Teoría de juegos
Juegos Repetidos

TEORÍA DE LAS DECISIONES

M. PAULA BONEL

¿Consultas clase anterior?

Motivación

- Las relaciones entre los agentes habitualmente duran más que un período (un juego).
- Por lo tanto, cuando existe el futuro, nos importan la reputación, los incentivos y los castigos.
  - Ejemplos: las relaciones de trabajo duran varios años, las relaciones entre los países implican muchas instancias de negociación, los competidores en una industria se encuentran repetidamente en el mercado.
- Los agentes toman las decisiones condicionados por la historia, por lo que ellos y los otros jugadores han hecho en el pasado. Las decisiones futuras van a estar influenciadas por lo que ocurra hoy.

Sticks and Carrots

La promesa de recompensas futuras (**zanahorias**) y la amenaza de castigos futuros (**palos**) puede proporcionar incentivos a comportarme de forma cooperativa hoy.

De juegos one-shot a repetidos…

- El juego se repite dos veces. El juego es simultáneo en cada período.
- Los jugadores maximizan la rentabilidad total (suma de pagos en ambos períodos).
- El resultado del primer período es observado públicamente antes de que se juegue la segunda vuelta.
- Por simplicidad los pagos a futuro no se descuentan (δ=1).
- ¿Hay posibilidad de cooperación?

	C	D
C	$5, 5$	$0, 6$
D	$6, 0$	$1, 1$

Repetición Dilema del Prisionero un número finito de veces

- Las acciones en la última etapa (t=1) son independientes de lo que sucede en la etapa 0.
- **Intuición:** El incentivo a cooperar se relaciona a la posibilidad de aumentar mis pagos en el futuro.
- En la última etapa del juego (t=1) no hay etapas futuras a tener en cuenta.
- Por lo tanto, esperamos que en t=1 se juegue (D,D) independientemente de lo que sucede en t=0**. Los pagos de la etapa 1 son (1,1).**

Repetición Dilema del Prisionero un número finito de veces

- Resolvemos ahora la etapa 0 por inducción hacia atrás, sumando 1 al pago de cada uno de los jugadores sobre el juego inicial:

	C	D
C	$6, 6$	$1, 7$
D	$7, 1$	$2, 2$

Repetición Dilema del Prisionero un número finito de veces

- **En el juego anterior encontramos que el ENPS =** $\{(DD, DD)\}$ **¿Qué sucede para un juego de un número de períodos arbitrario** n **?**
- La respuesta sigue siendo la misma. En el último período, n , independientemente de lo jugado en las rondas anteriores, existe un único equilibrio de Nash para los subjuegos resultantes donde cada jugador juega D.
- Las acciones del período $n - 1$ no tienen ningún efecto en lo que se jugará el día siguiente. Procediendo de esta manera hasta la fecha $t = 0$, encontramos que existe un único equilibrio perfecto en subjuegos: en cada período t y para cada historial de jugadas anteriores, cada jugador juega D.
- Aunque haya muchas repeticiones en el juego y las ganancias en el futuro pueden ser altas, cualquier plan de acción cooperativo se deshace porque los jugadores no pueden comprometerse con ningún plan de acción en la última ronda. ***Este es un resultado general para este tipo de juegos.***

¿Hay posibilidad de cooperación?

- El juego se Las reglas son iguales al caso anterior pero ahora en cada etapa se juega una matriz distinta:

	C	D
C	$5, 5$	$0, 6$
D	$6, 0$	$1, 1$

	S	R
S	4, 4	0, 2
R	2, 0	2, 2

¿Hay posibilidad de cooperación?

- El juego se Las reglas son iguales al caso anterior pero ahora en cada etapa se juega una matriz distinta:

	C	D
C	$5,5$	$0,6$
D	$6,0$	$1,1$

Etapa 1: Dilema del Prisionero

	S	R
S	4, 4	0, 2
R	2, 0	2, 2

Etapa 2: Caza del ciervo

¿Hay posibilidad de cooperación?

- **Observaciones clave:** ¿Cuántos EN hay en la caza del ciervo?¿Es creíble la amenaza de jugar conejo (R)?

