



Universidad
Zaragoza

Trabajo Fin de Grado

Campos colectivos en modelos sociales de
comportamiento corrupto

Autor

Jesús Sánchez Cornago

Directores

Luis Mario Floría Peralta

Francisco Juan Bauzá Mingueza

FACULTAD DE CIENCIAS
2021

No puedo empezar a escribir este trabajo sin dar las gracias a todas aquellas personas que han estado a mi lado y han hecho posible que esté donde estoy ahora. Primero, a mis padres y a mi hermano, que me han apoyado desde el principio; nunca han dudado de mí y se han sacrificado para que yo pueda estar donde estoy hoy. E incluso cuando les he fallado, me han dado una segunda oportunidad. También quiero dar las gracias a mis amigos, con los que he pasado tantas horas y que han hecho más llevadero el este reto que empecé hace cinco años. A Marta, que desde la distancia me ha acompañado hasta en los momentos más duros y me ha enseñado que siempre puedo mejorar. Y por último, a mis directores Mario Floría y Francisco Bauzá, quienes me han acompañado en mi primer trabajo de investigación. Han estado para ayudarme siempre que he necesitado, a la vez que me han dado la libertad para que yo pudiera avanzar y; en definitiva, sacar lo mejor de mí.

Aunque estas escasas palabras no hagan justicia a todo el apoyo que me habéis dado, quiero que sepáis que os estaré eternamente agradecido, y que, esté donde esté, siempre os llevaré en el recuerdo.

Índice

Introducción	1
1. El modelo HCO	2
2. Campos globales	3
3. Análisis teórico y simulaciones	4
3.1. Puntos fijos y estabilidad	4
3.2. Campos Lineales	8
3.3. Estados estacionarios en campo medio	10
3.4. Simulaciones de agentes	14
4. Biestabilidad inducida por campos globales	16
4.1. Estados estacionarios de campo medio	19
5. Conclusión	20
Bibliografía	22
A. Anexo I: Puntos fijos y estabilidad en diferentes situaciones	23
A.1. Campo global de corruptos sin wtw	23
A.2. Campo global de honestos sin wtw	25
A.3. Campo global de honestos con wtw	26
A.4. Ambos campos sin wtw	28
A.5. Ambos campos con wtw	30
B. Anexo II: Campos globales de honestidad lineales	32
C. Anexo III: resolución iterativa de las ecuaciones de campo medio para el campo de honestidad	33
D. Anexo IV: Biestabilidad para el campo de honestidad	35

Introducción

El propósito de este trabajo es la extensión de un modelo recientemente introducido [1], denominado HCO, de comportamientos sociales corruptos. Mientras otros estudios previos [2], [3] de este tipo de comportamientos utilizan el formalismo de la teoría matemática de juegos [4], el modelo HCO modeliza el fenómeno social de la corrupción como un proceso infeccioso, de modo que se inscribe en el marco formal de los modelos epidemiológicos [5].

En la sección 1 presentaremos en detalle el modelo HCO, donde los individuos de una sociedad (población) pueden estar en los estados H (honestidad), C (corrupción) u O (ostracismo). En este modelo, un individuo (o agente) en estado C es un individuo que incumple las normas sociales de honestidad, mereciendo como castigo su aislamiento social (exilio, cárcel) cuando es delatado por un individuo honesto. A su vez, un individuo H puede pasar al estado C por contacto con un individuo corrupto. Un individuo O, socialmente aislado, puede ser reinsertado en la sociedad (población socialmente activa) como individuo H. Además, un individuo corrupto, como consecuencia del miedo al castigo, puede pasar a ser honesto (conversión) si percibe un alto riesgo, cuantificado por la proporción en la población de individuos en estado O.

Las decisiones humanas sobre comportamientos sociales, no solo vienen condicionadas por los contactos sociales (interacciones) entre pares de individuos; en ocasiones el “clima social” influye de un modo determinante. Por ejemplo, la presencia en la sociedad de un elevado número de agentes corruptos genera un clima social que puede provocar, por imitación, la corrupción de individuos honestos, sin la mediación de contactos individuales. De modo análogo, una alta proporción de honestos (clima social de honestidad) puede determinar la conversión espontánea de individuos corruptos. En la sección 2 implementaremos en el modelo HCO original estos efectos, que consideramos socialmente relevantes, mediante la introducción de dos flujos adicionales (transiciones entre estados), $H \rightarrow C$ y $C \rightarrow H$, con probabilidades de transición dadas, respectivamente, por sendas funciones de la proporción de corruptos y de la proporción de honestos. En adelante nos referiremos a ellas como campo global de corrupción y campo global de honestidad, respectivamente.

En la sección 3 se incluye únicamente el campo global de corrupción al modelo HCO, y se estudia con detalle cómo la expresión del campo modifica los resultados obtenidos en [1], tanto analítica como numéricamente. Para acabar la sección, las predicciones teóricas se verifican con simulaciones Monte Carlo de agentes.

Por último, en la sección 4, se estudian las condiciones que deben cumplir las expresiones del campo global de corrupción para que se dé la coexistencia entre dos estados estables, y se presentan varias resoluciones numéricas para apoyar esto. Tanto el estudio que se realiza en esta sección como en la sección 3, se realiza para el campo global de honestidad en los anexos.

1. El modelo HCO

Consideremos una población de N agentes, el estado de cada uno de los cuales, σ_i ($i = 1, \dots, N$), puede ser H, C u O. En problemas de dinámica de poblaciones es habitual pensar en los estados como “compartimentos” que contienen a los agentes. La red de contactos sociales entre individuos viene dada por la llamada *matriz de adyacencia* \mathbf{A} , simétrica, cuyo elemento $A_{i,j}$ es 1 si los agentes (nodos) i, j son vecinos y 0 si no lo son. Se supondrá que los contactos permanecen fijos en el tiempo. Asimismo denotaremos mediante $k(i)$ al número de vecinos del individuo i .

Las transiciones entre estados, a las que nos referimos como flujos entre compartimentos, están determinadas por los estados de los vecinos:

- Flujo $H \rightarrow C$: La probabilidad (condicionada) de que un agente i honesto pase a corrupto por interacción con sus vecinos es

$$f_{\alpha}^{HC}(i, \{\sigma_j\}) = 1 - \prod_{j=1}^N (1 - A_{ij}\alpha\delta_{\sigma_j,C}), \quad (1)$$

donde $\{\sigma_j\}$ denota el microestado del sistema, $\delta_{x,y}$ es la delta de Kronecker, y α es la tasa de corrupción, es decir, la probabilidad de que un contacto H-C resulte en la corrupción del agente honesto.

- Flujo $C \rightarrow O$: La probabilidad de que un agente corrupto i sea delatado por su entorno es

$$f_{\beta}^{CO}(i, \{\sigma_j\}) = 1 - \prod_{j=1}^N (1 - A_{ij}\beta\delta_{\sigma_j,H}), \quad (2)$$

donde β es la tasa de delación, es decir, la probabilidad de que un contacto H-C resulte en la delación del agente corrupto.

- Flujo $O \rightarrow H$: La probabilidad de que un agente O sea reinsertado es un parámetro r , $0 \leq r \leq 1$.
- Flujo $C \rightarrow H$ *warning to wrongdoers (wtw)*: Con probabilidad dada por la proporción

$$\langle O \rangle = \frac{1}{N} \sum_{i=1}^N \delta_{\sigma_i,O} \quad (3)$$

de individuos castigados, un individuo corrupto se convierte a honesto, como efecto del temor a ser delatado.

Para poder estudiar el proceso de Markov asociado a este modelo, se realiza la siguiente aproximación: en vez de clasificar cada miembro en un grupo, se le asocia una cierta probabilidad de que sea *honesto*, *corrupto* u *ostracismo* y al haber N agentes, habrá N vectores asociados a estos, quedando en total $3N$ variables. Esta información se recoge en un vector cuyas componentes son las probabilidades de que el agente se encuentre en cada uno de los estados respectivamente:

$$\vec{\rho}(i, t) = (\rho_h(i, t), \rho_c(i, t), \rho_o(i, t)) \quad (4)$$

Las ecuaciones (1) y (2) quedan consecuentemente modificadas del modo siguiente:

$$f_\alpha(i, \{\vec{\rho}_j\}) = 1 - \prod_{j=1}^N [1 - A_{ij} \alpha \rho_c(j)] \quad (5)$$

$$f_\beta(i, \{\vec{\rho}_j\}) = 1 - \prod_{j=1}^N [1 - A_{ij} \beta \rho_h(j)] \quad (6)$$

Por otra parte, la ecuación (3) se reescribe quedando:

$$\langle \rho_o \rangle = \frac{1}{N} \sum_{i=1}^N \rho_o(i) \quad (7)$$

Cada paso temporal ($t \rightarrow t + 1$) es un turno en el que cada agente interacciona con sus vecinos modificándose las probabilidades de pertenecer a cada estado. De esta forma el estado de un nodo en un tiempo $t + 1$ $\vec{\rho}(i, t + 1)$ se relaciona con el estado en el tiempo t $\vec{\rho}(i, t)$ según:

$$\vec{\rho}(i, t + 1) = \mathbf{Q}_i \cdot \vec{\rho}(i, t), \quad (8)$$

donde \mathbf{Q}_i es la matriz

$$\mathbf{Q}_i = \begin{pmatrix} P_{H \rightarrow H} & P_{C \rightarrow H} & P_{O \rightarrow H} \\ P_{H \rightarrow C} & P_{C \rightarrow C} & P_{H \rightarrow C} \\ P_{H \rightarrow O} & P_{C \rightarrow O} & P_{H \rightarrow =} \end{pmatrix} = \begin{pmatrix} 1 - f_\alpha & \langle \rho_o \rangle & r \\ f_\alpha & (1 - \langle \rho_o \rangle)(1 - f_\beta) & 0 \\ 0 & (1 - \langle \rho_o \rangle)f_\beta & 1 - r \end{pmatrix} \quad (9)$$

En el artículo [1] se hace un estudio detallado del modelo, tanto analítica como numéricamente. Los resultados analíticos se obtienen mediante la aproximación de *campo medio* (véase sección 3), reflejando la existencia de tres tipos de puntos fijos de la dinámica del sistema: “Full H” (Todos honestos), “Full C” (Todos corruptos) y soluciones mixtas (*mixed*), en las que coexisten las tres clases de agentes. Para un conjunto de parámetros dado, solo hay una solución asintótica estable, de modo que es imposible que haya coexistencia entre estados asintóticos (*biestabilidad*) de cualquier tipo, y se determinan los valores críticos de los parámetros α y β , a los que ocurren las transiciones entre los diferentes tipos de soluciones asintóticas. También se observa el efecto del término *wtw*. Este juega un papel importante en las transiciones “Full C”-*mixed*, reduciendo el rango de estabilidad del estado “Full C”. Sin embargo, no tiene un papel relevante en las transiciones de “Full H”-*mixed*. Todos estos resultados se confirman con simulaciones de agentes.

2. Campos globales

Como se dijo en la introducción, extendemos aquí el modelo HCO, mediante la inclusión de los campos globales de corrupción y/o de honestidad. Las probabilidades de transición asociadas a estos flujos inducidos por el clima social, vienen dadas respectivamente por las funciones $g(\langle \rho_c \rangle)$ y $h(\langle \rho_h \rangle)$, que por ahora no especificamos, pero que supondremos funciones analíticas (desarrollables en serie de potencias). Por claridad, de ahora en adelante escribiremos estas funciones y los términos en las ecuaciones asociados a ellas en color rojo y verde, respectivamente. Vemos a continuación cómo se ve modificado

el elemento $P_{H \rightarrow C}$ de la matriz de la ecuación (9) al incluir el efecto del **campo global de corrupción**.

Un nodo honesto de la red interacciona con sus contactos. Como se ha descrito en la ecuación (1) se corromperá con una probabilidad f_α o se mantendrá honesto con una probabilidad $1 - f_\alpha$, pudiendo entonces pasar a corrupto con probabilidad $g(\langle \rho_c \rangle)$ de modo que

$$P_{H \rightarrow C} = f_\alpha + g(\langle \rho_c \rangle)(1 - f_\alpha) \quad (10)$$

Siguiendo este razonamiento para **el campo global de honestidad**, el elemento $P_{C \rightarrow H}$ de la matriz queda:

$$P_{C \rightarrow H} = \langle \rho_o \rangle + h(\langle \rho_h \rangle)(1 - \langle \rho_o \rangle) \quad (11)$$

Teniendo en cuenta los cambios de estos y otros elementos, la matriz final \mathbf{Q}_i asociada al vector de probabilidades de cada agente tendrá la siguiente expresión:

$$\mathbf{Q}_i = \begin{pmatrix} 1 - f_\alpha - g(\langle \rho_c \rangle)(1 - f_\alpha) & \langle \rho_o \rangle + h(\langle \rho_h \rangle)(1 - \langle \rho_o \rangle) & r \\ f_\alpha + g(\langle \rho_c \rangle)(1 - f_\alpha) & (1 - \langle \rho_o \rangle)(1 - h(\langle \rho_h \rangle))(1 - f_\beta) & 0 \\ 0 & (1 - \langle \rho_o \rangle)(1 - h(\langle \rho_h \rangle))f_\beta & 1 - r \end{pmatrix} \quad (12)$$

3. Análisis teórico y simulaciones

Como primer paso en el análisis de un fenómeno colectivo, siempre es muy recomendable llevar a cabo una aproximación de campo medio, cuya hipótesis esencial es la homogeneidad del sistema. En nuestro caso, esto quiere decir homogeneidad de probabilidades (ρ_h, ρ_c, ρ_o) de los estados de agente y de la red de contactos de manera que

- (a) todo agente se comporta como el promedio de todos ellos $\vec{\rho}_i(t) = \langle \vec{\rho} \rangle$ (agente promedio), y
- (b) los $k(i) = k$ vecinos del agente promedio pueden seleccionarse al azar de entre la población en cada paso de tiempo (*población bien mezclada*).

Dado que $0 \leq \rho_h, \rho_c, \rho_o \leq 1$ y $\rho_h + \rho_c + \rho_o = 1$, el espacio de fases del sistema se ve reducido al triángulo (o simplex) en el plano (ρ_h, ρ_c) cuyos vértices serán $(0,0)$, $(1,0)$ “Full H” y $(0,1)$ “Full C”. Las probabilidades de transición f_α y f_β vienen ahora dadas por

$$\begin{cases} f_\alpha = 1 - (1 - \alpha \rho_c)^k \\ f_\beta = 1 - (1 - \beta \rho_h)^k \end{cases} \quad (13)$$

3.1. Puntos fijos y estabilidad

Por simplicidad, presentamos el estudio de los puntos fijos del sistema y su estabilidad, incluyendo únicamente el campo global de corrupción y el efecto *wtw*. El análisis análogo para el campo global de honestidad y de ambos campos globales sin *wtw* se presenta de forma detallada en el anexo A. Al limitarnos al campo global de corrupción, la evolución del sistema sobre el simplex viene dada en tiempo discreto por:

$$\begin{cases} \rho_h(t+1) = (1 - f_\alpha - g(\rho_c)(1 - f_\alpha))\rho_h + \rho_o\rho_c + r\rho_o \\ \rho_c(t+1) = (f_\alpha + g(\rho_c)(1 - f_\alpha))\rho_h + [(1 - \rho_o)(1 - f_\beta)]\rho_c \end{cases} \quad (14)$$

que pasando a tiempo continuo y sustituyendo $\rho_o = 1 - \rho_h - \rho_c$ queda:

$$\begin{cases} F_h \equiv \dot{\rho}_h = -[f_\alpha + g(\rho_c)(1 - f_\alpha) + \rho_c + r]\rho_h + (r + \rho_c)(1 - \rho_c) \\ F_c \equiv \dot{\rho}_c = [f_\alpha + g(\rho_c)(1 - f_\alpha) + (1 - f_\beta)\rho_c]\rho_h + [(1 - f_\beta)\rho_c - 1]\rho_c, \end{cases} \quad (15)$$

donde denotamos mediante F_h y F_c las componentes del campo de velocidades. Así, un punto fijo satisface que $F_h = 0$ y $F_c = 0$. El lugar geométrico de los puntos que satisfacen $F_h = 0$ es la nulclina de F_h , y el de los puntos que cumplen $F_c = 0$ define la nulclina de F_c . En general, una nulclina consta de diversas líneas o ramas. Un punto fijo es necesariamente un punto de intersección de una rama de la nulclina de F_h y una rama de la nulclina de F_c . De estos puntos fijos, los casos más asequibles de estudiar son los puntos (1,0) “Full H” y (0,1) “Full C”. Primero, hay que confirmar si estos siguen siendo puntos fijos con la modificación incluida:

• “Full H”:

$$\begin{cases} F_h(1, 0) = -g(0) - r + r = -g(0) \\ F_c(1, 0) = g(0) \end{cases} \quad (16)$$

• “Full C”:

$$\begin{cases} F_h(0, 1) = (r + 1)(1 - 1) = 0 \\ F_c(0, 1) = [(1 - f_\beta(0)) \cdot 1 - 1] \cdot 1 = 0 \end{cases} \quad (17)$$

“Full C” es punto fijo siempre independientemente de $g(\rho_c)$. Sin embargo, **el punto (1,0) “Full H” será punto fijo solo cuando $g(0) = 0$** . Lo cual es de esperar. Si $g(0) > 0$, el agente tiene una tendencia innata a corromperse, de modo que en una población totalmente honesta existe una probabilidad no nula de que un agente honesto se corrompa.

