



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

PRINCIPIOS DE JUSTICIA DISTRIBUTIVA EN JUEGOS SECUENCIALES.

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN
ECONOMÍA APLICADA

PABLO FERREIRO GÓMEZ

**PROFESOR GUÍA
JUAN F. ESCOBAR CASTRO**

**MIEMBROS DE LA COMISIÓN
JAVIER BUSTOS SALVAGNO
JUAN ESCOBAR CASTRO
CARLOS NOTON NORAMBUENA**

SANTIAGO DE CHILE
2015

RESUMEN DE LA TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN ECONOMÍA APLICADA AUTOR: PABLO FERREIRO GÓMEZ FECHA: ENERO 2015 PROFESOR GUÍA: JUAN ESCOBAR CASTRO
--

PRINCIPLES OF DISTRIBUTIVE JUSTICE IN SEQUENTIAL GAMES.

Most economic models assume that all people are rational beings who pursue only their own material interests, without regard for other people or social welfare. However, there is experimental evidence that people not only maximize their individual utility, but also takes into consideration the welfare of those who belong to their reference group. In this context the theories of “Social Preferences” and in particular those of “Inequality Aversion” has relatively success at explaining the puzzling experimental evidence . These works are based on model that use certain values such as equality, envy, altruism, etc. And how these values determine the behavior of subjects in different situations. This paper seeks to incorporate notions of distributive justice through a simple model that has the flexibility for adapting to different principles of distributive justice that are studied here, namely equality, merit and need. The philosophical question behind is about the limitation of studying a single value versus the breadth and diversity of principles that constitute justice. The model is consistent in its internal logic and is able to explain the emergence of cooperation in games like prisoners dilemma or centipede. Empirical and experimental consistency is proposed for future research.

RESUMEN DE LA TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN ECONOMÍA APLICADA AUTOR: PABLO FERREIRO GÓMEZ FECHA: ENERO 2015 PROFESOR GUÍA: JUAN ESCOBAR CASTRO
--

PRINCIPIOS DE JUSTICIA DISTRIBUTIVA EN JUEGOS SECUENCIALES.

Usualmente los modelos económicos asumen que todas las personas son entes racionales que persiguen únicamente su interés material propio sin consideración a las otras personas o al bienestar social. Sin embargo, existe evidencia experimental que demuestra que las personas no solo maximizan su utilidad individual, sino que también tiene en consideración el bienestar de aquellos que conforman su grupo de referencia. En este marco las teorías de “Preferencias Sociales” y en particular las de “Aversión a la Desigualdad” tienen relativo éxito a la hora de explicar la evidencia experimental y el comportamiento de sujetos tanto egoístas como los que no se comportan de tal manera. Estos trabajos se basan en modelar ciertos valores como igualdad, envidia, altruismo, etc. Y cómo estos valores determinan el comportamiento de los sujetos en diversas situaciones. El presente trabajo busca incorporar las nociones de justicia distributiva a través de un modelo simple y a su vez flexible que permita adaptarse a los distintos principios de justicia distributiva que se estudian, a saber igualdad, mérito y necesidad. El cuestionamiento principal tiene que ver con la limitación de estudiar un único valor versus la amplitud y diversidad de principios que constituyen la justicia. El modelo es consistente en su lógica interna y es capaz de explicar el surgimiento de la cooperación en juegos como el dilema de los prisioneros o el ciempiés. La consistencia empírica y experimental queda propuesta para futuras investigaciones.

Agradecimientos

En primer lugar quisiera agradecer a mis padres, por entregarme una educación de excelencia y condiciones que me permitieran aprovecharla. En particular a mi madre quien me dio la posibilidad de hacer este magister con todas las facilidades posibles.

Me gustaría agradecer a todas las personas que estuvieron conmigo durante el magíster. A Felix, Javier, Diana, Andrea, Nicolás T., Nicolás I., Verónica, Felipe y Rodrigo, por ser grandes amigos, compañeros de estudio y ayudarme en mi aprendizaje. A Juan, por guiarme en el desarrollo de esta tesis. A Gonzalo por sus ayudas y comentarios. A Olguita quien siempre tuvo la mejor voluntad para ayudarme a resolver problemas.

Finalmente, agradecer a mi familia, compañeros y profesores que siempre tuvieron la mejor disposición y paciencia. A mis amigos, gracias por la magia.

Tabla de Contenido

1. Introducción	2
2. Literatura	6
2.1. La hipótesis del sujeto egoísta y las preferencias de consideración al otro . . .	6
2.2. Teorías de consideración al otro	9
2.3. Preferencias Sociales	10
2.3.1. Altruismo	10
2.3.2. Ingreso Relativo y Envidia	11
2.3.3. Aversión a la Desigualdad	11
2.4. Igualdad y Justicia	13
2.4.1. Principios de Justicia Distributiva	16
3. Modelo	18
3.1. Un modelo de Aversión a la Injusticia	18
3.2. El juego	19
4. Aplicación y Ejemplos	28
4.1. Un juego de merito e igualdad	28
4.2. Dilema de los prisioneros	30
4.3. Juego del ciempiés	33
5. Discusión	36
6. Conclusión	40
7. Bibliografía	42
A. Anexos	47

Capítulo 1

Introducción

La mayoría de los modelos económicos tienen como supuesto base que todas las personas son racionales y que dicha racionalidad consta de una persecución de intereses y fines materiales puramente egoístas. Todos los modelos usan diferentes supuestos que permiten estudiar una realidad en particular o situaciones que van de lo específico a lo general. No obstante, la hipótesis de los sujetos egoístas es un supuesto transversal a casi todos los modelos económicos. Este supuesto puede ser verdadero para (quizás) una gran cantidad de personas, pero ciertamente no lo es para todos.

Para el sociólogo alemán Max Weber este supuesto se trata de un tipo de *racionalidad con arreglo a fines*, donde el sujeto se comporta ajustando los medios de la manera más eficiente para obtener un determinado fin material. Sin embargo, dentro de las categorías de acción social que define Weber (1964), existe más de un tipo de racionalidad, además de acciones que no caen dentro de la categoría de racionales (como las acciones emotivas o tradicionales). La *racionalidad con arreglo a valores*, por ejemplo, define un tipo de comportamiento en la que el sujeto no tiene como objetivo un determinado fin material sino que la realización de un valor, ya sea religioso, ético, estético, etc. Así, una persona que tenga como valor cúspide la justicia buscara las acciones que permitan la realización de dicho valor, ya sea donando a beneficencia, cumpliendo la ley, ayudando o persiguiendo a víctimas o victimarios, etc. Este valor determinará sus acciones independientemente del campo social en el que se encuentre.

En esta línea existe una abundante evidencia experimental que muestra que una cantidad considerable de sujetos no se preocupan solamente por sus pagos materiales (Güth y Tietz, 1990; Camerer y Thaler, 1995). Es más, es posible sugerir que muchas veces las firmas tienen en consideración lo que sus trabajadores puedan percibir como justo a la hora de fijar sueldos (Blinder y Choi, 1990; Agell y Lundborg, 1995; Campbell y Kamlani, 1997). Esto explicaría por qué muchas veces, para no alterar la moral de trabajo, las empresas no recorten los sueldos

en una recesión por temor a que los trabajadores lo consideren injusto. Por otro lado, existe evidencia que muestra como casi todos los sujetos, bajo ciertas condiciones (experimentos de mercados y competencia), se comportan de manera egoísta (Smith y Williams, 1990; Güth, Marchard y Rulliere, 1997). Es decir, existe evidencia experimental contradictoria que tiende a refutar y otras veces a sustentar el supuesto del sujeto racional-egoísta. Un fenómeno similar ocurre cuando se estudia el problema de la cooperación en juegos con una estructura similar al dilema de los prisioneros, si bien el modelo económico estándar predice que no existe la cooperación, hay numerosos experimentos donde la cooperación se alcanza casi totalmente (Fehr y Gächter, 2000; Ostrom y Walker, 1991). Este tipo de evidencia también es confusa, puesto que alterando ciertas condiciones de los experimentos se pueden encontrar resultados que dicten no-cooperación.

Para dar cuenta de este problema, dentro de la economía del comportamiento, existe una rama de teorías conocidas como “preferencias de consideración al otro”. Este tipo de teorías incluyen, a través del marco de la *elección racional*, formas de comprender el actuar de sujetos que buscan algo más que solo su bienestar material individual. En este grupo de teorías se encuentran las teorías de “Preferencias Interdependientes”, de “Reciprocidad Basada en Intenciones” y las de “Preferencias Sociales”. Estas últimas asumen que la función de utilidad de un jugador no depende solo de su pago material, sino que también puede depender de como son distribuidos los recursos dentro de un grupo de referencia.

Las teorías de “Equidad, Reciprocidad y Competencia” de Bolton y Ockenfels (2000) y la “Teoría de Justicia, Competencia y Cooperación” más conocida como la de “Aversión a la Desigualdad” de Fehr y Schmidt (1999) son dos de los trabajos más importantes dentro del marco de las “Preferencias Sociales”. Ambas teorías buscan conciliar la evidencia contradictoria a través de un modelo único y simple. Estos modelos tienen como factor común que a los sujetos no les gusta la desigualdad. Por ende al sujeto no le interesa solamente que su ingreso pecuniario sea lo más alto posible, sino que además que este no sea muy distinto del de otros jugadores o su grupo de referencia. A través del uso de evidencia experimental Fehr y Schmidt calibran su modelo para luego probar como se ajusta en juegos típicos como el del Ultimátum, el Dictador, juegos de mercado, etc. Fijando ciertas restricciones sobre los parámetros de su modelo logran dar cuenta de algunos de los resultados contradictorios observados en los experimentos.

Sin embargo, cabe preguntarse si todas las personas tienen una preocupación por la desigualdad. Se podría pensar que si existe algún tipo de justificación con sentido para el sujeto, la desigualdad puede ser tolerable e incluso deseable. La pregunta que nos hacemos

entonces ¿Es la igualdad siempre justa? Es más ¿puede la justicia ser desigual¹? A juicio propio, la justicia es un concepto general que puede involucrar a la igualdad, pero ésta no es una condición *sine qua non* para la existencia de justicia. Este trabajo intentará entrar en la discusión sobre las “Preferencias Sociales” tratando de incorporar conceptos de justicia dentro de las dimensiones que conforman al sujeto y el proceso de toma de decisiones.

La ciencia política y la filosofía han intentado delimitar el concepto de justicia y sus dimensiones. Aristóteles señala que existen tres tipos de justicia: justicia procedimental, que dice de los procesos de administración de las normas y leyes; la justicia retributiva, que dicta la compensación por daños o beneficios, sanción de crímenes, etc.; y finalmente la justicia distributiva, que regula la asignación de beneficios y cargas en la actividad económica. Es en la justicia distributiva que la economía y la psicología social ponen su atención. La asignación de bienes escasos para necesidades múltiples y las labores productivas es el problema fundante de la ciencia económica. La construcción de modelos que apuntan a maximizar el bienestar social ha sido parte importante del estudio económico en el último siglo, pero originalmente no era la única. Tanto Adam Smith como gran parte de la ilustración escocesa y varios de sus intelectuales como Hume y Mandeville, tenían como una preocupación central el orden social, la moral y la conducta humana. Esto último tiene directa relación con los principios de la justicia distributiva, puesto que estos son principios normativos. La asignación de los bienes y deberes económicos pueden depender del conjunto de valores morales y la ética a la que adscribe un individuo y que constituyen a la sociedad. Es decir, la forma en que una asignación puede considerarse justa o correcta no es única.

De esta manera se pretende construir un modelo simple que permita dar cuenta de un sujeto cuyas preferencias y racionalidad van más allá del egoísmo individual, sino que también busque la justicia. El desafío principal está en que el modelo tenga la flexibilidad de poder modelar distintos tipos de justicia y a su vez explicar el comportamiento de los individuos en diversos tipos de juegos. Para esto se construye a un sujeto que se preocupa de su ingreso personal y a su vez “sufre” cuando los pagos que él considera justo (tanto para los demás como para sí mismo) difieren del pago real. Este trabajo busca hacer una contribución a la comprensión de un sujeto racional que no es exclusivamente egoísta desde una perspectiva multidisciplinaria, utilizando los aportes de la economía, la ciencia política, la psicología y la sociología. Es importante comprender que tanto el modelo que se propondrá como la literatura relacionada no constituyen una crítica a la racionalidad misma, sino que a un concepto más acotado, a saber, el egoísmo. Se asume que los sujetos siguen siendo racionales, solo que su preocupación abarca el bienestar propio y el de los demás sujetos de su grupo de

¹Justicia como la repartición de recursos o tareas y más que la aplicación de normas y leyes.

referencia. Por ahora solo se buscará la consistencia interna del modelo y su aplicabilidad a distintos juegos, el desarrollo y comparación con evidencia experimental quedará propuesto para futuras investigaciones.

Lo que resta de este trabajo se ordena de la siguiente manera. En el capítulo 2 se hará una revisión y discusión de la literatura relacionada a la economía del comportamiento en el marco de las teorías de “Preferencias Sociales” y de los principales conceptos de justicia distributiva. En el capítulo 3 se propondrá el modelo de justicia en el marco de la teoría de juegos, más específicamente, de juegos secuenciales. El capítulo 4 desarrollará, a través de ejemplos, como juega y se comporta un sujeto con preferencias por la justicia. El capítulo 5 se enfocará en la discusión de los distintos alcances de la teoría propuesta y las perspectivas de investigación futura. Finalmente el capítulo 6 presentará las conclusiones finales.

Capítulo 2

Literatura

2.1. La hipótesis del sujeto egoísta y las preferencias de consideración al otro

Desde los orígenes filosóficos de la economía en la antigua Grecia hasta su formalización moderna, muchos economistas e intelectuales influyentes como Gary Becker (1974), Paul Samuelson (1993) y Amartya Sen (1995) entre otros, han señalado que a menudo las personas tienen preocupaciones por el bienestar de otros. El padre de la economía moderna, Adam Smith (1759), en su obra “La teoría de los sentimientos morales” , exploraba los tipos de conductas que no tenían su origen en el egoísmo y que sin embargo tenían un impacto importante en la actividad económica. Sin embargo, la mayoría de los economistas todavía asumen que el interés material individual es la única motivación de todas las personas. Este supuesto contrasta fuertemente con la abundante evidencia presentada por economistas experimentales y psicólogos en las últimas dos décadas. Esta evidencia muestra que un porcentaje no despreciable de las personas están fuertemente motivadas por consideraciones hacia el bienestar de los demás. La reciprocidad, la justicia y la equidad también surgen como fuente de las interacciones sociales (Fehr y Schmidt, 2005).

