

UNIVERSIDAD AUTONOMA DE NUEVO LEON
FACULTAD DE ECONOMIA
DIVISION DE ESTUDIOS SUPERIORES

TESIS

**DETERMINANTES DE LA PIRATERIA:
EL CASO DEL SOFTWARE PARA EL 2009**

POR

MILTON ALEJANDRO LOPEZ SIFUENTES

Como requisito parcial para obtener el Grado de MAESTRIA EN
ECONOMIA con Especialidad en Economía Industrial

Marzo 2014

Índice

1: Introducción.....	3
2: Antecedentes.....	6
3: Causas de la Piratería.....	12
4: Marco Teórico	
a) Hipótesis.....	17
b) Factores de Mercado.....	18
5: Marco Empírico	
a) Modelo	22
b) Datos.....	24
c) Análisis de Datos y Estimaciones.....	28
6: Conclusión.....	33
7: Bibliografía.....	36
8: Anexo.....	39

RESUMEN

Esta investigación proporciona evidencia empírica sobre la relación entre la piratería, y factores mercado. Para ello, usamos un conjunto de datos para 61 países en el 2009, las tasas de piratería en relación a las institucionales, ingreso, precios, educación e infraestructura. Como principal conclusión, encuentro que los precios y la incorrupción son más importantes que la legislación para influir sobre la tasa de piratería.

1.- Introducción

El conocimiento que se ha desarrollado en este siglo se difunde con gran rapidez gracias a la tecnología. Sin embargo, las mismas innovaciones tecnológicas han propiciado los delitos como el robo de conocimientos debido a que han reducido las barreras para su protección y han facilitado el copiado. La piratería es un problema derivado de la violación de los derechos de propiedad intelectual (DPI)¹. Legalmente la Propiedad Intelectual (PI) ha sido definida como el “derecho exclusivo de explotar las producciones del ingenio y la creatividad”. Esta definición se identifica con los postulados de los regímenes constitucionales en los cuales las creaciones del intelecto humano se constituyen en un bien jurídico amplio que merece valor. De este modo cuando la sociedad valora la PI, crea instituciones que sostienen estos derechos (Becker L., 1993).

El consumidor piratea sólo algunos productos que presentan características especiales: son bienes que pueden reproducirse sin perder sus características originales de manera rápida y a un costo casi cercano a cero. Los productos con estas características que figuran entre los más pirateados son los siguientes: software, música, videos y libros.

El objetivo de la investigación es contribuir a la comprensión de la problemática de la piratería identificando los factores que la provocan, con el fin de que las autoridades que

¹ Los Derechos de Propiedad Intelectual (DPI) se dividen en dos categorías: propiedad industrial y derechos de autor. Los últimos se refieren a derechos concedidos a los creadores sobre sus obras, son derechos privados que crean un monopolio legal sobre la explotación comercial de la propiedad intelectual en un período específico de tiempo y son automáticos para el Autor, por el Convenio de Berna. Esta información se encuentra en la página web de la Organización Mundial de la Propiedad Intelectual (OMPI).

legislan dispongan de criterios para enfrentarla. Esta investigación se concentra en la piratería de software debido a que el análisis de toda la gama de DPI sería difícil. En parte, esta dificultad se debe a que los mecanismos para medir la violación de los DPI son diversos y las mediciones son realizadas por diferentes organizaciones. Es importante aclarar que usualmente el software no se vende. En realidad, los consumidores adquirimos una licencia, que permite el uso pero restringe la disposición del mismo. Así este estudio está enfocado en identificar y cuantificar el efecto de los principales factores que explican los diversos niveles de piratería de software en los países para el 2009. A través de esta investigación se encuentra que los precios y la incorrupción son más importantes que la legislación para influir sobre la tasa de piratería del software.

Los trabajos pioneros sobre piratería fueron meramente descriptivos y examinaban por qué los individuos incurren en esta práctica. Estudios posteriores utilizan datos de corte transversal, identificando al ingreso per cápita, así como factores culturales e institucionales como los principales determinantes de las tasas de piratería de software en distintos países (Gopal y Sanders, 1998; Husted, 2000; Simpson y Simpson 1994).

En el siguiente capítulo se hace un repaso de algunos trabajos previos sobre la piratería y algunas cifras internacionales sobre las tasas de piratería de software en distintos países. En el tercer capítulo se presenta una revisión de la literatura relacionada con los factores que inducen el fenómeno de la piratería a nivel internacional. En el cuarto capítulo se presentan las hipótesis del trabajo. En el quinto capítulo se describe el modelo empírico que se propone en este estudio para analizar el fenómeno señalado anteriormente, así como las variables a utilizar (base de datos) y posteriormente se presenta la estimación empírica,

utilizando el análisis de regresión lineal. En el sexto se presentan las conclusiones y limitaciones.

2.- Antecedentes

La piratería es considerada un delito dado que viola los DPI. Esto sucede cuando se reproduce o distribuye el material o producto sin autorización de quien tenga los DPI. En el caso específico del software sucede cuando se copia, descarga, comparte, vende o instala en múltiples equipos personales o de trabajo. Lo que muchas personas no advierten es que al adquirir software están comprando una licencia para usarlo, y no el software en sí².

Hay diversos estudios que enfatizan las desventajas de la piratería. Por ejemplo, los primeros estudios económicos sobre el tema, afirman que los DPI incentivan la innovación (Arrow, 1962; Nelson, Peck y Kalachek, 1967; Johnson, 1985; Novos y Waldman, 1984), lo cual genera crecimiento porque produce información que permite el progreso tecnológico (Dasgupta y Stiglitz, 1980). El progreso tecnológico, a su vez, genera desarrollo económico (Arrow, 1962; Nelson, Peck y Kalachek, 1967), pero el efecto final es limitado cuando existe piratería.

Existen estudios que no tienen una visión negativa de la piratería pues afirman que la creatividad no requiere estímulos legales o económicos (Lakhani y Wolf , 2005). Incluso algunos autores vinculan un aumento en protección a la propiedad intelectual con una pérdida de bienestar social. Las creaciones son una realización colectiva de distintas personas (McLeod, 2001), ya que la innovación se genera combinando el conocimiento ya existente con la creatividad de los autores (Landes y Posner, 1989).

²Tomada la definición de la página de Business Software Alliance (BSA): http://www.bsa.org/country.aspx?sc_lang=es-MX

Estas dos divisiones pueden generar un conflicto que se puede observar comparando países con distinto nivel de desarrollo. En particular, si el crear protección a los DPI produce resultados diferentes dependiendo del país en el que se aplique la política. De acuerdo con Robinson (1993), en un país desarrollado la protección a los DPI incentiva la innovación y con ello el crecimiento. Por el contrario, en un país en vías de desarrollo considera que proteger los DPI restringe el conocimiento existente y hace más difícil el desarrollo.

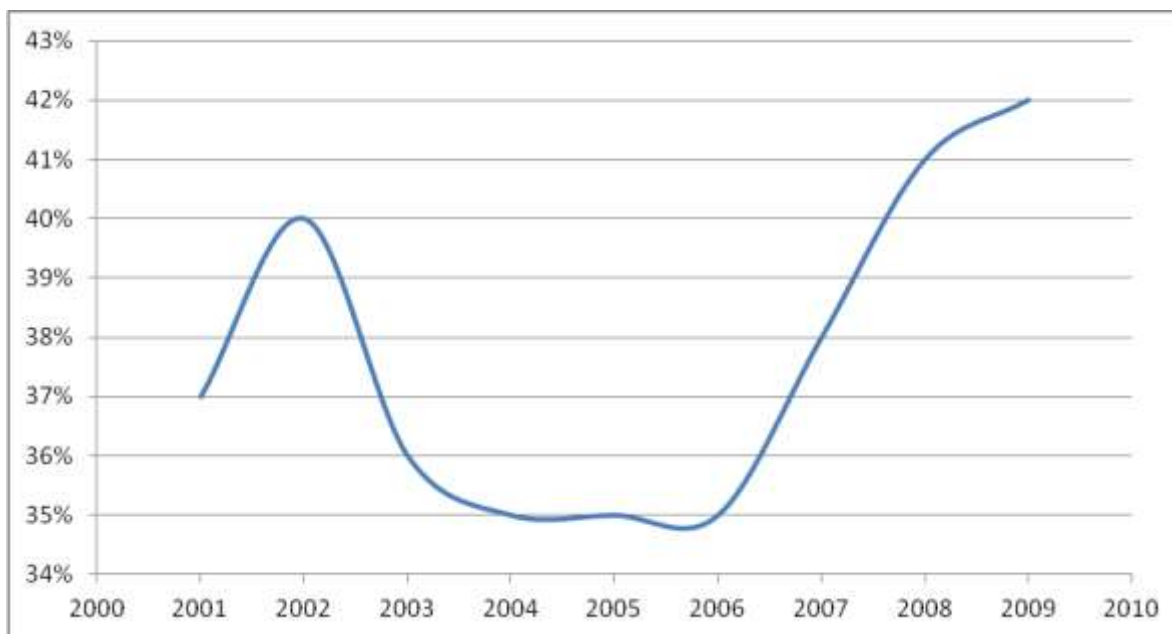
El desarrollo de un sistema de intercambio globalizado ha hecho que los contextos económicos, sociales y culturales que fundamentan los sistemas de regulación cedan frente a la unificación que favorece el comercio internacional. Esta unificación se traduce en compromisos en forma de tratados o acuerdos internacionales. Estos tratados son dirigidos a la protección del interés de países desarrollados, los cuales buscan no ver amenazados su inversión y el desarrollo en información y tecnología en términos de la competencia internacional (Ronkainen y Guerrero, 2001). Dada la aplicación de un sistema jurídico que surge de cierto contexto cultural, económico y político, sobre otro contexto poco similar, la consecuencia será el desarrollo de una regulación ineficiente, traduciéndose esto en una constante violación a las normas.

