



**PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
INSTITUTO DE ECONOMIA
MAGISTER EN ECONOMIA**

**TESIS DE GRADO
MAGISTER EN ECONOMIA**

Araya, Vergara, Dominique Delmar

Diciembre, 2017



**PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
INSTITUTO DE ECONOMIA
MAGISTER EN ECONOMIA**

Matching en el Mercado de Citas

Dominique Delmar Araya Vergara

Comisión

Nicolás Figueroa, Rodrigo Harrison

Santiago, Diciembre de 2017

Resumen: Se analizo un contexto de two-sided matching en donde para poder experimentar con las personas al otro lado del mercado, estas deben querer experimentar con uno. La motivación para este trabajo fue la aparición de páginas web que unen a pacientes con investigadores que conducen ensayos clínicos. De manera sencilla, se introdujo un proceso adaptativo para las creencias sobre la utilidad de salir con un hombre o mujer, y la creencia sobre la probabilidad de ser aceptados de los hombres. Se concluye que al menos localmente, es decir, para cuando las mujeres ya no se equivocan en elegir, el proceso anterior converge a un estado estacionario.¹

¹Tesis escrita como estudiante de Magister en el seminario de tesis de Microeconomía, Departamento de Economía, Pontificia Universidad Católica de Chile. Agradezco a mis profesores Rodrigo Harrison y especialmente a Nicolas Figueroa, por su apoyo. También, agradezco el instituto ISCI (Núcleo Milenio ICR RC130003) y a CONYCIT (CONYCIT-PFCHA/MagisterNacional/2017-22170613) por apoyo económico. Cualquier error o omisión es mi exclusiva responsabilidad.

1. Introducción

Para poder tomar una decisión, los agentes deben aprender sobre sus opciones. En muchos casos, para hacerlo, deben experimentar con ellas, – o alguien en su grupo social debe hacerlo– lo que conlleva un trade-off entre experimentación y explotación. Con este trade-off, el agente debe decidir si aprender más sobre el pago de los distintas opciones (experimentar) o simplemente elegir la opción que tiene el pago esperado más alto (explotar) (Bergemann y Valimaki, 2006).

Este trade-off se relaciona estrechamente con la literatura de multi-armed bandit problem, el cual es un modelo de decisión estadística en donde un agente intenta optimizar sus decisiones mientras simultáneamente va mejorando su información. En este problema, cada palanca de un traga monedas, si es elegida, resulta en un pago inmediato aleatorio, pero el conocimiento respecto al proceso que determina estos pagos va evolucionando durante el juego (Bergemann y Valimaki, 2006).

Sin embargo, hay veces en que solo se puede experimentar si el agente al otro lado del mercado quiere experimentar con uno. Por ejemplo, un trabajador solo puede conocer una empresa si esta lo contrata o una persona puede conocer a una pareja si esta accede a salir con ella. Esto es un contexto particular de experimentación estratégica. Bergemann y Valimaki, por ejemplo, endogenizan el costo de experimentación del consumidor al permitir la competencia en precios de los vendedores en un escenario en que un único consumidor compra bienes de forma progresiva de distintos vendedores en el tiempo, donde el valor de cada producto es desconocido y solo se puede ganar nueva información a través de experimentación (Bergemann y Välimäki, 1996).

Dado este contexto, el foco estará en situaciones en que se tienen two-side matching entre agentes que deben experimentar unos con otros. Un claro ejemplo de esto son las páginas web que unen a pacientes con investigadores que conducen ensayos clínicos. En estas plataformas, los pacientes eligen participar en un ensayo clínico, y otras personas con características similares y la misma enfermedad, puede actualizar sus percepciones sobre este. A su vez, el investigador va actualizando sus creencias sobre qué tipo de paciente tiene una mejor respuesta al tratamiento experimental. Una de las características importantes de este ejemplo, es la ausencia de transferencias entre pacientes e investigadores, y además, esta el hecho de que los pacientes deben decidirse por solo un ensayo clínico en el cual participar, con lo que existe un costo de oportunidad al tomar una decisión.

Otro posible ejemplo es el mercado de citas, en el que, en la mayoría de los casos, hombres y mujeres salen en múltiples citas que si bien entregan un pago inmediato a ambos agentes, también entregan información sobre el verdadero valor de la pareja.

De esta forma, basándome en el trabajo de Das y Kamenica, considero un juego repetido donde agentes van obteniendo un pago incierto por ser emparejados, a través de algún mecanismo de matching, con otro agente en particular cada periodo. Dado que ambos lados del mercado tienen agencia, Das y Kamenica se refieren a este problema como un “Two-Sided Bandit Problem” (Das y Kamenica, 2005).

Tal como Das y Kamenica indican, si bien se querría analizar el equilibrio bayesiano del comportamiento de los agentes, este equilibrio es prohibitivo de computar, no solo para el modelador sino también para los agentes. Dado lo anterior, el acercamiento que usan Das y Kamenica está relacionado con la teoría de aprendizaje en juegos (Fudenberg y Levine, 1998) y analizan la estabilidad asintótica cuando los agentes utilizan un e-greedy algoritmo de aprendizaje adaptado a distintos mecanismos de matching. Por otro lado, acá se considera un acercamiento donde los agentes siguen heurísticas simples basadas en el historial.

Siguiendo a Cominetti, Melo y Sorin (Cominetti et al., 2010) y la teoría de aprendizaje en juegos (Fudenberg y Levine, 1998), utilizo smooth fictitious play como proceso adaptativo y herramientas de aproximación estocástica para observar la convergencia de procesos estocásticos de tiempo discreto (Benaïm y Hirsch, 1999).

Si bien seguiré la metodología de Cominetti, Melo y Sorin (2010) para estudiar la convergencia de las dinámicas adaptativas, existen marcadas diferencias entre ambos trabajos. Por un lado, Cominetti, Melo y Sorin tienen un matching many-to-one, donde muchos agentes en un periodo pueden elegir una misma opción y, además, donde el pago esperado dependerá de la cantidad de agentes que la hayan elegido, creándose un concepto de congestión. A diferencia de estos, este trabajo tiene un match one-to-one, en que ambos lados del mercado toman decisiones, y donde una vez realizado el pago de la unión, este no depende directamente de la cantidad de personas que también eligieron la misma estrategia. Pero sí hay un efecto de la cantidad de agentes que eligen una estrategia, en la probabilidad de concretar el match.

Dentro de los mecanismo de matching, analizare uno de ofertas simultáneas, en que cada periodo un lado del mercado envía propuestas a un agente al otro lado del mercado, todos al mismo tiempo, y una vez recibidas estas, el otro lado del mercado elige, dentro de sus alternativas, a quien aceptar. Se puede imaginar una situación en que hombres envían por correo una invitación, todas al mismo tiempo, y luego las mujeres, dado los correos recibidos, eligen con quien tener una cita. Se supondrá que todos los agentes tienen un valor objetivo de salir con ellos, pero este valor se supone desconocido. Por lo tanto, se tienen creencias sobre el valor de salir con una persona del otro lado del mercado y, para el caso de los agentes que invitan, también creencias sobre la probabilidad de ser aceptado. Una vez que se produce la unión, cada persona recibe un pago que es el valor de la persona con la que salió más un ruido blanco. Este pago es comparado con el pago esperado de la unión y con esto, se actualizan las creencias de los agentes.

Por lo tanto, en este sistema los parámetros que van cambiando en el tiempo son simplemente valoraciones que las personas tienen sobre sus posibles parejas, y las creencias sobre la posibilidad de ser aceptado. Los procesos estocásticos de tiempo discreto son tales que el valor de los parámetros entre mañana y hoy, dependen de los pagos recibidos por la acción tomada. (Benaïm, 1999)

Para poder simular la experimentación por parte de un lado del mercado, utilizo una logit para definir la probabilidad de que se realicen invitaciones, de forma de que si bien la alternativa que maximiza su utilidad esperada tomando en cuenta solo el periodo presente tiene más probabilidad de ser invitada, la probabilidad de elegir cualquier alternativa es siempre positiva. La razón de que se usa una logit para generar las estrategias mixtas de un lado del mercado es que la regla logística es una heurística simple para capturar el trade-off entre experimentación y explotación, y que además tiene buenas propiedades matemáticas. El lado del mercado que recibe las invitaciones, por otro lado, toma decisiones de manera miope, eligiendo la alternativa que creen tiene un valor esperado más alto.

El proceso adaptativo que se propone para las creencias de los agentes es sencillo, y la cantidad de información que deben mantener los agentes es simplemente sus creencias durante el periodo anterior. Dado esta dinámica, mi pregunta de investigación es si esta forma de aprendizaje, basado en una información tan simple, es capaz de generar un sistema que alcance el equilibrio. Es decir, quiero analizar los parámetros que permiten la convergencia del sistema a valores específicos, y discutir que tan lejos están estos de los valores reales. Además, la relación de estos valores con el equilibrio de algún juego.

2. Literatura Relacionada

Das y Kamenica (2005) consideran un juego repetido en que las personas, separadas en hombres y mujeres, obtienen un pago incierto al ser emparejados entre ellos a partir de algún mecanismo de matching, debiendo enfrentarse al trade-off entre experimentación y explotación. Dado la complejidad de estos procesos, los autores utilizan un algoritmo llamado e -greedy que con probabilidad e selecciona la opción con mayor valor esperado y con $1 - e$ otra opción de forma aleatoria. A diferencia de los autores, utilizo una regla logística para simular la experimentación, que entrega las estrategias mixtas de los agentes.

Cominetti, Melo y Sorin (2010) está dentro de la literatura en dinámicas de juegos repetidos en donde los agentes toman decisiones basados en información adquirida en experiencias previas. En este paper, se asume que cada jugador tiene una percepción previa del pago de cada alternativa, en este caso de una ruta a elegir. Después que el pago es observado, este es utilizado para actualizar la percepción a través de un proceso de aprendizaje que depende únicamente de los pagos realizados. Aunque cada jugador observa solo sus pagos, estos se ven afectados por las decisiones de los demás agentes, ya que se produce un efecto de congestión. De esta forma, se tiene un trade-off entre experimentación versus cooperación, en vez de uno entre experimentación versus explotación. Para analizar lo anterior, ocupan herramientas usadas en la teoría de algoritmos estocásticos.

Los procesos estocásticos de tiempo discreto son tales que la diferencia entre el valor del parámetro entre mañana y hoy está dada por un pequeño paso y una regla o algoritmo expresado como función, que depende del valor del parámetro hoy y un error que puede interpretarse como la nueva información recibida. En cada momento, este sistema va recibiendo información nueva sobre el parámetro, que causa que se actualice siguiendo el algoritmo específico (Benaïm, 1999).