A partir de (15) podemos ver fácilmente que el eje de abscisas, en ausencia de campo global de corruptos, es una rama de nulclina de F_c , mientras que en presencia de este, deja de serlo, a menos que $g(0) = 0$, pues:

$$F_c(\rho_h, 0) = g(0)\rho_h \quad (18)$$

Por las mismas razones, *mutatis mutandi*, lo mismo ocurre con el campo global de honestos para el punto “Full C”. En este caso, $h(0) > 0$ indica que el agente tiene una tendencia innata a ser honesto.

Asumiendo que $g(0) = 0$ vamos a estudiar la estabilidad de los puntos fijos “Full H” y “Full C”. Para ello, calculamos el jacobiano, que viene dado por:

$$\mathbf{J} = \begin{pmatrix} \frac{\partial F_h}{\partial \rho_h} & \frac{\partial F_c}{\partial \rho_h} \\ \frac{\partial F_h}{\partial \rho_c} & \frac{\partial F_c}{\partial \rho_c} \end{pmatrix} \quad (19)$$

La expresión de las derivadas parciales es la siguiente:

$$\begin{cases} \frac{\partial F_h}{\partial \rho_h} = -(f_\alpha + g(\rho_c)(1 - f_\alpha) + \rho_c + r) \\ \frac{\partial F_h}{\partial \rho_c} = -\rho_h(f'_\alpha(1 - g(\rho_c)) + g'(\rho_c)(1 - f_\alpha) + 1) + 1 - r - 2\rho_c \\ \frac{\partial F_c}{\partial \rho_h} = f_\alpha + g(\rho_c)(1 - f_\alpha) + (1 - f_\beta)\rho_c - f'_\beta\rho_c\rho_h - f'_\beta\rho_c^2 \\ \frac{\partial F_c}{\partial \rho_c} = \rho_h(f'_\alpha(1 - g(\rho_c)) + g'(\rho_c)(1 - f_\alpha) + (1 - f_\beta)) + 2(1 - f_\beta)\rho_c - 1 \end{cases} \quad (20)$$

A continuación, evaluamos las derivadas en el punto fijo en cuestión. Teniendo en cuenta que la ecuación de autovalores es $\lambda^2 - T\lambda + D = 0$, para que el punto fijo sea estable se debe cumplir que la traza (T) del jacobiano sea negativa, y el determinante (D), positivo.

•(0, 1)

$$\mathbf{J}_c = \begin{pmatrix} -(f_\alpha(1) + g(1)(1 - f_\alpha(1)) + 1 + r) & -r - 1 \\ f_\alpha(1) + g(1)(1 - f_\alpha(1)) + 1 - f'_\beta(0) & 1 \end{pmatrix}$$

$$\begin{cases} T = -(f_\alpha(1) + g(1)(1 - f_\alpha(1)) + r) < 0 \text{ siempre} \\ D = r[f_\alpha(1) + g(1)(1 - f_\alpha(1))] - f'_\beta(0)(1 + r) \end{cases} \quad (21)$$

Entonces, para que el punto sea estable, $D \geq 0$, y por tanto:

$$f_\alpha(1) + g(1)(1 - f_\alpha(1)) > f'_\beta(0) \left(1 + \frac{1}{r}\right) \quad (22)$$

•(1, 0)

$$\mathbf{J}_h = \begin{pmatrix} -r & -r - f'_\alpha(0) - g'(0) \\ 0 & f'_\alpha(0) + g'(0) - f_\beta(1) \end{pmatrix} \Rightarrow \begin{cases} \lambda_1 = -r \\ \lambda_2 = f'_\alpha(0) + g'(0) - f_\beta(1) \end{cases} \quad (23)$$

Así, si $\lambda_2 < 0$ el punto fijo será estable, y para que esto se cumpla:

$$g'(0) < f_\beta(1) - f'_\alpha(0) \quad (24)$$

Con las condiciones obtenidas, podemos obtener las regiones para las cuales “Full H” y “Full C” son estables en el diagrama de fases (β, α) .

$$\alpha \leq \alpha_c = \frac{1 - (1 - \beta)^k - g'(0)}{k} \quad \text{Región “Full H” estable} \quad (25)$$

$$\beta \leq \beta_c = \left(\frac{r}{r+1}\right) \frac{1 - (1 - \alpha)^k(1 - g(1))}{k} \quad \text{Región “Full C” estable} \quad (26)$$

Apoyamos estos resultados con la figura 1, en la que se compara el modelo HCO con la

inclusión del campo global de corrupción.

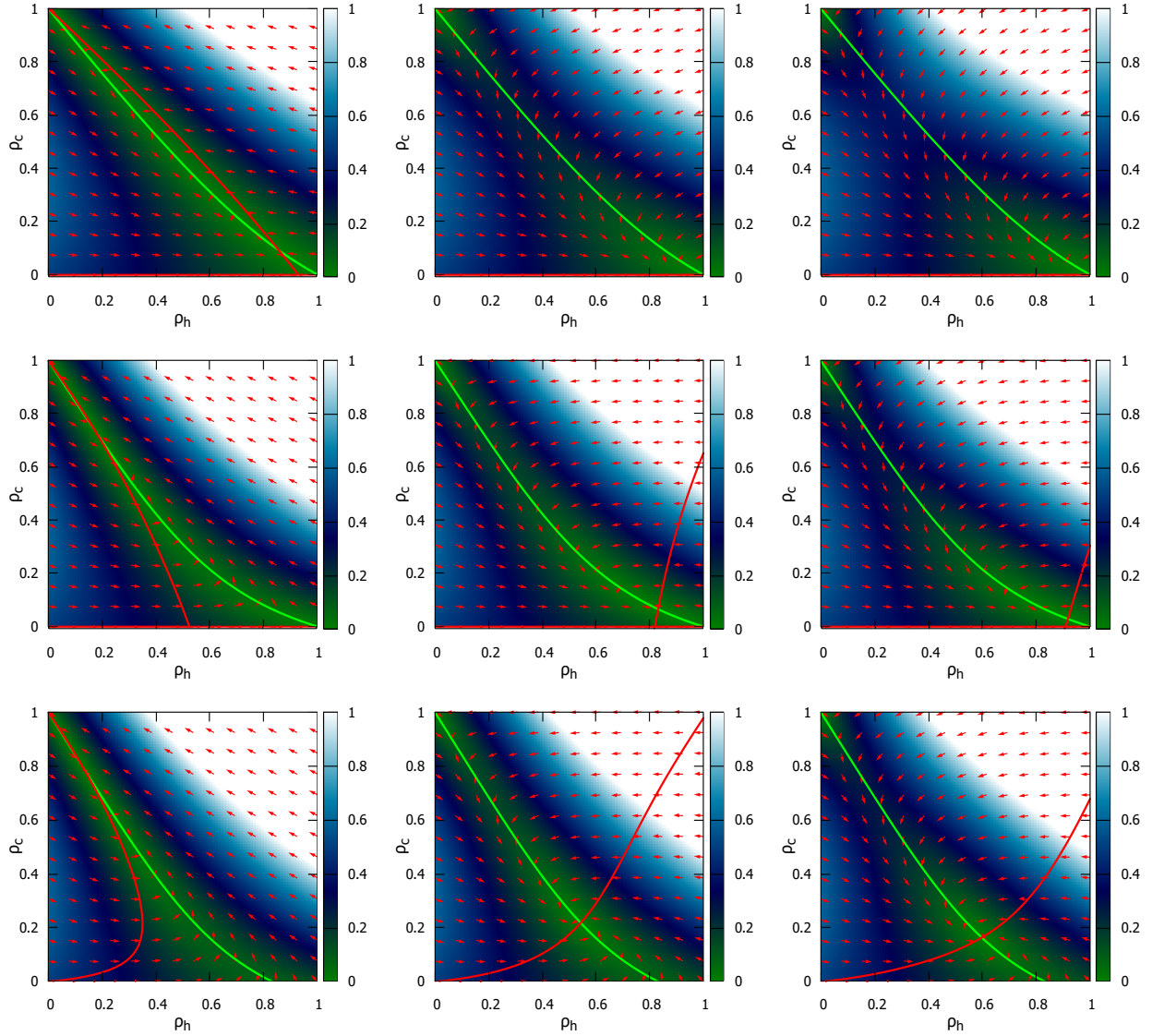


Figura 1: Nulclinas de F_h (verde) y F_c (roja) y dirección y módulo (barra de color) del flujo. $\alpha = 0.1$ y $r = 0.5$ para todas las gráficas. Cada columna representa un mismo valor de β , tomando de izquierda a derecha, los valores $\beta = 0.1, 0.5, 0.9$. Cada fila de gráficas representa la dinámica para una misma expresión de $g(\rho_c)$, siendo la primera fila la correspondiente al modelo HCO ($g(\rho_c) = 0$); la segunda fila, la correspondiente a $g(\rho_c) = 0.7\rho_c$; y la tercera, a $g(\rho_c) = 0.1 + 0.7\rho_c$. En las dos primeras filas se aprecia una rama de la nulclina de F_c sobre el eje de abscisas, mientras que para la última desaparece.

Para cada fila de gráficas vemos que la nulclina de F_h es siempre la misma. Esto es porque F_h es independiente de β . Además, en la última fila se aprecia que desaparece la rama $\rho_c = 0$ de la nulclina de F_c , en concordancia con el análisis anterior. La estabilidad de los puntos fijos depende de cómo se cortan las nulclinas, ya que hará que el campo de velocidades converja o diverja en el punto fijo como se explica en [1]. En el caso “Full H”, el punto fijo será estable cuando el corte una rama de la nulclina de F_c con el eje horizontal no esté en el simplex: denotando este punto por $(\rho_h^*, 0)$, cuando se cumpla que $\rho_h^* > 1$ el punto fijo “Full H” será estable, cumpliéndose la condición de estabilidad dada por $\lambda_2 < 0$ en (24). El punto fijo “Full C” será estable cuando la pendiente de la rama de

la nulclina de F_c sea menor que la pendiente de la rama de la nulclina de F_h . Dicho de otra forma, cuando la nulclina de F_c quede por debajo de la nulclina de F_h tras el corte, “Full C” será estable. Este mismo razonamiento se puede aplicar al punto fijo interno que aparece para algunos valores de los parámetros. En [6] se justifica cómo la geometría de las nulclinas cuando se cruzan da lugar a puntos fijos estables o inestables.

Una vez explicada la interpretación de las gráficas, queda estudiar el efecto de incluir el campo global de corrupción. La primera diferencia notable entre las dos primeras filas de gráficas y la última es que la nulclina de F_h ya no pasa por $(1,0)$, y ya no hay una rama de la nulclina de F_c que pase por el eje de abscisas. Consecuentemente queda probado gráficamente que cuando $g(0) > 0$ “Full H” ya no es un punto fijo, y en cambio, “Full C” siempre será punto fijo aunque en alguna de las gráficas no se observe el cruce entre las nulclinas de F_h y F_c . Esto se debe a que la rama de F_c que pasa por $(0,1)$ no entra en el simplex. Otro resultado que se desprende de incluir este campo global es un desplazamiento de la dinámica hacia $(0,1)$: el punto fijo interno que se observa en $\beta = 0.1$ se desplaza hacia esta esquina; y para $\beta = 0.5, 0.9$, aun cuando $g(0) = 0$, aparece un punto fijo estable que no está en el modelo original, para el que el punto fijo estable es $(1,0)$.

3.2. Campos Lineales

Al introducir el campo global de corrupción, hemos visto que los puntos fijos y su estabilidad dependen de $g(0)$, $g'(0)$ y $g(1)$ (se obtienen resultados análogos en caso de estudiar el campo global de honestos $h(\rho_h)$). En el apartado anterior, hemos utilizado funciones lineales para representar el campo global de corrupción en la figura 1. En este apartado queremos estudiar más detalladamente el impacto que tienen sobre el sistema. Como el efecto principal del término independiente ya está caracterizado lo omitimos en adelante, y la función del campo global de corrupción que adoptamos es:

$$g(\rho_c) = g_1 \rho_c, \quad (27)$$

de forma que $g'(0) = g(1) = g_1$. Queremos analizar si al introducir los campos globales se abre la posibilidad de que el sistema presente biestabilidad, es decir, estabilidad simultánea de dos puntos fijos para el mismo conjunto de valores de los parámetros, algo que en el modelo HCO original no ocurría. En concreto, estudiamos la posible biestabilidad de los puntos fijos “Full H” y “Full C”. Ello solo ocurrirá si las regiones del plano (α, β) determinadas por la ecuación (25) y la ecuación (26) (despejando α en función de β) se solapan. Para observar este posible solapamiento, estudiamos la monotonía y curvatura de ambas ecuaciones.

Así, para la ecuación (25):

$$\frac{\partial \alpha_c}{\partial \beta} = -(1 - \beta)^{k-1}(-1) = (1 - \beta)^{k-1} \Rightarrow \frac{\partial^2 \alpha_c}{\partial \beta^2} = -(k - 1)(1 - \beta)^{k-2} < 0 \quad (28)$$

Llamamos $\bar{\alpha}(\beta)$ a la función inversa de $\beta_c(\alpha)$; es decir, $\beta_c(\bar{\alpha}(\beta)) = \beta$. Entonces,

$$\beta = \frac{r}{r+1} \frac{1 - (1 - \alpha)^k (1 - g_1)}{k} \Rightarrow \bar{\alpha} = 1 - \left(\frac{1 - \beta k \left(\frac{r+1}{r} \right)}{1 - g_1} \right)^{\frac{1}{k}} \quad (29)$$

$$\frac{\partial \bar{\alpha}}{\partial \beta} = \left(\frac{1 - \beta k \left(\frac{r+1}{r} \right)}{1 - g_1} \right)^{\frac{1}{k}-1} \frac{r+1}{(1-g_1)r} \quad (30)$$

$$\frac{\partial^2 \bar{\alpha}}{\partial \beta^2} = \left(\frac{1 - \beta k \left(\frac{r+1}{r} \right)}{1 - g_1} \right)^{\frac{1}{k}-2} \frac{(k-1)(r+1)^2}{(1-g_1)^2 r^2} > 0 \quad (31)$$

Concluimos que $\alpha_c(\beta)$ es una función convexa creciente que corta el eje de abscisas en el valor de $\beta^* = 1 - (1 - g_1)^{1/k}$, con pendiente $(1 - g_1)^{\frac{k-1}{k}}$, mientras que $\bar{\alpha}(\beta)$ es una función cóncava creciente que corta el eje de abscisas en $\beta_c(0) = \frac{r}{r+1} \frac{g_1}{k} \leq \frac{g_1}{2k}$ (pues $r \leq 1$), con pendiente $\frac{r+1}{r(1+g_1)}$. Ha de notarse que esta última pendiente es estrictamente mayor que la de $\alpha_c(\beta)$ en β^* .

