

UNIVERSIDAD DE CONCEPCIÓN

FACULTAD DE INGENIERÍA

DEPARTAMENTO DE INGENIERÍA CIVIL INDUSTRIAL



Profesor Patrocinante:

**PhD. Pablo Catalán
Martínez**

Informe de Memoria de Título

para optar al título de:

Ingeniero Civil Industrial

Identificación del efecto de capacidades
científicas, tecnológicas y productivas en la
generación y desarrollo de nuevos productos

UNIVERSIDAD DE CONCEPCIÓN

Facultad de Ingeniería

Departamento de Ingeniería Civil Industrial

Profesor Patrocinante:

Ph.D. Pablo Catalán Martínez

Identificación del efecto de capacidades científicas, tecnológicas y productivas en la generación y desarrollo de nuevos productos

Nicolás Sebastián Pérez Mundaca

Informe de Memoria de Título
para optar al Título de

Ingeniero Civil Industrial

Agosto de 2019

Resumen

La presente investigación realiza un análisis de los efectos que poseen las capacidades científicas, tecnológicas y productivas con respecto a la generación y desarrollo de nuevos productos mediante la estimación de un modelo de regresión lineal múltiple a través del método de mínimos cuadrados ordinarios (MCO). Para realizar este modelo, es construida una base de datos de publicaciones científicas, patentes y exportación de productos provenientes de Web of Science, United States Patent and Trademark Office (USPTO) y el OEC del MIT Media Lab, quienes utilizan los datos de la United Nations International Trade Statistics Database (UN Comtrade), respectivamente, abarcando un total de 254 categorías en publicaciones científicas, 661 tipos de patentes y 782 categorías de productos para un total de 120 países en el periodo comprendido entre 1988 y el 2017.

De esta forma, se observa que las capacidades científicas, tecnológicas y productivas afectan positivamente la generación de ventajas comparativas en el desarrollo de productos, siendo la más relevante la relación entre la cartera productiva del país, seguido en menor medida por las capacidades tecnológicas y científicas del mismo. En base a estos resultados, se sugieren políticas públicas que fomenten el desarrollo y la innovación de ecosistemas productivos con la finalidad de aumentar las relaciones existentes entre los productos pertenecientes al portafolio productivo del país y así poseer una mayor cantidad de oportunidades de diversificación. Además, debido a la gran influencia que poseen las capacidades productivas por sobre las científicas y tecnológicas, es que las políticas públicas que deseen incentivar la generación de nuevos productos a través del desarrollo de ciencia y tecnología deben ser disruptivas, de manera de desarrollar ecosistemas productivos que fomenten la generación de un set de productos en específico con la finalidad de romper la fuerte influencia de las actuales capacidades productivas y generar así un mayor abanico de oportunidades de diversificación.

Tabla de Contenido

CAPÍTULO 1:	INTRODUCCIÓN	1
1.1.	OBJETIVOS DEL TEMA	1
1.1.1.	<i>Objetivo General</i>	1
1.1.2.	<i>Objetivos Específicos</i>	1
1.2.	JUSTIFICACIÓN DEL TEMA	2
1.3.	ALCANCE Y LIMITACIONES	2
1.4.	METODOLOGÍA	2
1.5.	ESTRUCTURA DE LA MEMORIA DE TÍTULO	3
CAPÍTULO 2:	MARCO TEÓRICO	4
CAPÍTULO 3:	METODOLOGÍA	10
3.1.	INDICADORES	10
3.1.1.	<i>Revealed Comparative Advantage (RCA)</i>	10
3.1.2.	<i>Proximidad</i>	11
3.1.3.	<i>Proximidad Cruzada</i>	12
3.2.	FUENTES DE DATOS	12
3.2.1.	<i>Publicaciones Científicas</i>	13
3.2.1.1.	Descarga de Publicaciones Científicas	13
3.2.1.2.	Preprocesamiento de Publicaciones Científicas	17
3.2.2.	<i>Patentes</i>	18
3.2.2.1.	Descarga de Patentes	20
3.2.2.2.	Preprocesamiento de Patentes	23
3.2.3.	<i>Productos</i>	24
3.2.3.1.	Descarga de Productos	25
3.3.	INTERVALOS DE ESTUDIO	26
3.4.	SELECCIÓN DE PAÍSES	26
3.5.	MODELO ECONÓMETRICO	27
3.5.1.	<i>Modelos de Regresión Lineal Múltiple (RLM)</i>	27
3.5.2.	<i>Estimación de Modelos RLM: Método de Mínimos Cuadrados Ordinarios (MCO)</i>	28
3.5.3.	<i>Variables del Modelo</i>	30
3.5.3.1.	Variable Dependiente	30
3.5.3.2.	Variables Explicativas	30
3.5.3.2.1.	Ventajas Comparativas	31
3.5.3.2.2.	Densidad Productiva	31
3.5.3.2.3.	Densidades Cruzadas	32
3.5.3.2.4.	Factores Fijos	34
3.6.	CONSTRUCCIÓN DEL MODELO DE REGRESIÓN	34
CAPÍTULO 4:	RESULTADOS Y ANÁLISIS	37
4.1.	MODELOS ESTADÍSTICOS	37
4.2.	MODELO DE EFECTOS FIJOS	38
4.3.	IMPPLICANCIA EN POLÍTICAS PÚBLICAS	40
CAPÍTULO 5:	CONCLUSIONES	42
CAPÍTULO 6:	BIBLIOGRAFÍA	44
ANEXO A.	LISTADO DE PAÍSES CON FORMATO ISO 3166-2	46

Lista de Tablas

Tabla 3.1: Campos seleccionados del análisis de resultados de <i>WoS</i>	16
Tabla 3.2: Ejemplo de jerarquización de patentes en el estándar <i>CPC</i> para el subgrupo E04D1/20.	19
Tabla 3.3: Secciones del estándar <i>CPC</i>	20
Tabla 3.4: Descripción de campos seleccionados de PatentsView	20
Tabla 3.5: Clasificación principal de la SITC, Rev. 4	25
Tabla 3.6: Descripción de los campos seleccionados del OEC.....	25
Tabla 3.7: Resumen de las bases de datos	26
Tabla 4.1: Análisis estadístico en el periodo 1988 – 2017	38

Lista de Figuras

Figura 3.1: Consulta realizada en búsqueda avanzada para el Ejemplo A en WoS.....	14
Figura 3.2: Ventana de Resultados de WoS.....	14
Figura 3.3: Análisis de resultados para la consulta en WoS.....	15
Figura 3.4: Formato de descarga de datos desde WoS.....	15
Figura 3.5: Dimensiones seleccionadas para eliminar sobreestimaciones.....	17
Figura 3.6: Formato de descarga de datos para eliminar sobreestimaciones.....	17
Figura 3.7: Formato de base de datos de publicaciones científicas	18
Figura 3.8: Diferencias entre <i>USPC</i> , <i>IPC</i> y <i>CPC</i>	19
Figura 3.9: Diagrama de flujo de algoritmo de descarga de datos	22
Figura 3.10: Formato de descarga de datos desde PatentsView.....	23
Figura 3.11: Formato de base de datos de patentes	24
Figura 3.12: Formato de descarga de los datos del OEC	26

Capítulo 1: Introducción

La economía geográfica, mediante la consolidación del área de *estudios espaciales*, ha contribuido a la discusión sobre desarrollo económico en función de su propuesta de representación y conceptualización de las capacidades endógenas territoriales. En forma más reciente, ha centrado su acción en determinar las causas de las brechas registradas entre países en lo referente a diversificación científica, tecnológica y de producto en consideración las capacidades endógenas registradas a nivel territorial. (Hidalgo, Klinger, Barabási, & Hausmann, 2007).

La conceptualización propuesta se sustenta en procesos de co-ocurrencia que derivan en espacios de representación en red de posibles interacciones y niveles de cercanía entre productos *Product Space* (Hidalgo et al., 2007), tecnologías *Technology Space* (Boschma, Balland, & Kogler, 2014) o áreas de conocimiento *Science Space* (Boschma, Heimeriks, & Balland, 2014). La metodología propuesta permite determinar los niveles de complejidad territorial, en base a consideraciones de diversidad y ubicuidad, además de calcular el efecto de portafolios de productos, tecnologías y áreas de conocimiento en la probabilidad de diversificación. Ahora bien, al día de hoy las estimaciones no consideran incidencias cruzadas, de manera de revisar, por ejemplo, el efecto de portafolios tecnológicos o científicos en procesos de diversificación de producto. De esta forma, las conclusiones emanadas de la presente Memoria de Título (MT) contribuyen a una mejor comprensión de los efectos de capacidades endógenas en procesos de diversificación de producto y de desarrollo económico.

1.1. Objetivos del Tema

1.1.1. Objetivo General

Identificar y determinar el efecto de las capacidades científicas, tecnológicas y productivas en procesos de diversificación de productos a nivel país.

1.1.2. Objetivos Específicos

- Construir bases de datos de publicaciones científicas, patentes y productos.
- Diseñar y testear un modelo econométrico que permita capturar el efecto de las relaciones ciencia-producto y tecnología-producto en el desarrollo de nuevos productos o industrias a nivel país.
- Generar algoritmos computacionales que permitan el procesamiento de altos volúmenes de datos, según las herramientas econométricas estudiadas.
- Determinar implicancias en políticas públicas en consideración a resultados obtenidos.

1.2. Justificación del tema

La realización de la presente Memoria de Título (MT) busca contribuir a la comprensión del efecto de capacidades científicas y tecnológicas en procesos de diversificación de producto a nivel país. En función de los resultados obtenidos, se describirán posibles implicancias de política pública de manera de profundizar la discusión realizada al día de hoy en la temática en cuestión.

1.3. Alcance y Limitaciones

La presente MT se limita a analizar las publicaciones científicas publicadas en *Web of Science*, las patentes concedidas por *United States Patent and Trademark Office (USPTO)* y el nivel de exportación total de productos obtenidos desde el *Observatory of Economic Complexity (OEC)*, el cuál utiliza los datos de la *United Nations Trade Statistics (UN Comtrade)*, descartando así el resto de oficinas gubernamentales y privadas que se dediquen a la publicación de artículos científicos, registros de patentes o de exportaciones a nivel internacional. Debido a la importancia que poseen las entidades antes mencionadas, las conclusiones de la presente Memoria de Título pueden ser consideradas como un indicador válido para analizar y comprender las relaciones entre los portafolios científicos, tecnológicos y productivos de los países y su capacidad para generar ventajas comparativas en la exportación de una categoría de producto.

Las publicaciones científicas y patentes pueden ser asignadas a más de un país simultáneamente, lo que dificulta la diferenciación entre los países publicadores o patentadores y, por lo tanto, se hace necesario realizar un preprocesamiento a aquellos datos con el fin de eliminar la sobreestimación de publicaciones científicas y patentes asociadas a cada país. Debido a las características de estas bases de datos y los años que presentan, es que existe una limitante temporal que permite realizar el estudio solo entre los años 1988 y 2017, años en los que las tres bases de datos presentan datos para un total de 120 países que se encuentran presentes en el estudio.

1.4. Metodología

La metodología aplicada considera los siguientes pasos. En primer término, la construcción y procesamiento de bases de datos de publicaciones científicas, patentes y productos, mediante algoritmos de descarga masiva de análisis de datos generados en el lenguaje de programación *Python*, en función de su versatilidad y funcionalidad. En segundo término, se aplica un modelo econométrico, basado en una serie de variables explicativas representativas de la complejidad de producto, densidades ciencia-producto y tecnología—producto, de forma de estimar el efecto de capacidades endógenas en procesos de diversificación de producto a nivel país.

1.5. Estructura de la Memoria de Título

La presente MT se estructura de la siguiente forma. En el Capítulo 2 se describen los estudios previos basados en el análisis de co-ocurrencia y el desarrollo de ventajas comparativas en descubrimientos científicos, generación de patentes y exportación de productos a través de la relación entre los portafolios científicos, tecnológicos y productivos con el objetivo de analizar la metodología utilizada en dichos análisis. En el Capítulo 3 se describe la metodología propuesta para ser utilizada en esta investigación, los índices y variables a utilizar además de describir los modelos econométricos aplicados. Además, se hace referencia a los métodos de obtención de datos y su posterior procesamiento para la generación de los indicadores presentados. En el Capítulo 4 se presentan los resultados obtenidos junto a sus respectivos análisis, además de indicar las implicancias de política pública. Finalmente, en el Capítulo 5 se muestran las principales conclusiones derivadas del presente estudio.

Capítulo 2: Marco Teórico

En la tarea por comprender el proceso de gestación de nuevas tecnologías y su respectivo desarrollo hacia la elaboración de nuevos productos a lo largo del tiempo, diversos autores han estudiado la relación existente entre la actividad científica, tecnológica y productiva a diversos niveles geográficos, utilizando para ello, proxies o símiles bibliométricos tales como publicaciones científicas, patentes y productos, respectivamente. Una de las corrientes económicas que se ha enfocado en realizar estos estudios es la *economía geográfica*, disciplina científica que relaciona las actividades económicas con el lugar físico en el cuál se llevan a cabo (país, región, ciudad, etc.), analizando así el desarrollo económico desde un punto de vista territorial. Para ello en base a estudios espaciales y temporales se busca comprender la relación entre productos, tecnologías, y áreas del conocimiento, además de su correspondiente efecto en procesos de diversificación a nivel de territorios. Tal concepción ha derivado en la definición del *Principio de Relación* (Hidalgo et al., 2018), el cual define la probabilidad de diversificación hacia una nueva actividad económica en función de la cercanía o proximidad que el portafolio del territorio presenta en relación con tal nueva actividad económica.

