equilibrium expected membership (*$\delta(p) < 0$*) for small positive* p *and decreases equilibrium expected membership (*$\delta(p) > 0$*) for large* p*.*

[5]The ability to pass the expectations operator through the functions h and W depends on the fact that the model is linear in γ. At the abatement stage (given that γ is uncertain) the IEA's decision depends on the number of members, which was determined at the earlier stage, and on the expectation of γ, $p\gamma_L + (1-p)\gamma_H$.

At the participation stage, each nation takes other nations' participation decisions as given, it takes as given the IEA's decision rule used in the next stage, and it and maximizes its own expected payoff. Because of the linearity of the model, its decision depends on the expectation of γ. As a consequence, the equilibrium number of members, $m^{\text{no learn}}$, also depends only on this expectation.

Aggregate welfare depends *directly* on γ, but since welfare is linear in γ, expected welfare depends only on the expectation of γ. Aggregate welfare also depends on the number of members and on the IEA's decision rule, but these depend only on the expectation of γ, not on its realization.

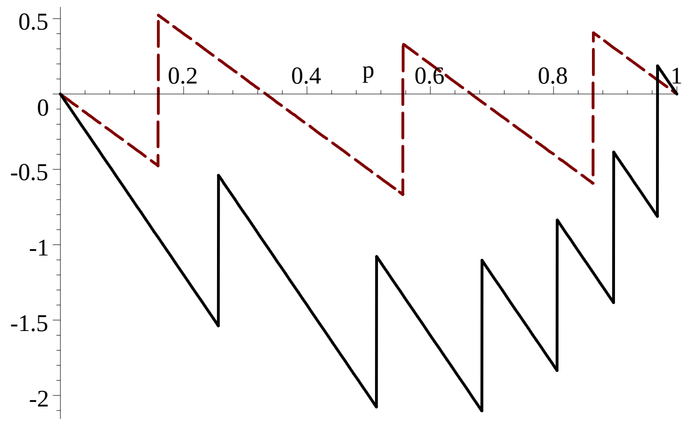

Figure 3: The graphs of δ as a function of p, for "large cost dispersion", $\gamma_L = 0.105$ and $\gamma_H = 0.3$ (the solid graph), and for "small cost dispersion", $\gamma_L = 0.105$ and $\gamma_H = 0.15$ (the dashed graph). For small p learning increases membership, and for large p learning reduces membership in both cases. With small dispersion, the effect of learning is ambiguous at intermediate values of p..

Proposition 2 *Adopt Assumption 1. (i) If* $h\left(\frac{1}{\gamma_H}\right) = h\left(\frac{1}{\gamma_L}\right)$ *then learning has no effect on equilibrium expected welfare. (ii) If* $h\left(\frac{1}{\gamma_H}\right) < h\left(\frac{1}{\gamma_L}\right)$ *then learning increases equilibrium expected welfare (*$\Delta(p) < 0$*) for small positive* p *and decreases equilibrium expected welfare (*$\Delta(p) > 0$*) for large* p*.*

In the interest of completeness, the Appendix discusses in more detail the relation between my results and those of earlier papers. The Appendix also contains proofs of Propositions 1 and 2, but the intuition is readily apparent from Figure 2. Let $\gamma_L = A$ and $\gamma_H = C$. These two values lie on either side of a point of discontinuity, as in Figure 2, if and only if learning has an effect on the equilibrium expected values of membership and welfare. For small p, the expectation of γ is close to point C, where the convex combination of $m(A)$ and $m(C)$, and of $W(A)$ and $W(C)$, lie above the graphs of $m(\gamma)$ and of $W(\gamma)$, respectively. Therefore, for small p, $Em(\gamma) > m(E\gamma)$ and $EW(\gamma) > W(E\gamma)$, where "E" denotes expectation. Large p reverses this ranking.

Figure 3 illustrates Proposition 1 ii; it shows the graphs of δ as a function of p, for two pairs of damage parameters. For $\gamma_L = 0.105$ and $\gamma_H = 0.3$ (the solid graph) the cost dispersion is "large", and for $\gamma_L = 0.105$ and $\gamma_H = 0.15$ (the dashed graph) the cost dispersion is "small".[6] Larger cost dispersion (e.g. a higher variance of the damage parameter, γ) means that there is

[6]The values of m given by equation (1), corresponding to $\gamma \in \{0.105, 0.15, 0.3\}$ are $\{10, 7, 4\}$. I show the

