**Referee Report – Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness**

**Germán Camilo Rodríguez Perilla**[1](#page1)

gecrodriguezpe@unal.edu.co



Universidad Nacional de Colombia

Economía Experimental y del Comportamiento

Colombia

24 Mayo 2021

* Estudiante pregrado Universidad Nacional de Colombia

Economía experimental y del comportamiento - Referee Report Universidad Nacional de Colombia

**Índice**

|  |  |
| --- | --- |
| [**1. El artículo está motivado por el estudio de un comportamiento que contradice un**](#page3) |  |
| [**supuesto bastante común en economía. ¿Puede indicar cuál es este supuesto, cuál**](#page3) |  |
| [**sería el comportamiento en contradicción y cuáles ejemplos de dicho comporta-**](#page3) |  |
| [**miento hemos discutido en clase a partir de evidencia experimental?**](#page3) | **1** |

1. [**El artículo dice criticar algunas teorías que pretenden explicar el comportamiento**](#page4)[**enunciado en 1. ¿Cómo son llamadas estas teorías? ¿Cuál es el supuesto a través**](#page4)

|  |  |
| --- | --- |
| [**del cual son capaces de explicar este tipo de comportamientos?**](#page4) | **2** |

1. [**El artículo hace tres ejemplos de comportamientos que difícilmente podrían ser ex-plicados por las teorías enunciadas en 2. ¿Cuáles son? ¿Por qué no son compatibles**](#page5)

|  |  |
| --- | --- |
| [**con las teorías en 2?**](#page5) | **3** |

1. [**En la línea de base (tratamiento control), se usa una tarea que hemos aprendido a**](#page6)[**estudiar en clase, ¿cuál es? La versión usada es bastante diferente de la que hemos**](#page6)

|  |  |  |  |
| --- | --- | --- | --- |
|  | [**visto, ¿cuál es la ventaja (y desventaja) de utilizar esta versión?**](#page6) | | **4** |
| [**5.**](#page8) | [**¿Cuáles son los tratamientos (condiciones experimentales) introducidos en este**](#page8) | |  |
|  | [**experimento?**](#page8) | | **6** |
|  | [5.1.](#page9) | [Hidden information treatment](#page9) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 7 |
|  | [5.2.](#page9) | [Multiple dictator treatment](#page9) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 7 |
|  | [5.3.](#page9) | [Plausible deniability treatment](#page9) . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . | 7 |
| [**6.**](#page10) | [**En los procedimientos para llevar a cabo el experimento, ¿cuáles elementos se**](#page10) | |  |
|  | [**podrían mejorar? ¿Por qué?**](#page10) | | **8** |

1. [**Piense en una manera para resumir los resultados, de manera gráfica. Puede ser**](#page12)[**una Tabla o una Gráfica. Tiene que ser autocontenida, es decir, tener toda la**](#page12)[**información para que alguien que ha leído superficialmente el paper sea capaz de**](#page12)

|  |  |  |
| --- | --- | --- |
|  | [**interpretar los resultados.**](#page12) | **10** |
| [**8.**](#page13) | [**Intente explicar en un párrafo cuál es la lección fundamental que se puede sacar**](#page13) |  |
|  | [**de estos resultados.**](#page13) | **11** |
| [**9.**](#page13) | [**¿Por qué los autores no han usado una tarea como el ultimátum o el juego de**](#page13) |  |
|  | [**confianza?**](#page13) | **11** |

Economía experimental y del comportamiento - Referee Report Universidad Nacional de Colombia

1. **El artículo está motivado por el estudio de un comportamiento que contradice un supuesto bastante común en economía. ¿Puede indicar cuál es este supuesto, cuál sería el comportamiento en contradicción y cuáles ejemplos de dicho comportamiento hemos discutido en clase a partir de evidencia experimental?**

Dentro de la *teoría neoclásica*[2](#page3) que se ha estandarizado para modelar el comportamiento de los agentes que hacen parte de una economía hay dos supuestos fundamentales: 1) El supuesto de *self-interest*[3](#page3)enel que los agentes actúan de acuerdo a un interés personal que tengan y 2) el supuesto de *racionalidad* en dónde los agentes actúan y toman decisiones de tal forma que maximizan su interéspersonal[4](#page3).

Algunos textos, denotan el supuesto de racionalidad como el *principio de optimización* dónde las personas mediante un análisis de beneficios-costos toman las decisiones que maximicen su utilidad o función objetivo (Varian, [2014)](#page14). Respecto al supuesto de *self-interest*, se puede entender como es una *fuerza motivacional* en el comportamiento individual en dónde los agentes actúan para obtener unbeneficio personal (Medema, [2009)](#page14).

Ahora bien, el artículo de Dana et al. [(2007)](#page14) pretende investigar a mayor profundidad el *comportamiento de justicia/equidad*[5](#page3)enlas personas donde aparentemente los agentes muestran preocupaciónpor el bienestar de los demás, más allá de haber un riesgo reputacional o algún castigo. Es de interés comprender el comportamiento de justicia de las personas, dado que dicho comportamiento parece contradecir el supuesto de *self-interest* lo que indicaría que los agentes no solo actúan en pro de su bienestar individual sino que también actúan de tal forma de favorecer o aumentar el bienestar de los demás.

Si bien, a lo largo del semestre se han realizado una serie de experimentos y se han revisado un conjunto extenso de artículos relacionados con economía experimental en donde se han explorado múltiples juegos en donde hay interacción estratégica entre agentes, en lo personal, consideró que de toda la literatura estudiada y trabajada solo hay tres ejemplos de clase en donde hemos discutido el comportamiento de justicia/equidad.

Lo anterior se debe a que, si bien tanto en el *juego del últimatum* como en el *juego de confianza* se hablo de comportamiento de altruismo/equidad, una de las conclusiones a las que se llegó en ambos juegos es que la única manera de aislar dicho comportamiento pro-social de un comportamiento de racional estratégica por parte del jugador que hace el envío es mediante un *juego del dictador* en donde se elimine la posibilidad de que el jugador destinatario tenga la posibilidad de interactuar de manera estratégica en el juego[6](#page3). De ahí, que solo considere tres ejemplos de *comportamiento de* *justica/equidad* que se hayan estudiado de manera aislada en el curso dado que para lograr estudiardicho comportamiento de manera aislada es necesario emplear el juego del dictador. Los tres ejemplos han sido:

1. **Experimento del dictador que se realizó en clase**: En dicho experimento el jugador comodictador tenía que escoger un envío de entre 0 a 10, donde cero implicaba no realizar ningún envío positivo y 10 implicaba enviar todo. La *utilidad/pagos* de cada jugador dependía de la cantidad de *fichas/tokens* que cada jugador tuviera al final del juego.