Lo que este trabajo hace es combinar las herramientas desarrolladas en Cominetti et al. (2010) con el ambiente descrito por Das y Kamenica (2005). Una de las aplicaciones de este trabajo es el caso de pacientes que utilizan plataformas de Internet para ser parte de ensayos clínicos. Unos de estos sitios web es OncoLink, manejado por University of Pennsylvania Cancer Center, que une a pacientes de cancer con ensayos clínicos (Metz et al., 2005). También existe TrialX, página web que une pacientes con ensayos clínicos extrayendo su información de plataformas de registros personales de salud como Google Health. Para ver como funciona esta página, refieranse a Patel et al. (2010). De manera más oficial, el US National Institutes of Health mantiene una página en que también se pueden encontrar ensayos clínicos para distintas enfermedades, y se entrega la información del investigador para que el paciente pueda ponerse en contacto y ver si puede ser partícipe del ensayo (of Health et al., 2012).

Los mercados en que se unen dos lados, o *two-sided matching markets*, son aquellos en que los agentes pertenecen a uno de dos grupos distintos disjuntos, como es el caso de firmas y trabajadores, donde la unión aparece en el hecho de que si un paciente trabaja en una firma, entonces la firma lo contrata (Roth y Sotomayor, 1992). Este tipo de situación fue analizada por Gale y Shapley en 1962, en el caso del llamado *marriage problem*. Cierta cantidad de hombres y mujeres ranquean a aquellos del sexo opuesto en términos de su preferencia para casarse. Se analiza la existencia de grupos de matrimonios estables y se presenta un procedimiento llamado *deferred acceptance*, que genera uniones, en este caso matrimonios, estables. (Gale y Shapley, 1962).

Luego, los autores expanden lo anterior al problema de admisión a las universidades, en donde alumnos deben decidir a que universidad postular y universidades, que tienen una cantidad limitada de cupos, deben decidir a qué estudiantes admitir. En este caso se trata de many-to-one matching. Gale y Shapley describen un proceso que sirve para asignar postulantes a universidades que satisface a ambos grupos (Gale y Shapley, 1962). La contribución de Gale y Shapley fue extendida por autores como Roth y Sotomayor (1989) Roth (1982), Roth (1985), quienes muestran propiedades que poseen las uniones estables en problemas de matching many-to-one, y Crawford (1991).

En modelos más dinámicos, se tiene a Ünver (2010), que estudia mecanismo óptimos de intercambios que minimizan los tiempos de espera, en el contexto del intercambio de riñones, donde la dinámica está en la entrada y salida de donantes y receptores del mercado, y donde los agentes son unidos una sola vez. En este mismo contexto de mercados dinámicos óptimos, donde agentes llegan de forma secuencial, se tiene el caso de Akbargpour et al. (2014) y también se tienen situaciones donde agentes llegan y se van de manera estocástica, como es el caso de Baccara et al. (2015).

A diferencia de lo anterior, en un contexto dinámico también se han analizado matchings one-to-one en

mercados donde las valoraciones van evolucionando en el tiempo debido al aprendizaje a través de experimentación, estudiando las propiedades de subastas de matching tales que los agentes ofertan sus valoraciones reales sobre todas sus posibles parejas (Fershtman y Pavan, 2015). Estos introducen una clase de mecanismos de matching que toman en cuenta el hecho de que en muchos ambientes, los valores de la unión van cambiando en el tiempo.

Este paper también se relaciona con la literatura de experimentación, y más específicamente, la literatura referente a *bandits problems*. Este problema fue originalmente descrito por Robbins (1952), donde un agente busca optimizar su decisión a la vez que intenta mejorar la información de la que dispone. En este contexto, el trade off entre el pago actual (explotación) y el valor de la información experimentación, juega un rol importante (Bergemann y Valimäki, 2006). En términos de experimentación y matching, por ejemplo, Jovanovic (1979) aplica un *bandit problem* a mercados de trabajo competitivos. Acá, los agentes son estratégicos y sus acciones determinan los pagos de distintas alternativas, como sucede en Bergemann y Välimäki (1996) y Bergemann y Välimäki (2000).

3. Modelo

Se considera un modelo de infinitos periodos con dos lados del mercado, donde jugadores en ambos lados perciben utilidad por ser unidos en cada periodo. La utilidad que cada unión entrega a agentes es desconocida por los agentes, quienes se forman creencias sobre estos valores. Dado que los valores son desconocidos, los jugadores deben decidir entre intentar unirse al jugador que perciben entregará un mayor pago esperado en el periodo, o buscar aprender más sobre el valor de otro jugador. Por lo tanto, la decisión del agente en cada periodo conlleva un trade off entre explotación y exploración. Cada jugador ajusta de forma iterativa su estrategia como una función de los resultados obtenidos en el pasado.

Sea $I = I_m \cup I_w$ el set de jugadores, donde $I_m = \{1, \dots, n\}$ es el conjunto de jugadores en el lado M del mercado, que llamaremos hombres, y $I_w = \{1, \dots, n\}$ el conjunto de jugadores en el lado W , que llamaremos mujeres.

Cada jugador $i \in I$ está caracterizado por un set finito S^i de estrategias puras y una función de pagos $G^i : S^i \times S^{-i} \rightarrow \mathbb{R}$, donde $S^{-i} = \prod_{k \neq i} S^k$.

Se define el conjunto de estrategias mixtas sobre S^i como Δ^i , y donde $\Delta = \prod_{i \in I} \Delta^i$. En cada periodo $t \in \{1, 2, 3, \dots\}$, cada hombre $m \in I_m$, de forma simultánea, invita a una mujer $w \in I_w$. Luego, cada mujer $w \in I_w$ elige a quien aceptar entre las invitaciones recibidas. Así,

$$\begin{aligned} S^m &= I_w \quad \forall m \in I_m \\ S^w &= \{i \in I_m \mid s^i = w\} \quad \forall w \in I_w \end{aligned}$$

Al final del periodo, cada agente observa el pago de la unión que es función del verdadero valor de la persona con la que salió más un ruido blanco. Se define el valor de salir con una persona s como v^s .² La ganancia de ser unido al jugador s en cada periodo es $g_t^{is} = v^s + \xi_t^{is}$, mientras que si un jugador no es unido a otro, el pago de ese periodo es 0.

En cada momento del tiempo, los agentes tienen creencias sobre el valor de los agentes al otro lado del mercado. Se define z^{mw} como la creencia sobre el valor que tendrá el salir con la mujer w y p^{mw} como la creencia sobre la probabilidad de ser aceptado por la mujer w . Donde $z_t^m \times p_t^m = (z_t^{mw} \times p_t^{mw})_{w \in I_w}$ es la percepción del pago de las estrategias puras disponibles para los hombres y es la relevante para ellos en la elección de su estrategia.

²Las personas aportan un valor a la relación que es independiente de la pareja, $v^{is} = v^s \quad \forall i \in I$

Análogamente, $x_t^w = (x_t^{wm})_{m \in S^w}$ es la percepción del pago de las estrategias puras disponibles para las mujeres, que en este caso es la creencia sobre el valor de ser emparejado con el hombre m que indica la estrategia s_t^w .

Un hombre $m \in I_m$, elige su estrategia aleatoriamente usando una regla logística:

$$\sigma^{mw}(z^m \times p^m) = \frac{\exp(\beta(z^{mw} \times p^{mw}))}{\sum_{a \in S^i} \exp(\beta(z^{ma} \times p^{ma}))} \quad \forall m \in I_m \quad (1)$$

Cabe notar que σ^{mw} es estrictamente positiva para todo $w \in S^m$, $m \in I_m$. En particular, siempre hay experimentación por parte de los hombres.

Las mujeres, por su parte, juegan su mejor respuesta de manera miope, tal que aceptan al hombre $m \in I_m$ que creen tiene un mayor valor entre los que la invitaron. En el caso que nadie la invite, se queda sola, con lo que su estrategia queda definida por:

$$s_t^w \in \begin{cases} \underset{\Phi}{\operatorname{argmax}}_{\{m \mid s^m = w\}} x_t^{wm} & \text{si } \{s^m = w\} \neq \emptyset \\ \text{si no} & \end{cases} \quad \forall w \in I_w \quad (2)$$

Por lo tanto, se puede definir la estrategia mixta de $i \in I_w$ ex-post, es decir, después de que se reciban las invitaciones, como un vector $(n+1) \times 1$ de unos y ceros:

$$\overline{\pi_t^{wm}} = \overline{\sigma^{wm}}(x_t^w) = \begin{cases} 1 & \text{si } m = s_t^w \\ 0 & \text{si no} \end{cases}$$

Si bien la estrategia que se elije tras que las invitaciones son recibidas es pura, se puede escribir la estrategia ex-ante de $w \in I_w$ como una mixta definida de la siguiente forma:

$$\pi_t^{wm} = \tilde{\sigma}^{wm}(x_t^w, z, p) = \sigma^{mw}(z_t^m, p_t^m) \prod_{\{m' \mid x_t^{wm'} > x_t^{wm}\}} (1 - \sigma^{m'w}(z_t^{m'}, p_t^{m'})) \quad \forall w \in I_w \quad (3)$$

Donde la probabilidad ex-ante de que $w \in I_w$ acepte a $m \in I_m$ es la probabilidad de que m le invite multiplicado por la probabilidad de que ningún otro agente que w considera mejor haya enviado una invitación. Además, se define la probabilidad de quedar solo, que equivale a la probabilidad de que nadie envíe una invitación a $w \in I_w$.

$$\pi_t^{w\Phi} = \prod_{m \in I_m} (1 - \sigma^{mw}(z_t^m, p_t^m))$$

Estas expresiones corresponden a una estrategia completamente mixta (todos los hombres son seleccionados con probabilidad estrictamente positiva) y corresponde a las estrategias de las mujeres desde un punto de vista ex-ante (al comienzo del periodo).

Considero que las percepciones de las mujeres se actualizan de la siguiente forma:

$$x_{t+1}^{wm} = \begin{cases} (1 - \gamma_{t+1})x_t^{wm} + \gamma_{t+1}g_t^i & \text{si } s = s_t^i \\ x_t^{wm} & \text{si no} \end{cases} \quad \forall w \in I_w$$

En donde la evolución de una variable depende exclusivamente de un shock aleatorio, que aparece si es que el hombre o mujer eligió cierta alternativa, y del valor anterior de esta. Se puede escribir la iteración en forma vectorial como:

$$x_{t+1} - x_t = \gamma_{t+1}[w_{t+1} - x_{t+1}] \quad \forall w \in I_w \quad (4)$$

con

$$w_t^{wm} = \begin{cases} g_t^{wm} & \text{si } s = s_t^i \\ x_t^{wm} & \text{si no} \end{cases}$$

Donde $\gamma_{t+1} \in (0,1)$ es una secuencia talque $\sum_t \gamma_{t+1} = \infty$ y $\sum_t (\gamma_{t+1})^2 < \infty$. Un candidato para $\gamma_t \in$ es simplemente $\frac{1}{t}$. (Cominetti et al., 2010)

Para los hombres, se separan los procesos de actualización de las dos percepciones y se tienen las siguientes actualizaciones para el valor de la unión y para la probabilidad de ser aceptado respectivamente

$$z_{t+1}^{mw} - z_t^{mw} = \begin{cases} \gamma_{t+1}(g_t^m - z_t^{mw}) & \text{si } s^m = w \text{ y } s^w = m \\ 0 & \text{otherwise} \end{cases}$$

$$p_{t+1}^{mw} = \begin{cases} p_t^{mw} & \text{si } s^m \neq w \\ p_t^{mw} \times (1 - \gamma_{t+1}) & \text{si } s^m = w \wedge s^w \neq m \\ p_t^{mw} \times (1 - \gamma_{t+1}) + \gamma_{t+1} & \text{si } s^m = w \wedge s^w = m \end{cases}$$

Se escribe la actualización de forma matricial.

$$\begin{bmatrix} z_{t+1} \\ p_{t+1} \end{bmatrix} - \begin{bmatrix} z_t \\ p_t \end{bmatrix} = \gamma_{t+1} \left(\begin{bmatrix} w_{z,t} \\ w_{p,t} \end{bmatrix} - \begin{bmatrix} z_t \\ p_t \end{bmatrix} \right) \quad \forall m \in I_m \quad (5)$$

Donde

$$w_{z,t}^{mw} = \begin{cases} g_t^m & \text{si } s^m = w \wedge s^w = m \\ z_t^{mw} & \text{otherwise} \end{cases}$$

$$w_{p,t}^{mw} = \begin{cases} 1 & \text{si } s^m = w \wedge s^w = m \\ 0 & \text{si } s^m = w \wedge s^w \neq m \\ p_t^{mw} & \text{si } s^m \neq w \end{cases}$$

Es importante notar que a los hombres solo les importa $z \cdot p$. Sin embargo, la historia de resultados $z_t \cdot p_t$ no es un estadísticos suficiente para actualizar las expectativas respecto a $z_{t+1} \cdot p_{t+1}$, como se ve en el siguiente lema.