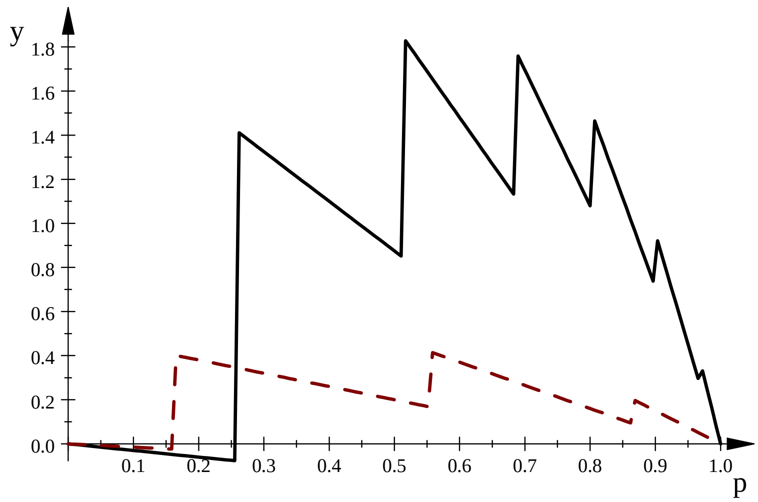

Figure 4: The graphs of Δ as a function of p for. $\theta_L = 0.105$ and $\theta_H = 0.3$ (the solid graph), and $\theta_L = 0.105$ and $\theta_H = 0.15$(the dashed graph). For small p learing increases welfare and for large p learning reduces welfare.

"more uncertainty".

For both sets of damage parameters, better information increases membership for small p ($\delta(p) < 0$)and decreases membership for large p ($\delta(p) > 0$), as Proposition 1 ii asserts. The figure also shows that for both sets of parameter values, the difference in equilibrium membership is a non-monotonic function of p. Thus, a greater probability of low damages might either increase or decrease the difference in equilibrium size, in the two scenarios with learning and without learning.

The more interesting feature is that $\delta(p)$ changes signs only once when cost dispersion is large ($\gamma_L = 0.105$ and $\gamma_H = 0.3$), but it changes signs five times when cost dispersion is small. Under the scenario with large cost dispersion, better information increases the equilibrium expected size of the IEA ($\delta < 0$) if and only if $p < 0.97$, i.e. for most of parameter space In contrast, when the cost dispersion is small ($\gamma_L = 0.105$ and $\gamma_H = 0.15$), there are three disjoint intervals of p over which learning increases the equilibrium expected size of the IEA, and thee disjoint intervals over which learning decreases the equilibrium expected size of the IEA. This reswitching between a positive and negative sign illustrates the difficulty of strengthening Proposition 1: although we know the effect of learning for large and for small p, we can say nothing in general about the effect for intermediate values of p.

Figure 4 illustrates Proposition 2 ii. It shows the graphs of Δ for the same two sets of damage parameters as above. For small values of p learning increases welfare, and for large

graphs in Figures 3 - 5 as continuous, for ease of viewing. In fact, the graphs are discontinuous where they are vertical.

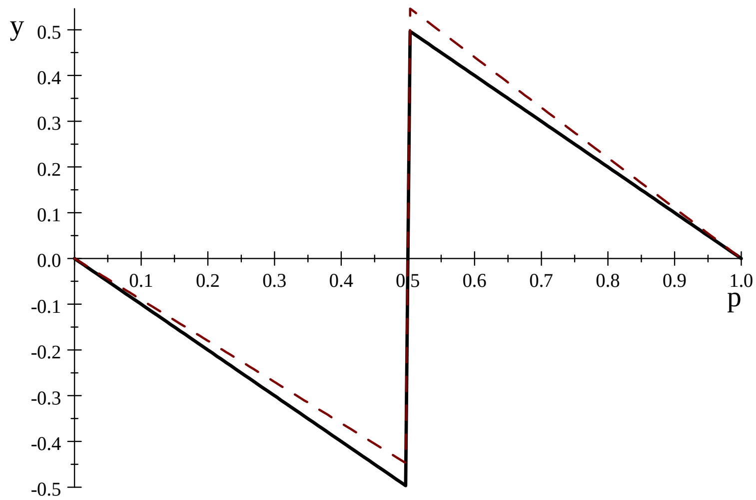

Figure 5: The graph of δ (solid) and Δ (dashed) for $\gamma_L = 0.19$ and $\gamma_H = 0.21$.

values of p learning lowers welfare.