Una de las primeras metodologías utilizadas para analizar si dos actividades productivas, tecnológicas o científicas se encuentran relacionadas, suponía observar si aquellas actividades requerían de las mismas capacidades o conocimiento para ser realizadas, aunque en la práctica, la medición de tales capacidades revestía un alto grado de complejidad, limitando la robustez de los resultados obtenidos (Leamer, 1984; Lall, 2000). Es por ello que en investigaciones más recientes se utilizan relaciones inferidas que no dependen de procesos de observación de los factores para ser utilizadas, como lo son, por ejemplo, la co-exportación de productos en un país (Hidalgo et al., 2007), la co-ocurrencia de productos en portafolios de firmas manufactureras que pertenecen a distintas industrias productivas (Neffke, Henning, & Boschma, 2011) o el flujo de mano de obra entre industrias que requieren capital humano de similares características (Neffke & Henning, 2013). Debido a esta flexibilidad metodológica, es que diversos investigadores han detectado una relación robusta y replicable entre la probabilidad de que una región desarrolle ventajas comparativas en una actividad productiva, tecnológica o de investigación científica y el número de actividades relacionadas a dicha actividad que se encuentran actualmente presentes en aquella región (Hidalgo et al., 2018).

Una de las principales investigaciones que ejemplifican dicho principio corresponde al análisis espacial propuesto por Hidalgo et al. (2007). En ella se introduce la idea del *Product Space*, representación gráfica en red que muestra la relación entre distintos tipos de productos existentes en las matrices económicas de los países. Esta representación es fundamentada por la idea de que, si dos productos se encuentran relacionados entre sí, ya sea porque requieren de similares infraestructuras, factores físicos, tecnologías o una combinación de estos elementos, entonces ellos

tenderán a ser producidos en tándem, y en su contraparte, productos que no se encuentren relacionados entre sí, tenderán a requerir capacidades distintas para su manufactura.

Para la construcción del *Product Space*, Hidalgo et al. (2007) trabajaron en base a 775 categorías de productos correspondientes al *Standar International Trade Classification* (SITC). En primera instancia obtienen el denominado *Revealed Comparative Advantage (RCA)*, índice introducido por Balassa (1964) que mide si una región posee ventajas comparativas con respecto a otras en la producción y exportación de un bien, donde un *RCA* mayor a uno, para un cierto producto, indica que el país posee una ventaja comparativa en la producción de dicho bien por sobre el resto de los países bajo estudio. De forma complementaria, el índice *RCA* es utilizado para medir la co-ocurrencia entre dos productos, considerando que existe co-ocurrencia cuando, en un país, ambos productos poseen un *RCA* mayor a uno en un mismo período de tiempo, pues se supone que ambos necesitan las mismas, o similares, capacidades para ser producidos. Gracias al *RCA*, Hidalgo et al. (2007) construyen dos índices que permiten confeccionar el *Product Space*. El primero de ellos corresponde a la *proximidad*, el cual formaliza la idea de que la habilidad de un país para producir un nuevo producto se encuentra relacionada con su habilidad de producir productos altamente relacionados a éste, obteniendo así las relaciones entre las 775 categorías de productos. El segundo indicador corresponde a la *densidad*, definida como la proximidad promedio de un nuevo producto con respecto al portafolio productivo actual de la región. Mediante la utilización de la *proximidad*, Hidalgo et al. (2007) son capaces de calcular la matriz de relación entre cada par de las 775 categorías de productos, construyendo así, el *Product Space*.

De forma heterogénea, *Product Space* concentra en su núcleo productos de mayor sofisticación e interacción entre sí, los cuales requieren de un alto grado de conocimiento técnico y teórico, además de una infraestructura de desarrollo similar para su producción. En contraparte, en su periferia se registran zonas de baja concentración de productos, los cuales demandan bajo conocimiento técnico para su desarrollo, requiriendo además de infraestructura específica durante su proceso de construcción.

En línea con los planteamientos de Hidalgo, diversos autores han realizado estudios en profundidad con el fin de explorar dinámicas espaciales y temporales en temáticas relacionadas a la generación y desarrollo de ventajas comparativas en campos productivos. Bajo dicha base, Hausmann and Klinger (2007) a través de la construcción de un modelo de capital humano sustentado en el conocimiento específico requerido en la producción de cada producto y el grado de sustituibilidad de los productos, concluyen que las oportunidades para futuras transformaciones a la matriz productiva de un país dependerán de las ventajas comparativas que ya ha desarrollado hasta la actualidad. En un periodo inicial existe un traspaso de información y aprendizaje desde un trabajador experto hacia uno inexperto en la producción de un bien, donde en un próximo periodo, el nuevo trabajador puede mantenerse produciendo dicho producto, o bien, moverse hacia la producción de

otro que se encuentre relacionado al mismo. De esta manera y, debido a la estructura heterogénea del *Product Space*, es que en general, países con portafolios de productos más conectados poseen mayores posibilidades de crecimiento en el futuro que aquellos que producen bienes poco conectados.

Del mismo modo, además de comprender las dinámicas productivas a nivel mundial, Neffke et al. (2011); Boschma, Minondo, and Navarro (2013) y Zhu, He, and Zhou (2017) analizan el comportamiento de la diversificación a nivel nacional, mediante el análisis de las matrices productivas de Suecia, España y China respectivamente. Entre los resultados obtenidos, es posible determinar que la probabilidad de que una industria irrumpa en alguna de las regiones estudiadas, depende directamente del número de industrias desarrolladas y relacionadas a ella, por lo que aquellas industrias que no poseen relación alguna con la matriz productiva de las regiones analizadas, tenderán a disminuir su participación en la matriz de productos de aquella zona, como lo fue el caso de la región de Linköping en Suecia, en donde un efecto de bola de nieve provocó la pérdida total de la industria textil y maderera en la región (Neffke et al., 2011).

Identificando una relación común entre los resultados de sus investigaciones, Hausmann e Hidalgo introducen el concepto de complejidad económica, esto es, la capacidad de los países de generar valor, basándose en la diversificación de sus exportaciones. Tal variable es indicativa de las perspectivas de crecimiento de cada país identificando variados autores, en forma posterior, la oportunidad de no acotarlo solamente a la capacidad productiva de un territorio, sino extendiéndolo a dinámicas de diversificación de capacidades científicas y tecnológicas (Colombelli, Krafft, & Quatraro, 2013; Essletzbichler, 2015; Quatraro, 2010; Saviotti & Frenken, 2008).

En esa línea, Boschma, Balland, et al. (2014), mediante el análisis de patentes concedidas por la *United States Patents and Trademark Office (USPTO)*, estudian la existencia de una dependencia relacional en los procesos de diversificación tecnológica a nivel estatal en los Estados Unidos, a través del cambio tecnológico de 355 ciudades estadounidenses para el período comprendido entre 1981 – 2010. Como resultado, se propone el *Technology Space*, representación gráfica en red que ilustra la relación de co-ocurrencia entre las distintas clases de patentes concedidas por USPTO. Entre sus principales resultados se observa que la probabilidad de entrada de una nueva tecnología aumenta en un 30% para un próximo período de tiempo si el nivel de relación entre esta nueva tecnología y el portafolio de tecnologías existentes en la ciudad aumenta en un 10% en el periodo actual, mientras que la probabilidad de salida de una tecnología disminuye en un 8%. Debido a esto, Boschma, Balland, et al. (2014) concluyen que existe suficiente evidencia para sostener que el aumento y la caída de conocimiento tecnológico, representado por la aparición y disminución de clases de patentes en ciudades, está condicionado por la existencia previa de una base de conocimiento tecnológico en las ciudades, encontrando que es más probable que una nueva tecnología irrumpa al portafolio tecnológico de la ciudad cuando está relacionada con otras

tecnologías existentes en el portafolio actual, y una tecnología existente tiene una alta probabilidad de migrar de la ciudad cuando no está relacionada con las tecnologías existentes actualmente (Koeger et al. 2013; Rigby, 2015).

Continuando la línea investigativa que proponen Boschma, Balland, et al. (2014), Petralia, Balland, and Morrison (2017) se interesan en investigar como los países son capaces de acumular y desarrollar sus capacidades innovadoras, que obstáculos deben superar para lograrlo y como les es posible identificar nuevas oportunidades para desarrollar áreas de especialización tecnológica. Para ello analizan los patrones de diversificación y especialización tecnológica de 65 países durante 15 años, observando la cantidad de patentes que registra cada país en la *USPTO*. Una de las principales conclusiones a las que llegan es que las patentes registradas por los países demuestran que el desarrollo de nuevas tecnologías es un proceso altamente acumulativo y dependiente, en donde los patrones de diversificación tecnológica emergen de las bases de conocimiento previas que poseen los países. En épocas tempranas de desarrollo, la diversificación se encuentra fuertemente limitada por las conexiones existentes entre las tecnologías pertenecientes al portafolio actual de los países, mientras que, en etapas más adelantadas, los países son capaces de generar saltos mas grandes y, por lo tanto, el desarrollo de nuevas tecnologías se ve cada vez menos relacionado con su portafolio tecnológico.

De manera análoga, Boschma, Heimeriks, et al. (2014) estudian el impacto de las relaciones de dinámicas de conocimiento en biotecnología, a nivel de ciudades, durante el período comprendido entre 1989 – 2008, utilizando las palabras presentes en títulos de publicaciones científicas en el campo de la biotecnología, para identificar la evolución de los tópicos dentro del área en cada una de las ciudades bajo estudio, con el fin de generar el *Scientific Space*, representación gráfica en red que ilustra la relación de co-ocurrencia entre los 1.028 tópicos científicos en biotecnología estudiados. Una de las conclusiones descritas por este estudio, es que los nuevos temas científicos en biotecnología tienden a surgir sistemáticamente en ciudades donde ya existen previamente temas científicos relacionados, mientras que temas existentes en la ciudad son más propensos a desaparecer de ella cuando estos se encuentran débilmente relacionados con la cartera científica de la misma. Asimismo, Guevara, Hartmann, Aristarán, Mendoza, and Hidalgo (2016) analizan las relaciones existentes en conocimiento científico basado en los registros de Google Scholar, utilizando una muestra significativa de publicaciones académicas y conectando los campos en donde cada autor había publicado, logrando así predecir cambios a nivel individual, organizacional y nacional de las trayectorias científicas.

El alcance de estudios espaciales de complejidad abarca también la perspectiva de los efectos institucionales y sociales en la producción de los países. En tal sentido, Boschma and Capone (2015) comparan sistemas institucionales, demostrando que países basados en economías de mercado coordinadas (e.g. Alemania) tienden a redistribuir sus recursos diversificando su matriz productiva

hacia productos de mayor relación con su portafolio actual, mientras que las economías de mercados liberales (e.g. Estados Unidos) poseen una mayor probabilidad de redistribuir sus recursos hacia productos de menor relación con su portafolio actual. De igual manera, Hartmann, Guevara, Jara-Figueroa, Aristarán, and Hidalgo (2017) determinan que países que han diversificado sus portafolios hacia productos más complejos han generado niveles más bajos de desigualdad de ingresos que aquellos que se han dedicado a exportar productos de menor complejidad. Bajo el mismo prisma, Mindruta, Albino-Pimentel, and Hoetker (2015) demuestran que un proceso evolutivo similar tiene lugar en el dominio de las patentes, y que las instituciones de coordinación determinan las diferencias entre países en materia de innovaciones exploratorias y de explotación.

Si bien es importante entender como se gestiona la ciencia, la tecnología y los productos con respecto a los portafolios actuales de los países, la relación que existe entre estos tópicos es mucho más interesante de analizar debido al tipo de información que estos pueden entregar, ya sea para entender dinámicas de crecimiento de los países o para la generación de nuevas políticas públicas que propongan soluciones conjuntas a las problemáticas actuales de desarrollo. Por ejemplo, de acuerdo a Artz, Norman, Hatfield, and Cardinal (2010), las patentes son frecuentemente consideradas como una buena fuente de anticipación de productos, lo cual ocurre debido a la suposición de que las tecnologías antes de ser introducidas al mercado, son patentadas por sus creadores¹.

Además, el traspaso desde el I&D hacia la creación de nuevas tecnologías y el posterior desarrollo de productos, genera un sistema tripartito que permite investigar, entre otros elementos, como los países líderes en los mercados actuales relacionan estas tres variables y si esto les genera, o no, ventajas comparativas con respecto a quienes solo en algunas de estas aristas, o quizás, en ninguna de ellas.

En línea con tal desafío de investigación, Catalán, Navarrete, and Figueroa (2019) analizan la relación entre las capacidades científicas y tecnológicas a nivel país en función de procesos de diversificación tecnológica, relacionando los desempeños científicos y tecnológicos de los países, mediante el constructo *scientific-technological cross-proximity*, el cual cuantifica la relación entre áreas científicas y clases tecnológicas. De manera similar, es desarrollado el índice *cross-density*, el cual se define como la proximidad promedio de una nueva tecnología con respecto al portafolio científico a nivel país y permite proyectar la probabilidad de diversificación hacia una nueva tecnología en función de los portafolios científicos y tecnológicos. Los autores ratifican que efectivamente la diversificación tecnológica de un país depende del portafolio tecnológico actual del mismo, pero la principal conclusión que obtienen es que existe una relación positiva entre el índice *cross-density* presentado y la diversificación tecnológica a nivel país, lo que significa que es más probable que una

¹ De acuerdo a la regulación de patentes, estas deben ser presentadas previo a la utilización en una invención, de lo contrario no sería conforme a los criterios de novedad (Gerken et al. 2014)

tecnología sea desarrollada en un país, cuando esta se encuentra más relacionada con su portafolio científico y tecnológico actual.

Los resultados obtenidos por Catalán et al. (2019) abren un abanico de nuevas preguntas de investigación en el campo de la economía geográfica, bajo la hipótesis de que la ciencia, tecnología y productos presentan capacidades y conocimiento característicos que permiten comprender, de mejor manera, el proceso de transición entre ellos. Así, en la presente memoria de título se analizará como afectan, además del portafolio actual de productos en un país, las capacidades científicas y tecnológicas en el desarrollo de nuevos productos en su cartera productiva, inspirados en la cross-density presentada en su investigación.

Capítulo 3: Metodología

En el presente capítulo se describe la metodología propuesta para la elaboración de un modelo econométrico que permita el análisis de las relaciones existentes entre las capacidades científicas, tecnológicas y productivas a nivel país. En primera instancia, se definen los indicadores necesarios para la construcción de las variables a utilizar en el modelo econométrico. Posteriormente, son detalladas las fuentes de datos utilizadas, indicando también la elección de los países seleccionados para la realización de esta memoria de título. Finalmente se detalla el modelo econométrico y el set de variables utilizado para la obtención de resultados.

