

EL RANGO DE VALORES DE MERCADO EN PRECIOS DE TRANSFERENCIA

Comentarios a la Nota de la AEAT del 24 de febrero de 2021

Antonio Barba (Coordinador)

José Manuel Calderón

David Cañabate

Albert Collado

José Luis Gonzalo

Ramón López de Haro

M^a Teresa Ruiz de Azúa

Víctor Viana

1. INTRODUCCIÓN

El pasado 24 de febrero la AEAT publicó una *Nota sobre determinadas cuestiones relativas al rango de plena competencia en materia de precios de transferencia* (“la Nota”), cuestión ésta, la del rango, que, como sabemos, es de todo menos pacífica en las comprobaciones de operaciones vinculadas, dado el automatismo de la Administración de ajustar precios y/o rentabilidades siempre a la mediana del rango de valores de mercado que el set de comparables refleja (sean estos internos o externos a resultas de un estudio de benchmarking con bases de datos aportado por el contribuyente y/o ajustado por la Administración).

La Nota –que representa una guía interpretativa de naturaleza administrativa- es bienvenida en el sentido de que siempre es positivo conocer la posición o criterio “oficial” de la AEAT en sus comprobaciones, máxime en temas en los que el riesgo de discrecionalidad es elevado y cuando además se incluyen reconocimientos claros sobre cuestiones que hasta no hace mucho la administración obviaba; A este respecto, es de celebrar cómo la Nota de la AEAT confirma la relevancia interpretativa de las Directrices OCDE de Precios de Transferencia y de las recomendaciones del Foro UE de PT, a los efectos de la aplicación de la normativa española (art.18 LIS) y de los convenios de doble imposición¹.

No obstante, se nota el contexto en el que se ha elaborado esta Nota, tras (o como reacción a) la SAN de 6 de marzo de 2019 (caso IKEA)², y el objetivo final que parece

¹ La AEAT no parece que haya integrado en su guía los principios derivados de la jurisprudencia del Tribunal Supremo (casos Stryker y Colgate) sobre las limitaciones al uso (creativo y retroactivo) de las Directrices de precios de transferencia y los Comentarios al modelo de convenio (*Soft-law*) de la OCDE.

² Cendoj 28079230022019100110.

buscarse que, intuimos, no es otro que el de mantener la praxis de la Administración de aplicar la mediana de forma cuasi automática en caso de ajuste, lo que, como se comentará, da lugar a contradicciones, no tanto con la propia sentencia sino principalmente con los propios apartados de las Directrices que interpreta³.

2. LA POSICIÓN DE LA AEAT

La Nota admite claramente la utilización del rango completo de valores, cuando los resultados sean relativamente iguales y muy fiables, una vez eliminados aquellos puntos del rango u observaciones que claramente se entienden que no son comparables o que adolecen de ciertas deficiencias. Por tanto, en estos casos si el resultado (precio o rentabilidad) obtenido por el contribuyente está dentro del rango completo, se cumple el Principio de Libre Competencia (“PLC”),⁴ dado que cualquier punto de este responde a un valor de mercado. Por el mismo motivo, en caso de caer fuera del rango completo, el ajuste sería al punto o valor más próximo a aquél.

Por otro lado, en la Nota la AEAT también parece admitir (excepcionalmente) situaciones donde los valores utilizados por el contribuyente están fuera del rango justificando que se trata de un resultado compatible con el principio de plena competencia⁵.

A continuación, la Nota recoge el segundo escenario posible, en el que los resultados no son relativamente iguales ni muy fiables, persistiendo defectos de comparabilidad que no pueden identificarse y/o cuantificarse, motivo por el cual y con base al artículo 17.7 RIS y el apartado 3.57 de las Directrices, se da entrada a la aplicación de medidas estadísticas de tendencia central como pueda ser el rango intercuartílico. Sería el caso de los análisis de comparabilidad realizados a través de bases de datos (estudios de benchmarking), donde las limitaciones en el acceso de información “pueden” dar lugar a rangos con valores más dispersos. En este sentido, es importante resaltar que precisamente se liga la persistencia de defectos de comparabilidad y la calificación de

³ Apartados 3.55 a 3.62, Directrices de la OCDE aplicables en materia de Precios de Transferencia a empresas multinacionales y administraciones tributarias (Julio 2017).