* Dicha teoría, se ha vuelto la teoría dominante para entender y modelar el comportamiento de los agentes dentro de una economía. Sus principales resultados se basan en que los agentes se comportan como *Homo economicus* que toman decisiones como lo sugiere la *Rational choice theory*.

3Mal introducido al español como *egoismo*. Una mejor traducción, sería interés personal, que sería más cercana a lo que pretender capturar dicho supuesto.

4Dicho interés personal, generalmente es representado automáticamente mediante unas preferencias y su maximización se da mediante la maximización de un objetivo que suele ser representado mediante una función de utilidad (la función de utilidad es otra forma de representar las preferencias de un agente según la teoría estándar neoclásica.)

5Otras personas, lo podrían ver como comportamiento de generosidad o altruismo dependiendo de como se traduzca la palabra *fairness*

6Lo anterior, hace que el *juego del dictador* no sea estrictamente un juego dado que se ha eliminado la componente de interacción en éste.

1

Economía experimental y del comportamiento - Referee Report Universidad Nacional de Colombia

* 1. **Artículo de Forsythe et al.** [**(1994)**](#page14):En el artículo se buscar saber si exclusivamente a partir decomportamiento de justicia/equidad es posible explicar envíos positivos en juegos de negociación entre dos personas y para ello conduce tanto juegos del ultimátum como juegos del dictador. Lo que se da cuenta Forsythe et al. [(1994)](#page14) es que una desventaja de utilizar el juego del ultimatum para estudiar el efecto del comportamiento de justicia en los envíos positivos es que no puedo aislar dicho efecto de un comportamiento de racionalidad estratégico por parte del jugador que hacer el envío y para poder estudiar de manera aislada como el comportamiento de justicia afecta los envíos positivos de manera aislada requiere eliminar la interacción estratégica del jugador que recibe y para ello es que requiere implementar el juego del dictador.
  2. **Artículo de Hoﬀman et al.** [**(1996)**](#page14):El otro artículo que trabaja el tema comportamientode justicia/equidad en el posible envío de montos positivos en una situación donde haya dos personas/individuo es el artículo de Hoﬀman et al. [(1996)](#page14). Este artículo, pretende extender los resultados de la literatura hasta el momento, incluyendo los resultados del artículo de Forsythe et al. [(1994)](#page14). Los autores en dicho estudio, buscan ver como la manera en la que se implementa el juego del dictador podría afectar los resultados del juego y por ende los envíos realizados por parte del *dictador*. De esta forma, Hoﬀman et al. [(1996)](#page14) hacen una serie de modificaciones de tal forma que miran como el entorno del experimento determina los resultados. Ahí, exploran como el *framing* del experimento puede afectar los envíos, como la introducción de un doble ciego puede afectar los envíos y entre las principales conclusiones, además de que la forma en la que se conduzcan los procedimientos del experimento puede afectar los resultados, es que a medida que se da mayor anonimidad al *dictador* cada vez se observa menos ese comportamiento de justicia/equidad y se empieza a obtener resultados más cercanos a los esperados por el supuesto de *self-interest*.

1. **El artículo dice criticar algunas teorías que pretenden explicar el comportamiento enunciado en 1. ¿Cómo son llamadas estas teorías? ¿Cuál es el supuesto a través del cual son capaces de explicar este tipo de comportamientos?**

Al conjunto de teorías que pretenden explicar dicho comportamiento de justicia que aparentemente contradice el supuesto de *self-interest* se les conoce por *teorías de preferencias sociales*[7](#page4). Dichas teorías pretenden capturar dicho comportamiento de justicia asumiendo los resultados reflejan preferencias por resultados equitativos o de bienestar social. Entre dichas teorías, destacan Dana et al. [(2007):](#page14)

Las teorías donde la gente comparte con otros porque tienen una utilidad creciente con los pagos de los demás.

Las teorías donde son adversos a ventajas en la diferencia de los pagos

Las teorías donde los agentes quieren maximizar los pagos sociales totales o el pago más bajo a cualquier contraparte.

Es decir, las *teorías de preferencias sociales* lo que intentan decir es que el comportamiento de justicia/equidad que parece verse en muchos experimentos y en trabajos de campo se debe más que todo a que las personas tienen preferencias pro-sociales en el sentido que prefieren resultados que consideren justos o equitativos.

El supuesto que asumen todos los modelos basados en *teorías de preferencias sociales* es que las preferencias de un *dictador* pueden ser caracterizadas considerando solo las distribuciones finales de riquezas (Dana et al., [2007)](#page14). Es decir, el comportamiento de justicia/equidad puede ser comprendido en su totalidad a partir de los resultados de la interacción[8](#page4) y que dichos resultados lo que muestran es que las personas tienen una preferencia por resultados equitativos o justos.

7En inglés serían *social preference theories*

* vistos éstos como montos monetarios en el caso de un experimento

2

Economía experimental y del comportamiento - Referee Report Universidad Nacional de Colombia

1. **El artículo hace tres ejemplos de comportamientos que difícil-mente podrían ser explicados por las teorías enunciadas en 2. ¿Cuáles son? ¿Por qué no son compatibles con las teorías en 2?**

No obstante, Dana et al. [(2007)](#page14) reconocen que otros motivos para dar pueden no capturarse adecuadamente bajo el supuesto de que los *pagos monetarios por si solos recogen el comportamiento* *de justicia de las personas*[9](#page5).

Frente a lo anterior, los autores reconocen que las personas pueden sentir la necesidad de dar en algunas circunstancias, a pesar de que prefieran el resultado de maximizar su pago personal, porque no quieren aparecer *egoístas*, ya sea consigo mismos o con los demás. Así, la motivación subyacente que mueve una parte importante del comportamiento de *justicia/equidad* en los agentes puede ser *interés personal*[10](#page5),junto a un deseo de mantener la ilusión de no ser egoísta (Dana et al., [2007)](#page14).

Los autores, proveen tres ejemplos en dónde es evidente que las *teorías de preferencias sociales* no son capaces de explicar por si mismas:

1. Un *dictador* preferiría no conocer las consecuencias de sus acciones, en lo posible, para no sentirse obligado/forzado a dar[11](#page5).
2. De igual forma, las personas podrían no sentirse obligados a donar en presencia de otros *dictadores* que potencialmente podrían donar a cualquier destinatario[12](#page5).
3. El otro ejemplo, tiene lugar cuándo un dictador puede explotar la incertidumbre sobre que, precisamente, genera el resultado injusto. Una persona que envia en circunstancias *normales* de un juego de dictador, podría no hacerlo y actuar de manera más auto-interesado en caso de que pudiera camuflar/justificar su ausencia de donación a circunstancias inciertas o fuera de su control[13](#page5).