La gráfica 1 muestra la tendencia de la tasa de piratería del software³, la progresión nos muestra que del 2002 al 2003 la tasa de piratería descendió 5 puntos porcentuales y después, durante 3 años, se mantuvo alrededor de la tasa de 35%. Ahora bien, el comportamiento de la gráfica no sólo reporta la disminución de la tasa de violación, sino que muestra tres ciclos, el primero ascendente, del 2001 al 2002, segundo descendente de 2002

³ Tasa de piratería= número de software piratas/ número total de software

al 2003 y tercero ascendente del 2006 al 2009. Las razones de este cambio en el ritmo cíclico se pueden deber, entre muchas otras razones, al auge del Internet en los países subdesarrollados y las diferentes políticas de la propiedad intelectual⁴.

Gráfica 1: Tasa de Piratería Mundial (%).



Fuente: BSA

En la gráfica 2 se observa la evolución del nivel de piratería que ha tenido cada región a través del tiempo⁵. Datos destacables son los de la región de Europa Central y

⁴Países tomados en cuenta para hacer las gráficas: Albania, Argelia, Argentina, Armenia, Australia, Austria, Azerbaijan, Bahrein, Bangla Desh, Belarús, Bélgica, Bolivia, Bosnia, Botswana, Brasil, Brunei, Bulgaria, Camerún, Canadá, Chile, China, Colombia, Costa Rica, Croacia, Chipre, República Checa, Dinamarca, República Dominicana, Ecuador, Egipto, El Salvador, Estonia, Finlandia, Francia, Georgia, Alemania, Grecia, Guatemala, Honduras, Hong Kong, Hungría, Islandia, India, Indonesia, Irak, Irlanda, Israel, Italia, Costa de Marfil, Japón, Jordania, Kazajstán, Kenia, Kuwait, Letonia, Líbano, Libia, Lituania, Luxemburgo, Malasia, Malta, Mauritania, México, Moldavia, Montenegro, Marruecos, Países Bajos, Nueva Zelanda, Nicaragua, Nigeria, Noruega, Omán, Pakistán, Panamá, Paraguay, Perú, Filipinas, Polonia, Portugal, Puerto Rico, Qatar, Reunión, Rumania, Rusia, Arabia Saudita, Senegal, Serbia, Singapur, Eslovaquia, Eslovenia, Sudáfrica, Corea del Sur, España, Sri Lanka, Suecia, Suiza, Taiwán, Tailandia, Túnez, Turquía, Ucrania, Reino Unido, Estados Unidos de América, Uruguay, Venezuela, Vietnam, Yemen, Zambia y Zimbabwe.

⁵ De acuerdo con BSA: Países que conforman cada región.

Asia y el Pacífico Australia, Bangla Desh, Brunei, China, Hong Kong, India, Indonesia, Japón, Malasia, Nueva Zelanda, Pakistán, Filipinas, Singapur, Corea del Sur, Sri Lanka, Taiwán, Tailandia y Vietnam.

Europa Central y Oriental Albania, Armenia, Azerbaijón, Bosnia, Bulgaria, Croacia, República Checa, Estonia, Georgia, Hungría, Kazajstán, Letonia, Lituania, Moldavia, Montenegro, Polonia, Rumania, Rusia, Serbia, Eslovaquia, Eslovenia y Ucrania.

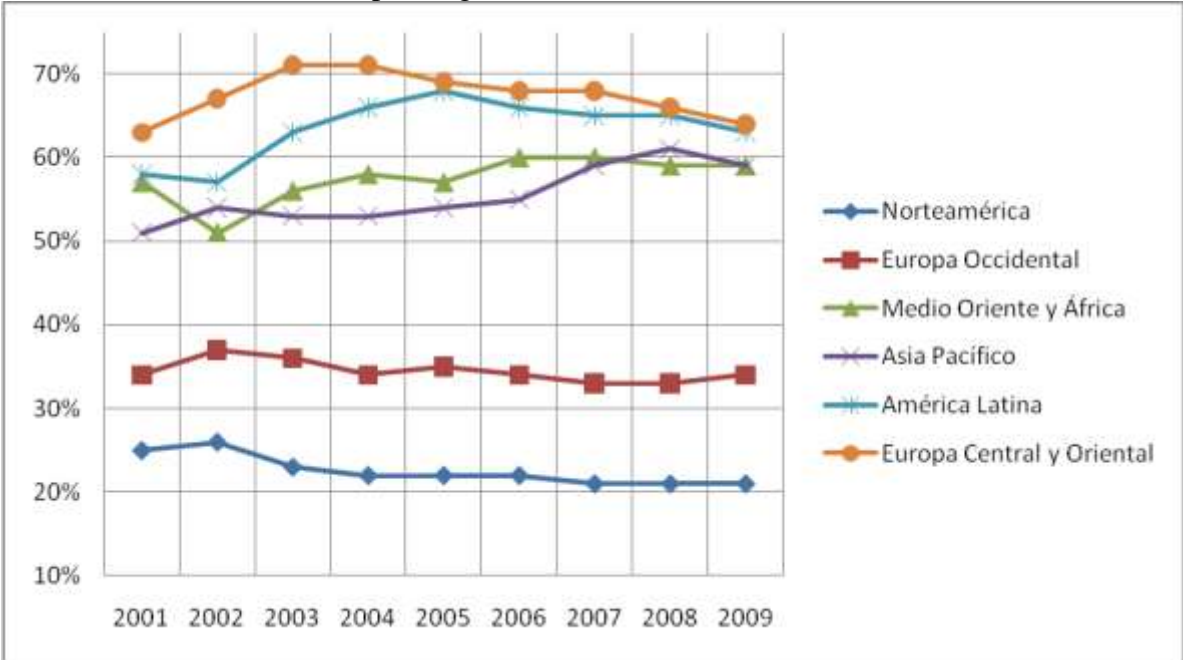
América Latina Argentina, Bolivia, Brasil Chile, Colombia, Costa Rica, República Dominicana, Ecuador, El Salvador, Guatemala, Honduras, México, Nicaragua, Panamá, Paraguay, Uruguay, Perú y Venezuela.

Oriente Medio y África Argelia, Bahrein, Botswana, Camerún, Egipto, Iraq, Israel, Costa de Marfil, Jordania, Kenia, Kuwait, Líbano, Libia, Mauricio, Marruecos, Nigeria, Omán, Qatar, Reunión, Arabia Saudita, Senegal, Sudáfrica, Túnez, Turquía, Emiratos Árabes Unidos, Yemen, Zambia y Zimbabwe.

América del Norte Canadá, Puerto Rico y Estados Unidos.

Oriental donde se ha observado la mayor incidencia de piratería, comparada con las otras cinco regiones. De acuerdo con la grafica en los primeros tres años del periodo analizado se observó un aumento en la piratería, pero a partir del año 2003 esta ha tenido un descenso paulatino. La región que ha tenido una menor participación en este delito es Norteamérica, donde su nivel de piratería ha tenido una leve variación del 2003 al 2009.

Gráfica 2: Nivel de Piratería por Región (%).



