POSTGRADO
ECONOMÍA Y NEGOCIOS
UNIVERSIDAD DE CHILE

Escasez y manipulación en mercados bilaterales sin opciones de salida

TESIS PARA OPTAR AL GRADO DE MAGISTER EN ECONOMIA

Alumno: Camilo Maximiliano Jaramillo Sirguiado
Profesor Guía: Juan Pablo Torres-Martínez

Abril 2024

ESCASEZ Y MANIPULACIÓN EN MERCADOS BILATERALES SIN OPCIONES DE SALIDA

CAMILO JARAMILLO SIRGUIADO†

Resumen. Manipulación y estabilidad son aspectos claves en el diseño de mercados de emparejamiento bilateral. Como consecuencia del Teorema de Imposibilidad de Roth (1982), parte de la literatura se ha concentrado en comparar mecanismos estables en base a su grado de manipulación. En esta tesis estudiamos el efecto de la ausencia de opciones de salida sobre la comparación de mecanismos estables, cuando ambos lados del mercado son estratégicos. Nuestra principal conclusión es que la escasez es determinante en la búsqueda de mecanismos estables que minimicen la manipulación.

1. Introducción

En las últimas décadas, la teoría de emparejamientos se ha consolidado como una de las áreas más activas de la teoría microeconómica. Esto se debe al gran número de aplicaciones que se ha conseguido implementar en diferentes áreas, como por ejemplo, sistemas de elección escolar, asignación de objetos, residencia de médicos y plataformas de intercambio de riñón. Uno de los principales problemas en los que se ha centrado la literatura es como asignar emparejamientos, induciendo que los participantes no tengan incentivos a actuar de manera estratégica. A los sistemas centralizados que cumplen esta propiedad se les denomina mecanismos *strategy-proof*.

Otra propiedad importante dentro de esta literatura es la *estabilidad*, la cual consiste en que ningún par de agentes tenga incentivos a juntarse para desviar del emparejamiento asignado. Lamentablemente, en mercados bilaterales uno-a-uno, no existe un mecanismo estable y strategy-proof (Roth, 1982). Este resultado de imposibilidad ha generado que la teoría de emparejamientos se enfoque en lograr que al menos un lado del mercado no tenga incentivos a actuar de manera estratégica, propiedad conocida como *one-side strategy-proofness*, la cual es satisfecha por un único mecanismo estable: el mecanismo de aceptación diferida (Alcalde & Barberà, 1994). Este ha sido un argumento clave en el cambio de sistemas de elección escolar llevado a cabo en los últimos 20 años, donde el mecanismo de aceptación diferida, con los estudiantes haciendo las propuestas, ha sido fundamental.

Aunque el mecanismo de aceptación diferida fue recomendado para la asignación de estudiantes a colegios, al momento de su implementación las autoridades descuidaron un detalle clave. En concreto, la mayoría de los sistemas escolares permitían reportar preferencias con un tope máximo en la cantidad de colegios. A este mecanismo, la literatura le denomina aceptación diferida *restringida*, el

† Agradezco a mi profesor guía, Juan Pablo Torres-Martinez, por mostrarme lo entretenido que es trabajar en teoría económica y su capacidad (quizás sin querer) de sacar lo mejor de sus estudiantes.

cual es estable y obviamente no es strategy-proof para los estudiantes (por el teorema de unicidad antes citado). Así, al adoptar este mecanismo, los estudiantes sí tenían incentivos a actuar de manera estratégica, una de las principales críticas hacia los sistemas escolares anteriores.

Con el fin de aliviar esta preocupación, en Pathak & Sönmez (2013) los autores muestran que el mecanismo de aceptación diferida restringida es menos manipulable por los estudiantes que su antecesor, el mecanismo de Boston (restringido). Este artículo abrió un nuevo tópico dentro de la teoría de emparejamientos: la comparación de mecanismos en términos de qué tan manipulables son.

Considerando esta área de investigación es natural preguntarse: ¿existe un mecanismo estable que sea el menos manipulable? De existir, podría ser una nueva solución al Teorema de Imposibilidad de Roth, en particular, en contextos en que el diseñador esta interesado en reducir la manipulación en ambos lados del mercado. Obviamente, para responder esta pregunta se necesita una métrica o criterio que permita comparar mecanismos en torno a qué tan manipulables son.

En Pathak & Sönmez (2013), un mecanismo es débilmente menos manipulable que otro, si en cada perfil de preferencias bajo el cual es manipulable, el otro también lo es (para una definición formal, véase la Sección 2 de esta tesis). A lo largo de este trabajo, llamaremos a este criterio **comparación en preferencias**. Aunque esta métrica es útil para comparar ciertos mecanismos, es poco informativa a la hora de comparar a los que cumplen estabilidad. Bonkoungou & Nesterov (2023, Proposición 1) muestran que bajo un perfil de preferencias cualquiera, todo mecanismo estable es manipulable o ninguno lo es. Esto implica que bajo el criterio de Pathak & Sönmez (2013) todas las reglas de asignación estables son igual de manipulables.

La cantidad de agentes manipuladores es otra medida que se ha utilizado en la literatura. En Bonkoungou & Nesterov (2023, Teorema 1) los autores comparan las versiones restringidas de los mecanismos de aceptación diferida y de Boston en términos de cuántos individuos los manipulan. A este métrica la llamaremos **comparación en cantidad de agentes** ¿Qué ocurre al comparar reglas de asignación estables bajo este criterio? Bonkoungou & Nesterov (2023, Corolario 1) muestra que las dos versiones del mecanismo de aceptación diferida (dependiendo del lado que hace las propuestas), minimizan la cantidad de agentes manipuladores en todo perfil de preferencias. Adicionalmente, Sirguiado (2024) caracteriza la familia de mecanismos estables que cumplen esta propiedad.

Nuestro objetivo es mostrar el efecto de la ausencia de opciones de salida sobre la comparación de mecanismos estables. Para esto, utilizaremos los dos criterios de comparación antes descritos. Diremos que un agente no posee opciones de salida cuando sus preferencias están definidas solo sobre el otro lado del mercado o cuando su opción de salida (de existir) es su peor alternativa.[1] Para que nuestro concepto de strategy-proofness haga sentido, asumiremos que el diseñador de mecanismos sabe que los participantes del mercado no poseen opciones de salida.

Para tener un contexto en mente, todos nuestros resultados son presentados en el siguiente escenario: un conjunto de estudiantes de pregrado debe emparejarse con un conjunto de profesores

[1]Abusaremos de esta imprecisión solo para facilitar la descripción de los resultados.

para que sean sus guías de tesis. Asumiremos que cada profesor puede tener a cargo a lo más un tesista. En este contexto, uno puede imaginar que tanto guías como estudiantes pueden no poseer opciones de salida. Por el lado de los profesores, es posible que estén obligados por contrato a considerar a todos los alumnos como aceptables. En el caso de los estudiantes, es poco creíble que prefieran no completar sus estudios, a tener que ser emparejado con un guía no deseado. De esta forma, hace sentido que no existan opciones de salida en este contexto.

Concentrándonos en mercados bilaterales uno-a-uno, cuando todos los agentes son estratégicos, el mercado es desbalanceado y nadie tiene opciones de salida, mostramos que:

(a) Bajo el criterio de Pathak & Sönmez (2013), el mecanismo de aceptación diferida, con el lado corto haciendo las propuestas, es débilmente menos manipulable que cualquier mecanismo estable. Además, existen perfiles de preferencias bajo los cuales este mecanismo no es manipulado y aceptación diferida, con el lado largo haciendo las propuestas, si lo es. (ver Teorema 1).
(b) Bajo el criterio de Bonkoungou & Nesterov (2023), el mecanismo de aceptación diferida, con el lado corto haciendo las propuestas, minimiza la cantidad de agentes manipuladores en todo perfil de preferencias. Además, esta propiedad no es alcanzada por ningún mecanismo estable y strategy-proof para el lado largo del mercado (ver Teorema 2).
(c) Existen mecanismos estables, diferentes de aceptación diferida, que también minimizan la cantidad de agentes manipuladores en cada perfil de preferencias (ver Proposición 1).
(d) Dentro de los mecanismos estables y strategy-proof para los agentes que pertenecen al lado largo del mercado, existe un mecanismo que minimiza la cantidad de agentes manipuladores en cada perfil de preferencias (ver Proposición 2).

Estos resultados muestran que la escasez del mercado, junto a la ausencia de opciones de salida, juegan un rol fundamental sobre la capacidad de manipulación de los individuos. Los items (a) y (b) muestran que en este contexto hace sentido comparar mecanismos estables bajo ambas métricas, algo que no ocurre en mercados bilaterales clásicos. Adicionalmente, el item (c) nos dice que es posible construir mecanismos que minimizan la cantidad de agentes manipuladores en cada perfil de preferencias, difieren de aceptación diferida, y son más justos en términos de bienestar para los agentes que pertenecen al lado largo del mercado. Notar que el Teorema 2 implica que tales mecanismos no son strategy-proof para ningún lado del mercado. Por último, el item (d) es importante en contextos en los que la autoridad desea minimizar la cantidad de agentes manipuladores sujeto a que el mecanismo sea strategy-proof para el lado del mercado que posee más participantes.

¿Qué ocurre cuando no hay un desbalance en el mercado? En este caso utilizamos la simetría del problema para obtener un resultado de imposibilidad. En concreto, en mercados bilaterales uno-a-uno balanceados, mostramos que:

- No existe un mecanismo estable y one-side strategy-proof, que minimice la cantidad de agentes manipuladores (ver Teorema 3 (i)).

- Las dos versiones del mecanismos de aceptación diferida no son comparables entre sí, bajo ninguno de nuestros criterios de comparación (ver Teorema 3 (ii)).

Esto sugiere que, para que la comparación de mecanismos estables sea útil para escoger sistemas centralizados de emparejamiento, es necesario algún tipo de ventaja para algún lado del mercado.

Para concluir nuestra investigación, nos concentramos en la comparación de mecanismos estables a nivel individual, concepto introducido en Pathak & Sönmez (2013) y generalizado para contextos muchos-a-muchos en Chen, Egesdal, Pycia & Yenmez (2016). En este último trabajo, los autores muestran que un mecanismo estable es menos manipulable que otro, desde la perspectiva del individuo h si y solo si es más preferido por h. Este resultado muestra una estrecha relación entre la capacidad de manipular y las preferencias de los agentes en mercados de emparejamiento bilateral.

¿Cómo se ve afectada esta equivalencia por la ausencia de opciones de salida? Siguiendo la estrategia de demostración del resultado original, mostramos que esta equivalencia es valida solo en mercados desbalanceados y solo para aquellos agentes que pertenecen al lado corto del mercado (ver Teorema 4). De hecho, ninguna de las direcciones es valida para los agentes que no pertenecen al lado corto.

El Cuadro 1 resume los principales resultados de esta tesis, para mercados bilaterales conformados por los conjuntos E y G. Se puede ver que, en contextos en que los agentes no poseen opciones de salida, el balance del mercado juega un rol fundamental en la comparación de mecanismos estables. Esto no ocurre cuando todos los participantes del mercado poseen opciones de salida.

Cuadro 1. El rol de la ausencia de opciones de salida sobre la comparación de mecanismos estables

Agentes poseen opciones de salida	Si	No	
Mercados bilaterales uno-a-uno		$\lvert E\rvert < \lvert G\rvert$	$\lvert E\rvert = \lvert G\rvert$
1. Comparación a la Pathak & Sönmez (2013)			
AD_E minimiza manipulación	✓	✓	×
AD_G minimiza manipulación	✓	×	×
AD_E y AD_G son comparables	✓	✓	×
Se cumple Chen et al. (2016, Teorema 1) para todo agente	✓	×	×
2. Comparación a la Bonkoungou & Nesterov (2023)			
AD_E minimiza manipulación	✓	✓	×
Hay un mecanismo en $\mathbb{MS} \cap \mathbb{SP}_{\mathbb{G}}$ que minimiza la manipulación	✓	×	×
AD_E y AD_G son comparables	✓	✓	×
Hay un mecanismo en $\mathbb{MS} \cap (\mathbb{SP}_{\mathbb{E}} \cup \mathbb{SP}_{\mathbb{G}})$ que minimiza la manipulación	✓	✓	×

Nota: En azul nuestros resultados. $\mathbb{MS}$ y $\mathbb{SP}_{\mathbb{H}}$ denotan los conjuntos de mecanismos estables y strategy-proof para el lado H, respectivamente.

Literatura relacionada. Al mejor de nuestro conocimiento, este trabajo es el primero en combinar dos tópicos de la literatura de *matching*: la comparación de mecanismos en términos de su manipulabilidad y los efectos de la ausencia de opciones de salida.

Respecto a la comparación de mecanismos, como ya lo hemos hecho notar, nuestro trabajo es fuertemente influenciado por Pathak & Sönmez (2013), Chen, Egesdal, Pycia & Yenmez (2016), y Bonkoungou & Nesterov (2023). Respecto al primero de estos trabajos, muchos artículos han utilizado su criterio de comparación para aplicarlo a diferentes contextos y mecanismos (Chen & Kesten (2017); Dur (2019); Dur, Hammond & Morrill (2019); Dur, Pathak, Song & Sönmez (2022)). En el caso de la comparación en cantidad de agentes, este fue utilizado por primera vez en el área de diseño de mecanismos en Andersson, Ehlers y Svensson (2014a), donde los autores comparan mecanismos justos y balanceados contando a los individuos que los manipulan (en un contexto de asignación de objetos y dinero).