En primer lugar, tenemos que probar que $\beta^* \geq \beta_c(0) \forall g_1$ t.q. $0 < g_1 < 1$, para lo que basta probar que las funciones $f_1(x) = 1 - (1 - x)^{1/k}$ y $f_2(x) = \frac{x}{2k}$ satisfacen que $f_1(x) > f_2(x) \forall x$ t.q. $0 < x \leq 1$ y $k > 1$. En efecto,

(a) $f_1(0) = f_2(0) = 0$

(b) $f_1'(x) = \frac{1}{k}(1-x)^{\frac{1}{k}-1}$; $f_2'(x) = \frac{1}{2k} \Rightarrow f_1'(0) = \frac{1}{k} > f_2'(0)$

(c) $f_1''(0) = \frac{1}{k} \left(1 - \frac{1}{k}\right) (1-x)^{\frac{1}{k}-2} > 0 \forall x$ t.q. $0 \leq x \leq 1$

De a), b) y c) se concluye que, para todos $0 < x \leq 1$ y $k > 1$, $f_1(x) > f_2(x)$. Por lo tanto, las regiones de “Full H estable” y “Full C estable” no se solapan en el eje de abscisas. Del carácter convexo y creciente de $\alpha_c(\beta)$ y cóncavo creciente de $\bar{\alpha}(\beta)$, junto con la desigualdad estricta de sus pendientes en el eje de abscisas, se concluye que tampoco hay solapamiento para cualquier valor de x .

Como visualización, mostramos a continuación tres gráficas en las que aparece el diagrama de fases (β, α) . La primera gráfica es la correspondiente al modelo original ($g_1 = 0$). Las dos siguientes muestran el diagrama de fases para dos valores de g_1 . Para estas representaciones se ha elegido $r = 0.5$.

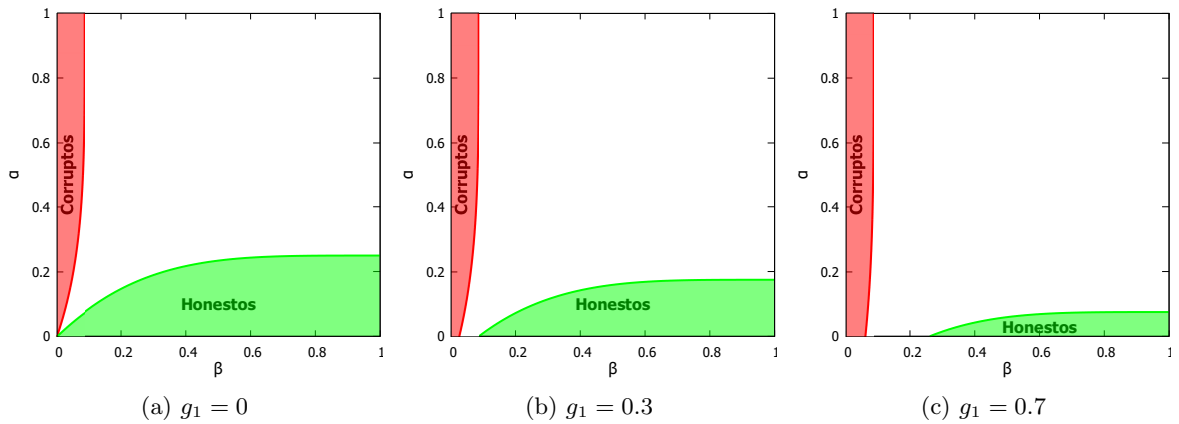


Figura 2: Diagramas de fases para distintos valores de g_1

En primer lugar se aprecia que al aumentar g_1 la distancia entre estas regiones aumenta, comprobándose gráficamente que no se produce solapamiento. Además, el tamaño de la

región de “todos corruptos” aumenta; y la región “todos honestos” disminuye, como cabía esperar, ya que el campo favorece la tendencia a la corrupción. Por otra parte, vemos que para $g_1 = 0.3$ la mayor diferencia con el modelo original se produce para valores de α y β bajos, de donde se deduce que cuando tenemos una interacción débil entre los agentes, el campo global juega un papel importante en la dinámica del sistema aunque el término lineal sea pequeño.

3.3. Estados estacionarios en campo medio

En este apartado presentamos la resolución numérica de la dinámica de campo medio y markoviana a través de la iteración de las respectivas ecuaciones y las comparamos entre sí y con el modelo HCO.

Observando las ecuaciones (15), nos damos cuenta que no es posible resolver estas ecuaciones ($F_h = 0$, $F_c = 0$) para obtener una expresión cerrada de los puntos fijos (ρ_h , ρ_c) en función de los parámetros del sistema. Por este motivo, hemos utilizado un método numérico de iteración de las ecuaciones, fijando el valor de los parámetros, para poder obtener el punto fijo como solución asintótica de las ecuaciones del modelo.

De esta manera, fijando el valor de r , la expresión de $g(\rho_c)$ y de uno de los parámetros α o β , dejando el variable, obtenemos la dependencia del punto fijo (ρ_h ó ρ_c) en función del parámetro variable. Cuando α es el parámetro variable (figura 4), se observan las transiciones $H - mixed$; mientras que si β es el parámetro variable (figura 5), las transiciones que se tienen lugar son $C - mixed$. Ocasionalmente si se escogen adecuadamente los valores de los parámetros fijos pueden verse transiciones $H - mixed - C$ cuando es α el parámetro que crece, o viceversa si es β el parámetro que crece.

La iteración de las ecuaciones markovianas se realiza sobre redes de 10000 nodos con $k = 4$ contactos por nodo. Las redes se eligen con dos estructuras diferentes: en una de ellas los nodos se disponen en una red cuadrada plana 100×100 en la que imponemos condiciones de contorno periódicas. Esta red la llamamos *Lattice*. En la segunda, que se conoce como *Random Regular Network* (RRN), a cada nodo se le asignan exactamente 4 vecinos al azar en la red. Ambas son redes homogéneas, y una de sus propiedades es que los resultados numéricos no difieren de los obtenidos por la iteración de las ecuaciones de campo medio, como se aprecia en la figura 3. Por lo tanto, se empleará la aproximación de campo medio a la hora de iterar las ecuaciones, ya que se ahorra una notable cantidad de tiempo en cada ejecución del programa. En el anexo C se muestra algunos ejemplos con el campo global de honestidad.

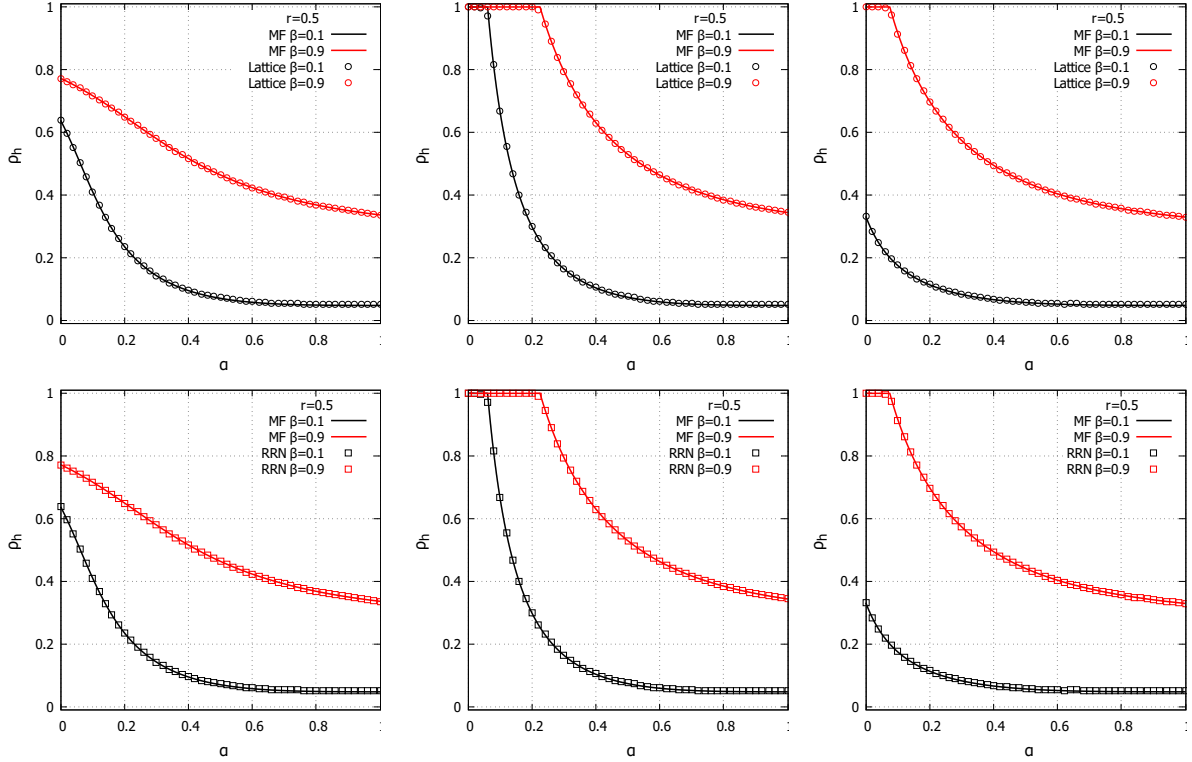
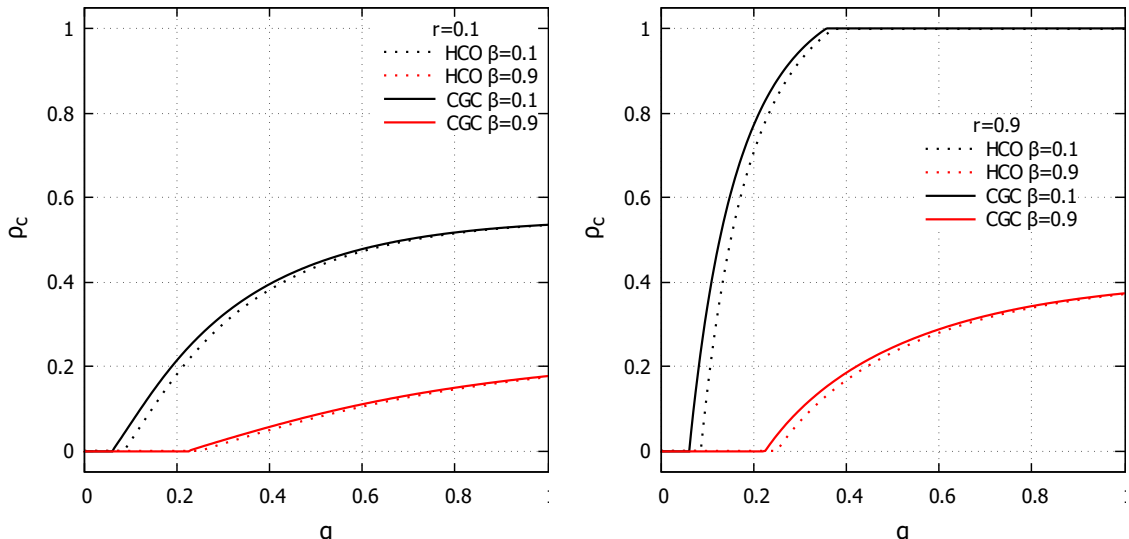


Figura 3: Porcentaje de honestos en función de la tasa de corrupción $\rho_h(\alpha)$. Los valores de los parámetros fijos escogidos son $r = 0.5$ y $\beta = 0.1$ (negro), $\beta = 0.9$ (rojo). Por columnas (de izquierda a derecha), las expresiones del campo global de corrupción son: $g(\rho_c)0.1 + 0.1\rho_c$, $g(\rho_c) = 0.1\rho_c$, $g(\rho_c)0.7\rho_c$. Se la resolución de campo medio (curvas continuas) frente a Markov para redes homogéneas, donde la red escogida en la primera fila es una *Lattice* (círculos); y en la segunda, una *RRN* (cuadrados).

A continuación estudiamos en detalle cómo la inclusión del campo global de corrupción modifica las predicciones del modelo original. De esta manera se analiza el valor de α y β críticos, el papel que juega r y cómo el campo global de corrupción (CGC) influye en los puntos fijos estables.



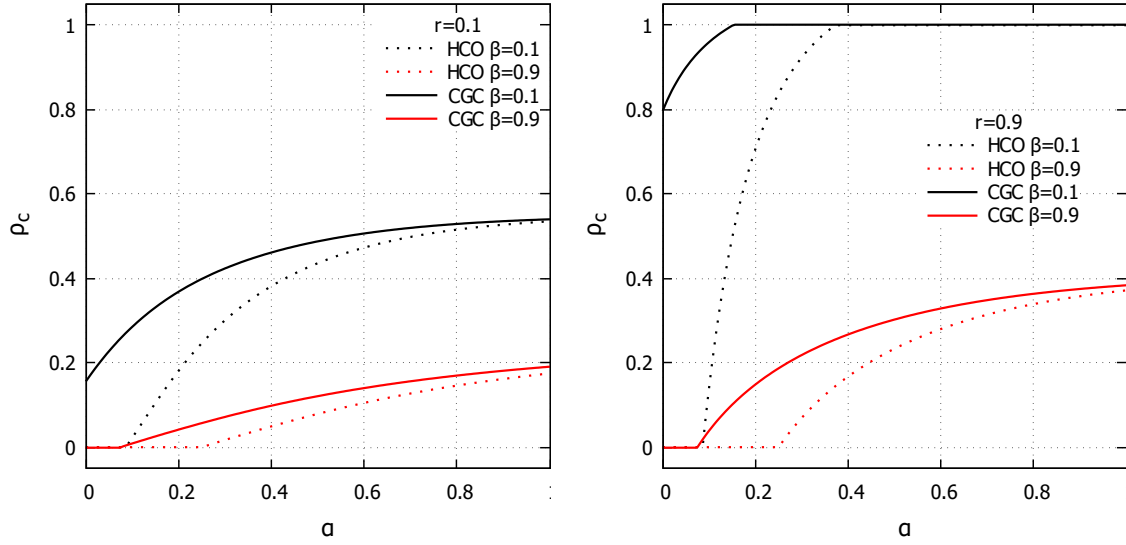


Figura 4: Porcentaje de corruptos en función de la tasa de corrupción $\rho_c(\alpha)$ fijando la tasa de delación a $\beta = 0.1$, $\beta = 0.9$. La tasa de reinsertión r toma los valores $r = 0.1$ en la primera columna y $r = 0.9$ en la segunda columna. Las líneas punteadas muestran las predicciones numéricas del HCO; y las líneas continuas, del campo global de corrupción (CGC). La expresión del (CGC) en la primera fila es $g(\rho_c) = 0.1\rho_c$; y en la segunda, $g(\rho_c) = 0.7\rho_c$.

La primera diferencia clara que se observa en la figura 4 la transición “Full H”-*mixed* tiene lugar para valores de α_c más bajos conforme aumenta el valor de g_1 . En la tabla 1 mostramos los valores que predice (25) en función los distintos conjuntos de parámetros de la gráfica 4 y hay un acuerdo total con los resultados obtenidos numéricamente.

CGC	$g(\rho_c) = 0$	$g(\rho_c) = 0.1\rho_c$	$g(\rho_c) = 0.7\rho_c$
$\beta = 0.1$	0.086	0.061	0.09
$\beta = 0.9$	0.25	0.22	0.075

Cuadro 1: Valores de α_c en función de β y g_1

También se observa que los valores numéricos no varían con r en esta transición, justo como predice (25). Las mayores diferencias entre las predicciones originales y las nuevas tienen lugar en el entorno de la región de transición “Full H”-*mixed*, donde la corrupción se ve claramente favorecida. Ya para el caso de $g(\rho_c) = 0.1\rho_c$ se observa que $\alpha_c(\text{CGC}) < \alpha_c(\text{HCO})$, y la curva $\rho_c(\alpha)$ para CGC queda por encima de la original; es decir, añadir el campo global de corrupción tiene el efecto esperado. Para valores altos de α , sin embargo, las curvas muestran resultados más similares, indicando que predomina la interacción agente-agente sobre el CGC.

La otra transición que se observa es *mixed-C*, que tiene lugar en las gráficas con $r = 0.9$ para $\beta = 0.1$. La ecuación (29) indica que esta transición tiene lugar en $\alpha_c = 0.36$ para $g(\rho_c) = 0.1\rho_c$; y en $\alpha_c = 0.15$ para $g(\rho_c) = 0.7\rho_c$. Y los resultados numéricos coinciden con estos valores. Sin embargo, esta última transición depende tanto de r como de g_1 , por lo que, para estudiar el efecto que tiene cada parámetro en la transición, en la figura 5 mostramos gráficas $\rho_c(\beta)$ para distintos valores de α y r .