Vernom Smith (1962) descubrió en los años 50 que mercados experimentales rápidamente convergen al equilibrio competitivo, si los sujetos intercambian bienes homogéneos y todos los aspectos de estos bienes son posibles de expresar en un contrato. Esto ha sido demostrado en múltiples otros experimentos de mercados experimentales (Davis y Holt, 1993) donde se parte de la premisa que todos los jugadores son exclusivamente egoístas, por lo que estos resultados se han interpretado como una confirmación de dicho supuesto.

Por otro lado, la validez de la hipótesis del sujeto exclusivamente egoísta se empieza a relativizar en los años 80, cuando los economistas experimentales empezaron a estudiar juegos

de negociación bilateral e interacciones en pequeños grupos en laboratorios con condiciones controladas (ver Roth, Malouf y Murningham, 1981; Güth, Schmittberger y Schwarze, 1982). Uno de los juegos experimentales que expresan más nítidamente las inconsistencias con la hipótesis del sujeto egoísta es el juego del Ultimátum (Güth, Schmittberger y Schwarze, 1982). Otros juegos típicos como el juego del Dictador, Intercambio de Dones, Juegos de Confianza, etc. también han contribuido con evidencia experimental que debilitan esta hipótesis. Todos estos juegos comparten la característica de ser simples, lo que permite a los jugadores entender el sentido, estrategias posibles y minimizar la posibilidad de que las conductas observadas sean producto de errores por no comprender el juego. Desde la óptica del sujeto egoísta todos estos juegos tienen una clara estrategia dominante fácil de descubrir y comprender, por lo que es válido interpretar que si no se juega dicha estrategia fue por una decisión consciente (Fehr y Schmidt, 2005).

En el juego del Ultimátum hay dos jugadores, el primero propone la división de una suma fija de dinero a repartir y el otro jugador debe decidir si acepta la división propuesta. En caso de ser rechazada la propuesta ambos jugadores quedan con cero, de ser aceptada, el dinero se reparte según la división propuesta por el primer jugador. Bajo los supuestos estándares de que los jugadores son seres racionales que piensan en su propio bienestar material y de que además ambos jugadores saben que el otro es racional y egoísta, el equilibrio perfecto en el subjuego es único: el jugador que debe decidir si acepta o rechaza la propuesta acepta cualquier suma positiva y el que propone ofrece la cantidad mínima posible y se queda con el resto.

Los resultados experimentales muestran con bastante consistencia que una gran mayoría de las propuestas de división van entre un 40 % y un 50 % del monto a repartir. Las propuestas que son bajo un 20 % son rechazadas con una probabilidad de entre 0.4 y 0.6. Además esta probabilidad es decreciente en el tamaño de la oferta (ver Güth, Schmittberger y Schwarze, 1982; Camerer y Thaler, 1995; Roth, 1995; Camerer 2003 o sus referencias). Se podría decir que el jugador que responde tiene un monto en su cabeza que considera *justo*, esta consideración pareciera imponerse a la hora de rechazar ofertas *injustas*, cosa que además es anticipada por el jugador que propone y que ofrece una cantidad mayor que la mínima.

Una explicación posible es que el jugador que propone tiene miedo a que le rechacen la oferta, por lo que ofrece una cantidad mayor de la que se esperaría según los supuestos estándares. Para esto Forsythe et al. (1994) compara los resultados del juego del Ultimátum con los del juego del Dictador. En este último juego la opción de aceptar o rechazar la oferta es removida y el segundo jugador acepta obligatoriamente cualquier oferta hecha por el primer jugador, por lo que el temor a ser rechazado desaparece. La predicción del modelo estándar es más radical aún que la del juego del Ultimátum, en el juego del Dictador el

primer jugador se queda con todo el dinero y no le da nada al otro jugador. Sin embargo, lo que encuentran los autores es que el jugador que reparte ofrece entre un 10 % y un 25 % del monto a repartir, con modas que se distribuyen entre 0 y el 50 %. Estos resultados muestran que si bien el temor al rechazo explica parte de las ofertas generosas del juego del Ultimátum, esta razón no es suficiente. Aún existe un elemento que lleva a los sujetos a ofrecer sumas positivas en el juego del Dictador.

El juego del Dictador es una forma de medir altruismo puro. El juego de Confianza tiene una estructura similar al juego del dictador pero que también intenta medir reciprocidad. Dos jugadores reciben una cantidad X de dinero, el jugador 1 o inversor entrega una porción $z \in [0, X]$ al jugador 2, el controlador del experimento triplica z , luego el jugador 2 debe decidir una cantidad entre cero y $X + 3z$ para devolverle al jugador 1. En teoría el jugador 2 no le devuelve nada al jugador 1, que anticipa esto y decide invertir 0. Lo que se puede observar en los experimentos (Berg, Dickhaut y McCabe, 1995) es que usualmente el jugador 1 invierte la mitad de su dotación inicial, la que suele ser devuelta (una vez más otras veces menos) por el jugador 2. Este juego muestra que los sujetos tienen un sentido de reciprocidad cuando el otro deposita una cierta confianza en ellos.

En el juego del Intercambio de Dones, un jugador ofrece un salario $w \in [\underline{w}, \bar{w}]$ a cambio de un esfuerzo $e \in [\underline{e}, \bar{e}]$, el esfuerzo es costoso para el segundo jugador, pero aumenta el pago de quién propone w . A diferencia del juego de Confianza, el jugador que decide el tamaño del excedente a repartir es el que responde. El equilibrio de dos jugadores egoístas implica que el segundo jugador siempre maximiza su utilidad dando el mínimo esfuerzo posible (independiente del w ofrecido), anticipando esto el jugador 1 siempre ofrece (y es aceptado) el mínimo w posible. Este juego replica de manera estilizada la relación agente-principal con contratos incompletos. Los experimentos han reportado que el esfuerzo promedio está positivamente correlacionado con el salario ofrecido (Fehr, Kirchsteiger y Riedl, 1998; Fehr y Falk, 1999; Gächter y Falk, 2001), constatando que ofertas generosas inducen esfuerzos generosos.

Como se señaló anteriormente, la simpleza de los juegos y la claridad de las posibles estrategias permiten suponer que el comportamiento de las personas en estos experimentos muestra un desvío consciente de la opción egoísta, develando la existencia de preferencias que van más allá del interés personal. Sin embargo, la evidencia de los experimentos de mercados competitivos se mantiene, provocando la existencia de resultados contradictorios que sustentan y rechazan las hipótesis del modelo egoísta. Como explicación a este problema surgen las llamadas *teorías de consideración al otro*, que hacen un esfuerzo por intentar conciliar la evidencia contradictoria ofreciendo otra estructura de incentivos y preocupaciones para los sujetos.

2.2. Teorías de consideración al otro

A continuación se hará una revisión de los tipos de modelos que, desde la perspectiva de las preferencias de consideración al otro, han intentado aproximaciones teóricas que buscan explicar el comportamiento observado en los distintos experimentos, siempre en el marco de la elección racional. Estos modelos se pueden clasificar en:

- i) “Preferencias Interdependientes”. Estos modelos parten de la base que las personas les preocupa el “tipo” que es el otro jugador, de manera que adecuan su comportamiento según como identifiquen al otro jugador. Así por ejemplo, si las personas pueden ser de tipo egoístas o altruistas, un jugador se comportará de forma altruista si descubre que el otro jugador también lo es. De lo contrario, si observa que su oponente es egoísta, sus preferencias también lo serán. En resumen las preferencias de un jugador dependen de las preferencias del oponente y viceversa. Estos modelos suelen ser relativamente complejos en relación a los de “Preferencias Sociales” (Fehr y Schmidt, 2005). Dentro de esta categoría se encuentra, por ejemplo, el modelo de “Altruismo y Rencor” de Levine (1998). Si bien la función de utilidad de este modelo puede explicar por qué algunas personas contribuyen en juegos de bienes públicos, tiene problemas a la hora de explicar por qué una persona podría ser altruista en una situación y rencorosa en otra.
- ii) “Reciprocidad Basada en Intenciones”. Estos tipos de modelo se centran en las intenciones (e interpretación de la intención en una acción determinada) de los sujetos, más que en el “tipo” de los jugadores o la asignación de los pagos. De esta manera los sujetos adecuan sus acciones de acuerdo a lo que “sienten” como trato de parte del otro. Así, si un jugador se siente tratado con amabilidad, responderá de manera amable. De la misma forma, si se siente herido, tratará de herir a su oponente. En estos modelos, la interpretación de las intenciones juega un rol central. Rabin (1993) presenta el primer modelo de “Reciprocidad Basada en Intenciones”, que luego es extendido por Dufwenberg y Kirchsteiger (2004) para juegos secuenciales. Sobre este último se profundizará en el capítulo 4 ya que funcionará como marco para el modelo de justicia que se propondrá. Para este tipo de modelos se requiere la incorporación de elementos de la *teoría de juegos psicológicos*.
- iii) “Preferencias Sociales”. Este tipo de modelos asume que la función de utilidad de un jugador depende del pago que reciban, pero además de la forma en que los recursos son asignados dentro de su grupo de referencia. Estas preferencias sociales se encuentran

dentro del marco de sujetos que se comportan racionalmente, por lo que se aplican los conceptos clásicos de teoría de juegos. Dentro de estos modelos destacan los de Bolton y Ockenfels (2000) y Fehr y Schmidt (1999) anteriormente señalados. Dado que el modelo de justicia que se propondrá se deriva de los modelos de “Preferencias Sociales”, en el siguiente apartado se hará una breve revisión de estos modelos.

2.3. Preferencias Sociales

Usualmente se asume que el tomador de decisiones tiene preferencias sobre la asignación de un bien material, que dichas preferencias son racionales y consistentes con ciertos requerimientos (como la transitividad y completitud). Sin embargo, la interpretación usual de este marco general asume que la preocupación por las asignaciones de un sujeto solo obedecen a lo que le corresponde a él. Los modelos de “Preferencias Sociales” expanden esta noción, por lo que el tomador de decisiones se preocupa de su asignación personal y además de lo que los otros reciben (Fehr y Schmidt, 2005).

Formalmente, X es el conjunto de asignaciones posibles de los $i \in \{1, \dots, n\}$ jugadores con $n \geq 2$. Por simpleza x_i es el pago monetario que le corresponde a i . Bajo los supuestos del modelo estándar la utilidad del jugador i solo depende de x_i , bajo las “Preferencias Sociales” la utilidad de i también depende de x_j , con $j \neq i$. La manera en que afecta x_j la utilidad del jugador i expresa algún tipo de preferencia social como el altruismo, la envidia, etc.

2.3.1. Altruismo

Según la descripción formal se entiende como un sujeto altruista aquel cuya primera derivada parcial de su utilidad $u(x_1, \dots, x_n)$ con respecto a x_1, \dots, x_n es positiva. Es decir, la utilidad del sujeto altruista es creciente en el pago de las demás personas. Como se señaló anteriormente el juego del Dictador es sumamente iluminador para medir esta noción de altruismo. Andreoni y Miller (2002) conducen una serie de experimentos del juego del Dictador, asignando distintos presupuestos y cantidades a repartir. En una primera instancia, los autores chequean violaciones al Axioma General de Preferencias Reveladas, constatando que los sujetos en su gran mayoría tienen preferencias que se pueden describir como funciones de utilidad cuasi-cóncavas y que se comportaban consistentemente de acuerdo a este chequeo básico de racionalidad.

Entre sus hallazgos Andreoni y Miller encuentran que cerca de un 30% de los sujetos repartían de manera que se igualaba el pago monetario entre ambos. Un 20% se comportaba como si dar un peso al otro o quedarse con él fuese lo mismo (sustitutos perfectos), maximi-

zando una suma ponderada de los pagos monetarios. Casi un 50 % se comportaba de manera egoísta, entregándole a los demás jugadores una cantidad despreciable del total a repartir. En otro experimento Andreoni y Miller encuentran que cerca de un 23 % de los sujetos se comportan de manera envidiosa, disminuyendo el pago del oponente cuando éste se encontraba en mejor posición. Finalmente concluyen que una mayoría de los sujetos parecieran tener preferencias de *consideración al otro* y que el comportamiento de estos sujetos, al menos en el juego del Dictador, es consistente con la racionalidad. Además destacan que existe heterogeneidad entre los individuos y solo una porción minoritaria (pero no despreciable) se comporta como altruistas incondicionales¹.

2.3.2. Ingreso Relativo y Envidia

Otra forma de incorporar “Preferencias Sociales” viene de la comparación relativa del sujeto con respecto a los demás. Es decir además del ingreso absoluto, importa en que posición se está en comparación a los otros jugadores. Bolton (1991) formaliza esta idea en un contexto experimental de negociación entre dos jugadores. Define una función de utilidad $U_i(x_i, x_j) = u_i(x_i, \frac{x_i}{x_j})$, que es estrictamente creciente en x_i y donde su derivada parcial respecto a $\frac{x_i}{x_j}$ es positiva cuando $x_i < x_j$ e igual a 0 cuando $x_i \geq x_j$. Es decir, el sujeto i sufre cuando el jugador j está mejor que él, pero no le importa cuando el jugador j está peor. Este modelo de envidia tiene relativo éxito en explicar los juegos de negociación que estudia Bolton, pero es incapaz de explicar la generosidad en el juego del Dictador, la reciprocidad en juegos de Confianza e Intercambio de Dones o las contribuciones voluntarias en los juegos de bienes públicos. Los mismos problemas aparecen en la aproximación de envidia de Kirchsteiger (1994). Pareciera ser que la envidia como preferencia social es muy específica y falla en explicar situaciones más generales (Fehr y Schmidt, 2005).

2.3.3. Aversión a la Desigualdad

Dentro de los modelos de “Preferencias Sociales”, las llamadas teorías de “Aversión a la Desigualdad” son los que han tenido un mayor éxito a la hora de explicar la evidencia experimental. Uno de estas teorías es la de “Equidad, Reciprocidad y Competencia” (ERC) de Bolton y Ockenfels (2000), su modelo básico se define como,

$$v_i = v_i(y_i, \sigma_i)$$

¹Para otros modelos de preferencias altruistas ver Cox, Sadiraj y Sadiraj (2001).

donde

$$\sigma_i = \sigma_i(c, y_i) = \begin{cases} \frac{y_i}{c} & \text{si } c > 0 \\ \frac{1}{n}, & \text{si } c = 0 \end{cases}$$

es la parte del pago relativo de i , y

$$c = \sum_{j=1}^n y_j$$

es la suma del pago pecuniario de todos los jugadores.