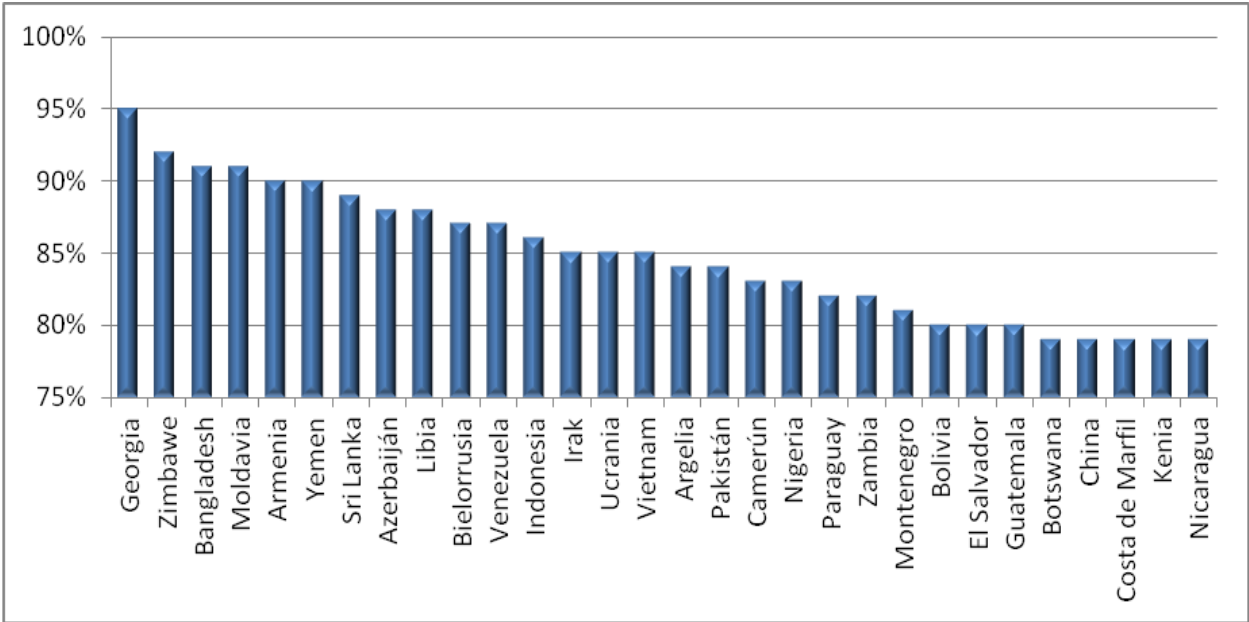
Fuente: BSA

El último reglón de la gráfica 2 nos muestra las tasas de piratería para el 2009. Las cifras oscilaron entre el 59% en Asia, 64% en Europa Central y del Este, 63% en Latinoamérica, 59% en África y Oriente Medio, 21% Norteamérica y 34% en Europa Occidental.

Europa Occidental Austria, Bélgica, Chipre, Dinamarca, Finlandia, Francia, Alemania, Grecia, Islandia, Irlanda, Italia, Luxemburgo, Malta, Países Bajos, Noruega, Portugal, España, Suecia, Suiza y Reino Unido.

En la gráfica 3 se destacan los niveles más altos de piratería en el mundo por país en el 2009, donde el nivel más alto de piratería lo ocupa Georgia con una tasa de 95%, mientras que el país latino con mayor problema de piratería es Venezuela con una tasa de 87%. Es necesario hacer la observación que los países con más altas tasas de piratería en su mayoría pertenecen a las regiones de Latinoamérica, África y Europa Central y Oriental.

Grafica 3: Tasas de Piratería más altas en el mundo en 2009 (%).



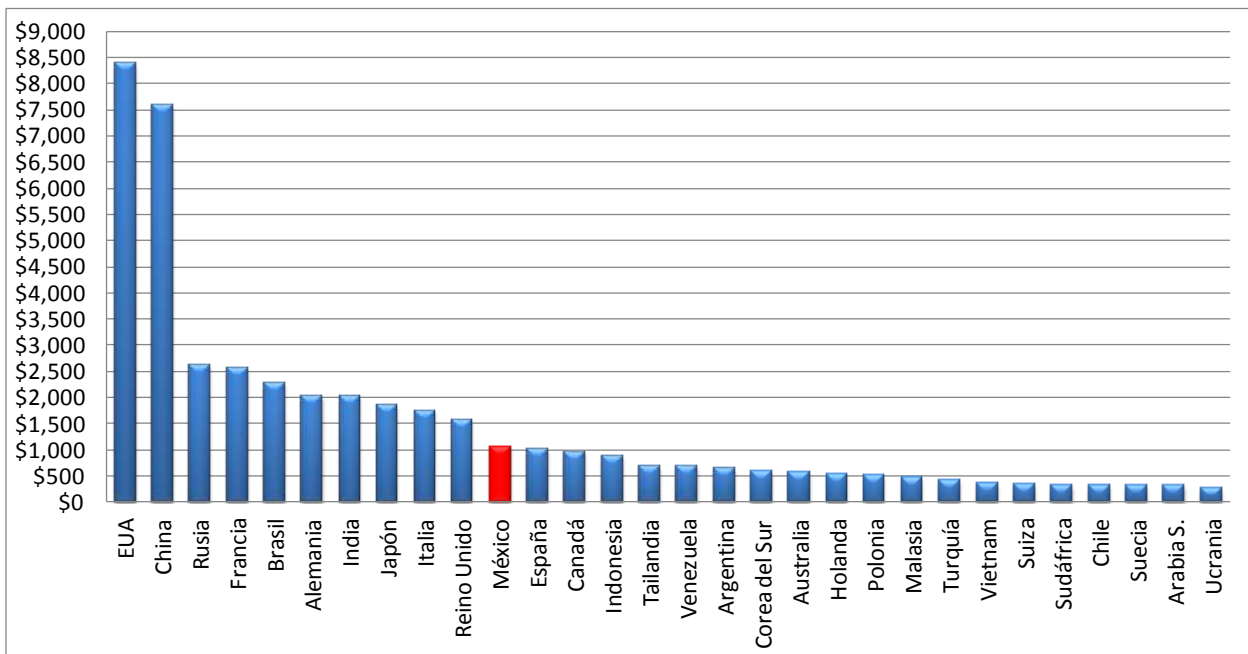
Fuente: BSA.

En la gráfica 4 se presentan los valores de piratería para el 2009⁶, entre los más altos del mundo se encuentra Estados Unidos. A pesar de tener una tasa de piratería del 20% solamente, el valor de mercado de las mismas es \$ 8,390 millones. La explicación de este fenómeno es el tamaño del mercado de software que tiene el país.

⁶ Valor de la Piratería = Número de software pirata por valor de software original.

México tiene pérdidas por este concepto de 1,000 millones de dólares con una tasa de piratería de 60%, mientras que España tiene una pérdida aproximadamente igual en valor monetario pero con una tasa de piratería de 42%, la diferencia entre estos dos países se da por el tamaño de mercado del software.

Gráfica 4. Países con valores más elevados de Piratería 2009 (millones de dólares).



Fuente BSA.

3.- Causas de la Piratería

Los primeros trabajos que buscan los factores que influyen en la violación de los DPI mencionan como principales variables el ingreso, ambiente legal y precio (Rapp y Rozek, 1990; Gopal y Sanders, 2000). Lo anterior puede parecer incongruente con algunos ejemplos como los que se exponen a continuación. Hay países como Rumania y Alemania que tienen un Índice de Entorno Legal similar⁷; sin embargo la tasa de piratería varía 40 puntos porcentuales entre ellas. De manera similar, entre Austria y Rusia la tasa de piratería varía 44 puntos porcentuales. Finalmente, en países como Estados Unidos de América y Hong Kong, donde el PIB per cápita (Poder de Paridad de Compra) es similar, la tasa de piratería varía 28 puntos porcentuales, mientras que entre Sudáfrica y Azerbaijón la tasa de piratería varía 55 puntos porcentuales.

Rapp y Rozek (1990) estudian los DPI con las patentes de invención. Demuestran que la violación a los derechos concedidos por patente está altamente correlacionada negativamente con el desarrollo económico de un país. De este modo encuentran que países con un alto grado de piratería tienden a ser economías en desarrollo, caracterizadas por un PIB de nivel bajo.

Gopal y Sanders (2000) muestran que la política de precios utilizada en la fabricación de software está basada solamente en el país de origen y que al exportar no toma en cuenta el ingreso per cápita de los países a los cuales envía los productos,

⁷ Alcance y transparencia de la legislación de propiedad intelectual, Estado de las leyes nacionales contra los delitos electrónicos, Estado de las leyes nacionales de privacidad de datos y contra el correo no deseado.

haciéndolos de algunas maneras inalcanzables para cierto nivel de usuarios. Por lo tanto, proponen una estrategia de discriminación global de precios para disminuir la piratería.

Swinyard, Rinne y Keng (1990) citan un conocido proverbio chino que nos puede dar una idea de los valores en esa cultura : “Aquel que comparte será recompensado y aquel que no, se condenará”. Sobre la base de estas diferentes perspectivas, se puede confirmar que los asiáticos tienen valores morales muy distintos con respecto a los derechos de autor y, si bien se puede decir que su conducta es ilegal, no se puede decir que sea inmoral. Se puede decir que sus valores morales son distintos de los occidentales. Estos hallazgos merecen ser profundizados en una perspectiva global, pues se propongan leyes universales de propiedad intelectual que, como se ve, no son aceptadas por todos. Por ello es importante estudiar la piratería tomando en cuenta estas diferencias de perspectiva sobre la moral.

Peace, Galleta y Thong (2003) hacen un estudio para Estados Unidos entre 201 estudiantes adultos de un programa MBA de tiempo parcial. Todos los participantes tenían conocimientos de computación y usaban computadoras diariamente en sus centros de trabajo. Encuentran que las actitudes, las normas subjetivas y el control de conducta son precursores de la intención de copiar software ilegalmente. A ello se suma que la severidad del castigo, la certeza del mismo y el precio del software tienen efectos directos en la actitud hacia la piratería, mientras que la certeza de ser castigado afecta fuertemente el control de la conducta.