Dentro del área de teoría de emparejamientos, los primeros en incorporar esta noción fueron Imamura & Tomoeda (2022), y Bonkoungou & Nesterov (2023). Ambos trabajos se llevaron a simultáneamente (de manera independiente) y compararon las versiones restringidas del mecanismo de Boston y de aceptación diferida, contando la cantidad de agentes manipuladores. Los autores muestran que el primero de estos mecanismos posee una mayor o igual cantidad de agentes manipuladores en todo perfil de preferencias. Adicionalmente, Bonkoungou & Nesterov (2023) muestra que la cantidad de individuos que manipulan aceptación diferida restringida decrece a medida que se hace menos restrictivo el mecanismo.

Dentro de la literatura se han utilizado otros criterios de comparación. En Bonkoungou & Nesterov (2021), los autores utilizan el concepto de accesibilidad estratégica para comparar el mecanismo de aceptación diferida restringida con una familia de mecanismos denominados *first-preference-first*. Adicionalmente, logran demostrar que contra menos se restrinja el mecanismo de aceptación diferida, menos estratégicamente accesible será. Otra métrica es la usado en Decerf & Van der Linden (2021), donde los autores utilizan la idea de *inclusión de preferencias dominantes*, introducida en Arribillaga and Massó (2016), para comparar los mecanismos restringidos de aceptación diferida y de Boston.

Recientemente, Nesterov, Rospuskova & Rubtcova (2024) plantearon el concepto de *robustez a manipulaciones*, criterio de comparación que va en linea con los de accesibilidad estratégica e inclusión de preferencias dominantes. Los autores muestran que todos los resultados obtenidos utilizando estos, así como el introducido por Pathak & Sönmez (2013), pueden ser replicados utilizado la robustez a manipulaciones.

La literatura de teoría de emparejamientos y opciones de salida es relativamente escasa. En Teo, Sethuraman, and Tan (2001) los autores muestran cómo se ve afectada la capacidad de manipulación de un individuo, cuando este no posee opciones de salida. En concreto, en un mercado bilateral balanceado sin opciones de salida, demuestran que cuando se aplica el mecanismo óptimo para las mujeres, existen hombres que no son capaces de alcanzar su mejor pareja estable mediante una manipulación. El Lema 1 de esta tesis, clave en mucha de nuestras pruebas, muestra que esto no es cierto cuando los hombres conforman el lado escaso.[2] Adicionalmente, Tadenuma & Toda (1998) muestran que existe un mecanismo estable y Nash implementable en un mercado balanceado si y solo si ambos lados del mercado poseen dos agentes y nadie posee opciones de salida.

En contextos de elección escolar, Kesten (2010) y Kesten & Kurino (2019) estudian cómo la ausencia de opciones de salida afecta la existencia de mecanismos strategy-proof para los estudiantes, que dominen al mecanismo de aceptación diferida. Recientemente, Sirguiado & Torres-Martinez (2024) mostraron que la ausencia de opciones de salida compromete el resultado de unicidad de Alcalde & Barberà (1994). Adicionalmente, los autores concluyen que dentro de la familia de mecanismos estables y strategy-proof para un lado del mercado, siempre existe un mecanismo que es el mejor para el otro lado. Por último, caracterizan los escenarios en que tal mecanismo no coincide con aceptación diferida. Como se verá más adelante, este resultado es clave en alguna de nuestras demostraciones (ver Proposición 2 y Teorema 4).

En un contexto de elección escolar, Akbarpour, Kapor, Neilson, Van Dijk & Zimmerman (2022) muestran que, aquellos estudiantes que poseen opciones de salida, siempre preferirán mecanismos manipulables antes que strategy-proof (independiente de si estos son estables o no). De esta forma, estudiantes que tienen opciones de salida (por ejemplo, que tienen el poder adquisitivo para pagar educación privada) se ven beneficiados por sistemas de elección escolar que son manipulables. Este resultado va en la dirección contraria a lo mostrado en Chen, Egesdal, Pycia & Yenmez (2016), donde los agentes que poseen opciones de salida siempre prefieren los mecanismos estables menos manipulables (ver Sección 4 de esta tesis).

El resto de esta tesis esta organizada de la siguiente forma. Mientras la Sección 2 describe nuestro modelo, la Sección 3 presenta nuestros principales resultados, y se divide en mercados desbalanceados y balanceados. La Sección 4 estudia la comparación de mecanismos a nivel individuo y la Sección 5 presenta las principales conclusiones de este trabajo. Por último, en el Apéndice se muestra un nuevo resultado en mercados con opciones de salida, sumado a algunas demostraciones omitidas en el cuerpo principal de esta tesis.

2. Modelo

Considere un mercado de emparejamientos bilateral uno-a-uno $[E, G, \succ]$ (o simplemente *un mercado*), en el cual un conjunto finito de estudiantes (E) deben emparejarse con un conjunto finito de guías de tesis (G). Denotaremos por $\emptyset$ el no ser emparejado y definiremos a $\succ = (\succ_h)_{h\in E\cup G}$ como un perfil de preferencias, bajo el cual:

- Para cada $e \in E$, $\succ_e$ es un orden lineal definido sobre G.
- Para cada $g \in G$, $\succ_g$ es un orden lineal definido sobre E.

Note que al no tener tener opciones de salida, las preferencias de cada estudiante están definidas solo sobre los guías, y no sobre $G \cup \{\emptyset\}$. Lo mismo ocurre para los guías. Esta es la única diferencia entre un mercado bilateral clásico y nuestro contexto.

[2]Este resultado es utilizado en la prueba del Teorema 2 de Sirguiado & Torres-Martinez (2024).

Diremos que un mercado $[E, G, \succ]$ es *desbalanceado* si la cantidad de estudiantes difiere de la cantidad de guías. Análogamente, un problema es *balanceado* solo si $|E| = |G|$.

Un *matching* es un mapeo $\mu : E \cup G \to E \cup G$ que distribuye estudiantes y guías en parejas de tal forma que:

(i) Para cada $e \in E$, $\mu(e) \in G \cup \{\emptyset\}$.
(ii) Para cada $g \in G$, $\mu(g) \in E \cup \{\emptyset\}$.
(iii) Para cada $h \in E \cup G$, $\mu(\mu(h)) = h$.

Un matching μ es:

- *bloqueado* si existe un par $(e, g) \in E \times G$ tal que $g \succ_e \mu(e)$ y $e \succ_g \mu(g)$.
- *estable* si no es bloqueado por ningún par de agentes.[3]

Sea $\mathcal{M}$ y $\mathcal{P}$ el conjunto de todos los matchings y de todos los perfiles de preferencias. Un *mecanismo* es un protocolo centralizado que asocia a cada perfil de preferencias $\succ \in \mathcal{P}$ un matching $\mu \in \mathcal{M}$.

Dado $\Omega : \mathcal{P} \to \mathcal{M}$, diremos que el agente $h \in E \cup G$ *manipula* el mecanismo Ω bajo el perfil $\succ$ si existe un perfil de preferencias $\succ' \in \mathcal{P}$ tal que $\Omega[\succ'_h, \succ_{-h}] \succ_h \Omega[\succ]$. Además, definimos las siguientes propiedades:

- Ω es *estable* cuando para cada $\succ \in \mathcal{P}$ el matching $\Omega[\succ]$ es estable bajo $\succ$.
- Dado un lado del mercado, $H \in \{E, G\}$, Ω es *strategy-proof para* H, cuando para todo $\succ \in \mathcal{P}$ ningún $h \in H$ manipula el mecanismo Ω bajo $\succ$.
- Ω es *manipulable* bajo un perfil de preferencias, si existe un agente que lo manipula en tal perfil.

Como nuestro objetivo es comparar mecanismos estables en términos de su grado de manipulación, necesitamos definir cuales serán los criterios que utilizaremos para realizar tal comparación. En esta tesis nos concentramos en dos criterios que previamente han sido usados en la literatura de manipulación de mecanismos. A continuación, definimos formalmente cada una de estas métricas de comparación, tanto en sus formas débiles, como en sus formas más estrictas. Para esto, denotaremos por $M^{\psi}[\succ]$ al conjunto de agentes que manipulan el mecanismo ψ bajo el perfil $\succ$.

Un mecanismo $\Omega : \mathcal{P} \to \mathcal{M}$ es:

- **Débilmente menos manipulable en preferencias** que el mecanismo φ si, para cada perfil de preferencias bajo el cual Ω es manipulable, φ también lo es.
- **Menos manipulable en preferencias** que el mecanismo φ si, es débilmente menos manipulable en preferencias que φ y existe $\succ \in \mathcal{P}$ tal que Ω no es manipulado bajo $\succ$ y φ si lo es.

[3]Note que nuestra definición de estabilidad no requiere incorporar la individualidad racional. Esto se debe a que esta propiedad es satisfecha trivialmente cuando ningún agente posee opciones de salida.

- **Débilmente menos manipulable en cantidad de agentes** que el mecanismo φ si, para cada perfil de preferencias $\succ \in \mathcal{P}$ se cumple $|M^{\Omega}[\succ]| \leq |M^{\varphi}[\succ]|$.
- **Menos manipulable en cantidad de agentes** que el mecanismo φ si, es débilmente menos manipulable en cantidad de agentes y existe un perfil $\succ \in \mathcal{P}$ tal que $|M^{\Omega}[\succ]| < |M^{\varphi}[\succ]|$.

Es fácil ver que débilmente menos manipulable en cantidad de agentes implica débilmente menos manipulable en preferencias.[4] Sin embargo, note que esto no es necesariamente cierto para la versión estricta de estos criterios.

Por último, describamos uno de los mecanismos más utilizados en la literatura, el cual usaremos a lo largo de esta tesis. Sea $H \in \{E, G\}$ un lado del mercado.

Mecanismo de Aceptación Diferida (AD_H) (Gale & Shapley (1962)):

- Paso 1: Cada agente del conjunto H realiza una propuesta a su individuo preferido del otro lado del mercado. Cada agente que obtuvo al menos una propuesta, escoge su oferta favorita entre las que recibió y la acepta momentáneamente.
- Paso $k > 1$: Cada agente del conjunto H que fue rechazado en la etapa $k-1$, realiza una propuesta a su individuo preferido dentro de los que no lo han rechazado (si lo hay). Cada agente escoge su oferta favorita entre las que recibió y la que mantenía (de existir), y la acepta momentáneamente.
- El algoritmo termina cuando no se realizan más propuestas.

3. Resultados

En este apartado mostraremos el efecto de la escasez sobre la comparación mecanismos estables según qué tan manipulables son. Como veremos, la cantidad de estudiantes que buscan un guía de tesis relativo a la cantidad de profesores disponibles, es un factor clave para saber cuales mecanismos estables son los menos manipulables. Esta sección parte por incorporar algunas herramientas que nos serán útiles en las demostraciones de nuestros principales resultados, los cuales serán presentados en dos subsecciones, dependiendo de si se aplican a mercados desbalanceados o balanceados.

Sabemos que, cuando todos los agentes del mercado poseen opciones de salida, si un mecanismo estable (cualquiera) es manipulable bajo un perfil de preferencias, entonces todo mecanismo estable lo será (Chen, Egesdal, Pycia & Yenmez (2016); Bonkoungou & Nesterov (2023)). De esta forma, en cada perfil existen solo dos opciones: todos los mecanismos estables son manipulables o ninguno lo es.

[4] Supongamos que el mecanismo φ es débilmente menos manipulable que el mecanismo ψ y no los es en preferencias. Luego, existe un perfil $\succ \in \mathcal{P}$ bajo el cual φ es manipulado y ψ no lo es. Sin embargo, esto implica que $|M^{\Omega}[\succ]| > |M^{\varphi}[\succ]|$, una contradicción.

Note que los únicos perfiles de preferencias bajo los cuales los mecanismos estables no son manipulados (cuando hay opciones de salida), son aquellos donde el conjunto de matchings estables es un singleton.[5] Esto es una consecuencia directa de que si un agente puede declarar a otros como inaceptables, entonces cada vez que un mecanismo no le asigne su pareja estable óptima, él podrá alcanzarla mediante una estrategia de manipulación. En efecto, Roth (1982, Lema 1) muestra que cuando se aplica un mecanismo estable a un mercado bilateral, es suficiente que un agente reporte como único agente aceptable a su mejor pareja estable, para que esta le sea asignada. ¿Qué ocurre cuando los individuos no tienen acceso a este tipo de estrategias?

El siguiente resultado, extiende Roth (1982, Lema 1) a contextos en los que ningún agente posee opciones de salida y el mercado es desbalanceado. Como veremos en la demostración del Teorema 1, este resultado no es valido para el lado del mercado que no es escaso y por lo tanto no se cumple en mercados balanceados, tal como se hace notar en Teo, Sethuraman y Tan (2001).

Lema 1. (Sirguiado & Torres-Martinez (2024), Proposición 1) *Sea φ un mecanismo estable, $[E, G, \succ]$ un mercado desbalanceado y $H \in \{E, G\}$ el lado corto del mercado. Si para algún $h \in H$ se cumple $\varphi[\succ](h) \neq \mathrm{AD}_H[\succ](h)$, entonces el agente h manipula φ bajo $\succ$.*

Demostración. Sin perdida de generalidad, supongamos que hay más guías que estudiantes. Como el matching $\mathrm{AD}_E[\succ]$ es E-óptimo, existe $e \in E$ tal que $\mathrm{AD}_E[\succ](e) \succ_e \varphi[\succ](e)$. Además, como $|E| < |G|$, existe un guía $g^* \in G$ tal que $\mathrm{AD}_E[\succ](g^*) = \emptyset$.