Esta función de motivación puede ser considerada como un tipo especial de función de utilidad esperada. Los siguientes supuestos caracterizan a esta función: I) v_i es continua y doblemente diferenciable² en (y_i, σ_i) . A1.1. Interés propio reducido: $v_{i1} \geq 0$, $v_{i11} \leq 0$, no se asume que es estrictamente creciente en el componente pecuniario para no descartar sujetos que les importe más su pago relativo; II) Dada dos opciones, donde $v_i(y_1, \sigma) = v_i(y_2, \sigma)$ e $y_1 > y_2$, el jugador i elige (y_1, σ) (quiebra empates); III) Efecto comparativo: $v_{i2} = 0$ para $\sigma_i = \frac{1}{n}$, y $v_{i22} < 0$. Esto es, manteniendo el argumento pecuniario fijo, la función es cóncava hacia abajo en el argumento relativo, con un máximo cerca de donde la asignación personal es similar al promedio. Aquí importa la significancia colectiva, por lo tanto se refiere a la asignación igualitaria como el punto de referencia social.

La heterogeneidad existente en la data es representada por la teoría postulando la tensión, o *trade-off*, entre acercarse al punto de referencia y la ganancia individual. Esta tensión se captura en como los individuos resuelven el *trade-off* mediante los umbrales en los cuales se difiere del comportamiento puramente egoísta. Cada jugador tiene dos umbrales, $r_i(c) = \arg \max_{\sigma_i} v_i(c\sigma_i, \sigma_i)$, $c > 0$ y $s_i(c) : v_i(cs_i, s_i) = v_i(0, \frac{1}{n})$, $c > 0$, $s_i \leq \frac{1}{n}$. En el juego del dictador r_i corresponde a la división que realiza i y en el juego del ultimátum s_i sería el umbral de rechazo de la oferta.

Este modelo explica bastante bien la amplia evidencia conflictiva como la generosidad en el juego del Dictador, Intercambio de Dones, juegos de Confianza y los rechazos en el juego del Ultimátum.

Otro de los trabajos más importantes en este ámbito es la “Teoría de Equidad, Competencia y Cooperación” más conocida como “Aversión a la Desigualdad” de Fehr y Schmidt (1999). Este modelo tiene la ventaja de ser simple, incluir consideraciones altruistas y de envidia hacia los otros jugadores desde un punto de referencia igualitario. Además el modelo trata de rescatar la heterogeneidad de los agentes. Así, la función de utilidad de un sujeto que es averso a la desigualdad se define como,

²Para este caso el subíndice numérico de v_i refleja las derivadas parciales del primer o segundo argumento de la función.

$$U_i(x) = x_i - \alpha_i \frac{1}{n-1} \sum_{j \neq i} \max[x_j - x_i, 0] - \beta_i \frac{1}{n-1} \sum_{j \neq i} \max[x_i - x_j, 0]$$

donde se asume que $\beta_i \leq \alpha_i$ y $0 \leq \beta_i < 1$.

Esta función de utilidad racionaliza acciones positivas y negativas entre jugadores que buscan la igualdad. El primer término de la ecuación señala el pago pecuniario del jugador i , el segundo término representa la pérdida de utilidad por una desigualdad desventajosa, mientras que el último término representa la pérdida de utilidad por una desigualdad que favorece al jugador i . Fehr y Schmidt muestran que la interacción de diferentes tipos (expresados en diversos valores de los parámetros α 's y β 's en diversos ambientes estratégicos pueden dar a lugar resultados muy equitativos o muy desiguales. Los autores eligen una distribución de sus parámetros que es capaz de explicar muy bien los resultados experimentales de juegos de Ultimátum, Dictador, etc. Y que además entrega excelentes predicciones para diversos juegos de negociación, mercados y de dilemas sociales. Sin embargo, los modelos de "Aversión a la Desigualdad" pueden derivar en soluciones sub-óptimas en problemas de agencia (Fehr, Klein y Schmidt, 2007). Aún así, la crítica que se realiza en este trabajo se refiere más al valor que da origen al modelo más que a la formulación misma de éste. En el presente trabajo se expande la noción de un sujeto que prefiere o no la igualdad, hacia un sujeto que busca la justicia.

Los modelos de Fehr y Schmidt y de Bolton y Ockenfels entregan resultados muy similares para situaciones de dos jugadores³, pero se diferencian para $n > 2$. Esto ocurre básicamente porque mientras que para Fehr y Schmidt el pago de referencia es siempre igual para todos los jugadores, para Bolton y Ockenfels el pago de referencia es el promedio de los jugadores. Es decir, el jugador i en vez de preocuparse de su posición relativa al jugador j se preocupa de su posición relativa al pago promedio.

2.4. Igualdad y Justicia

Dentro de los modelos de "Preferencias Sociales" son valores como el altruismo, envidia, igualdad y equidad los que más se repiten. Se trata de construir teorías que intentan replicar valores que a su vez dictan el comportamiento de las personas. Cabe preguntarse entonces sobre el rol que tiene la justicia como configuración de valores socialmente aceptados y que justifica la existencia de ciertas instituciones y estructuras sociales. Los modelos como el de "Aversión a la Desigualdad" traen implícitamente una restricción en el valor motriz

³Notar que para el caso de dos jugadores la definición funcional del modelo de Bolton y Ockenfels toma una forma muy similar a la de Fehr y Schmidt: $v_i(c\sigma_i, \sigma_i) = a_i c\sigma_i - \frac{b_i}{2}(\sigma_i - \frac{1}{2})^2$; $a_i \geq 0, b_i > 0$

que los origina. Al margen de la capacidad que tenga o no el modelo de explicar los diversos problemas, las personas y comunidades no necesariamente se guían por ese valor en específico. La igualdad no necesariamente es el valor que determina qué es justo y qué no lo es. El contexto, la cultura, la historia, etc. Puede dar origen a múltiples formas de justicia que a su vez determinan las estructuras sociales y las preferencias individuales. Este cuestionamiento filosófico es reconocido (implícitamente) por Fehr y Schmidt (1999) cuando señalan que la estructura del juego es determinante a la hora de permitir que distintos tipos de preferencias se expresen. Hablar de justicia, permite una flexibilidad mayor a la hora de interpretar las preferencias de los sujetos más allá de su egoísmo.

En una definición amplia la justicia se entiende como un valor determinado que es considerado un bien común para la sociedad y que da sentido y forma a sus normas jurídicas. Nace de la necesidad de mantener la convivencia armónica y orden entre los integrantes de una comunidad, a través de pautas y criterios que moderan las relaciones entre individuos, instituciones, etc. Esto generalmente se expresa a través del *permiso* y/o la *prohibición*.

La justicia posee un fundamento formal, expresado a través de un código escrito, imparcial, impersonal que dicta disposiciones para la interacción de sujetos e instituciones y que modera y resuelve los conflictos. Por otro lado existe el fundamento cultural, que se refiere a lo que una sociedad considera como bueno o malo, y de los lineamientos generales con los que se organiza dicho grupo. En este sentido, la justicia tiene un fundamento contingente a una época y lugar determinado. En los orígenes occidentales de la teoría del derecho y la justicia el aporte romano (y su origen griego) se puede considerar la piedra angular sobre la cual se construye la mayoría de los sistemas jurídicos del mundo occidental. El jurista romano Ulpiano elaboró una definición de justicia que hasta el día de hoy es citada: “La justicia es la constante y perpetua voluntad de dar (conceder) a cada uno su derecho(...) vivir honestamente, no hacer daño y *dar a cada uno lo que corresponde*.” El énfasis en la cita es propio y tiene como sentido no solo destacar una frase comúnmente conocida cuando se habla de justicia, sino que también expresar la contingencia de lo que es justo. Lo que “a cada uno corresponde” se constituye en torno a algún valor determinante en una sociedad. Por ejemplo, en la distribución de un bien, a cada uno puede corresponderle lo mismo, por lo tanto la igualdad es el valor central en torno a lo que algo se puede considerar justo. Pero la igualdad no es necesariamente el único valor que origina la justicia. Dar a cada uno según lo que aporte a la sociedad es otra forma de repartir el mismo bien, y que perfectamente puede ser considerado justo para alguna comunidad.

Filosófica e históricamente la *Ley del Talión*, comúnmente denominada como la ley del “ojo por ojo”, es de las formas de justicia más antigua de que se tenga registro. Basada en un principio de fuerza, poder y retribución, la Ley del Talión es cuestionada por los filósofos de

la Escuela de Atenas. En “La República” Platón define que una república es justa y que por ende sus ciudadanos deben ser justos. La justicia es entendida por Platón como que cada cual haga lo suyo (lo que le corresponde), que cada uno haga su trabajo, lo que implica la posesión de lo que cada quien necesita para hacer su parte. En Aristóteles el concepto se modifica en base a un criterio proporcional, dar a cada cual lo que corresponde según las necesidades del ciudadano, su aporte a la sociedad y sus méritos. Con Santo Tomás de Aquino, la justicia existe en base a derechos naturales que son otorgados por Dios, integrando a los criterios y valores de la religión en la concepción de justicia de una determinada comunidad. Desde el punto de vista del utilitarismo, la justicia es lo que beneficie a la mayoría. Bajo esta óptica el deber de las instituciones públicas es buscar la felicidad o maximizar la utilidad agregada. Con John Rawls la pregunta por lo que es justo va un paso más allá. Para Rawls la elección de principios por los cuales las personas rigen su acción es determinante, y por ende no pueden ser adoptados criterios de justicia desde una determinada posición, objetivo, ambiciones, etc. Que pueda tener una persona en particular o por la confrontación de ideas y principios entre los ciudadanos. Para esto “el velo de la ignorancia” priva a los ciudadanos de conocer su posición originaria, de esta manera se libera a los ciudadanos de los sesgos que su posición en la sociedad les pueda dar, dejando a las partes para que libres, iguales y racionalmente elijan sus principios de justicia.

Dentro de estas corrientes filosóficas es Aristóteles quien define tres tipos de justicia que han predominado hasta el día de hoy en el estudio del derecho. i) La justicia procedimental, que se refiere a las formas de creación y aplicación de las normas, leyes y sus procesos para resolver disputas. ii) La justicia retributiva, que define la respuesta o retribución proporcional y moralmente aceptable para un crimen o falta. iii) La justicia distributiva, que tiene que ver con la forma y principios normativos mediante los cuales se asignan los bienes y deberes de la actividad económica.

Es en el ámbito de la justicia distributiva que se realiza el presente trabajo. Como se señaló anteriormente, las teorías de las “Preferencias Sociales” han concentrado su atención en la existencia de valores o criterios como la igualdad o el altruismo, para explicar el comportamiento más allá del interés propio de los sujetos. El estudio de las formas mediante las cuales una sociedad distribuye sus bienes, recursos y tareas es también el problema fundamental de la economía. Por lo que aquí se pretende extender la idea de las “Preferencias Sociales” hacia los principios de la justicia distributiva.

2.4.1. Principios de Justicia Distributiva

Los principios de la justicia distributiva son las bases normativas que determinan la distribución de los beneficios y cargas de la actividad económica. En su teoría de la justicia Rawls (1971) propone un método mediante el cual sujetos racionales e imparciales seleccionan sus principios de justicia. Rawls señala que a través del “velo de la ignorancia”, los sujetos sin saber su posición en la estructura social y sin conocimiento de sus propias habilidades eligen la *igualdad* absoluta como principio de justicia y como norma de medida para otros principios distributivos (Michelbach, Scott, Matland y Bornstein, 2003). Si bien existe una primacía de la igualdad como principio, Rawls admite que se puede aceptar desigualdad cuando existe un incremento en la eficiencia de la producción de bienes y si esta beneficia a todos, en especial, a los que están peor. Este “principio de diferencia” (Rawls, 1971) como justificación para la desigualdad, esconde un principio de *necesidad* (Michelbach, Scott, Matland y Bornstein, 2003).

Debido a la estrecha relación conceptual entre igualdad y necesidad⁴ (Michelbach, Scott, Matland y Bornstein, 2003) se adopta la idea de Rawls en torno la necesidad desde un punto de vista absoluto, donde se considera un nivel mínimo necesario de bienes sociales. De esta manera se propone un principio re-distributivo basado en esta concepción de necesidad. Sin embargo se debe notar que este “nivel mínimo de bienes sociales” que necesita un sujeto puede variar de acuerdo a la riqueza y los estándares de calidad de vida del grupo social al que pertenece. Entonces este mínimo social definido de manera abstracta puede traducirse en un umbral máximo entre los bienes que recibe el sujeto más beneficiado y el menos beneficiado. El principio re-distributivo tiene como objeto corregir parte de esas diferencias que aparecen en el origen.

Finalmente el principio del *mérito* ha tenido un rol central en las discusiones sobre justicia distributiva. Dar a cada cual según su aporte a la sociedad, sus habilidades y capacidades (ejercidas y no inactivas), implica un reconocimiento de que lo que se aporta es valorado por la sociedad. Para Rawls (1971), Nozick (1974), Okun (1975) y Hayek (1976) el mérito es un principio menor puesto que las diferencias que puedan haber en torno a lo que dos sujetos entiendan como mérito y premio de sus acciones hasta las ventajas por la “lotería natural” de las habilidades y motivaciones, constituyen una fuente de inequidad que no se justifica. Por otro lado existe también un apoyo al principio Aristotélico del mérito (e.g., Sandel, 1982; Sher 1989; Miller, 1999) como principio legítimo que estimula la eficiencia, alinea los incentivos, motiva a los sujetos a alcanzar el potencial de sus habilidades y esfuerzos, etc. En sus experimentos Michelbach, Scott, Matland y Bornstein (2003) encuentran que para los

⁴Especialmente cuando se entiende la necesidad como relativa al grupo.

sujetos, las distribuciones basadas en el mérito, tienen un sustento de validez desde el punto de vista moral. En lo que concierne a este trabajo, se entenderá como mérito al aportar al grupo social algo que sea valorado y en el contexto económico esto se entiende como maximizar el pago social antes que el individual. Esta opción se fundamenta en tratar de diferenciar un comportamiento de un sujeto que se esfuerza solo para maximizar su pago, de otro sujeto que se esfuerza para aportar al grupo de referencia. Luego siguiendo la forma Aristotélica, el mérito es valorado y retribuido por la sociedad, por lo tanto esta forma de definir mérito solo tiene sentido en un contexto social que lo valore y premie.

Si bien hay muchos otros principios que se pueden definir, en base a la literatura estudiada se pueden reducir a estos tres principios básicos -igualdad, redistribución o necesidad y mérito- como los centrales en torno a los cuales se pueden componer formas más complejas de justicia distributiva (Garildo y Benitez, 2011). En el presente trabajo se estudiará un modelo de “Preferencias Sociales”, donde las personas se preocupan por ser justas de acuerdo a los tres principios de justicia distributiva recién definidos.

