



Facultad De Ciencias Económicas Y Empresariales

ANÁLISIS SOBRE EL FUTURO DEL TELETRABAJO EN EUROPA MEDIANTE UN MODELO PREDICTIVO: COMPARATIVA ENTRE ALEMANIA, ITALIA Y PORTUGAL

Autor: María de los Reyes Ruiz-Labourdette Gallardo

Director: Pedro César Martínez Morán

MADRID | Junio 2023

ÍNDICE

1. RESUMEN Y ABSTRACT	1
2. INTRODUCCIÓN Y OBJETIVOS	2
3. REVISIÓN DE LA LITERATURA	5
3.1. MARCO TEÓRICO GENERAL.....	5
3.2. ALEMANIA.....	7
3.3. ITALIA.....	9
3.4. PORTUGAL.....	11
3.5. FORMULACIÓN DE HIPÓTESIS.....	13
4. METODOLOGÍA.....	13
4.1. EXPLICACIÓN DE LA BASE DE DATOS.....	13
4.2. LA VARIABLE DEPENDIENTE	14
4.3. LAS VARIABLES INDEPENDIENTES	16
4.4. NORMALIZACIÓN DE LAS VARIABLES INDEPENDIENTES	17
4.5. MODELOS.....	18
4.6. ESTIMACIÓN DE VARIABLES.....	20
4.6.1. La media del diferencial del incremento.....	21
4.6.2. Estimación con regresión lineal de las variables independientes	25
4.6.3. Comparación de los resultados según el método de estimación.....	28
5. RESULTADOS DE LA INVESTIGACIÓN	33
5.1. EVALUACION DE LOS MODELOS.....	33
5.1.1. Modelo de Regresión Lineal.....	33
5.1.2. Modelo de Regresión Ridge.....	37
5.1.3. Modelo de Regresión Lasso.....	39
5.2. RESULTADO PREDICCIONES.....	40
5.2.1. Modelo de Regresión Lineal.....	41
5.2.2. Modelo de Regresión Ridge.....	42

5.2.3.	Modelo de Regresión Lasso.....	42
5.2.4.	Análisis comparativo de las predicciones de los tres modelos	43
6.	DISCUSIÓN	46
6.1.	DESCENSO PREVISTO PARA EL AÑO 2023	47
6.2.	ASCENSO PREVISTO PARA EL AÑO 2024.....	49
6.3.	ANÁLISIS DEL NIVEL DE ESTABILIZACIÓN DEL PORCENTAJE DE TELETRABAJO DESPUÉS DEL COVID.....	51
7.	CONCLUSIONES	53
8.	ÍNDICE DE GRÁFICOS, ILUSTRACIONES Y TABLAS	56
9.	ANEXOS.....	58
10.	REFERENCIAS	59

1. RESUMEN Y ABSTRACT

RESUMEN

Este trabajo de investigación se centra en la evolución del teletrabajo a lo largo de los años y principalmente en el futuro del teletrabajo. Parece que el Covid ha acelerado en gran medida la implantación del trabajo a distancia en muchos países del mundo. Sin embargo hay muchos factores que inciden directamente en la consolidación del trabajo a distancia, como la riqueza del país, la estructura industrial, la cultura o las inversiones tecnológicas.

El presente estudio pone el foco en tres países del ámbito de la Unión Europea, que son Alemania, Italia y Portugal. En cada uno de estos países se dan parámetros diferentes en cuanto a nivel de renta, estructura industrial e inversión tecnológica. Esto justifica los diferentes valores porcentuales de teletrabajo antes, durante y después de la crisis sanitaria.

Si no se dispone de los medios necesarios, la implementación rápida y precipitada del teletrabajo puede haber conducido a una mala primera toma de contacto con esta nueva modalidad laboral, lo que puede haber infundido su rechazo. Sin embargo, la proporción de los recursos imprescindibles y la garantía de técnicas adecuadas de control, podrían conducir a una aceptación generalizada del trabajo a distancia.