Figure 4 might give the impression that learning lowers welfare $(\Delta > 0)$ for "most" values of p, and in that respect one of the claims in the earlier papers might appear to be "approximately right". To correct that possible impression, Figure 5 shows the graphs of δ and Δ together, for $\gamma_L = 0.19$ and $\gamma_h = 0.21$. The point at which the graphs become vertical – where the actual graphs are discontinuous – occurs at $p = 0.5$. In this example, information increases membership and welfare for $p < 0.5$ and decreases membership and welfare for $p > 0.5$.

The graphs of δ and Δ have the same points of discontinuity. At a point of discontinuity, the graphs jump up (as p increases). Over intervals of continuity, both graphs are decreasing in p. However, the rates of decrease (the slopes of the graphs) and the sizes of the jumps are different for the two graphs. Figure 5 illustrates a situation where the two graphs have the same sign, so that information increases expected IEA size if and only if it also increases aggregate expected welfare. However, comparison of Figures 3 and 4 show that this relation is not general. For example, for $\gamma_L = 0.105$ and $\gamma_H = 0.15$, δ changes sign five times over the interval $(0, 1)$ whereas the graph of Δ changes sign only once. Thus, for these parameter values, there are intervals of p for which better information increases expected membership and decreases aggregate expected welfare.

4 Conclusion

Learning has an ambiguous effect on membership and welfare in an IEA game. The equilibrium must satisfy the integer constraint, the requirement that the number of countries that join the IEA be an integer. This constraint is central to the model, not an incidental feature of it. If nations were arbitrarily small (i.e. of measure 0) relative to the size of the world, then an individual nation's participation decision would have no effect on the equilibrium abatement decisions of IEA signatories. In that case, there would be no integer constraint. However, in that case, no nation would be "pivotal" and the only Nash equilibria that remain are trivial, i.e. they contain too few (possibly 0) members to induce abatement in the second stage of the game. In other words, the fact that nations have positive measure is responsible for the outcome that an equilibrium IEA has positive members. It is also responsible for the discontinuity of the graphs of equilibrium membership and welfare, as a function of a model parameter such as γ. This discontinuity (together with the slopes of the pieces) means that the graphs are neither concave nor convex. Therefore, the effect of learning has an ambiguous effect on both the equilibrium number of members and on equilibrium welfare.

Previous papers reached the wrong conclusion regarding the relation between (complete) learning and the equilibrium IEA size and aggregate welfare, because they relied on approximations that have properties not shared by the functions that are supposedly being approximated. The correct result, under the assumption of a two point distribution on the damage parameter, is that complete learning (when it has any effect) increases membership and welfare if the high damage outcome is "very likely"; learning decreases membership and welfare if the low damage outcome is "very likely". For intermediate probabilities, I showed by example that learning might either reduce or increase expected membership and welfare. In addition, learning can have different effects on expected membership and welfare, possibly increasing one and decreasing the other.

A Appendix: Related literature and proofs

Related literature Proposition 1 *does not* contradict Proposition 2 of Kolstad (2007), which states (translated into my notation) that $m^{\text{no learn}} - m^{\text{learn}} = \delta \leq 1$. (Indeed, my simulations are consistent with this inequality; see Figures 3 and 5.) In the proof of his proposition, Kolstad also notes that the membership function $m(\gamma)$ is not convex.[7] However, Kolstad and Ulph (2008a) do rely on the convexity of $m(\gamma)$ and they assert that $\delta < 0$; see the quotes in footnote 1 above.

Kolstad (2007) possibly misinterprets his result, which might explain the discrepancy between the two papers. He writes "...we find that learning tends to increase the size of the

[7]He writes that the function is not "locally" convex, but does not define this term. I doubt that a function can be convex without also being "globally convex".

cooperating coalition in an international environmental agreement." (page 78) This statement is weaker than the claim in Kolstad and Ulph (2008a), due to the inclusion of the qualifier "tends to". However, it is not clear what this qualifier means. Figure 5 provides an example where learning increases expected membership for $p < 0.5$ and decreases expected membership for $p > 0.5$, so it is hard to discern a "tendency" in one direction or the other.

Even though (according to Kolstad's Proposition 2) learning can decrease the level of *expected* membership by no more than 1 ($\delta \leq 1$), for some damage realizations membership without learning might be much greater than membership with learning. In the absence of learning, membership is not a random variable, and therefore it must be an integer. With learning, membership for each realization of random damages is also an integer. However, the *expected* level of membership is a convex combination of two integers and therefore in general takes non-integer values.

Proofs of Propositions **Proof.** (Proposition 1) From the definition of δ

$$\delta(0) = \delta(1) = 0. \tag{3}$$

If there is no uncertainty, there can be no learning.