Capítulo 3

Modelo

3.1. Un modelo de Aversión a la Injusticia

El modelo que se propone se basa en la teoría de “Aversión a la Desigualdad” de Fehr y Schmidt (1999), pero el concepto que se intenta explorar es el de justicia. En esencia este modelo puede considerarse una versión más general del diseñado por Fehr y Schmidt, puesto que los autores modelan desigualdad como diferencias absolutas respecto al propio pago. El concepto con que aquí se trabaja es el de *justicia*, en nuestro caso, justicia distributiva. Esto implica que el sujeto no solo se compara a si mismo en relación al resto, sino que compara a todos los jugadores en relación a una norma definida por el principio de justicia distributiva que se utilice. De esta manera, la utilidad de un sujeto que es averso a la injusticia está definida como:

$$U_i(X, \bar{X}_i) = x_i - \alpha_i \left(\frac{1}{n} \sum_{j=1}^n (x_j - \bar{x}_{ij})^2 \right) \quad (3.1)$$

Donde $x_i \in X$ es el pago pecuniario de los $i = \{1, 2, \dots, n\}$ jugadores. Con $\bar{x}_{ij} \in \bar{X}_i$ es el valor del pago que el jugador i considera justo para j . Estos pagos se definen según el principio de justicia al que el sujeto adscribe. El primer término de la ecuación representa la utilidad directa recibida por el pago correspondiente al jugador i . El segundo término de la ecuación representa la pérdida de utilidad por resultados injustos. Este factor depende del valor de \bar{x}_{ij} , el “Concepto Personal de Justicia”, que representa lo que el jugador i cree justo para los demás. Es decir, el sujeto compara los pagos de cada jugador (incluyéndose) con los pagos que considera justo para cada uno. La idea de que las diferencias tengan un peso mayor de acuerdo éstas se vayan ampliando es sugerida por los mismos Fehr y Schmidt (1999) dentro de las potenciales mejoras que sugieren para su modelo. Por último, el parámetro α_i

es el grado de “Aversión a la Injusticia” es personal y pondera cuanto afecta a la utilidad del jugador la existencia de injusticia.

Este modelo, al igual que el de Fehr y Schmidt o de Bolton y Ockenfels, cae dentro de la categoría de los modelos basados en el resultado. Es decir modelos que perturban las funciones de utilidad de los individuos, por lo que los pagos son observados de manera distinta y por ende cambian las estrategias. Por otro lado están los modelos como los de Rabin (1993) o Dufwenberg y Kirchsteiger (2004) donde los sujetos alteran su comportamiento durante el juego de acuerdo a las intenciones percibidas sobre los otros. Al trabajar con conceptos como mérito, necesidad e igualdad, se hace necesaria una mirada general sobre la estructura del juego. Esto quiere decir, que para poder valorar un tipo de comportamiento es necesario poder observar como se comportan los jugadores, las acciones y estrategias que siguen. Así, por ejemplo, si un jugador que valora la meritocracia desea comportarse así y además ser recíproco con aquellos que también lo son, debe ser capaz de poder analizar y responder a las acciones observadas de los otros. Es por esto que este modelo debe ser entendido en un contexto de juegos secuenciales que entrega las herramientas para poder desarrollar los tres principios de justicia distributiva que se han discutido.

3.2. El juego

A continuación se presenta un esquema de juego secuencial dentro del cual se desenvuelve un sujeto con preferencias por la justicia. Lo que motiva la acción de un jugador depende de la intención que tenga el jugador, pero también depende de la creencia sobre las repercusiones que tendrá su acción y como será interpretada por la contraparte. A su vez las motivaciones recíprocas dependen de creencias sobre creencias. Para incluir estas consideraciones se utiliza el marco de la *teoría de juegos psicológicos* (Geanakoplos et al., 1989), enfoque en el cual los pagos que reciben los jugadores dependen no solo de los perfiles de estrategias que se juegan, sino que además, de las creencias que tiene el jugador sobre las elecciones estratégicas y creencias de los otros jugadores.

Se parte de un juego estándar, visto como una descripción de situaciones estratégicas que involucran pagos. Luego, para incluir las repercusiones de las creencias se utiliza como referencia el enfoque de Rabin (1993), derivando a un *juego psicológico* donde las funciones son redefinidas para reflejar las consideraciones recíprocas, creencias y los valores personales. Sin embargo el modelo de Rabin se centra en juegos de forma normal, mientras que aquí se trabaja con juegos en forma extensiva¹ y se imponen requerimientos de racionalidad secuen-

¹Como se discutió anteriormente, las consideraciones de justicia que se pretenden estudiar se involucran en una dinámica que considera la estructura del juego y la observación del comportamiento de otros jugadores,

cial. En este sentido este trabajo sigue la línea de la *Teoría de Reciprocidad Secuencial* de Dufwenberg y Kirchsteiger (2004). Los autores construyen un modelo que intenta explicar el comportamiento recíproco entre agentes en un contexto de juegos secuenciales. Para esto utilizan funciones de utilidad que se evalúan acorde se van jugando nuevas etapas del juego. El presente trabajo utiliza el esquema de juegos secuenciales que desarrollan Dufwenberg y Kirchsteiger, sin embargo, el modelo que aquí se propone intenta replicar sujetos que quieren ser justos, en este sentido la función de utilidad es distinta a la que presentan Dufwenberg y Kirchsteiger².

Tal y como definen Fudenberg y Tirole (1991), se desarrolla un juego finito de múltiples etapas y sin naturaleza. El juego avanza en etapas en las cuales cada jugador, a lo largo del camino que llevó a ese nodo,

- i) conoce todas las elecciones previas
- ii) se mueve solo una vez
- iii) no obtiene información sobre las elecciones de los jugadores en esa etapa.

Excluyendo la naturaleza, la única fuente de información incompleta proviene de acciones simultáneas, la restricción hacia juegos secuenciales de múltiples etapas con acciones observadas no altera el foco del modelo y permite una aproximación teórica tanto para la descripción de las acciones y creencias que se definirán, como de los conceptos a estudiar. Formalmente, sea $N = \{1, \dots, n\}$ con $n \geq 2$, el conjunto de jugadores. H es el conjunto de los perfiles de elecciones o historias que llevan a los subjuegos (\emptyset pertenece a H y lleva al nodo inicial o raíz). Sea A_i el conjunto de estrategias de comportamiento de $i \in N$; cada estrategia asigna para cada historia $h \in H$ una distribución de probabilidad al set de posibles elecciones del jugador i en h . Si bien se puede notar que A_i permite la aleatorización, dada la interpretación de los conceptos a desarrollar se sigue la formulación de Dufwenberg y Kirchsteiger (2004), lo que implica que los jugadores solo pueden hacer elecciones puras. Por lo tanto, la aleatorización se interpreta como las frecuencias en que elecciones puras son hechas en una población.

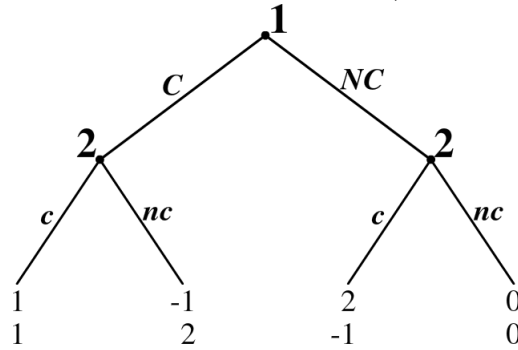
Se define $A = \prod_{i \in N} A_i$. Usando la asignación de pagos de los nodos finales y según el perfil de A que se jugó, se define la función de pagos materiales $\pi_i : A \rightarrow \mathbb{R}$. Sin embargo el sujeto no solo le interesa maximizar su utilidad material, sino que además quiere ser “consecuente”

por lo que juegos dinámicos y secuenciales se adaptan de mejor manera a las situaciones estratégicas que se pretende modelar.

²En su función de utilidad Dufwenberg y Kirchsteiger definen funciones que interpretan si una determinada acción de otro jugador se entiende como amable o agresiva, de manera que el jugador pierde utilidad si no se comporta de la misma manera que se comportan con él

con sus valores sobre la justicia y comportarse recíprocamente con aquellos que él considera que han sido justos o injustos con él. Al pago material se le suma un pago “valórico” que se construye en base a las inclinaciones personales por la justicia y las creencias del jugador i sobre las estrategias y creencias de los otros jugadores. Las creencias se representan como estrategias de comportamiento, de esta manera sea $B_{ij} = A_j$ el conjunto de posibles creencias de i sobre la estrategia de j . Análogamente $C_{ijk} = B_{jk} = A_k$ es el set de posibles creencias de i sobre las creencias de j sobre la estrategia de k .

Figura 3.1: Γ_1 . Dilema de los prisioneros secuencial (Dufwenberg y Kirchsteiger, 2004)



En esta formulación se debe tener en cuenta que la percepción y creencias sobre las intenciones de los otros jugadores pueden ir cambiando cuando se alcanza un nodo después de distintas historias. Por lo que se hace necesario mantener registro del cambio en el comportamiento, intenciones, creencias y la percepción sobre las intenciones de otros, para cada jugador en las distintas historias. Para esto se toma el juego Γ , como se definió previamente, para $a_i \in A_i$ y $h \in H$, sea $a_i(h)$ la estrategia actualizada que excluye las mismas elecciones que a_i , a excepción de las elecciones que definen h que se hacen con probabilidad 1. Notar que $a_i(h)$ esta definida únicamente para cada historia. De la misma forma, para las creencias $b_{ij} \in B_i$ o $c_{ijk} \in C_{ijk}$, se definen las creencias actualizadas como $b_{ij}(h)$ y $c_{ijk}(h)$.

Supongamos que el jugador i juega la estrategia $a_i \in A_i$, inicialmente cree que los otros jugarán $(b_{ij})_{j \neq i}$ y creerán $(c_{ijk})_{k \neq j}$. Después de la historia h el jugador i está jugando su estrategia $a_i(h) \in A_i$ creyendo que los otros jugarán $(b_{ij}(h))_{j \neq i}$ y creerán $(c_{ijk}(h))_{k \neq j}$. Esto implica que los jugadores renuncian a las creencias que incluyen aleatorización por creencias con elecciones puras a medida se van realizando las elecciones. Tomando el ejemplo del dilema de los prisioneros de la Figura 1, sea h el nodo donde le toca actuar al jugador 2 después de que el jugador 1 jugó NC, $b_{21} = (1 - \mu)C + \mu NC$ (la notación indica la aleatorización donde $\mu \in [0, 1]$ que es la probabilidad de que 1 juegue NC). Para cualquier valor de μ , se tiene que $b_{21}(h) = NC$, es decir que independiente de cuales sean las creencias del jugador

2 en la raíz del juego, cuando se le pide al jugador 2 que juegue en h, éste observa que se jugó NC con probabilidad 1. Si $\mu > 0$, esta forma de actualización es “Bayesiana”, dada la interpretación de la aleatorización de estrategias como frecuencias de estrategias puras más que una aleatorización consciente. Si $\mu = 0$ esta actualización aún es “Bayesiana”, pero de la forma particular en que los jugadores asumen intencionalidad de dicha acción. Es decir, el jugador 2 en h cuando observa NC, tratará al jugador 1 como que eligió NC y no como que eligió C y cometió un error. Es importante este supuesto ya que la teoría induce a comportamientos recíprocos o correctivos.

La actualización de las creencias permite lidiar con el problema de cómo cambian las preferencias cuando movidas inesperadas ocurren. Esta solución propuesta por Dufwenberg y Kirchsteiger (2004) no permite lidiar con el problema general de qué concluir cuando se juegan acciones que están fuera del equilibrio. Este es un problema que surge en la teoría de juegos estándar para cualquier concepto de racionalidad secuencial, y esta solución no pretende dar respuesta a este problema general.

Hasta ahora este juego se ha desarrollado de la misma forma que proponen Dufwenberg y Kirchsteiger. Sin embargo para incluir las preferencias por justicia se desarrollan las siguientes definiciones, según lo discutido en los capítulos previos.

El principio de igualdad absoluta de Rawls (1971) se construye en base a la ignorancia que cada sujeto tiene sobre sus capacidades y sobre su futuro. Dado que cada individuo racional en este estado de ignorancia optaría por una distribución igualitaria, es este principio de igualdad absoluta el que determina, para Rawls, una distribución justa y que sirve como punto de referencia comparativo para otros principios. Un jugador que tiene preferencias por la igualdad intentará minimizar las diferencias de pagos entre todos los jugadores.

Definición 1. Sea $a_i^I \in A_i$ el subset de acciones del jugador i tal que

$$a_i^I \in \arg \min_{a_i \in A_i} \sum_{j \in N} |\pi_i(a_i, (b_{ij})_{j \neq i}) - \pi_j(a_i, (b_{ij})_{j \neq i})|$$

Análogamente, b_{ij}^I son las creencias de lo que el jugador i considera como un comportamiento igualitario por parte de j.

El principio de necesidad en Rawls justifica una distribución desigual ante un incremento de la eficiencia, solo si esta tiene como norma distributiva el maximizar el bienestar del sujeto que está peor (*maximin*). El principio re-distributivo que se propone se basa en este principio de necesidad, pero se centra en el objetivo de corregir desigualdades de origen a través de

la re-distribución de la riqueza. Frohlich, Oppenheimer y Eavey (1987) en sus experimentos definen cuatro formas de operacionalizar los principios distributivos que propone Rawls: (i) maximizar el bienestar del que está peor; (ii) maximizar el bienestar del promedio; (iii) maximizar el bienestar del promedio con un piso mínimo para el que esta peor; y (iv) maximizar el bienestar del promedio con un rango máximo entre el que esta peor y el que esta mejor. En base a (iv) se construye el principio de re-distribución que se pretende desarrollar, tomando como diferencias de origen las desigualdades que la estructura de un juego origine. Es decir, si existen diferencias de pagos sistemáticas entre jugadores, independiente del perfil de estrategias que se juegue. De esta manera, un jugador con preferencias re-distributivas busca eliminar las posibles diferencias de pagos que perjudiquen estructuralmente a un jugador. En este caso, antes de iniciarse el juego el jugador i evalúa si los pagos del juego, independiente de las estrategias que se jueguen, perjudican a un jugador.