  ¿Se puede utilizar esa amenaza para inducir la cooperación desde el principio?

	C	D
C	$5, 5$	$0, 6$
D	$6, 0$	$1, 1$

Etapa 1: Dilema del Prisionero

	S	R
S	4, 4	0, 2
R	2, 0	2, 2

Etapa 2: Caza del ciervo

¿Hay posibilidad de cooperación?

- **Proponemos la siguiente estrategia:** "Juega a Cooperar (C) en la Etapa 1. Si tu compañero también eligió Cooperar (C), juegue Ciervo (S: Stag) en la Etapa 2. Si tu compañero no eligió cooperar (D), juegue Conejo (R: Rabbit) en la segunda etapa". **¿Es sostenible la estrategia?**

	C	D
C	$5, 5$	$0, 6$
D	$6, 0$	$1, 1$

Etapa 1: Dilema del Prisionero

	S	R
S	4, 4	0, 2
R	2, 0	2, 2

Etapa 2: Caza del ciervo

¿Hay posibilidad de cooperación?

- **Los pagos de cooperar son: 5 + 4 = 9**
- **Los pagos del desvío son: 6 + 2 = 8**

	C	D
C	5+4, 5+4	0+2, 6+2
D	6+2, 0+2	1+2, 1+2

	C	D
C	9, 9	2, 8
D	8, 2	3, 3

En resumen…

- Existe la posibilidad de alcanzar cooperación en la primera etapa ya que uno de los EN es **(CS, CS).** Para alcanzarlo, la decisión de juego en el futuro debe ser variable (sujeta a lo que sucede en el pasado)

- La posibilidad de no cooperar aún es un equilibrio **EN = (DR, DR)**

- La estrategia cooperativa puede requerir jugar un EN "malo" en la Etapa 2: **(R, R).**
  - Problema: renegociación / riesgo moral / rescates

Juegos Repetidos Infinitamente
Conceptos importantes

- Hasta aquí aprendimos que las decisiones en la etapa final de juego fueron cruciales. **Si el juego es finito, en la última etapa siempre se juega un EN.**

- ¿Qué pasa si no hay una etapa final (o no lo sé)?

- Considere ahora la repetición **infinita** algunos juegos.

Impaciencia: El tiempo importa

- No es lo mismo ganar \$1 hoy que ganar \$1 mañana.
- Costo de oportunidad: un peso invertido hoy me genera 1+r mañana (tasa de retorno=.
- Los pagos de distintos períodos se deben expresar en valor presente para tomar la decisión correcta.

Factor de descuento: $\boldsymbol{\delta = \frac{1}{1+r}}$

- 0 <δ < 1
- δ será más alto cuánto menor sea la impaciencia, cuánto más importe el futuro. Si soy muy impaciente le voy a dar una valuación muy baja a mis pagos en los períodos futuros.

Pagos: Valor presente

- En un juego repetido infinitamente, uno no puede simplemente sumar los pagos de cada etapa ya que la suma se vuelve infinita.
- Para estos juegos, asuma que los jugadores maximizan la suma descontada de los pagos de los juegos simultáneos. Es decir, los jugadores deciden en función del valor presente de los pagos.
- Para el factor de descuento δ y el vector de pagos $\pi = (\pi_0, \ldots., \pi_t, \ldots)$ el valor presente de una secuencia de pagos es:

$$\mathbf{VP}\ (\boldsymbol{\pi}, \boldsymbol{\delta}) = \boldsymbol{\pi}_0 + \boldsymbol{\delta}\,\boldsymbol{\pi}_1 + \boldsymbol{\delta}^2\boldsymbol{\pi}_2 + \ldots + \boldsymbol{\delta}^t\boldsymbol{\pi}_t + \ldots$$

- A partir de aquí comenzaremos a operar con este tipo de pagos. *Para más detalle pueden mirar la nota de clase sobre series geométricas.*