⁴ Tal reconocimiento por parte de la Administración es especialmente relevante cuando se dispongan de comparables internos (por ejemplo, si el contribuyente distribuye bienes análogos en condiciones similares tanto a terceros como a vinculados), por cuanto una vez eliminados los datos anómalos u operaciones específicas claramente no comparables, el rango de valores (precios, rentabilidades) del set de comparables resultante debería ser ya muy fiable e igual.

⁵ Este principio (recogido en el apartado 3.61 de las Directrices) normalmente no resulta observado por las administraciones tributarias. No obstante, no puede dejar de señalarse cómo la jurisprudencia del TJUE ha puesto en valor tal principio; en particular, el Tribunal de Justicia ha declarado que la aplicación de la normativa nacional (o convencional) reguladora del principio de plena competencia debe realizarse de forma acorde con el Derecho de la UE, en el sentido de que el contribuyente debe poder acreditar, sin soportar dificultades administrativas excesivas, la concurrencia de razones comerciales que justifican la operación.

los puntos o valores del rango como no muy fiables ni relativamente iguales, a la existencia de una “*amplia dispersión en el rango*”.

En este segundo escenario, si el resultado del contribuyente está dentro del rango intercuartílico, no da lugar a ajuste, lo que confirma también el cambio de posición de la Administración, que hasta no hace mucho seguía ajustando a la mediana en estos casos. Por el contrario, si el resultado del contribuyente cae fuera del rango intercuartílico, la Nota aboga por el ajuste a la mediana al entender que existen defectos de comparabilidad y en aplicación del apartado 3.62 de las Directrices en combinación con el 3.57 del mismo texto. Finalmente, la Nota recoge la Recomendación 6ª letra d) del Foro Conjunto de PT de la UE, al indicar que, en estos casos, sólo si el contribuyente o la Administración prueban debidamente que se debería tomar otro punto o valor del rango, sería factible no ajustar a la mediana.⁶

3. ASPECTOS CONTROVERTIDOS DE LA NOTA

En primer lugar debemos celebrar el reconocimiento expreso que la AEAT hace en la nota de que en aquellas situaciones donde existan defectos de comparabilidad que no pueden identificarse o cuantificarse, pueda mejorarse la fiabilidad del análisis, estrechando el rango de valores identificado utilizando como referencia herramientas estadísticas de tendencia central (rango intercuartil o los percentiles). Esta afirmación de la AEAT es especialmente bienvenida en los cada vez más frecuentes casos, donde la Administración en el marco de un procedimiento de inspección invoca “defectos genéricos de comparabilidad” para modificar o cambiar de forma radical la metodología utilizada por el contribuyente, dislocando todo el cálculo de la valoración de ingresos y gastos aplicado a las operaciones objeto de análisis. Parece pues que defectos de comparabilidad de “*bajo voltaje*” pueden depurarse a través del rango, sin que en estos casos esté justificado la “exclusión” de uno o varios comparables relativos o inexactos.

No obstante, si bien la Nota supone un esfuerzo relevante a efectos de sistematización y pretende ser una interpretación ordenada de los diversos apartados de las Directrices que regulan lo relativo al rango de valor, debemos lamentar que la nota no cierre del todo cuestiones que son relevantes, dejando la puerta abierta a una cierta arbitrariedad interpretativa por parte de la Administración. En particular, destacaríamos lo siguiente:

- La Nota no define cuándo la Administración entiende que hay una “*amplia dispersión del rango*”, lo que sería necesario dado que fundamenta en ello la existencia de defectos de comparabilidad que provoca que los valores no sean fiables ni relativamente iguales y, como consecuencia y de forma cuasi automática, (i) la necesidad de aplicar el rango intercuartílico en una primera

⁶ Es el supuesto que recoge la SAN 1125/2018, de 26 de febrero, en el caso MICROSOFT, en el que el tribunal confirma la bondad del ajuste al tercer cuartil realizado por la Administración atendiendo a las funciones y riesgos que asumía el contribuyente en la operativa analizada. Cendoj 280792230022018100090.

instancia⁷ y (ii) el ajuste a la mediana en caso de que el resultado del contribuyente caiga fuera.