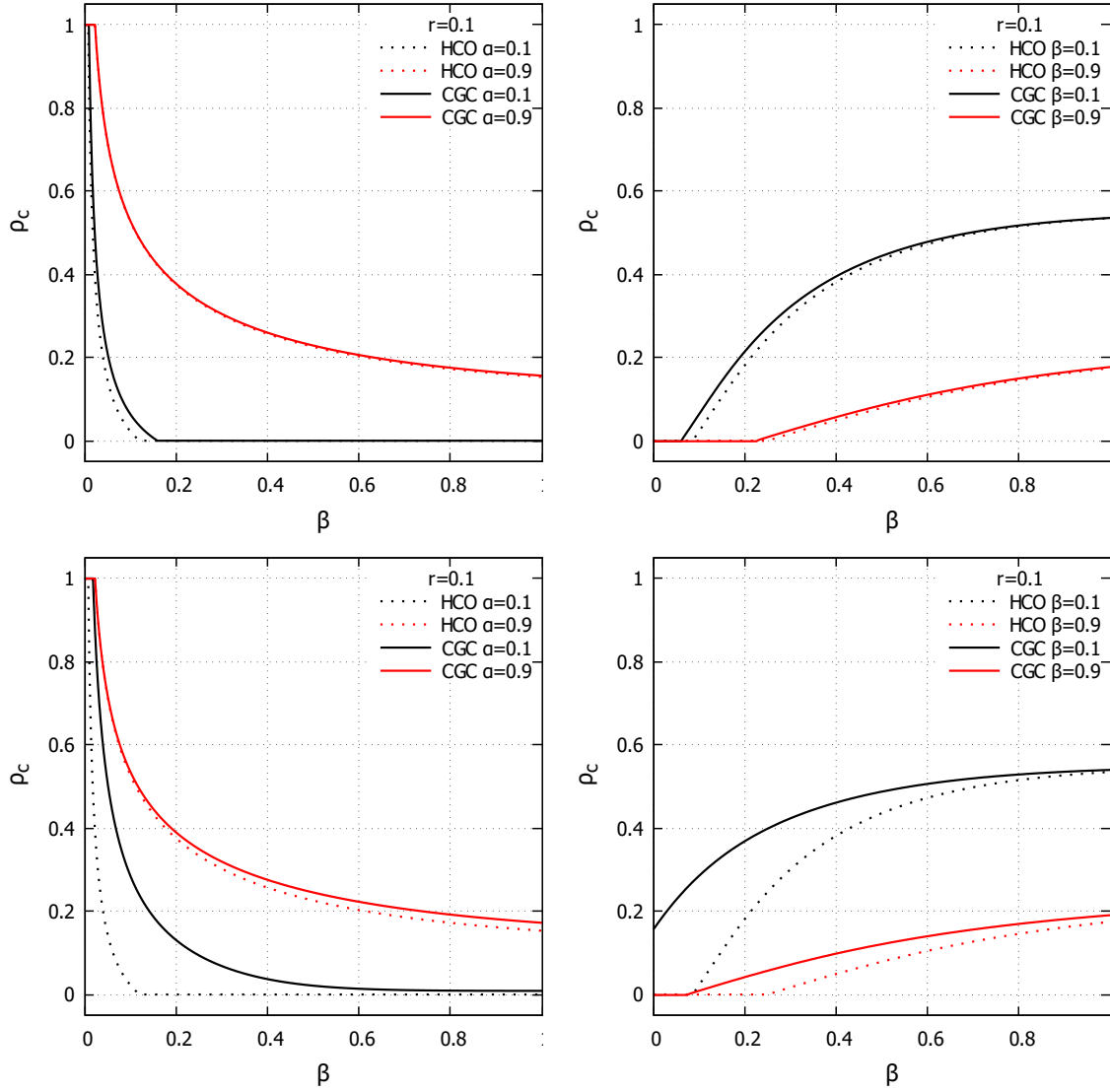


Figura 5: Porcentaje de corruptos en función de la tasa de delación $\rho_c(\beta)$ fijando la tasa de corrupción a $\alpha = 0.1$, $\alpha = 0.9$. La tasa de reinsertión r toma los valores $r = 0.1$ en la primera columna y $r = 0.9$ en la segunda columna. La expresión del (CGC) en la primera fila es $g(\rho_c) = 0.1\rho_c$; y en la segunda, $g(\rho_c) = 0.7\rho_c$.

De nuevo, la curva $\rho_c(\beta)$ CGC queda por encima de la HCO por el mismo motivo. Sin embargo, r sí juega un papel importante en la transición, ya que es capaz de limitar el valor de β_c de una forma mucho más acusada de lo que lo hacen α o g_1 , tal y como se desprende de (26). Además, en estas gráficas el valor de α resulta determinante a la hora de regular el efecto del campo global sobre el modelo HCO, ya que para $\alpha = 0.9$ la diferencia entre las predicciones originales y las nuevas es mínima, mientras que para $\alpha = 0.1$ existe una mayor disparidad.

Por último, se realiza una resolución numérica con una función lineal cuyo término independiente es no nulo. De esta forma, se observa la pérdida del macroestado “Full H” como solución asintótica, quedando como posibles puntos fijos “Full C” y un estado mixto. Como este tipo de campos distorsiona notablemente las predicciones originales, el valor máximo que pueda alcanzar y el término independiente escogido son bajos.

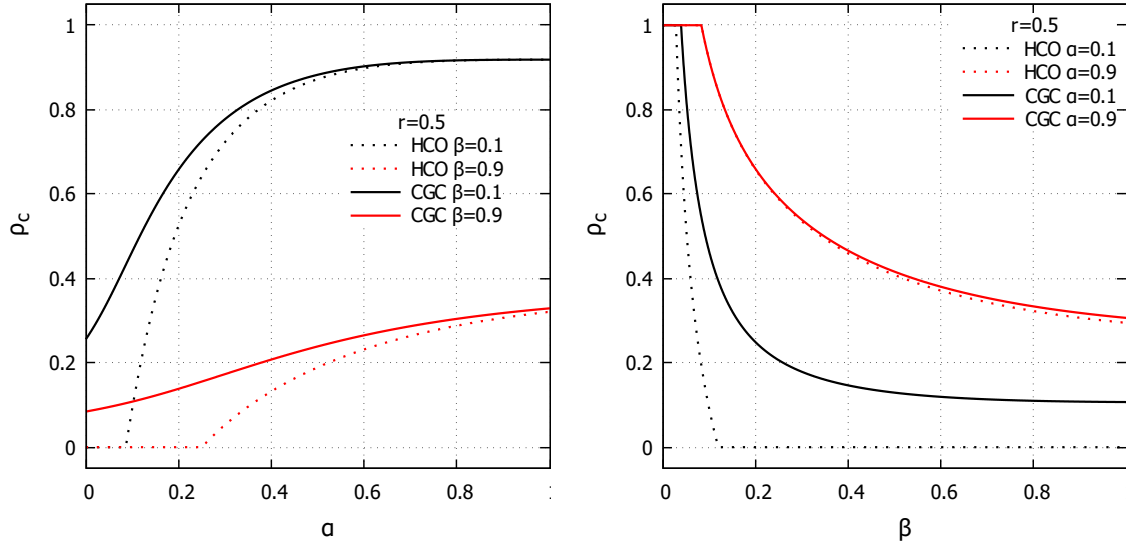


Figura 6: Porcentaje de corruptos en función de la tasa de corrupción $\rho_c(\alpha)$ fijando la tasa de corrupción a $\beta = 0.1$, $\beta = 0.9$. (izquierda) y porcentaje de corruptos en función de la tasa de delación $\rho_c(\beta)$ fijando la tasa de corrupción a $\alpha = 0.1$, $\alpha = 0.9$. (derecha). En ambas situaciones, $r = 0.5$ y la expresión del CGC es $g(\rho_c) = 0.1 + 0.1\rho_c$

3.4. Simulaciones de agentes

El estudio de campo medio realizado hasta ahora ha revelado cómo la inclusión del campo global de corrupción afecta a los resultados de [1] (en los anexos A, B y C se realiza el estudio para el campo global de honestidad). En concreto, modifica los valores críticos de α y β para los que tienen lugar las transiciones “Full H”-*mixed* cuando $g(0) = 0$ (ya que si $g(0) > 0$ “Full H” no es punto fijo), y “Full C”-*mixed*. La otra consecuencia de incluirlo es un aumento en el porcentaje de corruptos respecto al modelo *HCO* para los estados mixtos.

En este apartado pretendemos validar estos resultados realizando simulaciones de Monte Carlo (MC) sobre poblaciones en una red. Como se explica en el apartado anterior, se emplean dos tipos de redes homogéneas: *Lattice* y *RRN* con $k = 4$ vecinos por nodo, ambas con un tamaño de $N = 10000$ nodos para evitar los efectos de tamaño finito. El sistema empieza con un 10% de corruptos elegidos al azar y se deja evolucionar hasta llegar al estado estacionario. Cada simulación se repite 100 veces y se promedian los resultados para eliminar el ruido que las fluctuaciones del sistema puedan introducir.

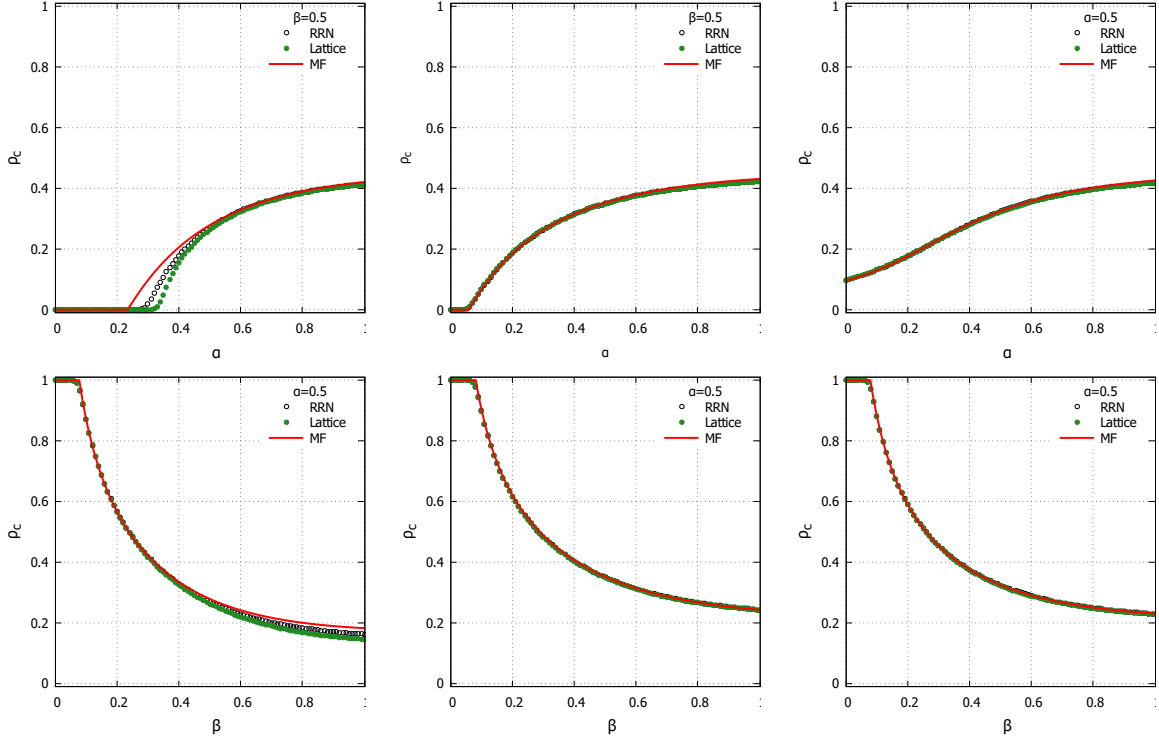


Figura 7: Fracción de corruptos ρ_c en función de α y β . Las curvas rojas muestran la aproximación de campo medio, y los puntos las simulaciones MC. Las columnas muestran varios valores del CGC siendo nulo para la primera, $g(\rho_c) = 0.7\rho_c$ para la segunda, $g(\rho_c) = 0.1 + 0.1\rho_c$ para la tercera. En todos los casos, $r = 0.5$.

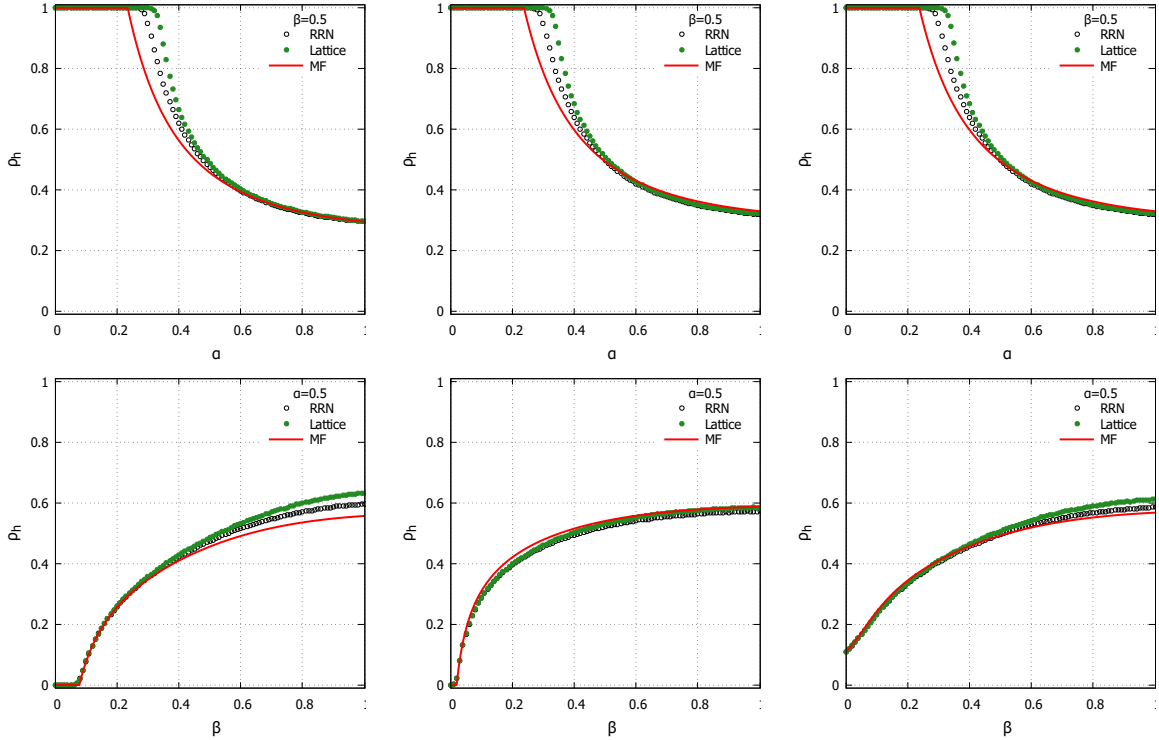


Figura 8: Fracción de honestos ρ_h en función de α y β . Las curvas rojas muestran la aproximación de campo medio, y los puntos las simulaciones MC. Las columnas muestran varios valores del CGH siendo nulo para la primera, $h(\rho_h) = 0.7\rho_h$ para la segunda, $h(\rho_h) = 0.1 + 0.1\rho_h$ para la tercera. En todos los casos, $r = 0.5$.

En las figuras 7 y 8 se compara los resultados de campo medio con los de las simulaciones de Monte Carlo. A partir de ellas, confirmamos las predicciones realizadas a lo largo de esta sección mediante la aproximación de campo medio. Además, se comparan los resultados del modelo HCO en ausencia (primera columna) y en presencia (segunda y tercera columna) de campos globales. Esta comparación arroja una serie de resultados interesantes. Para empezar, los campos globales tienden a reducir las diferencias entre los resultados de campo medio y de Monte Carlo. Sin embargo, es el campo global de corrupción (figura 7) el que lo consigue de forma más clara. En concreto, consigue que el valor de α_c predicho se cumpla en la transición “Full H”-*mixed*, contrarrestando las posibles correlaciones espaciales que introducen las topologías de las redes empleadas y que causaban estas diferencias.

4. Biestabilidad inducida por campos globales

Hasta ahora hemos estudiado funciones lineales con y sin término independiente y el efecto que tiene cada término en la dinámica del sistema y el diagrama de fases. Sin embargo, podemos preguntarnos qué efecto tendrán en la dinámica del sistema y especialmente en la posible biestabilidad de los puntos “Full H” y “Full C”, expresiones más complejas de las funciones de los flujos globales.