Dilema del Prisionero repetido infinitamente

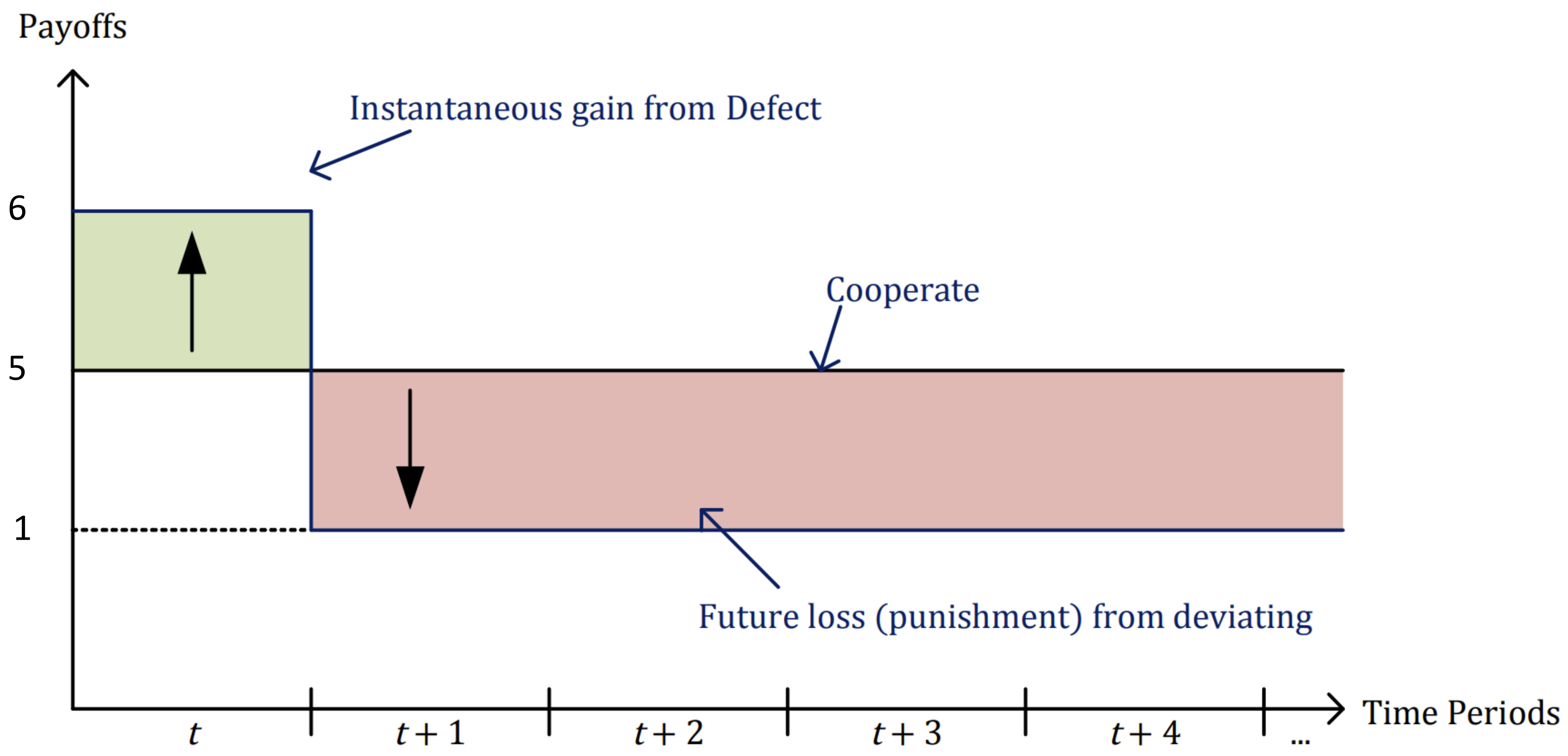
- No podemos usar inducción hacia atrás.