Para tratar de dar respuesta a tantas preguntas acerca del futuro del teletrabajo, se ha desarrollado un modelo predictivo con una base de datos descargada de Eurostat. Los resultados de la investigación muestran una tendencia al alza en el porcentaje de personas que teletrabajan. Sin embargo, los resultados arrojan una clara diferencia entre los países, teniendo en cuenta que las cifras no han alcanzado todavía la estabilidad post-pandemia.

ABSTRACT

This research paper focuses on the evolution of telework over the years focusing on its future. It appears that Covid has greatly accelerated the implementation of telework in many countries around the world. However, there are many factors that have a direct impact on the consolidation of telecommuting, such as country wealth, industrial structure, culture or technological investments.

This study focuses on three European Union countries, which are Germany, Italy and Portugal. In each of these countries there are different parameters in terms of income level, industrial structure and technological investment. This justifies the different percentage values of telework before, during and after the pandemic.

If the necessary means are not available, the rapid and hasty implementation of telework may have led to a poor first contact with this new work modality, which may have instilled its rejection. However, the provision of the indispensable resources and the guarantee of adequate control techniques could lead to a generalized acceptance of telework.

To try to answer the many questions about the future of telework, a predictive model has been developed using a database downloaded from Eurostat. The results of the research show an upward trend in the share of telework. However, the results show a clear difference between countries, bearing in mind that the figures have not yet reached post-pandemic stability.

2. INTRODUCCIÓN Y OBJETIVOS

El presente trabajo de fin de grado se centra en el estudio del teletrabajo y en su evolución esperada tras la pandemia mundial del Covid. Esta temática ha sido elegida porque con la crisis sanitaria ha habido un crecimiento del teletrabajo, y considero que su evolución es relevante para entender los cambios a nivel laboral y social que se están produciendo en este contexto. El trabajo a distancia ha incidido en la manera en la que trabajan los empleados eliminando la exigencia de desplazamiento físico al lugar de trabajo. El estudio de este fenómeno permite comprender cómo ha afectado a la vida laboral de los trabajadores y cómo éstos se han adaptado a esta nueva modalidad de trabajo.

El teletrabajo ha obligado a las empresas a implementar nuevas estructuras organizativas y nuevos métodos de control y de gestión de la productividad. Esto implica cambios en la organización empresarial, que traerán consigo nuevas oportunidades y retos para las empresas. Esta nueva modalidad de trabajo requiere el uso de tecnología y herramientas que permitan la comunicación a distancia, lo que puede suponer un inconveniente dada la inversión económica que algunas empresas tendrían que realizar. Por último, el principal

motivo por el que se ha elegido el presente tema es que, si bien es cierto que la pandemia ha acelerado la implementación del teletrabajo, no está del todo claro qué va a pasar después. Estudiar la evolución que se espera de esta nueva modalidad laboral permite conocer de antemano cómo se va a desarrollar en el futuro, y cómo se van a adaptar las empresas y los trabajadores a los cambios.

El objetivo principal del presente trabajo es estudiar el teletrabajo y su evolución esperada. En cuanto a objetivos secundarios, se analiza el impacto de la crisis sanitaria en el teletrabajo; el grado de aceptación y acogida que ha tenido el teletrabajo en los países que han sido seleccionados para el estudio; la adaptación de las empresas y trabajadores, que puede verse afectada por la estructura industrial de los países; así como algunos de los factores que pueden influir favoreciendo o dificultando la implantación del teletrabajo.

El trabajo se estructura en cuatro grandes partes; además del resumen, la introducción y las conclusiones. La primera es la revisión de la literatura, donde se he hecho un recopilatorio de lo que dice la ciencia en relación con lo que cabe esperar del trabajo a distancia en los años venideros. Este apartado se ha dividido en cuatro epígrafes, siendo el primero el marco teórico general. En este epígrafe se ha analizado de forma genérica en qué consiste el teletrabajo, cuál ha sido su evolución hasta la pandemia mundial, y cuáles son los países con cifras de trabajo a distancia más altas en el ámbito de la Unión Europea. Para el presente trabajo se ha hecho una selección de países en los que se ha enfocado el estudio para una mayor precisión. Los países seleccionados han sido Alemania, Italia y Portugal, por lo que se ha dedicado a cada uno un epígrafe que recoge la evolución antes del Covid, las cifras durante la pandemia y las expectativas a futuro. Por último, dentro del apartado de revisión de la literatura se han formulado las dos hipótesis a contrastar con los resultados del modelo predictivo.