Lema 1 Actualizar z y p de forma separada y actualizar $z \times p$ de manera conjunta no es equivalente.

Demostración Lema 1 La actualización cuando el hombre m no invita a la mujer w es equivalente entre ambas formas, ya que de forma conjunta se tiene que

$$(z \times p)_{t+1}^{mw} = (z \times p)_t^{mw}$$

mientras que siguiendo (5) se tiene:

$$z_{t+1}^{mw} = z_t^{mw}$$

$$p_{t+1}^{mw} = p_t^{mw}$$

con lo que se tiene que:

$$(z \times p)_{t+1}^{mw} = z_{t+1}^{mw} \times p_{t+1}^{mw} = z_t^{mw} \times p_t^{mw} = (z \times p)_t^{mw}$$

Por otro lado, cuando el hombre m sí invita a la mujer w , pero esta no lo acepta, la actualización de $z \times p$ es la siguiente:

$$(z \times p)_t^{mw} + 1 = (1 - \gamma_{t+1})(z \times x)_t^{mw}$$

y al separar los procesos se tiene:

$$(z \times p)_t^{mw} + 1 = z_{t+1}^{mw} \times p_{t+1}^{mw} = z_t^{mw} \times p_{t+1}^{mw}(1 - \gamma_{t+1}) = (1 - \gamma_{t+1})(z \times p)_t^{mw}$$

Sin embargo, cuando el hombre m invita a la mujer w y esta lo acepta, de forma conjunta la actualización es la siguiente:

$$(z \times p)_t^{mw} + 1 = (z \times p)_t^{mw} + \gamma_{t+1}[g_t^{mw} - (z \times p)_t^{mw}]$$

mientras que de manera separada la actualización es

$$\begin{aligned} (z \times p)_{t+1}^{mw} &= z_{t+1}^{mw} \times p_{t+1}^{mw} = [z_t^{mw} + \gamma_{t+1}(g_t^{mw} - z_t^{mw})][p_t^{mw}(1 - \gamma_{t+1} + \gamma_{t+1})] \\ &= z_t^{mw} p_t^{mw} (1 - \gamma_{t+1})^2 + p_t^{mw} \gamma_{t+1} (1 - \gamma_{t+1}) + \gamma_{t+1}^2 (g_t^{mw} - z_t^{mw}) \end{aligned}$$

lo que no es equivalente. ■

Por lo tanto, $p \times z$ no es un estadístico suficiente y si bien el agente $m \in I_m$ toma decisiones basadas en $p \times z$, lo que se va actualizando en realidad es el vector $\begin{bmatrix} p \\ z \end{bmatrix}$.

3.1. Aproximación Estocástica

En este trabajo, me interesa entender el comportamiento del sistema dinámico planteado anteriormente, en particular su comportamiento asintótico, para así entender que ocurre con las creencias de las personas en el largo plazo dentro de este modelo. Sin embargo, el sistema estocástico presentado es muy difícil, o casi imposible, de analizar pese a simpleza. Por lo tanto, con el fin de poder estudiar este comportamiento, una técnica que se puede utilizar es la que nos entrega la teoría de algoritmos estocásticos. Siguiendo la teoría de algoritmos estocásticos (Benaïm, 1999) (Benaïm et al., 2005), se pueden establecer conexiones entre procesos de aprendizaje con variable estado x cuando $t \rightarrow \infty$ y el comportamiento asintótico de una ecuación diferencial. Para las mujeres, la ecuación (4) puede aproximarse por:

$$\frac{dx}{dt} = E[w|x] - x$$

Para los hombres $m \in I_m$, la ecuación (5), se puede aproximar como:

$$\begin{bmatrix} \frac{dz}{dt} \\ \frac{dp}{dt} \end{bmatrix} = \begin{bmatrix} E[w_z|z] - z \\ E[w_p|p] - p \end{bmatrix}$$

Formalmente, se tiene la siguiente proposición.

Proposición 1 La dinámica continua puede ser expresada como:

$$\frac{dx^{wm}}{dt} = \tilde{\sigma}^{wm}(x^w, p, z)[v^m - x^{wm}] \quad \forall w \in I_w \quad (6)$$

$$\begin{bmatrix} \frac{dz^{mw}}{dt} \\ \frac{dp^{mw}}{dt} \end{bmatrix} = \begin{bmatrix} \sigma^{mw}(p^m, z^m) \tilde{\sigma}^{wm}(x^w, p, z)[v^w - z^{mw}] \\ \sigma^{mw}(p^m, z^m) [\tilde{\sigma}^{wm}(x^w, p, z) - p^{mw}] \end{bmatrix} \quad \forall m \in I_m \quad (7)$$

Demostración Proposición 1 Usando la definición de w , se puede computar el valor esperado $E(w|x)$.

$$E(w^{wm}|x) = \pi^{wm}v^m + (1 - \pi^{wm})x^{wm}$$

Usando lo definido anteriormente, con $\pi^{wm} = \tilde{\sigma}^{wm}(x^w, p, z)$ se tiene:

$$E(w^{wm}|x) = \tilde{\sigma}^{wm}(x^w, p, z)v^m + (1 - \tilde{\sigma}^{wm}(x^w, p, z))x^{wm}$$

Lo que si se remplacea en la dinámica, se obtiene la ecuación 6.

Para el caso de $m \in I_m$, se tiene lo siguiente:

$$\begin{aligned} E(w_z^{mw}|z) - z &= \pi^{mw}\pi^{wm}v^w + (1 - \pi^{mw})z^{mw} + \pi^{mw}(1 - \pi^{wm})z^{mw} - z^{mw} \\ &= \pi^{wm}\pi^{mw}(v^w - z^{mw}) \\ &= \sigma^{mw}(p^m, z^m)\tilde{\sigma}^{wm}(x^w, p, z)(v^w - z^{mw}) \end{aligned}$$

y

$$\begin{aligned} E(w_p^{mw}|p) - p &= \pi^{mw}\pi^{wm} \times 1 + (1 - \pi^{mw})p^{mw} + \pi^{mw}(1 - \pi^{wm}) \times 0 - p^{mw} \\ &= \pi^{mw}(\pi^{wm} - p^{mw}) \\ &= \sigma^{mw}(p^m, z^m)(\tilde{\sigma}^{wm}(x^w, p, z) - p^{mw}) \end{aligned}$$

que equivalen a la ecuación (7). ■

Lo que quiero estudiar es como se comportan los jugadores en el largo plazo. Se verá que los valores x, z, p del proceso de aprendizaje en el largo plazo van a converger a los puntos de reposo de las ecuaciones diferenciales.

Un punto de reposo es aquel en que $\frac{dx}{dt} = 0$, en donde la dinámica deja de cambiar. Para poder entender cómo son las creencias en el largo plazo, se definen los puntos de reposo de las ecuaciones diferenciales planteadas en la Propiedad 2.

Como $\sigma(z^m, p^m) > 0 \ \forall m \in I_m$ y tanto v^m como v^w son escalares, los puntos de reposo de las dos primeras ecuaciones diferenciales son los siguientes.

$$\begin{aligned} x^{wm} &= v^m \quad \forall w \in I_w \\ z^{mw} &= v^w \quad \forall m \in I_m \end{aligned}$$

Para el caso de la última ecuación, cuando β es lo suficientemente pequeño, $\tilde{\sigma}^{wm}(x^w, p, z)$ es contractante y se tiene un único punto fijo para la función $\tilde{\sigma}^{wm}(x^w, p, z)$, que corresponde al punto de reposo de la ecuación diferencial ya que $\sigma(z^m, p^m) > 0 \ \forall m \in I_m$.

$$p^{mw} = \tilde{\sigma}^{wm}(x^w, p, z) \quad \forall m \in I_m$$

Lema 2 Si $\beta \max v^w \leq 1$, entonces $p^{mw} = \tilde{\sigma}^{wm}(x^w, p, z) \ \forall m \in I_m$ existe y es único.

Demostración Lema 2 La demostración de existencia es directo del teorema de Brouwer. Utilizando el hecho de que $v^{ij} = v^j \ \forall i \in I$, y el que se puede establecer una jerarquía, se supondrá que los valores reales de los hombres son tales que $v^1 > v^2 > \dots > v^n$.

Para el caso del hombre con la valoración más alta, se puede calcular p^{1w} como:

$$p^{1w} = \sigma^{1w}(z^1, p^1)$$

Y para el hombre con la segunda valoración más alta, se tiene:

$$p^{2w} = \sigma^{2w}(z^2, p^2) \sigma^{1w}(z^1, p^1) = \sigma^{2w}(z^2, p^2) C$$

donde C es una constante que viene de la solución del problema anterior, y por lo tanto, $C \leq 1$. Para los siguientes se sigue de forma inductiva.

De esta forma, definiendo la función que sigue p^{mw} como $f_w : [0, 1]^w \rightarrow [0, 1]$ y se define $f : [0, 1]^n \rightarrow [0, 1]^n$, lo que por Brouwer me entrega existencia.