Recordando lo explicado en la sección 3.2, sabemos que existirá biestabilidad entre “Full H” y “Full C” si la región del plano representada por la ecuación (25) tiene zonas comunes con la región de la ecuación (26). Esto ocurre si se cumple

$$\alpha_c(\beta) \geq \bar{\alpha}(\beta) \quad (32)$$

para algún valor de los parámetros. Es decir, para un valor de β dado, el máximo valor de α para el que la región “Full H” es estable debe ser superior al mínimo valor de α para el que la región “Full C” es estable. Este último lo obtenemos al invertir la ecuación (26). Podemos ahora sustituir las expresiones y resolver la desigualdad:

$$\frac{1 - (1 - \beta)^k - g'(0)}{k} \geq 1 - \left(\frac{1 - \beta k \left(\frac{r+1}{r} \right)}{1 - g(1)} \right)^{1/k} \Rightarrow$$

Despejando $g'(0)$:

$$g'(0) \leq 1 - \left[1 - \left(\frac{1 - \beta k \left(\frac{r+1}{r} \right)}{1 - g(1)} \right)^{1/k} \right] k - (1 - \beta)^k \quad (33)$$

A partir de (29) encontramos que la acotación $0 \leq \bar{\alpha}(\beta) \leq 1$ equivale a

$$g(1) \leq \beta k \frac{r+1}{r} \quad \text{y} \quad \beta \leq \frac{r}{k(r+1)} \quad (34)$$

En la figura 9 representamos la cota superior de $g'(0)$ dada por (33) en función de $g(1)$ para diversos valores de la tasa de delación β , en los rangos de valores de ambos dados por (34).

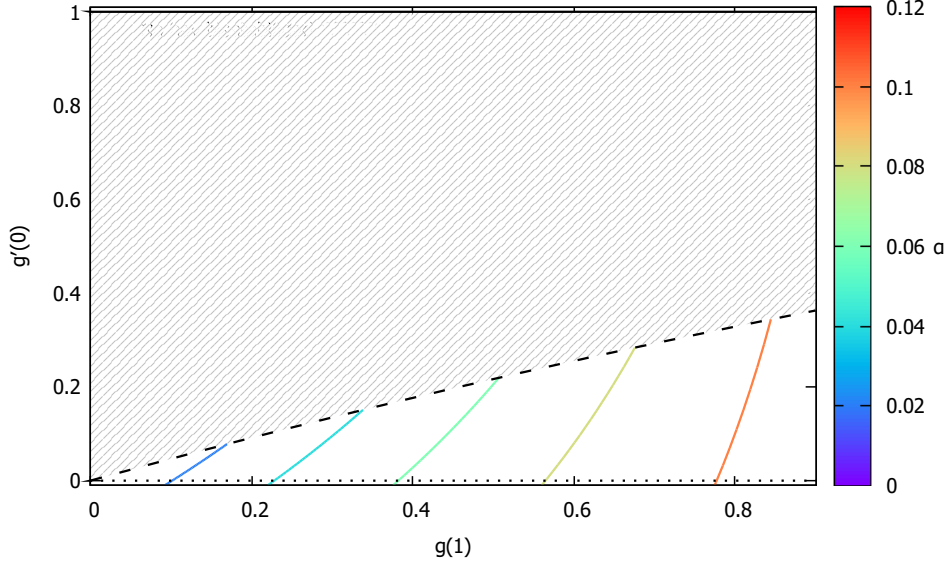


Figura 9: Cota superior de $g'(0)$ para distintos valores de β , indicados mediante código de color. La región rayada indica que para ese conjunto de pares $(g(1), g'(0))$ nunca se solapan las regiones “Full H” y “Full C”. La línea discontinua representa los puntos sobre las curvas donde $g(1)$ alcanza su cota superior.

La figura 9 muestra para cada valor de β la región de valores $(g(1), g'(0))$ para los que se puede dar solapamiento entre regiones. El efecto que tiene aumentar β es disminuir el número de valores de $g(1)$ y $g'(0)$ para los que se da el solapamiento. Vistas estas relaciones, pasamos a concretar la forma de las funciones que elegimos para poder realizar las simulaciones que se mostrarán más adelante. En la sección 2 se explica que elegimos funciones desarrollables en series de Taylor; ya hemos estudiado qué efectos tienen el término independiente y el término lineal, por lo que proseguimos estudiando funciones cuadráticas:

$$g(\rho_c) = g_1\rho_c + g_2\rho_c^2 \quad (35)$$

que cumplen para ciertos valores de g_1 y g_2 las condiciones necesarias para que se de la biestabilidad entre “Full H” y “Full C”. Esto se puede observar gráficamente en la figura 10.

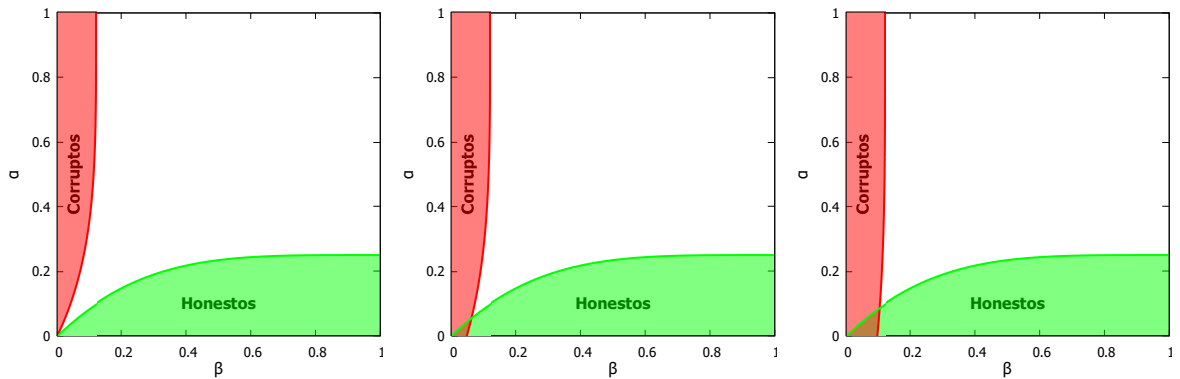


Figura 10: Diagramas de fases para $g_2 = 0, 0.4, 0.8$. En los tres casos se ha escogido $r = 0.9$. $g'(0) = 0$ para todos los casos, por lo que la región “Full H” dada por (25) es la misma en los tres casos. La región “Full C” descrita por (26) sí que cambia, ya que depende de $g(1) = g_2$.

Demostrado que puede darse la coexistencia entre los macroestados “Full H” y “Full C”, queda comprobar si es posible que haya coexistencia entre dos soluciones mixtas, o entre “Full H” o “Full C” y una solución mixta. Para ello, se estudia gráficamente cómo son las nulclinas cuando los parámetros α y β toman valores próximos a aquellos para los que se tiene la coexistencia de todos honestos con todos corruptos.

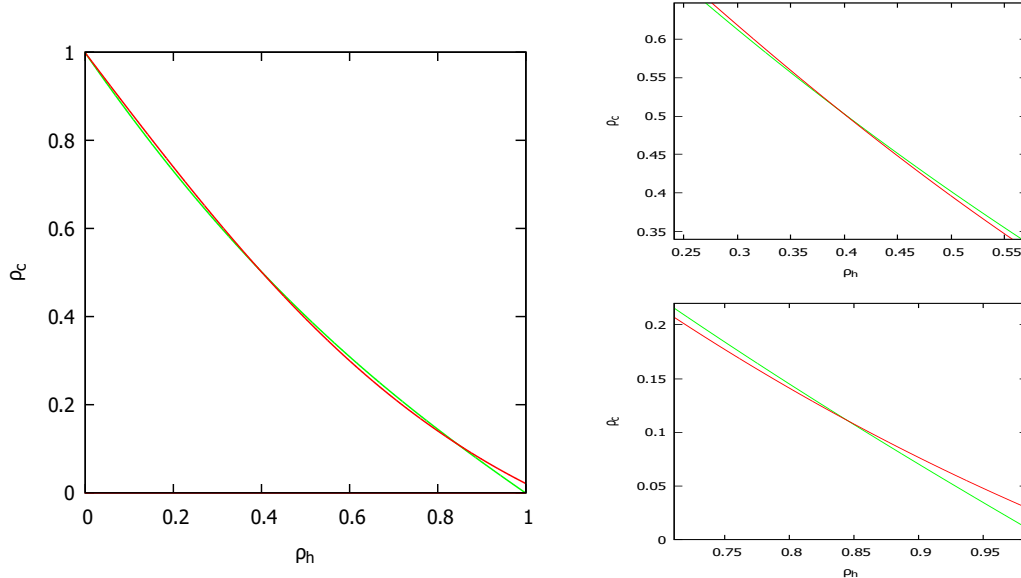


Figura 11: Nulclinas de F_h (curva verde) y F_c (curva roja) para $\alpha = 0.1$, $\beta = 0.13$ y $r = 0.9$. La expresión del campo de corrupción es $g(\rho_c) = 0.75\rho_c^2$. La gráfica de la izquierda muestra las nulclinas en el simplex entero. La primera gráfica de la derecha muestra la intersección entre nulclinas que da lugar al punto fijo estable, y la segunda gráfica, la intersección que da lugar al punto fijo inestable.

Las nulclinas de la figura 11 muestran cuatro cortes: $(1,0)$, $(0,1)$ y dos puntos internos: una de las ramas de la nulclina de F_c (curva roja) tiene una curvatura que le hace cortar dos veces sobre la nulclina F_h (curva verde). Así, se forman los dos puntos internos que se aprecian mejor en las gráficas de la derecha. Por cómo se cruzan las nulclinas (apartado 3.1), $(1,0)$ y $(0.4,0.5)$ serán estables, y $(0,1)$ y $(0.85,0.11)$ inestables. Ha quedado demostrado que sí es posible encontrar valores de los parámetros y de la expresión del CGC para los que hay biestabilidad entre “Full H” y estados *mixed*. Por tanto, queda ver qué conjunto de valores de α y β conforman la región de biestabilidad y qué curvas lo delimitan. Dos de ellas vienen dadas por (25) y (26). La tercera marca el mínimo valor de α para el que se da la coexistencia entre los estados macroscópicos de “todos honestos” y “soluciones mixtas” y se obtiene de forma numérica con la aproximación de campo medio. El procedimiento para encontrar el estado estacionario *mixed* es el siguiente:

- Fijamos un valor de β .
- Calculamos el valor de α_c y elegimos uno ligeramente inferior.
- Como condiciones iniciales escogemos una población con un porcentaje muy alto de corruptos (del 90%).
- Disminuimos α y vemos cómo cambia este punto fijo interno. Guardaremos el menor valor de α para el cual la solución del sistema no es $(1,0)$ y se disminuye β .
- Si ya en la primera ejecución la solución del sistema es $(1,0)$ se disminuye β .

Así, se recorre el sistema a lo largo de todo el dominio de β para obtener la curva. A continuación, se muestra el resultado final.

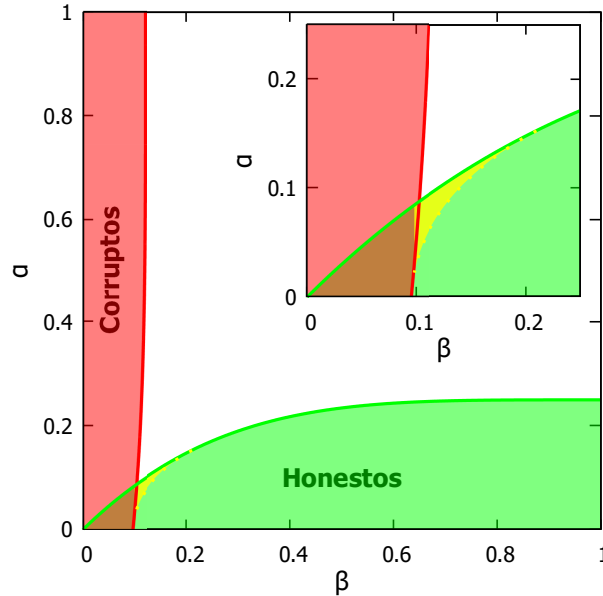


Figura 12: Diagrama de fases para $r = 0.9$ y $g(\rho_c) = 0.8\rho_c^2$. Se pueden observar 5 regiones diferentes: la región “Full H” en color verde, la región “Full C” en color rojo, la región *mixed* en blanco, la región de coexistencia “Full H”-“Full C” y la región de coexistencia “Full H”-*mixed* en color amarillo.

La línea punteada marca el límite inferior de la región para la que se tienen dos cortes internos de las nulclinas en el simplex (y se tiene coexistencia entre “Full H” y *mixed*). En esta región, la transición entre la situación de todos honestos y una situación mixta será abrupta. La línea verde, dada por la expresión (25), proporciona los valores de α y β para los que solo hay un corte interno en el simplex (“Full H” deja de ser estable). El punto en el que la línea discontinua y la línea verde se unen en el plano (β, α) indica que las nulclinas de F_h y F_c se cortan exactamente en $(1,0)$ y a partir de estos valores la transición “Full H”-*mixed* será suave.

4.1. Estados estacionarios de campo medio

De la misma forma que en el apartado 3.3, vamos a analizar los resultados de la iteración de las ecuaciones (15) centrándonos en los cambios cualitativos que provocan los polinomios de orden 2.

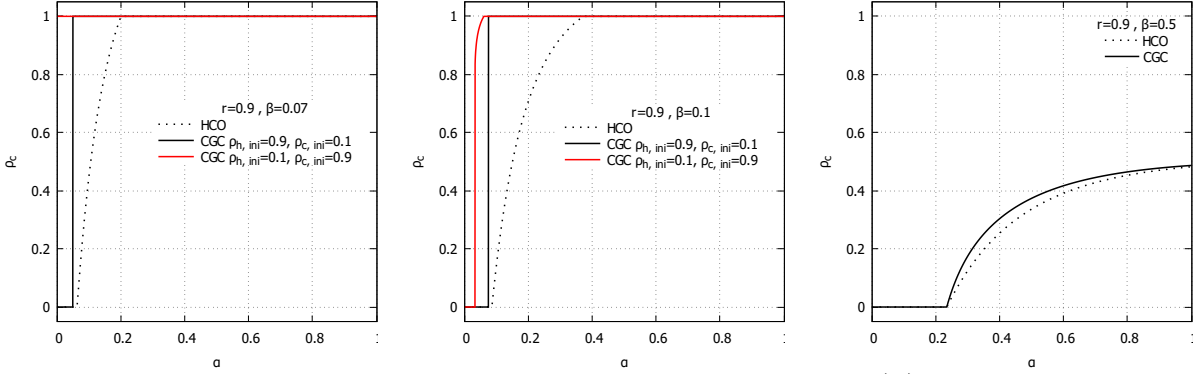


Figura 13: Fracción de corruptos en función de la tasa de corrupción $\rho_c(\alpha)$ para $\beta = 0.07$, $\beta = 0.1$, $\beta = 0.5$ y $r = 0.9$. La expresión de CGC en ambos casos es $g(\rho_c) = 0.8\rho_c^2$. La curva discontinua muestra las predicciones de *HCO*, la curva negra continua muestra las predicciones de CGC con condiciones iniciales $(\rho_{h,ini}, \rho_{c,ini}) = (0.9, 0.1)$ y la curva roja muestra las predicciones de CGC con condiciones iniciales $(\rho_{h,ini}, \rho_{c,ini}) = (0.1, 0.9)$.

La figura 13 muestra cómo evoluciona la fracción de corruptos dependiendo de si nos encontramos en la región de biestabilidad ($\beta = 0.07, 0.1$), o fuera de ella ($\beta = 0.5$). Para el primer caso hay dos soluciones estables. Así, las condiciones iniciales determinan cuál será el estado estacionario. Para $\beta = 0.07$ el sistema se encuentra (para valores de α bajos) en la región de biestabilidad “Full H”-“Full C”; de ahí que para las condiciones iniciales $(\rho_{h,ini}, \rho_{c,ini}) = (0.1, 0.9)$ la solución estable sea siempre “Full C”. Para $\beta = 0.1$, el sistema se encuentra en la región de biestabilidad “Full H”-*mixed*. Por ello, antes de desestabilizarse “Full H” aparece un estado estacionario *mixed*, coincidiendo con el diagrama de fases de la figura 12. En ambos casos, las condiciones iniciales $(\rho_{h,ini}, \rho_{c,ini})$, determinan el valor de α para el que se produce la transición (de forma abrupta) desde “Full H” al otro punto fijo estable debido a que aumentar α , dado un valor de β fijo, disminuye la cuenca de “Full H” y aumenta la cuenca del otro punto fijo estable. Así, cuanto más cerca estén las condiciones iniciales de $(1, 0)$, mayor será el valor de α para el que se da la transición abrupta.