El siguiente apartado del trabajo es la metodología, donde se proporciona una explicación detallada del método empleado para el desarrollo del estudio. Se ha desarrollado un modelo predictivo que permite estimar el porcentaje de teletrabajo para los años futuros. Por lo tanto, el apartado comienza con la explicación de la base de datos. En él se explica cómo se ha reestructurado la información contenida en ella, y cómo se han elegido la variable dependiente y las variables predictoras o independientes. En este apartado se explica también

el proceso de normalización de las variables para evitar que las diferentes unidades de medida distorsionen los resultados; y la forma en la que se han entrenado los tres modelos en R studio. Para terminar el apartado de metodología, se han explicado los dos métodos contemplados para la estimación de las variables independientes; además, se ha justificado la elección del método de regresión lineal para estimar tras comparar los resultados.

En el apartado titulado “Resultados de la investigación” se ha recogido por un lado la evaluación cada uno de los tres modelos, y por otro el resultado de sus predicciones. En el primer epígrafe se han analizado las métricas de error asociadas a cada uno de los modelos para así poder comparar los resultados y elegir el modelo con mejor rendimiento. En el segundo epígrafe se han recogido las predicciones hechas por cada uno de los tres modelos entrenados. Se han comparado los resultados para estudiar cuánto varían las predicciones de un modelo a otro y ver si las diferencias son significativas.

Por último, en el apartado con el título “Discusión” se han interpretado las predicciones hechas por el modelo que se ha considerado que tiene un mejor rendimiento. Se ha buscado una explicación a los ascensos y descensos en las cifras de trabajo a distancia de las predicciones. Por último, se han comparado los resultados del modelo con la literatura recopilada para cada uno de los tres países; además, se han extraído las tendencias esperadas de la tasa de teletrabajo para Alemania, Italia y Portugal.

Este trabajo puede aportar contribuciones significativas a la literatura científica a cerca del teletrabajo y su evolución tras la crisis sanitaria. Proporciona una revisión detallada de la evolución del trabajo a distancia, además de la proyección de su evolución. Esto puede servir de ayuda a los investigadores y profesionales para comprender cómo está cambiando el panorama laboral. En este trabajo se estima el porcentaje de teletrabajo que se espera para los años futuros desarrollando un modelo predictivo. Esta contribución puede ser de utilidad para quienes buscan comprender las tendencias futuras del teletrabajo. Por último, se hace una comparativa entre la literatura científica existente para Alemania, Italia y Portugal, y los resultados del modelo predictivo. Esto permite contextualizar y contrastar los resultados con estudios previos ya existentes.

3. REVISIÓN DE LA LITERATURA

3.1. MARCO TEÓRICO GENERAL

El teletrabajo ha sido definitivo por la OIT como el “uso de tecnologías de la información y las comunicaciones – como teléfonos inteligentes, tabletas, computadoras portátiles y de escritorio – para trabajar fuera de las instalaciones del empleador (OIT, 2020).

Desde hace diez años el porcentaje de teletrabajo en el ámbito de la Unión Europea ha crecido lentamente. De hecho, la media de personas que trabajaban a distancia de manera habitual en la Unión Europea se mantuvo bastante constante, en torno al 5,4%, entre los años 2009 y 2019. Sin embargo, la media para el grupo de personas que a veces trabaja a distancia pasó de un 5,2% en el año 2009 a un 9% en el año 2019 (Milasi et al., 2021).