3.1. Indicadores

3.1.1. Revealed Comparative Advantage (RCA)

El *Revealed Comparative Advantage (RCA)* es un índice introducido por Balassa (1964), el cual permite estimar la capacidad exportadora de una región y sus ventajas comparativas en relación al nivel de exportación del resto del mundo (Hidalgo et al., 2007).

El *RCA* del país c sobre una categoría de producto i en el periodo de tiempo t , se calcula considerando el total de exportaciones del producto i que el país c realiza en el periodo de tiempo t con respecto a la participación de dicho producto en el total de exportaciones a nivel mundial en el periodo t , tal como lo muestra la ecuación (1).

$$RCA_{i,c,t} = \frac{product_{i,c,t} / \sum_i product_{i,c,t}}{\sum_c product_{i,c,t} / \sum_i \sum_c product_{i,c,t}} \quad (1)$$

Donde:

$product_{i,c,t}$: Total de producto en la categoría i exportado por país c en el periodo de tiempo t .

$\sum_i product_{i,c,t}$: Total de productos exportados por el país c en el periodo de tiempo t .

$\sum_c product_{i,c,t}$: Total de producto exportado en la categoría i a nivel mundial en el periodo de tiempo t .

$\sum_i \sum_c product_{i,c,t}$: Total de productos exportados a nivel mundial en el periodo de tiempo t .

Por consiguiente, un $RCA_{i,c,t}$ superior a la unidad indica que el país c en el periodo de tiempo t posee ventajas comparativas en la exportación del producto i y, en caso contrario, se considera que el país posee desventajas comparativas en la exportación de dicho producto (Hidalgo et al., 2007). Debido a la versatilidad que presenta el RCA , diversos autores lo han utilizado para estimar ventajas comparativas de países en diferentes categorías, como publicaciones científicas o clases de patentes (Boschma & Capone, 2015; Rigby, 2015).

3.1.2. Proximidad

El índice de proximidad introducido por Hidalgo et al. (2007) mide la probabilidad mínima que un país posea ventajas comparativas en la exportación de un producto i , dado que posee ventajas comparativas en la exportación de un producto j , o viceversa, donde valores de proximidad más cercanos a la unidad indican mayor relación entre los productos. Además de su utilización en categorías de productos (Hidalgo et al., 2007), diversos autores han ampliado su uso a estudios sobre correspondencia entre clases de patentes y publicaciones científicas (Boschma & Capone, 2015; Boschma, Heimeriks, et al., 2014). La Ecuación (2) demuestra el cálculo de la proximidad entre el tipo de producto i y el tipo de producto j en el periodo de tiempo t .

$$\Phi_{i,j,t} = \min \left\{ \begin{array}{l} P(RCA_{product_{i,t}} > 1 | RCA_{product_{j,t}} > 1), \\ P(RCA_{product_{j,t}} > 1 | RCA_{product_{i,t}} > 1) \end{array} \right\} \quad (2)$$

Ejemplificando el cálculo de $P(RCA_{product_{i,t}} > 1 | RCA_{product_{j,t}} > 1)$, el numerador acumula el total de ocasiones donde ocurren simultáneamente ventajas comparativas en la exportación del producto i y la exportación del producto j en el periodo de tiempo t , y se obtiene verificando en el total de países si $RCA_{product_{i,c,t}} > 1$ dado que $RCA_{product_{j,c,t}} > 1$, registrando en un contador la cantidad de veces donde ocurre esta situación. Por su parte, el denominador acumula el total de ocasiones donde ocurren ventajas comparativas en la exportación del producto j , es decir, el total de países donde $RCA_{product_{j,c,t}} > 1$.

De forma análoga, el cálculo de $P(RCA_{product_{j,t}} > 1 | RCA_{product_{i,t}} > 1)$ se realiza iterando dentro de todos los países y verificando si $RCA_{product_{j,c,t}} > 1$ dado que $RCA_{product_{i,c,t}} > 1$, registrando la cantidad de veces en que ocurre esta situación. De la misma manera, el denominador identifica el total de países donde $RCA_{product_{i,c,t}} > 1$. Así, la proximidad entre el producto i y el producto j corresponde al mínimo de las probabilidades calculadas anteriormente.

3.1.3. Proximidad Cruzada

El índice de proximidad cruzada o *cross-proximity* es introducido por Catalán et al. (2019) y es utilizado para medir la correspondencia entre un área de desarrollo tecnológico l y una publicación científica de clase k , o viceversa. La proximidad cruzada mide la probabilidad mínima que un país presente ventajas comparativas en el desarrollo de patentes en el área l dado que posee ventajas comparativas en la generación de conocimiento en la categoría k , o viceversa, donde valores de proximidad cruzada más cercanos a la unidad indican una relación más potente entre el área de desarrollo tecnológico l y el tipo de conocimiento k . La Ecuación (3) muestra el cálculo de la proximidad cruzada entre el desarrollo tecnológico l y la publicación científica de clase k en el periodo de tiempo t .

$$\Phi^X_{l,k} = \min \left\{ \begin{array}{l} P(RCA_{patent_{l,t}} > 1 | RCA_{scientific_{k,t}} > 1), \\ P(RCA_{scientific_{l,t}} > 1 | RCA_{patent_{k,t}} > 1) \end{array} \right\} \quad (3)$$

Ejemplificando el cálculo de $P(RCA_{patent_{l,t}} > 1 | RCA_{scientific_{k,t}} > 1)$, el numerador acumula, registrando en un contador, el total de ocasiones donde ocurren simultáneamente ventajas comparativas en el desarrollo de tecnología en el área l y el generación de conocimiento de tipo k en el periodo de tiempo t y se obtiene verificando en el total de países si $RCA_{patent_{l,t}} > 1$ dado que $RCA_{scientific_{k,t}} > 1$. El denominador acumula el total de ocasiones donde ocurren ventajas comparativas en la generación de conocimiento científico de clase k , es decir, el total de países donde $RCA_{scientific_{k,t}} > 1$. El cálculo de $P(RCA_{scientific_{l,t}} > 1 | RCA_{patent_{k,t}} > 1)$ se realiza de manera análoga. Así, la proximidad cruzada entre la clase de patente l y el área de conocimiento k corresponde al mínimo de las probabilidades calculadas anteriormente. Este indicador puede ser utilizado de manera análoga para analizar las relaciones entre ciencia y tecnología, ciencia y productos y tecnología y productos.

3.2. Fuentes de Datos

Para poder analizar si existe una relación entre las capacidades científicas, tecnológicas y productivas de un país, utilizando para ello los índices y variables a describir, se requiere de una base de datos que represente dichas temáticas, como lo son bases de datos de publicaciones científicas, patentes y exportaciones de producto. A continuación, se describen las bases de datos seleccionadas, como también el método de obtención y el preprocesamiento de datos, en caso de ser necesario.

3.2.1. Publicaciones Científicas

Web of Science (WoS) es un servicio online de indexación de citas científicas producido originalmente por el *Instituto de Información Científica*² (*ISI por sus siglas en inglés*), el cual permite el acceso a múltiples bases de datos que contienen información sobre investigación multidisciplinaria de alto nivel académico y científico en tres grandes áreas de investigación: *Artes y Humanidades*, *Ciencias Sociales* y *Ciencia y Tecnología*. Cada área de investigación se encuentra sub-dividida en categorías, las cuales permiten detallar con mayor precisión los contenidos abarcados en cada una de las publicaciones. Es considerada una de las bases de datos en publicaciones científicas más importantes a nivel internacional, junto a Google Scholar, Scopus y PubMed (siendo esta última exclusiva de publicaciones médicas), teniendo como característica diferenciadora ser la que cuenta con los registros más antiguos (Falagas, Pitsouni, Malietzis, & Pappas, 2008).

En la presente memoria se utilizará la base de datos de *Web of Science* debido a que se permite el acceso a estudiantes de la Universidad de Concepción (de manera presencial y a través de un proxy³), además de contener datos disponibles desde 1988 hasta el 2018.

3.2.1.1. Descarga de Publicaciones Científicas

Debido a que *Web of Science* no cuenta con una *API*⁴ que permita la automatización de la descarga de datos, se debió realizar el proceso manualmente con el objetivo de obtener la información referente a 254 categorías de publicaciones científicas en el periodo comprendido entre los años 1988 y 2018. Para lograr esta labor, se realizaron consultas en búsqueda avanzada considerando dos dimensiones de búsqueda, las cuales son “*Web of Science Category (WC)*” y “*Publication Year (PY)*”, realizándose un total de 7874 consultas a la base de datos de *WoS* (correspondientes a realizar una consulta por año para cada una de las 254 categorías, durante el periodo comprendido entre 1988-2018). Para ejemplificar el procedimiento realizado, se explicará el caso particular de la consulta realizada para la categoría *WC=Engineering, Aerospace* y el año de publicación *PY=2005*, la cual será referenciada como Ejemplo A (Ver figura 3.1).

² El *Instituto de Información Científica* fue un servicio de publicación académica, fundado por Eugene Garfield en Filadelfia en 1960. *ISI* ofreció servicios de bases de datos cuantitativas y bibliográficas y su especialidad era la indexación y el análisis de citas.

³ Un proxy es un intermediario en la comunicación entre dos ordenadores, que media entre ellos con el fin de que la comunicación entre los otros no sea de manera directa. En este caso permite ingresar a *WoS* desde cualquier lugar con internet como si se estuviera ingresando desde la red de internet de la universidad.

⁴ *Application Programming Interface (API)*, conjunto de funciones y procedimientos, utilizadas como un software, que permite la obtención o procesamiento de datos desde una biblioteca particular.

Figura 3.1: Consulta realizada en búsqueda avanzada para el Ejemplo A en WoS

Basic Search Cited Reference Search **Advanced Search** Author Search

Use field tags, Boolean operators, parentheses, and query sets to create your query. Results will appear in the Search History table at the bottom of the page. ([Learn more about Advanced Search](#))

Example: TS=(nanotub* AND carbon) NOT AU=Smalley RE #1 NOT #2 [more examples](#) | [view the tutorial](#)

WC=Engineering, Aerospace AND PY=2005

Search

Restrict results by languages and document types:

All languages	All document types
English	Article
Afrikaans	Abstract of Published Item
Arabic	Art Exhibit Review

Fuente: Elaboración propia

Luego de realizada la consulta para el Ejemplo A, *Web of Science* despliega la ventana de resultados como se observa en la Figura 3.2. Allí, la opción “*Analyze Results*” permite analizar los resultados obtenidos en la consulta.

Figura 3.2: Ventana de Resultados de WoS

Web of Science Clarivate Analytics

Search Tools Searches and alerts Search History Marked List

Results: 2,796
(from Web of Science Core Collection)

You searched for: (WC=Engineering, Aerospace AND PY=2005) AND LANGUAGE: (English) AND DOCUMENT TYPES: (Article) ...[More](#)

[Create Alert](#)

Refine Results

Search within results for...

Filter results by:

☐ [Open Access](#) (68)

Sort by: **Date** [17](#) Times Cited Usage Count Relevance More ▼

1 of 280

☐ Select Page [Export...](#) [Add to Marked List](#)

☐ 1. **Computation of optimal Mars trajectories via combined chemical/electrical propulsion, Part 2: Minimum time solutions with bounded thrust direction**
By: Miele, A; Wang, T; Williams, PN
ACTA ASTRONAUTICA Volume: 57 Issue: 11 Pages: 819-828 Published: DEC 2005
[Full Text from Publisher](#) [View Abstract](#) ▼

☐ 2. **Computation of optimal Mars trajectories via combined chemical/electrical propulsion, Part 3: Compromise solutions**
By: Miele, A; Wang, T; Williams, PN
ACTA ASTRONAUTICA Volume: 57 Issue: 11 Pages: 829-840 Published: DEC 2005
[Full Text from Publisher](#) [View Abstract](#) ▼

Analyze Results
[Create Citation Report](#)

Times Cited: 3
(from Web of Science Core Collection)

Usage Count ▼

Times Cited: 3
(from Web of Science Core Collection)

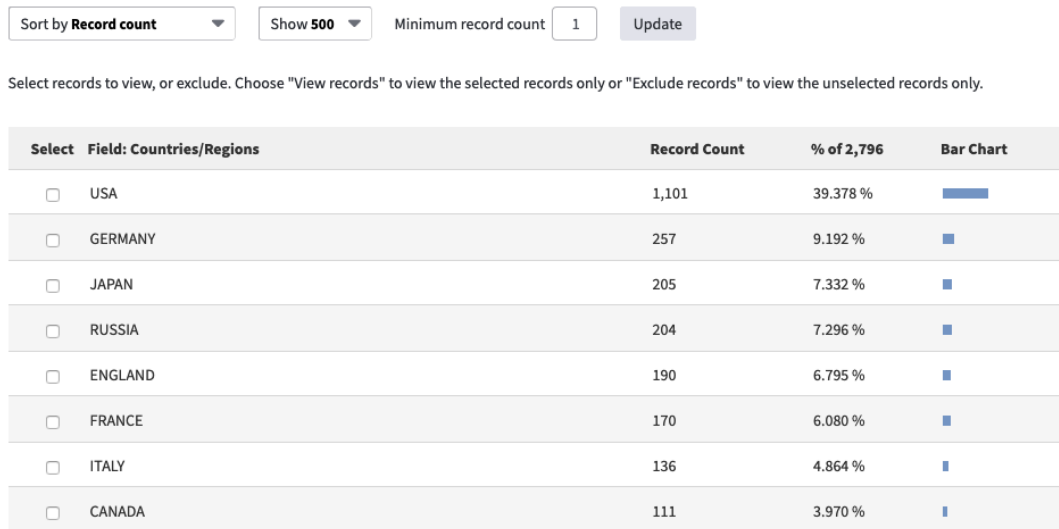
Usage Count ▼

Fuente: Elaboración propia

La Figura 3.3 muestra los resultados desplegados para la consulta realizada en el Ejemplo A. WoS permite desplegar hasta 500 registros de países que posean, a lo menos, una publicación científica

en dicha categoría en el año consultado, permitiendo así no excluir ningún artículo de los resultados. La descarga de datos se realizó a través de un algoritmo desarrollado en Python, el cual permite el almacenamiento de estos en formato pickle⁵, considerándose cuatro características para cada una de las tuplas de datos, como se muestra en la Figura 3.4

Figura 3.3: Análisis de resultados para la consulta en WoS



Fuente: Elaboración propia

Figura 3.4: Formato de descarga de datos desde WoS

	category	year	country	value
653	Engineering, Aerospace	2005	USA	1101
654	Engineering, Aerospace	2005	GERMANY	257
655	Engineering, Aerospace	2005	JAPAN	205
656	Engineering, Aerospace	2005	RUSSIA	204
657	Engineering, Aerospace	2005	ENGLAND	190

Fuente: Elaboración propia

Así, para el cálculo de indicadores y variables, se seleccionaron los campos category, year, country y value, los cuales se encuentran descritos en la Tabla 3.1.