Ahora bien, ninguno de los comportamientos anteriores podrían ser explicados por las *teorías de* *preferencias sociales* dado que estas teorías se basan en los resultados que surge de la interacción. Sise considera solo preferencias sobre los resultados asociados a dar o no dar, todos los tres ejemplos anteriores no podrían ser evaluados mediante un juego estándar del dictador (Dana et al., [2007)](#page14). La falta de compatibilidad entre los ejemplos anteriores y las *teorías de preferencias sociales* radica en el hecho de que las últimas teorías se justifican exclusivamente en los outcomes para argumentar la existencias de preferencias pro-sociales o pro-bienestar pero dichas teorías no tienen en cuenta el contexto en el que se dio la interacción, y por ende no analizan el entorno en el que el *dictador* decidió entregar al destinatario. Como se puede observar, todos los tres ejemplos anteriores, son situaciones en donde el contexto en dónde ocurre la entrega cambia ligeramente y en dónde no es *transparente* la relación entre las acciones del *dictador* y los posibles *resultados*. Ante dichas situaciones, en dónde existen formas en las que el dictador podría justificar no donar sin sentir culpa de ser egoísta ante sí mismo o ante los demás podría decidir no hacerlo y esa posibilidad no es contemplada en las *teorías de* *preferencias sociales* dado que se concentran solo en los resultados de la interacción y no del contextoen el que tuvo lugar.

* Bajo las teorías mencionadas anteriormente, la presencia de comportamiento de *justicia/equidad* se observa es direc-tamente en los resultados de la interacción, es decir, si el dictador decidió enviar o no, pero no toma en consideración bajo que circunstancias lo hizo (no tiene en cuenta el contexto en el que el dictador decidió hacer la donación y la influencia que éste podría tener en su decisión al solo tener en cuenta dichas teorías los pagos monetarios finales.)

10Self-interest

11Por ejemplo, un donante de medula osea que donaría si supiera que es compatible con alguien que requiera la donación pero que intentaría no estar en ningún registro para evitar enterarse que puede donar

12Por ejemplo, a la hora de manejar al lado de un persona con el vehículo barado, uno podría justificar no ayudar con el pretexto de que alguien más va a proveer la ayuda

13Por ejemplo, un CEO podría actuar de tal forma que maximizara el valor/utilidad de los propietarios de la empresa cuándo sea evidente que existe una relación entre sus acciones y los resultados sobre las empresas pero en caso en que existan fuerzas o condiciones externas que puedan afectar el resultado, como fuerzas de mercado, podría utilizar dichas condiciones externas para llevar a cabo comportamiento autointeresado a expensas de los propietarios o trabajadores de la empresa

3

Economía experimental y del comportamiento - Referee Report Universidad Nacional de Colombia

Ahora bien, la contradicción entre las *teorías de preferencias sociales* y los ejemplos anteriores radica en que para estas teorías las personas tienen unas preferencias por los resultados justos o equitativos o prosociales. Es decir, las personas intrínsecamente buscan resultados justos y por ende prefieren dichas situaciones a situaciones menos justas o más inequitativas. No obstante, cada uno de los tres ejemplos anterioes muestra que a pesar de que los sujetos podrían estar eligiendo dichas opciones o resultados que son más justos o equitativos deciden no hacerlo a favor de beneficios personales por lo que se entraría en contradicción con las *teorías de preferencias sociales* dado que en los ejemplos es evidente que las personas no están prefiriendo los resultados justos o equitativos sobre los injusto o inequitativos de beneficio personal.

1. **En la línea de base (tratamiento control), se usa una tarea que hemos aprendido a estudiar en clase, ¿cuál es? La versión usada es bastante diferente de la que hemos visto, ¿cuál es la ventaja (y desventaja) de utilizar esta versión?**

La tarea que se usa en la línea base, es decir, en el tratamiento control del experimento, fue el de un *juego del dictador* en donde un *dictador* hace una división de unas dotaciones entre él y un *destinatario anónimo* que debe aceptar la división (Dana et al., [2007)](#page14). La característica principal de este *juego*[14](#page6), es que no hay una situación de interacción estratégica entre los jugadores dado que un jugador es el que decide y el otro jugador es el que acepta sin posibilidad de interactuar de manera estratégica.

La versión que usan en el *juego del dictador* para la línea base del experimento del artículo, es la de una *versión binaria* del juego del dictador en dónde el *dictador* tiene dos opciones de elección, de ahí que sea una versión binaria del juego. La *opción A* es una opción que maximiza los pagos del *dictador* en el sentido de que ofrece el mayor pago posible a éste pero con la consecuencia de ofrecerel menor pago posible para su contraparte. Por otro lado, la *opción B* es una opción que genera una distribución equitativa de pagos en el sentido que tanto el *dictador* como su contraparte reciben ambos el mismo pago pero con la característica que dicho pago es menor al pago que recibiría el dictador si decidiera jugar A y al mismo tiempo sería un pago mayor al pago que recibiría el destinatario si el dictador decidiera jugar A. Por otro lado, la suma de los pagos de ambos jugadores en la *opción A* es menor a la suma de los pagos de ambos jugadores en la *opción B* por lo que el óptimo de Pareto para este juego se encontraría en la opción B, siendo ésta la *situación eficiente* del juego.

La representación gráfica del juego está dada por la figura [1](#page7) en donde se resume el párrafo anterior. Ahora bien, en clase, realizamos un experimento del dictador donde un *dictador* tenía que escoger un envío de entre 0 a 10, donde cero implicaba no realizar ningún envío positivo y 10 implicaba enviar todo. La utilidad/pagos de cada jugador dependía de la cantidad de fichas/tokens que cada jugador tuviera al final del juego. Como se puede observar, mientras que en la línea base del experimento del artículo el dictador solo tenía dos opciones en la versión del juego del dictador realizada en clase le dictador tenía un conjunto discreto pero de más de opciones para el envío de pagos al jugador

destinatario.

Entre las ventajas de utilizar una versión binaria del *juego del dictador* para la línea de base del experimento se encuentra que:

Hay una diferencia significativa y evidente para el dictador entre elegir la *opción A* y elegir la *opción B*. Si el dictador decide elegir la opción A sabe que estaría maximizando su pago a costo del pago del jugador destinatario que obtendría un mínimo pago en dicha situación. Por el contrario, si el dictador elige la *opción B* puede notar que dicha opción le genera un pago menor pero que aumenta el pago de su contraparte y además puede notar que es una situación equitativa en el sentido de que ambos jugadores obtendrían el mismo pago. Esa diferencia notoria entre los dos resultados posibles del juego hace más fuerte ese *moral wiggle* en el sentido de que para el

1. Nuevamente, estrictamente el juego del dictador no es un juego, dado que no tiene una situación de interacción estratégica entre jugadores.