Según los datos recogidos en el territorio de la UE en el año 2019, hay países que en aquel momento ya contaban con cifras superiores al 30%; mientras que había otros que apenas alcanzaban el 5%. En cuanto a los países del presente estudio, Alemania estaba en torno al 13%, Portugal superaba ligeramente el 15% e Italia no llegaba al 5%. Las diferencias se deben en parte a las distintas estructuras industriales de empleo en cada país. Por lo general, los países que cuentan con más puestos de trabajo enfocados a los servicios intensivos en conocimiento y TIC presentan cifras más altas de trabajo a distancia (INE, 2020). Según un estudio realizado por Dingel y Neiman, existe una relación positiva entre el porcentaje de empleos que pueden desarrollarse a distancia y los niveles de renta. Esto implica que el teletrabajo supone un reto más difícil para las economías emergentes (Dingel & Neiman, 2020).

Independientemente de que el análisis realizado en el presente trabajo se enfoque exclusivamente en Alemania, Italia y Portugal, parece necesario mencionar que, en el ámbito de la Unión Europea hay ciertos países, como es el caso de Países Bajos, Finlandia y Luxemburgo, que tienen niveles bastante más altos de teletrabajo en relación con el resto de estados miembros. La tasa de personas que trabajan a distancia de manera frecuente en estos países fue respectivamente del 14%, 13,3%, y 11% en el año 2009. Dentro de estos porcentajes, más del 40% son trabajadores por cuenta propia (Dingel & Neiman, 2020).

Dentro de las categorías profesionales, el grupo de los autónomos presenta las cifras más altas de trabajo a distancia. El porcentaje de teletrabajo entre los autónomos incrementó en un 6%, pasando de un 30% en 2009 a un 36% en 2019; y el porcentaje de teletrabajo para los trabajadores por cuenta ajena aumentó en un 3,5% entre los años 2009 y 2019. Queda constatada la prevalencia del teletrabajo de los autónomos, ya que el porcentaje de empleados por cuenta ajena que trabajaba a distancia en 2019 seguía en torno al 11% (Milasi et al., 2021).

Tabla 1. Porcentaje de teletrabajo en la UE según categoría profesional

	2009	2019
Autónomos	30%	36%
Por cuenta ajena	7,5%	11%

Por otro lado, la tendencia general en la mayoría de países es que el número de mujeres en situación de teletrabajo sea superior al de hombres. Concretamente en la Eurozona el porcentaje de hombres que trabajan a distancia es del 5,5%; mientras que el porcentaje de mujeres oscila alrededor del 6,1% según los datos recogidos el año 2019 (Milasi et al., 2021).

El presente estudio pone su enfoque en tres países: Alemania, Italia y Portugal. La razón por la que se han elegido estos tres países es porque se trata de estados miembros de la Unión Europea con distintos valores de PIB per cápita. Como se ha mencionado anteriormente, hay una correlación positiva entre el nivel de renta y el porcentaje de teletrabajo, ya que, a mayor nivel de renta, mayor proporción de puestos de trabajo que permiten esta nueva modalidad de empleo (Dingel & Neiman, 2020). Por tanto, se han elegido países con valores extremos en cuanto al PIB per cápita, siendo Alemania el extremo superior, Portugal el inferior y Italia el punto medio.

A continuación se expone un análisis de la situación actual de teletrabajo para cada uno de los países elegidos, así como las futuras expectativas.

3.2. ALEMANIA

En Alemania el teletrabajo es definido por primera vez en el año 2016, cuando se reforma el reglamento de puestos de trabajo. Así, queda definido como la modalidad laboral en la que el trabajador realiza al menos una parte de su actividad laboral fuera de la empresa habiendo sido el puesto de trabajo por parte del empresario, y desarrollando el trabajo con medios facilitados por la empresa (Ministerio de Empleo y Seguridad Social, 2017; Ministerio de Empleo y Seguridad Social, 2017).

Según una encuesta realizada por el Ministerio de Interior de Alemania en el año 2017, en torno a un 22% de las personas que trabajaban en empresas del sector privado con una plantilla de más de 50 personas trabajaba a distancia al menos ocasionalmente (Grunau et al., 2020). En el caso de Alemania, esta modalidad de trabajo es muy frecuente en profesiones de carácter administrativo. Según esta encuesta, el 30% de los empleados podrían trabajar desde casa para desempeñar los trabajos clásicos de oficina.