Definición 2. Sea $\delta_{kl} = \sum_{a \in A} |\pi_k(a) - \pi_l(a)|$, la suma de la diferencia de todos los pagos posibles entre algún par de jugadores k y l . Si se cumple que el máximo de los $\delta_{kl} \geq \epsilon_i$ para cualquier par de jugadores, entonces se define $a_i^R \in A_i$ tal que,

$$a_i^R \in \arg \min_{a_i \in A_i} |\pi_k(a_i, (b_{ik})_{k \neq i}) - \pi_l(a_i, (b_{il})_{l \neq i})|$$

De la misma forma, b_{ij}^R son las creencias de lo que i considera como un comportamiento redistributivo por parte de j . En este caso, si el jugador i observa que para uno o más jugadores existe una diferencia δ de la suma de los pagos posibles del juego mayor a un ϵ_i mínimo, intentará corregir esta diferencia quitándole al jugador que, previo al inicio del juego, se ve beneficiado por la estructura de este.

Cuando se habla de principio meritocrático lo que hay detrás es un principio de proporcionalidad de acuerdo a las contribuciones. Cuando el objetivo es la productividad, la norma de repartición obedece a la proporcionalidad de acuerdo a lo que se aporta, en cambio los principios de igualdad y necesidad son preferidos cuando la armonía y el bienestar son los objetivos (Barret-Howard y Tyler, 1986). Siguiendo esta línea el aumento de la productividad es algo valorado socialmente, sin embargo esto no siempre garantiza una mayor asignación para quién más aportó. Ciertamente la expresión de este principio tal como se describió implica realizar una serie de supuestos restrictivos sobre la realización de un juego y la distribución de los pagos. Para el contexto de juegos que se está desarrollando se toma la idea de aumentar la productividad a través de aumentar el pago social, y del reconocimiento de los pares de dicho esfuerzo a través de acciones similares (que apunten a aumentar el pago social). De esta forma se dice que un jugador tiene preferencias meritocráticas cuando

intenta maximizar el pago social antes que su pago individual.

Definición 3. En base a esto sea $a_i^M \in A_i$ el subset de acciones del jugador i tal que

$$a_i^M \in \arg \max_{a_i \in A_i} \sum_{i \in \bar{N}} \pi_i(a_i, (b_{ij})_{j \neq i})$$

Donde, $\bar{N} = \{j \in N : a_j^M[(b_{jk})_{k \neq j}] = b_{ij}\}$. A su vez b_{ij}^M son las creencias de lo que i considera como un comportamiento meritocrático por parte de j . Es decir, el comportamiento meritocrático del jugador i depende de que crea que el jugador j también lo es. Notar que si el jugador i cree que hay un jugador que no es meritocrático, no lo incluirá dentro del pago social que está maximizando.

Como se discutió en los capítulos anteriores, la meritocracia se entiende en un ambiente que es capaz de reconocer (premiar o castigar) los méritos. En este sentido es un concepto construido a nivel social (o al menos en una relación entre más de un sujeto y que no sea única en el tiempo). Es por esto que el comportamiento meritócrata para ser reconocido en una relación social como meritorio, debe aportar algo que en el contexto grupal sea valorado de lo contrario no se podría diferenciar de un comportamiento puramente egoísta. En un juego donde los pagos no se pueden relacionar directamente a un esfuerzo o a la habilidad, la meritocracia no puede aplicar solo a un esfuerzo de aumentar el pago individual. Si bien en la función de utilidad propuesta existe una preocupación por los pagos individuales, el esfuerzo y la habilidad que se valora debe ser contribuir al valor generado de la relación entre los agentes. Esto significa maximizar el valor generado de la relación social (o el resultado del juego), lo que puede llevar a que a veces el sujeto meritócrata sacrifique parte de su pago individual por el bienestar del grupo. En ese caso la diferencia con el altruismo es que el sujeto meritócrata espera que su comportamiento sea valorado y retribuido por los otros jugadores.

Finalmente, las definiciones para el conjunto de acciones y creencias que representa a cualquiera de los tres principios de justicia definidos también se extiende para la evaluación con respecto a un nodo específico. Así, por ejemplo $a_i^I(h)$ representa las acciones que busquen la igualdad desde el nodo h .

Definición 4. Tal como definen Dufwenberg y Kirchsteiger, la utilidad de i en la historia $h \in H$ es una función de

$$U_i : A_i \times \prod_{j \neq i} \left(B_{ij} \times \prod_{k \neq j} C_{ijk} \right) \rightarrow R$$

que para este caso se define

$$\begin{aligned}
 U_i(a_i(h), (b_{ij}(h), (c_{ijk}(h))_{k \neq j})_{j \neq i}) = \\
 \pi_i(a_i(h), (b_{ij}(h))_{j \neq i}) - \frac{\alpha_i}{n} \sum_{i \in N} (\pi_i(a_i^J(h), (b_{ij}^J(h), (c_{ijk}(h))_{k \neq j})_{j \neq i}) - \pi_i(a_i(h), (b_{ij}(h))_{j \neq i}))^2
 \end{aligned}
 \tag{3.2}$$

Donde α_i es un parámetro que pondera la importancia relativa de las preferencias por justicia del jugador i . Inicialmente α_i puede ser cualquier número real. Por simplicidad se denomina a_i^J y b_{ij}^J como las acciones y creencias correspondiente a alguno de los tipos de preferencias de justicias definidas anteriormente. Notar que esta función de utilidad es análoga a la inicialmente presentada pero contingente al juego en particular. Así el jugador compara los pagos que considera justo con lo que realmente se obtienen dentro de los pagos de un juego específico. Cabe recordar $c_{ijk}(h) = b_{jk}(h)$ con $k \neq j$, ya que para determinar el pago de referencia son necesarias las creencias del jugador i sobre las creencias del jugador j . Por ejemplo, si el jugador i y el jugador j son meritocráticos, para determinar el pago de referencia que maximiza el bienestar social el jugador i debe tener en cuenta las creencias de j sobre los demás jugadores³ que, como se verá más adelante, deben ser consistentes en el equilibrio.

Para definir el equilibrio se requiere de la siguiente notación adicional: para cualquier $a = (a_i)_{i \in N} \in A$ e historia $h \in H$, sea $A_i(h, a) \subseteq A_i$ el conjunto de estrategias de cada jugador i que prescribe las mismas elecciones que $a_i(h)$ para todas las otras historias que no sea h . Esto significa que $A_i(h, a)$ es el conjunto de estrategias que el jugador i podría usar si se comporta de acuerdo a $a_i(h)$ en todas las historias menos h , estando el jugador i libre de hacer cualquier elección en h . Notar que $a_i(h) \in A_i(h, a)$.

Definición 5. El perfil $a^* = (a_i^*)_{i \in N}$ es un Equilibrio de Justicia Secuencial (EJS) si para todo $i \in N$ y para cada historia $h \in H$ se cumple que

- 1) $a_i^*(h) \in \arg \max_{a_i \in A_i(h, a^*)} U_i(a_i(h), (b_{ij}(h), (c_{ijk}(h))_{k \neq j})_{j \neq i})$
- 2) $b_{ij} = a_j^* \forall j \neq i$.
- 3) $c_{ijk} = a_k^* \forall j \neq i, k \neq j$.

Por la condición (1) un EJS es un perfil de estrategias tal que en la historia h cada jugador hace elecciones que maximizan su utilidad dada sus creencias y que sigue su estrategia de

³Ver la descripción de \bar{N} en la definición 4.

equilibrio en otras historias. En la etapa inicial, las condiciones (2) y (3) garantizan que las creencias iniciales son correctas. A cualquier historia posterior, la condición (1) requiere que las creencias asignen probabilidad uno a la secuencia de elecciones que definen esa historia, pero son de otra manera como la creencia inicial.

Debido a (1), un EJS descarta cualquier desviación “local” rentable en cualquier etapa en particular. Sin embargo la definición no excluye la posibilidad de que una desviación conjunta en varias etapas sucesivas aumente la utilidad de un jugador evaluada en la primera etapa de la desviación. Se debe notar que las preferencias de un jugador entre dos diferentes estrategias podría cambiar dada la actualización que ocurre en cada historia. Supongamos que un jugador puede tener una desviación de múltiples etapas, que es rentable respecto a la evaluación que hace en una etapa de desvío temprana, incluso si su estrategia es parte de un EJS. Pero si este es el caso, entonces el jugador no tendría incentivos a llevar a cabo la desviación en alguna etapa posterior, puesto que la condición (1) debe aplicar a la historia donde supuestamente ocurriría el desvío posterior. Entonces, para que un desvío de múltiples etapas ocurra, el jugador debe ser capaz de obligarse a si mismo en la primera etapa de desvío a seguir una estrategia de desviación en las etapas posteriores. Si tal compromiso fuese posible por parte de un jugador, entonces debería ser modelado explícitamente, lo que cambiaría la forma en que los jugadores juegan, y por ende, llevaría a un árbol distinto.

Existencia El teorema de existencia del equilibrio se desprende directamente del desarrollado por Dufwenberg y Kirchsteiger (2004). En su teoría los autores intentan modelar la reciprocidad a través del concepto de amabilidad. La dificultad de probar existencia de un equilibrio viene de que tanto la amabilidad como la percepción de amabilidad en una historia h depende de las creencias y las creencias de segundo orden sobre las acciones que siguen a todas las historias. En particular, la amabilidad (o percepción de amabilidad) depende también de las acciones que no siguen a la historia h . De la misma forma, el comportamiento justo (y la percepción de un comportamiento justo) se evalúa tanto por el camino seguido como por el camino que no se siguió. Como en equilibrio las creencias deben ser consistentes con el perfil de estrategia de equilibrio, en un juego con incentivos de reciprocidad (o en este caso, justicia) no se puede determinar en general elecciones de equilibrio mirando juegos (y subjuegos) de manera aislada. Por lo tanto no se pueden aplicar métodos inductivos tradicionales para probar existencia (Dufwenberg y Kirchsteiger, 2004).

Esta dificultad se supera considerando todas las historias simultáneamente. En la Definición 5 se requiere que no existan desviaciones locales en alguna historia particular. Ahora si bien la utilidad de un jugador en una historia h se calcula considerando estrategias y creencias que se relacionan a elecciones en todas las historias, la maximización de utilidad se

realiza en torno a las acciones particulares en h . Esto permite la construcción de una correspondencia de mejor respuesta que considera tanto jugadores como las diferentes historias en las que se mueven. Con esto existe suficiente estructura para aplicar el teorema de punto fijo de Kakutani.

Teorema. *Existe un EJS en un juego psicológico con incentivos de justicia.*

La demostración de este teorema es análoga a la presentada por Dufwenberg y Kirchsteiger (2004) y se encuentra en el anexo 1 de este trabajo.

Capítulo 4

Aplicación y Ejemplos

En el siguiente capítulo se desarrollarán ejemplos para ilustrar el modelo. Se utilizarán juegos para mostrar cómo se comporta cada jugador de acuerdo a cada tipo o principio de justicia. Si bien se desarrollan tres principios de justicia distributiva, el principio de igualdad puede ser utilizado como referencia respecto al modelo de “Aversión a la desigualdad” de Fehr y Schmidt. Aunque las funciones de utilidad difieren, tanto el sujeto que es averso a la desigualdad como el que sigue un principio de justicia igualitario, buscan disminuir las diferencias y maximizan su utilidad cuando el resultado es igual para todos los jugadores.

4.1. Un juego de mérito e igualdad

El siguiente juego es de elaboración propia y tiene como objetivo detallar concretamente la forma en que jugadores con preferencias por la justicia juegan, cómo llegan a los respectivos equilibrios y qué hacen cuando están fuera de estos. Γ_2 es un juego secuencial de tres jugadores, por ahora no se darán valores a los parámetros de la función de utilidad, simplemente se desarrollará el juego dando las intuiciones correspondientes para cada tipo de sujeto. El equilibrio de Nash perfecto en el subjuego de Γ_2 viene dado por $a^* = (I, n, f)$ otorgando un pago de 1 al jugador 1, 1 al jugador 2 y 3 al jugador 3.

Desde la perspectiva de jugadores con preferencias meritocráticas es fácil ver que el set de acciones que maximiza el pago social viene dado por $a^M = (M, m, a)$, dando un pago social de 8. El jugador 1 analiza los pagos y juega M creyendo que los otros jugadores también son meritocráticos y tienen como referencia el mismo pago que prescribe a^M . Una vez en h1 el jugador 2 actualiza sus creencias viendo que son consistentes juega m creyendo que el jugador 3 también es meritócrata. En h3 el jugador 3 actualiza sus creencias ve que son consistentes, en este punto puede optar por jugar a y alcanzar el pago que maximiza el

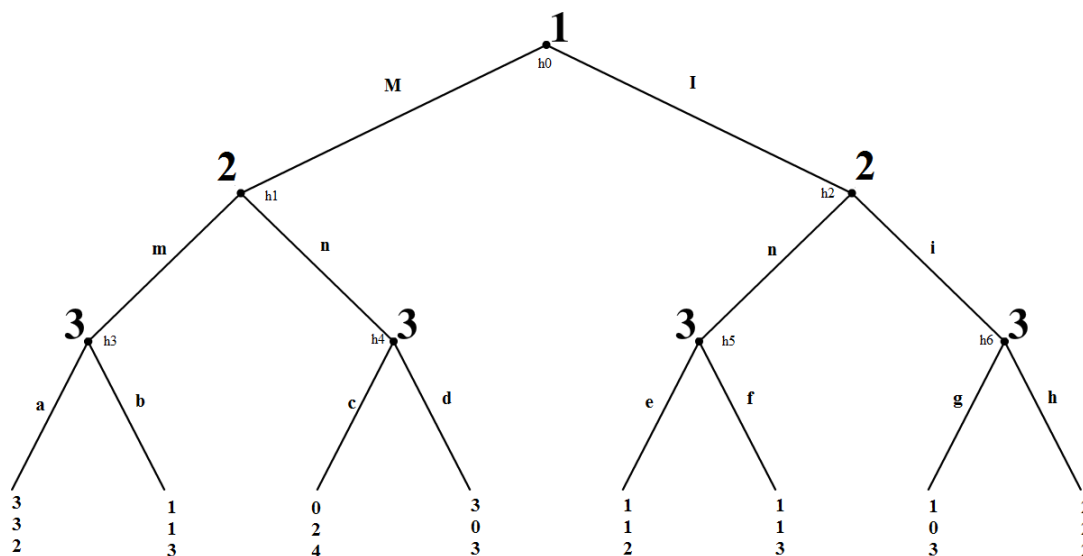


Figura 4.1: Γ_2 . Ejemplo

bienestar social o jugar b y alcanzar un pago que maximiza su pago individual. Por ahora supondremos que existe un α_3 tal que para el jugador 3 es preferible jugar a. De esta manera todos los jugadores maximizan su utilidad de acuerdo a sus preferencias y cuyas creencias son consistentes, por lo tanto se trata de un EJS.