Al alejarnos de la región de biestabilidad ($\beta = 0.5$) se observa que la transición “Full H”-*mixed*, tanto para *HCO* como cuando se incluye CGC, ocurre para el mismo valor de $\alpha_c = 0.235$ que coincide con el predicho por (25). Aparte, la curva descrita por CGC ya no depende de las condiciones iniciales, debido a que para estos valores de los parámetros solo hay un punto fijo estable.

5. Conclusión

Con el objeto de caracterizar, en el marco formal del modelo HCO de comportamientos sociales corruptos, la influencia del clima social, hemos extendido el modelo HCO original mediante la inclusión de “campos globales” de corrupción y honestidad dados por funciones analíticas de las fracciones de individuos corruptos y honestos, respectivamente.

La caracterización se realiza a través de la aproximación de campo medio, con la que se describe el comportamiento del “individuo promedio”. Los resultados obtenidos quedan avalados por la resolución numérica de las ecuaciones de Markov para redes homogéneas dada su total concordancia, y se emplean para comparar los resultados el modelo original

con los de las modificaciones propuesta. Por último, se validan mediante simulaciones de Monte Carlo, con las que se aprecia un enorme acuerdo cualitativo.

Si el desarrollo de Taylor de la función representativa del campo global de corrupción (resp. honestidad) posee término independiente no nulo (tendencia innata a la corrupción (resp. honestidad)) los únicos macroestados asintóticos posibles son el de corrupción total (resp. honestidad total) y estados de composición mixta.

En ausencia de término independiente y función representativa lineal, el diagrama de fases (macroestados asintóticos) en el espacio de parámetros muestra solo tres regiones de comportamiento social: honestidad total, corrupción total y macroestados mixtos, con líneas críticas que los separan. En esta situación, dado un conjunto de valores de los parámetros solo existe un equilibrio estable posible para el sistema.

Sin embargo, la presencia de términos no lineales en el desarrollo de las funciones representativas de los campos globales modifican sustancialmente el diagrama de fases del modelo, que en este caso presenta regiones de coexistencia entre estados de honestidad y corrupción totales y de estos con estados de composición mixta. Ello da lugar a comportamientos de histéresis frente a a variación lenta de los parámetros y transiciones discontinuas entre equilibrios sociales estables diferentes (crisis sociales).

De cara al futuro, se abren nuevas vías de estudio. Por una parte, se puede continuar con el estudio de funciones representativas más complejas de los “campos globales” cuya expresión trate de plasmar el clima social, siendo de potencial interés funciones que saturan. Por otra parte, se puede evaluar el modelo sobre redes heterogéneas: redes en las que el número de contactos por nodo sigue una ley de potencias y que replican de forma más precisa la estructura de las interacciones sociales. El resultado final será una descripción más fiel de los comportamientos sociales corruptos.

Bibliografía

- [1] Dan Lu, F. Bauza, D. Soriano-Paños, J. Gómez-Gardeñes, and L. M. Floría. Norm violation versus punishment risk in a social model of corruption. *Phys. Rev. E*, 101:022306, Feb 2020.
- [2] Prateek Verma and Supratim Sengupta. Bribe and punishment: An evolutionary game-theoretic analysis of bribery. *PloS one*, 10(7):e0133441, 2015.
- [3] Joung-Hun Lee, Karl Sigmund, Ulf Dieckmann, and Yoh Iwasa. Games of corruption: How to suppress illegal logging. *Journal of Theoretical Biology*, 367:1–13, 2015.
- [4] Von-Newman John and Oskar Morgenstern. Theory of games and economic behavior. *The Press of Princeton University, Princeton, NJ*, 1944.
- [5] Nino Boccara. *Modeling complex systems*. Springer Science & Business Media, 2010.
- [6] Rüdiger Seydel. *Practical bifurcation and stability analysis*, volume 5. Springer Science & Business Media, 2009.

A. Anexo I: Puntos fijos y estabilidad en diferentes situaciones

En el texto principal se ha estudiado la condición de punto fijo para el campo global de corruptos incluyendo el término *warning to wrongdoers*. Sin embargo, durante el proyecto se estudian otras situaciones cuyo análisis es similar (en cuanto a forma) al del texto principal.

A.1. Campo global de corruptos sin *wtw*

La matriz se simplifica al eliminar el término ρ_o , quedando

$$\mathbf{Q} = \begin{pmatrix} 1 - f_\alpha - g(\rho_c)(1 - f_\alpha) & 0 & r \\ f_\alpha + g(\rho_c)(1 - f_\alpha) & 1 - f_\beta & 0 \\ 0 & f_\beta & 1 - r \end{pmatrix} \quad (\text{A.1})$$

Las ecuaciones de la dinámica vienen dadas por:

$$\begin{cases} F_h = -[f_\alpha + g(\rho_c)(1 - f_\alpha) + r]\rho_h - r\rho_c + r \\ F_c = [f_\alpha + g(\rho_c)(1 - f_\alpha)]\rho_h - f_\beta\rho_c \end{cases} \quad (\text{A.2})$$

Es inmediato comprobar que el estado “Full C” es siempre solución del sistema. Pero no es el caso de “Full H”:

$$(\mathbf{1}, \mathbf{0}) : \begin{cases} F_h = -g(0) \\ F_c = g(0) \end{cases}$$

Además, para la nulclina $F_c = 0$ en este caso se pueden obtener dos ramas de forma analítica. Esto se hace considerando los casos $\rho_h = 0$ y $\rho_c = 0$:

$$\begin{cases} F_c(\rho_h, 0) = -g(0)\rho_h \\ F_c(0, \rho_c) = 0 \quad \forall \rho_c \end{cases}$$

Entonces, cuando $g(0) = 0$ obtenemos los mismos resultados que los obtenidos en el texto principal: “Full H” será solución del sistema y se habilita una recta horizontal ($\rho_c = 0$) como rama de la nulclina $F_c = 0$. Sin embargo, en este caso obtenemos una rama que no está al incluir el *wtw*. Esta es una recta vertical $\rho_h = 0$.

Ahora vamos a estudiar la estabilidad de los puntos de equilibrio. Para ello, calculamos el jacobiano:

$$\mathbf{J} = \begin{pmatrix} \frac{\partial F_h}{\partial \rho_h} & \frac{\partial F_c}{\partial \rho_h} \\ \frac{\partial F_h}{\partial \rho_c} & \frac{\partial F_c}{\partial \rho_c} \end{pmatrix} \quad (\text{A.3})$$

cuyas derivadas parciales son

$$\begin{cases} \frac{\partial F_h}{\partial \rho_h} = -(f_\alpha + g(\rho_c)(1 - f_\alpha) + r) \\ \frac{\partial F_h}{\partial \rho_c} = -r - \rho_h(f'_\alpha(1 - g(\rho_c)) + g'(\rho_c)(1 - f_\alpha)) \\ \frac{\partial F_c}{\partial \rho_h} = f_\alpha + g(\rho_c)(1 - f_\alpha) - f'_\beta \rho_c \\ \frac{\partial F_c}{\partial \rho_c} = -f_\beta + \rho_h(f'_\alpha(1 - g(\rho_c)) + g'(\rho_c)(1 - f_\alpha)) \end{cases} \quad (\text{A.4})$$

Estudiaremos los puntos de equilibrio $(0, 1)$ y $(1, 0)$ (cuando $g(0) = 0$).

• $(1, 0)$

$$\mathbf{J}_h = \begin{pmatrix} -r & -r - (f'_\alpha(0) + g'(0)) \\ 0 & -f_\beta + (f'_\alpha(0) + g'(0)) \end{pmatrix}$$

$$\begin{cases} \lambda_1 = -r < 0 \\ \lambda_2 = -f_\beta(1) + f'_\alpha(0) + g'(0) \end{cases} \quad (\text{A.5})$$

Cuando $g'(0) < f_\beta(1) - f'_\alpha(0)$ el punto $(1, 0)$ será estable.

$$\alpha \leq \alpha_c = \frac{1 - (1 - \beta)^k - g'(0)}{k} \quad (\text{A.6})$$

• $(0, 1)$

$$\mathbf{J}_c = \begin{pmatrix} -f_\alpha(1) - g(1)(1 - f_\alpha(1)) - r & -r \\ f_\alpha(1) + g(1)(1 - f_\alpha(1)) + f'_\beta(0) & 0 \end{pmatrix}$$

$$\begin{cases} T = -(f_\alpha(1) + g(1)(1 - f_\alpha(1)) + r) < 0 \text{ siempre.} \\ D = r[f_\alpha(1) + g(1)(1 - f_\alpha(1)) - f'_\beta(0)] \end{cases} \quad (\text{A.7})$$

Desarrollando más la ecuación para el determinante, vemos que la condición de estabilidad para la región “Full C” es

$$f_\alpha(1) + g(1)(1 - f_\alpha(1)) - f'_\beta \geq 0 \Rightarrow \beta \leq \beta_c = \frac{1 - (1 - \alpha)^k(1 - g(1))}{k} \quad (\text{A.8})$$

Podemos comparar estos resultados con los obtenidos en el texto principal. Vemos que el wtw no interviene en la estabilidad de “Full H”, ya que las expresiones obtenidas son idénticas. Sin embargo, juega un papel importante en la estabilidad de “Full C”. En concreto, se ve que

$$\beta_{crit}^{con} = \left(\frac{r}{r + 1} \right) \beta_{crit}^{sin} \quad (\text{A.9})$$

Recordemos que el wtw ofrece una posibilidad a los corruptos de convertirse en honestos por miedo a las represalias. Estas represalias se reflejan en el parámetro r . Cuanto mayor sea el castigo menor será r (mayor será el tiempo que los agentes permanecen relegados en el ostracismo) y viceversa. Así, es de esperar que al habilitar esta “fuga” de corruptos se reducen las situaciones para las cuales “Full C” es estable. Y esto es lo que se percibe

al comparar las ecuaciones para β_{crit} con y sin wtw .

A.2. Campo global de honestos sin wtw

Partimos de la matriz de la ecuación 12, y eliminamos los términos asociados al wtw y al campo global de corruptos. La matriz final queda

$$Q = \begin{pmatrix} 1 - f_\alpha & h(\rho_h) & r \\ f_\alpha & (1 - h(\rho_h))(1 - f_\beta) & 0 \\ 0 & (1 - h(\rho_h))f_\beta & 1 - r \end{pmatrix} \quad (A.10)$$

Así pues, las correspondientes ecuaciones que describen el flujo vienen dadas por:

$$\begin{cases} F_h = -(f_\alpha + r)\rho_h + (h(\rho_h) - r)\rho_c + r \\ F_c = f_\alpha\rho_h + [-h(\rho_h) - f_\beta + h(\rho_h)f_\beta]\rho_c \end{cases} \quad (A.11)$$

Volvemos a analizar cada una de las ecuaciones del flujo. Empezamos con F_c . Vemos que para $\rho_c = 0 \rightarrow f_\alpha = 0$ y por tanto, $F_c = 0 \forall \rho_h$. Es decir, hemos encontrado una rama de la nulclina. A diferencia de lo que sucede en el campo de corruptos, esta rama está presente para cualquier forma de $h(\rho_h)$. Con esta situación, vemos cómo queda la ecuación F_h :

$$F_h = -r\rho_h + r$$

Así pues, cuando $\rho_h = 1 \rightarrow F_h = 0$ y por tanto, **(1, 0) es un punto fijo, independientemente del valor que tome el flujo $h(\rho_h)$.**

Ahora, vemos si el otro punto fijo del modelo inicial (0, 1) sigue siendo un punto fijo o no:

$$\begin{cases} F_h = h(0) \\ F_c = -h(0) \end{cases} \quad (A.12)$$

Por tanto, si $h(0) = 0$, tendremos que (0, 1) será un punto fijo. Y además, esto habilitará la rama $\rho_h = 0$ de la nulclina $F_c = 0$; la rama $\rho_c = 0$ de esta nulclina no es afectada por el campo de honestos. Vemos que otra vez el comportamiento es el mismo que en el caso anterior, pero en este caso afecta a si “Full C” es solución del sistema. Es decir, tenemos un resultado análogo al obtenido para “Full H” y el campo de corruptos.

Por último, vemos cómo se comportan los puntos fijos. Para ello, calculamos el jacobiano:

$$\begin{cases} \frac{\partial F_h}{\partial \rho_h} = -(f_\alpha + r) + h'(\rho_h)\rho_c \\ \frac{\partial F_h}{\partial \rho_c} = -f_\alpha\rho_h + (h(\rho_h) - r) \\ \frac{\partial F_c}{\partial \rho_h} = f_\alpha + [f'_\beta(h(\rho_h) - 1) + h'(\rho_h)(f_\beta - 1)]\rho_c \\ \frac{\partial F_c}{\partial \rho_c} = f'_\alpha\rho_h + h(\rho_h)f_\beta - f_\beta - h(\rho_h) \end{cases} \quad (A.13)$$

•(1, 0)

$$\mathbf{J}_h = \begin{pmatrix} -r & -f'_\alpha(0) + h(1) - r \\ 0 & f'_\alpha(0) + h(1)f_\beta(1) - h(1) - f_\beta(1) \end{pmatrix}$$

Los dos autovalores son:

$$\begin{cases} \lambda_1 = -r \\ \lambda_2 = f'_\alpha(0) + h(1)(f_\beta(1) - 1) - f_\beta(1) \end{cases} \quad (\text{A.14})$$

Entonces, para que el punto fijo sea estable, las dos condiciones que se deben cumplir son:

$$\begin{cases} r > 0 \text{ siempre} \\ h(1) > \frac{f'_\alpha(0) - f_\beta(1)}{1 - f_\beta(1)} \end{cases} \quad (\text{A.15})$$

•(0, 1)

Recordamos que en este caso se tiene que cumplir $h(0) = 0$ para que sea un punto fijo.

$$\mathbf{J}_c = \begin{pmatrix} -r - f_\alpha(1) & -r \\ f_\alpha(1) - f'_\beta(0) - h'(0) & 0 \end{pmatrix}$$

$$\begin{cases} T = -f_\alpha(1) - r < 0 \text{ siempre} \\ D = r(f_\alpha(1) - f'_\beta(0) - h'(0)) \end{cases} \quad (\text{A.16})$$

Y para que sea un punto fijo, $D > 0$. Entonces:

$$h'(0) < f_\alpha(1) - f'_\beta(0)$$

De las condiciones de estabilidad para “Full H” y “Full C” obtenemos las curvas críticas que margan el límite de estas regiones:

$$h(1) > \frac{f'_\alpha(0) - f_\beta(1)}{1 - f_\beta(1)} \Rightarrow \alpha \leq \frac{1 - (1 - \beta)^k(1 - h(1))}{k} \quad (\text{A.17})$$

$$h'(0) > f_\alpha(1) - f'_\beta(0) \Rightarrow \beta \leq \frac{1 - (1 - \alpha)^k - h'(0)}{k} \quad (\text{A.18})$$

A.3. Campo global de honestos con wtw

Si ahora recuperamos los términos que dependen del wtw, la matriz de la dinámica queda:

$$\mathbf{Q} = \begin{pmatrix} 1 - f_\alpha & \rho_o + h(\rho_h)(1 - \rho_o) & r \\ f_\alpha & (1 - \rho_o)(1 - h(\rho_h))(1 - f_\beta) & 0 \\ 0 & (1 - \rho_o)(1 - h(\rho_h))f_\beta & 1 - r \end{pmatrix} \quad (\text{A.19})$$

Y las ecuaciones para los flujos quedan:

$$\begin{cases} F_h = -[f_\alpha + r + \rho_c(1-h(\rho_h))]\rho_h + [1 - \rho_c - r + h(\rho_h)\rho_c]\rho_c + r \\ F_c = [f_\alpha + (1 - f_\beta)(1 - h(\rho_h))\rho_c]\rho_h + [(1 - f_\beta)(1 - h(\rho_h))\rho_c - 1]\rho_c \end{cases} \quad (\text{A.20})$$

Con estas ecuaciones, evaluando en $(1, 0)$ y $(0, 1)$ vemos que el comportamiento es el mismo que en la sección 2.1. Es decir, $(1, 0)$ siempre es un punto fijo y $(0, 1)$ solo lo será cuando $h(0) = 0$.