Además, en 2019 se realizó de nuevo la misma encuesta pero esta vez preguntando cuáles eran las actividades que hacían habitualmente en un día de trabajo. Las respuestas se agruparon en tres grupos: comunicación, trabajo en ordenador y trabajo con herramientas o en máquinas y sistemas. Las dos primeras respuestas son a menudo digitales, mientras que el trabajo técnico con máquinas de momento es difícilmente digitalizable. Por lo tanto, la conclusión de la encuesta fue que, al menos el 88% de los empleados utilizaba tecnologías digitales (Grunau et al., 2020).

Por otra parte, en el año 2019 la Federación de Empresas de las TIC llevó a cabo una encuesta en la que se concluía que el 39% de las empresas contemplaba la posibilidad de trabajo a distancia, cifra que había aumentado en 9 puntos porcentuales desde 2016, año en que se regula esta modalidad laboral por primera vez. A pesar de esta facilidad contemplada por las empresas, en 2018 solo el 9% de los trabajadores optaron por la modalidad de trabajo a distancia (Ministerio de Empleo y Seguridad Social, 2017).

Con la llegada de la pandemia, el teletrabajo creció de forma exponencial llegando a alcanzar un 59% para el sector de la información y comunicación (TIC). Sin embargo, es necesario

tener en cuenta que hay sectores que no permiten el trabajo a distancia, por lo que sus niveles de teletrabajo fueron muy bajos, como por ejemplo la gastronomía, la sanidad o la construcción (Hermann, 2020).

Según un estudio publicado en 2020 por el IWD (Instituto Económico Alemán), el 43% de los empleados con acceso a internet trabaja a distancia al menos de vez en cuando; mientras que solo el 35% lo hacía antes del Covid. Además, sorprendentemente, en Alemania según los datos recogidos por la Oficina Federal de Estadística, trabajaban a distancia más hombres que mujeres, aunque la diferencia es solo del 2% (Hermann, 2020).

Un informe del año 2020 ha confirmado que, el porcentaje de ingresos de los trabajos que no permiten el teletrabajo es un 44% inferior a los ingresos asociados a trabajos que sí permiten trabajar a distancia. A nivel regional, este informe confirma una mayor implantación del teletrabajo en las zonas occidentales del país, como por ejemplo Múnich, que en la parte oriental. Esto se debe en parte a que la parte oriental cuenta con un PIB inferior y con menores rentas disponibles en los hogares. Si estas cifras se mantienen, puede suponer un problema ya que dificultaría el proceso de equiparación entre ambas partes del país haciendo más acusada las diferencias entre la mitad oriental y occidental (Irlacher & Koch, 2020).

En cuanto a las perspectivas a futuro, en Alemania parece que el teletrabajo tiende a instalarse. La aceptación del trabajo a distancia ha crecido considerablemente desde la pandemia. Esto se debe a que con el incremento del trabajo a distancia, y en vistas a su necesidad, se han corregido sus deficiencias con una mejor gestión del tiempo, mejora de aptitudes digitales y una mayor confianza por parte de los directivos (Hermann, 2020). Así, según un estudio publicado por el IFO, el 25% de los empleados en Alemania se inclina por el trabajo a distancia. Sin embargo, hay que tener en cuenta que este porcentaje no tiene en cuenta las grandes diferencias entre sectores (Agencia Informativa Latinoamericana, 2020).

Según los resultados de un estudio del IZA (Institute of Labor Economics), que recoge la opinión de los trabajadores, el 56% de los puestos de trabajo existentes en Alemania podrían hacerse de forma no presencial (Alipour et al., 2020). El estudio también revela las diferencias que hay entre Oriente y Occidente, ya que la media de puestos de trabajo que

permiten su realización a distancia en Occidente es de 59%; mientras que para Oriente apenas alcanza el 50% (Alipour et al., 2020).