⁵ El módulo pickle implementa un algoritmo fundamental, pero poderoso, para serializar y deserializar una estructura de objetos Python.

Tabla 3.1: Campos seleccionados del análisis de resultados de WoS

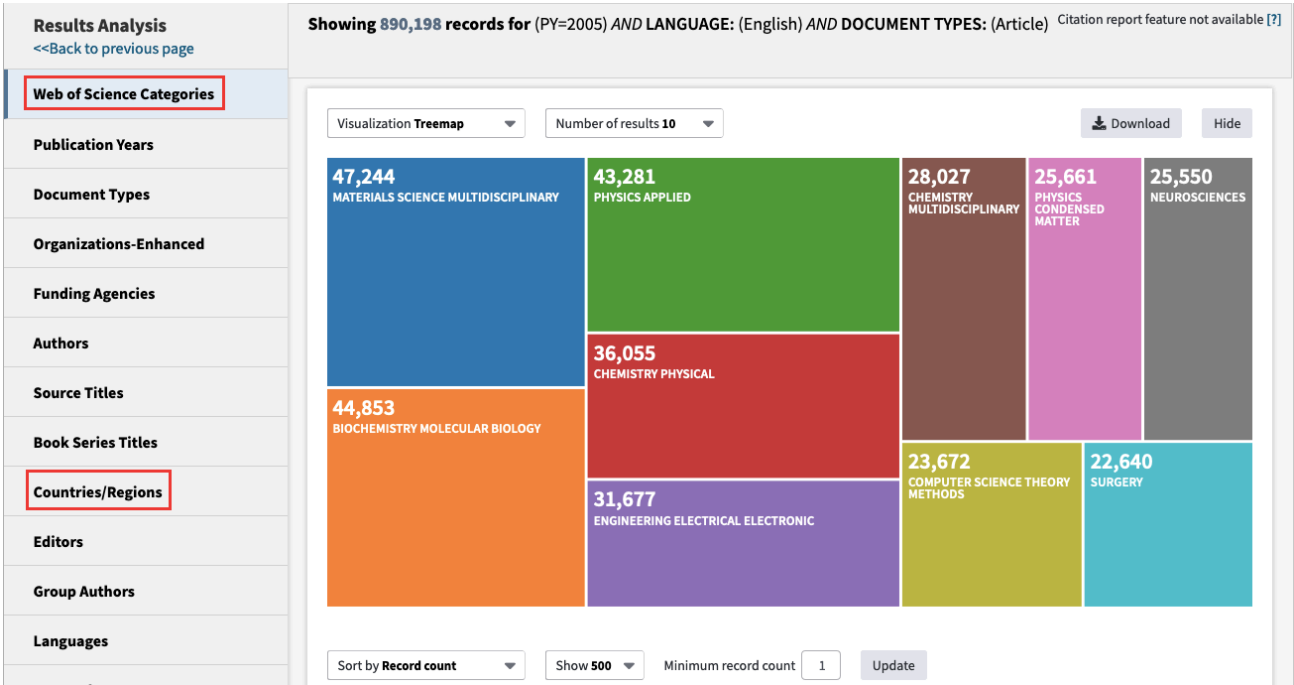
Campo	Descripción
category	Categoría de Web of Science
year	Total de publicaciones de la categoría i en el periodo de tiempo t
country	País del autor de la publicación científica
value	Año en el cual fue realizada la publicación científica

Fuente: Elaboración propia

Debido a que el sistema de clasificación de publicaciones científicas permite clasificarlas en más de una categoría y asignarlas a más de un país, se requiere de dos bases de datos adicionales con el fin de eliminar sobreestimaciones para no generar errores al calcular indicadores y variables.

De acuerdo al cálculo del RCA mostrado en la Ecuación (1), además de la base de datos explicada anteriormente, se requiere saber el número de publicaciones científicas que el país c realiza en el periodo de tiempo t y la cantidad de publicaciones de la categoría i que fueron publicadas en el periodo t , las cuales son obtenidas realizando una consulta en búsqueda avanzada con la dimensión “PY” para cada uno de los años comprendidos entre 1988-2018. El proceso de descarga de datos es similar al explicado anteriormente, pero al momento de visualizar el análisis de resultados presentando por WoS, se procede a descargar las dimensiones “Web of Science Category” y “Countries/Regions” para cada una de las consultas, como se muestra en la Figura 3.5. Las tuplas resultantes de estas consultas son mostradas en la Figura 3.6.

Figura 3.5: Dimensiones seleccionadas para eliminar sobreestimaciones



Fuente: Elaboración propia

Figura 3.6: Formato de descarga de datos para eliminar sobreestimaciones

Field: Web of Science Categories			Record Count	Field: Countries/Regions			Record Count
0	MATERIALS SCIENCE MULTIDISCIPLINARY		47244	0	USA		282712
1	BIOCHEMISTRY MOLECULAR BIOLOGY		44853	1	JAPAN		74052
2	PHYSICS APPLIED		43281	2	GERMANY		66342
3	CHEMISTRY PHYSICAL		36056	3	PEOPLES R CHINA		64478
4	ENGINEERING ELECTRICAL ELECTRONIC		31677	4	ENGLAND		62894
5	CHEMISTRY MULTIDISCIPLINARY		28027	5	FRANCE		47112

Fuente: Elaboración propia

3.2.1.2. Preprocesamiento de Publicaciones Científicas

Tal como se aprecia en las Figuras 3.4 y 3.6, los países vienen descritos en sus nombres completos o abreviaturas, por lo tanto, se procedió a actualizar estos datos, reemplazando los nombres entregados por WoS por su código ISO 3166-2 de tres dígitos respectivamente, resultando en la base de datos presentada en la Figura 3.7.

Figura 3.7: Formato de base de datos de publicaciones científicas

	category	year	country	value
146924	ENGINEERING AEROSPACE	2005	ARE	1
146949	ENGINEERING AEROSPACE	2005	ARG	11
146966	ENGINEERING AEROSPACE	2005	ARM	1
146992	ENGINEERING AEROSPACE	2005	AUS	43
147021	ENGINEERING AEROSPACE	2005	AUT	27
147058	ENGINEERING AEROSPACE	2005	BEL	23
147074	ENGINEERING AEROSPACE	2005	BGD	1
147102	ENGINEERING AEROSPACE	2005	BGR	9
147147	ENGINEERING AEROSPACE	2005	BRA	42

Fuente: Elaboración propia

3.2.2. Patentes

La *Oficina de Patentes y Marcas de los Estados Unidos (USPTO⁶ por sus siglas en inglés)* es una agencia del Departamento de Comercio de los Estados Unidos cuyo rol es la concesión de patentes para la protección de invenciones y el registro de marcas con el fin de promover la protección de la propiedad intelectual de innovadores y empresas mediante la colaboración con otras agencias de derecho a la propiedad intelectual en el mundo, como lo son la *Oficina Europea de Patentes (EPO⁷)* y la *Oficina Japonesa de Patentes (JPO⁸)*. Posee registros escritos a partir del 31 de julio de 1790 y en formato digital a partir del 1 de enero de 1976 hasta la actualidad.

USPTO trabaja con los tres sistemas de clasificación de patentes más importantes a nivel mundial: el *United States Patent Classification (USPC)* desarrollado por la propia USPTO, el *International Patent Classification (IPC)* desarrollado por la *Oficina de Propiedad Intelectual Mundial (WIPO⁹)* y el *Cooperative Patent Classification (CPC)* desarrollado en conjunto por USPTO y EPO con la finalidad de generar un sistema de clasificación común, el cual entró en vigor el 1 de enero de 2013.

Para do Canto Cavalheiro, Joia, and Van Veenstra (2016), el sistema de clasificación CPC nace como una respuesta a las demandas planteadas por las empresas innovadoras que requerían una clasificación de patentes más precisa con el fin de proporcionar una protección de mayor seguridad a sus nuevas tecnologías.

En esta memoria de título se trabaja en base a 661 tipos de patentes concedidas por USPTO entre los años 1976 y 2018 debido a la alta valoración que posee el mercado mundial sobre USPTO dada

⁶ *United States Patent and Trademark Office*

⁷ *European Patent Office*

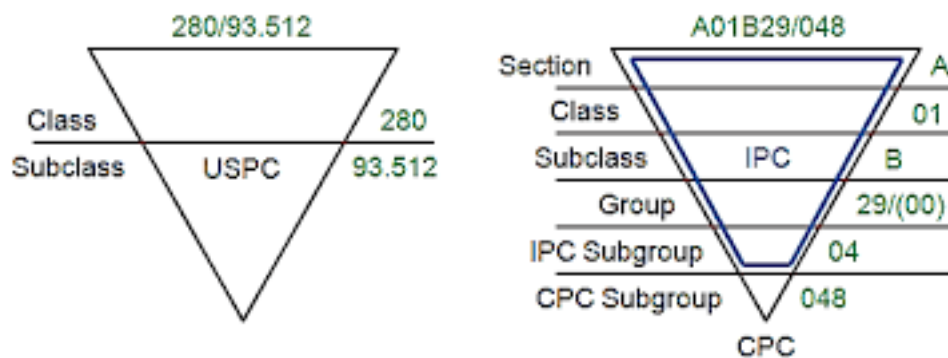
⁸ *Japan Patent Office*

⁹ *World Intellectual Property Organization*

la competitividad del mercado estadounidense a nivel global (Criscuolo, 2006), además de la facilidad de obtención de los registros de patentes con fines investigativos.

Sobre la clasificación de patentes, se decidió trabajar con *CPC* por sobre *USPC* e *IPC* debido a que dispone de una división más precisa en la definición de tecnologías, lo que permite una mejor delimitación y reducción del número de patentes en cada subgrupo (IPO, 2015). La Figura 3.8 muestra la diferencia entre los sistemas de clasificación *USPC*, *IPC* y *CPC*, mientras que la Tabla 3.2 ejemplifica la jerarquización del estándar *CPC*.

Figura 3.8: Diferencias entre *USPC*, *IPC* y *CPC*



Fuente: IPO (2015)

Tabla 3.2: Ejemplo de jerarquización de patentes en el estándar *CPC* para el subgrupo E04D1/20

Clasificación	Descripción
Sección E	Construcciones Fijas
Clase E04	Edificio
Subclase E04D	Cubiertas de techo; Tragaluz; Canaletas; Herramientas de techo
Grupo Principal E04D/00	Cubrimiento de techo utilizando tejas, pizarras, teja asfáltica u otros elementos de techo pequeños
Subgrupo E04D1/20	De plástico o de materiales fibrosos

Fuente: Elaboración propia a partir de datos de *USPTO*

Como se observa en la Figura 3.8, en el sistema de clasificación *CPC* las patentes se encuentran agrupadas en macro-áreas o secciones, las cuales representan las nueve categorías principales a las que pueden pertenecer las patentes. Esta agrupación se observa en la Tabla 3.3.

Tabla 3.3: Secciones del estándar *CPC*

Jerarquía	Sección
A	Necesidades Humanas
B	Operaciones y Transporte
C	Química y Metalúrgica
D	Textiles y Papel
E	Construcciones Fijas
F	Ingeniería Mecánica
G	Física
H	Electricidad
Y	Nuevas tecnologías intersectoriales

Fuente: Elaboración propia a partir de datos de *USPTO*

3.2.2.1. Descarga de Patentes

La descarga de la base de datos de patentes es realizada a través “*PatentsView*”, una plataforma externa apoyada por *USPTO* la cual provee la descarga de datos en formato *JSON*¹⁰ mediante el uso de una *API* pública, lo que permite la automatización del proceso a través de la utilización de un algoritmo recursivo desarrollado en el lenguaje de programación Python que posibilita la obtención de la información solicitada. El algoritmo obtiene todas las patentes concedidas por *USPTO* divididas según su país de origen, información que se obtiene a través de la nacionalidad de las organizaciones patrocinadoras de la patente denotado por *assignee_country*. La Tabla 3.4 muestra los campos resultantes en la descarga de datos al utilizar el algoritmo, los cuales serán utilizados para su posterior procesamiento.