Definamos la preferencia $\succ'_e$, donde los guías $\mathrm{AD}_E[\succ](e)$ y g^* son sus mejores opciones, en este orden. Como el matching $\mathrm{AD}_E[\succ]$ es estable bajo el perfil $(\succ'_e, \succ_{-e})$, sigue del Teorema del Hospital Rural (Gale & Sotomayor (1985, Teorema 1)) que g^* no es emparejado bajo ningún matching estable del problema de asignación $[E, G, (\succ'_e, \succ_{-e})]$. Esto implica que $\varphi[\succ'_e, \succ_{-e}] \succ_e g^*$ o de lo contrario, como e es aceptable para el guía g^*, el par (e, g^*) bloquearían el matching $\varphi[\succ'_e, \succ_{-e}]$. Esto implica que $\mathrm{AD}_E[\succ](e) = \varphi[\succ'_e, \succ_{-e}](e) \succ_e \varphi[\succ](e)$. Concluimos que el mecanismo φ es manipulado por el estudiante e. □

En la demostración anterior, se puede observar que, en contextos en los que nadie tiene opciones de salida, existe una nueva estrategia de manipulación para los agentes que pertenecen al grupo corto del mercado, la cual siempre les permite alcanzar su mejor pareja estable: fijarla como primera opción, seguido de un agente que no es emparejado. El Teorema del Hospital Rural, sumado a que nadie posee opciones de salida y que el mercado es desbalanceado, nos ayudan a asegurar que esta estrategia de manipulación es exitosa. El siguiente Remark es directo de la estrategia de demostración del Lema 1.[6]

[5]El conjunto de los estables es un singleton bajo el perfil $\succ$ si y solo si $\mathrm{AD}_E[\succ] = \mathrm{AD}_G[\succ]$.

[6]El Remark 1 es clave en nuestra demostración del Teorema 4 (ver Sección 4).

Remark 1. *Sean φ, ψ dos mecanismos estables y $\succ \in \mathcal{P}$ un perfil de preferencias cualquiera. Luego, si $\varphi[\succ](h) \succ_h \psi[\succ](h)$, entonces existe $\succ'_h$ tal que $\psi[\succ'_h, \succ_{-h}](h) = \varphi[\succ](h)$.*

La versión clásica del Lema 1 es uno de los ingredientes claves en muchas de las pruebas que se relacionan a incentivos en la literatura de emparejamientos. Ejemplo de ello son las clásicas demostraciones de resultados como la Imposibilidad de Roth (1982) o la unicidad de Alcalde y Barberà (1994). Más importante para nosotros, es que esta propiedad es utilizada recurrentemente en demostraciones relacionadas a la comparación de mecanismos estables que previamente se han hecho en la literatura (Pathak & Sönmez (2013); Chen, Egesdal, Pycia & Yenmez (2016); Bonkoungou & Nesterov (2021, 2023)). Deseamos responder qué ocurre al comparar mecanismos estables cuando Roth (1982, Lema 1) no es valido para un conjunto de individuos (aquellos agentes que no pertenecen al lado del mercado con menor cantidad de individuos).

Para llevar a cabo nuestro objetivo, necesitamos un resultado adicional que será de utilidad en muchas de nuestras pruebas. Este nos dice que, si el mecanismo AD_H, con $H \in \{E, G\}$, es manipulado bajo un perfil de preferencias, entonces todo mecanismo estable que implemente el emparejamiento H-óptimo en aquel perfil, también será manipulado.

Lema 2. *Sea φ un mecanismo estable tal que para algún $\succ \in \mathcal{P}$ se cumple $\varphi[\succ] = \mathrm{AD}_H[\succ]$. Luego, si AD_H es manipulado por $h \in E \cup G$ bajo $\succ$, entonces φ también lo es.*

Demostración. Realizaremos esta demostración para el mecanismo AD_E. Argumentos análogos pueden ser utilizados para mostrar qué AD_G cumple la propiedad.

Supongamos que AD_E es manipulado bajo $\succ \in \mathcal{P}$. Como este mecanismo es strategy-proof para los estudiantes (Dubins y Freedman (1981, Teorema 9)), sabemos que quien manipula AD_E bajo $\succ$ es un guía $g \in G$.

Sea φ un mecanismo estable tal que $\varphi[\succ] = \mathrm{AD}_E[\succ]$. Definamos la preferencia $\succ'_g$ como la mejor manipulación que puede realizar el guía g cuando se aplica AD_E al perfil $\succ$. De esta forma, $\succ'_g$ cumple $\mathrm{AD}_E[\succ'_g, \succ_{-g}](g) \succeq_g \mathrm{AD}_E[\succ''_g, \succ_{-g}](g)$ para toda preferencia $\succ''_g$. Adicionalmente, considere el conjunto $\widehat{E} \subseteq E$, el cual contiene a todos aquellos estudiantes que no son alcanzables por g cuando se aplica AD_E al perfil $\succ$, además del alumno que obtiene cuando reporta $\succ'_g$. Esto es,

$$\widehat{E} = \{e \in E : e \succeq_g \mathrm{AD}_E[\succ'_g, \succ_{-g}](g)\}.$$

Definamos $\succ^*_g$ como un ranking sobre E tal que:

- Los agentes en $\widehat{E}$ son las mejores alternativas y son ordenados de acuerdo a $\succ_g$.
- Los agentes en $E \setminus \widehat{E}$ son ordenados de acuerdo a $\succ'$.

Afirmamos que $\mathrm{AD}_E[\succ'_g, \succ_{-g}](g) = \mathrm{AD}_E[\succ^*_g, \succ_{-g}](g)$. Efectivamente, como $\succ'_g$ es la mejor manipulación del guía g, ningún estudiante $e \in \widehat{E} \setminus \{\mathrm{AD}_E[\succ'_g, \succ_{-g}](g)\}$ le hace una propuesta a g, independiente de la preferencia que reporte tal guía. De no ser así, g podría aceptar la propuesta, contradiciendo que $\succ'_g$ es la mejor manipulación que puede realizar. Luego, las únicas relaciones de preferencias relevantes al aplicar AD_E son las definidas sobre $(E \setminus \widehat{E}) \cup \{\text{AD}_E[\succ'_g, \succ_{-g}](g)\}$, alternativas donde $\succ^*_g$ coincide con $\succ'_g$. Así, las propuestas que acepta y rechaza g cuando se aplica AD_E son las mismas bajo los perfiles $(\succ^*_g, \succ_{-g})$ y $(\succ'_g, \succ_{-g})$, al igual que para el resto de los agentes. Por lo tanto, $\text{AD}_E[\succ'_g, \succ_{-g}](g) = \text{AD}_E[\succ^*_g, \succ_{-g}](g)$.[7]

Por Knuth (1976) sabemos que el estudiante $\text{AD}_E[\succ^*_g, \succ_{-g}](g)$ es la peor pareja estable bajo el perfil $(\succ^*_g, \succ_{-g})$. Luego, como el mecanismo φ es estable, $\varphi[\succ^*_g, \succ_{-g}](g) \succeq^*_g \text{AD}_E[\succ^*_g, \succ_{-g}](g)$. Sin embargo, como $\{\varphi[\succ^*_g, \succ_{-g}](g), \text{AD}_E[\succ^*_g, \succ_{-g}](g)\} \subseteq \widehat{E}$, la preferencia $\succ^*_g$ coincide con $\succ_g$ en aquel subconjunto de estudiantes, por lo cual

$$\varphi[\succ^*_g, \succ_{-g}](g) \succeq_g \text{AD}_E[\succ^*_g, \succ_{-g}](g) \succ_g \text{AD}_E[\succ](g) = \varphi[\succ](g)$$

Concluimos que φ es manipulado por el guía g bajo el perfil $\succ$. $\square$

El siguiente Remark es directo de la demostración del Lema 2.

Remark 2. *Si $h \in E \cup G$ manipula AD_H bajo $\succ \in \mathcal{P}$, mediante $\succ'_h$ y con $H \in \{E, G\}$, entonces todo mecanismo estable φ, tal que $\text{AD}_H[\succ'_h, \succ_{-h}](h) \succ_h \varphi[\succ](h) \succeq_h \text{AD}_H[\succ](h)$, también será manipulado por h.*

De esta forma, si un guía es capaz de manipular el mecanismo de aceptación diferida con los estudiantes haciendo las propuestas, entonces él manipulará cualquier mecanismo estable que no le asigne algo tan bueno como la pareja que alcanza al manipular AD_E. La intuición detrás de este resultado es que cuando un guía puede manipular el mecanismo que implementa su peor pareja estable, significa que el logra *cortar* la parte *baja* del lattice que conforma el conjunto de matchings estables correspondiente a un determinado perfil de preferencias.

Note que en contextos en qué hay opciones de salida, los agentes siempre pueden realizar este corte al conjunto estables, producto de Roth (1982, Lema 1). Por esta razón, es que el Lema 2 se cumple trivialmente en contextos clásicos. En contraste, cuando ningún agente posee opciones de salida, puede existir un perfil $\succ \in \mathcal{P}$ y un par de mecanismos estables, tal que estos asignan el mismo matching estable bajo $\succ$ y solo uno es manipulable.

Para un ejemplo de lo anterior, considere el siguiente perfil de preferencias:

$\succ'_{e_1}$	$\succ'_{e_2}$	$\succ'_{e_3}$	$\succ'_{g_1}$	$\succ'_{g_2}$	$\succ'_{g_3}$
g_1	g_2	g_3	e_2	e_1	e_3
g_2	g_1	$\vdots$	e_1	e_2	$\vdots$
g_3	g_3	$\vdots$	$\vdots$	$\vdots$	$\vdots$

Definamos el siguiente mecanismo:

[7]Un argumento similar se utiliza en la prueba del Corolario 5.1 de Roth (1982).

$$\Phi[\succ] = \begin{cases} \mathrm{AD}_G[\succ], & \text{cuando } \succ = \succ', \\ \mathrm{AD}_E[\succ], & \text{cuando } \succ \neq \succ', \end{cases}$$

Es claro que Φ es estable y que $\Phi[\succ'] = \mathrm{AD}_G[\succ']$. Afirmamos que AD_G no es manipulado bajo el perfil $\succ'$, mientras que Φ si lo es. Efectivamente, como AD_G es strategy-proof para los guías, sumado a que los estudiantes solo reciben una propuesta, ningún agente manipula este mecanismo. Por otro lado, el agente e_1 puede reportar la preferencia $e_1 \succ^*_{e_1} e_3 \succ^*_{e_1} e_2$ y así obtener $\Phi[\succ^*_{e_1}, \succ'_{-e_1}](e_1) = g_1 \succ'_{e_1} g_2 = \Phi[\succ'](e_1)$. Este sencillo ejemplo muestra que es posible que dos mecanismos asignen el mismo emparejamiento estable, pero solo uno es manipulado, una de las principales diferencias respecto a modelos de emparejamiento bilateral con opciones de salida.[8]

3.1 Mercados desbalanceados

Habiendo presentado los Lemas 1 y 2, ya somos capaces de presentar nuestros principales resultados para mercados desbalanceados. Sin perdida de generalidad, a lo largo de esta sección asumiremos que hay menos estudiantes que profesores guías, es decir, $|E| < |G|$. El Teorema 1 señala que, cuando ningún agente posee opciones de salida, el mecanismo de aceptación diferida con el lado corto del mercado haciendo las propuestas, es débilmente menos manipulable en preferencias que cualquier otro mecanismo estable. Además, esto se cumple de manera estricta cuando lo comparamos con AD_G .

Teorema 1. *En un mercado bilateral uno-a-uno desbalanceado, suponga que ningún agente posee opciones de salida. Luego,*

(i) el mecanismo AD_E *es débilmente menos manipulable en preferencias que cualquier otro mecanismo estable.*

(ii) el mecanismo AD_E *es menos manipulable en preferencias que* AD_G*.*

Demostración. (i) Por contradicción, supongamos no es así, es decir, existe $\varphi : \mathcal{P} \to \mathcal{M}$ estable y $\succ \in \mathcal{P}$ tal que AD_E es manipulable bajo $\succ$ y φ no lo es. Como el mecanismo AD_E es strategy-proof para los estudiantes (Dubins y Freedman (1981, Teorema 9)), sabemos que es un guía $g \in G$ quien manipula tal mecanismo bajo $\succ$. Existen dos casos posibles:

- $\varphi[\succ] = \mathrm{AD}_E[\succ]$. Como φ es estable y el mecanismo AD_E es manipulado bajo $\succ$, el Lema 2 nos asegura que φ también será manipulado por g bajo $\succ$. Una contradicción.
- $\varphi[\succ] \neq \mathrm{AD}_E[\succ]$. Como φ es estable, el Lema 1 implica que existe un estudiante $e \in E$ que manipula el mecanismo φ bajo $\succ$. Una contradicción.

Concluimos que AD_E es débilmente menos manipulable en preferencias que cualquier mecanismo estable.

[8]Note que $\Phi[\succ](e_1) \succeq_{e_1} \mathrm{AD}_G[\succ](e_1)$ no implica que e_1 puede alcanzar $\Phi[\succ^*_{e_1}, \succ'_{-e_1}](e_1)$ cuando se aplica el mecanismo AD_G. Este argumento es clave en nuestra demostración del Teorema 4 (ver Sección 4.)

(ii) Probaremos que AD_E es menos manipulable en preferencias que el mecanismo AD_G. Por el item (i), sabemos que el primero es débilmente menos manipulable en preferencias que cualquier mecanismo estable. Luego, para completar nuestra demostración basta encontrar un perfil de preferencias bajo el cual AD_E no es manipulado y el mecanismo AD_G si lo es.