- Parece ser que la Administración ya ha concluido que “*en la práctica es muy infrecuente disponer de un rango de valores que comprenda resultados muy fiables y relativamente iguales*”, lo que parece ser una interpretación apriorística y un tanto arbitraria que abre claramente la puerta al ajuste cuasi automático a la mediana en caso de estudio de benchmarking con base de datos.⁸
- Ligado con lo anterior, conviene recordar que más allá de las limitaciones de información que pueda haber en las bases de datos, las Directrices de la OCDE en su versión actual, recogen en sus Capítulos II (Métodos) y III (Análisis de comparabilidad) toda una serie de pautas, recomendaciones y opciones, de cara a precisamente mejorar la fiabilidad del análisis de comparabilidad – entre otras cuestiones abogando claramente por la exclusión de la muestra de aquellos potenciales elementos que no cumplan estrictamente los criterios de comparabilidad marcados, en lugar de incluirlos en la muestra y luego buscar un punto concreto en el rango intercuartílico. La AEAT en su Nota parece no considerar estas pautas, presumiendo *ab initio* que el rango de valores no va a ser suficientemente fiable.
- Teniendo en cuenta la rotundidad con la que la Nota se expresa en relación con otras cuestiones, es de lamentar que la misma no sea más precisa a la hora de plantear, aunque fuera a modo ejemplificativo, en qué supuestos podrían apreciarse por parte de la Administración defectos de comparabilidad que justifiquen la aplicación de medidas de tendencia central. La Nota se limita a señalar que la Administración deberá “*hacer referencia*” a tales defectos lo cual no parece ser suficiente a la luz de la jurisprudencia emanada de la de la sentencia IKEA, que parece exigir en estos casos algo más que meros argumentos.⁹

⁷ A este respecto, ni la Nota ni las Directrices entran en el debate sobre cuándo se entiende que hay un número suficiente de valores (“observaciones” en lenguaje matemático) que hagan viable aplicar medidas de tendencia central, pero viendo la apuesta clara de las Directrices actuales por comparables de calidad y no tanto por cantidad, no estaría de más el plantearlo. A este respecto, el TEAC, en su resolución de 28 de noviembre de 2013 (asunto 3799/2011), rechazó la aplicación del rango intercuartílico por parte de la AEAT en un caso en el que sólo había tres comparables.

⁸ Automatismo que más allá de otras consideraciones y como afirma CALDERON CARRERO, puede acabar provocando un alejamiento de los hechos y circunstancias económicas del contribuyente en la operación contrario al PLC. CALDERON CARRERO, J.M.: “Doctrina de la Audiencia Nacional sobre Ajustes de Precios de Transferencia y la aplicación del rango de precios de mercado derivado de los -benchmarking-”, *Revista Quincena Fiscal*, núm. 14, 2019, pp. 103-111.

⁹ En el caso IKEA, la Administración afirmaba que las diferencias entre los volúmenes de venta del contribuyente y las entidades seleccionadas como comparables suponían un defecto de comparabilidad relevante, extremo éste que ya fue rechazado en sede del TEAC y confirmado después por la AN, por lo

- Dado que la Administración parece abogar en su Nota por cierto automatismo en la aplicación de medidas de tendencia central como el rango intercuartílico para la fijación del rango de valores en estudios de benchmarking, la pregunta es ¿por qué no termina de precisar su criterio y aboga también por la aplicación de otras medidas estadísticas que le permitan concluir “objetivamente” si hay o no una dispersión amplia en el rango?¹⁰ Puestos a automatizar procesos, hágase hasta el final. De este modo se evitaría la arbitrariedad, se lograría mayor certeza y seguridad jurídica y el ajuste sería sólo al punto del rango más cercano.¹¹

En nuestra opinión, en la Nota emitida por la AEAT la Administración parece obviar el principio de proporcionalidad¹² que debe orientar el análisis de comparabilidad, de acuerdo con las Directrices de la OCDE, tanto en su versión de 2010 como en las actualmente vigentes del 2017. En efecto:

a) Como reconoce la Nota, los precios de transferencia no son una ciencia exacta y, tal y como las Directrices inciden de forma reiterada, no existe el comparable perfecto; de hecho, ni siquiera la no identificación de transacciones comparables entre independientes debe llevar a la conclusión de que se incumple el PLC.