Sabido esto, calculamos el hamiltoniano:

$$\begin{cases} \frac{\partial F_h}{\partial \rho_h} = -[f_\alpha + r + (1-h(\rho_h))\rho_c] + h'(\rho_h)\rho_c(\rho_c + \rho_h) \\ \frac{\partial F_h}{\partial \rho_c} = -(f'_\alpha + 1-h(\rho_h))\rho_h + 1 + 2\rho_c(h(\rho_h) - 1) - r \\ \frac{\partial F_c}{\partial \rho_h} = f_\alpha + (1 - f_\beta)(1-h(\rho_h))\rho_c + [f'_\beta(h(\rho_h) - 1) \\ \quad + h'(\rho_h)(f_\beta - 1)]\rho_c(\rho_h + \rho_c) \\ \frac{\partial F_c}{\partial \rho_c} = [f'_\alpha + (1 - f_\beta)(1-h(\rho_h))]\rho_h + 2(1 - f_\beta)(1-h(\rho_h))\rho_c - 1 \end{cases} \quad (\text{A.21})$$

• $(1, 0)$:

$$\mathbf{J}_h = \begin{pmatrix} -r & f'_\alpha(0) + h(1) - r \\ 0 & f'_\alpha(0) + (1 - f'_\beta(1))(1 - h(1)) - 1 \end{pmatrix}$$

Los dos autovalores son:

$$\begin{cases} \lambda_1 = -r \\ \lambda_2 = f'_\alpha(0) + h(1)(f'_\beta(1) - 1) - f'_\beta(1) \end{cases} \quad (\text{A.22})$$

Entonces, para que el punto fijo sea estable, las dos condiciones que se deben cumplir son:

$$\begin{cases} r > 0 \text{ siempre} \\ h(1) > \frac{f'_\alpha(0) - f'_\beta(1)}{1 - f'_\beta(1)} \end{cases}$$

• $(0, 1)$ ($h(0) = 0$):

$$\mathbf{J}_c = \begin{pmatrix} -(r + f_\alpha(1) + 1) + h'(0) & -1 - r \\ f_\alpha(1) + 1 - f'_\beta(0) - h'(0) & 1 \end{pmatrix}$$

$$\begin{cases} T = -(r + f_\alpha(1) + 1) + h'(0) + 1 = h'(0) - r - f_\alpha(1) \\ D = (f_\alpha(1) - h'(0)) - f'_\beta(0) \left(1 + \frac{1}{r}\right) \end{cases} \quad (\text{A.23})$$

Aplicando las condiciones que deben cumplir traza y determinante para que sea un punto fijo estable ($T < 0$, $D > 0$):

$$\begin{cases} h'(0) < r + f_\alpha(1) \\ h'(0) + f'_\beta(0) \left(1 + \frac{1}{r}\right) > f_\alpha(1) \end{cases} \quad (\text{A.24})$$

De este último sistema cabe resaltar que es inmediato ver que si se cumple una de las condiciones automáticamente se cumple la otra. Y los límites de las regiones “Full H” y “Full C” resultan ser, respectivamente:

$$\alpha \leq \alpha_c \frac{1 - (1 - \beta)^k (1 - h(1))}{k} \quad (\text{A.25})$$

$$\beta \leq \beta_c \left(\frac{r}{r+1}\right) \frac{1 - (1 - \alpha)^k - h'(0)}{k} \quad (\text{A.26})$$

Una vez analizado el efecto que tiene cada campo en el sistema por separado, los juntamos. Y analizaremos los resultados tanto con los términos correspondientes al wtw como sin ellos.

A.4. Ambos campos sin wtw

$$\mathbf{Q} = \begin{pmatrix} 1 - f_\alpha - g(\rho_c)(1 - f_\alpha) & h(\rho_h) & r \\ f_\alpha + g(\rho_c)(1 - f_\alpha) & (1 - h(\rho_h))(1 - f_\beta) & 0 \\ 0 & (1 - h(\rho_h))f_\beta & 1 - r \end{pmatrix} \quad (\text{A.27})$$

$$\begin{cases} F_h = -(f_\alpha + r + g(\rho_c)(1 - f_\alpha))\rho_h + r + (h(\rho_h) - r)\rho_c \\ F_c = (f_\alpha + g(\rho_c)(1 - f_\alpha))\rho_h + (-f_\beta - h(\rho_h) + f_\beta h(\rho_h))\rho_c \end{cases} \quad (\text{A.28})$$

Estudiamos los puntos (1,0) y (0,1):

$$(\mathbf{1}, \mathbf{0}) : \begin{cases} F_h = -g(0) \\ F_c = g(0) \end{cases} \quad (\mathbf{0}, \mathbf{1}) : \begin{cases} F_h = h(0) \\ F_c = -h(0) \end{cases}$$

y vemos que un campo afecta a “Full H” y el otro a “Full C”. Y lo hacen de la misma forma que al estudiarlos por separado. Además, si $\rho_c = 0$ y $g(\rho_c) = 0$ la rama que hemos visto anteriormente de $F_c = 0$ se mantiene.

El jacobiano para estudiar la estabilidad de los puntos fijos queda:

$$\begin{cases} \frac{\partial F_h}{\partial \rho_h} = -(f_\alpha + r + g(\rho_c)(1 - f_\alpha)) + h'(\rho_h)\rho_c \\ \frac{\partial F_h}{\partial \rho_c} = -[f'_\alpha(1 - g(\rho_c) + g'(\rho_c)(1 - f_\alpha))]\rho_h + h(\rho_h) - r \\ \frac{\partial F_c}{\partial \rho_h} = f_\alpha + g(\rho_c)(1 - f_\alpha) + [f'_\beta(h(\rho_h) - 1) + h'(\rho_h)(f_\beta - 1)]\rho_c \\ \frac{\partial F_c}{\partial \rho_c} = [f'_\alpha(1 - g(\rho_c)) + g'(\rho_c)(1 - f_\alpha)]\rho_h - f_\beta - h(\rho_h) + f_\beta h(\rho_h) \end{cases} \quad (\text{A.29})$$

• $(1, 0)$:

$$\mathbf{J}_h = \begin{pmatrix} -r & -r - (f'_\alpha(0) + g'(0)) + h(1) \\ 0 & -f_\beta(1) + f'_\alpha(0) + g'(0) - h(1) + h(1)f_\beta(1) \end{pmatrix}$$

Cuyos autovalores son:

$$\begin{cases} \lambda_1 = -r \\ \lambda_2 = -f_\beta(1) + f'_\alpha(0) + g'(0) - h(1) + h(1)f_\beta(1) \end{cases} \quad (\text{A.30})$$

Como en los casos anteriores, ambos deben ser negativos para que el punto sea fijo. Entonces, λ_1 siempre cumple es menor que 0, pero λ_2 no.

$$\begin{aligned} -f_\beta(1) + f'_\alpha(0) + g'(0) - h(1) + h(1)f_\beta(1) < 0 \rightarrow \\ \underbrace{f'_\alpha(0) + g'(0)}_{\text{términos de } \rho_c} < \underbrace{f_\beta(1)(1 - h(1)) + h(1)}_{\text{términos de } \rho_h} \end{aligned} \quad (\text{A.31})$$

Es decir, siempre que la aportación de los términos relacionados con $\rho_c = 0$ sea menor que la aportación de los términos de $\rho_h = 1$, tendremos que $(1, 0)$ es un punto fijo estable.

• $(0, 1)$:

$$\mathbf{J}_c = \begin{pmatrix} -r - f_\alpha(1) - g(1)(1 - f_\alpha(1)) & -r \\ f_\alpha(1) - f'_\beta(0) - h'(0) + g(1)(1 - f_\alpha(1)) & 0 \end{pmatrix}$$

$$\begin{cases} T = -r - f_\alpha(1) - g(1)(1 - f_\alpha(1)) < 0 \text{ siempre} \\ D = r \cdot (f_\alpha(1) - f'_\beta(0) - h'(0) + g(1)(1 - f_\alpha(1))) \end{cases} \quad (\text{A.32})$$

Y en este caso debe cumplirse que $D > 0$. Por tanto,

$$\begin{aligned} f_\alpha(1) - f'_\beta(0) - h'(0) + g(1)(1 - f_\alpha(1)) > 0 \rightarrow \\ \underbrace{f_\alpha(1) + g(1)(1 - f_\alpha(1))}_{\text{términos de } \rho_c} > \underbrace{f'_\beta(0) + h'(0)}_{\text{términos de } \rho_h} \end{aligned} \quad (\text{A.33})$$

Es decir, siempre que la aportación de los términos relacionados con $\rho_c = 1$ sea mayor que la aportación de los términos de $\rho_h = 0$, tendremos que $(0, 1)$ es un punto fijo estable. Por último, calculamos las curvas que marcan el límite de las regiones “Full H” y “Full

C'':

$$\alpha \leq \alpha_c \frac{1 - (1 - \beta)^k (1 - h(1)) - g'(0)}{k} \quad (\text{A.34})$$

$$\beta \leq \beta_c \frac{1 - (1 - \alpha)^k (1 - g(1)) - h'(0)}{k} \quad (\text{A.35})$$

A.5. Ambos campos con wtw

Ahora hacemos el estudio del sistema con todos los términos de la matriz.

$$\mathbf{Q} = \begin{pmatrix} 1 - f_\alpha - g(\rho_c)(1 - f_\alpha) & \rho_o h(\rho_h) & r \\ f_\alpha + g(\rho_c)(1 - f_\alpha) & (1 - \rho_o)(1 - h(\rho_h))(1 - f_\beta) & 0 \\ 0 & (1 - \rho_o)(1 - h(\rho_h))f_\beta & 1 - r \end{pmatrix} \quad (\text{A.36})$$

$$\begin{cases} F_h = -(f_\alpha + g(\rho_c)(1 - f_\alpha) + (1 - h(\rho_h))\rho_c)\rho_h + r + (r + \rho_c)(1 - \rho_c) + h(\rho_h)\rho_c^2 \\ F_c = (f_\alpha + g(\rho_c)(1 - f_\alpha) + (1 - f_\beta)(1 - h(\rho_h))\rho_c)\rho_h + ((1 - f_\beta)(1 - h(\rho_h))\rho_c - 1)\rho_c \end{cases} \quad (\text{A.37})$$

Vemos cuándo (1,0) y (0,1) son puntos fijos:

$$(\mathbf{1}, \mathbf{0}) : \begin{cases} F_h = -g(0) \\ F_c = g(0) \end{cases} \quad (\mathbf{0}, \mathbf{1}) : \begin{cases} F_h = h(0) \\ F_c = -h(0) \end{cases}$$

Y las conclusiones que sacamos son las mismas que en el apartado anterior. Queda estudiar la estabilidad de estos puntos y lo hacemos calculando el jacobiano:

$$\begin{cases} \frac{\partial F_h}{\partial \rho_h} = -[f_\alpha + g(\rho_c)(1 - f_\alpha) + r + (1 - h(\rho_h))\rho_c] + h'(\rho_h)\rho_c(\rho_c + \rho_h) \\ \frac{\partial F_h}{\partial \rho_c} = -[f'_\alpha(1 - g(\rho_c)) + g'(\rho_c)(1 - f_\alpha) + 1 - h(\rho_h)]\rho_h + 1 - r + 2\rho_c(h(\rho_h) - 1) \\ \frac{\partial F_c}{\partial \rho_h} = f_\alpha + g(\rho_c)(1 - f_\alpha) + (1 - f_\beta)(1 - h(\rho_h))\rho_c \\ \quad + [f'_\beta(h(\rho_h) - 1) + h'(\rho_h)(f_\beta - 1)]\rho_c(\rho_h + \rho_c) \\ \frac{\partial F_c}{\partial \rho_c} = [f'_\alpha(1 - g(\rho_c)) + g'(\rho_c)(1 - f_\alpha) + (1 - f_\beta)(1 - h(\rho_h))]\rho_h + 2(1 - f_\beta)(1 - h(\rho_h))\rho_c - 1 \end{cases} \quad (\text{A.38})$$

Y con esto, estudiamos los dos puntos fijos de interés:

• (1, 0):

$$\mathbf{J}_h = \begin{pmatrix} -r & -r - (f'_\alpha(0) + g'(0) - h(1)) \\ 0 & f'_\alpha(0) + g'(0) - f_\beta(1) - h(1) + h(1)f_\beta(1) \end{pmatrix}$$

Cuyos autovalores son:

$$\begin{cases} \lambda_1 = -r < 0 \text{ siempre} \\ \lambda_2 = f'_\alpha(0) + g'(0) - f_\beta(1) - h(1) + h(1)f_\beta(1) \end{cases} \quad (\text{A.39})$$

Entonces, para que el punto fijo sea estable se tiene que cumplir que:

$$\begin{aligned} f'_\alpha(0)+g'(0) - f_\beta(1)-h(1) + h(1)f_\beta(1) &< 0 \rightarrow \\ f'_\alpha(0)+g'(0) &< f_\beta(1)+h(1) - h(1)f_\beta(1) \end{aligned} \quad (\text{A.40})$$

• $(\mathbf{0}, \mathbf{1})$:

$$\mathbf{J}_c = \begin{pmatrix} -[f_\alpha(1)+g(1)(1-f_\alpha(1)) + r + 1]+h'(0) & -1-r \\ f_\alpha(1)+g(1)(1-f_\alpha(1)) + 1-f'_\beta(0) - h'(0) & 1 \end{pmatrix}$$

En este caso, analizamos la estabilidad según los valores que tomen traza (T) y determinante(D):

$$\begin{cases} T = -[f_\alpha(1)+g(1)(1-f_\alpha(1)) + r]+h'(0) \\ D = (1+r)(f_\alpha(1)+g(1)(1-f_\alpha(1)) + 1-f'_\beta(0) - h'(0)) \\ \quad - [f_\alpha(1)+g(1)(1-f_\alpha(1)) + r + 1]+h'(0) \end{cases} \quad (\text{A.41})$$

$$\begin{aligned} T < 0 &\rightarrow h'(0) < f_\alpha(1)+g(1)(1-f_\alpha(1)) + r \\ D = f'_\beta(0)(-1-r)-h'(0)r + (f_\alpha(1)+g(1)(1-f_\alpha(1)))r &> 0 \rightarrow \\ f_\alpha(1)+g(1)(1-f_\alpha(1)) &> h'(0) + f'_\beta(0) \left(1 + \frac{1}{r}\right) \end{aligned} \quad (\text{A.42})$$

Como en el campo global de honestos, si se cumple la condición de T, se cumple la condición para D y el punto fijo será estable. Por último, calculamos las curvas que marcan el límite de las regiones “Full H” y “Full C”:

$$\alpha \leq \alpha_c \frac{1 - (1 - \beta)^k(1-h(1))-g'(0)}{k} \quad (\text{A.43})$$

$$\beta \leq \beta_c \left(\frac{r}{r+1}\right) \frac{1 - (1 - \alpha)^k(1-g(1))-h'(0)}{k} \quad (\text{A.44})$$

Tanto en este caso como para el caso sin wtw, vemos que las curvas límite son combinación de las obtenidas en los estudios de cada campo por separado.

B. Anexo II: Campos globales de honestidad lineales

Vamos a estudiar si introducir el campo global de honestidad con una función lineal sin término independiente abre la posibilidad a que haya regiones de biestabilidad. En concreto, si es posible que haya biestabilidad entre los puntos fijos “Full H” y “Full C”. La expresión del campo global de honestidad es

$$h(\rho_h) = h_1 \rho_h \quad (\text{B.1})$$

Con $h'(0) = h(1) = h_1$, y vamos a aplicar el mismo razonamiento que en apartado 3.2. En este caso, las curvas que delimitan en el diagrama de fases (β, α) los límites de las regiones “Full H” y “Full C” vienen dadas por (A.25) y (A.26). Comenzamos estudiando la monotonía de cada curva.