4

Economía experimental y del comportamiento - Referee Report Universidad Nacional de Colombia

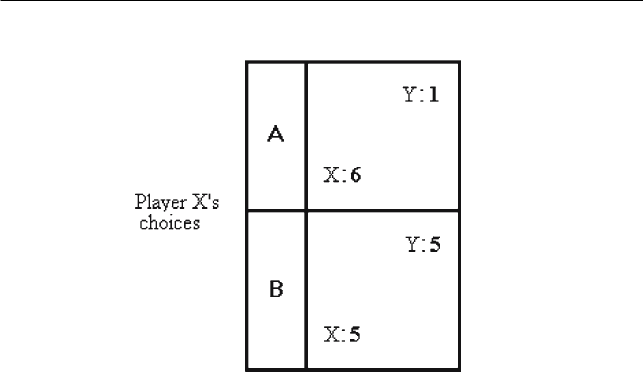


Figura 1: Representación gráfica de la versión binaria del juego del dictador empleado en la línea base del experimento

dictador es más evidente que tiene que hacer una elección entre una opción que lo favorece más frente a una opción más justa pero que le ofrece un menor pago. Lo importante, es que bajo la versión binario del juego del dictador ese dilema moral es más fuerte y evidente para el dictador.

De igual forma, el experimento además de contener una línea base, contiene unos tratamientos que resultan ser modificaciones de la línea de base. Utilizar juegos del dictador binarios, *facilita* la comparación entre tratamientos y permite realizar una lectura homogénea de los resultados. En el caso de un juego del dictador con más opciones, a pesar de que dicha comparación podría hacerse, sería más complicada, en especial en los tratamientos: *multiple dictator treatment* y *plausible deniability treatment*. En estos dos últimos tratamientos, dejaría de ser claro ese *moral wiggle* que intentan explorar los autores del artículo.

También, agregar más opciones complica el diseño del experimento, en particular en los trata-mientos adicionales, y tendría la de ventaja de complicar los cálculos que tiene que hacer los jugadores dado que ya no tienen que pensar en dos opciones sino en 10 o más opciones[15](#page7). Por ejemplo, el diseño del *tratamiento con cutoﬀ* se complicaría mucho en el sentido que habría que pensar muy bien para que no se confunda haber perdido la posibilidad de elegir por el punto de corte ya sea porque el jugador decidió voluntariamente que a si fuera o porque simplemente los cálculos eran muy complejos para el jugador y no tuvo el tiempo suficiente para decidir.

Finalmente, agregar más resultados complica el análisis de resultados porque no es lo mismo analizar que los jugadores en cierto tratamiento pasen de elegir la *opción A* a elegir la *opción B* a que ciertos jugadores pasen a elegir la opción 5 a la opción 6 y otros jugadores de la opción 4 a la opción 7 en el mismo tratamiento. Agregar múltiples opciones complica en el análisis y la inferencia de conclusiones.

Entre las desventajas de utilizar una versión binaria del *juego del dictador* para la línea de base del experimento se encuentra que:

La *opción B* que le da pagos iguales a ambos jugadores, y en particular que le de pagos de 5 a cada uno, podría ser una asignación focal ya sea por su simetría o porque 5 es la mitad de 10.

1. Peor aún si hubiera un continuo de opciones.

5

Economía experimental y del comportamiento - Referee Report Universidad Nacional de Colombia

En todo caso, podría llamar la atención del jugador, lo que haría que la elección de la *opción* *B* no fuera por un comportamiento de justicia/equidad sino porque dicha opción fuera focalvisualmente.

Los pagos para ambas opciones posibles son todos positivos lo que puede presentar problemas en la medida que si alguien ve un pago de uno, así sea poco, sentiría que está aportando algo al otro jugador. En el caso en que se incluyeran pagos nulos, el efecto del *moral wiggle* sería mucho más fuertes porque el *dictador* sería mucho más consciente que en una de las dos opciones que debería escoger el pago que el está dando al otro jugador es de cero, por lo que el conflicto moral sería más fuerte si se incluyera pagos nulos.

Otra desventaja de los pagos escogidos, es que en la opción B, al ofrecer un pago mayor para el jugador destinatario y además ser la opción con un valor total de la suma de ambos pagos mayor que en la opción A, en caso tal de que el dictador decidiera escoger la opción B no se podría estar seguro si el dictador escoge dicha opción ya sea porque aumentar el pago del otro jugador y por ende lo estaría beneficiando o porque es la situación que ofrece el mayor pago en total para la *sociedad* y por ende la razón fuera buscar el óptimo de Pareto en el juego. Por tanto, podría haber una confusión en los verdaderos motivos para que el dictador escogiera la *opción B*.

Finalmente, y de pronto la mayor desventaja de escoger un juego del dictador binario es que puede que el comportamiento de justicia/equidad no sea binario como lo pretende este juego. Es decir, puede que el comportamiento de justicia/equidad no sea que la persona es justa o no lo es sino que puede que exista una escala de comportamiento de justicia en donde algunas personas sean más justas que otras pero en donde puede que haya diferentes niveles en dicho comportamiento de justicia. Lo anterior es importante, dado que un juego del dictar binario podría no capturar esos grados/niveles/matizes de comportamiento de justicia/equidad al clasificar a las personas como justo o no justo cuando en realidad puede que las personas no sean completamente justas pero si tengan cierto grado de comportamiento de justicia a la hora de actuar y que no se este capturando adecuadamente por dicho diseño de juego del dictador.

1. **¿Cuáles son los tratamientos (condiciones experimentales) intro-ducidos en este experimento?**

La línea base, es decir, el tratamiento control para el experimento que resultó ser un juego del dictador estándar[16](#page8) tiene la característica de ser *transparente* en el sentido de haber una *relación uno-a-uno* entre las acciones del dictador y los resultados para ambas partes de la interacción. Bajo esassituaciones, es posible que se forcé a las personas a tener comportamiento de justicia/equidad (Dana et al., [2007)](#page14).

No obstante, como lo mostraron los ejemplo del punto 3, es posible que hayan situaciones, como aquellos 3 ejemplos, en donde dicha transparencia se pierda y pueda que los agentes puedan dejar de sentirse forzados a actuar con comportamiento de justicia/equidad y por el contrario empiecen a actuar más acorde a sus propios intereses. Para comprobar lo último, y ver cuál es el efecto de la transparencia entre decisión del dictador y resultado de la interacción en los comportamiento de justicia, los autores diseñan tres variaciones del juego del dictador de tal forma que cada variación se asocie y pueda representar cada uno de los ejemplos planteados en el punto 3 de tal forma que puedan ver el efecto que tiene la transparencia en el comportamiento de justicia y así tener una mejor compresión de este último comportamiento[17](#page8).