Tan (2002) estudia la conducta moral en la decisión de consumo de un producto pirata, asumiendo que la conducta es un reflejo de la intención de compra, y analiza algunos factores que influyen en ella, como la intensidad de la percepción moral, expresada por la magnitud de las consecuencias, la inmediatez en el tiempo y el consenso social; el riesgo percibido desde el

punto de vista financiero, de resultados, de ser perseguido y censurado por la sociedad; y el juicio moral entendido como juicio cognoscitivo y razonamiento moral. Los resultados indican que el riesgo percibido y el juicio moral influyen en la toma de una decisión ética, mientras que la variable de intensidad de la moral no es un factor determinante. Basado en sus resultados, el autor sugiere que para eliminar la piratería hay que levantar más medidas punitivas, pues ello transmite la información de que el acto de piratear es malo.

Simpson y Simpson (1994) hacen una investigación aplicando 209 encuestas a estudiantes pertenecientes a varias clases de negocios relacionados con informática. El cuestionario fue diseñado para evaluar las características demográficas, el comportamiento pasado, la percepción de la conducta ética de la piratería, y el potencial de factores de motivación para la piratería de software. Las variables demográficas obtenidas son la edad, sexo y las preferencias religiosas. Los autores descubren en su investigación que la decisión de consumos de bienes piratas es independiente de la percepción ética que tenga cada individuo. Existen situaciones que benefician a la piratería como la exclusividad en la distribución, así como el tiempo de retraso entre el pedido y la entrega.

Ginarte y Park (1997) llevan a cabo un estudio para 110 países con datos de 1960 a 1990. El objetivo de esta investigación es encontrar qué factores o características de las economías determinan los derechos de patentes. Lo innovador de esta investigación es la construcción de un índice compuesto por cinco categorías para cuantificar las patentes: extensión de la cobertura, miembros de acuerdos internacionales, disposiciones para la pérdida de protección, mecanismos de aplicación, y la duración de la protección. El estudio

revela que la piratería a patentes es menor en economías más desarrolladas, con mayor capital humano y con sectores grandes de investigación y desarrollo.

Husted (2000) lleva a cabo un estudio de 39 países. En particular contrasta el nivel de desarrollo económico, la desigualdad en el ingreso y las variables culturales de Hofstede (1983), con los porcentajes de violación de los Derechos de Propiedad Intelectual, medidos en términos de la piratería de Software. Este autor, encuentra que la piratería de software está correlacionada negativamente con el PIB per cápita, la desigualdad en el ingreso y el individualismo. Este estudio es valioso en tanto que analiza la cultura como determinante de la piratería de software. Concluye estableciendo que el individualismo, la desigualdad en el ingreso y el bienestar económico, son las únicas variables que tienen una alta relevancia en la piratería.

Rodríguez (2002) hace un estudio para 24 países con datos de 1994, 1997 y 2000. Así utiliza un panel de datos para probar que en el caso del software, el nivel de ingresos tiene una relación negativa con las tasas de piratería de los países. Finalmente concluyen que hay efectos fijos por países que afectan a la piratería de software.

Márquez (2003) hace un estudio para 15 países con datos del 2002. El objetivo de esta investigación es encontrar qué factores de las economías determinan la piratería. Para su trabajo toma variables de cultura, institucionales, de mercado y específicas por nación, que le permiten probar un modelo alternativo que explica la violación de los DPI derivados del software. Demuestra que la violación a los DPI es más un fenómeno cultural que un

fenómeno de mercado. Este modelo alternativo permitió definir como variables explicativas, la educación, la cultura y las instituciones.

4.- Marco Teórico

HIPÓTESIS

En el anterior capítulo se comentaron algunos modelos que explican la piratería. De ellos, se pueden derivar modelos alternativos que incluyen factores adicionales. De aquí surge la hipótesis general del trabajo:

Hay una relación directa entre la piratería y algunas variables complementarias a las de mercado tradicionalmente consideradas en el análisis de tipo de mercado⁸.

La pregunta que surge ahora, y que será nuestro objeto de estudio en la siguiente parte del texto, es la siguiente: ¿cuáles son las variables que complementan la piratería? Según las investigaciones anteriores, pueden agruparse en: institucionales, mercado, educación e infraestructura. Éstos sustentarán el modelo para explicar la piratería. A continuación se explica cada uno de los factores y las variables que podemos incluir en el modelo.

⁸ El término variable de mercado puede ser equivocado, debido a que las variables mencionadas también hacen parte de aquel tipo de interacción entre agentes que los economistas llaman mercado, pero lo usaré así para diferenciarlo de las demás variables.

Factores de Mercado

La piratería se puede explicar por variables que reflejen la situación de los mercados internos de cada país. Así, la piratería se presenta por los bajos ingresos de sus habitantes, por la escasa atención que se le presta en la política económica y por los altos precios de los productos, entre otros.

Ingresos

Se puede señalar que el nivel de ingresos en una economía, influye en la piratería. Tal como lo explican Rapp y Rozek (1990) los DPI son más fuertes en países en los que el ingreso per cápita es mayor. Existen dos explicaciones de esta relación directa entre DPI y el nivel de ingresos. Por una parte, se argumenta que los DPI favorecen el crecimiento (Ginarte y Park, 1997). Por otra parte, los individuos con altos ingresos son menos propensos a adquirir bienes piratas (Husted, 2000; Rodríguez 2002). En cualquier caso, se podría esperar que “un mayor ingreso, genere una menor tasa de piratería”.

Inversión en investigación y desarrollo

La protección a los DPI es más valorada en términos de la política económica (Ginarte y Park, 1997) en los países que tienen altas tasas de inversión en Investigación y Desarrollo (I y D). De no haber protección a tales derechos, se presenta un desincentivo a la inversión y con ello, se incurre en problemas de crecimiento económico. Así cuando se invierte en I y D, se espera que exista cierta protección a esa inversión. En este caso, se exigirá que entre mayor sea la inversión en I y D haya también mayor protección a los DPI.

Por lo anterior se puede considerar lo siguiente “A mayor nivel de inversión en investigación y desarrollo, la tasa de piratería de software podría ser menor”.

Precios del mercado

Los DPI otorgan al productor cierto poder de mercado que le permite controlar los precios del producto. El productor para maximizar sus beneficios, ofrece su producto a precios elevados. El alto precio puede incentivar al consumidor a adquirir el software ilegal que tiene un precio más bajo (Peace, Galleta y Thong, 2003; Seale, Polakowski y Schneider, 1998; Moores y Dhillon, 2000). El consumidor asume el riesgo que implica la compra de dicho bien ilícito teniendo en cuenta la probabilidad de ser atrapado y sancionado. El productor de software ilegal tiene unos costos bajos de producción pues no invierte en el desarrollo del software y la producción del mismo es muy sencilla debido a la naturaleza física de dichos sistemas. Así, hay una amplia diferencia entre el precio del software pirata y el precio del software legal, la cual incentiva al consumidor a comprar el software ilícito. Por ello, se puede esperar que “a menor precio del software legal, menor sería la tasa de piratería”.

Corrupción

En muy pocas ocasiones los administradores y planificadores se detienen a pensar cómo afecta el conjunto de las reglas básicas políticas y sociales. El principal postulado de la escuela institucionalista y neo-institucionalista es que las instituciones importan en el desarrollo del análisis económico. Tal como dice Coase: “Tiene poco sentido para un economista discutir el proceso de intercambio sin especificar el arreglo institucional en el que las transacciones tienen lugar dado que esto afecta los incentivos para producir y los

costos de transacción” (Coase, 1994). Con base en estudios anteriores se puede suponer que “A menor corrupción, menor tasa de piratería de software”.

Ambiente Legal

Los países productores de PI exigen sistemas de protección a los países con los que tienen intercambio. Las exigencias de estos sistemas se manifiestan a través de acuerdos internacionales que promueven la protección de los DPI. Los acuerdos internacionales con otros países son factores que afectan la protección a los DPI, puesto que el cambio de régimen y el compromiso internacional afectan la asignación de los recursos del Estado (Guerrero-Ronkainen, 2001; Burke, 1996). Los tratados reflejan el ambiente legal donde se ve la aplicación, el alcance y la transparencia de las leyes. Por lo tanto, se espera lo siguiente “A mayor ambiente legal, menor piratería de software” (Peace, Galleta y Thong, 2003; Tan, 2002; Rodríguez, 2002).

Factor de educación

La educación ha sido considerada como una característica fundamental que permite el progreso de la tecnología y, por ello, es una base para el desarrollo económico (Tilak, 1993). Podemos establecer que la educación es un factor que incide en la piratería debido a que ésta afecta la demanda de productos. Si los consumidores no tienen cierto número de años en el sistema educativo, entonces no tienen capacidad de demanda. (Ginarte y Park, 1997). La débil protección a los DPI en los niveles de educación baja, es debido a la poca valoración del trabajo en el sector de la información. Con todo, no podemos afirmar que a mayor nivel educativo mayor demanda de Propiedad Intelectual (PI) y mayor violación a los DPI. Se ha encontrado que en niveles de educación altos hay una valoración adicional a la

legalidad en algunos contextos culturales. Podemos suponer que “A mayor coeficiente de educación en un país, menor tasa de piratería”.

