Facultad de Ciencias Económicas y Empresariales

Trabajo de Fin de Grado

Doble Grado en Administración y Dirección de Empresas y Derecho

Módulo de Finanzas

Incertidumbre, tenis y apuestas deportivas:

La influencia del Covid-19 en la fijación de precios por parte de las casas de apuestas

Íñigo Ruiz Alén

Pamplona-Iruña, 21 de diciembre

Curso 2022/2023

Director: Luis Fernando Muga Caperos

RESUMEN

Ante la crisis producida por el Covid-19 tanto a nivel económico como social, este trabajo analiza la importancia del aspecto informacional en los mercados financieros, enfocándose en los mercados de apuestas deportivas, y más concretamente en el sector del tenis.

Analizaremos en primer lugar la existencia y empleo por parte de las casas de apuestas de un Sesgo del Favorito, para posteriormente tratar la fijación de precios. Para ello, nos centraremos en los márgenes establecidos por las casas de apuestas y su evolución en un periodo de incertidumbre como el que hemos vivido desde finales de 2019. De este modo, estudiaremos si realmente la incertidumbre generada por la pandemia, las asimetrías informativas y otros factores derivados de esta situación excepcional han tenido un impacto representativo sobre los precios de las apuestas.

Palabras clave: Horquilla, sesgo del favorito, tenis, incertidumbre, Covid-19.

ABSTRACT

In view of the crisis produced by Covid-19 at both the economic and social levels, this paper analyzes the importance of the informational aspect in financial markets, focusing on sports betting markets, and more specifically on the tennis sector.

We will first analyze the existence and use by bookmakers of a Favorites Bias, and then deal with pricing. To do so, we will focus on the margins set by bookmakers and their evolution in a period of uncertainty such as the one we have experienced since the end of 2019. In this way, we will study whether the uncertainty generated by the pandemic, information asymmetries and other factors derived from this exceptional situation have really had a representative impact on betting prices.

Key words: overround, favourite-longshot bias, tennis, uncertainty, Covid-19.

ÍNDICE DE CONTENIDOS

1. INTRODUCCIÓN 4

2. MARCO TEÓRICO 6

2.1. Hipótesis de los Mercados Eficientes 6

2.2. Comparativa mercados financieros - mercados de apuestas 10

3. SESGO DEL FAVORITO 11

4. EL MARGEN 13

5. BASE DE DATOS 14

5.1. Diferencias en las ganancias (o pérdidas) durante el Covid-19 por el Sesgo del Favorito 15

5.2. Análisis del margen 20

5.3. Regresión Lineal 28

6. CONCLUSIONES 34

7. REFERENCIAS 36

8. ANEXO 38

1. INTRODUCCIÓN

Este trabajo desarrolla el aspecto de la fijación de precios y los sesgos informativos dentro del mundo de las apuestas deportivas, orientado al tenis y en relación con la situación originada por el Covid-19. Se centrará por tanto en los mercados dirigidos por precios (*bookmakers*), ya que también existen los mercados dirigidos por órdenes (*betting exchanges*). Los mercados de apuestas deportivas dirigidas por precios son aquellos en los que hay un agente, el *bookmaker*, que se encarga de ofrecer la posibilidad de apostar, pero sin permitir la negociación de precios en función del volumen de apuesta. Tras haber analizado toda la información involucrada en el partido en cuestión, fija unos precios a cada suceso, que serán la inversa de las probabilidades de ocurrencia asignadas a dicho suceso.

Únicamente en el último trimestre de 2022, el sector del comercio electrónico de juegos de azar y apuestas generó en España una facturación de 740,83 millones de euros[1], con una tendencia alcista muy pronunciada en las últimas décadas. Los jugadores activos en el juego online, de hecho, aumentaron desde 2013 a 2018 un 130%[2], llegando a superar los 3 millones en este último año.

Si bien en 2019 por primera vez la facturación bruta en el mundo de las apuestas deportivas no creció (y de hecho en 2020 cayó levemente), entre octubre y diciembre de 2020 se superaron por primera vez en la historia de la economía española los 2.000 millones de euros de facturación total[3]. La proliferación de plataformas online para apostar e invertir en los últimos años ha tenido una gran influencia en la masificación de estas prácticas, lo que ha hecho que la competencia aumente y haya más información en juego. Los datos mencionados nos muestran la relevancia de las apuestas deportivas dentro de la economía española, teniendo en cuenta que se trata de una actividad de ocio, y en un marco en el que la crisis económica se ha agudizado por diversas causas, entre las que se encuentra el Covid-19, factor determinante sobre el que se sustenta este trabajo.

[1] https://es.statista.com/estadisticas/1030565/juegos-de-azar-y-apuestas-facturacion-comercio-electronico-espana/
[2] https://economipedia.com/definiciones/la-evolucion-de-las-apuestas-en-espana-en-los-ultimos-anos.html
[3] https://www.2playbook.com/mas-deporte/casas-apuestas-rompen-su-techo-pese-veto-gobierno-facturan-11000-millones-en-2021_8105_102.html

De hecho, resulta interesante que, en el momento en que más incierto podía ser el futuro de este tipo de actividades (ya que se ha llegado a poner en debate si se había alcanzado ya un techo que no podía superarse en términos relativos), el Covid-19 irrumpió en todos los ámbitos, tanto sociales como sanitarios y, por supuesto, en el ámbito económico.

En este contexto, el trabajo que abordamos no versará sobre cuestiones sociológicas (pese a que sí estudiaremos la existencia de un Sesgo del Favorito, sólo que desde el punto de vista del "*bookmaker*") sobre los efectos del Covid-19 en los comportamientos de las personas que apuestan. Nos centraremos en los efectos que ha tenido en las casas de apuestas, y, más concretamente, en la reacción que han tenido para adaptarse a la falta de información e inestabilidad producidas durante este periodo en lo que a fijación de precios respecta.

Trataremos, en primer lugar, el Sesgo del Favorito en las apuestas de tenis, observando las posibles variaciones que hayan tenido o no lugar a raíz del Covid-19. Así, veremos si realmente han aumentado las pérdidas al apostar al jugador con menos probabilidad asignada de victoria. Esto nos demostraría que, en periodos de incertidumbre, aumentaría el Sesgo del Favorito o "*Favourite-Longshot Bias*" (Cain, 2000). Esto se debe a que, las casas de apuestas, para adaptarse ajustan los precios, pues en caso de darse el resultado menos probable, estarían desembolsando cantidades muy grandes de dinero.

Estas situaciones llevan a incrementar el margen que las casas de apuestas establecen para cubrir gastos, al tener que protegerse ante situaciones de incertidumbre y asimetrías informativas y generar un beneficio. Como expone Lahvička (2014), "los *bookmakers* podrían potencialmente perder mucho dinero si infravaloran a los *longshots* y esta valoración errónea es explotada por las personas con información privilegiada o por el público general reaccionando más rápido que ellos a la nueva información".

Por ello, estudiaremos las variaciones en los márgenes de las casas de apuestas a raíz del Covid-19, teniendo en cuenta los factores que pueden hacer que éste varíe. De este modo, podremos estudiar la adaptación necesaria por parte de las casas de apuestas a raíz de la pandemia, teniendo en cuenta que es extrapolable a cualquier situación de los mercados financieros, en general, y de apuestas, en particular, y cuya causa tiene un valor fundamental en la fijación de precios: la incertidumbre.

Comenzaremos, por tanto, con una introducción a la teoría que sustenta este trabajo. Comentaremos la Hipótesis de los Mercados Eficientes, y su importancia en la fijación de precios. Posteriormente, hablaremos del Sesgo del Favorito y la implicación que tiene en periodos de mayor incertidumbre. Para ello, contrastaremos la rentabilidad obtenida al implementar dos estrategias diferenciadas: apostar siempre al que más probabilidades tiene de ganar, y apostar siempre, al contrario. En segundo lugar, analizaremos el comportamiento de los márgenes en la época Covid y no-Covid, además de estudiar las variables de control y su impacto relativo en el margen de manera unifactorial. Por último, pondremos en común las diferentes variables que creemos afectan al margen u horquilla de manera conjunta, en un análisis de regresión lineal simple. De esta forma, habremos estudiado de diferentes formas la incertidumbre como parte del proceso de fijación de precios, y la importancia de la información en el mundo de los mercados financieros, en general, y en el mundo de las apuestas deportivas, en particular.

2. MARCO TEÓRICO

2.1. Hipótesis de los Mercados Eficientes

En el mundo de las apuestas, así como en el de los mercados de valores, los jugadores o inversores tratan de maximizar su beneficio, sacando una rentabilidad del mercado mediante la predicción de un resultado. La Hipótesis de los Mercados Eficientes, desarrollada por Fama (1970), y aplicada a los mercados de apuestas, determina que dicho beneficio, a largo plazo, no es posible. El concepto fundamental de la teoría de Fama es el siguiente: *"Un mercado financiero eficiente es aquel en el cual los precios de los valores reflejan plenamente la información disponible".*

A lo que se refiere esta frase es que la eficiencia que anteriormente podría entenderse en el sentido de Pareto ("en equilibrio no puede mejorar la situación de uno sin empeorar la de otro"), a nivel informacional implica que los precios de un activo deben incorporar constantemente la totalidad de la información relevante que existe sobre dicho activo. Así, no cabría obtener un beneficio del empleo de la información para especular

sobre un activo, ya que, en el momento que se emplease, estaría ya incorporada al precio del mismo.

En su trabajo, Fama propone lo siguiente. Para conocer el precio que se espera de una acción: “el valor en el momento t+1 es igual a uno más el retorno esperado por esa acción en el momento t+1, en tanto en cuanto toda la información sobre dicha acción (o valor) está totalmente reflejada en su precio en el instante en que se observa”. La conclusión de sería que, teniendo en cuenta que dicha rentabilidad es incierta y está dada por una variable aleatoria, el valor futuro de la acción no podrá predecirse (Riaño, 2018). Siendo así, la probabilidad de ganancias haciendo un riguroso análisis de la información sería idéntica a la probabilidad de beneficiarse mediante una elección al azar.

De este modo, la teoría muestra la imposibilidad de sacar una rentabilidad en caso de darse las condiciones establecidas, ya que la suma entre beneficios y pérdidas de los que apuestan a favor o en contra de un suceso que finalmente se da, es cero. Este significado de eficiencia se produce en un sentido de Pareto, implicando que no se puede aumentar la “utilidad” de un sujeto sin reducir la del otro.

Para comenzar, revisaremos el marco conceptual de esta teoría, que necesita del cumplimiento de ciertos supuestos indispensables para darse en la realidad: la inexistencia de costes de transacción, la accesibilidad a toda la información relevante sobre los valores que están en juego para todos los jugadores sin coste y el consenso en las implicaciones de la información mencionada sobre el valor futuro de los activos (Fama, 1971).