Para la ecuación (A.26):

$$\frac{\partial \beta_c}{\partial \alpha} = \left(\frac{r}{r+1} \right) (1-\alpha)^{k-1} \Rightarrow \frac{\partial^2 \beta_c}{\partial \alpha^2} = - \left(\frac{r}{r+1} \right) (1-\alpha)^{k-2} (k-1) \quad (\text{B.2})$$

Para la ecuación (A.25) llamamos $\bar{\beta}(\alpha)$ a la función inversa de $\beta_c(\alpha)$; es decir, $\alpha_c(\bar{\beta}(\alpha)) = \alpha$:

$$\alpha_c = \frac{1 - (1-\beta)^k (1-h_1)}{k} \Rightarrow \bar{\alpha} = 1 - \left(\frac{1-k\alpha}{1-h_1} \right)^{\frac{1}{k}} \quad (\text{B.3})$$

Ahora, calculamos su segunda derivada:

$$\frac{\partial \bar{\beta}}{\partial \alpha} = \left(\frac{1-k\alpha}{1-h_1} \right)^{\frac{1}{k}-1} \frac{1}{1-h_1} \Rightarrow \frac{\partial^2 \bar{\beta}}{\partial \alpha^2} = \left(\frac{1-k\alpha}{1-h_1} \right)^{\frac{1}{k}-2} \frac{k-1}{(1-h_1)^2} \quad (\text{B.4})$$

Teniendo en cuenta que β recorre el eje de abscisas y α el de ordenadas, se concluye que $\beta_c(\alpha)$ es una función cóncava creciente que corta al eje de ordenadas en el valor de $\alpha^* = 1 - (1-h_1)^{\frac{1}{k}}$, con pendiente $\left(\frac{r}{r+1} \right) (1-h_1)^{\frac{k-1}{k}} \leq \frac{(1-h_1)^{\frac{k-1}{k}}}{2}$ (pues $r \leq 1$), mientras que $\bar{\beta}(\alpha)$ es una función convexa creciente que corta al eje de abscisas en $\alpha_c(0) = \frac{h_1}{k}$. Esta pendiente es estrictamente mayor que la de $\beta_c(\alpha)$ en α^* .

Ahora probamos que $\alpha^* \geq \alpha_c(0)$ para $0 < h_1 < 1$ para lo que basta ver que las funciones $f_1(x) = 1 - (1-x)^k$ y $f_2(x) = \frac{x}{k}$ satisfacen que $f_1(x) > f_2(x)$ para $0 < x \leq 1$:

(a) $f_1(0) = f_2(0) = 0$

(b) $f_1'(x) = \frac{1}{k}(1-x)^{\frac{1}{k}-1}$; $f_2'(x) = \frac{1}{k} \Rightarrow f_1'(0) = \frac{1}{k} = f_2'(0)$

(c) $f_1''(0) = \frac{1}{k} \left(1 - \frac{1}{k}\right) (1-x)^{\frac{1}{k}-2} > 0 \forall x \text{ t.q. } 0 \leq x \leq 1$

De lo que se concluye que para todos $0 < x \leq 1$ y $k > 1$, $f_1(x) > f_2(x)$, y las regiones de “Full H” y “Full C” estables no se solapan en el eje de ordenadas. Y gracias al estudio de la monotonía de las curvas y a que la pendiente de A.26 es mayor que la pendiente de A.25 en $\beta = 0$, se deduce que no se produce solapamiento entre dichas regiones.

C. Anexo III: resolución iterativa de las ecuaciones de campo medio para el campo de honestidad

Como se ha dicho en 3.3, no es posible obtener expresiones analíticas para los puntos fijos $(F_h, F_c)=(0,0)$. Así, tratamos de hallar los valores de ρ_h y ρ_c de forma numérica mediante la resolución de las ecuaciones (A.20) para distintos valores de los parámetros del sistema. Primero, se prueba que estos resultados concuerdan totalmente con los de la iteración de las ecuaciones markovianas para redes homogéneas, como se muestra en la figura C.1

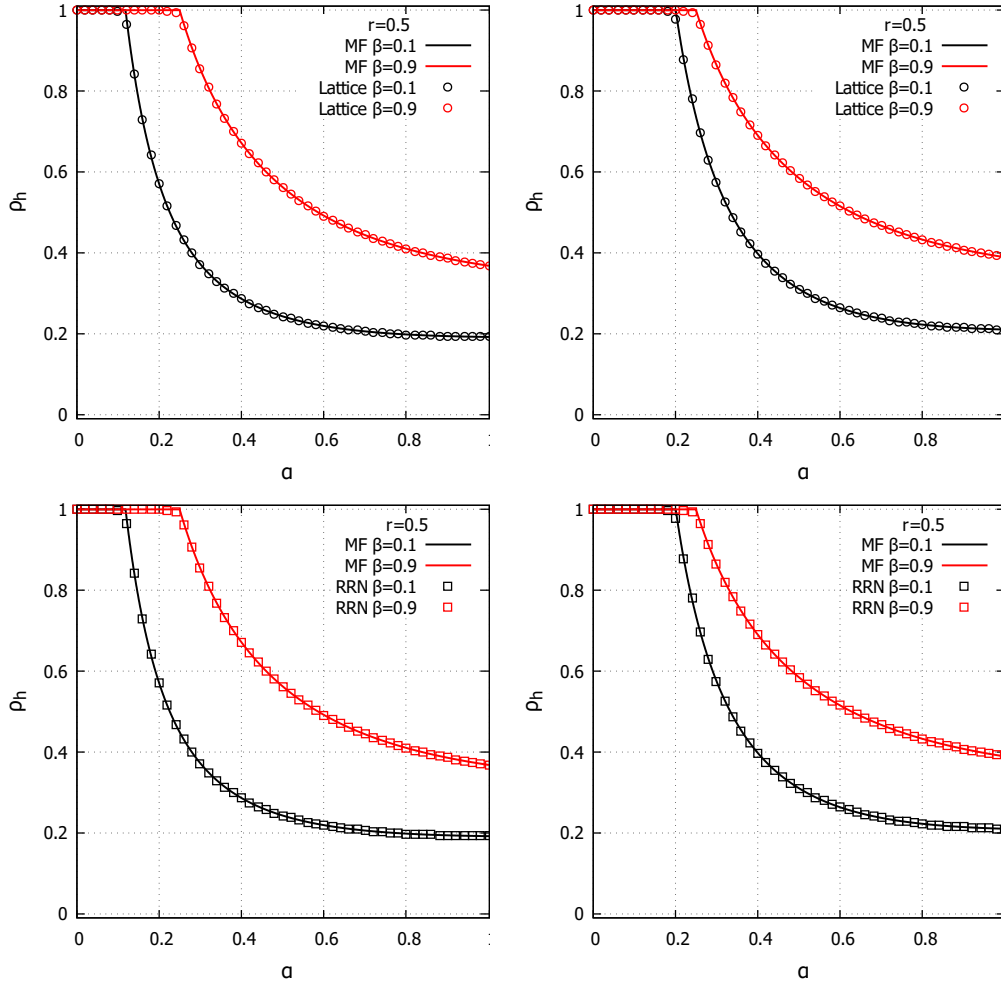


Figura C.1: Fracción de honestos en función de la tasa de corrupción $\rho_c(\alpha)$ con $r = 0.5$ y $\beta = 0.1, 0.9$. La línea continua representa la resolución de campo medio, los puntos circulares, la resolución de Markov para una RRN, y los puntos cuadrados, la resolución de Markov para una *Lattice*. La expresión del campo global de honestidad para la primera columna es $h(\rho_h) = 0.1 + 0.1\rho_h$, y para la segunda columna, $h(\rho_h) = 0.7\rho_h$.

Dado el total acuerdo entre estas dos resoluciones numéricas, pasamos a mostrar un análisis más detallado de la evolución de los estados estacionarios de campo medio en función de un parámetro variable.

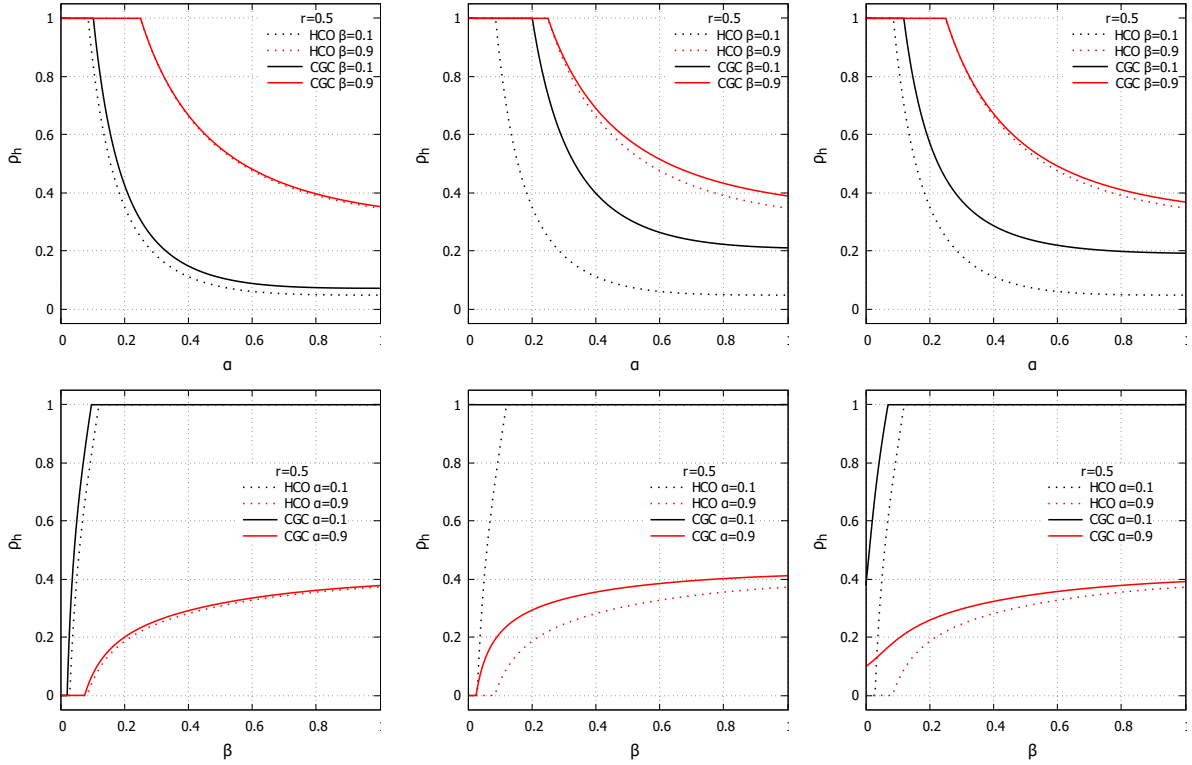


Figura C.2: Primera fila: fracción de honestos en función de la tasa de corrupción $\rho_h(\alpha)$ para $\beta = 0.1, 0.9$. Segunda fila: fracción de honestos en función de la tasa de delación $\rho_h(\beta)$ para $\alpha = 0.1, 0.9$. Las líneas puntadas muestran la resolución del modelo original y las líneas continuas, la resolución con el campo global de honestidad (CGH). La expresión del CGH es $h(\rho_h) = 0.1\rho_h$ para la primera columna, $h(\rho_h) = 0.7\rho_h$ para la segunda, y $h(\rho_h) = 0.1 + 0.1\rho_h$ para la tercera. En todas las gráficas $r = 0.5$

En la figura C.2 se muestra el efecto de incluir el campo global de honestidad (líneas continuas) sobre los resultados originales (líneas discontinuas). Como era de esperar, en todas las comparaciones, al incluir el CGH la fracción de honestos es mayor que en el resultado original. También se puede comprobar que efectivamente las transiciones “Full H”-*mixed* siguen (D.1), y las transiciones “Full C”-*mixed*, (D.2) salvo la última curva de la segunda fila. En esta, como $h(0) > 0$, “Full C” nunca es un punto fijo estable, tal y como se ha deducido en el apartado A.3.

D. Anexo IV: Biestabilidad para el campo de honestidad

Seguiremos el mismo razonamiento que en el texto principal, pero ahora las curvas que delimitan las regiones de "todo honestos" y "todo corruptos" son las obtenidas en el anexo A.3:

$$\alpha_c = \frac{1 - (1 - \beta)^k(1-h(1))}{k} \quad \text{para "todos honestos"} \quad (\text{D.1})$$

$$\beta_c = \frac{1 - 1(1 - \alpha)^k - h'(0)}{k} \left(\frac{r}{r+1} \right) \quad \text{para "todos corruptos"} \quad (\text{D.2})$$

Llamamos $\bar{\beta}(\alpha)$ a la función inversa de $\alpha_c(\beta)$; es decir, $\alpha_c(\bar{\beta}(\alpha)) = \alpha$:

$$\bar{\beta}(\alpha) = 1 - \left(\frac{1 - k\alpha}{1-h(1)} \right)^{1/k} \quad (\text{D.3})$$

Pero para hacer más sencillo el estudio, fijamos α . Así pues, para que haya solapamiento entre regiones el mínimo valor de β para el cual la región "Full H" es estable debe ser inferior al máximo valor de β para el cual la región "Full C" es estable. Imponiendo estas condiciones en las ecuaciones anteriores,

$$\begin{aligned} \bar{\beta} > \beta_c &\Rightarrow \frac{1 - 1(1 - \alpha)^k - h'(0)}{k} \left(\frac{r}{r+1} \right) > 1 - \left(\frac{1 - k\alpha}{1-h(1)} \right)^{1/k} \\ h'(0) &< 1 - (1 - \alpha)^k - \left[1 - \left(\frac{1 - k\alpha}{1-h(1)} \right)^{1/k} \right] \left(\frac{r+1}{r} \right) k \end{aligned} \quad (\text{D.4})$$

Además de esta condición, (B.3) introduce dos nuevas restricciones, dado que está acotada: $0 \leq \bar{\beta}(\alpha) \leq 1$, y se deduce que:

$$\left\{ \begin{array}{l} \leq \bar{\beta}(\alpha) \Rightarrow h(1) \leq k\alpha \\ \bar{\beta}(\alpha) \leq 1 \Rightarrow \alpha_{max} = \frac{1}{k} \end{array} \right. \quad (\text{D.5})$$

Sustituyendo la primera restricción en (D.4) se llega a la curva que limita el máximo valor que puede tomar $h'(0)$ en función de α :

$$h'(0) = 1 - (1 - \alpha)^k = 1 - \left(1 - \frac{h(1)}{k} \right)^4 \quad (\text{D.6})$$

La figura D.1 muestra el conjunto de valores de $h'(0)$ y $h(1)$ que dan lugar a biestabilidad entre "Full H" y "Full C" para distintos valores de α .

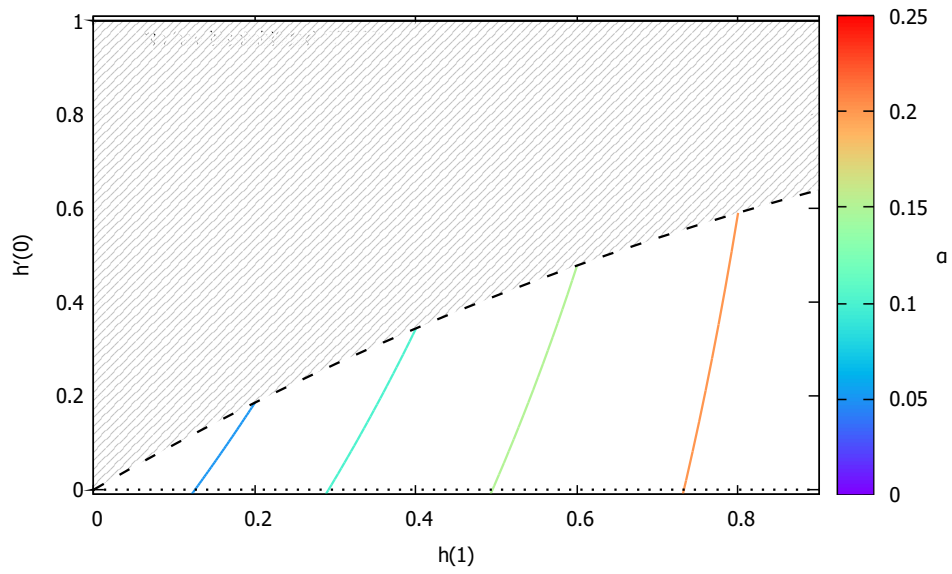


Figura D.1: Curva $h'(0)$ para distintos valores de α con $r = 0.9$.