Ahora supongamos que por alguna razón el jugador 2 quiere castigar al jugador 1 y además cree que el jugador 3 es puramente egoísta por lo que en h1 el jugador 2 juega n. Es evidente que las creencias ya no son consistentes, por lo que se está fuera de un EJS. En este caso el jugador 3 en h4 actualiza sus creencias y observa que el jugador 2 no se está comportando de manera meritocrática. En este punto el jugador 3 vuelve a calcular las utilidades y pagos de referencia. Ahora busca una nueva estrategia que optimice el pago social¹ pero que no incluya al jugador 2. En h4 esto ocurre cuando juega d, puesto que la suma de los pagos del jugador 1 y 3 es 6 mientras que si jugase c sería 4. Nuevamente ocurre que el pago individual del jugador 3 es mayor cuando juega c que cuando juega d, por lo que nuevamente debemos asumir que existe un α_3 que haga más atractivo para el jugador 3 jugar d.

En el caso de jugadores que buscan la igualdad, esta se alcanza cuando se juega $a^I = (I, i, h)$ y le da un pago de 2 a todos los jugadores. Si se asume que existe un α_i para cada jugador tal que a^I es un EJS, entonces el análisis de qué sucede cuando se está fuera del

¹Cabe recordar que en la formulación de a_i^M se requiere que los jugadores que se suman al pago social también se comporten de manera meritocrática, o al menos que así lo crea el jugador i.

equilibrio varía un poco en comparación al caso de la meritocracia. En este juego en particular no existen otros pagos que otorguen igualdad para todos los jugadores, por lo que al estar fuera del equilibrio los jugadores buscaran el resultado que minimice las diferencias y que a su vez maximice la utilidad. La diferencia sustancial con el caso meritócrata es que la igualdad no se preocupa del proceso que da a lugar un resultado sino de que este resultado sea igualitario. Como se discutió previamente, la meritocracia requiere de un ambiente social que valore el esfuerzo que un sujeto hace y lo retribuya. Es decir existe una preocupación por el cómo se está jugando, por lo que hay otros elementos de una relación social que están en juegos, como por ejemplo, la reciprocidad. En cambio la igualdad solo requiere de que el resultado sea idéntico para todos los jugadores independiente del comportamiento de estos. Es por esto que cuando un jugador no se comporta de manera igualitaria el jugador que si lo hace seguirá jugando de manera que se alcance el resultado más igualitario posible.

Finalmente, supongamos que para los jugadores que siguen un principio re-distributivo se tiene que $\epsilon_1 = \epsilon_2 = \epsilon_3 = 6$. Es fácil notar que Γ_2 beneficia estructuralmente al jugador 3, es decir que para para todos los perfiles de estrategias que se jueguen, a excepción de uno, el jugador tiene igual o mayor pago que los otros jugadores. Luego se tiene que el máx $\delta_{kl} = \delta_{23} = 11 \geq \epsilon_i$, por lo que los jugadores intentarán reducir las diferencias de pago entre los jugadores 2 y 3. Esto se logra cuando se juega el perfil $a^R = (I, i, h)$, que sería el EJS de asumir un α lo suficientemente grande para cada jugador. Ahora supongamos que los jugadores 1 y 2 no son del tipo re-distributivo o que simplemente tienen un ϵ que es muy alto (mayor que δ_{23} . En este caso, si los jugadores 1 y 2 saben que 3 es un jugador re-distributivo el resultado del juego será $a = (M, m, a)$, ya que si asumimos un α_3 lo suficientemente grande el mismo jugador 3 querrá disminuir la brecha que posee con el jugador 2.

4.2. Dilema de los prisioneros

Volviendo al ejemplo inicial del dilema de los prisioneros, un caso ampliamente estudiado en teoría de juegos y que representa una situación donde el comportamiento racional puede generar equilibrios que no son pareto-eficiente. Por simple que pueda aparentar ser el juego toca un problema que es crítico en el estudio de la sociedad. La cooperación suele ser lo mejor para el grupo (y los sujetos que lo componen) pero costoso para el individuo que coopera. Esta dualidad se observa en que los grupos con mayores niveles de cooperación tienden a sobrevivir y prosperar con mayor probabilidad que aquellos grupos con menores niveles de cooperación (Boyd y Richerson, 2005). Sin embargo esto permite la aparición de sujetos que

prefieren no cooperar y aún así beneficiarse del grupo, el fenómeno comúnmente conocido como *free riding*, por lo que el grupo genera mecanismos o instituciones que castigan este tipo de conductas (Guzmán, Rodríguez-Sickert y Rowthorn, 2007).

Este juego presenta particular interés para el estudio del modelo propuesto ya que permite comparar los costos y beneficios tanto individuales como grupales de la cooperación. De esta manera, se tiene un juego simple para dos jugadores (sin pérdida de generalidad) que permite calibrar el parámetro del modelo e ilustrar el cómo sujetos con preferencias por la justicia jugarían y los equilibrios resultantes. Se debe señalar que, dado que el dilema de los prisioneros que aquí se presenta tiene pagos simétricos, no es posible aplicar un principio re-distributivo.

Es fácil ver que en el dilema de los prisioneros ilustrado en Γ_1 el único equilibrio posible es aquel en que ninguno de los jugadores coopera y por lo tanto otorga un pago de 0 a cada jugador. Como se señaló este equilibrio no es pareto-eficiente, el resultado de la cooperación otorga pagos superiores a 0 a cada jugador, además de ser el que maximiza el bienestar social. La pregunta siguiente es entonces, ¿existe un valor para α_i que permita sustentar la cooperación como equilibrio?

Como se discutió en los capítulos previos un sujeto meritócrata intentará maximizar el pago social esperando que dicho esfuerzo sea valorado por los otros jugadores. Es fácil ver que en Γ_1 ese resultado ocurre cuando ambos jugadores cooperan. Es decir, utilizando la definición 3 del capítulo anterior se tiene que, $a^M = [C, c]$, esto implica que en el inicio del juego $\pi_1(a_1^m, b_{12}^m) = 1$ y $\pi_2(a_1^m, b_{12}^m) = 1$, lo que es análogo para el jugador 2 (cuando este también tiene preferencias meritocráticas). Una vez que ya se sabe cuáles son los pagos de referencia para el jugador meritocrático se busca un α_1 y un α_2 que, dentro de la función de utilidad propuesta, permita sustentar a la cooperación como un equilibrio.

Un cálculo simple permite determinar que un $\alpha > 0,4$ para ambos jugadores basta para sustentar un equilibrio cooperativo². Como se puede observar en la figura 2, Γ_3 es el mismo dilema de los prisioneros secuencial que Γ_1 con la diferencia que los pagos corresponden a los observados por dos jugadores con preferencias por justicia meritocrática. Al inicio del juego el jugador 1 sabe que su pago de referencia (el que maximiza el bienestar social) viene dado por que ambos jugadores cooperen y que otorga un pago de 1 a cada jugador. La creencia inicial del jugador 1 es que el jugador 2 también es meritócrata y dado los pagos observados el jugador 1 jugará C, esperando que su “esfuerzo”(o riesgo) sea valorado por el jugador 2 que a su vez le retribuirá cooperando, es decir jugando c. Este resultado del juego, si se cumple que las creencias de los jugadores son correctas, maximiza la utilidad de ambos y es

²Cabe señalar que el resultado del juego no cambia en su versión simultánea.

por definición un EJS.

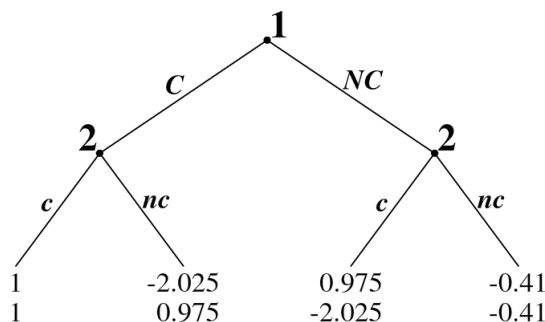


Figura 4.2: Γ_3 . Dilema de los prisioneros y meritocracia

Una vez que se conoce cuál es el equilibrio cabe preguntarse qué pasa cuando se está jugando fuera del equilibrio. En este caso, si el jugador 1 juega NC, el jugador 2 actualiza sus creencias sobre el jugador 1. Esto implica que ahora su pago de referencia cambió a $\pi_1(a_2^M(h), b_{21}(h))$ y $\pi_2(a_2^M(h), b_{21}(h))$. Donde h señala el nodo en el que el jugador 2 juega después de que el jugador 1 jugó NC. Sin embargo la definición de a_i^M requiere de que el jugador i crea que j también es meritocrático, es decir que $b_{ij}(h)^M = a_j(h)$, al no cumplirse esta condición con un jugador éste no entra en el calculo del pago social que maximiza el jugador meritócrata. En el ejemplo el jugador 2 cuando ve que el jugador 1 ha optado por no cooperar solo le queda maximizar su propia utilidad individual (puesto que hay sólo dos jugadores), lo que implica que finalmente se jugará el resultado en que ninguno de los jugadores coopera.

Para el caso de la justicia igualitaria el resultado varía un poco. En este caso un jugador con preferencias por la igualdad busca que los pagos entre los jugadores sean iguales, independiente si el resultado finalmente es óptimo para el conjunto o no. En el caso del ejemplo desarrollado hasta ahora, en Γ_1 existen dos pagos que son perfectamente iguales, el que resulta de la cooperación de ambos jugadores y que otorga un pago de 1 a cada jugador, y el que ocurre cuando ningún jugador coopera y que da un pago de 0 a cada jugador.

Los jugadores con preferencias igualitarias siguen teniendo como punto focal la colaboración, como se observa en Γ_4 [C,c] ofrece un pago que maximiza su utilidad y que además es igual para ambos jugadores. De la misma forma que para los jugadores con preferencias meritocráticas (y utilizando el mismo valor de α_i) la cooperación es también un EJS. Cabe señalar que a diferencia del caso anterior, cuando se está jugando fuera del equilibrio el jugador no cambia mayoritariamente su comportamiento, intentará minimizar las diferencias de pagos entre los jugadores. Si se observa en el subjuego de Γ_4 donde el jugador 1 jugó no cooperar, el jugador 2 no evalúa este comportamiento como egoísta, igualitario o meri-

tochrático, sino que simplemente buscará el pago que le permita maximizar su utilidad de acuerdo a sus preferencias.

En este caso, dada la estructura de pagos y la calibración del parámetro del modelo el jugador 2 optará por no colaborar, acabando con el resultado que minimiza el pago social, pero que es igual para ambos.

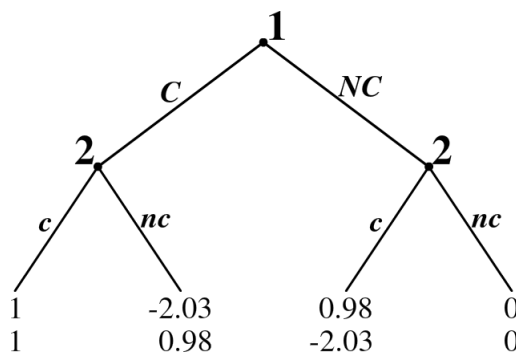


Figura 4.3: Γ_4 . Dilema de los prisioneros e igualdad.

Parte de lo que se ha discutido en este trabajo sobre la igualdad y la meritocracia se puede analizar a través del juego hasta ahora desarrollado. Si bien, para ambos tipos de preferencias (mérito e igualdad) la estructura del juego (pagos y estrategias) como también las preferencias que definen los equilibrios y desviaciones entregan resultados similares, el proceso en que se llegan a estos resultados es distinto. Mientras que el equilibrio para ambas preferencias es el mismo, esto ocurre puesto que en este juego maximizar el pago social e igualar los pagos es lo mismo. Con otro tipo de pagos esto no necesariamente es igual. El jugador meritócrata busca el pago que beneficie a todos, pero su motivación tiene como requisito que este comportamiento sea valorado y retribuido, existe una cierta preocupación por el cómo se jugó. El jugador igualitario se centra principalmente en el resultado final. Una de las diferencias principales que se han argumentado respecto a la justicia igualitaria y la justicia meritocrática.

4.3. Juego del ciempiés

El juego del ciempiés modela una situación en la cual seguir jugando es beneficioso para todos pero que a su vez los incentivos individuales están puestos en finalizarlo, es decir la ventaja la obtiene el que opta por finalizar el juego. Como se puede observar en Γ_5 , el jugador 1 inicia el juego y puede optar por finalizar el juego o cederle la opción al jugador 2, que a su vez tiene las mismas dos opciones. Cada vez que se avanza en los nodos el pago a repartir se hace más grande, sin embargo aquel que finaliza el juego es el que obtiene el pago mayor.

El único equilibrio de Nash de este juego es aquel en que el jugador 1 finaliza el juego en el primer nodo. Por inducción a la inversa el resultado es trivial, por mucho que en el nodo final el jugador 1 recibirá un pago igual o superior que en el primer nodo, el asumir un comportamiento egoísta por parte de ambos jugadores lleva a que al final siempre se quiera finalizar el juego inmediatamente.

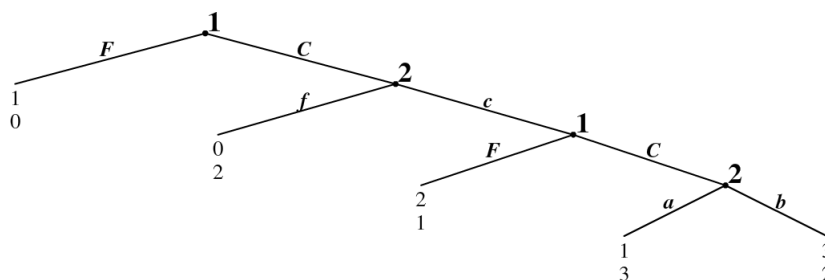


Figura 4.4: Γ_5 . Juego del Ciempiés (Dufwenberg y Kirchsteiger, 2004).

Observando el mismo juego con preferencias meritocráticas se puede ver que el pago de referencia es cuando el jugador 2 juega b en el nodo final. En este caso, utilizando el mismo valor de α_i del ejemplo anterior, para el jugador 1 es siempre conveniente seguir jugando, luego en el segundo nodo el jugador 2 actualiza su creencia y siendo también meritócrata opta por continuar jugando. Esto se repite hasta que se llega al ultimo nodo, donde finalmente el jugador 2 elije b que es el resultado que maximiza el pago social (ver Γ_6 en el anexo 2).