Por último, según una encuesta llevada a cabo por el IAO, todo apunta a que el teletrabajo adoptado en la pandemia va a traer consigo un cambio muy significativo en la organización del trabajo. En este sentido, la media de los resultados de las empresas encuestadas estima que el 57% de sus reuniones con los trabajadores se desarrollan de manera telemática, y que el 73% de sus reuniones con clientes son también desarrolladas de forma virtual. El 42% de las empresas encuestadas afirman contemplar la posibilidad de ampliar la modalidad de trabajo a distancia; y el 56% afirma la viabilidad del trabajo a distancia sin perjudicar a los procesos de empresa (Ministerio de Empleo y Seguridad Social, 2017).

3.3. ITALIA

En Italia el teletrabajo se introdujo por primera vez en el año 1998, aunque desde su primera incorporación se han regulado nuevas variantes. En 2017 se reguló la ley del trabajo ágil, aunque no debe confundirse con el teletrabajo, ya que se trata de una forma de trabajo a distancia en la que se trabaja por ciclos y objetivos. En Italia se entiende por teletrabajo el trabajo que se realiza en cualquier lugar que sea considerado idóneo, situado fuera del lugar de trabajo, con el uso de tecnologías de la información y comunicación que es de elección voluntaria para el trabajador y el empleador (Torniamo et al., 2022).

No existe mucha información sobre el trabajo a distancia en Italia para antes de la pandemia, ya que se trata de una modalidad apenas implantada. Entre los años 2015 y 2019, las cifras de teletrabajo en Italia se mantuvieron muy constantes y bajas rondando el 3%. Esto puede ser en parte debido a la estructura empresarial italiana, ya que en este país, según datos recogidos para el año 2018, las grandes y medianas empresas empleaban solo al 15% de los trabajadores de sectores que permiten el teletrabajo (Milasi et al., 2021).

Es importante tener en cuenta las prácticas de organización y gestión comúnmente empleadas en Italia. El impacto de las prácticas organizativas se manifiesta en la diferencia existente entre las cifras de teletrabajo de distintos países dentro de los mismos sectores. Por ejemplo, entre los profesionales de las TIC, la proporción de teletrabajadores en 2018 fue del 32% en

Alemania y del 11% en Italia. Esto indica que los trabajadores de una determinada ocupación en un país pueden tener más acceso al teletrabajo que en otros países. Italia tiene una cultura laboral vinculada a la presencialidad, y sus empresas son generalmente pequeñas. Esto puede influir a la hora de implementar eficazmente el teletrabajo por falta de recursos (Milasi et al., 2021).

Durante la pandemia, el teletrabajo aumentó significativamente llegando a alcanzar el 21,6% en las medianas empresas, y el 31,4% en las grandes empresas. Los sectores con cifras más altas en cuanto a la implementación del teletrabajo fueron los servicios de la información y comunicación (TIC), que pasaron del 5% al 48,8%; y las actividades profesionales, científicas y técnicas, que pasaron del 4,1% al 36,7% (Faramondi & Costa, 2020).

A diferencia de lo esperado para Alemania, parece que en Italia el teletrabajo ha ido desapareciendo con el fin de la pandemia. Según un estudio llevado a cabo por el Departamento de Geografía y Medio Ambiente de la LSE a finales del año 2022, las predicciones sobre el fin de la oficina presencial en Italia no son tan drásticas como se preveía. Igual que ha sucedido en Alemania, se han magnificado las desigualdades entre las regiones más pobres y las pequeñas y medianas empresas; y las regiones más ricas y las grandes empresas (Millán, 2022).

Según los resultados de esta investigación, solo el 12% de los trabajadores podía trabajar desde casa en el momento más crítico de la pandemia. El estudio también pone de manifiesto que el teletrabajo es más frecuente en las grandes ciudades. Además, según la investigación, el 70% de las grandes empresas contaba por lo menos con un trabajador a distancia a finales del año 2022. Sin embargo, solo el 1% de las empresas con menos de 10 empleados habían implantado medidas de teletrabajo. Las empresas que permiten el trabajo a distancia suelen tener sus sedes en las regiones más ricas del país, por ejemplo Lacio y Lombardía, donde están situadas las ciudades de Roma y Milán. El 21% de los trabajadores de estas regiones teletrabajaban al cierre del año 2022; mientras que solo el 2% lo hacía en la zona del sur de Italia, donde están las regiones más pobres (Millán, 2022).