b) Es en ese sentido que debe entenderse el apartado 3.55 del mismo texto, que sólo cita la Nota, en cuanto al rango de mercado se refiere, cuando acaba afirmando que “*en general la aplicación del PLC sólo da lugar a una aproximación [ya sea mayor o menor] a las condiciones que habrían sido establecidas entre independientes*”.

c) No existiendo el comparable perfecto, tampoco puede exigirse al contribuyente que lo aporte, siendo lo relevante el lograr un set de comparables lo más fiables e iguales posibles, aplicando los criterios y metodologías que fijan las propias Directrices y buscando esa razonabilidad y proporcionalidad en la aplicación del PLC.¹³ Todo ello con

cuanto la existencia de dichos defectos no había resultado acreditada y, por tanto, no era posible ajustar el resultado del contribuyente a nivel de la mediana.

¹⁰ Tales como los coeficientes de variación, la desviación típica o el teorema de la desigualdad de Chebyshev, entre otros.

¹¹ En derecho comparado podemos encontrar también jurisprudencia que es contraria al automatismo en la aplicación de la mediana y aboga por el ajuste al punto del rango más cercano, como sería el caso de las sentencias del *Bundesfinanzhof* (BFH) alemán, de 17 de octubre de 2001 (BStBl II 2004 171), de 6 de junio de 2003 (BFHE 202, 494) y 9 de julio de 2003 (BFHE 2003, 77), o la del Tribunal de Apelación de Versalles de 29 de diciembre de 2012 en el caso TSL Belgium (<https://www.legifrance.gouv.fr/affichJuriAdmin.do?idTexte=CETATEXT000033782686>)

¹² Ciertamente es que el alcance de las condiciones derivadas del test de proporcionalidad en materia de precios de transferencia que ha establecido el TJUE a través de su jurisprudencia no resulta diáfano; pensamos que el Tribunal de Justicia admite en determinados casos (razones comerciales sustantivas) que una operación intragrupo se haya estipulado en condiciones que no se ajustan a una aplicación mecánica del principio de plena competencia, tanto desde un punto de vista cuantitativo (precio) como cualitativo (operación no realizada por entidades independientes). Véanse las SSTJUE en los asuntos *SGI* C-311/18, *Hornbach-Baumark* C-382/16, o *Impresa Pizzarotti* C-558/19.

¹³ De ahí, como decíamos, el giro que dio la OCDE en la versión de sus Directrices de 2010 al poner más énfasis en reforzar las herramientas y criterios para el análisis de comparabilidad, así como la utilización

el objetivo de lograr la mejor aproximación posible a lo que debería ser un rango de valor de mercado y posibilitar que cualquier punto de este cumpla el PLC.

Por el contrario, de acuerdo al contenido de la Nota analizada parece llegarse a una interpretación forzada de estos apartados de las Directrices llegando a la doble conclusión de que (i) en general siempre habrá defectos de comparabilidad, no siendo los valores del rango muy fiables ni relativamente iguales y (ii) los mismos establecen la exigencia del comparable perfecto, lo que *de facto* lleva siempre al automatismo de aplicar cuasi automáticamente los mecanismos de tendencia central, para ajustar a la mediana.

Tales conclusiones parecen ajenas al objetivo buscado por la OCDE al regular los apartados 3.55 a 3.62 de las Directrices, siendo además susceptibles de llevar a interpretaciones contradictorias, lo que por otra parte se pone de manifiesto en el propio ejemplo que incluye la Nota.¹⁴

4. NUESTRA POSICIÓN

Con relación a la utilización del rango completo cuando los resultados sean muy fiables y relativamente iguales, estamos de acuerdo y coincidimos con la posición de la Administración que refleja la Nota, con la precisión, ya realizada, de que tal calificación (de fiabilidad e igualdad) no puede ser discrecional o arbitraria, y menos dejarse sólo a decisión de una de las partes.¹⁵

de medidas estadísticas para ajustar los defectos de comparabilidad no identificables ni cuantificables pero que se ponían de manifiesto por la amplia dispersión del rango.