Tabla 3.4: Descripción de campos seleccionados de *PatentsView*

Campo	Descripción
assignee_country	Territorio/país de la patente según estándar <i>ISO 3166-2</i>
cpc_group_id	Subclase del estándar de clasificación <i>CPC</i>
patent_number	Número de patente concedida. Es utilizado posteriormente para evitar elementos duplicados en la base de datos.
patent_year	Año en que es concedida la patente

Fuente: Elaboración propia a partir de datos de *PatentsView*

¹⁰ *JavaScript Object Notation*

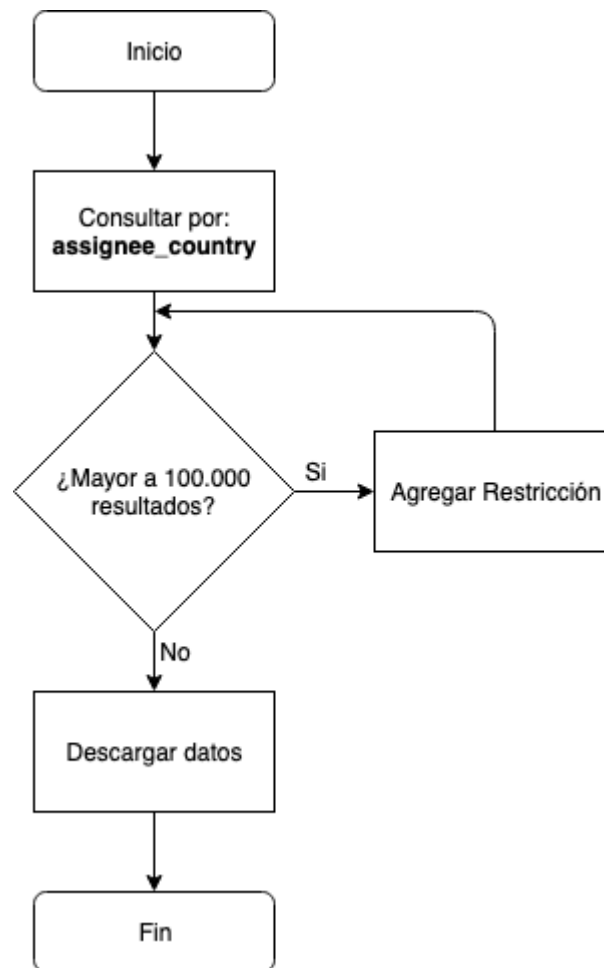
La *API* de descarga que provee PatentsView permite un máximo de 100.000 resultados por cada consulta, lo que genera limitaciones en ciertos casos particulares. Si un país supera los 100.000 resultados, entonces se procede a dividir la consulta mediante la utilización de los campos `assignee_country` y `patent_year` en conjunto, lo que significa restringir la búsqueda por país de origen y por el año en el cual la patente fue concedida. Estados Unidos es un caso particular ya que desde el año 2012 genera más de 100.000 resultados anualmente, teniendo que realizar una subdivisión adicional incluyendo el campo `cpc_section` que representa a cada una de las secciones contenidas en la Tabla 3.3.

La Figura 3.9 representa el diagrama de flujo del algoritmo de descarga de datos y la Figura 3.10 muestra el formato de descarga desde la API de PatentsView.

En resumen, el procedimiento de descarga de datos que realiza el algoritmo es el siguiente:

1. Obtener todos los resultados de patentes por país mediante el uso del campo `assignee_country`.
2. En caso de que la selección exceda los 100.000 resultados, añadir restricción con el campo `patent_year` a la consulta para realizar un filtro adicional con el año de concesión.
3. En caso de que la selección aún exceda los 100.000 resultados, añadir restricción con el campo `cpc_section` a la consulta para realizar un filtro adicional con la sección de cada categoría *CPC*.

Figura 3.9: Diagrama de flujo de algoritmo de descarga de datos



Fuente: Elaboración propia

Figura 3.10: Formato de descarga de datos desde PatentsView

```

{
  "patents": [
    {
      "patent_number": "7851177",
      "patent_year": "2010",
      "assignees": [
        { "assignee_country": "CL" }
      ],
      "cpcs": [
        { "cpc_group_id": "C12Q" }
      ],
    },
    {
      "patent_number": "7915031",
      "patent_year": "2011",
      "assignees": [
        { "assignee_country": "CL" }
      ],
      "cpcs": [
        { "cpc_group_id": "C22B" },
        { "cpc_group_id": "Y02P" }
      ],
    }
  ]
}

```

Fuente: Elaboración propia

3.2.2.2. Preprocesamiento de Patentes

Debido a que el sistema de clasificación *CPC* permite adjudicar la patente a más de un inventor, existen patentes que se encontrarán inscritas en más de un país luego de descargar los datos, generando con ello una duplicidad que podría alterar los resultados. Para evitar esta duplicidad se realiza un preprocesamiento de los datos, utilizando la librería Pandas¹¹ y siguiendo el siguiente procedimiento:

1. Unir todos los archivos *JSON* obtenidos de las descargas realizadas a PatentsView.
2. Ordenar las patentes por el campo `patent_number`.
3. Generar un vector de datos que contengan los valores únicos pertenecientes al campo `patent_number`.
4. Eliminar las tuplas que posean el campo `patent_number` duplicado, teniendo en consideración mantener una tupla dentro de la base de datos.
5. Asignar cada una de las categorías *CPC* a cada uno de los países asignados en la patente.

¹¹ Pandas es una biblioteca de código abierto con licencia BSD que proporciona estructuras de datos de alto rendimiento y fáciles de usar, además de herramientas de análisis de datos para el lenguaje de programación Python.

Debido a que cada registro de patente se encuentra normalizado según la norma ISO 3166-2 de dos dígitos, es necesario realizar la transformación a ISO 3166-2 de tres dígitos mediante el uso de un diccionario de Python. Como las categorías se encuentran normalizadas en formato CPC, no es necesario realizar un procedimiento adicional, resultando en la base de datos presentada en la Figura 3.11.

Figura 3.11: Formato de base de datos de patentes

	category	country	year	value
1857	A01G	CHL	2016	1.0
1858	A01G	CHL	2017	1.0
3329	A01K	CHL	2001	1.0
3330	A01K	CHL	2003	1.0
3331	A01K	CHL	2014	1.0
3332	A01K	CHL	2016	1.0

Fuente: Elaboración propia

3.2.3. Productos

El *Observatorio de Complejidad Económica (OEC*¹² por sus siglas en inglés) es actualmente la herramienta de visualización de datos en comercio internacional líder en el mundo, poniendo a disposición de los usuarios la recopilación de más de 50 años en datos. Liderado por Alexander Simoes y Cesar Hidalgo, el *OEC* es un proyecto perteneciente al *Massachusetts Institute of Technology Media Lab (MIT Media Lab)* y permite obtener información histórica del intercambio de productos entre países desde el año 1962 hasta el año 2017, donde los datos reportados entre el año 1962 y el año 2000 provienen del *Centro de Datos Internacionales de Robert Feenstra*¹³, mientras que los datos más recientes (2000-2017) provienen del *UN COMTRADE*¹⁴. Los datos se encuentran categorizados en base al *Standar International Trade Classification (SITC), Rev. 4*, la cual permite generar un estándar de bienes exportados e importados que hagan posible una comparación entre países a través del tiempo. Dicha versión fue aprobada en marzo de 2006 durante la 37ª sesión de la Comisión Estadística de la Organización de las Naciones Unidas (*ONU*), la cual recomienda su uso para el análisis del comercio internacional por parte de países y organizaciones interesadas. La *SITC, Rev. 4* clasifica los productos en 10 categorías principales, las cuales se presentan en la Tabla 3.5.

¹² *Observatory of Economic Complexity*

¹³ Feenstra, R. C., et al. (2005). World Trade Flows, 1962–2000. NBER working paper 11040

¹⁴ Base de datos de estadísticas de comercio internacional de las Naciones Unidas

Tabla 3.5: Clasificación principal de la SITC, Rev. 4

Nivel Superior	Categoría
0	Alimentos y animales vivos
1	Bebidas y tabaco
2	Materiales brutos no comestibles, excepto combustibles
3	Combustibles minerales, lubricantes y materiales relacionados
4	Aceites vegetales y animales, grasas y ceras
5	Productos químicos y relacionados
6	Bienes manufacturados, clasificados principalmente por material
7	Maquinaria y equipo de transporte
8	Diversos artículos manufacturados
9	Materias primas y transacciones no clasificadas en otra parte del SITC

Fuente: Elaboración propia a partir de datos de la División Estadística de las Naciones Unidas

3.2.3.1. Descarga de Productos

El OEC posee una API pública que permite la descarga de datos en formato TSV¹⁵ con una serie de indicadores para cada categoría de producto según su país de origen y año. De los indicadores que entrega la API para los datos descargados, se seleccionaron los campos de year, origin, sitc y export_val, descritos en la Tabla 3.6.

Tabla 3.6: Descripción de los campos seleccionados del OEC

Campo	Descripción
year	Año de referencia de la información expresada
origin	Identificador de país exportador en formato ISO 3166-2 de tres dígitos
sitc	Identificador de la categoría SITC correspondiente al producto exportado
export_val	Valor de la exportación de producto particular en el año y país correspondiente

Fuente: Elaboración propia a partir de datos del OEC

La Figura 3.12 muestra el formato de descarga de los datos desde la API del OEC. Debido a que los países ya vienen en formato ISO 3166-2 de tres dígitos, no se requieren procesamiento adicionales.

¹⁵ Extensión para un archivo delimitado por tabuladores realizado por un software de hoja de cálculo

Figura 3.12: Formato de descarga de los datos del OEC

	year	origin	sitc	export_val
0	1962	AGO	6940	1000.0
1	1962	AGO	7000	2000.0
2	1962	AGO	7130	4000.0
3	1962	AGO	7244	2000.0
4	1962	AGO	7271	1000.0

Fuente: Elaboración propia

3.3. Intervalos de Estudio

La Tabla 3.7 resume las características principales de las tres bases de datos presentadas anteriormente. En base a los periodos de tiempo limitantes, se obtiene un periodo de estudio de 30 años comprendidos entre 1988 y 2017.

Tabla 3.7: Resumen de las bases de datos

Fuente de datos	Año de inicio	Año de término	Número de categorías	Número de registros
Publicaciones	1988	2018	254 tipos de publicaciones científicas	41.198.765 registros
Patentes	1976	2018	661 tipos de patentes	4.494.620 registros
Productos	1962	2017	986 tipos de productos	2.351.542 registros

Fuente: Elaboración propia

3.4. Selección de Países

Debido a que cada base de datos presenta una distinta cantidad de países, se requiere procesar los datos de manera que no existan países dentro del estudio que no posean datos en alguna de las tres bases anteriormente presentadas. De esta manera, se procede a seleccionar los países de acuerdo al siguiente tratamiento:

1. En una primera instancia, seleccionar los países que poseen al menos una observación en cada una de las tres bases de datos (publicaciones científicas, patentes y productos).
2. De aquellos países, seleccionar solo los estados miembros de las Naciones Unidas, integrando los territorios dependientes al país respectivo que lo gobierna.

3. Finalmente, seleccionar todos los países que poseen una población mayor 1.25 millones de personas, misma metodología que la utilizada por OEC.

De esta manera, son 120 los países que serán utilizados en este estudio, los cuales se encuentran detallados en el Anexo A.

3.5. Modelo Económico

3.5.1. Modelos de Regresión Lineal Múltiple (RLM)

Una de las herramientas estadísticas más importantes es el análisis de la relación o dependencia entre dos o más variables con la finalidad de conocer el efecto que una o varias variables puedan causar sobre otra. En este sentido, los modelos de regresión lineal intentan explicar dicha dependencia entre una *variable dependiente* (Y) y el resto de *variables explicativas* (X), siendo el modelo de regresión lineal simple uno de los más básicos, el cuál posee la siguiente forma:

$$Y = \beta_0 + \beta_1 X + \varepsilon \quad (4)$$

Donde:

Y : Variable dependiente o explicada.

X : Variable explicativa o predictora.

ε : Término de error o perturbación.

β_0 : Parámetro de intersección o término constante.

β_1 : Parámetro de la pendiente entre las variables X e Y , el cuál indica cómo cambia Y al incrementar X en una unidad marginal, manteniendo constantes los factores contenidos en ε .

Si bien la aplicación de este modelo permite el análisis de dependencia entre las dos variables, indirectamente se está admitiendo que todos los factores que influyen a la variable respuesta Y pueden dividirse únicamente en dos grupos: el primero contiene a la variable explicativa X y el segundo contiene a todo el conjunto restante de factores que no son controlados en la regresión, englobados bajo el término ε . Esto provoca que la dependencia explicada entre las variables no sea perfecta y se encuentre sujeta a un nivel alto de incertidumbre. Un modelo más adecuado que minimiza el número de factores no controlados es el modelo de regresión lineal múltiple, el cuál permite disminuir el nivel de incertidumbre mediante la inclusión de una mayor cantidad de variables explicativas, las cuales afectan simultáneamente a la variable dependiente. La ecuación (5) muestra el modelo general de regresión lineal múltiple con k variables explicativas:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \cdots + \beta_k X_k + \varepsilon \quad (5)$$

Donde:

Y : Variable dependiente o explicada.

X_k : k -ésima variable explicativa o predictora.

ε : Término de error o perturbación.

β_0 : Parámetro de intersección o término constante.

β_1 : Parámetro de la pendiente entre las variables X_k e Y , el cuál indica cómo cambia Y al incrementar X_k en una unidad, manteniendo constantes los factores contenidos en ε .

Dada la inclusión de una mayor cantidad de variables explicativas en el modelo de regresión lineal múltiple, este tipo de análisis permite explicar de mejor forma la variabilidad presentada en la variable dependiente, obteniendo así, mejores modelos predictivos en comparación con el modelo de regresión lineal simple. Aun así, sin importar cuantas variables explicativas sean integradas en el modelo, siempre existirán factores no cuantificables que afecten a la variable explicada, los cuales serán contenidos en el término de error ε .

La Ecuación (5) demuestra el comportamiento lineal que existe entre la variable dependiente y el set de k variables explicativas a nivel poblacional, por lo que se hace necesario generar estimadores óptimos que permitan analizar el comportamiento de los factores haciendo uso de muestras aleatorias.

3.5.2. Estimación de Modelos RLM: Método de Mínimos Cuadrados Ordinarios (MCO)

El modelo de regresión lineal múltiple basado en la estimación por Mínimos Cuadrados Ordinarios (MCO) busca minimizar la suma de los cuadrados de las distancias verticales entre los valores observados en una muestra y los de la regresión estimada, es decir, minimiza la suma de los residuos al cuadrado, utilizando como residuo la diferencia entre el valor observado y el estimado. Su formulación se encuentra detallada en la ecuación (6).

$$\sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i,1} - \hat{\beta}_2 x_{i,2} - \dots - \hat{\beta}_k x_{i,k})^2 \quad (6)$$

Donde:

i : Subíndice asociado al número de la observación en la muestra. $i \in [1, n]$

$\hat{\beta}_k$: Valor estimado de β_k .