Dado que el trabajo a distancia requiere de dispositivos tecnológicos, la inversión en I+D es relevante a la hora de estudiar los diferentes niveles de teletrabajo en cada país.

Concretamente en Italia dicha inversión no es elevada, ya que el porcentaje de PIB destinado a I+D el año 2020 fue 1,51. Si se compara esta cifra con la de Alemania, la diferencia resulta ser bastante grande ya que, en 2020, el porcentaje del PIB destinado a I+D fue 3,13%. Esto explicaría los valores tan bajos de trabajo a distancia en Italia, ya que la inversión en I+D nunca ha sido alta, por ejemplo en el año 2010 el porcentaje del PIB destinado a I+D fue 1,22% (Instituto Vasco de Estadística, 2022).

Dado que la estructura empresarial en Italia está compuesta mayoritariamente por empresas pequeñas y sus inversiones en I+D son bajas, parece difícil que se mantengan las cifras de teletrabajo que había durante la pandemia.

3.4. PORTUGAL

A diferencia de Italia, en Portugal el teletrabajo estaba relativamente extendido antes de la pandemia. Según un estudio de la Comisión Europea, en 2009, la cifra de teletrabajo rondaba el 5%, y en 2019 había alcanzado el 15%. Esto situaba a Portugal a medio camino entre los países que tenían cifras más altas en Europa, como Países Bajos, y los que tenían cifras muy bajas como Italia (Milasi et al., 2021). A pesar de que tanto Italia como Portugal cuentan con muchas pequeñas y medianas empresas, la inversión en I+D en Portugal es más alta que en Italia. Por ejemplo en el año 2010 Portugal destinó un 1,54% de su PIB a I+D, y en el año 2020 invirtió un 1,62%. Esto explica la existencia de un porcentaje más alto de trabajo a distancia en Italia (Instituto Vasco de Estadística, 2022).

Según un estudio realizado por el gobierno español, en Portugal el porcentaje de personas que había teletrabajado antes de la pandemia llegaba al 31,3%. De este porcentaje un 6,7% lo hacían a diario y un 6,2% lo hacían varias veces a la semana. Según este mismo estudio, el trabajo a distancia en Portugal ha sido generalmente aceptado. Una vez terminada la pandemia, al 16,5% de los trabajadores les gustaría teletrabajar a diario, al 26,7% a menudo y al 23,8% varias veces al mes (Velasco, 2022).

Según los datos recogidos por el Instituto Nacional de Estadística de Portugal (Statistics), durante el período de abril a diciembre de 2020, el porcentaje de teletrabajo en Portugal alcanzó de media el 15,6%. Sin embargo, parece importante mencionar que incluso en

tiempos de pandemia, las cifras más altas de teletrabajo se concentraron en el área metropolitana de Lisboa llegando a alcanzar el 27,9%. Lisboa llegó a concentrar el 48% de la población del país que trabajaba a distancia.

De nuevo, las cifras del teletrabajo están vinculadas a las regiones más ricas del país y a determinadas ramas de actividades económicas. Por ejemplo, la proporción de teletrabajo en las Actividades de Información y Comunicación alcanzó el 66,9%; las actividades financieras el 47,6%; y las actividades de consultoría, científicas y técnicas alcanzaron el 32,5% (Statics Portugal, 2021).

En diciembre de 2021 se aprobó la regulación del nuevo régimen de trabajo a distancia para Portugal, que entró en vigor en enero de 2022. Con esta regulación se crea una nueva definición del concepto de teletrabajo, de forma que queda definido como “la prestación de trabajo en régimen de subordinación jurídica del trabajador a un empleador, en un lugar no determinado por este, utilizando tecnologías de información y comunicación” (Belzuz Abogados, 2021).