Factores infraestructura

La infraestructura computacional ha facilitado la violación a dichos derechos, ya que la red simplifica y agiliza la distribución de software pirata a un precio cercano a cero. Las redes permiten traspasar el documento más fácilmente y sin detección, facilitando el proceso de producción de las obras ilegales (Seale, Polakowski y Schneider, 1998; Moores y Dhillon 2000). Por lo tanto, se puede esperar que “A mayor infraestructura, mayor tasa de piratería”.

5.-Marco Empírico

Modelo

Se construye un modelo de regresión lineal para estudiar la piratería de software. Debido a la carencia de datos de series de tiempo, se emplearán datos de corte transversal del año 2009⁹. A continuación se especifica el modelo empírico base que se pretende estimar.

$$PIR_i = \alpha_i + \beta_1 INC_i + \beta_2 INF_i + \beta_3 CAP_i + \beta_4 LEG_i + \beta_5 R \& D_i + \beta_6 PIB_i + \beta_7 P_i + \varepsilon_i \quad (1)$$

El cuadro 1 define las abreviaciones de las variables y presenta los signos esperados de los coeficientes correspondientes. Por ejemplo, la variable P_i es el precio del software y se espera que un precio más alto provoque una mayor tasa de piratería. De manera similar, la variable INC_i es un índice de incorrupción y se espera que un mayor valor de este índice genere menor tasa de piratería. El resto de las variables en el cuadro se interpretan de forma similar.

⁹ Países utilizados en la muestra para la investigación: Argelia, Argentina, Australia, Austria, Azerbaijón, Bélgica, Brasil, Bulgaria, Canadá, Chile, China, Colombia, Republica Checa, Dinamarca, Ecuador, Egipto, Estonia, Finlandia, Francia, Alemania, Grecia, Hong Kong, Hungría, India, Indonesia, Irlanda ,Israel, Italia, Japón, Kazajstán, Letonia, Lituania, México, Países Bajos, Nueva Zelanda, Nigeria, Noruega, Pakistán, Perú, Filipinas, Polonia, Portugal, Rumania, Rusia, Arabia Saudita, Singapur, Eslovaquia, Eslovenia, Sudáfrica, España, Sri Lanka, Suecia, Suiza, Taiwán, Tailandia, Turquía, Ucrania, Reino Unido, Estados Unido, Venezuela y Vietnam.

Cuadro 1. Definición de variable y signos esperados.

VARIABLES		Signo Esperado
PIR	Tasa de piratería de software	
P	Precio del software representativo	+
INC	Índice de Incorrupción	-
INF	Índice de Infraestructura	+
CAP	Índice de Capital Humano	-
LEG	Índice de Ambiente Legal	-
R&D	Índice de Investigación y Desarrollo	-
PIB	Ingreso Per cápita medido en PPP	-

Fuente: Elaboración propia.

Es importante adelantar que este modelo base será sometido a las pruebas econométricas estándar. Por lo tanto, es posible que sea necesario realizar algunas modificaciones cuando se lleven a cabo las estimaciones.

Datos

Antes de desarrollar el análisis econométrico propuesto, debemos explicar qué tipo de datos utilizaremos para medir los efectos de cada variable en el modelo. Dado que el modelo propuesto consta de variables reales y *proxy*, es necesario hacer una explicación de las fuentes y las consideraciones que sustentan la utilización de cada variable. El análisis que realizaremos se aplicará a 61 países, debido a la ausencia de datos completos para los demás.

Ingreso. Para medir el ingreso se tomará PIB per cápita valuado en Poder de Paridad de Compra (PPP por sus siglas en inglés). El PIB per cápita en PPP se calcula teniendo en cuenta el producto interno bruto convertido a dólares estadounidenses usando las tasas de poder de paridad de compra y dividido por el total de la población. Esto indica que, un dólar, para una persona, tiene el mismo poder de compra sobre el PIB en relación con el poder de compra de un dólar norteamericano, independientemente de donde viva. Usamos estos datos en lugar del PIB per cápita simple, dado que queremos determinar el efecto del ingreso sobre la demanda, y el PIB per cápita simple no refleja el poder adquisitivo del mismo modo que el PIB per cápita PPP. La serie de datos que usaremos tiene su fuente en el Banco Mundial 2010.

Investigación y Desarrollo. Para medir este indicador se utilizará el índice de Investigación y Desarrollo. Este índice está compuesto por gasto bruto del gobierno en investigación y desarrollo, gasto bruto del sector privado en investigación y desarrollo, cantidad de nuevas patentes nacionales registradas por residente y recepción de regalías y aranceles por licencias. La serie de datos que usaremos es tomada de la BSA en su estudio

“Resistencia en medio de la confusión: Benchmarking la competitividad en la industria de tecnología de la información 2010”.

Precios. La variable que pareciera ser la más importante en el análisis de la piratería es la más difícil de encontrar debido a la falta de un índice de precios del software. Como no se tiene un índice se tomará como precio representativo el precio de software más copiado ilegalmente, en este caso es el MS Office Profesional 2007, en su edición estándar para nuevo usuario. La base fue tomada de la página Microsoft.

Factor de Incurrupción. Para esta variable no hay una medida precisa, es por tanto necesario usar una variable de tipo *proxy*. La variable que se elige es Índice de Transparencia Internacional¹⁰, llamado International Corruption Perceptions Index (Transparency International, 2009). Este indicador establece cuál es la percepción de la sociedad internacional del manejo de los negocios del sector público y privado. El índice establece una calificación dada por expertos y por el público en general con respecto a la corrupción. Es una medida negativa pues la corrupción implica un ambiente de inseguridad jurídica, lo cual desincentiva el desarrollo de contratos a largo plazo, y aumenta los costos de transacción de los individuos y las firmas. Esta situación impide el adecuado funcionamiento de las reglas básicas de la política, de lo social y de lo legal que establecen la base para la producción, el intercambio y la distribución.

¹⁰ Este índice es sugerido por Ronkainen y Guerrero-Cusumano (2001) y fue usado por Marron y Steel (2001) en un estudio similar.

Factor educacional. Para medir la educación usamos el Índice de Capital Humano. El Índice de Capital Humano está compuesto por la cantidad de estudiantes de nivel universitario, inscripción en programas de ciencias de nivel universitario, empleo en el nivel tecnológico y la capacidad del sistema educativo de formar tecnólogos con conocimiento de negocios. Nos indicará aproximadamente el nivel educativo de la población y nos permitirá comparar con otras naciones. El conjunto de datos que usaremos es tomado de la página BSA.

Factor Legal. Para medir el ámbito jurídico se usa el Índice de Ambiente Legal. El Índice de Ambiente Legal está compuesto por la aplicación de la legislación de PI, el estado de la legislación de PI, el estado de las leyes nacionales de privacidad de datos y contra el correo no deseado y contra los delitos electrónicos. El conjunto de datos que usaremos es tomado de la página BSA.

Factores de Infraestructura. Para medir el ámbito de Infraestructura se usa el Índice de Infraestructura de Tecnología de la Información (TI). Este índice está conformado por gasto del mercado en hardware, software y servicios tecnológicos, computadoras de escritorio y portátiles por cada 100 personas, conexiones de banda ancha por cada 100 personas, servidores de internet seguros por cada 100,000 personas y la penetración de telefonía móvil por cada 100 personas. El conjunto de datos que usaremos es tomado de la página BSA.

Como variable dependiente de este estudio he decidido usar la tasa de violación a los DPI derivados de la creación de Software. Esta tasa nos dice en qué porcentaje de unidades

de software ilícito, hay del total de software instalados. El conjunto de datos que usaremos es tomado de la página BSA. La estimación de la tasa hecha por la BSA es la más confiable conocida sobre violación a los Derechos de Propiedad Intelectual derivados del software. Su estimación se basa en la comparación de la instalación de software nuevo frente a las ventas de software legítimo. Para estimar la instalación de software, los grupos de distribuidores de software recogen datos respecto de la cantidad de computadores personales vendidos a usuarios, sea para el hogar o para otro tipo de destino. Una fracción de éstos es destinada a reemplazar computadores viejos, y dicha tasa es también estimada. Así, usando los estimados de nuevos computadores menos los que se dirigen a reemplazar computadores antiguos, determinan la cantidad de instalaciones de nuevo software. Para estimar las ventas de software legítimo, los grupos de distribución usan datos referentes al envío de software, desagregados por país y por el tipo de aplicación. Teniendo lo anterior en cuenta, el número de copias piratas en cada país es entonces calculada como la diferencia entre el número de instalaciones de software y la cantidad de software legítimamente vendida, de modo que la tasa de piratería de software es la razón de copias piratas frente a las instalaciones de software.

Análisis de los Datos y Estimaciones

El cuadro 2 muestra los estadísticos descriptivos de la base de datos para un total de 61 observaciones. La Tasa de Piratería más alta la tiene Sri Lanka con 89%. En este país, como sugiere la teoría económica, el precio del software es muy alto con relación al ingreso medio. El software costaría aproximadamente el 16% del ingreso medio anual de las personas de ese país.