De este modo, la teoría actual establece adicionalmente que los mercados son eficientes cuando existe información perfecta, los inversores son racionales y no hay límites al arbitraje.

El arbitraje se entiende como la posibilidad de compraventa en diferentes mercados al mismo tiempo un mismo activo. Sharpe y Alexander (1990) desarrollan teorías previas sobres cómo los arbitrajistas serían los que aprovechan las posibles infravaloraciones o sobrevaloraciones de los activos por parte de los fijadores de precios, comerciando con ellos en diferentes mercados aprovechando las diferencias de valoración. Este efecto hace que a largo plazo los precios vuelvan a su valor teórico, al equilibrarse la actuación de todos los arbitrajistas, que venden cuando los activos están

sobrevalorados (mediante créditos o derivados financieros) y compran cuando consideran que los activos están infravalorados y piensan que en un futuro va a subir el valor.

En las casas de apuestas dirigidas por precios es más complicado el arbitraje, ya que a la hora de fijarse los precios se tienen en cuenta muchos factores, y no es la "mano invisible" la que se encarga de equilibrar los valores. El volumen de apuesta tiene al fin y al cabo mucha influencia en las cuotas que establecen los "*bookmakers*".

Respecto a la racionalidad de los inversores, la teoría financiera define esta como el comportamiento de los mismos en busca de una maximización de su beneficio ante alternativas del mismo nivel de riesgo. Los inversores deben ser racionales, y si no lo son, perderán en favor de los primeros (Villalba, 2003).

Pese a esta hipótesis, la realidad empírica nos resigna a aceptar que existen en el mercado inversores irracionales que negocian con cierta aleatoriedad o movidos por sus emociones.

Hay inversores que reaccionan a la información relevante en su justa medida, analizándola y utilizándola de una manera lógica para sus predicciones, pero muchos de ellos se centran más en el ruido, y sobre-reaccionan a hechos con mucho valor mediático pero que no tienen una influencia tan directa en la realidad (De Carlos, 2021).

Sin embargo, la hipótesis de los mercados eficientes no se desmorona por la mera existencia de inversores irracionales. El hecho que tendría que darse para que esta irracionalidad sea eliminada por compensación es la correlación nula o aleatoriedad entre los comportamientos irracionales. Al fin y al cabo, cuando hay volúmenes tan altos de inversores, la cancelación entre estrategias de inversión irracionales hace que los precios no se desvíen de su valor fundamental (Shleifer, 2020).

La evidencia empírica, sin embargo, muestra cómo los comportamientos de los inversores irracionales atienden a unos sesgos muchas veces comunes, que la psicología financiera ha estudiado ampliamente *("behavioral finance").* Este sector de los participantes del mercado es conocido como "Negociadores ruidosos no relacionados" (Abinzano, Muga y Santamaría, 2020). Uno de los sesgos más comunes en los mercados de apuestas es el conocido como *"Favourite-Longshot Bias"*, que hace que las apuestas a favor del favorito tengan un mayor valor esperado que el de el contrario *("longshots"*). Al igual que este sesgo, hay muchos otros que nos podrían llevar a la conclusión de que

no se puede dar la eficiencia cuando hay inversores irracionales con comportamientos homogéneos.

Sin embargo, los arbitrajistas funcionan como contrapeso frente a este hecho, ya que al sobrevalorarse un valor en el mercado por la compra excesiva de los inversores irracionales movidos por sesgos ("*Favourite-Longshot Bias*", "Sesgo de Día Festivo", etc.), los inversores formados o arbitrajistas venderán en corto y comprarían otros valores similares para cubrir riesgos. A la contra ocurriría lo mismo, y es que los arbitrajistas tratan de obtener un beneficio, pero además así consiguen que los precios no se desvíen de su valor fundamental. Reducirán el precio con sus ventas si el efecto del arbitraje se incorpora rápidamente al precio de los activos.

Como vemos, la participación irracional es compatible con la teoría de los mercados eficientes en la medida que la irracionalidad es compensada entre los que invierten en un sentido o en el contrario cuando sus comportamientos no están correlacionados. Y en caso de haber correlación entre estos comportamientos irracionales, los arbitrajistas sabrán aprovechar las ventajas informacionales para sacar rédito económico, balanceando de nuevo los precios y evitando que se desvíen de su valor fundamental.

En cuanto a la información perfecta, en este trabajo es fundamental, ya que los presupuestos requeridos por Fama para su teoría requieren de una incorporación inmediata de toda la información existente sobre los activos en su precio. Si se produce una incorporación rápida, los inversores que lleguen mínimamente tarde no deberían ser capaces de obtener beneficios (De Carlos, 2021).

Dentro de este marco teórico comentado, Fama (1970) distingue entre la forma débil, la forma semi-fuerte y la forma fuerte de eficiencia de mercado. La forma débil indica que no se pueden obtener beneficios del mercado únicamente con la información de los precios pasados.

La semi-fuerte o semi-débil dicta que no pueden obtener rendimientos superiores ajustados a su riesgo al utilizar, además de los precios pasados, cualquier información pública, ya que dicha información debería incorporarse inmediatamente a los precios según la teoría.

Por último, la forma fuerte de eficiencia de mercado dejaría en juego únicamente la información privilegiada de cara a obtener un beneficio. Aun así, la teoría defiende que este tipo de información tiende a difundirse rápidamente y tampoco permitiría sacar rédito económico, pero la forma que nos interesa para este trabajo es la semi-fuerte.

Lo que queremos con este trabajo es comprobar si el mercado responde de una forma inmediata y eficaz a los cambios de información, a las asimetrías y a la incertidumbre. Por ello, intentaremos entender cómo funciona el margen u horquilla ("*overround*") sobre el precio de las apuestas, y veremos su evolución en tiempos de incertidumbre como puede ser la época del coronavirus. Analizaremos el funcionamiento de los mercados en el sentido de eficiencia semi-fuerte, lo que nos permitirá entender la implicación de la información en la fijación de precios, su incorporación a los mismos y la adaptación por parte de los "*bookmakers*" ante circunstancias que generen resultados más volátiles en su predicción.

2.2. Comparativa mercados financieros - mercados de apuestas

Los mercados de apuestas poseen mecanismos económicos muy similares a los mercados financieros convencionales. De hecho, las apuestas podrían entenderse como contratos de opción o derivados de crédito, en los que el valor de los mismos vendría definido por el resultado de un futuro incierto (Lyocsa y Fedorko, 2016). Además, se caracterizan por tener muchos participantes en el mercado que generan un flujo constante de operaciones, y también por la existencia de información pública sobre estos "contratos", que permite a los participantes (apostadores) especular sobre el resultado futuro de su contrato.

Los mercados de apuestas pueden entenderse en los mismos términos que los mercados financieros, de hecho. El jugador puede considerar como más probable que en un partido de tenis gane el que está mejor posicionado en el ranking ATP. Podría apostar a que va a ganar el partido, el primer set, o que va a ganar más de 12 juegos, por ejemplo. Hay muchísimas combinaciones de sucesos que las casas de apuestas ofertan a un precio X, sobre el que inviertes la cantidad de dinero que desees, lo que llegado al extremo podría llegar a entenderse como la compra de participaciones de una empresa.

El apostante puede pensar que un suceso va a darse, y "comprar la apuesta". Cuenta con información pública que puede analizar para, ante una situación de incertidumbre, obtener un beneficio. Esta información, sin embargo, presenta diferencias también entre ambos tipos de mercados. Al fin y al cabo, la CNMV española tiene una fuerte presencia fiscalizadora de la actividad en las bolsas de valores, y cualquier tipo de información privilegiada tiene menos posibilidades de ser utilizada para la obtención de un beneficio.

No obstante, en el mundo deportivo se mueve mucha información por parte de inversores expertos, y tanto agentes como medios de comunicación, analistas e incluso los propios jugadores pueden utilizar información privilegiada sobre la condición de los jugadores, la química de los equipos y sus aspiraciones, etc. (Bernal, 2020), sin que exista una entidad supervisora similar a la CNMV en los mercados financieros.

La diferencia sustancial entre ambos mercados es que en los financieros se negocia sobre precios, mientras que en los mercados de apuestas se negocia (aunque manejando precios también) con probabilidades de que se produzca un resultado. Diferente sería hablar de la formación de precios en los mercados de apuestas conocidos como "*Betting Exchange*", cuyo funcionamiento hace que, para crearse un precio, se contraponga la demanda de apuesta hacia cada uno de los resultados en términos de volumen.

3. SESGO DEL FAVORITO

A nivel informacional, como hemos visto, existen sesgos poblacionales que las casas de apuestas emplean en su favor, por diversas razones. En este trabajo nos centraremos en el Sesgo del Favorito, más conocido por su nombre en inglés "*Favourite-Longshot Bias"* (Franck et. al. 2010). Esta tendencia poblacional es muy conocida en los partidos de tenis, cuyo resultado va en favor siempre de uno de los dos jugadores y no cabe el empate.

Este se debe a la propensión al riesgo de los *"small bettors"*, que generalmente son los agentes desinformados, de modo que las casas de apuestas se aprovechan de esta sobreestimación de la capacidad del menos favorito, y al existir la posibilidad de que pierdan mucho dinero si infravaloran al "*longshot*", ofrecen precios mayores para los

favoritos de los que, por mera información, serían los fundamentales o de equilibrio (Lahvička, 2014).

La implicación de esto a la hora de estudiar los precios es que las casas de apuestas o "*bookmakers*" van a fijar precios superiores a los que estiman matemáticamente para el jugador con mayores probabilidades de ganar el partido, infravalorando las probabilidades de los resultados más probables.

Los estudios muestran cómo esta variación de las casas de apuestas se acentúa principalmente en situaciones de incertidumbre, y se da por diferentes causas. En primer lugar, cuando el volumen de apuestas a favor de uno de los dos jugadores es muy alto, los "*bookmakers*" deben intentar compensar el efecto de dicho volumen. En este sentido, lo que hacen es otorgar probabilidades sesgadas, aumentando los precios de las apuestas con mayor volumen, aumentando también las probabilidades implícitas del suceso contrario.

Conociendo que los apostantes menos formados tienden a sobreestimar las posibilidades del menos favorito, bajan las cuotas de los no favoritos, por lo que dan a entender que las probabilidades implícitas de que ganen son mayores de lo que realmente son. Así, en caso de que las asimetrías informativas sean grandes y finalmente se dé el resultado menos esperado, tendrán que desembolsar una cantidad menor de dinero (Elizalde, 2015).