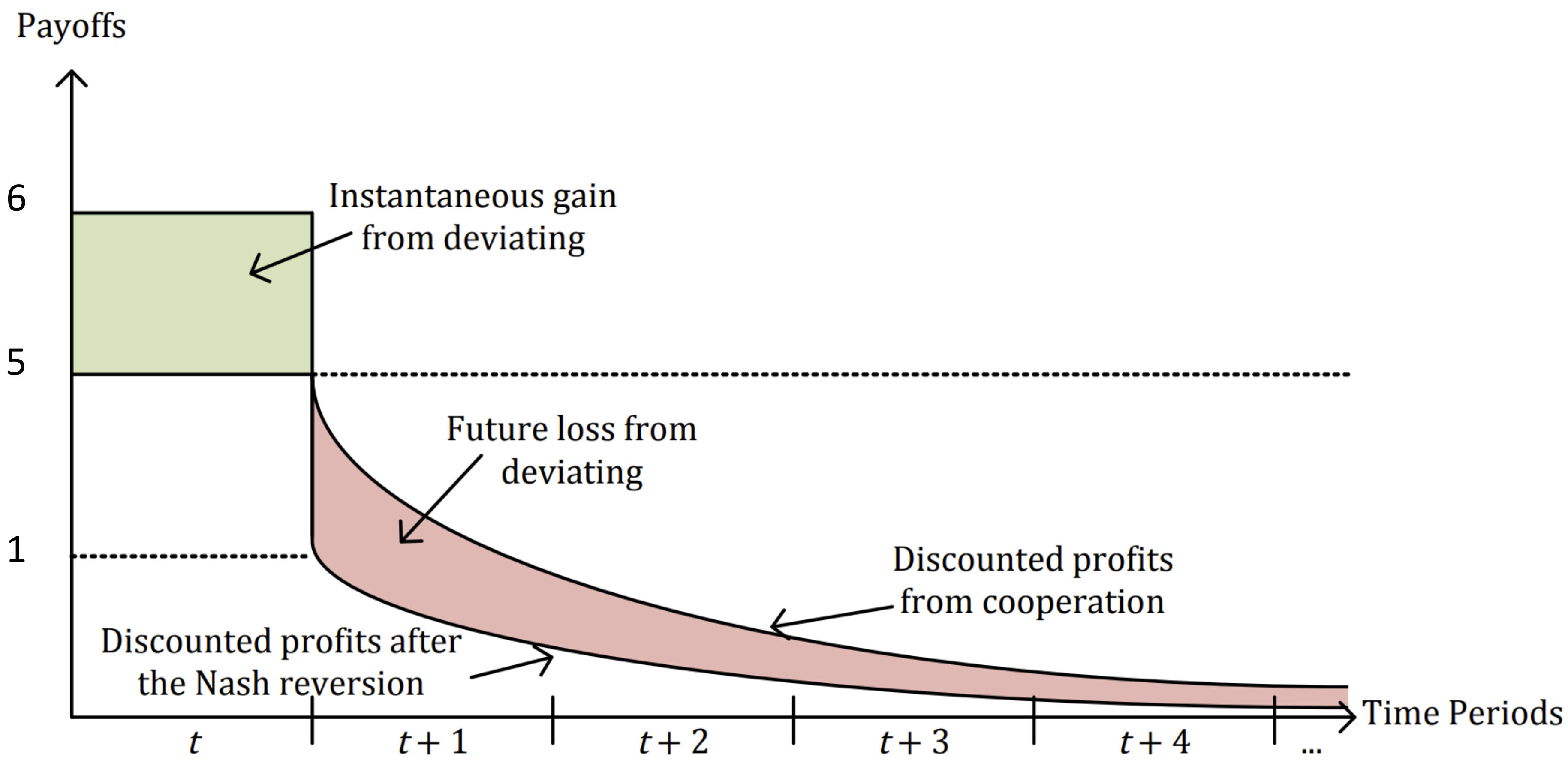
$$\begin{array}{c|c|c|}
 & C & D \\ \hline
C & 5,5 & 0,6 \\ \hline
D & 6,0 & 1,1 \\ \hline
\end{array},$$

Estrategia Grim-Trigger (Castigo eterno)

Perfil de estrategias Grim-Trigger:

- Comenzar jugando C en $t = 0$.
- Continuar jugando C siempre y cuando ambos jugadores hayan jugado C en el período anterior.
- Si algún jugador no jugó C en el período anterior, entonces ambos jugadores jugarán D para siempre.


Payoffs
Instantaneous gain from Defect
6
5
1
Cooperate
Future loss (punishment) from deviating
Time Periods
t
t + 1
t + 2
t + 3
t + 4
...