El precio más alto se encuentra en Ucrania (1142.6 dólares), cuya tasa de piratería se encuentra cerca de la media. Sin embargo, el precio más alto como proporción del ingreso se encuentra en Pakistán (40%).

Por otra parte, los países con los índices más altos en cuanto a incorrupción (corrupción más baja) son Nueva Zelanda y Dinamarca, cuyas tasas de piraterías son bajas. En contraste, Ecuador y Colombia son los países con índice de ambiente legal más bajo y tasa de piratería en la media.

Cuadro 2. Sumario de Estadísticas Descriptivas.

Variable	Observaciones	Media	Desviación Estándar	Mínimo	Máximo
PIR	61	51.885	21.487	20	89
P	61	518.851	235.923	196.798	1142.259
INC	61	52.967	24.017	19	94
INF	61	41.421	28.580	1.9	93.8
CAP	61	40.484	12.410	16.1	75.6
LEG	61	23.452	19.515	0.4	64.2
RD	61	65.123	16.648	33	92
PIB	61	22296.83	14272.99	2203.348	55672.12

Fuente: Elaboración propia con datos de BSA, Banco Mundial, Transparencia Internacional y Microsoft.

El cuadro 3 muestra la matriz de correlaciones de las variables que se emplean en este trabajo. En términos generales todas las variables tienen una alta correlación. Entre las mayores correlaciones están PIR y RD con -.901. Otras variables con correlación superiores a .8 son INC, INF y PIB. Otras variables con alta correlación son las que conforman los siguientes pares: (INF, INC) (RB, INC) (INF, RD) y (PIB, INF).

Cuadro 3. Matriz de Correlaciones.

Abreviación	PIR	INC	INF	CAP	LEG	RD	PIB	P	DES
PIR	1.000								
INC	-0.886	1.000							
INF	-0.890	0.930	1.000						
CAP	-0.707	0.654	0.714	1.000					
LEG	-0.711	0.733	0.787	0.711	1.000				
RD	-0.901	0.920	0.925	0.756	0.717	1.000			
PIB	-0.854	0.885	0.932	0.684	0.738	0.892	1.000		
P	0.784	-0.761	-0.798	-0.518	-0.505	-0.742	-0.815	1.000	

Fuente: Elaboración propia con datos de BSA, Banco Mundial, Transparencia Internacional y Microsoft.

La multicolinealidad es una situación que puede presentarse cuando existe una fuerte correlación entre variables explicativas del modelo. Se dice que hay un problema cuando una o más variables son una combinación lineal de otra. Para ver si hay problemas de multicolinealidad de las variables se utilizó el factor de inflación de la varianza (FIV) y la tolerancia (T). De acuerdo con estos criterios, las variables INC, INF, RD y PIB pueden ocasionar problemas de multicolinealidad. Para solucionar este problema se puede optar por juntar o eliminar algunas de estas variables. (En este caso se eliminara INF y RD)

Otro problema que encontramos es la heterocedasticidad. Es decir, la varianza del error no es constante a lo largo de las observaciones. Esto implica el incumplimiento de una de las hipótesis básicas sobre las que se asienta el modelo de regresión lineal múltiple. El

comportamiento no constante de las perturbaciones causa que el estimador mínimo cuadrático deje de tener varianza mínima, llevando con ello a que las pruebas estadísticas utilizadas pierdan validez.

Se utilizaron las pruebas de White y el multiplicador de LaGrange de Breusch-Pagan para detectar si hay problemas de heterocedasticidad. Las pruebas demostraron que hay problemas de heterocedasticidad. Por lo tanto, se procedió a resolver estos problemas, por la corrección de White.

Cuadro 4. Modelos explicativos de la piratería.

Variable	Modelo 1	Modelo 2
INC	-0.572 *** 0.069	-0.443 *** 0.116
CAP	-0.287 * 0.154	-0.314 ** 0.134
LEG	-0.044 0.080	-0.065 0.076
COST	46.141 *** 15.906	
P		0.022 ** 0.010
PIB		-0.000 0.000
Cons	92.022 *** 6.357	79.791 *** 11.034
R ²	0.834	0.842
F	97.1	92.11
***Significativo al 99%; **Significativo al 95%; *Significativo al 90%		
Entre paréntesis se encuentra el Error Estándar.		

Fuente: Elaboración propia con datos de BSA, Banco Mundial, Transparencia Internacional y Microsoft.

En el cuadro 4 se presentan los resultados de las estimaciones de los modelos empíricos sobre la tasa de piratería. A continuación se explica la diferencia entre los dos modelos estimados. En el primer modelo se incluye la variable COST, mientras que el segundo incluye P y PIB. La variable COST es el precio del software como proporción del PIB per cápita del país.

Analizando los resultados se puede señalar que la capacidad de los modelos para explicar la variabilidad de la piratería (R^2) se aproxima al 84%.

Cuadro 5. Signos del modelo

Variable	Signo Esperado	Resultados Modelo 1	Resultados Modelo 2
INC	-	-	-
CAP	-	-	-
LEG	-	0	0
COST	+	+	
P	+		+
PIB	-		0

Fuente: Elaboración propia con datos de BSA, Banco Mundial, Transparencia Internacional y Microsoft.

En términos generales, como se aprecia en el cuadro 5, los signos de las regresiones son los esperados. Una mayor incorrupción y educación del país reducen la tasa de piratería. Por el contrario, a mayor precio del software o mayor precio como proporción del PIB, aumentan la tasa de piratería. El PIB y Ambiente Legal no son significativas. Por lo tanto,

en estricto sentido estadístico, no podemos decir que esta variable tenga efecto sobre la piratería.

Las variables que resultaron significativas y que tienen el signo esperado son los siguientes: índice de incorrupción, índice de capital humano, así como COST y precio, en el modelo respectivo.

Con respecto a la magnitud de influencia de las distintas variables en la tasa de piratería, se encuentra lo siguiente. Con base en el Modelo 1, un incremento de 5 puntos en el índice INC hace que la piratería disminuya en 2.86%. Un incremento similar en el índice CAP hace que la piratería disminuya en 1.35%. Si el ingreso per cápita se dobla o el precio del software baja a la mitad, la piratería disminuye en 1.06%. En el Modelo 2 un aumento de 5 puntos en el Índice de INC y CAP, resultaría con una disminución 2.2% y 1.57%, respectivamente. Si el precio del software disminuyera en \$100 dólares, la piratería disminuiría en 2.2%. Con esto concluimos que los mayores cambios están dados por INC y el precio del software.

Es interesante notar que el marco legal, entendido como el alcance y transparencia de la legislación de propiedad intelectual, tiene efecto nulo sobre la tasa de piratería.

Los resultados en conjunto indican que la incorrupción y los precios son más importantes que la legislación para influir sobre la tasa de piratería. En este sentido, se puede decir que es poco relevante tener marco legal adecuado cuando las leyes no se hacen respetar.

6.- Conclusión

Estudios recientes han señalado que los factores que explican los diferentes niveles de tasas de piratería de software en distintos países del mundo han sido el ingreso per cápita, precio del producto (factores de mercado) así como factores institucionales y culturales (de no mercado o complementarias).

En la última década hemos visto cómo las tasas de piratería mundial han mostrado un comportamiento cíclico, en el cual del 2002 al 2003 la tasa de piratería descendió de un 40% al 35%, y los tres años posteriores se mantuvo en ese nivel; sin embargo, del 2006 en adelante se nota un repunte que llega a niveles cercanos al 42%. Entre los motivos que se han señalado de este repunte ha sido un auge del uso de internet en los países desarrollados y las diferentes políticas de propiedad intelectual.

En particular en este trabajo de investigación nos interesó determinar cómo se ha afectado la tasa de violación a los Derechos de Propiedad Intelectual (DPI) derivados de la creación del software, explicado por factores de mercado y no mercado (complementarias), para una muestra de 61 países del mundo con datos de corte transversal para el 2009.

Se usó el factor de infraestructura como una *proxi* del uso de internet, mientras que de las políticas de propiedad intelectual hemos usado variables como factor de incorrupción y un índice de ambiente legal.

En términos generales, según los resultados econométricos, se puede señalar que tanto los factores de mercado como las variables complementarias son relevantes para determinar el comportamiento de la tasa de piratería para esta muestra de países.

Se probó la robustez de los resultados utilizando dos modelos: en el primero se incluyen todas las variables de mercado y complementarias, excepto las variables Precio y PIB (Ingreso Per cápita). Para el segundo modelo se incluyeron dichas variables, pero se quitó la variable Costo (precio del software medido como proporción del PIB per cápita).

De acuerdo con los resultados, todas las variables mantienen los signos esperados, aunque no todas resultaron significativas. Por una parte, un precio elevado del software o bien un costo elevado del mismo, provoca tener más incentivos a copiar software legales, lo que conduce a una mayor tasa de piratería. Por otra parte, un alto índice de incorrupción y mayores niveles de educación de un país, se asocian con menores tasas de piratería. Sin embargo, lo que parece no afectar en nada el comportamiento de la piratería es el marco legal o el Ingreso, ya que estas variables no resultaron significativas.