Con esto, consiguen que menos gente apueste al favorito (que ellos sí han calculado que tiene más probabilidades de ganar), y bajan los precios de los menos favoritos ("*longshots*") para que, si se da la sorpresa, no tengan que pagar tanto dinero. Así se cubren de cualquiera de las dos formas.

Como vemos, la posible existencia de agentes mejor informados y las situaciones de incertidumbre hacen los mercados mucho más volátiles, y, por tanto, los "bookmakers" deben cubrirse ante estas situaciones. Para ello existe el margen, que es la porción pagada por el apostante que va a destinarse a cubrir asimetrías informativas, costes de transacción y beneficios del empresario, en este caso la casa de apuestas.

4. EL MARGEN

A la hora de fijar los precios, las casas de apuestas contraponen las probabilidades de que se dé un suceso con las probabilidades de que no se dé, siendo el precio de cada apuesta un juego de suma 0. No obstante, la cantidad total que nos daría el sumatorio de dividir 1 entre el precio de la apuesta a favor y en contra de un suceso sería superior a 1. Esta cantidad que sobrepasa la unidad se denomina margen, y se trata de la cantidad que la casa de apuestas necesita para cubrir posibles pérdidas derivadas de resultados imprevisibles y obtener un beneficio del juego.

Además de funcionar como margen de beneficio, el margen ("*overround*") cubre también los costes que supone la actividad de las apuestas. Entre ellos estaría el coste de mano de obra, los costes de transacción, informáticos y recursos mentales empleados en cada apuesta. En ese sentido, el "*overround*" funcionaría de forma similar a la horquilla de los mercados financieros dirigidos por precios (Barrutiabengoa et. al. 2022).

Por tanto, el corredor de apuestas en primer lugar calcularía las probabilidades que tiene un suceso de darse (ganar un partido, hacer X faltas, ganar X juegos…). Asignaría en función de la información existente en cada momento una cuota para cada uno de los sucesos, y calcularía el dinero extra que requiere para realizar la operación y sacar un beneficio, de donde sacará el margen.

Este margen sirve realmente para combatir las ineficiencias del mercado. Cabe la posibilidad de que se den asimetrías informativas y que los apostadores experimentados las utilicen en su favor, o que un resultado no sea el calculado como más probable, por lo que las casas de apuestas deben cobrar esta comisión, por la provisión de liquidez, la asunción de riesgos y la obtención de ganancias.

Sin embargo, a esto hay que añadir que las casas de apuestas tienen en cuenta otros factores además de la información existente sobre un activo o apuesta. Como hemos comentado, los sesgos también influyen en este margen, ya que la existencia de apostantes irracionales permite que las casas de apuestas manipulen levemente los precios para lucrarse. No obstante, los inversores formados pueden actuar a la inversa y ver cuándo una casa de apuestas está sobrevalorando o infravalorando un activo en base a la demanda o los sesgos, aprovechando para vender o comprar y obtener mayor beneficio (Lyócsa y Fedorko, 2016).

La lógica dice, por tanto, que, ante una mayor falta de información por parte de las casas de apuestas, o un periodo de mayor incertidumbre, los márgenes que establecen aumentarán, ya que tienen que cubrirse en mayor medida de las asimetrías. Además, utilizarán este aumento teniendo en cuenta los sesgos mencionados, por lo que la incertidumbre debería hacer que, según el "*Favourite-Longshot Bias*", no sólo los precios en general aumenten, sino que los de los jugadores con mayor probabilidad asignada de ganar lo hagan también en mayor medida.

Por otro lado, el aumento de la competencia producido durante las últimas décadas en el mundo de las apuestas deportivas ha provocado que las casas de apuestas bajen sus márgenes, para captar una mayor cuota de mercado.

Por ello, este trabajo estudia los efectos producidos por la situación del Covid-19 en los márgenes de las apuestas. En esta situación de pandemia, se produjo un parón de 174 días en los partidos de competiciones oficiales de ATP (Asociación de Tenistas Profesionales), que tuvo comienzo el 1 de marzo de 2022, en el que se jugó el último partido, y se reanudó el 22 de agosto del mismo año. A raíz de ese momento, la incertidumbre se convirtió en una parte importante de los partidos, ya que en cualquier partido un deportista podía entrar en protocolo, no participar o tener peores condiciones de salud por estar recuperándose del Covid-19. Además, la falta de público también influye en los deportes (Flurin et. al., 2020), ya que hay jugadores a los que afecta más la presión o el apoyo del público, y esto ha generado una incertidumbre enorme a la hora de analizar los mercados deportivos (Fischer y Haucap, 2020).

Es por ello que trabajamos la horquilla de precios u "*overround*" establecida por las casas de apuestas como potencial reflejo de la incertidumbre que la situación generaba. Al fin y al cabo, como hemos explicado, los precios son fijados en función de muchas circunstancias, y la falta de información a la hora de predecir probabilidades hace que las casas de apuestas tengan que trabajar con horquillas más o menos amplias, para paliar los efectos de dicha incertidumbre e, incluso, tratar de beneficiarse de ellos.

5. BASE DE DATOS

El estudio realizado recoge datos de 6267 partidos disputados en las temporadas 2019, 2020 y 2021, de la ATP. Se ha realizado una limpieza de los partidos sobre los que

no existía información relevante, por faltar variables de interés que podrían impedir un correcto análisis estadístico.

Así, ha quedado una muestra con un total de 2560 partidos disputados en 2019 (tras una limpieza de 50 partidos jugados ese año a los que les faltaban datos), 1247 partidos en 2020 (tras eliminar 20 partidos de ese año) y 2460 en 2021 (quitando 29). La consecuencia lógica del Covid-19 fue una disminución a prácticamente la mitad de los partidos en 2020, debido a los protocolos, confinamientos y medidas restrictivas que impidieron el desarrollo común de la actividad deportiva. De hecho, el confinamiento generalizado llevó a una suspensión de las competiciones desde inicios de marzo de 2020 hasta finales de agosto del mismo año.

Los datos recogidos provienen de la página http://www.tennis-data.co.uk/, que contiene una de las bases de datos más robusta sobre cuotas de tenis a nivel mundial. En el archivo obtenido se incluyen los nombres de los jugadores, tipo de torneo, ronda, superficie de la pista, fecha, ranking de cada jugador y sus puntos, ganador y perdedor, cuotas pagadas a la victoria de cada uno al empezar el partido, etc. La misma página ha sido empleada por autores como Barrutiabengoa et. al. (2022) o De Carlos (2021), para emplear datos sobre precios de apuestas deportivas de diferentes deportes (tenis el primero, fútbol el segundo).

Esta información de las cuotas nos servirá para calcular el mencionado margen, y lo realizaremos en base a las cuotas de las casas de apuestas Bet365 y Pinnacle Sports, dos de las mayores firmas en este ámbito que, además, presentan estrategias de formación de precios diferentes. La primera basada en información y la segunda basada en volúmenes (Barrutiabengoa et. al., 2022) También estudiaremos las horquillas establecidas en promedio por las principales casas de apuestas, para ver un reflejo del funcionamiento del mercado de las apuestas a nivel global.

5.1. Diferencias en las ganancias (o pérdidas) durante el Covid-19 por el Sesgo del Favorito

Para contrastar la existencia de Sesgo del Favorito, se va a observar cuál sería la pérdida o beneficio que obtiene un apostante al apostar 100 euros en todos y cada uno de los partidos jugados durante los 3 años a que va a ganar el tenista al que la casa de apuestas

ha otorgado mayores probabilidades implícitas de victoria. En segundo lugar, haremos lo mismo, pero apostando a que gana el tenista con menos probabilidades implícitas. Posteriormente compararemos ambos resultados.

De este modo, vamos a ver la evolución con el Covid-19 de los resultados promedio que obtendría este jugador hipotético. Como hemos comentado, el sesgo del favorito incide en mayor proporción en épocas de mayor incertidumbre. Las casas de apuestas otorgan mayores probabilidades de victoria al "longshot" de las que realmente tiene si se hace un cálculo estadístico ajustado a todas las circunstancias que rodean el partido.

Hemos creado una variable llamada "BºFX"[4], con la que representamos el beneficio o pérdida por partido si apostásemos al favorito para Bet365 (el que menor cuota tiene según esta casa), y hemos hecho lo propio con Pinnacle Sports, para el que hemos generado la variable "BºFY". Con ello, vemos la evolución a lo largo de los años 2019, 2020 y 2021 del resultado al llevar a cabo esta estrategia de apostar siempre al favorito.

Por otro lado, hemos seguido simultáneamente una estrategia de apostar 100 euros en todos los partidos al jugador con menos probabilidades de ganar asignadas por parte de las casas de apuestas. Así, creamos "BºLX" y "BºLY", con las que representamos los resultados si apostásemos en cada partido al "longshot".

Así, el resultado medio que nos dé en cada casa de apuestas y en cada año, lo entenderemos como porcentaje de pérdida (ya que apostamos 100 euros por partido), y nos va a mostrar si realmente apostando siempre al menos probable vamos a tener más pérdidas por partido, y si estas pérdidas aumentan en periodos en los que existe mayor incertidumbre, por los ajustes en el precio de las casas de apuestas. Como ya hemos comentado anteriormente, esto supondría que el intermediario se cubre de posibles resultados adversos aumentando las probabilidades del "longshot" y bajando el precio de éste para no tener que desembolsar tanto dinero ante situaciones inesperadas.

[4] En este trabajo, consideraremos la X para la casa de apuestas Bet365, y la Y para Pinnacle Sports, con la intención de abreviar y facilitar la lectura de las variables.

TABLA 1: Estadísticos de grupo

	AgTotX	N	Media	Desv. típ.	Error típ. de la media
BºTFX	1	6267	-7,415	67,757	,856
	0	6267	-9,645	168,292	2,126

TABLA 2: Estadísticos de grupo

	AgTotY	N	Media	Desv. típ.	Error típ. de la media
BºTFY	1	6267	-2,816	69,769	,881
	0	6267	-4,870	160,788	2,031

Las tablas 1 y 2 recogen los estadísticos descriptivos principales de la variable "beneficios" para Bet365 y PS, diferenciando entre la estrategia de apostar al favorito y la de apostar al "*longshot*". Fuente: Elaboración propia.

En las tablas 1 y 2, las agrupaciones realizadas han sido de "1" para la estrategia de apostar por el favorito, y "0" para la estrategia contraria. Como vemos, para Bet365 (X) nos saldría que las pérdidas por partido si apostásemos siempre al favorito serían de un 7,42%, con una desviación típica de 67.76. En caso de apostar al "*longshot*", las pérdidas aumentan a un 9,65%, con una desviación muy superior. Para Pinnacle Sports (Y), las pérdidas medias son de un 2,82% al apostar al favorito, y de 4,87% al apostar por el que menos probabilidades tiene asignadas de victoria.