En este sentido, las Directrices vigentes entre 1995 y 2009 no tenían tan desarrollado aspectos más concretos del análisis de comparabilidad, en especial en lo que a la utilización de medidas de tendencia central ni otras cuestiones más concretas del análisis de los comparables del rango se refiere. De hecho, se tendía más hacia un enfoque cuantitativo que cualitativo, dando mayor relevancia a la justificación por parte del contribuyente o la Administración del punto o valor de ese rango en el que se debería ubicar el primero conforme a su perfil funcional y de riesgos en la operación. En este sentido, el antiguo apartado 1.48 de las Directrices abogaba por que en caso de ajuste, el mismo se hiciera al punto dentro del rango que mejor reflejase los hechos y circunstancias de la operación. Este apartado 1.48 ya desaparece en las Directrices de 2010.

¹⁴ No tiene sentido alguno que el set de comparables sea “fiable” en 2018 pero no en 2017, en función de si el contribuyente está o no dentro del rango. Los comparables son fiables o no independientemente de dónde se sitúe el resultado del contribuyente. En sentido similar se pronuncia la Audiencia Nacional en la sentencia IKEA (FD 3º).

¹⁵ Incluso cabría plantearse en algunos supuestos especiales la utilización del rango completo aun cuando hubiera una elevada dispersión del rango. Sería el caso, por ejemplo, de supuestos en los que realmente es difícil identificar comparables con un elevado grado de comparabilidad al no haber entidades independientes en el sector o ser escasas, razón por la que el número de observaciones del rango fuera demasiado escaso para poder aplicar herramientas estadísticas de ningún tipo. Otro supuesto sería aquél en el que dicha dispersión fuese una característica del sector, por responder a ciclos económicos de gran volatilidad.

Respecto al segundo escenario, más habitual, que presenta la Nota en el que se hace necesario refinar los valores del rango al entenderse que no son plenamente fiables ni iguales, más allá de lo ya indicado, en nuestra opinión son dos las claves de inicio a tener en cuenta:

- La primera de ellas es que tal y como establecen los apartados 3.56, 3.57 y 3.59 de las Directrices, el punto de partida debe ser que se ha realizado un análisis de comparabilidad completo, utilizando la metodología de valoración más adecuada y produciéndose las eliminaciones de los comparables y/o datos o valores (“observaciones”) que tienen menor grado de comparabilidad, observando para ello todos los criterios y recomendaciones que las Directrices establecen y que, como ya se ha dicho, se desarrollaron más ampliamente tanto en la versión de 2010 como especialmente en la de 2017.
- La segunda clave es que, con base al punto anterior, el set de comparables seleccionados ya debería reportar un alto grado de comparabilidad, que debería resultar en un rango de valores también con una elevada fiabilidad e igualdad, sin perjuicio de supuestos excepcionales que siempre se pueden dar (vid. Nota 14).

Partiendo de lo anterior, es aquí donde ya entra en juego el apartado 3.57 de las Directrices cuando igualmente recomienda la aplicación de medidas de tendencia central (habitualmente el rango intercuartílico, pero pueden ser otras) para acabar de limar los defectos de comparabilidad que pudieran subsistir por las limitaciones en el acceso de información, y que no se pueden identificar ni cuantificar pero que se reflejan en la dispersión del rango.¹⁶ En este sentido, dichos **defectos**, por lo general, no pueden por más que calificarse de **menores**, por cuanto de lo contrario difícilmente no serían identificables ni cuantificables.

Aplicado el rango intercuartílico para eliminar los defectos menores en la comparabilidad, los valores ubicados dentro de dicho rango ya deberían ser muy fiables y relativamente iguales, por lo que cualquier punto de ese rango correspondería a un valor de mercado. Por tanto, y alineándonos con las Directrices y la posición de la Nota, si el resultado del contribuyente se ubica dentro del rango intercuartílico, se cumple el PLC, y de no ser así, se debería ajustar al valor o punto más próximo a aquél.

En esta última situación (fuera de rango), se plantea a continuación la pregunta sobre **¿cuándo entra en juego el apartado 3.62 y el consiguiente ajuste a la mediana?**

¹⁶ En caso de que el rango no fuera muy disperso, aquí se plantearía la posibilidad de aplicar el rango completo aun cuando se hubiese obtenido el mismo a través de un estudio de benchmarking, por cuanto ¿cómo puede defenderse la existencia de defectos de comparabilidad cuando no se pueden identificar, ni cuantificar, ni suponen una elevada dispersión del rango? ¿Dónde están los defectos? De nuevo aplica aquí el principio de proporcionalidad y el objetivo buscado en todo análisis de Precios de Transferencia, que no es otro que lograr la mejor aproximación posible y razonable a lo que deberían ser condiciones de mercado.