Payoffs
6
5
1
Instantaneous gain from deviating
Future loss from deviating
Discounted profits from cooperation
Discounted profits after the Nash reversion
Time Periods
t
t + 1
t + 2
t + 3
t + 4
...

Estrategia Trigger (Castigo eterno)

- **¿Me conviene cooperar si el otro jugador también juega la estrategia trigger?** Esto se responderá de forma afirmativa si el pago por cooperar es mayor al pago del desvío.

- Si cooperamos, obtenemos \$5 siempre

$$5 + \delta\, 5 + \delta^2\, 5 + \delta^3\, 5 \;\ldots = \frac{5}{1-\delta}$$

- Si hay desvío, obtenemos \$6 una vez y luego \$1 siempre

$$6 + \delta\, 1 + \delta^2\, 1 + \delta^3\, 1 \;\ldots = 6 + 1\,\frac{\delta}{1-\delta}$$

Estrategia Trigger (Castigo eterno)

- Me conviene cooperar si:

$$\frac{5}{(1-\delta)} \geq 6 + 1\frac{\delta}{(1-\delta)}$$

$$\delta \geq \frac{1}{5}$$

Al adoptar la estrategia gatillo, si ambos son suficientemente pacientes ($\delta \geq 1/5$) podrán alcanzar el óptimo (5, 5) en vez de llegar al resultado del Dilema del Prisionero del juego de un período (o estático).

Estrategia Trigger (Castigo eterno)

- **Unicidad de los resultados:** ¿Es el acuerdo que analizamos el único acuerdo que es un equilibrio si los jugadores son suficientemente pacientes?
- No. “Teoremas Folk”.
- El resultado donde los jugadores eligen (D, D) en todos los períodos de tiempo también se puede sostenerse como un resultado de equilibrio para todos los valores de δ.
- Dificultad en la predicción de los resultados.

Estrategia castigo temporario
Forgiving trigger

Perfil de estrategias Forgiving trigger:

- Comenzar jugando C en $t = 0$.
- Continuar jugando C siempre y cuando ambos jugadores hayan jugado C en el período anterior.
- Si algún jugador no jugó C en el período anterior, entonces ambos jugadores juegan D por una cantidad k de períodos.
- En el período $k + 1$ vuelven a cooperar.

Estrategia castigo temporario
Forgiving trigger

- Suponga que castigo se da por dos períodos $(\boldsymbol{k} = \mathbf{2})$. ¿Podemos alcanzar la cooperación si $\boldsymbol{\delta} = \frac{\mathbf{1}}{\mathbf{5}}$? ¿ y con $\boldsymbol{\delta} = \frac{\mathbf{3}}{\mathbf{5}}$?

- Nos ponemos en los pies del jugador 1. Partimos de la **historia** donde ambos jugadores cooperaron en $t = 0$. ¿Me conviene seguir cooperando?

$$
\begin{array}{c|c|c|}
 & C & D \\
\hline
C & 5,5 & 0,6 \\
\hline
D & 6,0 & 1,1 \\
\hline
\end{array},
$$

Estrategia castigo temporario

- **¿Me conviene cooperar si el otro jugador también juega la estrategia forgiving trigger?** Esto se responderá de forma afirmativa si el pago por cooperar es mayor al pago del desvío.

- Si cooperamos, obtenemos $5 siempre

$$5 + \delta 5 + \delta^2 5 + \delta^3 5 + \delta^4 5 \; \ldots.$$

- Si hay desvío, obtenemos $6 una vez y luego $1 por dos períodos. A partir del período siguiente volvemos a obtener $5.

$$6 + \delta 1 + \delta^2 1 + \delta^3 5 + \delta^4 5 \; \ldots.$$

Estrategia castigo temporario

- Me conviene cooperar si:

$$5 + \delta\, 5 + \delta^2\, 5 \geq 6 + \delta\, 1 + \delta^2\, 1$$
$$-1 + 4\delta + 4\delta^2 \geq 0$$

- Si $\boldsymbol{\delta = \frac{1}{5} = 0.2}$ reemplazo y obtenemos:

$$-1 + 4\frac{1}{5} + 4\left(\frac{1}{5}\right)^2 < 0$$
$$-1 + 0.8 + 0.16 < 0$$
$$-0.04 < 0$$