Esta diferencia entre ambos "*bookmakers*" se debe en gran medida a los márgenes que manejan, siendo muy inferiores en PS, ya que tiene una estructura y funcionamiento diferentes a los de Bet365, como señalan Barrutiabengoa et al. (2022), que muestra que Bet365 se basa en información para la fijación de precios, mientras que PS se basaría en volumen de apuesta en uno u otro sentido. No obstante, la diferencia que nos interesa es la que se realiza por año. De esta forma, la tabla a analizar es la siguiente:

TABLA 3: Diferencia media por año entre estrategias (Bet365)

	Bº Apostar Fav.	**Bº Apostar Long.**	**DIFERENCIA**
TOTAL	-7.42%	-9.65%	2.23%
2019	-8.49%	-10.81%	2.32%
2020	-6.62%	-10.50%	3.87%
2021	-6.70%	-8.01%	1.31%

En la tabla 3, observamos que la mayor diferencia entre las pérdidas en una u otra estrategia se produce en 2020, año en que se produjo la pandemia del Covid-19, por lo que a priori podríamos confiar en la existencia del "*Favourite-Longshot Bias*", y su

influencia en tiempos de incertidumbre sobre las mecánicas de las casas de apuestas. De igual manera, al realizar el estudio con Pinnacle Sports, nos saldría que las diferencias entre estrategias cada año se dan de la siguiente forma:

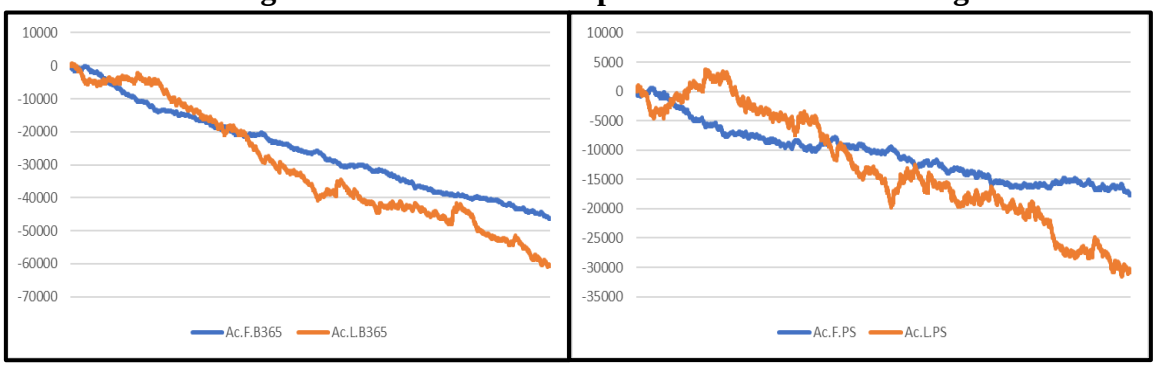
TABLA 4: Diferencia media por año entre estrategias (PS)

	Bº Apostar Fav.	Bº Apostar Long.	DIFERENCIA
TOTAL	-2.82%	-4.87%	2.05%
2019	-3.51%	-3.69%	0.18%
2020	-2.27%	-5.42%	3.15%
2021	-2.37%	-5.82%	3.45%

A partir de la tabla 4 podemos observar cómo PS también podría decirse que defiende una política de aumentar el Sesgo del Favorito en tiempos de incertidumbre, ya que la diferencia entre ambas estrategias es superior en 2020 a 2019, si bien en este caso en 2021 sigue aumentando la diferencia (este año también se ha visto muy afectado por la incertidumbre).

Estos datos dan que pensar que, llevando tanto una estrategia conservadora como una más arriesgada, la casa de apuestas calcula las probabilidades con la intención de obtener un beneficio siempre. Por eso la teoría nos lleva a que nunca se puede tener una ganancia a largo plazo en las apuestas por parte del apostante promedio. Los "*bookmakers*" tienen en cuenta no solo toda la información pública, sino también las conductas de los apostantes, y solo es posible lucrarse de manera continuada en este tipo de juegos si se posee información privilegiada, muy difícilmente alcanzable para la gran mayoría del público. Este gráfico nos mostraría la evolución económica del apostante hipotético al apostar según ambas estrategias para cada casa de apuestas:

Figura 1: Evolución de las pérdidas en ambas estrategias



Los gráficos muestran la evolución de las ganancias o pérdidas del apostante al seguir la estrategia de apuesta al favorito (Azul) o al *longshot* (Naranja). La de la izquierda es la obtenida para Bet365 y la de la derecha, para PS.

Comentado este punto, el estudio de las medias a lo largo del tiempo requiere de un análisis más exhaustivo, ya que tendremos que conocer si, ante esta comparación, existe una significatividad del efecto del Sesgo del Favorito sobre los márgenes establecidos por las casas de apuestas. Así, tendremos que realizar la prueba T de Student, para ver si se asumen varianzas iguales y, posteriormente, si las diferencias de medias que hemos observado son estadísticamente significativas.

TABLA 5: Prueba T para muestras independientes

	Prueba Levene		Prueba T-Student		
	F	Sig.	t-Student	Sig.	Dif. Medias
F/L.TotX	1839,399	0,000	0,973	0,331	2,229
F/L.TotY	2884,027	0,000	0,928	0,354	2,054
F/L.2019X	1342,540	0,000	0,739	0,460	2,315
F/L.2019Y	1297,900	0,000	0,053	0,957	0,181
F/L.2020X	240,998	0,000	0,662	0,508	3,874
F/L.2020Y	437,609	0,000	0,592	0,554	3,148
F/L.2021X	671,215	0,000	0,341	0,733	1,306
F/L.2021Y	1235,294	0,000	0,991	0,322	3,450

La tabla nos muestra las pruebas realizadas para analizar la diferencia de medias para las pérdidas obtenidas al apostar siempre al favorito (F) o al "longshot" (L). Fuente: Elaboración propia.

Como podemos ver en la tabla, al realizar comparaciones de medias entre la estrategia que apuesta siempre a favor del jugador con más probabilidades implícitas asignadas y la que apuesta siempre al "*longshot*", nos sale tanto en el cómputo total ("F/L.TotX" y "F/L.TotY") como en los análisis de cada año que no es estadísticamente significativa esta diferencia de ganancias. Pese a que habíamos visto cómo realmente aumentaba la diferencia de medias en los años del Covid-19, al plantear la hipótesis nula "se asume igualdad entre las medias", y sacar el p-valor, a un nivel de significación del 5% no tenemos evidencia estadística suficiente como para rechazarla, tras haber realizado la Prueba de Levene y haber rechazado la hipótesis de existencia de igualdad de varianzas.

Por tanto, resulta interesante observar las diferencias de pérdidas entre las dos estrategias y en los años en que se realiza la distinción, pero no sería correcto atribuir dichas diferencias de medias a la existencia de un sesgo del favorito, ya que los datos extraídos no son estadísticamente significativos, y sería un error afirmar dicha hipótesis.

5.2. Análisis del margen

En este estudio, la variable esencial sería el margen u horquilla que fijan las casas de apuestas para cubrir gastos, protegerse de resultados inesperados y generar un beneficio. Al fin y al cabo, las casas de apuestas estudian las probabilidades de un evento de darse, calculan teniendo en cuenta todos los factores que afectan al partido esa probabilidad, y generan una cuota haciendo la inversa de la probabilidad calculada. La suma entre dicho margen, la probabilidad de darse un suceso (A) y la probabilidad de darse el suceso contrario (B) deberían dar como resultado la unidad:

$$1 = \frac{1}{cuota\ asignada\ al\ suceso\ A} + \frac{1}{cuota\ asignada\ al\ suceso\ B} - Margen \qquad (1)$$

De este modo, un partido en el que la victoria del jugador A se paga a 1,36 euros por euro apostado, y la derrota de dicho jugador (o victoria del jugador B) se paga a 3,00 euros, la fórmula para calcular el margen sería la siguiente:

$$Margen = \frac{1}{1{,}36} + \frac{1}{3{,}00} - 1 \qquad (2)$$

El margen que se lleva la casa de apuestas de este cálculo es 0.0686 euros por euro apostado, que vendría a ser lo mismo que un 6,86% en términos relativos. Por tanto, la probabilidad (en porcentaje) que la casa de apuestas habría asignado a la victoria del jugador A (suceso a), teniendo en cuenta que, como hemos visto, las casas de apuestas desvían sus precios en función de la demanda y los sesgos de los inversores ruidosos *(noisy bettors)*. En este caso, sería de $\left(\frac{1}{1{,}36}\right) x100 = 73{,}53\%$, y la de victoria del jugador B, de $\left(\frac{1}{3}\right) x100 = 33.33\%$. En el tenis no cabe el empate, por lo que con la asignación de dichas dos probabilidades resultaría suficiente para el cálculo de las cuotas.

De este modo, suponiendo que la casa de apuestas tiene en cuenta toda la información relevante y disponible en el mercado, y también la existencia de sesgos de los que puede beneficiarse, por lo que incorpora toda esta información a sus precios, generando un margen de beneficio y coste con el que cubrirse, por otro lado, de las asimetrías informativas que puedan darse. En diferentes casas de apuestas, como veremos, el margen puede ser diferente para un mismo partido, debido a mecánicas

similares a las que existen en cualquier sector económico (barreras, cuota de mercado, fidelidad de los usuarios, etc.).

La variable de estudio a emplear va a ser la "*dummy*" creada para diferenciar entre la época "no-Covid" y la época "Covid", que sería la parte de 2020 posterior al parón que se hizo debido al confinamiento. Hemos tenido en cuenta esta fecha para el estudio de las horquillas, ya que para el ejercicio comentado sobre las pérdidas y el Sesgo del Favorito, lo hemos realizado por años (2019, 2020 y 2021). A esta variable independiente la hemos llamado "DCovid", y toma el valor 0 para los partidos jugados antes del parón realizado el 1 de marzo de 2020 y en 2021, y el valor 1 para los partidos jugados tras la reanudación, el 22 de agosto del mismo año, hasta diciembre de 2020 (comprendiendo el periodo total del 1 de enero de 2019 al 21 de noviembre de 2021).

De esta forma, aislamos el periodo en el que las restricciones por la pandemia fueron más duras, y las casas de apuestas aún estaban en proceso de adaptación a la nueva situación. Al fin y al cabo, estas circunstancias hicieron que la competencia cambiase, y la fijación de precios fuera diferente. Un análisis aislado de la variable dependiente a lo largo de este periodo de tiempo puede resultar útil e interesante, pero de poco vale interpretar los resultados que nos muestra su estudio si no realizamos simultáneamente un estudio de las variables de control que también pueden afectar al margen, independientemente de la "época Covid".