Como una primera conclusión se puede mencionar que los resultados en conjunto indican que la incorrupción y los precios son más importantes que la legislación para influir sobre la tasa de piratería. En este sentido, se puede decir que es poco relevante tener un marco legal adecuado cuando las leyes no se hacen respetar. Otra consecuencia importante que se desprende de estos resultados es que el nivel educativo con el que cuente la población repercute en la tasa de piratería del país y si el costo como proporción del ingreso per cápita es elevado, habrá incentivos para la piratería.

De esta manera, las políticas económicas dirigidas a reducir los niveles de piratería deben estar encaminadas a reducir los precios del software, alentado a una mayor competencia del mercado. Además las autoridades deben orientar políticas que reduzcan la inseguridad jurídica y los costos de transacción entre individuos y empresas.

Como una de las limitantes del presente trabajo, tenemos, en primer lugar que, los datos de la tasa de piratería pudieran ser criticables, en el sentido de que su fuente es una organización de productores de software, por lo que podría existir cierto sesgo en sus reportes de niveles de piratería. Otra de las limitaciones es que se cuenta con pocas observaciones, debido a que en ciertos países no se reportan algunas variables para construir los índices que se utilizaron.

7.- Bibliografía

- Arrow, K. 1962. “Economic Welfare and the Allocation of Resources for Invention”, *The rate and direction of inventive activity: Economic and social Factors*. Princeton, N, J. National Bureau of Economic Research. p: 609-626.
- Becker, L. 1993. “Deserving to Own Intellectual Property”, *Chicago-Kent Law Review*. p: 609-629.
- Burke, A. 1996. “How Effective are International Copyright Conventions in the Music Industry?”, *Journal of Cultural Economics*. p: 51– 66.
- Business Software Alliance. 2009. Seventh Annual BSA and IDC Global Software Piracy Study; Consultado: 01/01/2013 de www.bsa.org/usa/antipiracy/internet/.
- Coase, R. 1993. “The Institutional Structure of Production”, *American Economic Review*. p: 713-719.
- Dasgupta, P. y Stiglitz, J. 1980. “Industrial Structure and the Nature of Innovative Activity”, *Economic Journal*. p: 266-293.
- El problema de la colinealidad. 2010. Consultado: 01/01/2013 de www.uv.es/uriel/material/multicolinealidad3.pdf
- Ginarte, J. y Park, W. 1997. “Determinants of Patent Rights: A Cross - National Study”, *Research Policy*. p: 283-301.
- Gopal R. y Sanders L. 1998. “Internacional Software Piracy; Analysis of Key Issues and Impacts”, *Information Systems Research*. p: 380-398.
- Gopal, R. y Sanders L. 2000. “Global Software Piracy: You Can’t Get Blood Out of a Turnip”, *Communication of the ACM*. p: 82-89.

Hofstede, G. 1983 “Dimensions of National Cultures in Fifty Countries and Three Regions”.

Explications in cross-cultural psychology. Editado por: Deregowski, S. et. al.

Netherlands : Swets and Zeitlinger. p: 335-355.

Husted, B. 2000. “The Impact of National Culture on Software Piracy”, *Journal of*

Business Ethics. p: 197-211.

Johnson, W. 1985. “The Economics of Copying”, *Journal of Political Economy*.

p: 158-174.

Lakhani, K. y Robert G. 2005. “Why Hackers Do What They Do: Understanding

Motivation Effort in Free/Open Source Software Projects”, In Feller, J., Fitzgerald,

B., Hissan, S. and Lakhani, K.R. (eds), *Perspectives on Free and Open Source*

Software. Cambridge, MA: MIT Press. p: 3–21

Landes W. y Posner R. 1989. “An Economic Analysis of Copyright Law”, *Journal of Legal*

Studies. p: 325-353

Márquez P. 2003, “Violación a los Derechos de Propiedad Intelectual en América Latina: El

Caso de la Piratería de Software”. Departamento de Economía, Universidad Javeriana,

Bogotá.

Marron D. y Steel D. 2000. “Which Countries Protect Intellectual Property? The Case of

Software Piracy”, *Economic Inquiry*. p: 159–174.

McLeod K. 2001 "Owning Culture: Authorship, Ownership and Intellectual Property Law",

New York: Peter Lang Publishing, Inc.

Moore, T. y Gurpreet D. 2000. “Software Piracy: A View from Hong Kong”,

Communication of the ACM. p: 88–93.

Nelson R., Peck M. y Kalachek E. 1967. “Technology, Economic Growth, and Public

Policy”, *Journal of Finance*. p:703-704.

- Novos, I. y Michael W. 1984. "The Effects of Increased Copyright Protection: An Analytic Approach", *Journal of Political Economy*. p: 236-246.
- Peace A., Galleta D. y Thong J. 2003. "Software Piracy in the Workplace: A Model and Empirical Test", *Journal of Management Information Systems*. p: 153-177
- Rapp, R. y Rozek, R. 1990. "Benefits and Costs of Intellectual Property Protection in Developing Countries", *Journal of World Trade*. p: 75-102.
- Robinson, J. 1933. "The Economics of Imperfect Competition", MacMillan: Londres.
- Rodriguez A. 2002. "The European Software Piracy: An Empirical Application." Mimeo. Department of Economics, University of Southern Denmark.
- Ronkainen I. and Guerrero J. 2001. "Correlates of Intellectual Property Violations", *Multinational Business Review*. p: 59-65.
- Seale D., Polakowsky M. y Schneider S. 1998. "It's not Really Theft. Personal and Workplace Ethics that Enable Software Piracy", *Behavior & Information Technology*. p: 27-40.
- Simpson P., Banerjee D. Y Simpson, C. 1994. "Softlifting: A Model of Motivating Factors", *Journal of Business Ethics*. p: 431-438.
- Swinyard W., Rinne H. y Keng A. 1990. "The Morality of Software Piracy: A Cross Cultural Analysis", *Journal of Business Ethics*. p: 655-664.
- Tan B. 2002. "Understanding Consumer Ethical Decision Making with Respect to Purchase of Pirated Software", *Journal of Consumer Marketing*. p: 96-111.
- Tilak, J. 1993. "Education and Agricultural Productivity in Asia: A Review", *Indian Journal of Agricultural Economic*. p: 187-200.

8.- Anexo

Posibles problemas y técnicas de tratamiento.

Los datos utilizados serán de corte transversal. Estos se recolectan en un momento del tiempo con base en una o algunas variables que se obtienen como muestras de poblaciones específicas. Estos datos representan gran diversidad de características socioeconómicas y a diferencia de los estudios con series de tiempo, el análisis de corte transversal en este caso proporciona evidencia de la estructura, para hacer inferencia a un nivel microeconómico, regularmente para una población definida de forma específica. Dos posibles problemas que estos datos presentan son la colinealidad y perturbaciones heterocedásticas.

La colinealidad implica que una variable X_1 sea combinación lineal de otra X_2 , significando que ambas están relacionadas por la expresión $X_1 = \beta_1 + \beta_2 X_2$, siendo β_1 y β_2 constantes, por lo tanto el coeficiente de correlación entre ambas variables será 1.

Del mismo modo, que una variable X_1 sea combinación lineal de otra X_2, \dots, X_i con $i > 2$, significa que dichas variables están relacionadas por la expresión $X_1 = \beta_1 + \beta_2 X_2 + \dots + \beta_i X_i$, siendo β_1, \dots, β_i constante y, por tanto, el coeficiente de correlación múltiple $R_{X_1/X_2, \dots, X_i}$ también será 1.

Otro modo de definir la existencia de colinealidad es cuando alguno de los coeficientes de correlación simple o múltiples entre algunas de las variables independientes es 1, es decir, cuando algunas variables independientes están correlacionadas entre sí.

En la práctica, esta colinealidad exacta raras veces ocurre, pero sí surge con cierta frecuencia la llamada *casi-colinealidad*, o por extensión, simplemente colinealidad en que alguna variable es casi-combinación lineal de otra u otras, o dicho de otro modo, algunos coeficientes de correlación simple o múltiple entre las variables independientes están cercanos a 1, aunque no llegan a dicho valor.

En este caso la matriz $X'X$ es casi-singular, o sea su determinante no es cero pero es muy pequeño. Como para invertir una matriz hay que dividir por su determinante, en esta situación surgen problemas de precisión en la estimación de los coeficientes, ya que los algoritmos de inversión de matrices pierden precisión al tener que dividir por un número muy pequeño, siendo además inestable.