Sean $E = \{e_1, e_2, \ldots, e_s\}$ y $G = \{g_1, g_2, \ldots, g_r\}$ los conjuntos de estudiantes y guías respectivamente, con $s < r$ por suposición. Considere el perfil $\succ \in \mathcal{P}$ que cumple las siguientes condiciones:

- Dado $i \in \{1, 2\}$, g_i considera a e_i como su mejor pareja posible.
- Dados $i, j \in \{1, 2\}$ con $i \neq j$, e_i considera a g_j como su mejor pareja posible.
- Dado $i \in \{3, \ldots, s\}$, los agentes g_i y e_i se consideran mutuamente como su mejor pareja posible.

Bajo este perfil de preferencias existen solo dos matchings estables:

$$\begin{aligned} \mathrm{AD}_E[\succ] &= \{(e_1, g_2), (e_2, g_1), (e_3, g_3), \ldots, (e_s, g_s), g_{s+1}, \ldots g_r\}, \\ \mathrm{AD}_G[\succ] &= \{(e_1, g_1), (e_2, g_2), (e_3, g_3), \ldots, (e_s, g_s), g_{s+1}, \ldots g_r\}. \end{aligned}$$

Sabemos que ningún estudiante manipula AD_E. Además, como los guías $g_3, g_4, \ldots g_s$ son emparejados con su mejor pareja posible, ninguno de ellos tiene incentivos a manipular AD_E bajo este perfil. Además, los guías $g_{s+1}, \ldots, g_r$ no reciben ninguna propuesta, por lo que no pueden manipular. Luego, como los guías no poseen opciones de salida, y además g_1, g_2 solo reciben una propuesta, estos no pueden hacer más que aceptarlas. Concluimos que AD_E no es manipulado bajo $\succ$.

Por otro lado, como AD_G es estable y $\mathrm{AD}_G[\succ](e_i) \neq \mathrm{AD}_E[\succ](e_i)$, con $e_i \in \{e_1, e_2\}$, el Lema 1 nos asegura que los estudiantes e_1 y e_2 manipulan el mecanismo AD_G bajo el perfil $\succ$. Concluimos que AD_E no es manipulado bajo $\succ$, mientras que AD_G si lo es. □

Este resultado muestra qué, en contextos en que ningún participante del mercado tiene opciones de salida y existe un desbalance entre los grupos, comparar mecanismos estables utilizando el criterio de Pathak & Sönmez (2013), es informativo. Esto es, existen mecanismos estables que son menos manipulables en preferencias que otros. Adicionalmente, no existe un perfil bajo el cual el mecanismo de aceptación diferida, con el lado corto haciendo las propuestas, sea manipulado y algún mecanismo estable no lo sea. Esto no ocurre cuando las propuestas son hechas por el lado largo.

A pesar de que este resultado es relevante por sí mismo, un planificador puede estar preocupado no solo en si, bajo cierto perfil de preferencias, un mecanismo es manipulable y otro no lo es. Más allá de esto, él puede darle una mayor importancia a la cantidad de agentes que actúan de manera estratégica. ¿Por qué? Considere dos mecanismos estables φ y ψ y asuma que el primero es menos manipulable en preferencias que el segundo. Esta información lo único que nos dice es que existe un perfil bajo el cual φ no es manipulado y ψ sí lo es, sumado a que no existe un perfil donde ocurra lo contrario. Sin embargo, ¿qué ocurre si ψ es manipulado por un solo agente en el resto de perfiles de preferencia, mientras que φ lo es por todos los agentes del mercado? ¿Hace sentido implementar el mecanismo φ con el objetivo de reducir el comportamiento estratégico?

Nuestro siguiente objetivo es mostrar qué ocurre cuando comparamos mecanismos estables, considerando la cantidad de agentes que los manipulan. Antes de ir con este resultado, considere la siguiente idea.

Sea $[E, G, \succ]$ un mercado desbalanceado. Por Dubins & Freedman (1981, Teorema 9), sabemos que los mecanismos $\mathrm{AD}_E, \mathrm{AD}_G$ son strategy-proof para E y G, respectivamente. Luego, uno podría pensar que implementar el mecanismo estable óptimo para el lado del mercado con mayor cantidad de agentes, minimiza el número de individuos manipuladores. Esta intuición no es correcta y el siguiente resultado muestra el por qué.

Lema 3. *En un mercado bilateral uno-a-uno, considere un mecanismo estable φ y un agente $h \in E \cup G$, tal que $\varphi[\succ](h) = \emptyset$ para algún perfil $\succ \in \mathcal{P}$. Luego, h no manipula el mecanismo φ bajo $\succ$.*

Demostración. Por contradicción, supongamos que no es así. Como φ es un mecanismo estable, el Teorema del Hospital Rural (Gale & Sotomayor (1985, Teorema 1)) nos asegura que $\mathrm{AD}_H[\succ](h) = \emptyset$, con $H \in \{E, G\}$ y $h \in H$. Luego, como el mecanismo φ es manipulado por h bajo $\succ$, sabemos que existe una preferencia $\succ'_h$ tal que $\varphi[\succ'_h, \succ_{-h}](h) \succ_h \varphi[\succ](h) = \mathrm{AD}_H[\succ](h) = \emptyset$. Sin embargo, el Teorema de los Limites de la Manipulación (Demange, Gale & Sotomayor, 1987) nos asegura que no existe matching estable bajo el perfil $(\succ'_h, \succ_{-h})$ que sea más preferido por h que el emparejamiento H-óptimo bajo $\succ$. Una contradicción. $\square$

Este resultado nos entrega una cota superior sobre la cantidad de agentes que pueden manipular un mecanismo estable bajo un perfil de preferencias cualquiera. En efecto, el Lema 3 implica que a mayor cantidad de parejas estables se formen en un mercado, mayor será la cantidad de agentes que pueden llegar a manipular un mecanismo. El caso extremo de lo anterior se produce cuando todos los agentes del lado corto del mercado (si lo hay) son emparejados. Esta intuición conlleva el siguiente Remark.

Remark 3. *En un mercado bilateral uno-a-uno, sea φ un mecanismo estable cualquiera. Luego, para todo $\succ \in \mathcal{P}$ se cumple $|M^{\varphi}[\succ]| \leq 2 \cdot min\{|E|, |G|\}$.*

En un mercado bilateral, donde ambos lados del mercado son estratégicos y todos los agentes poseen opciones de salida, para cada lado del mercado solo existe un único mecanismo estable y strategy-proof, el cual es una versión del mecanismo de aceptación diferida (Alcalde & Barberà, 1994). Sin embargo, esto deja de ser cierto cuando ningún participante del mercado posee opciones de salida (Sirguiado & Torres-Martinez, 2024). En efecto, en ese contexto existe un único mecanismo estable y strategy-proof solo para el lado corto del mercado.

El siguiente resultado muestra que el mecanismo de aceptación diferida, con el lado corto haciendo las propuestas, minimiza la cantidad de agentes manipuladores en todo perfil de preferencias, es decir, es débilmente menos manipulable en cantidad de agentes que cualquier otro mecanismo estable. Adicionalmente, mostramos que esta propiedad no se cumple para ningún mecanismo estable y strategy-proof para el lado largo del mercado.

Teorema 2. *En un mercado bilateral uno-a-uno desbalanceado, suponga que ningún agente posee opciones de salida y sea E el lado corto del mercado. Luego,*

(i) el mecanismo AD_E es débilmente menos manipulable en cantidad de agentes que cualquier otro mecanismo estable.

(ii) el mecanismo AD_E es menos manipulable en cantidad de agentes que cualquier mecanismo estable y strategy-proof para el lado largo del mercado.

Demostración. (i) Por contradicción, supongamos que AD_E no es débilmente menos manipulable en cantidad de agentes que cualquier otro mecanismo estable. Esto es, existe $\varphi : \mathcal{P} \to \mathcal{M}$ y un perfil $\succ \in \mathcal{P}$ tal que $|M^{\varphi}[\succ]| < |M^{\mathrm{AD}_E}[\succ]|$.

Supongamos que $\mathrm{AD}_E[\succ] = \varphi[\succ]$. Sabemos que AD_E es strategy-proof para los estudiantes (Dubins y Freedman (1981, Teorema 9)), por lo que $M^{\mathrm{AD}_E}[\succ] \subseteq G$. Luego, como φ es estable, el Lema 2 nos asegura que todo profesor guía que manipula AD_E bajo $\succ$ también manipulará φ. Esto es, $M^{\mathrm{AD}_E}[\succ] \subseteq M^{\varphi}[\succ]$. Una contradicción con $|M^{\varphi}[\succ]| < |M^{\mathrm{AD}_E}[\succ]|$.

Ahora, supongamos que $\mathrm{AD}_E[\succ] \neq \varphi[\succ]$. Sean $E^{\varphi}[\succ] \subseteq E$ y $G^{\varphi}[\succ] \subseteq G$ subconjuntos tales que $E^{\varphi}[\succ] \cup G^{\varphi}[\succ] = M^{\varphi}[\succ]$. Análogamente, se definen los subconjuntos $E^{\mathrm{AD}_E}[\succ] \subseteq E$ y $G^{\mathrm{AD}_E}[\succ] \subseteq G$, respecto a $M^{\mathrm{AD}_E}[\succ]$. A partir de estos conjuntos formamos nuevos subconjuntos. Consideremos un conjunto $K \in \{E^{\varphi}[\succ], G^{\varphi}[\succ], E^{\mathrm{AD}_E}[\succ], G^{\mathrm{AD}_E}[\succ]\}$ y definamos los subconjuntos:

$$\widehat{K} = \{k \in K : \mathrm{AD}_E[\succ](k) = \varphi[\succ](k)\}, \quad \overline{K} = \{k \in K : \mathrm{AD}_E[\succ](k) \neq \varphi[\succ](k)\}.$$

Nuestro objetivo es llegar una contradicción al probar que $|M^{\mathrm{AD}_E}[\succ]| \leq |M^{\varphi}[\succ]|$, lo cual es equivalente a mostrar

$$|\widehat{E}^{\mathrm{AD}_E}[\succ]| + |\overline{E}^{\mathrm{AD}_E}[\succ]| + |\widehat{G}^{\mathrm{AD}_E}[\succ]| + |\overline{G}^{\mathrm{AD}_E}[\succ]| \leq |\widehat{E}^{\varphi}[\succ]| + |\overline{E}^{\varphi}[\succ]| + |\widehat{G}^{\varphi}[\succ]| + |\overline{G}^{\varphi}[\succ]|.$$

Como ningún estudiante manipula el mecanismo AD_E, sabemos que $\widehat{E}^{\mathrm{AD}_E}[\succ] = \overline{E}^{\mathrm{AD}_E}[\succ] = \emptyset$. Además, como φ es estable, el Lema 2 nos asegura que todo $g \in G$ que manipula AD_E, también manipula φ cuando $\mathrm{AD}_E[\succ](g) = \varphi[\succ](g)$. Esto es, $\widehat{G}^{\mathrm{AD}_E}[\succ] = \widehat{G}^{\varphi}[\succ]$.

Afirmamos que $\widehat{E}^{\varphi}[\succ] = \emptyset$. Efectivamente, de no ser así existiría un estudiante que manipulando el mecanismo φ alcanza un profesor más preferido que su mejor emparejamiento estable. Sin embargo, esto junto al Teorema de los Limites de la Manipulación (Demange, Gale & Sotomayor, 1987) contradicen la estabilidad de φ.

En conjunto, estos argumentos implican es suficiente mostrar que $|\overline{G}^{\mathrm{AD}_E}[\succ]| \leq |\overline{E}^{\varphi}[\succ]| + |\overline{G}^{\varphi}[\succ]|$ para concluir la demostración. Probaremos que $|\overline{G}^{\mathrm{AD}_E}[\succ]| \leq |\overline{E}^{\varphi}[\succ]|$. Note que la cantidad de estudiantes y guías, que no son emparejados con el mismo agente en los matchings $\mathrm{AD}_E[\succ]$ y

$\varphi[\succ]$, es la misma. Denotando por $E^* \subseteq E$ y $G^* \subseteq G$ a los conjuntos de aquellos estudiantes y guías, tenemos que $|E^*| = |G^*|$. Por el Lema 1, sabemos que todos los estudiantes que no son emparejados con su mejor guía estable, cuando se aplica φ al perfil $\succ$, manipulan tal mecanismo. Esto es, $E^* = \overline{E}^{\varphi}[\succ]$.

Luego, $|\overline{G}^{\mathrm{AD}_E}[\succ]| \leq |G^*| = |E^*| = |\overline{E}^{\varphi}[\succ]|$. Con lo cual, concluimos que $|\overline{G}^{\mathrm{AD}_E}[\succ]| \leq |\overline{E}^{\varphi}[\succ]|+|\overline{G}^{\varphi}[\succ]|$. Una contradicción con que φ es menos manipulable que AD_E bajo el perfil $\succ$. Concluimos que AD_E es débilmente menos manipulable en cantidad de agentes que cualquier mecanismo estable.

(ii) Sea φ un mecanismo estable y strategy-proof para los guías. Deseamos probar que el mecanismo de aceptación diferida con los estudiantes haciendo las propuestas es menos manipulable en cantidad de agentes que φ. Como el mecanismo AD_E es débilmente menos manipulable en cantidad de agentes que cualquier otro mecanismo estable, basta probar que existe un perfil bajo el cual AD_E es manipulado por menos agentes que el mecanismo φ.