Respecto a la unicidad, se sabe que

$$\|f(x) - f(y)\| \leq \max_{i \in I_w} |f_i(x) - f_i(y)|$$

por teorema del valor medio y por la desigualdad de Cauchy-Schwarz, se tiene

$$|f_i(x) - f_i(y)| \leq |\nabla f_i(\xi)| |x - y|$$

De esta forma, una condición suficiente para que f sea una contracción es que:

$$|\nabla f_i(x)| \leq \delta \leq 1$$

Así, la derivada de $f_i(x)$ con respecto a $x^i = p^{mi}$ es:

$$\frac{\partial f_i(x^i)}{\partial p^{mi}} = \beta z^{mi} \frac{e^{\beta z^{mi} p^{mi}}}{\sum_{j \in I_w} e^{\beta z^{mj} p^{mj}}} \frac{\sum_{h \neq i} e^{\beta z^{mh} p^{mh}}}{\sum_{j \in I_w} \beta z^{mj} p^{mj}} C$$

Por lo tanto, la condición para contracción se cumple si $\beta \max v^w \leq 1$, y de esta forma, se tiene un único punto fijo. ■

Se definen $X = \begin{bmatrix} x \\ z \\ p \end{bmatrix}$ y $\bar{X} = \begin{bmatrix} \bar{x} \\ \bar{z} \\ \bar{p} \end{bmatrix}$, donde este ultimo es el vector de puntos de reposos de las ecuaciones diferenciales.

3.2. Convergencia asintótica de las dinámicas

Se esta buscando un atractor global de la dinámica continua a la que converge el proceso estocástico. Para esto, se sigue el procedimiento de Cominetti, Melo y Sorin (2010), que se basa en el hecho de que como los pagos están acotados, también lo está cualquier secuencia X_t generada por las dinámicas. Combinando Benaim (1999, prop 4.1 y 4.2) y el limit set theorem (Benaim, 1999), se tiene que el conjunto de puntos de acumulación del proceso aleatorio de tiempo discreto X_n es casi seguro un *internally chain transitive set* (ICT) para la dinámica determinística de tiempo continuo. Según Benaim (1999, Corollary 5.4), si (6) y (7) tienen un único punto de reposo \bar{X} que es un atractor global entonces es la única ICT y X_n converge casi seguro a \bar{X} .

Dado lo anterior, lo que se quiere encontrar es que las ecuaciones diferencias tienen un único punto fijo. Para esto, se quiere encontrar una función de Lyapunov que nos entregue el que el punto fijo es un atractor global. En particular, para esto no se necesita imponer cosas sobre las estrategias de los hombres, pero si sobre las estrategias de las mujeres.

Proposición 2 Si $\tilde{\sigma}^w : R^{2(|M| \times |W|) + |M|} \rightarrow (0, 1) \ \forall w \in I_w$ es una función de Lipschitz, entonces el único punto fijo \bar{X} es un atractor global para las dinámicas adaptativas y los procesos de aprendizaje convergen casi seguro a \bar{X} .

Demostración Proposición 2 Se quiere encontrar una función Lyapunov estricta con un único mínimo en \bar{X} . Se propone que $\Phi(X) = \|X - \bar{X}\|_\infty$ tiene esta propiedad.

Como $\Phi(X) = \|X - \bar{X}\|_\infty = \max_{is}(\|x - \bar{x}\|_\infty, \|z - \bar{z}\|_\infty, \|p - \bar{p}\|_\infty)$, se puede tener tres posibles casos. Uno en que el máximo este en los x, en los z o en los p. Se puede definir:

$$\Phi \begin{bmatrix} x(t) \\ z(t) \\ p(t) \end{bmatrix} = \max_{is} \{ \varphi_x(x(t)), \varphi_z(z(t)), \varphi_p(p(t)) \}$$

donde $\varphi_v(v(t)) = \|v - \bar{v}\|_\infty$

Se tiene que

$$\frac{d\Phi(X(t))}{dt} \in \left\{ \frac{d\varphi_x(x(t))}{dt}, \frac{d\varphi_z(z(t))}{dt}, \frac{d\varphi_p(p(t))}{dt} \right\}$$

Sea $w \in I$ y $m \in S^i$ talque $\Phi(X(t)) = |x^{wm} - \bar{x}^{wm}|$, entonces, si $x^{wm} \geq \bar{x}^{wm}$ se tiene que:

$$\frac{d[x^{wm}(t) - \bar{x}^{wm}]}{dt} = \tilde{\sigma}^{wm}(x^w, p, z)[h^m - x^{wm} - h^m + \bar{x}^{wm}] = -\tilde{\sigma}^{wm}(x^w, p, z)|x^{wm} - \bar{x}^{wm}|$$

Sea $m \in I$ y $w \in S^i$ talque $\Phi(X(t)) = |z^{mw} - \bar{z}^{mw}|$, entonces, si $z^{mw} \geq \bar{z}^{mw}$ se tiene que:

$$\begin{aligned} \frac{d[z^{mw}(t) - \bar{z}^{mw}]}{dt} &= \sigma^{mw}(z^m, p^m)\tilde{\sigma}^{wm}(x^w, p, z)[h^w - z^{mw} - h^w + \bar{z}^{mw}] \\ &= -\sigma^{mw}(z^m, p^m)\tilde{\sigma}^{wm}(x^w, p, z)|z^{mw} - \bar{z}^{mw}| \end{aligned}$$

Sea $m \in I$ y $w \in S^i$ talque $\Phi(X(t)) = |p^{mw} - \bar{p}^{mw}|$, entonces, si $p^{mw} \geq \bar{p}^{mw}$ se tiene que:

$$\frac{d[p^{mw}(t) - \bar{p}^{mw}]}{dt} = \sigma^{mw}(z^m, p^m)[\tilde{\sigma}^{wm}(x^w, z, p) - p^{mw} - \tilde{\sigma}^{wm}(\bar{x}^w, \bar{z}, \bar{p}) + \bar{p}^{mw}]$$

Sea $\ell \in [0, 1)$ una Lipschitz constant para $\sigma^w(\cdot)$, se tiene que:

$$\begin{aligned} \frac{d[p^{mw}(t) - \bar{p}^{mw}]}{dt} &= \tilde{\sigma}^{mw}(z^m, p^m)[\tilde{\sigma}^{wm}(x^w, z, p) - p^{mw} - \tilde{\sigma}^{wm}(\bar{x}^w, \bar{z}, \bar{p}) + \bar{p}^{mw}] \\ &\leq \sigma^{mw}(z^m, p^m)[\ell \left\| \begin{bmatrix} x^w \\ z \\ p \end{bmatrix} - \begin{bmatrix} \bar{x}^w \\ \bar{z} \\ \bar{p} \end{bmatrix} \right\|_\infty - p^{mw} + \bar{p}^{mw}] \end{aligned}$$

Como se esta en el caso en que $|p^{mw} - \bar{p}^{mw}|$ maximiza $\|X - \bar{X}\|_\infty$, se tiene que

$$\begin{aligned} \frac{d[p^{mw}(t) - \bar{p}^{mw}]}{dt} &\leq \sigma^{mw}(z^m, p^m)[\ell\|p - \bar{p}\|_\infty - p^{mw} + \bar{p}^{mw}] = -\sigma^{mw}(z^m, p^m)(1 - \ell)\varphi_p(p(t)) \\ &= -\sigma^{mw}(z^m, p^m)(1 - \ell)\Phi(X(t)) \end{aligned}$$

Se sigue el mismo camino para cuando $x^{wm} < \bar{x}^{wm}$, $z^{mw} < \bar{z}^{mw}$ o $p^{mw} < \bar{p}^{mw}$, y se tiene que

$$\frac{d\Phi(X(t))}{dt} \leq -\min_{is} \sigma^{is}(X^i(t))[1 - \ell]\Phi(X(t))$$

Como $v^m \forall m \in I_m$, $v^w \forall w \in I_w$ y $\sigma^{is}(\cdot)$ estan acotados y este último esta entre 0 y 1, lo mismo ocurre para $X(t)$, talque $\frac{d\Phi(X(t))}{dt} \leq -\epsilon\Phi(X(t))$, para algún $\epsilon < 0$. Esto implica que Φ es una función de Lyapunov que decrece a 0, y como \bar{X} es el único punto en que $\Phi(\bar{X}) = 0$, este es un atractor global de la dinámica adaptativa y los procesos de aprendizaje convergen casi seguro a \bar{X} . ■

Sin embargo, para esto la estrategia de las mujeres debe cumplir una condición. Esto es que $|\tilde{\sigma}^{wm}(x_1^w, p_1, z_1) - \tilde{\sigma}^{wm}(x_2^w, p_2, z_2)|$ debe estar acotada por $C \times \|X - \bar{X}\|_\infty$, donde $C \leq 1$ es una constante. Sin embargo, esto no es posible.

Lema 3 $\tilde{\sigma}^w : R^{2(|M| \times |W|) + |M|} \rightarrow (0, 1) \forall w \in I_w$ no es una función de Lipschitz.

Demostración Lema 3

$$\begin{aligned} |\tilde{\sigma}^{wm}(x_1^w, p_1, z_1) - \tilde{\sigma}^{wm}(x_2^w, p_2, z_2)| &= |\sigma^{mw}(z_1^m, p_1^m) \prod_{\{m' | x_1^{wm'} > x_1^{wm}\}} [1 - \sigma^{m'w}(z_1^{m'}, p_1^{m'})] \\ &\quad - \sigma^{mw}(z_2^m, p_2^m) \prod_{\{m' | x_2^{wm'} > x_2^{wm}\}} [1 - \sigma^{m'w}(z_2^{m'}, p_2^{m'})]| \end{aligned}$$

sumando y restando

$$\sigma^{mw}(z_1^m, p_1^m) \prod_{\{m' | x_2^{wm'} < x_2^{wm}\}} [1 - \sigma^{m'w}(z_2^{m'}, p_2^{m'})]$$

y usando desigualdad triangular, se tiene que:

$$\begin{aligned} |\tilde{\sigma}^{wm}(x_1^w, p_1, z_1) - \tilde{\sigma}^{wm}(x_2^w, p_2, z_2)| &\leq |\sigma^{mw}(z_1^m, p_1^m) \prod_{\{m' | x_1^{wm'} > x_1^{wm}\}} \sigma^{m'w}(1 - (z_1^{m'}, p_1^{m'})) \\ &\quad - \prod_{\{m' | x_2^{wm'} > x_2^{wm}\}} (1 - \sigma^{m'w}(z_2^{m'}, p_2^{m'}))| \\ &\quad + |\prod_{\{m' | x_2^{wm'} > x_2^{wm}\}} (1 - \sigma^{m'w}(z_2^{m'}, p_2^{m'})) [1 - \sigma^{mw}(z_1^i, p_1^i)] - (1 - \sigma^{mw}(z_2^i, p_2^i))| \end{aligned}$$

En el caso en que para las creencias (z_2, x_2, p_2) no se tiene una persona preferida por $w \in I_w$ a $m \in I_m$, y se tiene solo una para las creencias (z_1, x_1, p_1) , talque $z_1^{m'} \times p_1^{m'}$ sea tal que $\sigma^{m'w}(z_1^{m'}, p_1^{m'})$ sea cercano a 0, se tiene que:

$$|\sigma^{mw}(z_1^m, p_1^m) \prod_{\{m' | x_1^{wm'} > x_1^{wm}\}} \sigma^{m'w}(1 - (z_1^{m'}, p_1^{m'})) - \prod_{\{m' | x_2^{wm'} > x_2^{wm}\}} (1 - \sigma^{m'w}(z_2^{m'}, p_2^{m'}))| \leq |1 - \sigma^{mw}(z_1^m, p_1^m)| \leq 1$$

es tan cercana a uno como quiera y no puede ser acotada por $C \|X - \bar{X}\|$. ■

El problema viene dado por la discontinuidad que existe entre x_1 y x_2 , es decir, pequeños cambios en estos valores causan que el orden de las preferencias cambien fuertemente, causando que no se pueda acotar la diferencia entre las estrategias aunque las creencias sean muy cercanas. Dado que no se ha podido obtener un atractor global, se buscara un atractor local.