Los tres tratamientos que resultaron de modificaciones al juego del dictador estándar empleado en la línea base fueron: 1) *hidden information*, *multiple dictator* y *plausible deniability*. A continuación se da una descripción más profunda de cada uno.

1. Versión binaria para el caso de este experimento
2. Es importante resaltar, que cada una de los tratamientos son variaciones de un juego del dictador binario

6

Economía experimental y del comportamiento - Referee Report Universidad Nacional de Colombia

**5.1. Hidden information treatment**

En este tratamiento, la idea es que el *dictador* permanezca ignorante sobre las consecuencias precisas sobre los destinatarios. Por tanto, los pagos para el dictador son conocidos y son los mismos que en la línea base pero los pagos a los destinatarios son desconocidos por el dictador. Lo interesante del diseño, es que los pagos al destinatario están aleatorizados y pueden ser los mismos pagos que en el experimento de línea base o pueden estar *invertidos*[18](#page9) (Dana et al., [2007)](#page14).

Dentro del diseño, se les aviso a los participantes que los verdaderos pagos no serían revelados públicamente, pero que el dictador tendría la opción de revelar los pagos del destinatario con un click si optaría por hacer eso y sin costo. La decisión del dictador de revelar o no revelar se mantendría privada del jugador destinatario. El jugador desetinatario se le dió la tarea de hace una elección hipótetica de los pagos (usando la matriz basada en los pagos del exprimento de línea base).

Ahora bien, este tratamiento tiene la ventaja de que permite ver si un dictador preferiría no conocer las consecuencias de sus acciones, en lo posible, para no sentirse obligado/forzado a dar. Es decir, este primer tratamiento funciona como un experimento que busca testar el primer ejemplo de los tres mencionados en el punto 3. Esto es así, dado que si fuera cierto que entregar en el juego del dictador de la línea base reflejará preferencias por una distribución equitativa, entonces la proporción de dictadores que darían en la línea base debería ser igual a la proporción de dictadores que revela sus pagos verdaderos y escoge la acción más equitativa en el *hidden information treatment*. No obstante, si los dictadores buscaran una escusa para no sentirse obligados a dar, entonces espeariamos que escogieran permanecer no informados y escoger la opción de mayor pago, es decir la opción A, permaneciendo ignorantes sobre los posibles pagos del jugador destinatario (Dana et al., [2007)](#page14).

**5.2. Multiple dictator treatment**

En este tratamiento, se agrega un segundo dictador como modificación al juego del dictador del juego base, así eliminando la responsabilidad a un dictador individual de ser el único responsable del resultado injusto/inequitativo. No obstante, cualquier de los dos dictadores podría implementar independientemente el resultado justo/equitativo. A todos los jugadores se les informó, que el pago dependería de las decisiones conjuntas de los jugadores X y Y que resultaron ser los dictadores. Mientras los jugadores X y Y hacían sus elecciones, a los jugadores Z, jugador destintario sin posibilidad de elegir, se le asignaba la tarea de indicar cuál era la opción que pensaba había sido seleccionado la mayoría de veces por la mayoría de jugadores (Dana et al., [2007)](#page14).

Este tratamiento tiene la ventaja de que permite ver si las personas podrían no sentirse obligados a donar en presencia de otros dictadores que potencialmente podrían donar a cualquier destinatario. Es decir, este segundo tratamiento funciona como un experimento que busca testar el segundo ejemplo de los tres mencionados en el punto 3. Ello se debe a que, dado que ambos dictadores tiene que escoger la opción A, la opción inequitativa, para obtener resultados injustos/inequitativos en el juego, la adición de un segundo dictador no hace nada para impedir a los sujetos asegurar un resultado justo/equitativo si así lo prefirieran por el hecho de que si alguno de los dos dictadores escogiera B, entonces podrían imponer el resultado equitativo/justo. De lo anterior, se infiere que si las *teorías* *de preferncias sociales* fueran correctas, entonces se esperaría la misma proporción de elecciones Ben este tratamiento que en el tratamiento base. Lo interesante de este tratamiento, es que se rompe la transparencia, la acción auto-interesada de escoger A ya no garantiza el resultado injusto para el jugador pasivo. Así, al escoger A, un jugador estratégico puede permitirse la opción auto-interesada de que ocurra, sin haberla implementado directamente (Dana et al., [2007)](#page14).

**5.3. Plausible deniability treatment**

En el último tratamiento, se permite al dictador la posibilidad de perder la posibilidad de elegir, así se permite que sea plausible obtener resultados por causas distintas a las acciones del dictador. En este tratamiento, se le agregó al tratamiento base una línea de corte/umbral[19](#page9). Los sujetos, se les

1. Nuevamente, al ser desconocidos dichos pagos para el *dictador* este no puede saber si los pagos para el otro jugador son los mismos que en la línea base o si se encuentran invertidos
2. Cut-oﬀ en inglés

7

Economía experimental y del comportamiento - Referee Report Universidad Nacional de Colombia

informó que tendrían un intervalo de 10 segundos en dónde ingresarían sus elecciones, pero en el caso que no hubiera hecho una elección en un punto seleccionado aleatoriamente en el intervalo, el software eliminaría su posibilidad de decidir o tomar una decisión y sería el mismo software el que elegiría con igual probabilidad entre la opción A o la opción B. Los sujetos no conocían el punto preciso donde se encontraba el punto de corte, solo que dicho punto de corte podría ocurrir en cualquier punto del intervalo de 10 segundos. Solo el dictador sería notificado si había sido eliminado su posibilidad de elección y el software había elegido por él, de tal forma que destinatarios no podrían estar seguros si el pago había sido determinado por la elección del dictador o del programa. El intervalo de elección en dónde se encontraba el punto de corte tuvo lugar después de que los sujetos se les diera un minuto para considerar sus elecciones y no empezaría hasta que los sujetos le dieran la opción de iniciar juego. Por su parte, los destinatarios, tenían que hacer unas elecciones hipotéticas que también tenían la posibilidad de perder su elección luego de un punto de corte. Sujetos que hayan perdido esa posibilidad de elección hipotética, se les preguntaba que indicaran como habrían elegido si no hubiera sido eliminada su opción de elegir[20](#page10) (Dana et al., [2007)](#page14).

Se destaca que ningún dictador perdió la posibilidad de elegir antes de 4 segundos, es decir, no hubo ninguna línea corte antes de los 4 segundos, lo que significa que un dictador que realmente estuviera interesado en realizar alguna elección tendría suficiente tiempo par hacerlo[21](#page10). Es así, que agregar el punto de corte en este tratamiento debería ser irrelevante si el comportamiento del dictador estuviera realmente dirigido a preferencias sobre la distribución de pagos finales.