Sean $E = \{e_1, e_2, \ldots, e_s\}$ y $G = \{g_1, g_2, \ldots, g_r\}$ los conjuntos de estudiantes y guías respectivamente, con $r < s$ por suposición. Considere el perfil $\succ \in \mathcal{P}$ dado por:

$\succ_{e_1}$	$\succ_{e_2}$	$\succ_{e_3}$	$\succ_{e_4}$	$\succ_{e_5}$	$\cdots$	$\succ_{e_s}$	$\succ_{g_1}$	$\succ_{g_2}$	$\succ_{g_3}$	$\succ_{g_4}$	$\succ_{g_5}$	$\cdots$	$\succ_{g_s}$	$\cdots$	$\succ_{g_r}$
g_1	g_2	g_1	g_4	g_5	$\cdots$	g_s	e_2	e_1	e_3	e_4	e_5	$\cdots$	e_s	$\cdots$	$\vdots$
g_2	g_1	g_3	$\vdots$	$\vdots$		$\vdots$	e_1	e_2	e_2	$\vdots$	$\vdots$		$\vdots$		$\vdots$
g_3	g_3	$\vdots$	$\vdots$	$\vdots$		$\vdots$	e_3	e_3	$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$

Bajo este perfil de preferencias existen solo dos emparejamientos estables:

$$\begin{aligned}
\mathrm{AD}_G[\succ] &= \{(e_1, g_2), (e_2, g_1), (e_3, g_3), \ldots, (e_s, g_s), g_{s+1}, \ldots, g_r\}, \\
\mathrm{AD}_E[\succ] &= \{(e_1, g_1), (e_2, g_2), (e_3, g_3), \ldots, (e_s, g_s), g_{s+1}, \ldots, g_r\}.
\end{aligned}$$

Supongamos que $\varphi[\succ] \neq \mathrm{AD}_G[\succ]$. Afirmamos que el guía g_1 puede reportar la preferencia $e_2 \succ'_{g_1} e_3 \succ'_{g_1} e_1 \ldots$ para manipular φ. Efectivamente, bajo el perfil de preferencias $(\succ'_{g_1}, \succ_{-g_1})$ se cumple $\mathrm{AD}_E[\succ'_{g_1}, \succ_{-g_1}] = \mathrm{AD}_G[\succ'_{g_1}, \succ_{-g_1}]$, por lo cual solo existe un matching estable bajo este perfil. Más aún, como φ es un mecanismo estable, se cumple que $\varphi[\succ'_{g_1}, \succ_{-g_1}] = \mathrm{AD}_G[\succ'_{g_1}, \succ_{-g_1}] = \mathrm{AD}_G[\succ]$. Sin embargo, esto implica que $e_2 = \varphi[\succ'_{g_1}, \succ_{-g_1}](g_1) \succ_{g_1} \varphi[\succ](g_1) = e_1$, una contradicción con que φ es strategy-proof para los guías.

Teniendo en cuenta que $\varphi[\succ] = \mathrm{AD}_G[\succ]$, mostraremos que bajo el perfil $\succ$ la cantidad de agentes que manipulan AD_E es menor que la cantidad de individuos que manipulan mecanismo φ. Primero, note que para $e_i \in \{e_1, e_2\}$ se cumple $\varphi[\succ](e_i) \neq \mathrm{AD}_E[\succ](e_i)$. Por el Lema 1, sabemos que todo estudiante que no sea emparejado con su mejor pareja estable, manipulará cualquier mecanismo estable. Luego, como el mecanismo φ es strategy-proof para los profesores, son 2 los agentes que manipulan φ (solo los estudiantes lo pueden manipular).

Por otro lado, note que al aplicar AD_E, el único guía que recibe más de una propuesta es g_1. Luego, como los guías no poseen opciones de salida, el único que podría manipular AD_E bajo $\succ$ es

g_1 (el cual efectivamente lo hace, como ya mostramos al suponer que $\varphi[\succ] \neq \mathrm{AD}_G[\succ]$). Luego, la cantidad de agentes que manipulan AD_E bajo el perfil de preferencias $\succ$ es solamente 1.

Concluimos que AD_E es menos manipulable en cantidad de agentes que φ. □

La prueba del item (i) del Teorema 2 sigue la misma estrategia de demostración que Bonkoungou & Nesterov (2023, Teorema 1), pero requiere de herramientas adicionales producto de estar en un contexto sin opciones de salida (Lemas 1 y 2). Aunque uno podría pensar que él item (ii) del Teorema 2 también vale bajo la comparación en preferencias, no es claro que esto sea cierto. Para entender el por qué, consideremos los aspectos relevantes de la demostración del item (ii). En ella, tomamos un perfil de preferencias tal que:

a) Existe más de un matching estable.
b) El guía g_1 manipula cualquier mecanismo estable que no le asigna su pareja estable óptima.
c) No existe un matching estable diferente del G-óptimo, tal que g_1 alcance a su pareja estable óptima.

Por los items b) y c) sabemos que todos los mecanismos estables y strategy-proof para los guías coinciden en aquel perfil de preferencias. El aprender cual matching implementan estos mecanismos, nos permite saber quienes los manipulan bajo el perfil de preferencias, para así compararlos con AD_E en cantidad de agentes. Adicionalmente, el item a) permite que este último mecanismo no coincida con los mecanismos estables y strategy-proof para los guías (de lo contrario, habría un único matching estable y por lo tanto, ningún agente manipularía).

¿Cuál es la diferencia respecto a la comparación en preferencias? En este caso buscamos un perfil bajo el cual AD_E no sea manipulable y que cualquier mecanismo estable y strategy-proof para los guías sí lo sea. Por esta razón, no podemos utilizar el item b), el cual es clave para saber qué matching se implementa bajo el mecanismo con el que estamos comparando a AD_E.

De forma más precisa, estamos buscando un perfil de preferencias bajo el cual:

- Ningún mecanismo estable y strategy-proof para los guías asigna el emparejamiento estable óptimo para los estudiantes bajo el perfil.
- Ningún guía puede manipular AD_E bajo el perfil.

La existencia de un perfil de preferencias que cumpla estas características es una pregunta que es dejada para investigación futura.

¿Existen otros mecanismos estables que minimicen la cantidad de agentes manipuladores? Note que de existir, estos no podrían ser one-side strategy-proof, porque para el lado corto del mercado el único mecanismo estable y strategy-proof es AD_E (Sirguiado & Torres-Martinez, 2024), mientras que para el lado largo no existen mecanismos estables y strategy-proof que minimicen la manipulación (ver Teorema 2 (ii)). El siguiente resultado responde esta pregunta.

Proposición 1. *En un mercado bilateral uno-a-uno desbalanceado, suponga que ningún agente posee opciones de salida y que E es el lado corto. Luego, existen mecanismos estables que son igual* *de manipulables, en preferencias y en cantidad de agentes, que* AD_E.

Demostración. Ver Apéndice A.

La intuición detrás de este resultado es que las dos versiones del mecanismos de aceptación diferida (con los guías y alumnos haciendo las propuestas), son manipulados por 2 agentes bajo el perfil $\succ^*$. Luego, al asignar el emparejamiento $\mathrm{AD}_G[\succ^*]$ al perfil de preferencias $\succ^*$, el mecanismo Ω posee le misma cantidad de agentes manipuladores que AD_E en ese perfil. Adicionalmente, como en el resto de perfiles de preferencias ambos mecanismos coinciden, ambos mecanismos son manipulados por la misma cantidad de agentes en el dominio $\mathcal{P}$.

Supongamos que un planificador desea implementar el mecanismo estable, que minimiza la cantidad de agentes manipuladores dentro del conjunto de mecanismos strategy-proof para el lado largo del mercado. ¿Existe un mecanismo que satisfaga este objetivo?

En Sirguiado & Torres-Martinez (2024), los autores muestran que, cuando ningún agente posee opciones de salida, existe un mecanismo estable y strategy-proof para el lado largo del mercado, que es el mejor para los agentes del lado escaso (dentro de la familia de mecanismos que cumplen estas dos propiedades). Denotando a tal mecanismo por Ω^E, el siguiente resultado muestra que este mecanismo es débilmente menos manipulable que cualquier otra regle de asignación estable y strategy-proof para el lado largo del mercado.[9]

Proposición 2. *En un mercado bilateral uno-a-uno desbalanceado, suponga que ningún agente posee opciones de salida y que G es el lado largo. Luego, Ω^E es débilmente menos manipulable en cantidad de agentes y en preferencias que cualquier otro mecanismo estable y strategy-proof para G.*

Demostración. Ver Apéndice A.

3.2 Mercados Balanceados

En esta sección, nos concentraremos en mercados balanceados, es decir, aquellos que cumplen $|E| = |G|$. Como veremos, la simetría del mercado (en cantidad de agentes y en preferencias) puede generar que los resultados presentados en la Sección 3.1 dejen de ser validos.

En efecto, el siguiente Teorema muestra que, al contrario de lo que ocurre en mercados desbalanceados, no existe un mecanismo estable y one-side strategy-proof que minimice la cantidad de agentes manipuladores en todo perfil de preferencias. Note que en este contexto, hay muchos mecanismos estables y strategy-proof para ambos lados del mercado (Sirguiado & Torres-Martinez, 2024), por lo cual probaremos que ninguno de estos logra minimizar la manipulación.

Adicionalmente, mostramos que las dos versiones del mecanismo de aceptación diferida no son comparables entre si bajo ninguna de nuestras métricas. Esto es, existen perfiles de preferencias

[9] Note que esto se cumple trivialmente en contextos en los que solo existe un único mecanismo estable y strategy-proof para el lado largo del mercado.

donde AD_E es manipulable y AD_G no lo es, y otros donde ocurre lo contrario.

Teorema 3. *En un mercado bilateral uno-a-uno balanceado, donde ningún agente posee opciones de salida:*

(i) No existe un mecanismo estable y one-side strategy-proof que sea débilmente menos manipulable en cantidad de agentes que cualquier otro mecanismo estable.

(ii) Los mecanismos AD_E *y* AD_G *no son comparables en preferencias ni en cantidad de agentes.*

Demostración. (i) Por contradicción, supongamos que existe φ estable, strategy-proof para los estudiantes y débilmente menos manipulable en cantidad de agentes que cualquier otro mecanismo estable. Considere el perfil $\succ \in \mathcal{P}$ dado por

$\succ_{e_1}$	$\succ_{e_2}$	$\succ_{e_3}$	$\succ_{e_4}$	$\succ_{e_5}$	$\cdots$	$\succ_{e_n}$	$\succ_{g_1}$	$\succ_{g_2}$	$\succ_{g_3}$	$\succ_{g_4}$	$\cdots$	$\succ_{g_n}$
g_2	g_1	g_1	g_2	g_5	$\cdots$	g_n	e_1	e_2	e_1	e_4	$\cdots$	e_n
g_1	g_2	g_3	g_4	$\vdots$		$\vdots$	e_2	e_1	e_3	$\vdots$		$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$

Bajo este perfil de preferencias existen solo dos matching estables:

$$\begin{aligned}
\text{AD}_E[\succ] &= \{(e_1, g_2), (e_2, g_1), (e_3, g_3), \ldots, (e_n, g_n)\},\\
\text{AD}_G[\succ] &= \{(e_1, g_1), (e_2, g_2), (e_3, g_3), \ldots, (e_n, g_n)\}.
\end{aligned}$$

Afirmamos que $\varphi[\succ] = \text{AD}_E[\succ]$. Efectivamente, de no ser así el estudiante e_1 puede reportar la preferencia $g_2 \succ'_{e_1} g_3 \ldots$, con lo cual se cumple $\text{AD}_E[\succ'_{e_1}, \succ_{-e_1}] = \text{AD}_G[\succ'_{e_1}, \succ_{-e_1}]$, por lo que solo existe un emparejamiento estable bajo el perfil $(\succ'_{e_1}, \succ_{-e_1})$. Note que este matching coincide con $\text{AD}_E[\succ]$. Luego, como φ es estable se tiene que $g_2 = \varphi[\succ'_{e_1}, \succ_{-e_1}](e_1) \succ_{e_1} \varphi[\succ](e_1) = g_1$, una contradicción con que φ es strategy-proof para los estudiantes.

Ahora, mostraremos que el mecanismo φ posee más agentes manipuladores que AD_G bajo el perfil $\succ$. Como $\varphi[\succ] \neq \text{AD}_G[\succ]$, sabemos que el estudiante e_1 manipula el mecanismo AD_G bajo $\succ$ (ver parrafo anterior). Además, como aquel estudiante es el único alumno que recibe más de una propuesta, cuando los guías las realizan, ningún otro estudiante puede manipular AD_G. Luego, producto de que este mecanismo es strategy-proof para los guías, solo un agente manipula AD_G bajo $\succ$.

Por otro lado, al aplicar el mecanismo φ al perfil $\succ$, los profesores guías g_1 y g_2 pueden manipularlo, reportando las preferencias $e_1 \succ'_{g_1} e_3 \ldots$ y $e_2 \succ'_{g_2} e_4 \ldots$, respectivamente. Para ver esto, note que bajo los perfiles $(\succ'_{g_1}, \succ_{-g_1})$ y $(\succ'_{g_2}, \succ_{-g_2})$, el único matching estable es $\text{AD}_E[\succ]$. Luego, como φ es un mecanismo estable, para cada caso se tiene que $e_1 = \varphi[\succ'_{g_1}, \succ_{-g_1}](g_1) \succ_{g_1} \varphi[\succ](g_1) = e_2$ y $e_2 = \varphi[\succ'_{g_2}, \succ_{-g_2}](g_2) \succ_{g_2} \varphi[\succ](g_2) = e_1$. Concluimos que φ es manipulado por más agentes que el mecanismo AD_G bajo el perfil $\succ$, una contradicción con que φ minimiza la manipulación en todo perfil de preferencias.