TABLA 6: Estadísticos descriptivos de las variables de control

Variable	Media	Mediana	Máximo	Mínimo	Desv. Típica
DCovid	0,11	1,00	1,00	0,00	0,32
DClay	0,30	0,00	1,00	0,00	0,46
DGrass	0,10	0,00	1,00	0,00	0,30
DATP500	0,15	0,00	1,00	0,00	0,35
DM1000	0,20	0,00	1,00	0,00	0,40
DSlam	0,22	0,00	1,00	0,00	0,42
D2ndRd	0,29	0,00	1,00	0,00	0,45
D3rdRd	0,07	0,00	1,00	0,00	0,26
D4thRd	0,02	0,00	1,00	0,00	0,14
DQF	0,10	0,00	1,00	0,00	0,30
DSF	0,05	0,00	1,00	0,00	0,22
DF	0,03	0,00	1,00	0,00	0,16
RKD	3,55	3,66	7,65	0,00	1,18
TREND	1,98	2,00	3,00	1,00	0,90

Así, tendremos que tener en cuenta por ejemplo la superficie en la que se juegan los partidos. La evidencia empírica ha mostrado que partidos jugados en hierba, tierra o superficie dura han tendido a influir en las decisiones de las casas de apuestas a la hora de fijar sus márgenes (Barrutiabengoa et. al., 2022). Esto se debe, en gran parte, a la existencia de ciertos jugadores más especializados en superficies como tierra y hierba. Al fin y al cabo, la superficie dura es la más barata de producir, y en la que más partidos se juegan a lo largo del año. Es mucho más fácil que haya una sorpresa en un partido en pista dura que en uno en tierra batida o hierba, por lo que las casas de apuestas tratan de cubrirse ante esta incertidumbre aumentando las horquillas, y con ellas los precios.

Por ello, se crean las *dummies* "DGrass" (hierba) y "DClay" (tierra batida), omitiendo "DHard" para la regresión lineal. En caso de que el partido sea jugado en hierba o tierra batida, tomará el valor 1 en la *dummy* correspondiente.

Por otro lado, tendríamos el tipo de torneo, en el que tenemos las *dummies* "DATP500", "DM1000" y "DSlam". DATP500 otorga 1 a los partidos disputados en la categoría ATP 500, y 0 al resto de partidos de otras categorías. Pasaría lo mismo con DM1000, para los partidos de Masters 1000, y DSlam para los torneos Grand Slam. En este caso, eliminaríamos para la regresión la *dummy* "DATP250", que tomaría valor 1 en los partidos de las categorías inferiores a las mencionadas.

Este conjunto de *dummies* lo hemos establecido como variables de control porque, de nuevo, la horquilla calculada en el precio por las casas de apuestas se ve afectada por tratarse de un tipo de categoría u otro. Al fin y al cabo, sobre los partidos más mediáticos (los Grand Slam, y posteriormente Masters 1000, ATP500 y ATP250) va a haber una cantidad de información disponible superior, por lo que resultaría más complicado que los inversores formados se aprovechasen de asimetrías informativas para conseguir un beneficio a costa de la casa de apuestas. Así, en partidos de categorías inferiores, los "*bookmakers*" tienen que aumentar las horquillas en el precio, ya que es posible que, al no cubrirse la información con tanta intensidad, haya más sorpresas que jueguen en contra de las probabilidades asignadas por el intermediario.

De la misma forma, la ronda que se esté disputando en el torneo correspondiente adquiere relevancia en el mismo sentido, ya que los partidos de finales o semifinales van a ser mucho más relevantes mediáticamente que los de primera o segunda ronda, haciendo que las horquillas, en teoría, bajen. No sólo las condiciones del partido son más conocidas,

sino que también los jugadores que tienden a llegar a últimas rondas suelen estar más en el radar de los apostantes, existiendo una mayor cantidad de información pública influyente en las probabilidades de un suceso u otro.

Para tener en cuenta estas circunstancias, creamos las *dummies* “D2ndRd”, “D3rdRd”, “D4thRd”, “DQF”, “DSF” y “DF”, para los partidos de segunda, tercera y cuarta ronda, cuartos de final, semifinales y final, respectivamente. Una vez más, eliminaríamos para la regresión la *dummy* “D1stRd”, correspondiente a los partidos que adquieren un valor 1 si son de primera ronda, y 0 si son de otra ronda.

Habrá que tener en cuenta también la variable “TREND”, que otorga un valor de 1 a los partidos jugados en 2019, 2 a los jugados en 2020 y 3 a los jugados en 2021. Al estudiar esta variable, lo que vamos a tener en cuenta es el efecto del paso del tiempo sobre el margen (Lyocsa y Fedorko, 2016). Como comentábamos al inicio del trabajo, el desarrollo de las plataformas online ha permitido un auge exponencial en la demanda de los juegos de azar y apuestas deportivas. La consecuencia lógica de este aumento de la demanda ha sido que la oferta crezca, por lo que las casas de apuestas tratan de competir por ofrecer los mejores servicios al mejor precio. Ello hace que, muchas veces, esta horquilla de la que estamos hablando, disminuya, ya que incluye también los beneficios por apuesta que pretende obtener el intermediario. En un mercado tan masivo como el de las apuestas deportivas, muchas casas de apuestas sacrificarán en cierta medida sus precios para obtener una cuota de mercado mayor, y tratar de fidelizar a sus clientes al competir en precios por encima de otros factores. En nuestro caso, el parón producido por el Covid-19 pudo hacer que algunas casas de apuestas iniciaran una estrategia de competencia en precios para no perder la clientela anterior.

Por último, también hemos calculado la diferencia de ranking (RKD) existente entre los dos adversarios en cada partido, haciendo el logaritmo neperiano del valor absoluto de la resta entre el ranking de ambos, para posteriormente crear nuevas variables que recojan el efecto multiplicador que puede tener la variable “Round” con respecto a la variable “RKD”. Esto se debe a que, al pasar de ronda, es más probable que los jugadores mejor posicionados en el ranking de la ATP puedan enfrentarse entre ellos. Al fin y al cabo, en primera ronda no pueden enfrentarse un número 1 y un número 2, por la forma en la que está configurado el torneo (para favorecer la meritocracia y el espectáculo en rondas avanzadas).

Por ello, creamos las variables “RKD*2nd”, “RKD*3rd” y así pasando por todas las rondas, hasta “RKD*F”, en función de la ronda que se está jugando. Como podemos observar, volvemos a eliminar la primera ronda para ser congruentes con el resto del estudio. Estas variables otorgarán el valor “1 x Ln (Diferencia de ranking)” a cada partido jugado en cierta ronda, y 0 a los partidos jugados en otra ronda diferente. Con ello, nos evitamos alterar los resultados por culpa de la configuración de los torneos, que restaría aleatoriedad en la muestra.

La cuestión a la que queremos dar respuesta e interpretación es si realmente existe una variación en la variable “Margen” a raíz del Covid-19 y el consecuente aumento de las asimetrías informativas. Trataremos de entender por qué se da ese cambio en caso de efectivamente haber una tendencia diferente durante el año 2020, que es cuando realmente se desencadenaron las medidas institucionales para paliar los contagios. De este modo, formularemos la siguiente hipótesis:

H_0: Los márgenes no han variado en época Covid-19

H_1: Los márgenes sí han variado en época Covid-19

Para empezar, los resultados sacados de la tabla 7 nos muestran las diferencias entre las medias de horquillas para el período previo al Covid y el posterior al confinamiento, tomando como fecha de referencia de ese parón el 1 de marzo de 2020, y retomándose la actividad el 20 de agosto del mismo año. Se han obtenido resultados de Bet365 (MgX), Pinnacle Sports (MgY) y la media de las principales casas de apuestas a nivel mundial (MgAvg). La variable *dummy* “DCovid” ha tomado el valor 0 en el periodo previo, y 1 en el posterior.

TABLA 7: Estadísticos de grupo

	DCovid	N	Media	Desviación típ.	Error típ. de la media
MgX	1	707	0,0627	0,0115	0,0004
	0	5560	0,0603	0,012	0,0002
MgY	1	707	0,0248	0,0038	0,0001
	0	5560	0,0252	0,0038	0,0001
MgAvg	1	707	0,0524	0,0042	0,0002
	0	5560	0,0534	0,0054	0,0001

Así, comprobamos que la media de horquillas del periodo No-Covid ha resultado, para Bet365, superior al periodo Covid, como se podría esperar de acuerdo a la lógica que hemos aportado al inicio del trabajo. Para PS, habría sido de media inferior en el periodo Covid, y para la media de las casas de apuestas, también inferior. A priori, podríamos entender que en estos dos últimos casos no ha supuesto un problema la situación de incertidumbre a la hora de fijar los precios, pero al tratarse de un análisis univariante, no podemos extraer conclusiones, ya que hay otros factores que afectan a la horquilla.

Posteriormente veremos los resultados del análisis multivariante, pero la tendencia de los precios debido al aumento de la competencia puede tener mucho que ver en dicha disminución de los márgenes. En un periodo tan inestable, en el que cualquier negocio podía enfrentarse al declive e incluso la ruina, las casas de apuestas podrían haber optado por bajar sus horquillas, generando menos beneficio por apuesta para ellas y adaptando su modelo de negocio hacia una mayor asequibilidad para el cliente, buscando mantener o incluso aumentar su cuota de mercado.

De la misma forma, la estructura que presenta PS a la hora de fijar sus precios está mucho menos guiada por aspectos informacionales (lo cual podría tener que ver con que no afecte tanto la incertidumbre). En este caso, es superior la importancia del volumen de demanda en las casas de apuestas, y no tanto la de la incertidumbre.