Este tratamiento tiene la ventaja de que permite ver si un dictador puede explotar la incertidumbre sobre que, precisamente, genera el resultado injusto. Es decir, el tratamiento permite ver si en caso de poder camuflar/justificar su ausencia de donación por elementos aleatorios o exógenos a su control el dictador/envidador decidiría utilizar dicha incertidumbre para no elegir y permitir que otro factor exógeno decidiera por él. Es decir, este último tratamiento funciona común experimento que busca testear el tercer ejemplo de los tres mencionados en el punto 3. Esto es evidente, si se tiene en cuenta que si los destinatarios nunca son capaces de saber si sus pagos se deben a elecciones del dictador o de un elemento aleatorio exógeno al dictador, el dictador podría escoger más fácimente la opción A de *self-interes* sin tener la preocupación de no tener que mostrarse egoista ante el otro jugador. Por otrolado, para dictadores que en la línea base deciden donar para evitar mostarse egoistas así mismos, entonces dichos dictadores podrían *vacilar* y permitir que se les quite la posibilidad de elección para que el computador eliga por ellos. En esta última situación, con la mitad de probabilidad el software escogerá el resultado justo/equitativo que se hubiera sentido forzados/obliados a escoger, pero con la misma probabilidad el resultado *selfish* habría sido obtenido y el dictador podría mantener la ilusión de no ser el responsable director de que ocurriera dicho outcome (Dana et al., [2007)](#page14).

1. **En los procedimientos para llevar a cabo el experimento, ¿cuáles elementos se podrían mejorar? ¿Por qué?**

Si bien considero que en su gran mayoría el diseño implementado por Dana et al. [(2007)](#page14) es brillante en el sentido que son capaces de: 1) representar adecuadamente en sus tratamientos los ejemplos que parecen contradecir las *teorías de preferencias sociales* y 2) cumplen con el objetivo de mostrar que las las *teorías de preferencias sociales* por sí solas son incompletas e incapaces de capturar totalmente el comportamiento de justicia/equidad observado en las personas, existen algunas modificaciones en la implementación de los tratamientos que podrían generar resultados más evidentes y que podrían reafirmar a un mejor el punto que los autores están intentando mostrar.

Una primera mejora en los procedimientos que podían conducir a mejores resultados está rela-cionado con la inclusión de más juegos en donde se homogenicen la suma de todos los pagos de los participantes a 10 en las dos opciones. Esto, dado que en la implementación original, la suma de los pagos en la opción A eran 7 mientras que en la opción 10 era B por lo que podría no ser claro si la elección de un dictador de escoger B se deba a que le da un pago mayor a su contraparte o porque en

1. Los puntos de corte, fueron sacados de una distribución normal discretizada, el objetivo era permitir a los dictadores tener una cantidad razonable de tiempo para elegir antes de perder esa opción y que el programa eligiera por ellos.
2. El tiempo de reacción promedio en los pre-test fue de alrededor de dos segundos.

8

Economía experimental y del comportamiento - Referee Report Universidad Nacional de Colombia

general dicha opción da pagos mayores a la opción A, alcanzando en la opción B un óptimo de Pareto. Es decir, un dictador podría escoger la opción B no por los pagos del otro jugador sino porque es el resultado que alanza el óptimo de Pareto. De esta forma, la adición de juegos donde se homogenizen la suma de pagos totales de los jugadores a 10 para ambas opciones podría mostrar si existe algún cambio en la elección de las opciones y así ver si la elección de B es por aumentar el bienestar del otro participante o si la idea es alcanzar un óptimo de Pareto para la sociedad en su conjunto.

Otro detalle que se podría explorar, es explorar un conjunto de pagos diferentes a los jugadores a los empleados en el experimento de tal forma que se pudiera observar si modificar los pagos afectan las decisiones entre A o B. Por ejemplo, para el caso base y los tratamientos darle un pago de 15 a la opción A y un pago de 11 a la opción B y cambiando los pagos del jugador destinatario de 1 a 11, puede que los resultados del experimento cambien y sea posible observar más elecciones de la opción A, inclusive en el caso base. Esto se debe a que la diferencia entre los pagos originales del dictador 6 a 5 puede ser muy pequeña para incentivar a comportarse de manera *self-interest* pero un cambio de 4 unidades en el caso que los pagos fueran de 11 a 15 podrían generar incentivos más fuertes a comportarse *self-interest*.

Otra modificación interesante es agregar la opción del *doble ciego* para el caso base y todos los tratamientos en lugar de usar computadores en una sola. Si bien los autores manifiestan que para el tratamiento de *Hidden information treatment* se ha implementado dicho tratamiento con doble ciego en otros estudios obteniendo el mismo resultado, sería interesante ver si lo mismo ocurre para la línea base y los otros tratamientos. Esto, dado que es posible que las personas se sientan presionados a actuar de cierta forma si no se encuentran totalmente anatomizados de toda observación, ya sea de experimentalistas u otras participantes, por lo que el doble ciego protegería de este efecto.

Otra modificación que se me ocurre valdría la pena explorar es el de colocar pagos de cero en el destinatario para la opción A. Un pago de 1, así sea muy pequeño, sigue siendo un pago positivo por lo que podría sesgar al dictador a que está *aportando* al jugador destinatario y por ende no se encuentra en un estado de *moral wiggle*. Ahora bien, si se colocara un pago de cero podría gener el efecto en el dictador de que no está aportando nada al otro jugador si escogiera la opción A por lo que el efecto del *moral wiggle* sería más pronunciado en ese caso, en el sentido de que puede obtener un pago mayor escogiendo A pero no le aportaría nada al otro jugador.

Finalmente, y la última modificación menos importante podría estar en la forma en la que se recluta a las personas. Dado que se coloca un aviso y se recluta a los estudiantes de pregrado por dicho aviso, existe la posibilidad de que un grupo grande de amigos o compañeros atiendan al experimento pero conociéndose, por lo que a pesar de que haya aleatorización puede que aún quede un sesgo a actuar den manera justa/equitativa porque puede que la persona con la que quede emparejada sea mi amigo o colega. Una mejor forma de reclutar estudiantes, sería de una lista totalmente aleatoria y diseñada por los investigadores se enviaran invitaciones personales, podría ser por correo, de tal forma que se minimice la posibilidad de que grupos de amigos grandes entre al experimento en el sentido de que las invitaciones ahora van a ser enviados por el investigador y de manera personal.