Proposición 3 Si $\tilde{\sigma}^w : R^{2(|M| \times |W|) + |M|} \rightarrow (0, 1) \forall w \in I_w$ es una función de Lipschitz cuando $\|x - \bar{x}\| < \varepsilon$, entonces el único punto fijo \bar{X} es un atractor local para las dinámicas adaptativas y los procesos de aprendizaje convergen casi seguro a \bar{X} .

Demostración Proposición 3 La demostración es equivalente a la demostración de la proposición 2. ■

Note que el problema era la pérdida en el orden en las creencias de las mujeres.

Sin pérdida de generalidad, se asume que $v^1 > v^2 > \dots > v^n$, y se denota $\varepsilon = \min \frac{|v^m - v^{m+1}|}{2}$. Suponga que

$$\|x - \bar{x}\| < \varepsilon$$

Cualquier mujer con creencias en esta vecindad va a preferir al hombre 1 sobre el 2, al 2 sobre el 3, y así sucesivamente. Quiero demostrar que en esta vecindad, tengo una función de Lyapunov.

Para esto, se debe demostrar que cuando los agentes en I_w no se equivocan al ordenar, $\tilde{\sigma}^w : R^{2(|M| \times |W|) + |M|} \rightarrow (0, 1) \forall w \in I_w$ es una función de Lipschitz.

Proposición 4 Si $nK\beta(1 + \max_{w \in I_w} v^w) \leq 1$, $\tilde{\sigma}^w : R^{2(|M| \times |W|) + |M|} \rightarrow (0, 1) \forall w \in I_w$ es una función de Lipschitz local.

Demostración Proposición 4 Centrándose en la diferencia que nos falta acotar

$$\left| \prod_{\{m' | x_1^{wm'} > x_1^{wm}\}} \sigma^{m'w}(1 - (z_1^{m'}, p_1^{m'})) - \prod_{\{m' | x_2^{wm'} > x_2^{wm}\}} (1 - \sigma^{m'w}(z_2^{m'}, p_2^{m'})) \right|$$

Cuando $m = 1$, la diferencia es

$$|1 - 0 - (1 - 0)| = 0$$

Entonces

$$|\tilde{\sigma}^{w1}(x_1^w, p_1, z_1) - \tilde{\sigma}^{w1}(x_2^w, p_2, z_2)| \leq |(1 - \sigma^{1w}(z_1^i, p_1^i)) - (1 - \sigma^{1w}(z_2^i, p_2^i))| = |\sigma^{1w}(z_1^i, p_1^i) - \sigma^{1w}(z_2^i, p_2^i)|$$

Cuando $m = 2$, se tiene

$$|(1 - \sigma^{1w}(z_1^1, p_1^1)) - (1 - \sigma^{1w}(z_2^1, p_2^1))| = |\sigma^{1w}(z_1^1, p_1^1) - \sigma^{1w}(z_2^1, p_2^1)|$$

entonces

$$|\tilde{\sigma}^{w2}(x_1^w, p_1, z_1) - \tilde{\sigma}^{w2}(x_2^w, p_2, z_2)| \leq |\sigma^{1w}(z_1^1, p_1^1) - \sigma^{1w}(z_2^1, p_2^1)| + |\sigma^{2w}(z_1^2, p_1^2) - \sigma^{2w}(z_2^2, p_2^2)|$$

Cuando $m = 3$, la diferencia es

$$\begin{aligned} & |(1 - \sigma^{1w}(z_1^1, p_1^1))(1 - \sigma^{2w}(z_1^2, p_1^2)) - (1 - \sigma^{1w}(z_2^1, p_2^1))(1 - \sigma^{2w}(z_2^2, p_2^2))| \\ & \leq |\sigma^{1w}(z_1^1, p_1^1) - \sigma^{1w}(z_2^1, p_2^1)| + |\sigma^{2w}(z_1^2, p_1^2) - \sigma^{2w}(z_2^2, p_2^2)| \end{aligned}$$

entonces

$$|\tilde{\sigma}^{w3}(x_1^w, p_1, z_1) - \tilde{\sigma}^{w3}(x_2^w, p_2, z_2)| \leq |\sigma^{1w}(z_1^1, p_1^1) - \sigma^{1w}(z_2^1, p_2^1)| + |\sigma^{2w}(z_1^2, p_1^2) - \sigma^{2w}(z_2^2, p_2^2)| + |\sigma^{3w}(z_1^3, p_1^3) - \sigma^{3w}(z_2^3, p_2^3)|$$

Finalmente, cuando $m = n$, la diferencia es:

$$\sum_{k=1}^{n-1} |\sigma^{kw}(z_2^k, p_2^k) - \sigma^{kw}(z_1^k, p_1^k)|$$

entonces

$$|\tilde{\sigma}^{wn}(x_1^w, p_1, z_1) - \tilde{\sigma}^{wn}(x_2^w, p_2, z_2)| \leq \sum_{k=1}^n |\sigma^{kw}(z_2^k, p_2^k) - \sigma^{kw}(z_1^k, p_1^k)|$$

Sea $\begin{bmatrix} z_3^k \\ p_3^k \end{bmatrix} \in \left(\begin{bmatrix} z_2^k \\ p_2^k \end{bmatrix}, \begin{bmatrix} z_1^k \\ p_1^k \end{bmatrix} \right)$, utilizando el teorema del valor medio, se tiene que:

$$\begin{aligned} |\nabla \sigma^{kw}(z_3^k, p_3^k) \begin{bmatrix} z_2^k \\ p_2^k \end{bmatrix} - \begin{bmatrix} z_1^k \\ p_1^k \end{bmatrix}| &= |\sigma^{kw}(z_2^k, p_2^k) - \sigma^{kw}(z_1^k, p_1^k)| \\ &\leq \left| \sum_{w=1}^n \sigma^{kw}(z_3^k, p_3^k)' \cdot \left(\begin{bmatrix} z_2^k \\ p_2^k \end{bmatrix} - \begin{bmatrix} z_1^k \\ p_1^k \end{bmatrix} \right) \right| \\ &\leq \sum_{w=1}^n |\sigma^{kw}(z_3^k, p_3^k)'| \left| \begin{bmatrix} z_2^k \\ p_2^k \end{bmatrix} - \begin{bmatrix} z_1^k \\ p_1^k \end{bmatrix} \right| \end{aligned}$$

Usando Coachy-Schwarz

$$\begin{aligned} &\leq \sqrt{\sum_{w=1}^n |\sigma^{kw}(z_3^k, p_3^k)|^2 \cdot \sum_{w=1}^n \left| \begin{bmatrix} z_2^k \\ p_2^k \end{bmatrix} - \begin{bmatrix} z_1^k \\ p_1^k \end{bmatrix} \right|^2} \\ &\leq \sqrt{\|\sigma^{kw}(z_3^k, p_3^k)'\|^2 \left\| \begin{bmatrix} z_2^k \\ p_2^k \end{bmatrix} - \begin{bmatrix} z_1^k \\ p_1^k \end{bmatrix} \right\|^2} \\ &\leq \|\sigma^{kw}(z_3^k, p_3^k)'\|_2 \left\| \begin{bmatrix} z_2^k \\ p_2^k \end{bmatrix} - \begin{bmatrix} z_1^k \\ p_1^k \end{bmatrix} \right\|_2 \\ &\leq K \|\sigma^{kw}(z_3^k, p_3^k)'\|_\infty \left\| \begin{bmatrix} z_2^k \\ p_2^k \end{bmatrix} - \begin{bmatrix} z_1^k \\ p_1^k \end{bmatrix} \right\|_\infty \end{aligned}$$

Donde, si $z_3^{kw} < p_3^{kw}$

$$\begin{aligned} \sup \sigma^{kw}(z_3^k, p_3^k)' &= \frac{d\sigma^{kw}(z_3^k, p_3^k)}{dz_3^{kw}} \\ &= \beta p^{kw} \pi^{kw} (1 - \pi^{kw}) \end{aligned}$$

y si $z_3^{kw} > p_3^{kw}$

$$\sup \sigma^{kw}(z_3^k, p_3^k)' = \frac{d\sigma^{kw}(z_3^k, p_3^k)}{dp_3^{kw}}$$

$$= \beta z^{kw} \pi^{kw} (1 - \pi^{kw})$$

De esta forma, en el primer caso, se tiene que

$$|\sigma^{kw}(z_2^k, p_2^k) - \sigma^{kw}(z_1^k, p_1^k)| \leq K\beta \left\| \begin{bmatrix} z_2 \\ p_2 \end{bmatrix} - \begin{bmatrix} z_1 \\ p_1 \end{bmatrix} \right\|_\infty$$

y en el segundo se tiene que

$$|\sigma^{kw}(z_2^k, p_2^k) - \sigma^{kw}(z_1^k, p_1^k)| \leq K\beta z^{kw} \left\| \begin{bmatrix} z_2 \\ p_2 \end{bmatrix} - \begin{bmatrix} z_1 \\ p_1 \end{bmatrix} \right\|_\infty$$

En el caso que interesa, $\left\| \begin{bmatrix} x \\ z \\ p \end{bmatrix} - \begin{bmatrix} \bar{x} \\ \bar{z} \\ \bar{p} \end{bmatrix} \right\|_\infty = |p - \bar{p}|$. Por lo tanto, se puede escribir lo anterior como:

$$|\sigma^{kw}(z^k, p^k) - \sigma^{kw}(\bar{z}^k, \bar{p}^k)| \leq K\beta z^{kw} \Phi(X(t))$$

y como para el caso en que $\max z^{kw} \geq \bar{z}^{kw}$

$$\max z^{kw} - \bar{z}^{kw} \leq \|p - \bar{p}\|_\infty$$

$$\max z^{kw} - \bar{z}^{kw} \leq 1$$

$$\max z^{kw} \bar{z}^{kw} \leq 1 + \bar{z}^{kw}$$

$$\max z^{kw} \leq 1 + \max \bar{z}^{kw}$$

entonces, se tiene que

$$|\sigma^{kw}(z^k, p^k) - \sigma^{kw}(\bar{z}^k, \bar{p}^k)| \leq K\beta(1 + \max_{w \in I_w} v^w) \Phi(X(t))$$

y

$$|\tilde{\sigma}^{wn}(x_1^w, p_1, z_1) - \tilde{\sigma}^{wn}(x_2^w, p_2, z_2)| \leq \sum_{k=1}^n |\sigma^{kw}(z_2^k, p_2^k) - \sigma^{kw}(z_1^k, p_1^k)| \leq nK\beta(1 + \max_{w \in I_w} v^w) \Phi(X(t))$$

■

Con el análisis anterior se obtiene el resultado de que existe convergencia local de las dinámicas estocásticas, cuando las creencias de las mujeres son tales que les permiten realizar la elección correcta. Es decir, hay convergencia cuando la jerarquía de las creencias de las mujeres genera que estas elijan efectivamente al hombre cuyo valor real es mayor.