(ii) Probaremos que existe un perfil $\succ \in \mathcal{P}$ tal que AD_G es manipulado bajo tal perfil y el mecanismo AD_E no lo es. Luego, utilizando la simetría del problema de asignación y de los mecanismos AD_E y AD_G, concluiremos que ambos mecanismos no son comparables en términos de manipulabilidad.

Considere el perfil de preferencias $\succ \in \mathcal{P}$ dado por

$\succ_{e_1}$	$\succ_{e_2}$	$\succ_{e_3}$	$\succ_{e_4}$	$\succ_{e_5}$	$\cdots$	$\succ_{e_n}$	$\succ_{g_1}$	$\succ_{g_2}$	$\succ_{g_3}$	$\succ_{g_4}$	$\succ_{g_5}$	$\cdots$	$\succ_{g_n}$
g_2	g_1	g_3	g_4	g_5	$\cdots$	g_n	e_1	e_2	e_1	e_4	e_5	$\cdots$	e_n
g_1	g_2	g_2	$\vdots$	$\vdots$		$\vdots$	e_2	e_1	e_3	$\vdots$	$\vdots$		$\vdots$
g_3	g_3	$\vdots$	$\vdots$	$\vdots$		$\vdots$	e_3	e_3	$\vdots$	$\vdots$	$\vdots$		$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$

Al aplicar el mecanismo inducido por el algoritmo de aceptación diferida, con los estudiantes haciendo las propuestas, se obtiene el matching $\text{AD}_E[\succ] = \{(e_1, g_2), (e_2, g_1), (e_3, g_3), \ldots, (e_n, g_n)\}$. Como AD_E es strategy-proof para los estudiantes, los profesores no pueden declarar a otros como inaceptables y cada uno de ellos recibe una y solo una propuesta, AD_E no es manipulado bajo $\succ$.

Por otro lado, cuando los guías son quienes hacen las propuestas el emparejamiento obtenido es $\text{AD}_G[\succ] = \{(e_1, g_1), (e_2, g_2), (e_3, g_3), \ldots, (e_n, g_n)\}$. Sin embargo, el estudiante e_1 puede mejorar al reportar la preferencia $g_2 \succ'_{e_1} g_3 \ldots$, con lo cual obtiene $g_2 = \text{AD}_E[\succ'_{e_1}, \succ_{-e_1}](e_1) \succ_{e_1} \text{AD}_G[\succ](e_1)$. Concluimos que AD_G es manipulado bajo $\succ$.

De manera análoga, la simetría de los dominios de preferencias de E y G, sumado a la de los mecanismos AD_E y AD_G, nos asegura que existe un perfil de preferencias bajo el cual el primero de estos mecanismos es manipulado, mientras que el segundo no lo es. De esta forma, concluimos que los mecanismos AD_E y AD_G no son comparables en preferencias, lo cual implica que tampoco lo son en cantidad de agentes. $\square$

4. Comparación a nivel individuo

En esta sección seguimos lo hecho en Chen, Egesdal, Pycia & Yenmez (2016) para estudiar la relación entre comparar mecanismos estables en base a nivel de manipulación y hacerlo en preferencias. Para comprender mejor esta equivalencia, partamos por establecer algo de lenguaje.

Definamos al **set de mejoras** como el conjunto que viene dado por

$$K[h, \varphi, \succ] = \{\varphi[\succ'_h, \succ_{-h}](h) \in E \cup G : \varphi[\succ'_h, \succ_{-h}](h) \succ_h \varphi[\succ](h)\},$$

el cual contiene todas las parejas que mejoran al individuo h y que puede alcanzar mediante alguna estrategia de manipulación $\succ'_h$, cuando se aplica el mecanismo φ al perfil de preferencias $\succ \in \mathcal{P}$.

Los siguientes criterios permiten comparar a dos mecanismos estables en términos de qué tan manipulados son por un agente cualquiera. Al igual que en definiciones anteriores, existe una definición débil y una estricta.

Sean ψ, φ dos mecanismos estables. Diremos que φ es:

- **Débilmente menos manipulable** que ψ para el agente h si:

$$\forall \succ \in \mathcal{P}, \quad K[h, \varphi, \succ] \subseteq K[h, \psi, \succ].$$

- **Menos manipulable** que ψ para el agente h si:

$$\forall \succ \in \mathcal{P}, \quad K[h, \varphi, \succ] \subseteq K[h, \psi, \succ] \quad \wedge \quad \exists \succ \in \mathcal{P}, \quad K[h, \varphi, \succ] \subset K[h, \psi, \succ].$$

Utilizando estos criterios de comparación de mecanismos estables, Chen, Egesdal, Pycia & Yenmez (2016) muestran que la siguiente propiedad se cumple en mercados donde todos los agentes poseen opciones de salida.

Chen, Egesdal, Pycia & Yenmez (2016, Teorema 1). *En mercados bilaterales uno-a-uno, en los que todos poseen opciones de salida, comparar mecanismos según su manipulabilidad es equivalente a compararlos en preferencias para todo $h \in E \cup G$. Esto es,*

Para todo perfil $\succ$: $\varphi[\succ](h) \succeq_h \psi[\succ](h) \iff \varphi$ es débilmente menos manipulable que ψ para h.

Esta propiedad señala que, cuanto mejor sea un mecanismo para un agente, menos manipulado será por él. Aunque es posible definir esta equivalencia de una forma más estricta, los resultados bajo esta definición no son novedosos respecto a los que presentaremos en esta tesis y las demostraciones se pueden obtener fácilmente a partir de la que mostramos a continuación.[10]

El siguiente resultado muestra que, en contextos en los que nadie posee opciones de salida, la equivalencia descrita anteriormente es valida solo en mercados desbalanceados y solo para aquellos agentes que pertenecen al lado corto del mercado.

Teorema 4. *En un mercado bilateral uno-a-uno, donde ningún agente posee opciones de salida:*

(i) Para aquellos agentes que pertenecen al lado corto del mercado, la equivalencia en Chen, Egesdal, Pycia & Yenmez (2016, Teorema 1) es valida.

(ii) Para aquellos agentes que no pertenecen al lado corto del mercado, la equivalencia en Chen, Egesdal, Pycia & Yenmez (2016, Teorema 1) no es valida en ninguna de sus direcciones.

[10] Considere las siguientes propiedades para el agente $h \in E \cup G$: i) $\forall \succ \in \mathcal{P} : \varphi[\succ](h) \succeq_h \psi[\succ](h)$ y ii) $\exists \succ \in \mathcal{P} : \varphi[\succ](h) \succ_h \psi[\succ](h)$. Luego, comparar mecanismos según su grado de manipulabilidad es equivalente a compararlos en preferencias de forma estricta para el agente h, cuando: i) y ii) se cumplen para $h \iff \varphi$ es menos manipulable que ψ para h.

Demostración. Sean φ y ψ dos mecanismos estables.

(i) Asumiendo que $|E| < |G|$, supongamos que $\varphi[\succ](h) \succeq_h \psi[\succ](h)$ para cada $\succ \in \mathcal{P}$, con $h \in E$. Además, tomemos un matching $\mu \in K[h, \varphi, \succ]$. Como $\mu(h) \succ_h \psi[\succ](h)$, el Remark 1 nos asegura que existe una preferencia $\succ'_h$ tal que $\psi[\succ'_h, \succ_{-h}](h) = \mu(h)$, lo cual implica que $\mu \in K[h, \psi, \succ]$. Concluimos que $K[h, \varphi, \succ] \subseteq K[h, \psi, \succ]$.

En la dirección contraria, asumamos que φ es débilmente menos manipulable que ψ para el estudiante h. Por contradicción, supongamos que existe $\succ \in \mathcal{P}$ tal que $\psi[\succ](h) \succ_h \varphi[\succ](h)$. Luego, el Remark 1 implica que $\psi[\succ](h) \in K[h, \varphi, \succ]$. Esto, sumado a que φ es débilmente menos manipulable que ψ, implica que $\psi[\succ](h) \in K[h, \psi, \succ]$. Una contradicción con que $K[h, \psi, \succ]$ es el set de mejoras del estudiante h.

(ii) Sea $H \in \{E, G\}$ el lado del mercado que no es el corto. En Sirguiado & Torres-Martinez (2024), los autores muestran que para cada $h \in H$ existe un mecanismo Ω^h estable y strategy-proof para H, tal que para algún perfil $\succ \in \mathcal{P}$ se cumple $\mathrm{AD}_H[\succ](h) \succ_h \Omega^h[\succ](h)$. Luego, como los mecanismos Ω^h, AD_H son strategy-proof para el conjunto H, sabemos que Ω^h es débilmente menos manipulable que AD_H para el guía h en cada perfil de preferencias. Sin embargo, para al menos un perfil $\succ \in \mathcal{P}$ se cumple $\mathrm{AD}_H[\succ](h) \succ_h \Omega^h[\succ](h)$. Concluimos que débilmente menos manipulable no implica dominancia débil en preferencias para los agentes que no pertenecen al lado corto.

En la otra dirección, sabemos que es posible que dos mecanismos estables asignen el mismo emparejamiento y que un agente h manipule solo a uno de ellos (ver footnote 8). Concluimos que dominancia débil en preferencias no implica menos manipulable en su forma débil. $\square$

Este resultado muestra que en ausencia de opciones de salida, basta preguntarle a un agente del lado escaso del mercado que mecanismo prefiere para saber cuál mecanismo es menos manipulable por él. Adicionalmente, este sencillo test no es valido para agentes que pertenecen al grupo de mayor tamaño (ni para agentes que participan en mercados balanceados). De esta forma, los agentes pertenecientes al lado escaso siempre prefieren mecanismos strategy-proof para su lado del mercado antes que mecanismos manipulables, lo cual complementa a lo mostrado en Akbarpour, Kapor, Neilson, Van Dijk & Zimmerman (2022), quienes prueban que los agentes que no poseen opciones de salida siempre prefieren mecanismos strategy-proof antes que manipulables.[11]

[11]Los autores utilizan el conocido modelo de Abdulkadiroglu et al. (2011) para probar sus resultados.

6. Conclusión

En mercados bilaterales uno-a-uno, cuando todos los agentes se comportan de forma estratégica, hemos mostrado el efecto de la ausencia de opciones de salida sobre la comparación de mecanismos estables. Para esto, hemos utilizado dos criterios que previamente han sido usados en la literatura. Dentro de mercados desbalanceados, encontramos que el mecanismo de aceptación diferida, con el lado corto haciendo las propuestas, minimiza la manipulación bajo ambas métricas de comparación. Note que esta propiedad también es alcanzada por otros mecanismos estables. Es importante señalar que esto no es cierto para el mecanismo de aceptación diferida, cuando el lado largo realiza las propuestas. Más aún, no existe un mecanismos estable y strategy-proof para el lado largo del mercado, que minimice la cantidad de agentes manipuladores. Esto se debe a la ventaja que genera la escasez dentro de los mercados bilaterales, la cual ya ha sido reportada en trabajos previos (c.f. Ashlagi, Kanoria & Leshno (2016)). Adicionalmente, mostramos que dentro de los mecanismos estables y strategy-proof, hay uno que es el menos manipulable.

En mercados balanceados, hemos demostrado que las dos versiones del mecanismo de aceptación diferida no son comparables entre sí bajo ninguna de nuestras métricas de comparación. Además, no existe un mecanismo estable y one-side strategy-proof que minimice la cantidad de agentes manipuladores. En este contexto, existen dos preguntas que no han sido respondidas: (i) ¿existen mecanismos estables que minimicen la cantidad de agentes manipuladores en todo perfil de preferencias? (ii) Dentro de los mecanismos estables y strategy-proof para un lado del mercado ¿existe uno que sea el débilmente menos manipulable en cantidad de agentes? Las respuestas a estas preguntas quedan abiertas para investigación futura.

Cuando estudiamos la comparación de mecanismos estables a nivel individual, logramos demostrar que la equivalencia entre comparar mecanismos en base a preferencias y según su nivel de manipulabilidad, solo se cumple en mercados desbalanceados y solo para aquellos participantes que pertenecen al lado corto. De esta forma, estos agentes siempre preferirán mecanismos estables y strategy-proof (para su lado del mercado) antes que manipulables. Note que esto no es cierto para individuos que no pertenecen al lado largo del mercado, ya que ninguna de las direcciones de la equivalencia es valida para ellos.

Nuestros resultados pueden ser adaptados a modelos muchos-a-uno y muchos-a-muchos, donde ambos lados del mercado actúan estratégicamente. Sin embargo, como muchos de nuestros resultados requieren del Teorema del Hospital Rutal, del Teorema de los Limites de la Manipulación y de la estructura de lattice del conjunto de los emparejamientos estables, es de esperar que en contextos más generales necesitemos asumir cierta estructura sobre las preferencias (c.f. Echenique & Oviedo (2006)). La principal razón para escribir esta tesis en el modelo uno-a-uno fue para mantener el lenguaje y las demostraciones lo más simple posible. Es de esperar que futuras versiones de este trabajo aborden contextos más generales.