No obstante, como hemos hecho en el ejercicio hipotético sobre los beneficios o pérdidas, lo que nos interesa aquí es ver si el Covid-19 ha sido un factor estadísticamente significativo en la fijación de los márgenes. La tabla 8 condensa dicha información:

TABLA 8: Prueba de muestras independientes

		Prueba de Levene		Prueba T para la igualdad de medias		
		F	Sig.	t	Sig. (bilateral)	Diferencia de medias
MgX	Se han asumido varianzas iguales	,128	,720	5,105	,000	0,00244
MgY	Se han asumido varianzas iguales	1,508	,219	-2,473	,013	0,00015
MgAvg	Se han asumido varianzas iguales	,033	,857	-5,049	,000	0,00021

Empezamos estudiando la prueba de Levene para la igualdad de varianzas, que nos determina si se puede rechazar o no la hipótesis de que existe realmente igualdad de varianzas. En este caso vemos que, para las tres variables expuestas (MgX, MgY y MgAvg), el p-valor (0,720, 0,219 y 0,857 respectivamente) nos sale superior a 0,05 (trabajamos con una significación del 5%). Esto nos muestra que no podemos rechazar la hipótesis nula, por lo que no podemos desechar que haya una igualdad de varianzas. Sin embargo, al pasar a la prueba T de Student, observamos que los p-valores son inferiores en todos los casos (0,000, 0,013, 0,000). Por tanto, podríamos rechazar la hipótesis nula, que formulaba que había igualdad de medias entre el periodo pre-Covid y el periodo post-Covid (durante el Covid). De esta forma, concluimos con que sí hay una significatividad estadística entre la diferencia de medias en período Covid y no-Covid, y la evidencia empírica nos ha mostrado que las horquillas en promedio aumentarían en periodo Covid.

Una vez realizado este análisis, pasaremos a estudiar los posibles efectos del resto de variables *dummy* sobre la variable independiente "Margen".

Para empezar, el tipo de pista sobre el que se juega el partido resulta interesante para este análisis, ya que en numerosos estudios ha sido demostrado que las pistas de tierra batida o de hierba pueden generar una menor incertidumbre que las pistas duras. Esto se debería, teóricamente, a que hay mayor cantidad de especialistas en estos tipos de superficies, mientras que las pistas duras son las más jugadas, por lo que generalmente son más propensas a las sorpresas que las pistas en las que hay jugadores mucho más dominantes.

TABLAS 9 y 10: Estadísticos de grupo (Variables DClay y DGrass)

	DClay	N	Media	Desv. típ.	Error típ. de la media
MgX	1	1883	,06056	,01187	,00027
	0	4384	,06058	,01203	,00018
MgY	1	1883	,02503	,00439	,00010
	0	4384	,02516	,00353	,00005
MgAvg	1	1883	,05330	,00421	,00010
	0	4384	,05333	,00575	,00009

	DGrass	N	Media	Desv. típ.	Error típ. de la media
MgX	1	617	,05753	,01034	,00042
	0	5650	,06091	,01210	,00016
MgY	1	617	,02526	,00339	,00014
	0	5650	,02510	,00385	,00005
MgAvg	1	617	,05331	,00422	,00017
	0	5650	,05332	,00544	,00007

Como vemos en las tablas 9 y 10, tanto en los partidos de tierra batida como en los de hierba habría una diferencia en la media de horquillas para estas casas de apuestas. Por lo general, establecerían márgenes inferiores cuando se juega en estas superficies, ya que, como comentábamos, hay jugadores que están más especializados en estas pistas, haciendo que sea inferior el "factor sorpresa". Al realizar la prueba T, nos sale que no es representativo estadísticamente (salvo DGrass para Bet365), pero sí que se trata de un factor a tener en cuenta al realizar una regresión sobre la variable margen u horquilla.

Por otro lado, el tipo de torneo que se juega presenta también diferencias en las medias. De hecho, los partidos de Grand Slam tienen horquillas inferiores por lo general al resto de partidos. Esto sucedería en sentido contrario en los torneos de menor calado (salvo en los Masters 1000, en los que hay variedad de diferencias de horquillas en función de la casa de apuestas). Las pruebas T de Student realizadas aparte (Anexo II) nos muestran que sí hay una diferencia estadísticamente significativa entre los partidos de Grand Slam y otros, para las horquillas establecidas. Como hemos comentado, podría darse por lo mediáticos que son unos u otros y la cantidad de información disponible de la que podrían beneficiarse los inversores formados. Por ello, serán variables *dummy* a tener en cuenta a la hora de hacer la regresión.

TABLA 11: Estadísticos de grupo (Variable DSlam)

	DSlam	N	Media	Desv. típ.	Error típ. de la media
MgX	1	1391	,05195	,00728	,00020
	0	4876	,06304	,01192	,00017
MgY	1	1391	,02411	,00244	,00007
	0	4876	,02541	,00407	,00006
MgAvg	1	1391	,04955	,00420	,00011
	0	4876	,05440	,00513	,00007

En función de la ronda también hay diferencias significativas estadísticamente hablando. Al realizar el estudio en la variable DF, que diferencia los partidos de finales de los que no lo son, se ha observado que las diferencias de medias aumentan hasta un 0,00662%, 0,00325% y 0,00136% para Bet365, Pinnacle Sports y la media de las casas de apuestas respectivamente.

TABLA 12: Estadísticos de grupo (Variable DF)

	DF	N	Media	Desv. típ.	Error típ. de la media
MgX	1	159	,05412	,00702	,00056
	0	6108	,06074	,01203	,00015
MgY	1	159	,02195	,00324	,00026
	0	6108	,02520	,00379	,00005
MgAvg	1	159	,05200	,00398	,00032
	0	6108	,05336	,00536	,00007

Esto sigue los mismos criterios que las diferencias de torneo. En el Anexo adjuntaremos también estadísticos descriptivos sobre rondas anteriores, en los que se ha podido observar que, para semifinales, cuartos de final, cuartas y terceras rondas, hay unos márgenes inferiores a los de segunda, primera ronda y clasificatorias de nivel anterior. Para todos los análisis ha salido estadísticamente significativa esta diferencia, salvo para la media de las casas de apuestas en partidos de cuartos de final, en los que igualmente las medias mostraban superiores horquillas frente a otros encuentros.

Al fin y al cabo, conforme pasan las rondas tienden a participar los mejores jugadores, y no cabe tanto la sorpresa como en partidos de primeras rondas. Por ello también introducimos las variables multiplicadoras de la diferencia de ranking por la ronda que se está jugando. Al ser más probable que jugadores de ranking bajo jueguen con los de ranking alto en las primeras rondas, esta variable permitirá que no quede sesgado el estudio.

5.3. Regresión Lineal

Una vez realizado el análisis univariante, si queremos saber realmente cómo ha funcionado la horquilla a raíz del Covid-19, tenemos que hacer una puesta en común de todos los factores que pueden afectarle, para explicar el modelo en la mayor medida

posible. Para ello, hemos creado un modelo de regresión lineal con el que ver el efecto conjunto de cada variable sobre el margen.

El modelo utilizado, vendrá dado por la siguiente fórmula:

$$\text{Margen} = \beta_0 + \beta_1\ \text{DCovid} + \beta_2\ \text{DClay} + \beta_3\ \text{DGrass} + \beta_4\ \text{DATP500} + \beta_5\ \text{DM1000} + \beta_6\ \text{DSlam} + \beta_7\ \text{D2ndRd} + \beta_8\ \text{D3rdRd} + \beta_9\ \text{D4thRd} + \beta_{10}\ \text{DQF} + \beta_{11}\ \text{DSF} + \beta_{12}\ \text{DF} + \beta_{13}\ \text{RKD2ND} + \beta_{14}\ \text{RKD3RD} + \beta_{15}\ \text{RKD4TH} + \beta_{16}\ \text{RKDQF} + \beta_{17}\ \text{RKDSF} + \beta_{18}\ \text{RKDF} + \beta_{19}\ \text{RKD} + \beta_{20}\ \text{TREND} \quad (4)$$

De este modo, y como comentábamos al principio del trabajo, el modelo que trata de explicar la variable dependiente "Margen", quedaría determinado por las variables de control expuestas en la fórmula. Esto nos permite en cierto modo aislar el efecto de cada una de ellas sobre la horquilla, y así conocer las diferencias producidas por el efecto Covid-19, a través de la variable DCovid.

TABLA 13: Resumen del modelo

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,641[a]	,411	,409	,0029

Modelo	Coeficientes no estandarizados		Coef. tipificados	t	Sig.
	B	Error típ.	Beta		
(Constante)	,027	,000		137,356	,000
DCovid	-7,554E-005	,000	-,006	-,634	,526
DClay	,000	,000	-,017	-1,698	,089
DGrass	-9,018E-006	,000	-,001	-,069	,945
DATP500	,000	,000	,016	1,565	,118
DM1000	-,002	,000	-,260	-23,560	,000
DSlam	-,002	,000	-,268	-23,678	,000
D2ndRd	-,002	,000	-,287	-7,708	,000
D3rdRd	-,003	,001	-,191	-5,474	,000
D4thRd	-,004	,001	-,149	-4,361	,000
DQF	-,004	,000	-,347	-10,499	,000
DSF	-,006	,000	-,326	-12,640	,000
DF	-,005	,001	-,189	-8,628	,000
RKD2ND	-6,941E-005	,000	-,033	-,858	,391
RKD3RD	4,164E-005	,000	,010	,283	,777
RKD4TH	,000	,000	-,022	-,653	,514
RKDQF	,000	,000	-,113	-3,430	,001
RKDSF	7,109E-007	,000	,000	,005	,996

RKDF	,000	,000	-,052	-2,381	,017
RKD	8,846E-005	,000	,027	1,977	,048
TREND	,001	,000	,137	13,991	,000

a. Variable dependiente: MgPS

Los resultados para Pinnacle Sports, recogidos en la tabla 13, nos ofrecen un R cuadrado de 0,411. Al observar las variables de control introducidas, y empezando por las variables dummy de tipo de pista sobre el que se juega (DClay y DGrass), vemos que los partidos en tierra batida y hierba presentan un efecto muy pequeño sobre el margen, siendo curioso que ninguno ofrece un p-valor inferior a 0,05. DClay estaría mostrando la diferencia en el margen en partidos jugados sobre tierra batida respecto de partidos jugados en cualquier otro tipo de pista, y lo mismo ocurriría con DGrass, pero respecto a partidos jugados en hierba. No obstante, como decimos, no aportan un gran valor informacional a nivel estadístico, ya que ninguna de las dos es significativa, y apenas muestran diferencias en el margen.

Además, los diferentes tipos de torneos poseen, tras hacer el análisis, diferentes efectos sobre el margen. En primer lugar, Los partidos de Grand Slam (a través de la variable DSlam) y Masters1000 (DM1000) tienen un efecto negativo sobre la horquilla, y con significatividad estadística. Esto podría darse, como hemos comentado anteriormente, a que en estos partidos hay un foco mediático mucho mayor, y la información relativa a los torneos en cuestión es mucho más abundante, por lo que cabe menos la posibilidad de que apostantes se beneficien de las asimetrías informativas. Por su lado, los partidos de ATP500 tienen un efecto prácticamente nulo sobre el margen, sin ser estadísticamente significativo, por lo que no podríamos realizar afirmaciones al respecto, pero sí que podemos observar una tendencia a aumentar los márgenes en torneos menores (los ATP250 tendrán seguramente efectos positivos sobre el margen, al tener las demás *dummies* efectos negativos o nulos).