Lo esencial es entender que son estas estrategias que se juegan en el límite. Es posible establecer una conexión con la estrategia de equilibrio en un juego estático perturbado jugado por los hombres. En la siguiente sección, se establece una conexión entre los puntos de reposo \bar{X} y el equilibrio de un juego perturbado.

4. Puntos Fijos y Juego Perturbado

Sea $\Omega = \prod_{m \in I_m} R^{s^i}$ el espacio de percepciones y $j^{mw} = \frac{\exp(\beta(z^{mw} \times p^{mw}))}{\sum_{a \in S^i} \exp(\beta(z^{ma} \times p^{ma}))} \quad \forall m \in I_m$ un mapeo que nos dice, dadas las percepciones, como se actuará. Luego, recordando que \bar{X} son los puntos de reposo, $j(\bar{X})$ indica como se está jugado cuando los hombres tiene las percepciones de estado estacionario. Sea G^m un juego perturbado en que el hombre m escoge su estrategia a jugar y obtiene pagos dependiendo de su decisión.

$$G^m(\pi) = \langle \pi^m, v \rangle - \frac{1}{\beta} \sum_{w \in I_w} \pi^{mw} [\ln \pi^{mw} - 1] \quad \forall m \in I_m, v \in (v^w)_{w \in I_w}$$

Este juego perturbado está siempre definido, y tiene un equilibrio en estrategias mixtas. Cuando los hombres deciden su estrategia a través de una regla logística, esto puede interpretarse como que en realidad los hombres juegan de forma racional el juego especificado anteriormente. De esta forma, solo los hombres enfrentan este juego perturbado ya que las mujeres elijen de manera miope.

Proposición 5 Si el mapa $\sigma^m(\cdot)$ está dado por una Logit $\forall m \in I_m$ y $\sigma^i(\cdot)$ es un vector $1 \times n$ compuesto de ceros y un uno en la posición m , talque $m = s^w$, entonces $\Sigma(\bar{X})$ es el conjunto de equilibrio de Nash del juego perturbado G , donde

$$G^m(\pi) = \langle \pi^m, v \rangle - \frac{1}{\beta} \sum_{w \in I_w} \pi^{mw} [\ln \pi^{mw} - 1] \quad \forall m \in I_m, v \in (v^w)_{w \in I_w}$$

Demostración Proposición 5 Una conocida caracterización de las probabilidades Logit está dada por:

$$\sigma^m(z^m \times p^m) = \operatorname{argmax}_{\pi^m \in \Delta^m} \langle \pi^m, z^m \times p^m \rangle - \frac{1}{\beta} \sum_{w \in I_w} \pi^{mw} [\ln \pi^{mw} - 1]$$

Si se busca el equilibrio para el juego perturbado, se tiene que:

$$\pi^m = \operatorname{arg máx} \langle \pi^m, v \rangle - \frac{1}{\beta} \sum_{w \in I_w} \pi^{mw} [\ln \pi^{mw} - 1] \quad \forall m \in I_m, v \in (v^w)_{w \in I_w}$$

Se sabe que $\Sigma(\varepsilon)$ está caracterizado por $\pi^m = \sigma^m(z^m \times p^m)$ y $z^m = v, v \in (v^w)_{w \in I_w}$. Remplazando en la solución dada por la Logit, se obtiene la solución del juego perturbado para $m \in I_m$.

Luego, el punto fijo es el conjunto de equilibrios de Nash para el juego perturbado. ■

5. Algoritmo Gale-Shapley

Manteniendo el mismo modelo, se puede realizar un cambio en el mecanismo de matching que utilizan los agentes. De esta forma, en vez de tratar con invitaciones simultáneas, se tiene que cada periodo los jugadores presentan un gente presenta un ranking S^i de preferencias sobre los agentes al otro lado del mercado. Usando estos rankings, una plataforma genera las uniones en base al algoritmo Gale Shapley de *Deferred Acceptance* talque:

$$G : \otimes_{m=1, \dots, n} S_{I_m} \times \otimes_{i=1, \dots, n} S_{I_w} \longrightarrow \{f : \{1, n\} \rightarrow \{1, n\}\}$$

con f una función biyectiva.

Lo que hace este algoritmo es lo siguiente (Gale y Shapley, 1962):

1. Cada hombre sin pareja invita a la mujer más alta en su ranking, entre las que no ha invitado antes.
2. Cada mujer acepta temporalmente al hombre más alto en el ranking entre aquellos que le enviaron invitaciones y su pareja actual, y rechaza a los demás.
3. Los hombres que fueron rechazados, vuelven al paso 1. Si ya no hay hombres rechazados sin mujeres a las que aun no hayan invitado, el algoritmo termina.

Las mujeres, tal como en el caso anterior, generan su ranking de manera miope, talque sus creencias sobre el valor de los hombres, que se definen como x^{wm} , son ordenadas de mayor a menor.

Por otro lado, los hombres van a estar experimentando. Para introducir la experimentación, estos agentes realizan un ranking de mayor a menor no sobre sus creencias sobre el valor de los agentes al otro lado del mercado, que se denominaron z^{mw} , si no sobre

$$\omega^{mw} = z^{mw} + \delta^{mw}$$

donde δ distribuye Gumbel con locación μ y escala β . Donde la media es $\mu + \beta\gamma$, γ es una Euler–Mascheroni constant, y la varianza es $\frac{\pi^2\beta^2}{6}$. Se escogen μ de tal forma que la media sea 0.

Por McFadden et al. (1973) se sabe que la fórmula de la regla logística necesariamente implica que la utilidad no observada distribuye siguiendo valores extremos, un ejemplo de ellos siendo la distribución Gumbel. Además, se tiene que la escala de esta distribución es equivalente al β que acompaña a la logit que se usa para el mecanismo de matching en las secciones anteriores.

Al final del periodo, cada agente observa el pago que es función del verdadero valor de la persona con la que salió más un ruido blanco. Como antes, se define el valor de salir con una persona s como v^s . La ganancia de la estrategia s en cada periodo es $g_t^i = v^s + \xi$, y una vez observada, cada mujer y hombre actualizan sus percepciones de la siguiente forma, respectivamente.

$$x_{t+1}^{wm} - x_t^{wm} = \begin{cases} \gamma_{t+1}(g_t^{mw} - x_t^{wm}) & \text{si } G(S_I)(w) = m \\ 0 & \text{si } G(S_I)(w) \neq m \end{cases}$$

$$z_{t+1}^{mw} - z_t^{mw} = \begin{cases} \gamma_{t+1}(g_t^{mw} - z_t^{mw}) & \text{si } G(S_I)(m) = w \\ 0 & \text{si } G(S_I)(m) \neq w \end{cases}$$

A diferencia del caso en que el mecanismo de matching se basaba invitaciones simultáneas, ya no se tienen creencias sobre la probabilidad de ser aceptado, ya que ahora siempre se termina uniéndose a alguien. Al usar el algoritmo dado por Gale y Shapley, los rankings, y por consiguientes las creencias sobre el valor de los otros jugadores, es lo importante.

6. Simulaciones

Al realizar simulaciones del caso original con invitaciones simultáneas, se eligen los siguientes valores para los distintos parámetros. Se supone que ξ distribuye como una Normal $(0, \sigma^2)$ y que existen 5 agentes en cada lado del mercado. Se asume que $\sigma^2 = 0,1$ y que el parámetro de la regla logit, β , es igual a 0.5. Se crean los valores reales de los agentes talque $v \in \{1, 2, 3, 4, 5\}$ y las creencias a priori son obtenidas de

forma aleatoria a partir de una uniforme con soporte $[0,5]$ para el caso de los valores de los agentes al otro lado del mercado y de una uniforme con un soporte $[0,1]$ para la probabilidad de ser aceptado. Sin pérdida de generalidad, se indexaran a los hombres y mujeres según la utilidad que entregan a la unión. De esta forma, la mujer 1 tiene un valor real de 1.

Además, en relación a γ , guiándome por Bravo (2016), se utilizan en las simulaciones $\gamma^{ij} = \frac{1}{\text{numero de uniones de } i \text{ con } j}$, y no $\gamma^{ij} = \frac{1}{t}$. La primera razón es que si no se realiza este cambio, la convergencia de las dinámicas se vuelve muy lenta, y segundo, parece más razonable que las personas actualicen sus creencias respecto a otra persona particular dependiendo de las veces en que ha salido con ella, y no el tiempo que lleva en el mercado.

6.1. Benchmark

Como benchmark, se utilizó el caso en que se tiene información completa sobre los valores reales de los agentes, pero se sigue usando una regla logit para definir las invitaciones realizadas, y por consiguiente las uniones. Se quiere comparar las probabilidades de ser aceptado y de invitar de convergencia, usando este caso de información completa, con el caso original en que se desconocen los valores reales de los agentes. Una característica de este benchmark, es que se converge al equilibrio del juego perturbado especificado en la sección 5, para ciertos valores de parámetros, en lo que respecta a la probabilidad de invitar y de ser aceptado de cada hombre. Es decir, la creencia sobre la probabilidad de ser aceptado, p , converge al equilibrio del juego perturbado.

A continuación se muestran los valores de equilibrio del juego perturbado.

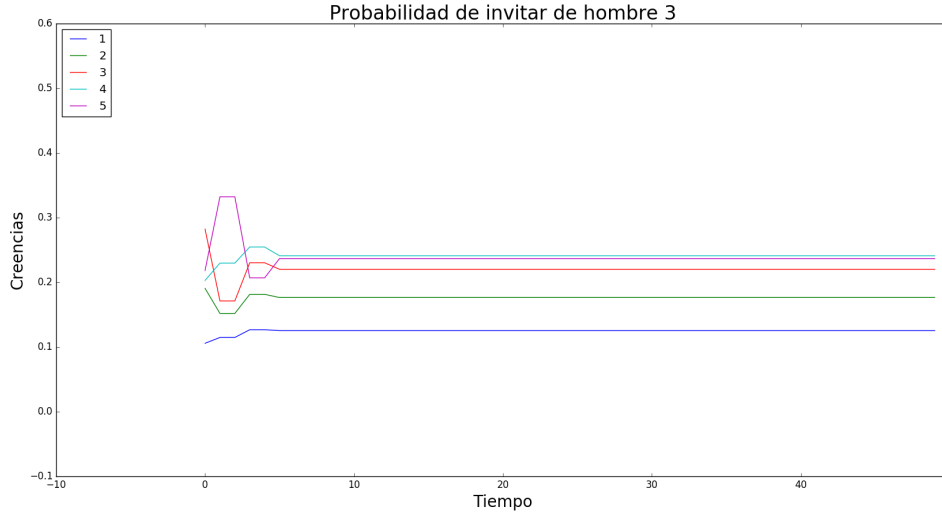
Hombre	Mujer	1	2	3	4	5
1		0.15921365	0.19350158	0.21197575	0.21824281	0.21706621
2		0.14493939	0.18794511	0.21593404	0.22662266	0.22455881
3		0.12541007	0.17652523	0.22015919	0.24111453	0.23679098
4		0.0990196	0.15273096	0.21871807	0.27160219	0.25792919
5		0.05801222	0.09564598	0.15769356	0.25999272	0.42865553

Cuadro 1: Probabilidad de hombre m de invitar a mujer w . $\beta = 0,5$

Hombre	Mujer	1	2	3	4	5
1		0.63469049	0.51238392	0.40238007	0.31635326	0.25092027
2		0.74227543	0.63097203	0.51319671	0.40905421	0.32358388
3		0.84871253	0.76623117	0.65807880	0.53901968	0.42397806
4		0.94198778	0.90435402	0.84230644	0.74000728	0.57134447
5		1	1	1	1	1

Cuadro 2: Probabilidad de hombre m de ser aceptado por mujer w . $\beta = 0,5$

Y donde las simulaciones entregan el siguiente gráfico para el hombre 3.