9

Economía experimental y del comportamiento - Referee Report Universidad Nacional de Colombia

1. **Piense en una manera para resumir los resultados, de manera gráfica. Puede ser una Tabla o una Gráfica. Tiene que ser auto-contenida, es decir, tener toda la información para que alguien que ha leído superficialmente el paper sea capaz de interpretar los resultados.**

Una forma de resumir los resultados principales del artículo de Dana et al. [(2007)](#page14) es mediante la gráfica [2](#page12)

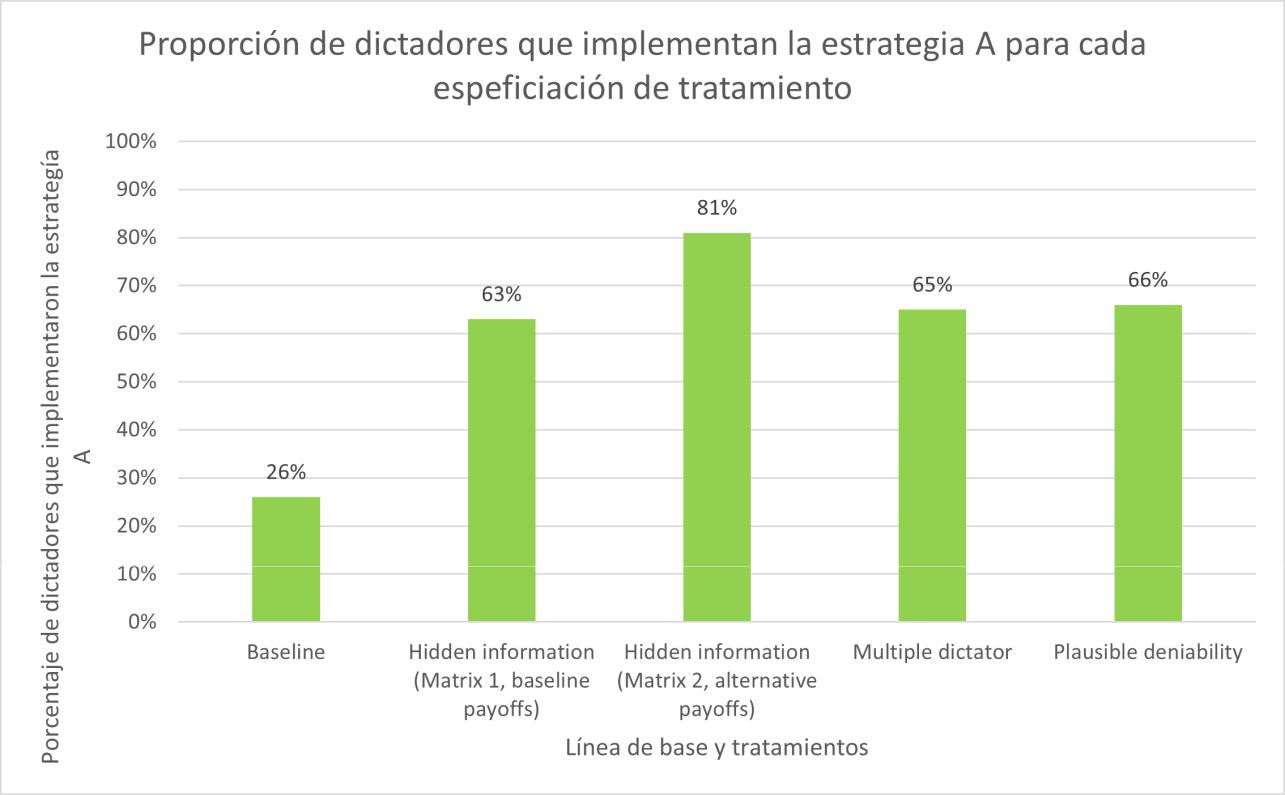


Figura 2: La gráfica muestra el porcentaje de dictadores que escogieron la *opción A*, que es la opción que se asocia al comportamiento de *self-interest*, tanto para el tratamiento de línea de base como para los otros tratamientos. Ahora bien, lo primero que se nota es que la línea de base es la que tiene el menor porcentaje de dictadores que eligen esa opción. Ellos se debe a que dada la transparencia, entendida está como la relación uno-a-uno entre decisiones del dictador y resultados, en la línea de base los dictadores no quieren parecer egoístas ante si mismos o ante los demás. Ahora bien, cuándo deja de ser evidente dicha transparencia como es en el caso de los demás tratamientos, se vuelve evidente que las personas empieza a dejar de tener un comportamiento de justicia/equidad y empiezan a tener un comportamiento auto-interesado. En cada tratamiento, se explora uno de los 3 ejemplos en dónde parecen no ser válidas las *teorías de preferencias sociales*. Para el tratamiento de *hidden information* existen dos opciones asociados cada una a una matriz de pagos distinta. En eltratamiento de *multiple dictator* lo que se hizo fue agregar un segundo dictador para que la eleccion de la opción A no implicara directamente el resultado inequitativo. Finalmente, se resalta que en el tratamiento de *plausible deniability* pueden haber dos justificaciones de porque las personas decidieron escoger la opción A. La primera es que el elemento de aleatorida de dicho tratamiento hiciera que las personas dejaran de tener la preocupación de no mostrarse egoístas ante los demás y escogieron la opción A. La otra alternativa, es que los dictadores vacilaran y dejaran que el software escogieran por ellos de tal forma que no parecieran egoístas para sí mismos.

10

Economía experimental y del comportamiento - Referee Report Universidad Nacional de Colombia

1. **Intente explicar en un párrafo cuál es la lección fundamental que se puede sacar de estos resultados.**

A mi juicio, la lección fundamental del artículo es que experimentalmente la eliminación de transpa-rencia en el juego del dictador, entendida esta como el conocimiento común de una relación uno-a-uno entre acciones y resultados, cambia sustancialmente el comportamiento de justicia/equidad que exhi-ben los agentes que enviar, es decir, los dictadores. A nivel teórico, lo anterior se entiende a que el comportamiento de generosidad o justicia que se observa frecuentemente en los experimentos no ne-cesariamente sea producto de unas preferencias de los agentes por resultados justos o eficientes, como es la interpretación usual que se da en disciplina como economía comportamental, sino que pueden deberse más bien a que las personas se sietan obligadas/forzadas a dar por las circunstancias de su entorno, pero que en realidad no están haciendo una valoración de los resultados globales y sobre todo del otro jugador (Dana et al., [2007)](#page14). Es decir, el artículo de Dana et al. [(2007)](#page14) intenta dejar como lección que la generosidad que se observa en los experimentos y en muchos estudios de campo no solo se debe a un deseo de resultados justos sino a miedo/repulsión a parecer egoístas ante los demás o ante si mismos, como lo muestran los tratamientos en donde se elimina la transparencia. Lo anterior es fundamental, dado que indica que es importante replantear los modelos y las teorías estándar que utiliza la economía para interiorizar y explicar el comportamiento de generosidad dado que gran parte de ese comportamiento pro-social o pro-bienestar no se debe a las razones usualmente explicados sino a explicaciones distintas como el miedo a parecer egoístas[22](#page13).