Para estimar el valor de cada $\hat{\beta}$ es necesario resolver el problema de minimización presentado en la Ecuación (6) mediante el uso de herramientas de cálculo multivariable sobre un set de $k+1$ ecuaciones y $k+1$ incógnitas. Su representación matricial se observa en la Ecuación (7).

$$\hat{\beta} = (X'X)^{-1}X'y \quad (7)$$

Donde:

- $\hat{\beta}$: Vector de estimación de los coeficientes β_k .
- X : Matriz de las variables explicativas $x_{i,k}$.
- X' : Matriz traspuesta de X .
- y : Vector de variable dependiente y_i .

Además, podemos considerar la regresión presentada en la ecuación (6) en términos diferenciales, tal como se muestra en la ecuación (8).

$$\Delta\hat{y} = \hat{\beta}_1\Delta x_1 + \hat{\beta}_2\Delta x_2 + \hat{\beta}_3\Delta x_3 + \dots + \hat{\beta}_k\Delta x_k \quad (8)$$

De esta forma, los coeficientes $\hat{\beta}_k$ obtenidos por el método de MCO mediante la Ecuación (7) pueden ser interpretados como efectos parciales o *ceteris paribus*¹⁶ según la formulación presentada en la Ecuación (8), permitiendo así interpretar que el coeficiente $\hat{\beta}_1$ mide cómo cambia la variable respuesta \hat{y} en el aumento o disminución en una unidad marginal en el valor de x_1 , manteniendo constantes al resto de variables explicativas incluidas en el modelo. De la misma manera, se pueden realizar una interpretación similar para los demás coeficientes.

Los modelos de regresiones lineales poseen diversos supuestos que deben ser corroborados con el objetivo de obtener estimadores insesgados al utilizar la metodología de MCO. El primero de ellos consta en verificar que las variables explicativas propuestas no son constantes y que, además, no existe una relación lineal entre ellas (es decir, que una variable no pueda ser descrita como una combinación lineal de otras). La verificación se requiere para evitar que exista multicolinealidad entre las variables explicativas. Un segundo supuesto, denominado media condicional cero, verifica que no exista correlación entre las variables explicativas y el término de error. Finalmente, se debe verificar el supuesto de homocedasticidad de la regresión, es decir que independiente de los valores que tomen las variables explicativas, el termino de error posee varianza constante.

Si se cumplen estos tres supuestos, entonces el método de estimación de MCO asegura la obtención de estimadores insesgados, pero debido a las características de las variables a utilizar en esta

¹⁶ Supuesto económico que explica el impacto de una variable sobre otra mientras todos los demás factores permanecen constantes

memoria de título (variables dicotómicas), es probable que el estimador obtenido al utilizar el método de MCO no posea mínima varianza, lo cual genera problemas de homocedasticidad en el modelo de regresión lineal. Es por esto que es necesario utilizar la corrección de White para asegurar que los coeficientes estimados sean insesgados y de mínima varianza.

3.5.3. Variables del Modelo

En base al estudio realizado por Hausmann and Klinger (2007) en conjunto con los indicadores y descubrimientos realizados por Catalán et al. (2019), la presente Memoria de Título analiza el efecto que poseen las capacidades científicas, tecnológicas y productivas sobre la generación de ventajas comparativas en la exportación de productos en los países.

3.5.3.1. Variable Dependiente

Según Hausmann and Klinger (2007), la generación de ventajas comparativas respecto a las capacidades productivas del país c en el periodo de tiempo t se verán reflejadas en un periodo de tiempo posterior $t+1$. Debido a esto, la variable dependiente utilizada en esta Memoria de Título refleja la existencia de ventajas comparativas en la exportación del producto i en el país c en el periodo de tiempo $t+1$, es decir, cuando el RCA del país c en la exportación del tipo de producto i en el periodo de tiempo $t+1$ sea mayor a uno, definiendo así la variable $M_{i,c,t+1}$ la cuál se encuentra descrita en la Ecuación (9):

$$M_{i,c,t+1} = \begin{cases} 1, & \text{si } RCA_{i,c,t+1} \geq 1 \\ 0, & \text{en otro caso} \end{cases} \quad (9)$$

Donde:

$RCA_{i,c,t+1}$: Índice *Revealed Comparative Advantage* para la exportación del tipo de producto i realizada por el país c en el periodo de tiempo $t+1$

De esta manera, un valor de $M_{i,c,t+1}$ de uno indicará que el país c posee ventajas comparativas en la exportación del producto i en el periodo de tiempo $t+1$ y un valor de cero indicará que no las posee.

3.5.3.2. Variables Explicativas

Las variables explicativas contenidas en el modelo tienen por objetivo identificar las relaciones existentes entre ciencia, tecnología y producto mediante el uso de un set de indicadores que permiten la caracterización y el análisis de aquellas relaciones.

3.5.3.2.1. Ventajas Comparativas

Una de las variables mas intuitivas al momento de analizar la generación de ventajas comparativas en la exportación del tipo de producto i para el país c en el periodo de tiempo $t+1$ es observar si el país ya posee estas ventajas en el periodo de tiempo t , debido a que se puede dar el caso de que el país exporte dicho producto, pero no con una ventaja comparativa con respecto al resto del mundo. Es por esto que se define la variable $M_{i,c,t}$, la cual refleja la existencia de ventajas comparativas en la exportación del producto i en el país c en el periodo de tiempo t , cuya formulación se encuentra detallada en la Ecuación (10):

$$M_{i,c,t} = \begin{cases} 1, & \text{si } RCA_{i,c,t} \geq 1 \\ 0, & \text{en otro caso} \end{cases} \quad (10)$$

Donde:

$RCA_{i,c,t}$: Índice *Revealed Comparative Advantage* para la exportación del tipo de producto i realizada por el país c en el periodo de tiempo t

Así, un valor de $M_{i,c,t}$ de uno indicará que el país c posee ventajas comparativas en la exportación del producto i en el periodo de tiempo t y un valor de cero indicará que no las posee. De esta manera estamos exigiendo que las exportaciones del producto i para el país c en el periodo de tiempo t sean realmente sustanciales, revelando si el país posee ventajas comparativas en la exportación del producto i en el periodo de tiempo t o no.

3.5.3.2.2. Densidad Productiva

La Densidad Productiva cuantifica la probabilidad de producir con ventajas comparativas un nuevo producto en un país específico en base a su portafolio productivo actual. Toma valores entre cero y uno, donde valores cercanos a la unidad indican la presencia de mayores competencias necesarias para la producción de un nuevo producto i en el país c para el periodo de tiempo t . Su formulación se encuentra presentada en la Ecuación (11).

$$dens_{i,c,t} = \frac{\sum_i \Phi_{i,j,t} x_{j,c,t}}{\sum_i \Phi_{i,j,t}} \quad (11)$$

Con:

$$x_{j,c,t} = \begin{cases} 1, & \text{si } RCA_{j,c,t} \geq 1 \\ 0, & \text{en otro caso} \end{cases} \quad (12)$$

Donde:

- $RCA_{j,c,t}$: Índice *Revealed Comparative Advantage* para la categoría de producto j en el país c en el periodo de tiempo t .
- $\Phi_{i,j,t}$: Proximidad entre la categoría de producto i y j en el periodo de tiempo t . Su cálculo se presenta en la Sección 3.1.2.
- $dens_{i,c,t}$: Densidad de la categoría de producto i en el país c en el periodo de tiempo t .

Así, una categoría de producto que tenga una densidad cercana a uno determinará una alta relación entre aquella y el portafolio productivo del país c en el periodo de tiempo t , evidenciando así un entorno productivo altamente desarrollado y conectado con el producto i . Por otra parte, una densidad cercana a cero determinará una baja relación entre aquella categoría y el portafolio productivo del país c en el periodo de tiempo t , evidenciando así, un entorno productivo poco conectado al producto i .

3.5.3.2.3. Densidades Cruzadas

La variable *densidad cruzada* o *cross-density* es presentada por Catalán et al. (2019) y es utilizada para combinar la relación entre los campos científicos y tecnológicos con la experiencia de un país en el desarrollo de tecnologías, con el fin de cuantificar la proximidad de una nueva tecnología a la especialización científica del país. Esta toma valores entre cero y uno y en términos analíticos, la *densidad cruzada* ciencia-tecnología corresponde a la proximidad cruzada promedio entre una clase de tecnología l y el portafolio científico del país c durante el periodo de tiempo t .

Para la presente memoria de título se calcularán las densidades cruzadas ciencia-producto (kp) y tecnología-producto (tp) de manera análoga a la presentada por los autores. Además, diversos autores han señalado la existencia de rezagos temporales entre la generación de conocimiento científico y su traspaso hacia la generación de nuevas tecnologías, y desde la generación de nuevas tecnologías hacia la producción de nuevos bienes (Finardi, 2011; Gerken, Moehrle, & Walter, 2015; Hausmann & Klinger, 2007; Wang & Hagedoorn, 2014), por lo cuál se hace necesario integrar a la dimensión del tiempo en las densidades cruzadas un término que cuantifique este rezago temporal, denominado *lag*, el cuál es identificado para las densidades cruzadas ciencia-producto y tecnología-producto por separado. Dichas formulaciones se encuentran presentadas en las Ecuaciones (13) y (15).

- **Densidad Cruzada Ciencia-Producto (kp):**

$$xdens_{i,c,t-lag'}^{kp} = \frac{\sum_k \Phi_{i,k,t-lag'}^x x_{k,c,t-lag'}}{\sum_k \Phi_{i,k,t-lag'}^x} \quad (13)$$

Con:

$$x_{k,c,t-lag'} = \begin{cases} 1, & \text{si } RCA_{k,c,t-lag'} \geq 1 \\ 0, & \text{en otro caso} \end{cases} \quad (14)$$

Donde:

$RCA_{k,c,t-lag'}$: Índice *Revealed Comparative Advantage* para el t3pico cient3fico k en el pa3s c en el periodo de tiempo $t-lag'$.

$\Phi_{i,k,t-lag'}^x$: Proximidad cruzada ciencia-producto entre el tipo de producto i y el t3pico cient3fico k en el periodo de tiempo $t-lag'$.

$xdens_{i,c,t-lag'}^{kp}$: Densidad cruzada ciencia-producto para la categor3a de producto i en el pa3s c en el periodo de tiempo $t-lag'$.

- **Densidad Cruzada Tecnolog3a-Producto (tp):**

$$xdens_{i,c,t-lag''}^{tp} = \frac{\sum_l \Phi_{i,l,t-lag''}^x x_{l,c,t-lag''}}{\sum_l \Phi_{i,l,t-lag''}^x} \quad (15)$$

Con:

$$x_{l,c,t-lag''} = \begin{cases} 1, & \text{si } RCA_{l,c,t-lag''} \geq 1 \\ 0, & \text{en otro caso} \end{cases} \quad (16)$$

Donde:

$RCA_{l,c,t-lag''}$: Índice *Revealed Comparative Advantage* para la tecnolog3a l en el pa3s c en el periodo de tiempo $t-lag''$.

$\Phi_{i,l,t-lag''}^x$: Proximidad cruzada tecnolog3a-producto entre el tipo de producto i y la tecnolog3a l en el periodo de tiempo $t-lag''$.

$xdens_{i,c,t-lag''}^{tp}$: Densidad cruzada tecnolog3a-producto para la categor3a de producto i en el pa3s c en el periodo de tiempo $t-lag''$.

De esta manera, valores de densidad cruzada más cercanos a la unidad indicarán que el país c posee un alto grado de desarrollo científico o tecnológico en áreas cercanas al tipo de producto i durante el periodo de tiempo $t-lag$, mientras que, a su vez, valores más cercanos al cero evidenciarán que el país c posee un bajo grado de desarrollo científico o tecnológico en áreas cercanas al tipo de producto i durante el periodo de tiempo $t-lag$. Los valores de rezagos temporales lag' y lag'' fueron calculados empíricamente, utilizando un máximo de 10 años de rezago y obteniéndose 6 años de lag para la relación ciencia-producto y 3 años de lag para la relación tecnología-producto con respecto al año $t+1$.

3.5.3.2.4. Factores Fijos

Las variables previamente mencionadas capturan el efecto que generan las capacidades científicas, tecnológicas y productivas en la generación de ventajas comparativas en el desarrollo de nuevos productos, pero existen además otros factores de interés, intrínsecos a las características propias de cada categoría de producto, país y año que es importante agregar al modelo para observar como afectan a la generación de ventajas comparativas. Estas tienen como finalidad capturar los efectos asociados a cada una de estas características y su incorporación al modelo es a través del uso de variables dummy¹⁷, utilizando para ello 781 variables para cada tipo de producto específico (categorías en las cuales se presentan datos para el año t y el año $t+1$), 119 para cada país en específico y 29 para los años analizados en el estudio.