Apéndice A: Demostraciones omitidas

Demostración de la Proposición 1. Sean $E = \{e_1, e_2, \ldots, e_s\}$ y $G = \{g_1, g_2, \ldots, g_r\}$ los conjuntos de estudiantes y guías, con $s < r$. Denotemos por $\succ^* \in \mathcal{P}$ al perfil de preferencias dado por:

$\succ^*_{e_1}$	$\succ^*_{e_2}$	$\succ^*_{e_3}$	$\succ^*_{e_4}$	$\succ^*_{e_5}$	$\cdots$	$\succ^*_{e_s}$	$\succ^*_{g_1}$	$\succ^*_{g_2}$	$\succ^*_{g_3}$	$\succ^*_{g_4}$	$\cdots$	$\succ^*_{g_s}$	$\cdots$	$\succ^*_{g_r}$
g_2	g_1	g_1	g_2	g_5	$\cdots$	g_s	e_1	e_2	e_3	e_4	$\cdots$	e_s	$\cdots$	$\vdots$
g_1	g_2	g_3	g_4	$\vdots$		$\vdots$	e_2	e_1	$\vdots$	$\vdots$		$\vdots$		$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$

Es fácil comprobar que bajo $\succ^*$ los únicos matchings estables son:

$$\begin{aligned} \mathrm{AD}_E[\succ^*] &= \{(e_1, g_2), (e_2, g_1), (e_3, g_3), \ldots, (e_s, g_s), g_{s+1}, \ldots, g_r\}, \\ \mathrm{AD}_G[\succ^*] &= \{(e_1, g_1), (e_2, g_2), (e_3, g_3), \ldots, (e_s, g_s), g_{s+1}, \ldots, g_r\}. \end{aligned}$$

Considere el siguiente mecanismo estable:

$$\Omega[\succ] = \begin{cases} \mathrm{AD}_G[\succ], & \text{cuando } \succ = \succ^*, \\ \mathrm{AD}_E[\succ], & \text{cuando } \succ \neq \succ^*, \end{cases}$$

Mostraremos que la cantidad de agentes qué manipula AD_E bajo $\succ^*$ coincide con la de Ω. Primero note que ningún guía manipula Ω bajo $\succ^*$, de lo contrario habría una contradicción con la estabilidad de Ω, producto del Teorema de los Limites de la Manipulación. Por otro lado, el Lema 1 implica que este mecanismo es manipulado por e_1 y e_2 bajo el perfil $\succ^*$. Note que ningún otro estudiante manipula, ya que son emparejados con su mejor pareja estable.

Por otro lado, como AD_E es strategy-proof para los estudiantes, sabemos que los únicos que pueden manipular este mecanismo son los guías. De hecho, los agentes g_1 y g_2 pueden reportar los perfiles $e_1 \succ'_{g_1} e_3 \ldots$ y $e_2 \succ'_{g_2} e_4 \ldots$, con el fin de ser emparejados con $\mathrm{AD}_E[\succ'_{g_1}, \succ^*_{-g_1}](g_1) = e_1$ y $\mathrm{AD}_E[\succ'_{g_2}, \succ^*_{-g_2}](g_2) = e_2$, respectivamente. Dado que el resto de guías son emparejados con su mejor pareja estable, ningún otro manipula o habría una contradicción con el Teorema de los Limites de la Manipulación. Concluimos que la cantidad de agentes manipuladores bajo ambos mecanismos es igual a dos.

Para cerrar la demostración, mostraremos que para el resto de perfiles de preferencias un agente manipula el mecanismo Ω si y solo si manipula AD_E. Primero, note que en estos perfiles ambos mecanismos asignan el matching estable óptimo para los estudiantes, por lo cual, ningún agente $e \in E$ manipula. De lo contrario, existiría una contradicción con la estabilidad de los mecanismos, producto del Teorema de los Limites de la Manipulación.

Ahora, supongamos que un agente $g \in G$ manipula alguno de los dos mecanismos bajo $\succ \neq \succ^*$, mediante la preferencia $\succ'_g$. Existen dos casos relevantes:

- Si $(\succ'_g, \succ_{-g}) \neq \succ^*$, entonces $\Omega[\succ] = \mathrm{AD}_E[\succ]$ y $\Omega[\succ'_g, \succ_{-g}] = \mathrm{AD}_E[\succ'_g, \succ_{-g}]$. Luego, si uno de los mecanismos es manipulado, el otro también lo será.
- Si $(\succ'_g, \succ_{-g}) = \succ^*$, sabemos que las preferencias de todos los agentes diferentes de g coinciden en los perfiles $\succ$ y $\succ^*$, es decir, $\succ_{-g} = \succ^*_{-g}$. Note que si $g \in \{g_3, \ldots, g_r\}$, tenemos que $\mathrm{AD}_E[\succ$

$](g) = \mathrm{AD}_G[\succ](g)$ para todo perfil $\succ = (\succ_g, \succ^*_{-g})$. Luego, el Teorema de los Limites de la Manipulación nos asegura que $g \in \{g_1, g_2\}$. Supongamos que $g = g_1$. Note que para todo perfil $\succ_{g_1}$ tal que $e_1 \succ_{g_1} e_2$ y $e_2 \succ_{g_1} e_3$, el profesor guía g_1 manipulará AD_E y Ω, reportando el perfil $\succ'_{g_1}$, el cual fue definido anteriormente (el conjunto de estables bajo $\succ$ coincide con el de $\succ^*$). Por otro lado, si $e_2 \succ_{g_1} e_1$ o $e_3 \succ_{g_1} e_2$, entonces $\mathrm{AD}_E[\succ](g_1) = \mathrm{AD}_G[\succ](g_1)$, una contradicción con el Teorema de los Limites de la Manipulación. Adicionalmente, si $g = g_2$ note que para todo perfil $\succ_{g_2}$ tal que $e_2 \succ_{g_2} e_1$ y $e_1 \succ_{g_2} e_4$, el guía g_2 manipulará AD_E y Ω, reportando el perfil $\succ'_{g_2}$. Por otro lado, si $e_1 \succ_{g_2} e_2$ o $e_4 \succ_{g_2} e_1$, entonces $\mathrm{AD}_E[\succ](g_2) = \mathrm{AD}_G[\succ](g_2)$, una contradicción con el Teorema de los Limites de la Manipulación.

Concluimos que AD_E es igual de manipulable en cantidad de agentes que el mecanismo Ω, lo cual implica que también lo es en preferencias (ver footnote 4). $\square$

Demostración de la Proposición 2. Sea ψ un mecanismo estable y strategy-proof para G diferente de Ω^E y $\succ$ un perfil de preferencias cualquiera. Utilizando que $M^{\psi}[\succ]$, $M^{\Omega^E}[\succ] \subseteq E$, probaremos que $M^{\Omega^E}[\succ] \subseteq M^{\psi}[\succ]$. Supongamos que $e^* \in E$ manipula Ω^E. Esto implica que $\mathrm{AD}_E[\succ](e) \neq \Omega^E[\succ](e)$, de lo contrario existiría una contradicción con el Teorema de los Limites de la Manipulación. Además, como Ω^E es el mejor mecanismo para los estudiantes, dentro de los estables y strategy-proofs para los guías, sabemos que $\mathrm{AD}_E[\succ](e) \succ_e \Omega^E[\succ](e) \succeq_e \psi[\succ](e)$ para todo $e \in E$. Esto último, junto al Lema 1 implican que el estudiante e^* también manipula el mecanismo ψ.

Concluimos que $M^{\Omega^E}[\succ] \subseteq M^{\psi}[\succ]$, por lo que Ω^E es débilmente menos manipulable que ψ en cantidad de agentes (lo cual implica que también en preferencias). $\square$

Apéndice B: Mecanismos óptimos y mercados con opciones de salida

En este breve apéndice probaremos un nuevo resultado que aplica a mercados bilaterales uno-a-uno clásicos, es decir, cuando ambos lados del mercado son estratégicos y poseen opciones de salida. De esta forma, los estudiantes tienen preferencias definidas sobre $G \cup \{\emptyset\}$, mientras que las de los guías lo están sobre $E \cup \{\emptyset\}$. Denotaremos por $\mathcal{P}$ al conjunto de todos estos perfiles de preferencias.

Como ya hemos mencionado, en este contexto no hace sentido comparar mecanismos estables en preferencias, ya que en cualquier perfil todo mecanismo estable es manipulado o ninguno lo es. ¿Qué ocurre en términos de la cantidad de agentes que manipulan? En Bonkoungou & Nesterov (2023, Corolario 1) muestra que la cantidad de agentes que manipulan los mecanismos AD_E y AD_G minimizan la cantidad de agentes estratégicos en todo perfil de preferencias.

¿Existen mecanismos estables, diferentes de AD_E y AD_G que sean débilmente menos manipulable en cantidad de agentes que cualquier otro mecanismo estable? Para responder esta pregunta, diremos que un mecanismo Ω es **óptimo** si para todo $\succ \in \mathcal{P}$ se cumple $\Omega[\succ] \in \{\mathrm{AD}_E[\succ], \mathrm{AD}_G[\succ]\}$. Así, un mecanismo óptimo Ω puede ser representado de la siguiente forma:

$$\Omega[\succ] = \begin{cases} \mathrm{AD}_G[\succ], & \text{cuando } \succ \in K, \\ \mathrm{AD}_E[\succ], & \text{cuando } \succ \in K^c, \end{cases}$$

donde $K \cup K^c = \mathcal{P}$. De esta forma, cuando $K = \emptyset$ se tiene que $\Omega = \mathrm{AD}_E$ y cuando $K = \mathcal{P}$, $\Omega = \mathrm{AD}_G$. El siguiente resultado caracteriza completamente cuales son los mecanismos estables que minimizan la cantidad de agentes manipuladores.

Teorema 5. *En un mercado con opciones de salida, sea φ un mecanismo óptimo. Luego, φ minimiza la cantidad de agentes manipuladores en todo perfil de preferencias.*

Demostración. Denotemos por E^φ y G^φ a la cantidad de estudiantes y guías, respectivamente, que no son emparejados con su mejor pareja estable al aplicar φ (análogo para los mecanismos AD_E y AD_G). Sea Ω un mecanismo óptimo. Fijemos un perfil $\succ \in \mathcal{P}$ y sin perdida de generalidad, supongamos que $\Omega[\succ] = \mathrm{AD}_E[\succ]$. Afirmamos que ningún estudiante manipula el mecanismo Ω bajo $\succ$. Efectivamente, de no ser así, existiría $e \in E$ que manipula un mecanismo estable (Ω) para alcanzar un guía mejor que su pareja estable óptima. Sin embargo, esto último, junto al Teorema de los Limites de la Manipulación implicarían que Ω no es estable. Una contradicción.

Por otro lado, dado que $\Omega[\succ] = \mathrm{AD}_E[\succ]$, tenemos $G^\Omega = G^{\mathrm{AD}_E}$ bajo el perfil $\succ$. Luego, como todo guía manipula aquellos mecanismos estables en los que no es emparejado con su mejor pareja estable (Roth 1982, Lema 1), los guías que manipulan los mecanismos AD_E y Ω bajo el perfil $\succ$, coinciden.

Por simetría, lo mismo ocurre bajo perfiles de preferencias en los que se cumple $\Omega[\succ] = \mathrm{AD}_G[\succ]$. Luego, cualquier mecanismo óptimo posee la misma cantidad de agentes manipuladores que las dos versiones del mecanismo de aceptación diferida. Cómo estos últimos mecanismos minimizan la cantidad de agentes manipuladores en todo perfil de preferencias (Bonkoungou & Nesterov, 2023), todo mecanismo óptimo también cumplirá esta propiedad. □

Este resultado implica que si un planificador central desea implementar un matching estable, minimizando la cantidad de agentes manipuladores, es suficiente escoger un mecanismo óptimo. ¿Cómo hacerlo? Basta que el planificador defina el conjunto K mediante alguna regla. A continuación, comentamos brevemente dos criterios que pueden ser utilizados para escoger el conjunto K.

1) **Agentes preferenciales**. Supongamos que el planificador desea implementar el mejor matching estable para los estudiantes, sujeto a que un subconjunto de guías siempre reciba su mejor pareja estable. Contextos como este podrían ocurrir cuando algunos profesores tienen contratos especiales que le aseguran trabajar con los mejores estudiantes dentro de lo posible (mientras se mantenga la estabilidad). Con este criterio en mente, denotemos al conjunto de guías preferenciales como G_c y definamos el siguiente subconjunto de preferencias:

$$K = \{\succ \in \mathcal{P} :\ \exists g \in G_c,\ \mathrm{AD}_G[\succ](g) \neq \mathrm{AD}_E[\succ](g)\}$$

Así, el mecanismo óptimo que es inducido por este subconjunto de preferencias asigna el mejor estable óptimo para los estudiantes solo en aquellos perfiles de preferencias en que todos los guías preferenciales están indiferentes entre las dos versiones del mecanismo de aceptación diferida. Denotemos por Ω a este mecanismo.

¿Cuanto mejoran en promedio los estudiantes al implementar este mecanismo óptimo en vez de AD_G? Para responder esta pregunta simularemos mercados de emparejamiento aleatorios, siguiendo lo hecho por Ashlagi, Kanoria & Leshno (2016). Esto es, generamos perfiles de preferencias siguiendo una distribución uniforme y aplicamos los mecanismos de interés a cada uno de estos perfiles. Denotemos por $Rank_e(g) = |\{g' \in G : g' \succeq g\}|$ al ranking que el estudiante e le asigna al guía g.

Dado un matching μ, diremos que el ranking promedio de guías asignados a estudiantes viene dado por:

$$R_E(\mu) = \frac{1}{|E|} \sum_{e \in E} Rank_e(\mu(e)).$$

Analogamente, definiremos $R_G(\mu)$ como el ranking promedio de estudiantes obtenidos por guías bajo el matching μ. La Figura 1 muestra como varía, en promedio, las métricas R_E y R_G al modificar la proporción de profesores preferenciales.