Las diferentes rondas en las que se juegan los partidos también ofrecen variedad de resultados. Las variables dummy empleadas para analizarlo son: “D2ndRd” (variable que diferencia los partidos de segunda ronda respecto de los demás), D3rdRd (lo mismo para los de tercera ronda), D4thRd (cuarta ronda), DQF (cuartos de final), DSF (semifinales) y DF (partidos de final). En este caso, los efectos desde la segunda ronda a la final resultarían negativos sobre el margen (por lo que podríamos deducir que en la primera ronda habría un efecto positivo), pero además resulta interesante que, siendo

todos estadísticamente significativos, disminuye la beta conforme avanzan las rondas. Esto podría entenderse en tanto en cuanto la información disponible sobre los jugadores y partidos que llegan a ulteriores rondas aumenta, y disminuyen las asimetrías informativas.

Dicho esto, hemos introducido también la variable TREND, que muestra la tendencia de las horquillas a lo largo del tiempo (debido principalmente al efecto competencia). Esta variable, que otorga un valor de 1 a los partidos jugados en 2019, 2 para los jugados en 2020 y 3 para los jugados en 2021, permite analizar la evolución de los márgenes a lo largo del tiempo, lo cual puede resultarnos interesante al hablar de los efectos de la variable Covid en el margen, pues vendría también influida por una posible tendencia ascendente o decreciente de las horquillas. Así, la influencia de TREND sobre la variable margen es positiva (0,001), y estadísticamente significativa, por lo que en teoría el margen debería aumentar para Pinnacle Sports a lo largo de los años.

Respecto a la variable Covid, que es la que más nos interesa, de los resultados extraemos que (a través de la *dummy* "DCovid") posee un efecto negativo sobre el margen. Esto significa que, en partidos de periodo Covid, la horquilla de los precios habría disminuido comparativamente, pero sin significatividad estadística. Este resultado, a priori, parece no encajar con la teoría que hemos expuesto, ya que el periodo Covid debería haber hecho aumentar las horquillas debido a la incertidumbre. Sin embargo, como ya hemos comentado, Pinnacle Sports funciona con un sistema de fijación de precios diferente al de Bet365, y la información sobre los partidos no tiene esa importancia de la que hablábamos, como sí la tendría, en este caso, el volumen de apuesta y la demanda de un resultado u otro. Esto podría explicar, viendo los resultados de nuestro modelo, cómo algunas variables que a priori son importantes a nivel informacional (DClay, DGrass, DCovid) no son significativas estadísticamente hablando.

Numerosos estudios muestran la existencia de casas de apuestas como Pinnacle que basan parte de su estrategia de fijación de precios en la demanda de uno u otro resultado (Lyocsa y Fedorko, 2016). Por ello, teniendo en cuenta todos los factores podemos intuir que las asimetrías informativas que en teoría deberían influir en las decisiones de los creadores de mercado, a veces no tienen tanto peso como la competitividad y el volumen de apuesta tienen. Pinnacle Sports es un ejemplo de ello, y por esta razón el Covid no habría tenido una influencia a priori coherente ni significativa sobre el margen que establece la casa de apuestas.

En segundo lugar, analizaremos el mismo modelo regresivo en Bet365, que sí que se ha demostrado que posee una estructura más determinada por la información y no tanto por el volumen. En el modelo de Bet365, toma especial importancia la variable tendencia, debido a la contraposición que podríamos encontrar entre el efecto del aumento de la competencia en el margen (disminución), frente al efecto sobre el mismo de la incertidumbre generada por la pandemia del Covid-19 (aumento).

TABLA 14: Resumen del modelo

Modelo	R	R cuadrado	R cuadrado corregida	Error típ. de la estimación
1	,646[a]	,417	,415	,0092

Modelo	Coeficientes no estandarizados		Coef. tipificados	t	Sig.
	B	Error típ.	Beta		
(Constante)	,081	,001		132,097	,000
DCovid	,004	,000	,112	11,366	,000
DClay	-,001	,000	-,027	-2,689	,007
DGrass	-,001	,000	-,028	-2,750	,006
DATP500	-,002	,000	-,062	-5,953	,000
DM1000	-,003	,000	-,113	-10,313	,000
DSlam	-,013	,000	-,458	-40,650	,000
D2ndRd	,002	,001	,067	1,815	,070
D3rdRd	-,008	,002	-,169	-4,851	,000
D4thRd	-,002	,003	-,020	-,593	,553
DQF	-,010	,001	-,244	-7,425	,000
DSF	-,011	,001	-,203	-7,923	,000
DF	-,013	,002	-,171	-7,848	,000
RKD2ND	-,001	,000	-,090	-2,371	,018
RKD3RD	,000	,000	,024	,703	,482
RKD4TH	-,001	,001	-,029	-,845	,398
RKDQF	,000	,000	-,019	-,577	,564
RKDSF	,000	,000	,017	,674	,500
RKDF	,000	,001	,013	,618	,536
RKD	-,001	,000	-,118	-8,587	,000
TREND	-,005	,000	-,354	-36,499	,000

a. Variable dependiente: MgB365

En este caso, las variables de control siguen una lógica similar a la que hemos expuesto para Pinnacle Sports. La tabla 14 muestra cómo las variables de tipo de pista presentan una influencia negativa y significativa, mientras que el tipo de ronda genera también una influencia negativa (salvo D2ndRd), con una progresión descendente, lo que

significa que el margen en teoría disminuiría en los torneos de mayor calado internacional. Esto podría explicarse porque la *dummy* de tipo de ronda excluida (D1stRd) posee un efecto positivo sobre el margen, al ser los partidos de primera ronda en los que menos información disponible hay. Cabe destacar de esto que la variable que recoge los partidos de cuarta ronda no sigue tanto esta progresión, pero puede entenderse por la escasez de partidos de este tipo (únicamente se disputan en las competiciones de Grand Slam) y la falta de significación estadística de la misma.

Las variables de tipo de torneo, por su parte, tienen una implicación negativa, y mucho más acentuada en los partidos de Grand Slam. Este aspecto también sería explicable según la lógica que hemos planteado al comienzo del trabajo, ya que los torneos más importantes generan una mayor mediatización y una exposición informativa mucho mayor.

Respecto a la variable tendencia, TREND posee una influencia negativa sobre el margen, con significatividad estadística. Esto tiene sentido, ya que, como comentábamos, la competencia se acentúa con el paso del tiempo, debido a la proliferación de plataformas de apuesta. Las casas de apuestas tienen por tanto que limitar sus beneficios y costes, para poder ser más competitivas que otras empresas del entorno.

Llegados a este punto, estudiamos el efecto de la variable DCovid en la horquilla. Podemos observar cómo produce un efecto positivo en la misma (0,004). A priori este resultado cuadraría con las expectativas, ya que el Covid-19 ha supuesto un incremento de la incertidumbre, y las casas de apuestas, bajo nuestra suposición, deberían incrementar los márgenes para paliar dicha falta de información. Además, el p-valor es 0,000 para DCovid, por lo que podemos considerar la afirmación anterior como estadísticamente significativa.

Como hemos visto, ante la hipótesis planteada anteriormente, para Bet365 el p-valor de la variable DCovid ha sido inferior a 0,05, lo que nos permite rechazar la hipótesis nula (hay igualdad de medias en periodo Covid y no-Covid), y afirmamos que en periodo Covid hay en promedio un aumento de las horquillas fijadas por las casas de apuestas. Para DCovid aplicada a Pinnacle Sports, por su lado, hay un p-valor superior a 0.05, lo que impide rechazar la hipótesis nula, no pudiendo afirmar nada sobre las diferencias de medias en periodo Covid.

Por tanto, en contraposición con Pinnacle Sports, Bet365 sí que mantiene una política de fijación de precios basada en la información existente sobre los partidos. Esto se puede ver reflejado en la significatividad de las principales *dummies* (salvo D2ndRd y D4thRd), que en parte inciden sobre la horquilla debido a su valor informacional. Por ello, observamos que el modelo de Bet365 queda mejor explicado por las variables incluidas (41,70%).

La evidencia empírica, por tanto, acompaña a las teorías formuladas por Fama y muchos otros autores, que trabajaron para mostrar las implicaciones que tiene la información sobre los mercados financieros. La falta de información genera incertidumbre, y cuando hablamos de márgenes tan pequeños sobre el precio, cualquier error de cálculo de probabilidades puede suponer una pérdida de dinero para los *bookmakers*. Es por ello que las casas de apuestas tratan de cuidarse ante periodos más inestables, pues de lo contrario las consecuencias serían fatales si no se diesen los resultados predichos. Para eso es para lo que aumentan las horquillas, cuando no controlan en la medida habitual la información disponible sobre un suceso.

6. CONCLUSIONES

Siendo el objetivo de la investigación realizada conocer las diferencias que han podido darse a raíz de la pandemia del Covid-19 sobre los precios de las casas de apuestas, los resultados nos han mostrado que la información y la incertidumbre juegan papeles fundamentales en los mercados financieros (y de igual manera al extrapolarse al mundo de las apuestas deportivas).

Si bien los datos no han respaldado la existencia de un Sesgo del Favorito en el periodo contemplado, hemos visto cómo las pérdidas medias del apostante aumentaban en situación de incertidumbre, lo que tendría sentido desde el punto de vista informacional. Al fin y al cabo, numerosos estudios han expuesto la estrategia de las casas de apuestas de sobrevalorar ciertos sucesos e infravalorar a conveniencia otros respecto de las verdaderas probabilidades implícitas de que ocurra dicho suceso.

Cabe poner en valor los resultados obtenidos de este ejercicio teórico sobre el Sesgo del Favorito. Siguiendo tanto una estrategia como la otra, el resultado era

claramente negativo para el apostante, lo que debería servir para entender que las apuestas deportivas no consisten en puro azar. Las casas de apuestas juegan con los precios en busca de maximizar beneficios, objetivo frontalmente opuesto al que tienen los apostantes. Los sistemas de fijación de precios que emplean son altamente eficientes, e incorporan la información existente al precio de las apuestas pendientes de una forma prácticamente instantánea, por lo que tratar de vencer a "la banca" a largo plazo en este ámbito parece prácticamente imposible. Hay muy pocos jugadores lo suficientemente informados y preparados para sacar rédito de las apuestas de una forma continuada, por lo que este trabajo anima a evitar las apuestas y los erróneamente denominados "juegos de azar".