(i) Under Assumption 1, if

$$h\left(\frac{1}{\gamma_H}\right) = h\left(\frac{1}{\gamma_L}\right) = x \tag{4}$$

for an integer x, then

$$\frac{1}{x-1} > \theta_H \geq \theta_L > \frac{1}{x}.$$

Define the expected value of damages as $\bar{\gamma} = p\gamma_L + (1-p)\gamma_H$ for $0 \leq p \leq 1$, so in general

$$h\left(\frac{1}{\gamma_H}\right) \leq h\left(\frac{1}{\bar{\gamma}}\right) \leq h\left(\frac{1}{\gamma_L}\right). \tag{5}$$

Equations (4) and (5) imply that $h\left(\frac{1}{\bar{\gamma}}\right) = x$, so $h\left(\frac{1}{\bar{\gamma}}\right)$ is independent of p. In this case,

$$\frac{d\delta}{dp} = -\left(h\left(\frac{1}{\gamma_L}\right) - h\left(\frac{1}{\gamma_H}\right)\right) = 0. \tag{6}$$

Equations (3) and (6) imply that $\delta = 0$ for all p.

(ii) Assumption 1 means that

$$\frac{dh\left(\frac{1}{\bar{\gamma}}\right)}{dp} = 0 \text{ for } p \approx 0 \text{ or } p \approx 1,$$

so δ is differentiable in the neighborhood of both $p = 0$ and $p = 1$. Therefore, for $p \approx 0$ or $p \approx 1$,

$$\frac{d\delta}{dp} = -\left(h\left(\frac{1}{\gamma_L}\right) - h\left(\frac{1}{\gamma_H}\right)\right) < 0. \tag{7}$$

Equation (3) and (7) imply that $\delta < 0$ for $p = \varepsilon$ and $\delta > 0$ for $p = 1 - \varepsilon$, where ε is a small positive number. ■

Proof. (Proposition 2) (i) In this case, by Proposition 1 i, better information has no effect on membership. Holding membership fixed, welfare is linear in γ, by equation (2). Therefore uncertainty about γ does not affect expected welfare.

(ii) Learning has no effect on welfare when there is no uncertainty:

$$\Delta(0) = \Delta(1) = 0. \tag{8}$$

Using the definition of Δ and simplifying gives

$$\Delta(p) = -\left((1 - \bar{\gamma}N)\, h\left(\frac{1}{\bar{\gamma}}\right) - p\,(1 - \gamma_L N)\, h\left(\frac{1}{\gamma_L}\right) - (1-p)\,(1 - \gamma_H N)\, h\left(\frac{1}{\gamma_H}\right)\right).$$

Consider first the behavior of Δ in the neighborhood of $p = 0$. By virtue of Assumption 1, a small change in p in the neighborhood of $p = 0$ leaves $h\left(\frac{1}{\bar{\gamma}}\right)$ unchanged, equal to $h\left(\frac{1}{\gamma_L}\right)$. In addition,

$$\frac{d\bar{\gamma}}{dp} = \gamma_L - \gamma_H < 0 \tag{9}$$

for all p. These two facts imply that in the neighborhood of $p = 0$, the derivative of $\Delta(p)$ exists and

$$\frac{d\Delta}{dp} = (1 - \gamma_H N)\left(h\left(\frac{1}{\gamma_L}\right) - h\left(\frac{1}{\gamma_H}\right)\right) < 0. \tag{10}$$

Now consider the neighborhood of $p = 1$. In this neighborhood, a small change in p leaves $h\left(\frac{1}{\bar{\gamma}}\right)$ unchanged, equal to $h\left(\frac{1}{\gamma_H}\right)$. Therefore in this neighborhood

$$\frac{d\Delta}{dp} = (1 - \gamma_L N)\left(h\left(\frac{1}{\gamma_L}\right) - h\left(\frac{1}{\gamma_H}\right)\right) < 0. \tag{11}$$

In view of equations (8), (10) and (11), $\Delta < 0$ for small positive p, and $\Delta > 0$ for p close to 1.

■

References

KOLSTAD, C. D. (2007): "Systematic uncertainty in self-enforcing international environmental agreements," *Journal of Environmental Economics and Management*, 53, 68–79.

KOLSTAD, C. D., AND A. ULPH (2008a): “Learning and international environmental agreements,” *Climatic Change*, 89, 125–41.

——— (2008b): “Uncertainty, learning and heterogeneity in international environmental agreements,” unpublished working paper.

ULPH, A. (2004): “Stable international environmental agreements with a stock pollutant, uncertainty and learning,” *Journal of Risk and Uncertainty*, 29, 53–73.