- Si $\boldsymbol{\delta = \frac{3}{5} = 0.6}$ reemplazo y obtenemos:

$$-1 + 4\frac{3}{5} + 4\left(\frac{3}{5}\right)^2 \geq 0$$
$$-1 + 2.4 + 1.44 \geq 0$$
$$2.84 \geq 0$$

Estrategia castigo temporario
Comentarios

- Cuando $(\boldsymbol{k}=\mathbf{2})$**,** un nivel de impaciencia de $\boldsymbol{\delta}=\frac{\mathbf{1}}{\mathbf{5}}$ no alcanza para sostener el acuerdo. Para que el acuerdo con dos períodos de castigo sea sostenible necesitamos que los jugadores sean más pacientes.
- Para $\boldsymbol{k}=\mathbf{2}$ necesitamos aproximadamente una tasa de impaciencia de al menos $\boldsymbol{\delta}=\mathbf{0.21}$.
- La posibilidad de alcanzar un acuerdo en estos contextos depende tanto de la tasa de impaciencia $(\boldsymbol{\delta})$ como de la dureza del castigo $(\boldsymbol{k})$.
- A medida que $\boldsymbol{k}\rightarrow\infty$, el nivel de impaciencia mínimo necesario para alcanzar el acuerdo tiende a $\frac{\mathbf{1}}{\mathbf{5}}$ (resultado estrategia castigo eterno).

Desconocimiento de la duración del juego

		Burger King	
		$20	$26
Mc Donald's	$20	288 , 288	360 , 216
	$26	216 , 360	324 , 324

- La clase pasada asumimos que los juegos eran finitos o infinitos y los jugadores conocían esto.
- Supongamos ahora que no sabemos la duración del juego, pero sí conocemos la probabilidad p de que el juego se repetirá en el siguiente período.

Desconocimiento de la duración del juego

		Burger King	
		$20	$26
Mc Donald's	$20	288 , 288	360 , 216
	$26	216 , 360	324 , 324

- Si cooperan, obtienen $324 siempre

$$324 + p\delta\, 324 + p^2\delta^2\, 324 + p^3\delta^3\, 324 \;\ldots = \frac{324}{1 - p\delta}$$

- Si hay desvío, obtienen $360 una vez y luego $288 siempre

$$360 + p\delta\, 288 + p^2\delta^2\, 288 + p^3\delta^3\, 288 \;\ldots = 360 + 288\, \frac{p\delta}{1 - p\delta}$$

- Me conviene cooperar si:

$$\frac{324}{(1 - p\delta)} \geq 360 + 288\, \frac{p\delta}{(1 - p\delta)}$$

$$p\delta \geq \frac{1}{2}$$

Desconocimiento de la duración del juego

		Burger King	
		$20	$26
Mc Donald's	$20	288 , 288	360 , 216
	$26	216 , 360	324 , 324

- Si cooperan, obtienen $324 siempre

$$324 + p\delta\, 324 + p^2\delta^2\, 324 + p^3\delta^3\, 324 \; \ldots. = \frac{324}{1 - p\delta}$$

- Si hay desvío, obtienen $360 una vez y luego $288 siempre

$$360 + p\delta\, 288 + p^2\delta^2\, 288 + p^3\delta^3\, 288 \; \ldots. = 360 + 288\, \frac{p\delta}{1 - p\delta}$$

- Me conviene cooperar si:

$$\frac{324}{(1 - p\delta)} \geq 360 + 288\, \frac{p\delta}{(1 - p\delta)}$$

$$p\delta \geq \frac{1}{2}$$

El comportamiento cooperativo sostenido por la estrategia gatillo puede romperse si hay una probabilidad suficientemente grande de que el juego repetido termine en el próximo período de juego, es decir, cuando hay un valor suficientemente pequeño de p.

Desconocimiento de la duración del juego

- Elegir cooperar está asociado a la probabilidad de continuidad del juego. Cuanto mayor sea la probabilidad de que el juego continúe, mayor será el incentivo a continuar cooperando.
- Si el juego dura un período, o pocos períodos, mayor será el incentivo a desviarse.