Respecto al margen, hemos visto contrapuestos dos sistemas muy diferentes de fijación de precios por dos casas de apuestas que realmente compiten en el mismo mercado. Así, la evidencia científica (también respaldada por diferentes investigaciones), nos muestra cómo Bet365 funciona de acuerdo a un sistema de fijación de precios basado en la información existente sobre los partidos. Ante un partido, esta empresa recogería todos los datos públicos y relevantes que puedan influir en el potencial resultado del partido o suceso dentro del partido, e integraría esta información con la mayor inmediatez posible en los precios. Así, una vez establecido un margen que recoge sus beneficios esperados, consigue que los agentes externos no puedan beneficiarse a largo plazo del mercado, ya que los ajustes de precios se adelantan a las decisiones de los apostantes.

Además, hemos tenido en cuenta para este análisis la tendencia de los márgenes a lo largo del tiempo, que refleja el aspecto de la competencia existente entre las casas de apuestas, que bajan sus márgenes ante la proliferación de potenciales competidores para aumentar su cuota de mercado. Este efecto en el margen nos ha sido muy útil estudiarlo junto al efecto de la información, pues son dos de los pilares sobre los que se basa la fijación de precios en el mundo de las apuestas.

Pinnacle Sports, por su lado, se basa fundamentalmente en volúmenes de apuesta en uno u otro sentido, dándole mucha menor importancia a la información pública sobre el suceso expuesto. Por ello, hemos visto que los resultados tenían mucho sentido para esta casa de apuestas. Al estar hablando de variables relacionadas con la información e incertidumbre, los resultados sobre PS no han sido por lo general significativos estadísticamente, y sobre todo lo podemos ver en la variable DCovid, que era la que más nos interesaba.

Las variables de control nos han servido para explicar el modelo, ya que su valor informacional permitía distinguir márgenes muy diferentes ante ciertas situaciones u otras. Esto nos ha permitido aislar el efecto de la variable Covid-19 sobre el margen en la medida de lo posible, acotando el mismo al periodo de tiempo más influido por la pandemia y las consiguientes restricciones (de marzo de 2020 a diciembre del mismo año), y en el que las casas de apuestas tenían que aprender a adaptarse a la nueva situación.

En conclusión, la incertidumbre generada por la crisis del Covid-19 ha provocado un aumento de los márgenes por parte de Bet365, que ha tenido que adaptar sus precios para paliar la escasez de información pública sobre los partidos ATP. Esta investigación refuerza la Hipótesis de los Mercados Eficientes, y pone en especial valor el aspecto informacional en la fijación de precios. Las casas de apuestas deben ser precisas al establecer sus márgenes, y la incertidumbre forma parte del mercado, por lo que adaptar los precios a periodos en los que disminuye la información en juego resulta un gran reto al que los "*bookmakers*" deben enfrentarse.

7. REFERENCIAS

ABINZANO, I.; MUGA, L.; SANTAMARIA, R.: "*The role of small bettors in price formation in betting exchanges*". *Applied Economics*, 2020. Vol. 52, no 33, p. 3637-3650

BARRUTIABENGOA, J.M.; CORREDOR, P.; MUGA, L.: "*Does the betting industry price gender? Evidence from professional tennis*". *Journal of Sports Economics*. 2022. Vol. 23, no 7, p. 881-906.

BERNAL LUCAS, S.: "*Las apuestas deportivas como alternativa de inversión*". Trabajo de Fin de Grado, Facultad de Administración y Dirección de Empresas. Universidad Politécnica de Valencia. 2020.

CAIN, M.; LAW, D.; PEEL, D.: "*The favourite-longshot bias and market efficiency in UK football betting*". *Scottish Journal of Political Economy*. 2000. Vol. 47, no 1, p. 25-36.

DE CARLOS ALGARRA, I.: "*Análisis de la ventaja del equipo local en el fçútbol. Eficiencia de los mercados de apuestas deportivas. El caso de la primera división española*". Trabajo de Fin de Grado. Facultad de Ciencias Económicas y Empresariales. Universidad Pública de Navarra. 2021.

ELIZALDE LALLANA, A. "*Microestructura y formación de precios en mercados financieros y de apuestas deportivas*". 2015. Trabajo de Fin de Grado. Facultad de Ciencias Económicas y Empresariales. Universidad Pública de Navarra. 2021.

FAMA, E.: "*Efficient capital markets: A review of theory and empirical work*". *The journal of Finance.* 1970. Vol. 25, no 2, p. 383-417.

FAMA, E.F.; LAFFER, A.B.: "*Information and capital markets*". *Journal of business*. 1971. P. 289-298.

FISCHER, K.; HAUCAP, J.: "*Betting market efficiency in the presence of unfamiliar shocks: The case of ghost games during the COVID-19 pandemic*". 2020.

FRANCK, E.; VERBEEK, E.; NÜESCH, S. "*Prediction accuracy of different market structures—bookmakers versus a betting Exchange*". *International Journal of Forecasting*. 2010, vol. 26, no 3, p. 448-459.

LAHVIČKA, J.: "*What causes the favourite-longshot bias? Further evidence from tennis". Applied Economics Letters*, 2014, vol. 21, no 2, p. 90-92.

LYÓCSA, S.; FEDORKO, I.: "*What drives intermediation costs? A case of tennis betting market". Applied Economics*. 2016, vol. 48, no 22, p. 2037-2053.

MEIER, P.F.; FLEPP, R.; FRANCK, E.: "*Are sports betting markets semistrong efficient? Evidence from the COVID-19 pandemic*". International Journal of Sport Finance, 2021, vol. 16, no 3, p. 111-126.

RIAÑO, D.I.: "*Hipótesis de Mercados Eficientes y la Sabiduría de las Masas*". Trabajo de Fin de Grado. Universidad Pontificia de Comillas (2018).

SHLEIFER, A.: "*Innefecient markets: An introduction to behavioral finance"*. OUP Oxford, 2000.

VILLALBA VILÁ, D.: "*¿Son los inversores racionales?*". Bolsa de Madrid. 2003.

8. ANEXO

ANEXO I: Estadísticos descriptivos (beneficio por partido en 2019, 20 y 21 para Bet365 y PS).

Tabla a): Estadísticos de grupo

	Ag19B365	N	Media	Desv. típ.	Error típ. de la media
BºF19B365	1	2560	-8,490	68,333	1,351
	0	2560	-10,805	143,042	2,827

Tabla b): Estadísticos de grupo

	Ag20B365	N	Media	Desviación típ.	Error típ. de la media
BºF20B365	1	1247	-6,623	66,345	1,879
	0	1247	-10,497	195,683	5,541

Tabla c): Estadísticos de grupo

	Ag21B365	N	Media	Desviación típ.	Error típ. de la media
BºF21B365	1	2460	-6,698	67,874	1,369
	0	2460	-8,005	177,396	3,577

Tabla d): Estadísticos de grupo

	Ag19PS	N	Media	Desviación típ.	Error típ. de la media
BºF19PS	1	2560	-3,507	70,785	1,399
	0	2560	-3,688	156,207	3,087

Tabla e): Estadísticos de grupo

	Ag20PS	N	Media	Desviación típ.	Error típ. de la media
BºF20PS	1	1247	-2,269	68,365	1,936
	0	1247	-5,417	174,975	4,955

Tabla f): Estadísticos de grupo

	Ag21PS	N	Media	Desviación típ.	Error típ. de la media
BºF21PS	1	2460	-2,373	69,428	1,400
	0	2460	-5,822	158,014	3,186

ANEXO II: Comparaciones de medias para DSlam y DF

Prueba de muestras independientes (Variable DSlam)

		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias		
		F	Sig.	t	Sig. (bilateral)	Diferencia de medias
MgX	Se han asumido varianzas iguales	561,821	,000	-33,003	,000	-,01109
	No se han asumido varianzas iguales			-42,784	,000	-,01109

MgY	Se han asumido varianzas iguales	209,322	,000	-11,280	,000	-,00129
	No se han asumido varianzas iguales			-14,768	,000	-,00129
MgAvg	Se han asumido varianzas iguales	15,351	,000	-32,249	,000	-,00484
	No se han asumido varianzas iguales			-36,024	,000	-,00484

Prueba de muestras independientes (Variable DF)

		Prueba de Levene para la igualdad de varianzas		Prueba T para la igualdad de medias		
		F	Sig.	t	Sig. (bilateral)	Diferencia de medias
MgX	Se han asumido varianzas iguales	57,722	,000	-6,909	,000	-,00662
	No se han asumido varianzas iguales			-11,467	,000	-,00662
MgY	Se han asumido varianzas iguales	,329	,567	-10,722	,000	-,00325
	No se han asumido varianzas iguales			-12,437	,000	-,00325
MgAvg	Se han asumido varianzas iguales	,012	,913	-3,166	,002	-,00136
	No se han asumido varianzas iguales			-4,198	,000	-,00136

ANEXO III: Estadísticos descriptivos para las variables de control de diferencia de ronda.

Estadísticos de grupo

	DSF	N	Media	Desviación típ.	Error típ. de la media
MgX	1	313	,05565	,00644	,00036
	0	5954	,06084	,01215	,00016
MgY	1	313	,02198	,00297	,00017
	0	5954	,02529	,00378	,00005
MgAvg	1	313	,05233	,00375	,00021
	0	5954	,05337	,00540	,00007

Estadísticos de grupo

	DQF	N	Media	Desviación típ.	Error típ. de la media
MgX	1	621	,05502	,00885	,00036
	0	5646	,06119	,01212	,00016
MgY	1	621	,02190	,00321	,00013
	0	5646	,02547	,00370	,00005
MgAvg	1	621	,05301	,00384	,00015
	0	5646	,05336	,00547	,00007

Estadísticos de grupo

	D4thRd	N	Media	Desviación típ.	Error típ. de la media
MgX	1	119	,05279	,00738	,00068
	0	6148	,06073	,01200	,00015
MgY	1	119	,02094	,00284	,00026
	0	6148	,02520	,00378	,00005
MgAvg	1	119	,04902	,00417	,00038
	0	6148	,05341	,00532	,00007

Estadísticos de grupo

	D3rdRd	N	Media	Desviación típ.	Error típ. de la media
MgX	1	439	,05331	,00815	,00039
	0	5828	,06112	,01204	,00016
MgY	1	439	,02342	,00293	,00014
	0	5828	,02525	,00384	,00005
MgAvg	1	439	,05111	,00415	,00020
	0	5828	,05349	,00538	,00007

Estadísticos de grupo

	D2ndRd	N	Media	Desviación típ.	Error típ. de la media
MgX	1	1788	,06264	,01242	,00029
	0	4479	,05975	,01170	,00018
MgY	1	1788	,02454	,00336	,00008
	0	4479	,02535	,00395	,00006
MgAvg	1	1788	,05383	,00770	,00018
	0	4479	,05312	,00400	,00006