En nuestra opinión, el ajuste a la mediana y no al punto del rango más próximo al resultado del contribuyente sólo debería aplicar cuando “persistan” los defectos de comparabilidad, esto es, cuando a pesar de haberse aplicado el rango intercuartílico, exista todavía una elevada dispersión en los valores de éste,¹⁷ lo que denotará que los defectos de comparabilidad persisten o, más bien, que en esa operativa concreta no es posible lograr un elevado grado de comparabilidad en el análisis (cuestión esta que es muy habitual en determinados sectores). Por tanto, el apartado 3.62 de las Directrices actuaría en combinación con el 3.57 como segundo filtro o refinamiento del rango de valores, para los casos en que el resultado del contribuyente cayera fuera del rango intercuartílico y éste continuara reflejando una amplia dispersión entre sus valores. Sólo bajo esta interpretación tiene sentido la aplicación conjunta de ambos apartados y se evita que ambos entren en contradicción, dándose situaciones como las que refleja el ejemplo que recoge la Nota o la propia sentencia IKEA (vid. Nota 10).

5. REFLEXIÓN FINAL

Como decíamos al inicio de este artículo, la Nota tiene aspectos muy positivos, es sistemática y didáctica y en este sentido es bienvenida al aportar claridad y certidumbre.

Es especialmente loable cómo la nota de la AEAT refleja una aplicación del principio de plena de competencia (PLC) a partir de un enfoque de “rango de valores de mercado” y no de “precio de mercado” en sentido estricto, con carácter general; la funcionalidad del rango, como ya hemos comentado, se materializa a través de su uso para “estrechar el rango” (defectos de comparabilidad), así como para determinar un punto del rango de valores de mercado (utilización de medidas de tendencia central: mediana, media ponderada, media). A este respecto, es de celebrar el reconocimiento por parte de la AEAT de que en situaciones donde existan defectos de comparabilidad que no pueden identificarse o cuantificarse, se pueda mejorar la fiabilidad del análisis, estrechando el rango, utilizando como referencia herramientas estadísticas de tendencia central sin hacer cuestionamientos más radicales del método de valoración utilizado por el contribuyente o de comparables concretos incluidos dentro del rango.

Sin embargo, vemos en los criterios interpretativos expuestos en la Nota un intento de fundamentar jurídicamente con base a las Directrices la existencia permanente de defectos de comparabilidad en el rango de comparables que aporten los contribuyentes, con el objetivo de permitir a la Administración continuar aplicando de forma automática ajustes a la mediana, llevando al límite la interpretación de los apartados 3.55 a 3.62 de las Directrices, en contra del principio de proporcionalidad que subyace en los mismos y del fin propio para el que fueron regulados.

¹⁷ De ahí la importancia de aplicar medidas estadísticas para verificar la existencia de tal dispersión (vid. Nota 8) y la oportunidad perdida de haber desarrollado en la Nota los pautas a considerar a efectos de establecer umbrales máximos de dispersión.

Por tanto, si consideramos que la praxis más habitual de la Administración en sus comprobaciones es la de cuestionar el análisis de comparabilidad del contribuyente, imponiendo el suyo propio, habitualmente basado en un benchmarking alternativo que

refleje que el valor aplicado por el contribuyente se encuentre fuera de rango; para posteriormente ajustar su resultado a la mediana del rango alternativo propuesto por la Administración,¹⁸ resulta claro que a partir de ahora habrá que reforzar todavía más el análisis de comparabilidad, aplicar medidas estadísticas adicionales para refinar el rango de valores y tratar siempre de quedar dentro del rango intercuartílico.

La reproducción, copia, uso, distribución, comercialización, comunicación pública o cualquier otra actividad que se pueda realizar con el contenido de este documento, incluida su publicación en redes sociales, queda condicionada a previa autorización de la AEDAF.

¹⁸ Lo que no dejará de ser curioso, por cuanto de aceptarse la interpretación de la Administración, resultará que ésta ajustaría a la mediana por persistir defectos de comparabilidad en un rango de valores seleccionado por ella misma en su propio estudio de benchmarking.