3.6. Construcción del Modelo de Regresión

Luego de definidas las variables a considerar se construye el modelo de regresión lineal múltiple a aplicar, el cuál se ajusta por medio de mínimos cuadrados ordinarios, para analizar los efectos que tienen las capacidades científicas, tecnológicas y productivas sobre la generación de ventajas comparativas en la exportación de productos a nivel país, el cuál es presentado en la Ecuación (17):

$$M_{i,c,t+1} = \beta_1 * M_{i,c,t} + \beta_2 * dens_{i,c,t} + \beta_3 * xdens_{i,c,t-5}^{kp} + \beta_4 * xdens_{i,c,t-2}^{tp} + C * \underline{\gamma} + P * \underline{\delta} + A * \underline{\varphi} + \varepsilon_{i,c,t} \quad (17)$$

Con:

$$\underline{\gamma} = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_{781} \end{pmatrix}_{(781 \times 1)}, \underline{\delta} = \begin{pmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_{119} \end{pmatrix}_{(119 \times 1)}, \underline{\varphi} = \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_{29} \end{pmatrix}_{(29 \times 1)} \quad (18)$$

¹⁷ Una variable dummy es una variable binaria, es decir que toma los valores cero o uno, cuyo objetivo es describir información cualitativa sobre el modelo de regresión.

$$\begin{aligned}
 C &= \begin{bmatrix} C_1 & 0 & 0 & 0 \\ 0 & C_2 & \cdots & 0 \\ 0 & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & C_{781} \end{bmatrix}_{(781 \times 781)} \\
 P &= \begin{bmatrix} P_1 & 0 & 0 & 0 \\ 0 & P_2 & \cdots & 0 \\ 0 & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & P_{119} \end{bmatrix}_{(119 \times 119)} \\
 A &= \begin{bmatrix} A_1 & 0 & 0 & 0 \\ 0 & A_2 & \cdots & 0 \\ 0 & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_{29} \end{bmatrix}_{(29 \times 29)}
 \end{aligned} \tag{19}$$

Donde:

- i : Índice de categorías de productos. En la presente Memoria de Título, $i \in [1, 781]$.
- c : Índice de países. En la presente Memoria de Título, $c \in [1, 119]$.
- t : Índice de años. En la presente Memoria de Título, $t \in [1, 29]$.
- $M_{i,c,t+1}$: Existencia de ventajas comparativas en la exportación de la categoría de producto i por el país c en el periodo de tiempo $t+1$.
- $M_{i,c,t}$: Existencia de ventajas comparativas en la exportación de la categoría de producto i por el país c en el periodo de tiempo t .
- $dens_{i,c,t}$: Densidad de la categoría de producto i en el país c en el periodo de tiempo t .
- $xdens_{i,c,t-5}^{kp}$: Densidad cruzada ciencia-producto para la categoría de producto i en el país c en el periodo de tiempo $t-5$.
- $xdens_{i,c,t-2}^{tp}$: Densidad cruzada tecnología-producto para la categoría de producto i en el país c en el periodo de tiempo $t-2$.
- $\underline{\gamma}$: Vector de coeficientes γ_i asociados a categoría de producto i . En la presente Memoria de Título, $i \in [1, 781]$.

$\underline{\delta}$:	Vector de coeficientes δ_c asociados a cada país c . En la presente Memoria de Título, $c \in [1,119]$.
$\underline{\varphi}$:	Vector de coeficientes φ_t asociados a cada año t . En la presente Memoria de Título, $t \in [1,29]$.
C :	Matriz de variables dummy c_i asociadas a cada categoría de producto i . En la presente Memoria de Título $i \in [1,781]$.
P :	Matriz de variables dummy P_c asociadas a cada país c . En la presente Memoria de Título, $c \in [1,119]$.
A :	Matriz de variables dummy A_t asociadas a cada año t . En la presente Memoria de Título, $t \in [1,29]$.
$\varepsilon_{i,c,t}$:	Error de estimación.

De esta forma, el modelo busca capturar el efecto de las capacidades científicas, tecnológicas y productivas sobre la generación de ventajas comparativas en la exportación de productos.

Capítulo 4: Resultados y Análisis

El presente capítulo muestra los resultados obtenidos con respecto a la influencia que poseen las capacidades científicas, tecnológicas y productivas en la generación de ventajas comparativas en la exportación de productos en los países, analizándolas con respecto a la aplicación del análisis estadístico descrito en la sección anterior.

4.1. Modelos Estadísticos

En la tabla 4.1 se presentan los resultados obtenidos por tres modelos econométricos realizados. En el primer escenario (Modelo 1) se incluyen las variables correspondientes a la presencia de ventajas comparativas en la exportación de la categoría de producto i para el país c en el periodo de tiempo t ($M_{i,c,t}$) junto con la proximidad promedio que posee una categoría de productos i con respecto al portafolio productivo del país c en el periodo de tiempo t ($dens_{i,c,t}$), de forma de aislar el efecto que poseen solo los productos en la generación de ventajas comparativas.

Para el segundo escenario (Modelo 2) se incluyeron a dichas variables la proximidad promedio que posee una categoría de productos i con respecto tanto al portafolio científico como tecnológico del país c en el periodo de tiempo t , a través de las variables $xdens_{i,c,t-5}^{kp}$ y $xdens_{i,c,t-2}^{tp}$, con la finalidad de cuantificar los efectos de las capacidades científicas, tecnológicas y productivas en conjunto. Para ambos escenarios se utilizó un modelo pooled OLS con corrección de White.

Además, cómo es de interés analizar como afectan las capacidades intrínsecas de las categorías de productos, países y años presentes en el estudio, se realizó un tercer escenario (Modelo 3) que incorporara a las variables previamente ingresadas (Modelo 2), los efectos fijos de estas capacidades con la finalidad de cuantificar dichos efectos. Este escenario utiliza un modelo de efectos fijos, también con corrección de White, el cual será analizado en el apartado posterior debido a que es el que más información provee además de ser el que posee un mayor R^2 ajustado, es decir, el que mayor precisión posee de los tres escenarios.

Tabla 4.1: Análisis estadístico en el periodo 1988 – 2017

	Modelo 1	Modelo 2	Modelo 3
$M_{i,c,t}$	0.81053 (0.00062)	0.81247 (0.00067)	0.78598 (0.00083)
$dens_{i,c,t}$	0.19562 (0.00086)	0.19908 (0.00180)	0.42645 (0.00474)
$xdens_{i,c,t-5}^{kp}$		0.00253 (0.00097)	0.04992 (0.00302)
$xdens_{i,c,t-2}^{tp}$		-0.19718 (0.00140)	0.08534 (0.00481)
R ² ajustado	0.7643	0.7670	0.7685
Efectos Fijos por Categoría	No	No	Si
Efectos Fijos por País	No	No	Si
Efectos Fijos por Año	No	No	Si
Nº de Observaciones	2.192.761	1.840.655	1.840.655

Fuente: Elaboración propia

4.2. Modelo de Efectos Fijos

Tanto la existencia de ventajas comparativas en la exportación de productos, como también la densidad de productos y las densidades cruzadas ciencia-producto y ciencia-tecnología poseen efectos positivos y significativos con respecto a la generación de ventajas comparativas en la exportación de productos en un periodo de tiempo posterior.

Respecto al análisis de los coeficientes obtenidos en el modelo 3, la estimación indica que una gran parte de la existencia de ventajas comparativas en el periodo de tiempo $t+1$ se ve explicada por la existencia de ventajas comparativas en el periodo de tiempo anterior t , indicando que los países que posean estas ventajas comparativas tendrán un 78.59% más de probabilidades de generar ventajas comparativas en la exportación del producto en el año posterior, inferencia similar a la que encuentran Hausmann and Klinger (2007). Este efecto concuerda con lo observado en la realidad debido a que países que se encuentran actualmente exportando productos con ventajas comparativas, como es el caso de Estados Unidos y Corea del Sur con su dominio en el mercado tecnológico o el caso de Chile con la exportación del cobre, se han mantenido relevantes con el paso del tiempo luego de haber generado estas ventajas de manera inicial.

En relación a los efectos que las capacidades científicas, tecnológicas y productivas que los países pudiesen tener con respecto a la generación de ventajas comparativas en la exportación de productos, los resultados indican que un aumento en un 10% en la densidad de productos reflejará un aumento de un 4.26% en la probabilidad de que en el país c obtenga ventajas comparativas en un periodo de tiempo posterior $t+1$. De igual forma, un aumento de un 10% en la densidad cruzada ciencia-producto y tecnología-producto reflejará un aumento de 0.49% y 0.85% respectivamente en la probabilidad de que en el país c demuestre ventajas comparativas en un periodo de tiempo posterior $t+1$.

De aquello se puede desprender que, para aumentar la generación de ventajas comparativas en la exportación de la categoría de productos i , los países se verán beneficiados si aumentan la cantidad de productos que exportan con ventajas comparativas que estén relacionados a aquella categoría de producto, ya que eso aumentaría la proximidad promedio (densidad) de aquella categoría con el portafolio productivo del país y , por consiguiente, aumentaría la probabilidad de exponer ventajas comparativas en la producción del bien i , razonamiento análogo a los presentados por Neffke et al. (2011), Hausmann and Klinger (2007), Boschma et al. (2013) o Zhu et al. (2017) entre otros.

De igual manera, el desarrollo de ciencia y tecnología relacionada a la categoría de producto i que el país c integre a su portafolio científico y tecnológico aumentara las densidades cruzadas ciencia-producto y tecnología-producto respectivamente, ampliando así la probabilidad de exponer ventajas comparativas en la producción del producto i . Eso si, estas probabilidades se verán reflejadas en menor medida con relación al portafolio productivo del país debido a la menor influencia que poseen sobre la generación de ventajas competitivas en la exportación de productos.

4.3. Implicancia en Políticas Públicas

Los resultados obtenidos en esta Memoria de Título presentan varias implicancias que pueden ser utilizadas para generar estrategias de desarrollo en países a través de la implementación de políticas públicas que fomenten la diversificación de los portafolios científicos, tecnológicos y productivos con la finalidad de aumentar la cantidad de áreas en las que un país posee ventajas comparativas, promoviendo así su desarrollo y crecimiento.

Antes que nada, los países deben ser capaces de mantener las ventajas comparativas que poseen en la producción de bienes debido a que siempre existe una probabilidad de que, sin considerar algún efecto que podría ocurrir en el mercado de aquel bien, si no se preocupan de mantenerse eficientes a través del tiempo, puedan llegar incluso a perder las ventajas que poseían y dejar de ser un país líder en aquel mercado como ocurrió en el caso de Linköping en Suecia presentado en Neffke et al. (2011).

Además, los países deben buscar mantener las ventajas comparativas que poseen en los productos donde se detecten mayor cantidad de conexiones entre su portafolio productivo y el resto de bienes con la finalidad de generar, a largo plazo, mayores oportunidades de diversificación. De lo contrario, mantenerse líderes en productos con pocas conexiones podría llevar al país a un estancamiento en su desarrollo a largo plazo. Las estrategias que busquen mantener dentro del portafolio productivo bienes que aumenten las posibilidades de diversificación deben generar ecosistemas productivos que fomenten la generación de productos relacionados con la finalidad de obtener fuentes de ingreso mas variadas que produzcan crecimiento y desarrollo en el país y, a su vez, proyectar el crecimiento a través del tiempo mediante el desarrollo de productos mas rentables y conectados.

Por otro lado, si bien las relaciones existentes entre las capacidades científicas y tecnológicas con la generación de productos obtenidas en esta Memoria de Título son bajas en comparación a la existente entre los mismos productos, estas son significativas y, por lo tanto, es recomendado para los países fomentar también la generación de ciencia y tecnología relacionadas con las capacidades productivas que poseen actualmente. A su vez, es recomendado incursionar en ciencia y tecnología que posea bastante conexiones con productos dentro del mercado que sean altamente llamativos y desarrollados, con la finalidad de aumentar las probabilidades de generar ventajas comparativas en la producción de aquellos bienes. En este punto, los entes participantes ya no son solo las empresas del país, debido a que se pueden realizar concursos y convenios que fomenten la generación de dicho conocimiento en universidades, institutos de formación técnica o cualquier institución que se dedique a generar conocimiento, con la idea de conectar ambos mundos y poder generar un aumento en las capacidades innovativas dentro de los países en relación a las áreas que sean definidas por la política pública.

Finalmente, las políticas públicas planificadas para generar crecimiento y desarrollo en los países deben ser enfocadas en fomentar ecosistemas innovadores los cuales contengan productos, tecnologías y ciencias que se relacionen entre si para aumentar la probabilidad de generar ventajas competitivas en dicho ecosistema. Como se ha observado en los resultados de esta Memoria de Título, además de las conclusiones a las que han llegado Hidalgo et al. (2018), Catalán et al. (2019), Neffke et al. (2011), Hausmann and Klinger (2007), Boschma et al. (2013), Boschma, Heimeriks, et al. (2014), entre otros, las conexiones que existen entre las distintas categorías de conocimientos científicos, tecnologías o productos son un motor fundamental en el proceso de diversificación hacia nuevas áreas de desarrollo, generando así mayores probabilidades de éxito al momento de enfrentarse a la decisión de desarrollar alguno de estos tópicos en contraparte a hacerlo desarrollando solamente un tipo de conocimiento científico, tecnológico o productivo.

Capítulo 5: Conclusiones

El presente estudio analizó las relaciones existentes entre las capacidades científicas, tecnológicas y productivas, y la generación de ventajas comparativas con respecto a la exportación de 782 categorías de productos en 120 países durante un periodo de 30 años (1988-2017). Para esto, se examinaron las proximidades promedio que tenían los portafolios científicos, tecnológicos y productivos, utilizando como proxies los registros de publicaciones científicas, patentes tecnológicas y exportaciones de productos respectivamente con el objetivo de aplicar metodologías bibliográficas mediante la creación de un set de variables que cuantifiquen dichas relaciones. Las publicaciones científicas fueron obtenidas de la base de datos de *WoS*, las patentes tecnológicas desde *USPTO* y las exportaciones de productos desde el *OEC*, proyecto perteneciente al *MIT MediaLab*.

Con ese propósito, se realizó un análisis cuantitativo a través de un modelo de regresión lineal múltiple con efectos fijos con el objetivo de identificar y cuantificar como afectan las capacidades científicas, tecnológicas y productivas a la generación de nuevas ventajas comparativas en la exportación de productos. Para ello, las variables del modelo reflejan la existencia de ventajas comparativas en la exportación de productos para un periodo t , la relación entre la categoría de producto i y el portafolio productivo del país y también de dicha categoría con el portafolio científico y tecnológico del país (cada una por separado), además de los efectos intrínsecos que cada categoría de productos, país y año posee. Entre los resultados obtenidos se establece que el principal factor que aumenta la probabilidad de poseer ventajas comparativas en la exportación del producto i en el periodo de tiempo $t+1$ es que el país ya tuviera ventajas comparativas en el periodo de tiempo t , seguido por la cercanía de la categoría de producto hacia el portafolio productivo del país y, en menor medida, las relaciones entre esta categoría y los portafolios tecnológicos y científicos.