Además, como la matriz de varianzas de los estimadores es proporcional a $X'X$, resulta que en presencia de colinealidad los errores estándar de los coeficientes son grandes (hay imprecisión también en sentido estadístico). Por consiguiente, a la hora de plantear modelos conviene estudiar previamente la existencia de casi-colinealidad (la colinealidad exacta no es necesario estudiarla previamente, ya que todos los algoritmos la detectan, de hecho no se puede efectuar la estimación). Como medida de la misma hay varios estadísticos propuestos, los más sencillos son los coeficientes de determinación de cada variable independiente con todas las demás, es decir

$$R_i^2 = R_{X_i | X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_k}^2 \quad i = 1, \dots, k$$

y, relacionados con ellos, el factor de inflación de la varianza (FIV) y la tolerancia (T), definidos como

$$FIV_i = \frac{1}{1 - R_i^2} \quad T_i = \frac{1}{FIV_i} = 1 - R_i^2$$

Una regla empírica, citada por Kleinbaum, consiste en considerar que existen problemas de colinealidad si algún FIV es superior a 10, que corresponde a algún R_i^2 0,9 y $T_i < 0,1$.

Aunque puede existir colinealidad con FIV bajos, además puede haber colinealidades que no impliquen a todas las variables independientes y que, por tanto, no son bien detectadas por el FIV .

En principio, el problema de la multicolinealidad está relacionado con deficiencias en la información muestral. Sin embargo, la aproximación cuantitativa a los conceptos teóricos puede ser inadecuada, haciendo que en el término de perturbación se absorban errores de especificación. Veamos a continuación algunas de las soluciones propuestas para resolver el problema de la multicolinealidad.

Eliminación de variables

La multicolinealidad puede atenuarse si se eliminan los regresores que son más afectados por la multicolinealidad. El problema que plantea esta solución es que los estimadores del nuevo modelo serán sesgados en el caso de que el modelo original fuera el correcto. Sobre esta cuestión conviene hacer la siguiente reflexión. El investigador está

interesado en que un estimador sea preciso (es decir, que no tenga sesgo o que éste sea muy pequeño) y con una varianza reducida. El error cuadrático medio (*ECM*) recoge ambos tipos de factores. Así para el estimador $\hat{\beta}_j$, el *ECM* se define de la siguiente manera:

$$ECM(\hat{\beta}_j) = \left[\text{sesgo}(\hat{\beta}_j) \right]^2 + \text{Var}(\hat{\beta}_j)$$

Si un regresor es eliminado del modelo, el estimador de un regresor que se mantiene (por ejemplo, $\hat{\beta}_j$) será sesgado, sin embargo, su *ECM* puede ser menor que el correspondiente al modelo original, debido a que la omisión de una variable puede hacer disminuir suficientemente la varianza del estimador. En resumen, aunque la eliminación de una variable no es una práctica que en principio sea aconsejable, en ciertas circunstancias puede tener su justificación cuando contribuye a disminuir el *ECM*.

Aumento del tamaño de la muestra teniendo en cuenta que un cierto grado de multicolinealidad acarrea problemas cuando aumenta ostensiblemente la varianza muestral de los estimadores, las soluciones deben ir encaminadas a reducir esta varianza. Existen dos vías: por un lado, se puede aumentar la variabilidad a lo largo de la muestra de los regresores colineales introduciendo observaciones adicionales. Esta solución no siempre es viable, puesto que los datos utilizados en las contrastaciones empíricas proceden generalmente de fuentes estadísticas diversas, interviniendo en contadas ocasiones el investigador en la recogida de información. Por otro lado, cuando se trate de diseños experimentales, se podrá incrementar directamente la variabilidad de los regresores sin necesidad de incrementar el tamaño de la muestra. Finalmente, conviene no olvidar que el término de perturbación no debe contener ningún factor que sea realmente relevante para la explicación de las

variaciones del regresando, con el fin de reducir todo lo posible la varianza del término de perturbación.

Otra posibilidad es la utilización de información extramuestral, bien estableciendo restricciones sobre los parámetros del modelo, o bien aprovechando los estimadores procedentes de otros estudios. El establecimiento de restricciones sobre los parámetros del modelo reduce el número de parámetros a estimar y, por tanto, valida las posibles deficiencias de la información muestral. En cualquier caso, para que estas restricciones sean útiles deben estar inspiradas en el propio modelo teórico o, al menos, tener un significado económico. En general, un inconveniente de esta forma de proceder es que el significado atribuible al estimador obtenido con datos de corte transversal es muy diferente del obtenido con datos temporales. A veces, estos estimadores pueden resultar realmente «extraños» o ajenos al objeto de estudio. Por otra parte, al estimar las varianzas de los estimadores obtenidos en la segunda regresión hay que tener en cuenta la estimación previa.

Por otro lado las perturbaciones heterocedásticas pueden tener diferentes orígenes. Uno de ellos ocurre cuando los individuos, empresas o unidades económicas de los cuales se toma la muestra no presentan un comportamiento homogéneo.

La heterocedasticidad significa que la varianza de las perturbaciones no es constante a lo largo de las observaciones y supone una violación de uno de los supuestos sobre los que se sienta el modelo de regresión lineal.

En un plano puramente analítico, la matriz de varianzas-covarianzas de las perturbaciones de un modelo heterocedástico se representaría del siguiente modo:

$$E(UU') = \begin{bmatrix} E(u_1)^2 & & & \\ E(u_1u_2) & E(u_2)^2 & & \\ & & \dots & \\ E(u_1u_n) & E(u_2u_n) & \dots & E(u_n)^2 \end{bmatrix} = \begin{bmatrix} E(u_1)^2 & 0 & \dots & 0 \\ 0 & E(u_2)^2 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & E(u_n)^2 \end{bmatrix} \neq \sigma_i^2 I_n = \sigma^2 \Sigma$$

La heterocedasticidad implica la heterogeneidad de los datos al asumir que provienen de distribuciones de probabilidad con distintas varianzas, es decir que no presentan el supuesto de homocedasticidad donde la varianza de cada término de perturbación u_i condicional con los valores escogidos de las variables explicativas es un numero constante igual a σ^2 .

Bajo condiciones heterocedásticas no se pierden las propiedades de insesgamiento y consistencia de los estimadores mínimos cuadrados ordinarios. Sin embargo, estos estimadores no poseen varianza mínima, por lo tanto no son eficientes y dejan de ser los Mejores Estimadores Linealmente Insesgados (Gujarati, 1992). Existen diferentes pruebas para detectar la heterocedasticidad como la prueba de heterocedasticidad general de White, la prueba de multiplicadores de LaGrange de Breusch-Pagan y la prueba de Goldfeld-Quandt, entre otras. En el caso de este análisis tomaremos como base las dos primeras para determinar la presencia de este problema en los datos.

H. White desarrolla en 1980 un método para obtener estimadores consistentes de las varianzas y covarianzas de los estimadores de mínimos cuadrados ordinarios e incluye una prueba directa de heterocedasticidad bajo la hipótesis:

$H_0: \sigma_i^2 = \sigma^2$ para todo i .

H_a : no se verifique la H_0

La forma de realizar la prueba se basa en estimar la regresión de los residuales al cuadrado, que constituyen el indicativo de las varianzas de las perturbaciones, frente a un término independiente, los regresores, sus cuadrados y productos cruzados.

Para interpretar la prueba observamos el poder explicativo de las variables incluidas en la regresión sobre los errores al cuadrado. Para que las perturbaciones resulten homocedásticas, no debe existir tal poder explicativo y en consecuencia el coeficiente de determinación R^2 de la estimación debiera ser pequeño.

Así, se rechazará la hipótesis nula de homocedasticidad cuando la probabilidad de rechazar la hipótesis nula siendo cierta sea menor al 5%, dado que por consecuencia el valor muestral del estadístico es suficientemente alto.

La segunda prueba a utilizar, el contraste de LaGrange de Breusch-Pagan, consiste en comprobar si se puede encontrar un vector de variables z que sirvan para explicar la evolución de la varianza de las perturbaciones, estimada a partir del cuadrado de los errores del modelo inicial sobre el que se busca probar heterocedasticidad. Para realizar esto, se obtienen los errores de la estimación del modelo inicial. Posteriormente se calcula una serie de errores al cuadrado estandarizados del modelo estimado. Se estima la regresión con el error calculado explicado por una constante y el conjunto de variables z , y se obtienen el coeficiente de determinación y la varianza estimada. El contraste propuesto en el caso de un modelo homocedástico, se distribuye como una χ^2 con “ r ” grado de libertad igual al conjunto de las variables z y un determinado nivel de significancia, con lo que, si el valor del

estadístico de prueba calculado supera el valor de tablas, se rechaza la hipótesis nula verificando que el modelo presenta heterocedasticidad.

En el caso de detectarse problemas por heterocedasticidad se requieren transformar la matriz de varianza covarianza. White prueba que, bajo ciertas condiciones generales, la matriz:

$$S_0 = (1/2) \sum_{i=1}^n e_i^2 x_i x_i'$$

Donde el i -ésimo residuo mínimo cuadrático e_i es un estimador consistente de Σ , por lo tanto el estimador de White :

$$\text{VarEst}(b) = n(X'X)^{-1} S_0 (X'X)^{-1}$$

Puede usarse como un estimador de la verdadera varianza del estimador mínimos cuadrático (Green, 2003).

De aquí que se pueda estimar por medio de mínimos cuadrados ordinarios este modelo transformado con la matriz varianza covarianza de las perturbaciones.