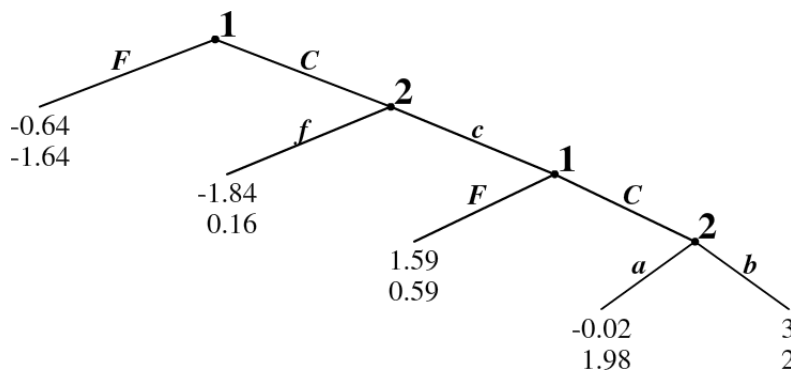


Figura 4.5: Γ_6 . Juego del Ciempiés y meritocracia.

En el caso de los jugadores con preferencias por la igualdad el proceso es un poco distinto, en este caso existen varios pagos de referencia que minimizan las diferencias entre los jugadores. Aún así, solo uno de estos pagos maximiza la utilidad o es Pareto-superior, lo que lo hace un punto focal más atractivo. En el primer nodo el jugador 1 tiene como pago de referencia 1 que es cuando juega F, que desde esa posición es el que minimiza las diferencias de pagos. Sin embargo, al observar al jugador 2 en el segundo nodo el pago de referencia

cambia, el pago que minimiza las diferencias entre ambos viene dado por que el jugador 2 juegue c y en el tercer nodo 1 juegue F otorgando pagos más altos para ambos. Una vez en el tercer nodo los jugadores actualizan sus creencias, ahora el pago de referencia es el que viene dado cuando el jugador 2 juega b en el nodo final. Se debe notar que en el nodo final, con la calibración utilizada, el jugador 2 está indiferente entre ambas opciones. Esto se ilustra en la Figura 7 del juego Γ_7 .

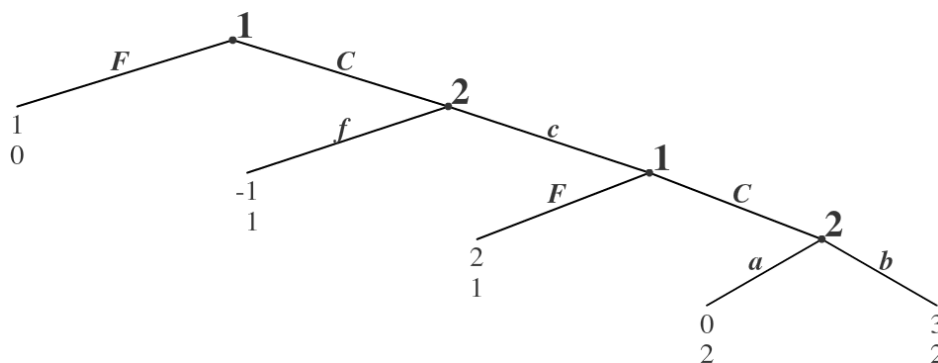


Figura 4.6: Γ_7 . Juego del Ciempiés y e igualdad.

Finalmente desde el punto de vista de jugadores con preferencias re-distributivas, en este juego la suma de los pagos favorece al jugador 2. Sin embargo, por tratarse de un juego de dos jugadores los resultados no son muy distintos al caso del principio igualitario. Reducir la brecha de pagos (asumiendo que $\delta_{12} \geq \epsilon_1 = \epsilon_2 = 1$) entre dos jugadores equivale a igualar los pagos entre todos los jugadores. Esto no ocurre para $n > 2$.

Capítulo 5

Discusión

La teoría que se ha desarrollado combina dos elementos utilizados por los modelos de “Preferencias Sociales”, por un lado la alteración de las preferencias a través de la perturbación de la función de utilidad y por el otro la utilización de un esquema de juego secuencial para explicar cambios en las estrategias de comportamiento de los sujetos. De esta manera se obtiene un marco más completo para entender el comportamiento de sujetos que actúan en concordancia a un determinado valor, en esta caso la justicia y los distintos principios distributivos que la componen.

El aporte del presente trabajo es proponer una función de utilidad que sea capaz de expresar más de un tipo de preferencia social (de justicia en este caso) sin modificar su estructura. Se trata de una función de utilidad simple en la que el sujeto compara los posibles resultados del juego con determinadas situaciones “ideales”, que son interpretadas del juego mismo en base a un escala de valores personal. El modelo es capaz de explicar distintas situaciones estratégicas, en ese sentido posee una consistencia interna. Sin embargo a nuestro juicio existe una discusión subyacente más profunda. Es difícil determinar cuál es la escala de valores que mejor explica el comportamiento social. Varios de los modelos que se han desarrollado en este trabajo tienen cierta capacidad para explicar fenómenos, sin embargo uno podría preguntarse por el trasfondo filosófico que da origen a un determinado modelo de “Preferencias Sociales”. ¿Es la igualdad el motor de las acciones humanas? ¿Son todos los sujetos altruistas? ¿O es acaso el altruismo o la envidia? Este trabajo pretende aportar incorporando un valor que no se había estudiado en el marco de la economía del comportamiento, la justicia. La justicia como valor o como ideal tiene un arraigo profundo en la historia misma de la humanidad, es un valor al que se aspira y que está presente en todas las formas de convivencia social. Es cierto que se trata de un valor contingente y que no es necesariamente universal. Este es el marco del segundo aporte de este trabajo, incorporar la

noción de justicia desde una perspectiva multidisciplinaria de las ciencias sociales al campo de la economía del comportamiento.

La estructura simple del juego secuencial como las características de la función de utilidad permiten determinar fácilmente la existencia de un equilibrio. Por otra parte los juegos elegidos para ejemplificar como funciona el modelo tienen como razón buscar las situaciones donde la cooperación y la reciprocidad es posible, aunque la teoría estándar diga lo contrario. Además, la simpleza de estos juegos permite explicar cómo se construyen estos equilibrios y cómo es el comportamiento del sujeto cuando está fuera de éste. La decisión de no utilizar juegos como el Dictador, Ultimátum o de Confianza, pasa principalmente por buscar juegos secuenciales que permitieran al sujeto evaluar las acciones de los demás para ver como comportarse de acuerdo a sus principios. Más concretamente en el juego del Dictador, por ejemplo, lo que se mide es altruismo o generosidad, es imposible a través de este juego medir comportamiento meritócrata de acuerdo a la definición utilizada. El dictador reparte una suma que fue otorgada sin esfuerzo mediante. Si no existió esfuerzo y la suma que se reparte socialmente es constante, entonces no hay posibilidad de maximizar el pago social, así como tampoco existe una medición, dentro del juego mismo, que permita establecer un parámetro de necesidad y por ende de redistribución. Lo mismo ocurre para los juegos de suma cero, en estos juegos solo se podría aplicar un criterio de igualdad, lo que ya fue ampliamente estudiado por Fehr y Schmidt. El marco general de la teoría permite explorar este tipo de juegos, pero presentan poco interés para estudiar los conceptos desarrollados hasta ahora.

Entonces lo que resta por trabajar es la construcción y el diseño de experimentos (o juegos) que permitan poner a prueba la consistencia externa del modelo, es decir, la capacidad de explicar teórica y empíricamente los alcances de la teoría propuesta. Una propuesta podría ser el desarrollo de un juego en que los sujetos se vean expuestos a situaciones estratégicas de distinta índole (cooperación, competencia, colaboración, trabajo, etc.). Por ejemplo, se podría elaborar un juego de computadora en el cual los participantes no compiten entre ellos directamente, sino que deberán “trabajar” para vencer a un oponente común. Este “trabajo” implica un costo individual para cada jugador. Dentro del juego mismo se genera un contexto donde cooperar facilite la tarea de derrotar al oponente, pero el logro individual otorgue un mayor premio. Antes de empezar a jugar habría una fase previa al inicio del juego, donde los jugadores puedan discutir y acordar normas para delimitar el comportamiento (o reparticiones) y estrategias que serán socialmente aceptadas. En base a esto se puede explorar cómo y qué a consenso se puede llegar, que sea interpretado como lo “justo” si esta característica es suficiente para que se cumpla lo acordado sin un agente externo que lo imponga. La variedad de juegos posibles se pueden desprender directamente de los distintos principios de justicia que se han trabajado. Como se discutió en secciones previas, distintas ramas de las ciencias

sociales han desarrollado el tema de la justicia, cómo esta se logra y es aprendida por los sujetos. En este sentido incorporar distintas metodologías tanto cualitativas como cuantitativas, para obtener más detalle y mejor información, abriría el campo de la presente investigación hacia un trabajo multidisciplinario. La ciencia política y la psicología experimental han abordado la discusión sobre qué principios de justicia es conveniente utilizar para orientar el orden político-económico de una comunidad. Como también la forma en que las personas pueden orientar y jerarquizar los principios de justicia por los cuales guían su actuar. Si bien dentro de la amplitud de criterios, los principios de igualdad, necesidad y equidad (en una orientación meritocrática) sobresalen como los criterios base de justicia distributiva (Sabbagh, Dar y Resh, 1994), Frohlich, Oppenheimer y Eavey (1987) creen que sujetos, desde una posición de ignorancia Rawlsiana, construyen criterios de justicia compuestos por más de un principio. Esto, entrega otra perspectiva para continuar la investigación desde la economía del comportamiento, llevando a la teoría a un grado de complejidad mayor que permita modelar sujetos que actúan con múltiples principios de justicia a la vez.

Otros factores a considerar tienen que ver con la estructura de los juegos y de las variables extra-económicas que inciden en el comportamiento económico de los sujetos. Fehr y Schmidt (1999) son enfáticos en señalar que la estructura de los juegos/experimentos es central para determinar qué tipo de preferencias son las que terminan predominando finalmente en los equilibrios resultantes. En una línea similar Cook y Hegtvedt (1983) mencionan que las dimensiones afectiva de las personas junto con la duración de la relación que puedan entablar dos sujetos son factores a considerar que pueden influir en la elección de principios de distribución. En su estudio Austin (1980) sugiere que entre amigos suele predominar una norma de igualdad, mientras que entre no-amigos la norma suele ser la de equidad o de interés personal. Cook y Yamagishi (1983) constatan que si el valor de las asignaciones es fijo, múltiples principios distributivos emergen, pero cuando este depende de la contribución de las personas, se adoptan normas más equitativas. Cook y Hegtvedt (1984) resumen que los factores que influyen las preferencias por principios distributivos en las personas incluyen: las características de la relación, el número de *inputs* relevantes y otros factores menores como la personalidad, contexto, religión, cultura, etc.

Estas ideas sustentan la necesidad de probar la validez externa del modelo con nueva evidencia de experimentos diseñados específicamente para identificar el comportamiento de los individuos según cada principio. Más aún, en el desarrollo teórico y el diseño experimental nace la necesidad de la colaboración multidisciplinaria. Tanto en la ciencia política como en la sociología y la psicología experimental se han realizado múltiples investigaciones donde el aporte y la formalización que tiene la economía del comportamiento ha sido más bien escasa. El modelo que se ha propuesto en este trabajo abarca una parte del proceso que explicaría

el comportamiento de sujetos con preferencias por la justicia. Más particularmente, cómo el individuo racional se comporta de acuerdo a un determinado valor o principio de justicia distributiva. Mientras que la ciencia política ha desarrollado la existencia de dichos principios, la raíz básica de cada uno de ellos y las diferentes combinaciones que se pueden lograr, la sociología y la psicología experimental han profundizado en la forma en que estos principios son elegidos y jerarquizados por los individuos (ver Cook y Hegtvedt, 1983; Sabbagh, Dar y Resh, 1994; Giraldo y Benitez, 2011; Shepelak y Alwin, 1986). El desafío entonces es expandir el ámbito de este trabajo (y muchos otros) en la discusión de otras disciplinas de modo de avanzar en el problema de entender el comportamiento humano y de la sociedad.

Finalmente en el área de teoría de contratos en Organización Industrial aparece un campo donde los modelos de preferencias sociales tienen terreno para trabajar. Por ejemplo Bartling y Von Siemens (2004) desarrollan un modelo de riesgo moral con múltiples agentes basado en la Aversión a la Desigualdad de Fehr y Schmidt. Los autores encuentran que la aversión a quedar perjudicado incrementa los costos de agencia para proveer incentivos, lo que también ocurre cuando al agente no le gusta estar mejor, al no ser que se puedan observar los costos del esfuerzo. Esta alza en los costos de agencia puede disminuir la eficiencia en dos sentidos. Primero, la aversión a la desigualdad puede hacer que los contratos con salarios planos se vuelvan una opción cuando los contratos con incentivos con agentes egoístas son óptimos. Segundo, para evitar la comparación social, el principal puede querer contratar solo a un sujeto, perdiendo la provisión de esfuerzo eficiente del segundo agente. En la misma línea Fehr, Gächter y Kirchsteiger (1997) usan evidencia experimental para mostrar que la reciprocidad funciona como elemento para hacer cumplir el contrato. Desde el punto de vista del diseño de contrato, existe evidencia experimental que apoya las teorías de equidad e igualdad frente al modelo estándar, tanto desde la perspectiva del principal como del esfuerzo del agente (ver Fehr y Schmidt, 2004; Fehr, Klein y Schmidt, 2007).

Capítulo 6

Conclusión

El presente trabajo ha profundizado en el problema de los alcances que puede tener el modelo estándar clásico de economía y su supuesto principal de agentes racionales que maximizan su propia utilidad según su interés egoísta. No cabe duda que éste supuesto ha permitido la utilización de las matemáticas, mecánica clásica, la lógica y el cálculo como herramienta sumamente fértiles para modelar el comportamiento humano en el campo de la economía. Y si bien su validez actual está lejos de ser descartada, si ha surgido una creciente rama en la economía del comportamiento que intenta ampliar los horizontes de la racionalidad económica.

Dentro de la economía del comportamiento aparecen las teorías de las “Preferencias Sociales”, que amplían el espacio de interés del sujeto desde el egoísmo hacia una preocupación que incluye al otro. Este tipo de teorías modifican las preferencias del sujeto y su forma de evaluar el comportamiento ajeno, pero no abandona el marco de un agente que intenta maximizar su “felicidad”. Por ende no es la racionalidad propiamente tal la que se discute, sino qué es lo que motiva al agente.

En las “Preferencias Sociales” los modelos de “Aversión a la Desigualdad” de Fehr y Schmidt o el de Bolton y Ockenfels son los que han tenido mayor impacto por su capacidad de dar cuenta de la evidencia experimental conflictiva del modelo estándar. Estos modelos se basan en que los agentes toman en cuenta la distribución de los pagos en relación a un determinado grupo de referencia. Sienten envidia (cuando están peor) o generosidad (cuando están mejor) si las asignaciones de los beneficios o recursos son dispares. Es decir, los agentes tienen como referencia la igualdad.