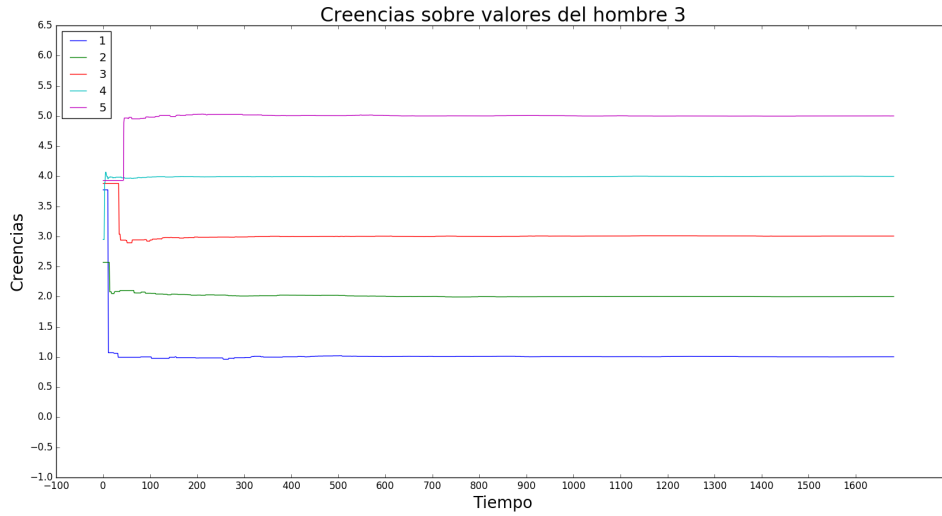


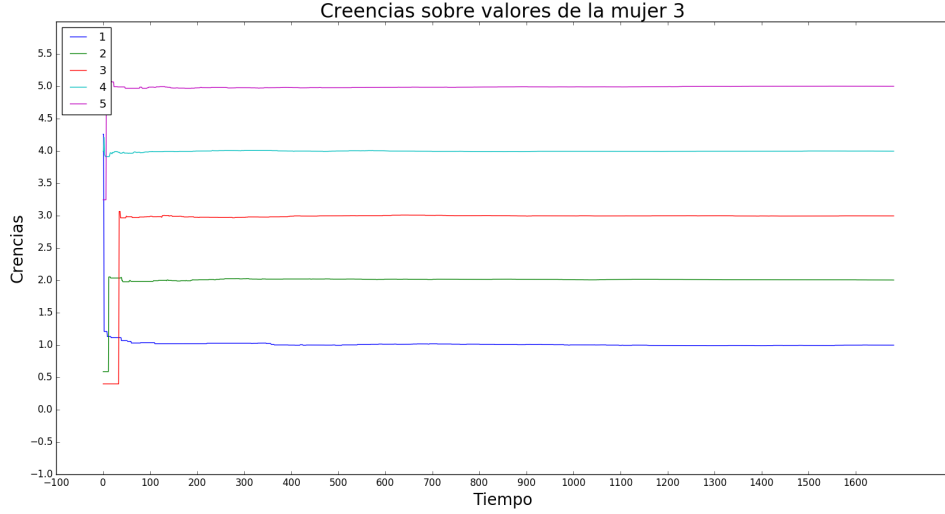
6.2. Invitaciones simultáneas

Ahora se observa el caso original en que no se conocen los valores reales de los agentes, y se tienen creencias sobre ellos.

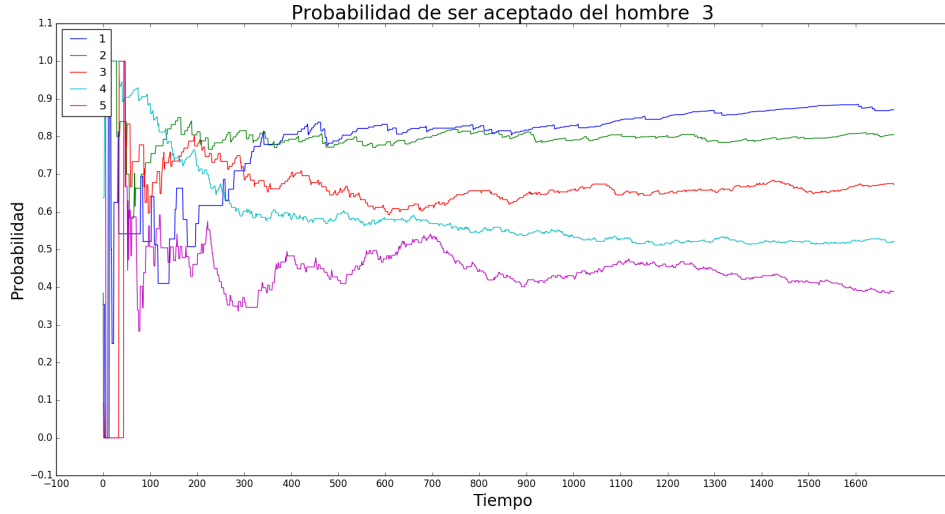
Se ve que para ciertos valores de β y σ^2 las creencias de las mujeres sobre los hombres, denotada por x y la creencias de los hombres sobre las mujeres, denotada por z convergen a sus valores reales.

De esta forma, cuando $\beta = 0,5$ y $\sigma^2 = 0,1$, se observa lo siguiente usando una regla de detención endógena con un criterio de convergencia de $|x_t - x_{t-1}| \leq 0,0001$ y $|z_t p_t - z_{t-1} p_{t-1}| \leq 0,0001$.





Note que convergencia se produce generalmente antes de las 100 iteraciones , sin embargo, $p \times z$ converge de forma más lenta, y dado que esto es lo que manda la estrategia de los hombres, es lo que importa que converja. La razón de esto es que la creencia sobre la probabilidad de ser aceptado se actualiza tanto cuando se es aceptado como cuando no, lo cual genera muchos más cambios en este valor que en los otros. Por ejemplo, el hombre 3 en la simulación presenta la siguiente probabilidad de ser aceptado en el tiempo:



Para el detalle de todos los agentes, refiérase al apéndice. Esta convergencia va disminuyendo en velocidad a medida que más alto es σ^2 , como se puede ver en la siguiente tabla

σ^2	Tiempo de Convergencia promedio ³
0.1	2021
0.2	3059
0.3	4087

Cuadro 3: Tiempos de convergencia para distintos σ^2 . $\beta = 0,5$

En términos de un primer mejor, dado que la utilidad que la sociedad obtiene de un match es aditiva, a un planificado central no le interesa que parejas se están formando.

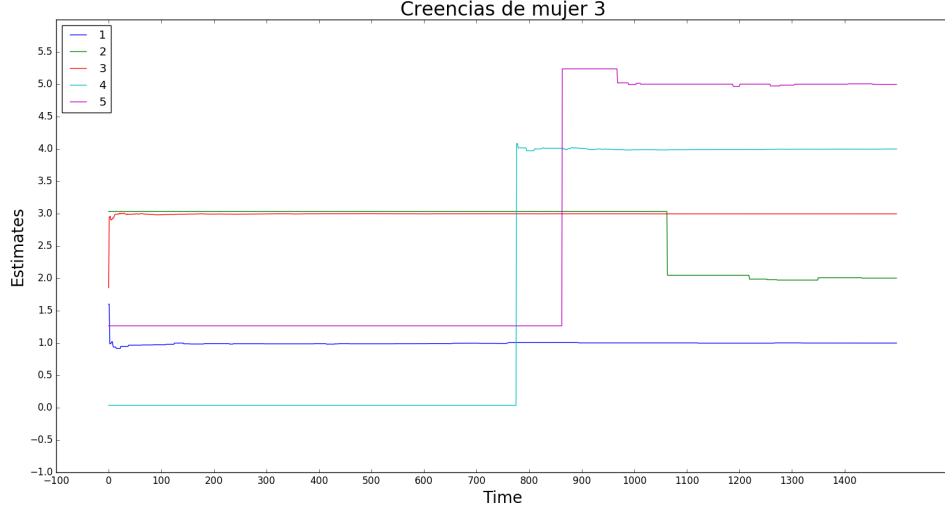
6.3. Algoritmo Gale-Shapley

Al realizar simulaciones del juego usando un mecanismo Gale-Shapley, tal como en el caso anterior se asume que ξ distribuye $normal(0, \sigma^2)$, y que existen 5 agentes. Además, se tiene que δ distribuye $Gumbel(\mu, \beta)$, con media 0 y $\beta = 0,5$.

Todos los otros valores son obtenidos de la misma forma que en el mecanismo anterior. Así, para los mismos parámetros, se tiene los siguientes gráficos para la convergencia de las creencias del hombre y la mujer 3. ⁴



⁴El resto de los gráficos se encuentran en el apéndice.



6.4. Comparación entre mecanismos

Note que si bien el algoritmo Gale-Shapley asigna mejor a los individuos, ya que en este ningún agente queda solo, a diferencia del mecanismo simultáneo, al cambiar el mecanismo de matching a uno en que se utiliza un algoritmo Gale-Shapley, bajo $\beta = 0,5$ y $\sigma^2 = 0,1$, se pierde la convergencia de las creencias sobre el valor de los agentes a sus valores reales. Una posible explicación es que tras cierto tiempo los rankings de los agentes quedan fijo, ya que la experimentación no es lo suficientemente fuerte - la varianza de la Gumbel no es la suficiente - como para cambiarlos. Esto último lleva a que sucedan cosas como la que se ven en el último gráfico.

7. Conclusión

Dentro de un contexto de two-side matching en donde los agentes deben experimentar unos con otros para poder aprender el verdadero valor de cada agente, la novedad de este trabajo es que para poder efectivamente experimentar, la persona al otro lado del mercado debe también querer experimentar con uno.

De esta manera, se introdujo un proceso adaptativo sencillo para las creencias sobre la utilidad de salir con un hombre o mujer, y la creencia sobre la probabilidad de ser aceptados de los hombres, que evoluciona de tal forma que la evolución de una variable depende de forma exclusiva del valor anterior de esta más un shock aleatorio, y se consideró una estrategia logit por parte de un lado del mercado para introducir experimentación.