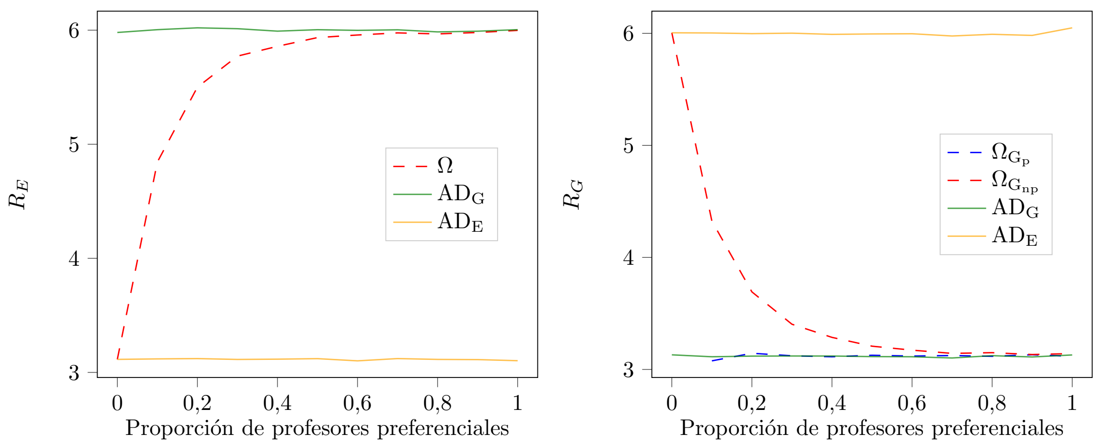

FIGURA 1. Ranking promedio de estudiantes y profesores al variar el conjunto de profesores preferenciales. Las curvas muestran el promedio de 10000 realizaciones, fijando el tamaño del mercado en 20 agentes en cada lado. Por simplicidad, asumimos que para cada agente su opción de salida es su alternativa menos preferida. En el panel de la derecha, la curva Ω_{G_p} y $\Omega_{G_{np}}$ denotan el ranking promedio de profesores preferenciales y no preferenciales, respectivamente.

En el gráfico de la izquierda se logra observar que, cuando no existen guías preferenciales, el mecanismo Ω coincide con AD_E en términos de R_E y a medida que aumenta la proporción de guías en esa condición, el mecanismo optimo va convergiendo a AD_G. Por otro lado, se puede ver que los guías preferenciales obtienen el mismo ranking bajo Ω y AD_G, y son los guías que no poseen ese beneficio quienes pagan el costo para mejorar a los estudiantes bajo el mecanismo óptimo.

2) **Preferencias correlacionadas**. En la práctica, se ha observado que las personas tienden a entregar una gran importancia a quedar con sus primeras preferencias. Por ejemplo, en contextos de elección escolar, una de las grandes críticas que han hecho los padres a la implementación del mecanismo de aceptación diferida ha sido el no quedar dentro de sus colegios más preferidos (cf. Dur, Mennle & Seuken (2019)). Esto implica que, la cantidad de agentes que son emparejados con su primera preferencia es un criterio de elección razonable a la hora de definir el conjunto K.

Denotemos por $Prop_{pp}(\mu)$ a la proporción de agentes del mercado que son emparejados con su primera preferencia bajo el matching μ. Esto es,

$$Prop_{pp}(\mu) = \frac{|\{e \in E : Rank_e(\mu(e)) = 1\}| + |\{g \in G : Rank_g(\mu(g)) = 1\}|}{|E \cup G|}$$

Mediante simulaciones, mostraremos que cuando las preferencias de los estudiantes están más correlacionadas que las de los profesores, implementar el matching óptimo para los guías logra una mayor proporción de agentes emparejados con sus primeras preferencias que al asignar el emparejamiento óptimo para los estudiantes. La intuición de este resultado es que existe una fuerte competencia entre los estudiantes por obtener a algunos profesores y una baja competencia entre estos últimos por obtener a los primeros.

Para generar preferencias correlacionadas, usaremos una adaptación del modelo de utilidad aleatoria de Hitsch et al. (2010), nuevamente siguiendo lo hecho en Ashlagi, Kanoria & Leshno (2016). Cada agente h posee dos caracteristicas, x_h^A y x_h^B. Luego, la utilidad h de ser emparejado con el agente j viene dada por:

$$u_h(j) = \beta^H x_j^A - \gamma(x_h^B - x_j^B) + \epsilon_{hj},$$

donde ϵ_{ij} es un término ideosincrático del par (h, j), distribuido de manera independiente por una distribución logistica. Además, $x_h^A \sim U[0, 1]$ y $x_h^B \sim U[0, 1]$, de manera independiente entre los agentes. Mientras el parametro γ mide que tan simetricas son las preferencias entre los lados del mercado, el parametro β^H mide que tan correlacionadas son las preferencias del lado del mercado $H \in \{E, G\}$.

Utilizando esta notación, definamos el subconjunto $K = \{\succ \in \mathcal{P} : \beta^E > \beta^G\}$, y definamos como Ω al mecanismo óptimo que es inducido por K. De esta forma, Ω asigna el mejor matching estable para los guías solo cuando las preferencias de estos están a lo más tan correlacionadas como la de los estudiantes. Caso contrario, se asigna el matching $\mathrm{AD}_E[\succ]$. En el gráfico de la izquierda de la Figura 2, fijando el parámetro $\beta^G = 5$ y variando β^E, se observa como Ω logra aumentar la proporción de agentes emparejados con su primera preferencia respecto a las dos versiones del mecanismo de aceptación diferida.

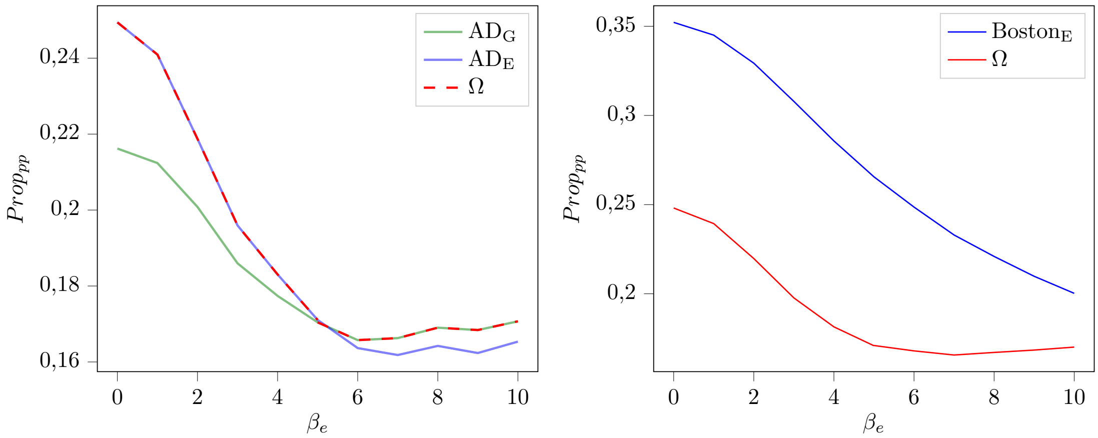

Figura 2. Proporción de agentes emparejados con su primera preferencia al variar qué tan correlacionadas son las preferencias de los estudiantes. Fijamos $\beta^G = 5$. Las curvas muestran el promedio de 10000 realizaciones, fijando el tamaño del mercado en 20 agentes en cada lado. Por simplicidad, asumimos que para cada agente su opción de salida es su alternativa menos preferida.

Sabemos que el mecanismo de Boston (Abdulkadiroğlu & Sönmez, 2003) pertenece a una familia de mecanismos denominados *first choice mechanisms* (Dur, Mennle & Seuken (2019)). Estos mecanismos se caracterizan porque ningún estudiante forma un par de bloqueo con su primera preferencia reportada y por maximizar la cantidad de estudiantes emparejados con su primera opción. ¿Qué tan bueno es Ω emparejando agentes con su primera preferencia comparado con el mecanismo de Boston? En el panel derecho de la Figura 2 se observa que el último de estos mecanismos domina a nuestro mecanismo óptimo para cualquier valor de β^E. Esto muestra que Ω esta lejos de ser el mejor mecanismo si es que la autoridad desea maximizar la proporción de agentes emparejados con sus primera preferencias. Sin embargo, es importante notar que la mejora que hace este mecanismo óptimo respecto a las versiones de aceptación diferida, es manteniendo la estabilidad y minimizando la cantidad de agentes manipuladores en todo perfil de preferencias, propiedades que el mecanismo de Boston está lejos de cumplir.

Referencias

[1] Abdulkadiroğlu, A., & T. Sönmez (2003): "School choice: a mechanism design approach," *American Economic Review*, 93, 729-747.

[2] Abdulkadiroğlu, A., Che, Y. K., & Yasuda, Y. (2011): "Resolving conflicting preferences in school choice: The Boston mechanism reconsidered". *American Economic Review*, 101(1), 399-410.

[3] Alcalde, J., & S. Barberà (1994): "Top dominance and the possibility of strategy-proof stable solutions to matching problems," *Economic Theory*, 4, 417-435.

[4] Andersson, T., Ehlers, L., & Svensson, L. G. (2014): "Budget balance, fairness, and minimal manipulability," *Theoretical Economics*, 9(3), 753-777.

[5] Akbarpour, M., A. Kapor, C. Neilson, W. Van Dijk, & S. Zimmerman (2022): "Centralized school choice with unequal outside options," *Journal of Public Economics*, 210, 104644.

[6] Arribillaga, R. P., & Massó, J. (2016): "Comparing generalized median voter schemes according to their manipulability," *Theoretical Economics*, 11(2), 547-586.

[7] Ashlagi, I., Kanoria, Y., & Leshno, J. D. (2017): "Unbalanced random matching markets: The stark effect of competition," *Journal of Political Economy*, 125(1), 69-98.

[8] Bonkoungou, S., & Nesterov, A. (2021): "Comparing school choice and college admissions mechanisms by their strategic accessibility," *Theoretical Economics*, 16(3), 881-909.

[9] Bonkoungou, S., & Nesterov, A. (2023): "Incentives in matching markets: Counting and comparing manipulating agents," *Theoretical Economics*, 18(3), 965-991.

[10] Chen, P., Egesdal, M., Pycia, M., & Yenmez, M. B. (2016): "Manipulability of stable mechanisms," *American Economic Journal: Microeconomics*, 8(2), 202-214.

[11] Chen, Y., & Kesten, O. (2017): "Chinese college admissions and school choice reforms: A theoretical analysis," *Journal of Political Economy*, 125(1), 99-139.

[12] Decerf, B., & Van der Linden, M. (2021): "Manipulability in school choice," *Journal of Economic Theory*, 197, 105313.

[13] Demange, G., Gale, D., & Sotomayor, M. (1987): "A further note on the stable matching problem," *Discrete Applied Mathematics*, 16(3), 217-222.

[14] Dubins, L.E., & D.A. Freedman (1981): "Machiavelli and the Gale-Shapley algorithm," *The American Mathematical Monthly*, 88, 485-494.

[15] Dur, U. (2019): "The modified Boston mechanism," *Mathematical Social Sciences*, 101, 31–40.

[16] Dur, U., Hammond, R. G., & Morrill, T. (2019): "The secure Boston mechanism: Theory and experiments," *Experimental Economics*, 22, 918-953.

[17] Dur, U., Pathak, P. A., Song, F., & Sönmez, T. (2022): "Deduction dilemmas: The Taiwan assignment mechanism," *American Economic Journal: Microeconomics*, 14(1), 164-185.

[18] Gale, D., & L. Shapley (1962): "College admissions and the stability of marriage," *The American Mathematical Monthly,* 69, 9-15.

[19] Gale, D., & M. Sotomayor (1985): "Some remarks on the stable matching problem," *Discrete Applied Mathematics*, 11, 223-232.

[20] Echenique, F., & Oviedo, J. (2004): "A theory of stability in many-to-many matching markets," *Theoretical Economics*, 1, 233–273

[21] Imamura, K., & Tomoeda, K. (2022): "Measuring manipulability of matching mechanisms," *SSRN 4000419.*

[22] Hitsch, G. J., Hortaçsu, A., & Ariely, D. (2010): "Matching and sorting in online dating," *American Economic Review*, 100(1), 130-163.

[23] Kesten, O. (2010): "School choice with consent," *The Quarterly Journal of Economics*, 125, 1297-1348.

[24] Kesten, O., & M. Kurino (2019): "Strategy-proof improvements upon deferred acceptance: A maximal domain for possibility," *Games and Economic Behavior,* 117, 120-143.

[25] Knuth, D. (1976): "Mariages stables et leurs relations avec d'autres problèmes combinatoires," Les Presses de l'Université de Montréal, Montréal

[26] Nesterov, A., Rospuskova, O. & Rubtcova, S (2024): "Robustness to manipulations in school choice," *Social Choice and Welfare*.

[27] Roth, A. (1982): "The economics of matching: stability and incentives," *Mathematics of Operations Research*, 7, 617-628.

[28] Pathak, P. A., & Sönmez, T. (2013): "School admissions reform in Chicago and England: Comparing mechanisms by their vulnerability to manipulation," *American Economic Review*, 103(1), 80-106.

[29] Sirguiado, C. (2024): "A Characterization of stable mechanisms that minimize manipulation," disponible en SSRN: https://ssrn.com/abstract=4784280.

[30] Sirguiado, C., & Torres-Martínez, J. P. (2024): "Strategic behavior without outside options," disponible en SSRN: https://ssrn.com/abstract=4704049.

[31] Tadenuma, K., & M. Toda (1998): "Implementable stable solutions to pure matching problems," *Mathematical Social Sciences*, 35, 121-132.