Lo que en este trabajo se discute es que los sujetos más que de la igualdad, se preocupan de la justicia. La igualdad es un elemento central cuando se tiene ignorancia sobre los otros y sobre la posible posición que uno pueda tener, pero este criterio no es necesariamente

universal. Lo que se propone es algo más general que la igualdad. Como se ha señalado la justicia distributiva analiza el problema sobre los principios (o valores) que fundan una norma para la distribución de los recursos, tareas y beneficios de una determinada sociedad. Como se discutió en los capítulos anteriores, estos principios varían en cada sociedad, momento histórico, contexto inmediato o la estructura de un juego en particular.

El modelo que se ha presentado tiene la ventaja de ser simple, basado en la teoría de “Aversión a la Desigualdad” de Fehr y Schmidt, pero con énfasis en los principios de justicia distributiva. Esto impone la necesidad de que el modelo pueda adaptarse a los principios de justicia estudiados (a saber igualdad, mérito y necesidad) sin modificar su estructura. Fehr y Schmidt (1999) señalan que la expresión de las preferencias por la igualdad o el egoísmo puro están estrechamente ligadas a la estructura que tiene el juego. Esto supone una limitación para su modelo, limitación que además se extiende al valor motriz que lo inspira: la igualdad. Como ya se ha señalado, la justicia es un valor más general que la igualdad y sus distintos principios permiten también expresar la heterogeneidad de las personas. La flexibilidad del modelo propuesto para expresar la preferencia por distintos valores, como también la opción por adoptar a la justicia como valor motriz, son los dos principales aportes de este trabajo. Para lograrlo se recurre a un contexto de juegos secuenciales que permite el análisis del comportamiento de los jugadores y sus posibles respuestas. Se determina la existencia de equilibrio a través de los análisis de punto fijo típicos en teoría de juegos y se prueba su consistencia en ejemplos tomados de juegos clásicos como el dilema del prisionero. Queda como objetivo para futuros trabajos el desarrollo del modelo en su relación con la evidencia experimental y su comparación con otras teorías, además de su exploración en otros campos de la economía, como la teoría de contratos, y otras ramas de las ciencias sociales.

La propuesta tiene un largo camino por recorrer, pero las perspectivas y el campo por explorar es muy amplio, dentro de la ciencia económica como también desde el punto de vista multidisciplinario. La comprensión y el estudio del comportamiento humano ha tenido grandes avances con los métodos y las herramientas de la economía clásica. Pero la complejidad del ser humano no puede ser aislada y explicada solo desde la racionalidad material-egoísta. Las múltiples dimensiones que componen el vivir tienen repercusiones en las diferentes esferas en que se desenvuelve el sujeto. He aquí un campo amplio y fértil de estudio para la economía del comportamiento y el resto de las ciencias sociales.

Bibliografía

- [1] AGELL, J., AND LUNDBORG, P. Theories of pay and unemployment: survey evidence from swedish manufacturing firms. *The Scandinavian Journal of Economics* (1995), 295–307.
- [2] ANDREONI, J., AND MILLER, J. Giving according to garp: An experimental test of the consistency of preferences for altruism. *Econometrica* 70, 2 (2002), 737–753.
- [3] ANDRÉS GUZMÁN, R., RODRÍGUEZ-SICKERT, C., AND ROWTHORN, R. When in rome, do as the romans do: the coevolution of altruistic punishment, conformist learning, and cooperation. *Evolution and Human Behavior* 28, 2 (2007), 112–117.
- [4] AUSTIN, W. Friendship and fairness effects of type of relationship and task performance on choice of distribution rules. *Personality and Social Psychology Bulletin* 6, 3 (1980), 402–408.
- [5] BARRETT-HOWARD, E., AND TYLER, T. R. Procedural justice as a criterion in allocation decisions. *Journal of Personality and Social Psychology* 50, 2 (1986), 296.
- [6] BARTLING, B., AND VON SIEMENS, F. Inequity aversion and moral hazard with multiple agents. Tech. rep., Citeseer, 2003.
- [7] BECKER, G. A theory of social interactions, 1974.
- [8] BERG, J., DICKHAUT, J., AND MCCABE, K. Trust, reciprocity, and social history. *Games and economic behavior* 10, 1 (1995), 122–142.
- [9] BLINDER, A. S., AND CHOI, D. H. A shred of evidence on theories of wage stickiness. *The Quarterly Journal of Economics* 105, 4 (1990), 1003–1015.
- [10] BOLTON, G. E. A comparative model of bargaining: Theory and evidence. *The American Economic Review* (1991), 1096–1136.

- [11] BOLTON, G. E., AND OCKENFELS, A. Erc: A theory of equity, reciprocity, and competition. *American economic review* (2000), 166–193.
- [12] BOWLES, S. *Social Preferences and Public Economics: Are good laws a substitute for good citizens?* Università degli Studi di Siena, 2007.
- [13] BOYD, R., AND RICHESON, P. J. Solving the puzzle of human cooperation. *Evolution and culture* (2005), 105–132.
- [14] CAMERER, C. *Behavioral game theory: Experiments in strategic interaction*. Princeton University Press, 2003.
- [15] CAMERER, C., AND THALER, R. H. Anomalies: Ultimatums, dictators and manners. *The Journal of Economic Perspectives* (1995), 209–219.
- [16] CAMPBELL III, C. M., AND KAMLANI, K. S. The reasons for wage rigidity: evidence from a survey of firms. *The Quarterly Journal of Economics* (1997), 759–789.
- [17] COOK, K. S., AND HEGTVEDT, K. A. Distributive justice, equity, and equality. *Annual review of sociology* (1983), 217–241.
- [18] COOK, K. S., AND YAMAGISHI, T. Social determinants of equity judgments: The problem of multidimensional input. *Equity theory: Psychological and sociological perspectives 95* (1983), 126.
- [19] COX, J., SADIRAJ, K., AND SADIRAJ, V. Trust, fear, reciprocity. Tech. rep., and altruism. Mimeo, University of Arizona, 2001.
- [20] DAVIS, D., AND HOLT, C. Experimental economics, 1993. *Princeton Univ. press, Princeton*.
- [21] DUFWENBERG, M., AND KIRCHSTEIGER, G. A theory of sequential reciprocity. *Games and economic behavior* 47, 2 (2004), 268–298.
- [22] FEHR, E., AND FALK, A. Wage rigidity in a competitive incomplete contract market. *Journal of political Economy* 107, 1 (1999), 106–134.
- [23] FEHR, E., AND GÄCHTER, S. Cooperation and punishment in public goods experiments. *American Economic Review* (2000), 980–994.

-
- [24] FEHR, E., GÄCHTER, S., AND KIRCHSTEIGER, G. Reciprocity as a contract enforcement device: Experimental evidence. *Econometrica: journal of the Econometric Society* (1997), 833–860.
- [25] FEHR, E., KIRCHSTEIGER, G., AND RIEDL, A. Gift exchange and reciprocity in competitive experimental markets. *European Economic Review* 42, 1 (1998), 1–34.
- [26] FEHR, E., KLEIN, A., AND SCHMIDT, K. M. Fairness and contract design. *Econometrica* 75, 1 (2007), 121–154.
- [27] FEHR, E., AND SCHMIDT, K. M. A theory of fairness, competition, and cooperation. *The quarterly journal of economics* 114, 3 (1999), 817–868.
- [28] FEHR, E., AND SCHMIDT, K. M. Fairness and incentives in a multi-task principal-agent model. *The Scandinavian Journal of Economics* 106, 3 (2004), 453–474.
- [29] FEHR, E., AND SCHMIDT, K. M. The economics of fairness, reciprocity and altruism—experimental evidence and new theories. Tech. rep., SFB/TR 15 Discussion Paper, 2005.
- [30] FORSYTHE, R., HOROWITZ, J. L., SAVIN, N. E., AND SEFTON, M. Fairness in simple bargaining experiments. *Games and Economic behavior* 6, 3 (1994), 347–369.
- [31] FROHLICH, N., OPPENHEIMER, J. A., AND EAVEY, C. L. Choices of principles of distributive justice in experimental groups. *American Journal of Political Science* (1987), 606–636.
- [32] FUDENBERG, D., AND TIROLE, J. Game theory. *MIT Press Books* 1 (1991).
- [33] GÄCHTER, S., AND FALK, A. Reputation or reciprocity? an experimental investigation. Tech. rep., CESifo Working Paper, 2001.
- [34] GEANAKOPOLOS, J., PEARCE, D., AND STACCHETTI, E. Psychological games and sequential rationality. *Games and Economic Behavior* 1, 1 (1989), 60–79.
- [35] GIRALDO, J. J., AND BENÍTEZ, E. O. Psicología de la justicia distributiva: antecedentes y tipos de estudios para su comprensión. *Acta Colombiana de Psicología* 14, 2 (2011), 91–101.
- [36] GÜTH, W., MARCHAND, N., AND RULLIERE, J.-L. *On the reliability of reciprocal fairness: An experimental study*. Humboldt-Univ., Wirtschaftswiss. Fak., 1997.

-
- [37] GÜTH, W., SCHMITTBERGER, R., AND SCHWARZE, B. An experimental analysis of ultimatum bargaining. *Journal of economic behavior & organization* 3, 4 (1982), 367–388.
- [38] GÜTH, W., AND TIETZ, R. Ultimatum bargaining behavior: A survey and comparison of experimental results. *Journal of Economic Psychology* 11, 3 (1990), 417–449.
- [39] HAYEK, F. A. *The sensory order*. University of Chicago Press, 1976.
- [40] KIRCHSTEIGER, G. The role of envy in ultimatum games. *Journal of economic behavior & organization* 25, 3 (1994), 373–389.
- [41] LEVINE, D. K. Modeling altruism and spitefulness in experiments. *Review of economic dynamics* 1, 3 (1998), 593–622.
- [42] MICHAEL, S. Liberalism and the limits of justice. *Cambridge University* (1982), 55–57.
- [43] MICHELBACH, P. A., SCOTT, J. T., MATLAND, R. E., AND BORNSTEIN, B. H. Doing rawls justice: An experimental study of income distribution norms. *American Journal of Political Science* 47, 3 (2003), 523–539.
- [44] MILLER, D. *Principles of social justice*. Harvard University Press, 1999.
- [45] NOZICK, R. *Anarchy, state, and utopia*, vol. 5038. Basic books, 1974.
- [46] OKUN, A. M. *Equality and efficiency, the big tradeoff*. Brookings Institution Press, 1975.
- [47] OSTROM, E., AND WALKER, J. Communication in a commons: cooperation without external enforcement. *Laboratory research in political economy* (1991), 287–322.
- [48] RABIN, M. Incorporating fairness into game theory and economics. *The American economic review* (1993), 1281–1302.
- [49] RAWLS, J. *Teoría de la justicia*. Fondo de cultura económica, 2012.
- [50] ROTH, A. Bargaining experiments, kagel jh, roth ae, handbook of experimental economics, 1995, 253-348.
- [51] ROTH, A. E., MALOUF, M. W., AND MURNIGHAN, J. K. Sociological versus strategic factors in bargaining. *Journal of Economic Behavior & Organization* 2, 2 (1981), 153–177.

- [52] SABBAGH, C., DAR, Y., AND RESH, N. The structure of social justice judgments: A facet approach. *Social Psychology Quarterly* (1994), 244–261.
- [53] SAMUELSON, P. A. Altruism as a problem involving group versus individual selection in economics and biology. *The American Economic Review* (1993), 143–148.
- [54] SEN, A. K., BRAVO, A. M., AND SCHWARTZ, P. *Nuevo examen de la desigualdad*. Alianza Madrid, 1995.
- [55] SHEPELAK, N. J., AND ALWIN, D. F. Beliefs about inequality and perceptions of distributive justice. *American Sociological Review* (1986), 30–46.
- [56] SHER, G. *Desert*. Princeton University Press, 1989.
- [57] SMITH, A. The theory of moral sentiments.
- [58] SMITH, V. L. An experimental study of competitive market behavior. *The Journal of Political Economy* (1962), 111–137.
- [59] SMITH, V. L., AND WILLIAMS, A. W. The boundaries of competitive price theory: Convergence, expectations, and transaction costs. *Advances in behavioral economics 2* (1990), 3–35.
- [60] WEBER, M. *Economía y sociedad: esbozo de sociología comprensiva*, vol. 2051. Fondo de cultura económica, 1964.

Apéndice A

Anexos

A1. Existencia de EJS

Sea $X_i(h)$ el conjunto de elecciones (o posiblemente distribuciones de probabilidad) de i en la historia $h \in H$. Si $x \in X_i(h)$, sea $a_i(h) \setminus x$ la estrategia de i que es igual a $a_i(h)$ pero que especifica la elección x en h . Se definen las correspondencias $\beta_{i,h} : A \rightarrow X_i(h)$ y $\beta : A \rightarrow \prod_{(i,h) \in N \times H} X_i(h)$ como

$$\beta_{i,h}(a) = \arg \max_{x \in X_i(h)} U_i(a_i(h) \setminus x, (a_j(h), (a_k(h))_{k \neq j})_{j \neq i}),$$

$$\beta(a) = \prod_{(i,h) \in N \times H} \beta_{i,h}(a)$$

Notar que el conjunto $\prod_{(i,h) \in N \times H} \beta_{i,h}(a)$ es topológicamente equivalente a A , por lo que β es equivalente a una correspondencia $\gamma : A \rightarrow A$ definida de manera conveniente para que los puntos fijos sean EJS's. Para esto se debe notar que $\beta_{i,h}$ sirve como la condición (1) de la Definición 5, tomando las creencias correctas como dicta las condiciones (2) y (3). Por lo tanto $\beta_{i,h}$ efectivamente encuentra las estrategias óptimas en $A_i(h, a)$, de acuerdo a la condición (1), salvo que se utilizan las elecciones óptimas en $X_i(h)$. β y γ son correspondencias de mejor respuesta combinadas, y como γ se define de A hacia A , es posible hacer un análisis de punto fijo.

Por el principio de máximo de Berge se garantiza que $\beta_{i,h}$ es no vacío, cerrado, y semi-continuo superior, puesto que $X_i(h)$ es no vacío y compacto y U_i es continua (dado que π_i es continua). $\beta_{i,h}$ es convexa, dado que $X_i(h)$ es convexo y U_i es cuasi-cóncava en las propias elecciones de i . Por lo tanto $\beta_{i,h}$ es no vacío, cerrado, semi-continuo superior, y convexa. Estas propiedades se extienden a β y γ . Por las propiedades del teorema de punto fijo de

Kakutani se desprende que γ tiene un punto fijo.