Se concluye que al menos localmente, es decir, para cuando las mujeres ya no se equivocan en elegir, el proceso anterior converge a un estado estacionario, que coincide con los valores reales de las personas y donde la creencia sobre la probabilidad de ser aceptado, coincide con la probabilidad de aceptar de las mujeres. De esta forma, se encontraron parámetros bajo los cuales se logra la convergencia.

Si bien el resultado teórico es local - lo que deja para el futuro la tarea de un resultado global-, en las simulaciones se puede ver que existe una convergencia global a los valores reales. No solo eso, ya que al

realizarse simulaciones de los procesos, el equilibrio del juego perturbado coincide con la convergencia del benchmark, en que se mantiene la experimentación pero partiendo de los valores reales de los agentes. Finalmente, se realizó una comparación entre el modelo en que el mecanismo de matching son invitaciones simultáneas y en aquel en que se tiene un Gale y Shapley, y se ve que, al menos en los resultados de las simulaciones, pese a que el algoritmo Gale-Shapley es más eficiente (nunca hay nadie solo), las creencias no convergen a los valores reales para los mismos parámetros que en invitaciones simultáneas.

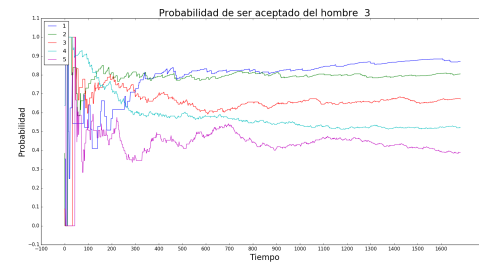
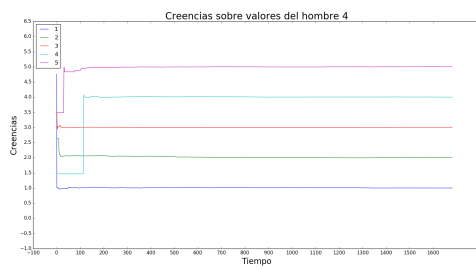
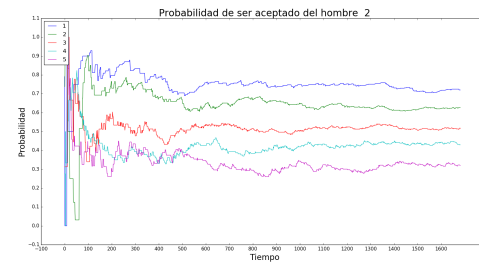
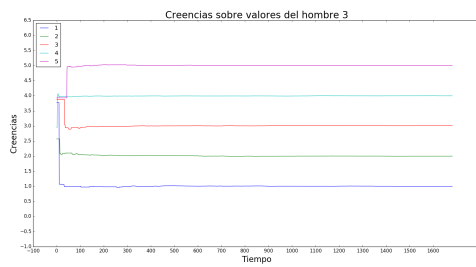
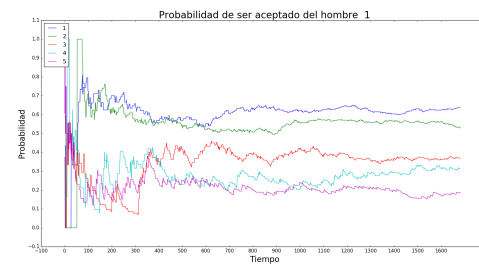
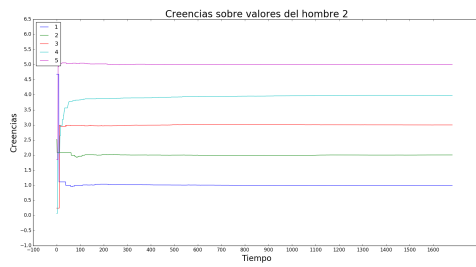
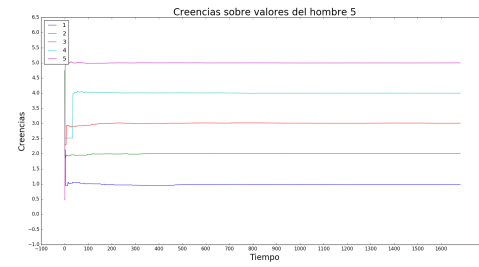
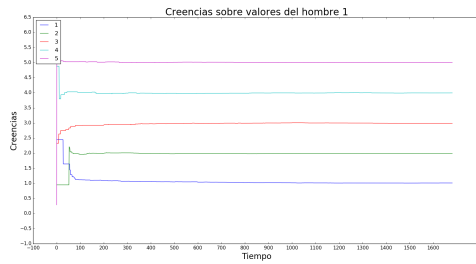
Referencias

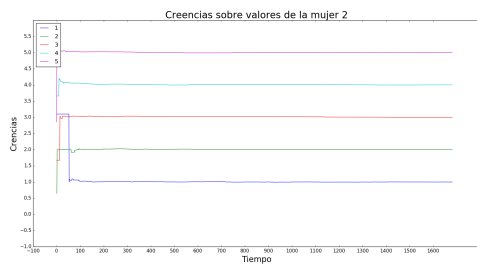
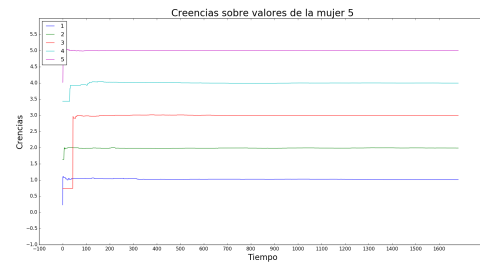
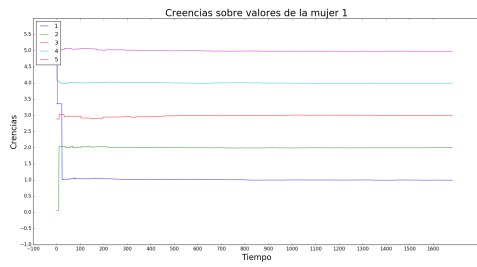
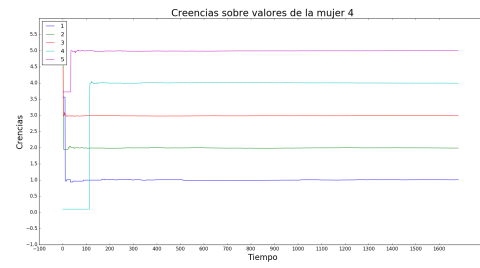
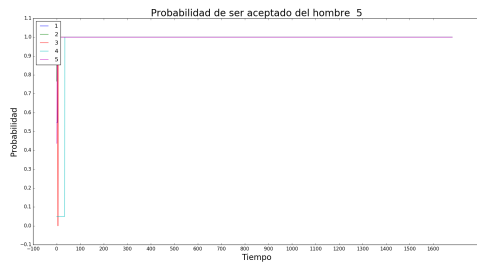
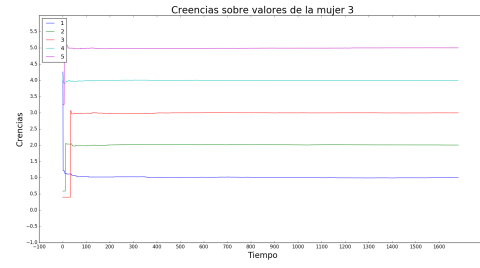
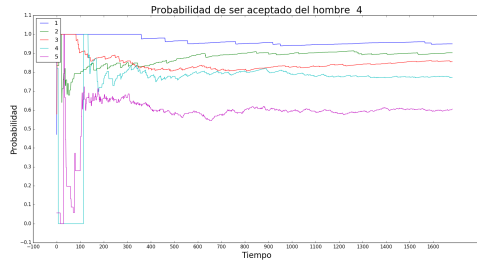
- Akbarpour, M., Li, S., y Gharan, S. O. (2014). Dynamic matching market design. *arXiv preprint arXiv:1402.3643*.
- Baccara, M., Lee, S., y Yariv, L. (2015). Optimal dynamic matching.
- Benaim, M. (1999). Dynamics of stochastic approximation algorithms. 1709.
- Benaim, M. y Hirsch, M. (1999). Mixed equilibria and dynamical systems arising from fictitious play in perturbed games. 29:36–72.
- Benaim, M., Hofbauer, J., y Sorin, S. (2005). Stochastic approximations and differential inclusions. 44.
- Bergemann, D. y Välimäki, J. (1996). Learning and strategic pricing. *Econometrica: Journal of the Econometric Society*, pages 1125–1149.
- Bergemann, D. y Välimäki, J. (2000). Experimentation in markets. *The Review of Economic Studies*, 67(2):213–234.
- Bergemann, D. y Valimaki, J. (2006). Bandit problems.
- Bravo, M. (2016). An adjusted payoff-based procedure for normal form games. *Mathematics of Operations Research*, 41(4):1469–1483.
- Cominetti, R., Melo, E., y Sorin, S. (2010). A payoff-based learning procedure and its application to traffic games. 70:71–83.
- Crawford, V. P. (1991). Comparative statics in matching markets. *Journal of Economic Theory*, 54(2):389–400.
- Das, S. y Kamenica, E. (2005). Two-sided bandits and the dating market. In *IJCAI*, volume 5, page 19.
- Fershtman, D. y Pavan, A. (2015). Matching auctions. *The Economist*.
- Fudenberg, D. y Levine, D. K. (1998). *The theory of learning in games*, volume 2. MIT press.
- Gale, D. y Shapley, L. S. (1962). College admissions and the stability of marriage. *The American Mathematical Monthly*, 69(1):9–15.
- Jovanovic, B. (1979). Job matching and the theory of turnover. *Journal of political economy*, 87(5, Part 1):972–990.
- McFadden, D. et al. (1973). Conditional logit analysis of qualitative choice behavior.

- Metz, J. M., Coyle, C., Hudson, C., y Hampshire, M. (2005). An internet-based cancer clinical trials matching resource. *Journal of medical Internet research*, 7(3).
- of Health, U. N. I. et al. (2012). Clinicaltrials. gov.
- Patel, C., Gomadam, K., Khan, S., y Garg, V. (2010). Trialx: Using semantic technologies to match patients to relevant clinical trials based on their personal health records. *Web Semantics: Science, Services and Agents on the World Wide Web*, 8(4):342–347.
- Robbins, H. (1952). Some aspects of the sequential design of experiments. *Bulletin of the American Mathematical Society*, 58(5):527–535.
- Roth, A. E. (1982). The economics of matching: Stability and incentives. *Mathematics of operations research*, 7(4):617–628.
- Roth, A. E. (1985). The college admissions problem is not equivalent to the marriage problem. *Journal of economic Theory*, 36(2):277–288.
- Roth, A. E. y Sotomayor, M. (1989). The college admissions problem revisited. *Econometrica: Journal of the Econometric Society*, pages 559–570.
- Roth, A. E. y Sotomayor, M. (1992). Two-sided matching. *Handbook of game theory with economic applications*, 1:485–541.
- Ünver, M. U. (2010). Dynamic kidney exchange. *The Review of Economic Studies*, 77(1):372–414.

8. Apéndice

8.1. Gráficos Mecanismo Simultáneo





8.2. Gráficos Mecanismo Gale-Shapley