De igual forma a los análisis presentados por Hausmann and Klinger (2007), Hidalgo et al. (2007) o Neffke et al. (2011), entre otros, además de inspirados por las relaciones presentadas en Catalán et al. (2019), y en concordancia con el “*Principle of Relatedness*” presentado por Hidalgo et al. (2018), se detecta que el principal motor que genera mayores probabilidades de diversificación exitosa a los países son las relaciones existentes entre las categorías de productos presentes en su portafolio productivo. Si bien existe una relación positiva entre la generación de ventajas comparativas en la exportación de productos y los portafolios científicos y tecnológicos, estas son bajas en comparación al portafolio productivo y, por lo tanto, requieren de un esfuerzo e inversión mucho mayor para ser relevantes. Es por ello que es importante generar políticas públicas que fomenten la generación de un ecosistema productivo interconectado entre si con la finalidad de poseer productos con alta relación en los portafolios productivos de los países y, de esta manera, aumentar la probabilidad de éxito en la generación de ventajas comparativas en la exportación de productos.

Considerando las relaciones obtenidas en este trabajo, se hace necesario mencionar que este es uno de los primeros análisis que se dedican a relacionar las capacidades científicas, tecnológicas y productivas con respecto a la generación de ventajas comparativas en la producción de bienes y, por lo tanto, se debe seguir profundizando en la generación de conocimiento en el área con el fin de evaluar e identificar otros factores que relacionen las tres variables y sean relevantes en la generación de ventajas comparativas en la producción y desarrollo de nuevos productos.

Capítulo 6: Bibliografía

- Artz, K. W., Norman, P. M., Hatfield, D. E., & Cardinal, L. B. (2010). A Longitudinal Study of the Impact of R&D, Patents, and Product Innovation on Firm Performance. *Journal of Product Innovation Management*, 27(5), 725-740. Retrieved from <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-5885.2010.00747.x>. doi:10.1111/j.1540-5885.2010.00747.x
- Boschma, R., Balland, P.-A., & Kogler, D. F. (2014). Relatedness and technological change in cities: the rise and fall of technological knowledge in US metropolitan areas from 1981 to 2010. *Industrial and Corporate Change*, 24(1), 223-250. Retrieved from <https://doi.org/10.1093/icc/dtu012>. doi:10.1093/icc/dtu012
- Boschma, R., & Capone, G. (2015). *Institutions and diversification: Related versus unrelated diversification in a varieties of capitalism framework* (Vol. 44).
- Boschma, R., Heimeriks, G., & Balland, P.-A. (2014). Scientific knowledge dynamics and relatedness in biotech cities. *Research Policy*, 43(1), 107-114.
- Boschma, R., Minondo, A., & Navarro, M. (2013). The Emergence of New Industries at the Regional Level in Spain: A Proximity Approach Based on Product Relatedness. *Economic Geography*, 89(1), 29-51.
- Catalán, P., Navarrete, C., & Figueroa, F. (2019). *Scientific and technological Cross-Space: Is technological diversification driven by scientific endogenous capacity?* Center of Innovation System Studies.
- Colombelli, A., Krafft, J., & Quatraro, F. (2013). The emergence of new technology-based sectors at the regional level: a proximity-based analysis of nanotechnology.
- Criscuolo, P. (2006). The 'home advantage' effect and patent families. A comparison of OECD triadic patents, the USPTO and the EPO. *Scientometrics*, 66(1), 23-41. Retrieved from <https://doi.org/10.1007/s11192-006-0003-6>. doi:10.1007/s11192-006-0003-6
- do Canto Cavalheiro, G. M., Joia, L. A., & Van Veenstra, A. F. (2016). Examining the trajectory of a standard for patent classification: An institutional account of a technical cooperation between EPO and USPTO. *Technology in Society*, 46, 10-17.
- Essletzbichler, J. (2015). Relatedness, Industrial Branching and Technological Cohesion in US Metropolitan Areas. *Regional Studies*, 49(5), 752-766. Retrieved from <https://doi.org/10.1080/00343404.2013.806793>. doi:10.1080/00343404.2013.806793
- Falagas, M. E., Pitsouni, E. I., Malietzis, G. A., & Pappas, G. (2008). Comparison of PubMed, Scopus, Web of Science, and Google Scholar: strengths and weaknesses. *The FASEB Journal*, 22(2), 338-342. Retrieved from <https://www.fasebj.org/doi/abs/10.1096/fj.07-9492LSF>. doi:10.1096/fj.07-9492LSF
- Finardi, U. (2011). Time relations between scientific production and patenting of knowledge: the case of nanotechnologies. *Scientometrics*, 89(1), 37. Retrieved from <https://doi.org/10.1007/s11192-011-0443-5>. doi:10.1007/s11192-011-0443-5
- Gerken, J. M., Moehrle, M. G., & Walter, L. (2015). One year ahead! Investigating the time lag between patent publication and market launch: insights from a longitudinal study in the automotive industry. *R&D Management*, 45(3), 287-303. Retrieved from <https://onlinelibrary.wiley.com/doi/abs/10.1111/radm.12085>. doi:10.1111/radm.12085

- Guevara, M. R., Hartmann, D., Aristarán, M., Mendoza, M., & Hidalgo, C. A. (2016). The Research Space: using the career paths of scholars to predict the evolution of the research output of individuals, institutions, and nations. *arXiv e-prints*. Retrieved from <https://ui.adsabs.harvard.edu/abs/2016arXiv160208409G>
- Hartmann, D., Guevara, M. R., Jara-Figueroa, C., Aristarán, M., & Hidalgo, C. A. (2017). Linking Economic Complexity, Institutions, and Income Inequality. *World Development*, 93, 75-93. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0305750X15309876>. doi:<https://doi.org/10.1016/j.worlddev.2016.12.020>
- Hausmann, R., & Klinger, B. (2007). *The Structure of the Product Space and the Evolution of Comparative Advantage*.
- Hidalgo, Balland, P.-A., Boschma, R., Delgado, M., Feldman, M., Frenken, K., . . . Zhu, S. (2018). The Principle of Relatedness: Proceedings of the Ninth International Conference on Complex Systems. In (pp. 451-457).
- Hidalgo, Klinger, B., Barabási, A.-L., & Hausmann, R. (2007). The product space conditions the development of nations. *Science*, 317(5837), 482-487.
- Mindruta, D., Albino-Pimentel, J., & Hoetker, G. (2015). *Institutions and the Direction of Innovative Search: Change and Persistence Between and Within Countries*.
- Neffke, F., & Henning, M. (2013). Skill relatedness and firm diversification. *Strategic Management Journal*, 34(3), 297-316. Retrieved from <https://onlinelibrary.wiley.com/doi/abs/10.1002/smj.2014>. doi:10.1002/smj.2014
- Neffke, F., Henning, M., & Boschma, R. (2011). How do regions diversify over time? Industry relatedness and the development of new growth paths in regions. *Economic Geography*, 87(3), 237-265.
- Petralia, S., Balland, P.-A., & Morrison, A. (2017). Climbing the ladder of technological development. *Research Policy*, 46(5), 956-969.
- Quatraro, F. (2010). Knowledge coherence, variety and economic growth: Manufacturing evidence from Italian regions. *Research Policy*, 39(10), 1289-1302.
- Rigby, D. L. (2015). Technological relatedness and knowledge space: entry and exit of US cities from patent classes. *Regional Studies*, 49(11), 1922-1937.
- Saviotti, P. P., & Frenken, K. (2008). Export variety and the economic performance of countries. *Journal of Evolutionary Economics*, 18(2), 201-218.
- Wang, N., & Hagedoorn, J. (2014). The lag structure of the relationship between patenting and internal R&D revisited. *Research Policy*, 43(8), 1275-1285.
- Zhu, S., He, C., & Zhou, Y. (2017). How to jump further and catch up? Path-breaking in an uneven industry space. *Journal of Economic Geography*, 17(3), 521-545.

Anexo A. Listado de Países con formato ISO 3166-2

Nombre	alpha-2	alpha-3
Angola	AO	AGO
United Arab Emirates	AE	ARE
Argentina	AR	ARG
Armenia	AM	ARM
Australia	AU	AUS
Austria	AT	AUT
Azerbaijan	AZ	AZE
Belgium	BE	BEL
Burkina Faso	BF	BFA
Bangladesh	BD	BGD
Bulgaria	BG	BGR
Bahrain	BH	BHR
Bosnia and Herzegovina	BA	BIH
Belarus	BY	BLR
Bolivia (Plurinational State of)	BO	BOL
Brazil	BR	BRA
Botswana	BW	BWA
Canada	CA	CAN
Switzerland	CH	CHE
Chile	CL	CHL
China	CN	CHN
Côte d'Ivoire	CI	CIV
Cameroon	CM	CMR
Colombia	CO	COL
Costa Rica	CR	CRI
Cuba	CU	CUB
Czechia	CZ	CZE
Germany	DE	DEU
Denmark	DK	DNK
Dominican Republic	DO	DOM
Algeria	DZ	DZA
Ecuador	EC	ECU
Egypt	EG	EGY
Spain	ES	ESP

Estonia	EE	EST
Finland	FI	FIN
France	FR	FRA
Gabon	GA	GAB
United Kingdom of Great Britain and Northern Ireland	GB	GBR
Georgia	GE	GEO
Greece	GR	GRC
Guatemala	GT	GTM
Hong Kong	HK	HKG
Honduras	HN	HND
Croatia	HR	HRV
Haiti	HT	HTI
Hungary	HU	HUN
Indonesia	ID	IDN
India	IN	IND
Ireland	IE	IRL
Iran (Islamic Republic of)	IR	IRN
Iraq	IQ	IRQ
Israel	IL	ISR
Italy	IT	ITA
Jamaica	JM	JAM
Jordan	JO	JOR
Japan	JP	JPN
Kazakhstan	KZ	KAZ
Kenya	KE	KEN
Kyrgyzstan	KG	KGZ
Korea, Republic of	KR	KOR
Kuwait	KW	KWT
Lao People's Democratic Republic	LA	LAO
Lebanon	LB	LBN
Liberia	LR	LBR
Sri Lanka	LK	LKA
Lithuania	LT	LTU
Latvia	LV	LVA
Morocco	MA	MAR
Madagascar	MG	MDG
Mexico	MX	MEX
North Macedonia	MK	MKD

Mali	ML	MLI
Myanmar	MM	MMR
Mauritius	MU	MUS
Malaysia	MY	MYS
Niger	NE	NER
Nigeria	NG	NGA
Netherlands	NL	NLD
Norway	NO	NOR
New Zealand	NZ	NZL
Oman	OM	OMN
Pakistan	PK	PAK
Panama	PA	PAN
Peru	PE	PER
Philippines	PH	PHL
Papua New Guinea	PG	PNG
Poland	PL	POL
Korea (Democratic People's Republic of)	KP	PRK
Portugal	PT	PRT
Paraguay	PY	PRY
Qatar	QA	QAT
Romania	RO	ROU
Russian Federation	RU	RUS
Saudi Arabia	SA	SAU
Senegal	SN	SEN
Singapore	SG	SGP
Sierra Leone	SL	SLE
El Salvador	SV	SLV
Serbia	RS	SRB
Slovakia	SK	SVK
Slovenia	SI	SVN
Sweden	SE	SWE
Syrian Arab Republic	SY	SYR
Chad	TD	TCD
Togo	TG	TGO
Thailand	TH	THA
Trinidad and Tobago	TT	TTO
Tunisia	TN	TUN
Turkey	TR	TUR

Taiwan, Province of China	TW	TWN
Uganda	UG	UGA
Ukraine	UA	UKR
Uruguay	UY	URY
United States of America	US	USA
Uzbekistan	UZ	UZB
Venezuela (Bolivarian Republic of)	VE	VEN
Viet Nam	VN	VNM
South Africa	ZA	ZAF
Zimbabwe	ZW	ZWE

UNIVERSIDAD DE CONCEPCION – FACULTAD DE INGENIERIA RESUMEN DE MEMORIA DE TITULO

Departamento de Ingeniería		Industrial	
Título		Identificación del efecto de capacidades científicas, tecnológicas y productivas en la generación y desarrollo de nuevos productos	
Nombre Memorista		Nicolás Sebastián Pérez Mundaca	
Modalidad	Investigación	Profesor(es) Patrocinante	
Concepto		Pablo Eduardo Catalán Martínez	
Calificación			
Fecha		Ingeniero Supervisor	Institución
Comisión (Nombre y Firma)			
Rodrigo de la Fuente Gallegos			
Resumen			
<p>La presente investigación realiza un análisis de los efectos que poseen las capacidades científicas, tecnológicas y productivas con respecto a la generación y desarrollo de nuevos productos mediante la estimación de un modelo de regresión lineal múltiple a través del método de mínimos cuadrados ordinarios (MCO). Para realizar este modelo, es construida una base de datos de publicaciones científicas, patentes y exportación de productos provenientes de Web of Science, United States Patent and Trademark Office (USPTO) y el OEC del MIT Media Lab, quienes utilizan los datos de la United Nations International Trade Statistics Database (UN Comtrade), respectivamente, abarcando un total de 254 categorías en publicaciones científicas, 661 tipos de patentes y 782 categorías de productos para un total de 120 países en el periodo comprendido entre 1988 y el 2017.</p> <p>De esta forma, se observa que las capacidades científicas, tecnológicas y productivas afectan positivamente la generación de ventajas comparativas en el desarrollo de productos, siendo la más relevante la relación entre la cartera productiva del país, seguido en menor medida por las capacidades tecnológicas y científicas del mismo. En base a estos resultados, se sugieren políticas públicas que fomenten el desarrollo y la innovación de ecosistemas productivos con la finalidad de aumentar las relaciones existentes entre los productos pertenecientes al portafolio productivo del país y así poseer una mayor cantidad de oportunidades de diversificación. Además, debido a la gran influencia que poseen las capacidades productivas por sobre las científicas y tecnológicas, es que las políticas públicas que deseen incentivar la generación de nuevos productos a través del desarrollo de ciencia y tecnología deben ser disruptivas, de manera de desarrollar ecosistemas productivos que fomenten la generación de un set de productos en específico con la finalidad de romper la fuerte influencia de las actuales capacidades productivas y generar así un mayor abanico de oportunidades de diversificación.</p>			