1. **¿Por qué los autores no han usado una tarea como el ultimátum o el juego de confianza?**

Una de las razones por las que se escogió un *juego del dictador* como tarea principal en la línea base y en los 3 tratamientos adicionales del experimento, fue que dicho *juego* tiene la característica de eliminar la interacción estratégica entre los jugadores dado que el segundo jugador, el que recibe, es pasivo en todo momento y solo acepta las decisiones del jugador que envía, lo que hace que todo elemento interactivo del juego desaparezca[23](#page13).

Ahora bien, a la hora de introducir *tareas/juegos* como el *juego del ultimatum* o como el *juego de* *confianza* se agrega un elemento de interacción que antes no estaba presente en el juego del dictadoren el sentido de que el otro jugador ya tiene poder de decisión y por ende a la hora de decidir en el envió el jugador que envía no solo tiene que pensar en su decisión sino también en la decisión del otro jugador, de ahí que surja una situación interactiva.

Una de las ventajas del juego del dictador es que al eliminar dicha situación interactva, elimina la posibilidad de que por ejemplo la *reciprocidad* o la *percepción de las intenciones de los demás* puedan afectar los resultados sociales del experimento. De esta forma, al hacer al destinatario pasivo, se logra que efectos de reciprocidad y las intenciones de los demás dejen de ser relevantes, es decir, se logra controlar por dichos posibles efectos.

Siendo más específicos, en el *juego del ultimatum* se observa que las personas deciden no enviar envíos tan pequeños como lo predeciria el equilibrio perfecto en subjuetos. Si bien esto se puede deber a comportamiento de equidad/justica, también se puede deber a que las personas, teniendo racionalidad

22Como una conclusión adicional y a parte a la que se acaba de dar pero que considero que también es importante mencionar se encuentra que otra lección, ya como experimentalistas, es que hay que tener mucho cuidado a la hora de interpretar los resultados de los experimentos y más aún a la hora de construir teorías basados en los resultados de los experimentos. Si bien, la mayoría de experimentos parecieran indicar que las personas tienen una fuerte preferencia por resultados pro-sociales lo que indicaría la existencia de preferencias pro-sociales en las personas, lo que justificaría le diseño de una teoría basada en preferencias altruistas, los experimentos realizados por Dana et al. [(2007)](#page14) muestran claramente que este no es el caso siempre y que hay ejemplos en dónde lo anterior no se cumple. Por tanto, una conclusión importante para un experimentalista es la importancia de realizar más experimentos de tal forma que se pueda obtener una compresión mejor y más profunda de un tema, en este caso entender el comportamiento de justicia/equidad que exhiben las personas, dado que es por medio de estos nuevos experimentos donde se pueden validar o no los resultados de experimentos anteriores y así construir teorías más completas y mejores

1. Esto, independiente si se utilizara una versión binaria, con varias opciones o un continuo de opciones en el juego del dictador.

11

Economía experimental y del comportamiento - Referee Report Universidad Nacional de Colombia

estratégica, sabe que si envian muy poquito puede que la contraparte pueda decidir rechazarle la oferta. Dado que, en el *juego del ultimatum*, es imposible separar estas dos explicaciones dado la situación de interacción presente en el juego, fue que se dio origen al uso del juego del dictador en economía experimental para poder aislar ambos efectos como lo muestran los estudios de Forsythe et al. [(1994)](#page14) y Hoﬀman et al. [(1996)](#page14).

Por otra parte, en el juego de confianza[24](#page14) ocurre lo mismo que en el *juego del ultimatum* en el sentido de que hay dos efectos que no son separables dado que envíos positivos pueden deberse ya sea a comportamiento de altruismo o de equidad o a comportamiento de racional estratégica. En el caso del *trust game* la motivación estratégica de un envió positivo podría alguna *creencia en la confiabilidad* del otro. Dicha racionalidad estratégica en el juego de confianza nuevamente no se puede aislar del comportamiento de generosidad/justicia que podría haber en un envío positivo por lo que la única manera de quitar ese componente de racionalidad estratégica es quitando la posibilidad de que la contraparte interactúe y eso se logra nuevamente mediante un *juego del dictador*, en donde el jugador destinatario no tiene capacidad de interactuar estratégicamente con su contraparte.

Como se puede ver, el *juego del dictador* surge como una solución a aislar el elemento de raciona-lidad estrátegica tanto en el *juego del ultimatum* como en el *juego de confianza* a la hora de observar envios positivos.

Por tanto, a mí juicio, escoger el juego del dictador es la elección adecuada porque lo que se busca es ver si los agentes tiene comportamiento pro-social o de bienestar social o si por el contrario tienen comportamiento más autointerezado y por ende se cumple el supuesto teórico de *self-interest*. Es decir, el juego del dictador, junto a sus diferentes variaciones como se mostró en el experimento, me permite testear directamente el supuesto teórico de *self-interest* y así ver si sí o si no hay comportamiento de justicia/equidad en los individuos. En caso de tener situaciones estratégicas, es más difícil testear este supuesto dado que el jugador que envía no solo tiene que pensar en sus propias acciones sino en las acciones de los demás por lo que sus decisiones dependerán de lo que otro decida. Por ejemplo, la persona que envía puede que decida no enviar muy poquito dado que si lo hace puede que el que recibe lo castigue o decida enviar más dado que espera que haya un efecto retributivo por parte de la contra parte en el juego. Todo eso, no ocurre en el jugo del dictador, por lo que es tan solo en este juego que en verdad se puede testear tanto el supuesto de *self-interest* como el comportamiento que aparentemente lo contradice de justicia/equidad aislando posibles factores de confusión o explicaciones alternativas plausibles.

**Referencias**

Dana, J., Weber, R. A., & Kuang, J. X. (2007). Exploiting moral wiggle room: Experiments demon-strating an illusory preference for fairness. *Economic Theory*, *33* (1), 67–80.

Forsythe, R., Horowitz, J. L., Savin, N. E., & Sefton, M. (1994). Fairness in simple bargaining experiments. *Games and Economic behavior*, *6* (3), 347–369.

Hoﬀman, E., McCabe, K., & Smith, V. L. (1996). Social distance and other-regarding behavior in dictator games. *The American economic review*, *86* (3), 653–660.

Medema, S. G. (2009). *The hesitant hand: Taming self-interest in the history of economic ideas*.

Princeton University Press.

Varian, H. R. (2014). *Intermediate microeconomics with calculus: A modern approach*. WW Norton & Company.

1. Trust Game en inglés

12