Esta nueva regulación sigue manteniendo la necesidad de acuerdo escrito que ya existía en la regulación anterior. Sin embargo, aumenta la seguridad jurídica ya que se amplían los elementos de detalle obligatorio en el contrato, como por ejemplo la periodicidad, el horario o el lugar de trabajo. También se ofrecen garantías para el empleado, ya que si el empleador propone el teletrabajo, el empleado puede negarse sin justificación. Además, el empleador ha de proporcionar los equipos y materiales necesarios para el teletrabajo (Belzuz Abogados, 2021).

El hecho de que exista un marco legal que delimita el teletrabajo en Portugal hace que las predicciones apunten a la consolidación del teletrabajo en Portugal, si bien con cifras más bajas que durante la pandemia.

A continuación se muestra una tabla resumen con las principales características de cada país.

Tabla 2. Tabla resumen

	Alemania	Italia	Portugal
Primera regulación del teletrabajo	2016	Primera regulación 1998 Nueva regulación 2017	Nuevo régimen 2021
Teletrabajo antes del COVID	En 2019 estaba en torno al 13%	En 2019 presentaba cifras muy bajas en torno al 3%	En 2019 había alcanzado el 15%
Teletrabajo durante el COVID	El 43 % de los empleados trabaja a distancia al menos de vez en cuando	Medianas empresas 21, 6 % Grandes empresas 31,4 %	Alcanzó una tasa de 27, 9 %.
Perspectivas a futuro	Consolidación del trabajo a distancia	Consolidación muy lenta	Consolidación del trabajo a distancia
Concentración del teletrabajo en sectores más propensos al trabajo a distancia	Servicios de la Información y Comunicación 60%	Servicios de Información y Comunicación 48,8 %	Servicios de Información y Comunicación 66,9%
Grado de aceptación del trabajo a distancia	Muy alto	Bajo Existe un gran vínculo a la presencialidad	Alto
Desigualdades regionales	Sí	Sí	Sí

3.5. FORMULACIÓN DE HIPÓTESIS

Hipótesis 1. El teletrabajo seguirá aumentando de cara a los años futuros en Alemania, Italia y Portugal.

Hipótesis 2. El ritmo de aumento de las cifras de teletrabajo en Alemania, Italia y Portugal va a sufrir una desaceleración tras el crecimiento tan rápido que han experimentado las cifras durante la pandemia mundial.

4. METODOLOGÍA

4.1. EXPLICACIÓN DE LA BASE DE DATOS

La base de datos utilizada en el presente trabajo se llama “Employed persons working from home as a percentage of the total employment, by sex, age and professional status” y ha sido descargada de Eurostat (Eurostat, 2023). El conjunto de datos contiene los valores históricos del teletrabajo desde 1992 hasta 2022 desglosados según el sexo, la frecuencia del teletrabajo, tramos de edad y categorías laborales. La base de datos recién descargada contenía una serie de letras y símbolos que han sido suprimidos para evitar problemas a la hora de importar y leer el conjunto de datos con R studio.

En el conjunto original de datos, las variables sexo y categoría laboral son categóricas nominales; mientras que las variables tramos de edad y la frecuencia son categóricas ordinales. Para este estudio se ha filtrado el conjunto de datos de forma que la variable dependiente es el porcentaje del total de personas empleadas que teletrabajan a menudo. Es decir, no se ha hecho distinción por sexo (TOTAL). Además, del conjunto original de datos se han seleccionado la frecuencia “usually” (USU), el tramo de edad de 26 a 65 años, la categoría laboral “personas empleadas” (EMP), y los años 2015 hasta 2022. El proceso de filtrado se ha hecho con R studio una vez importada la base de datos en formato *csv* una vez eliminados los valores perdidos.

4.2. LA VARIABLE DEPENDIENTE

Los países elegidos para el estudio son Alemania (DE), Italia (IT) y Portugal (PT); que han sido elegidos por razón del PIB per cápita de los países del año 2022. Como se ha explicado previamente, el criterio de elección de los países han sido sus distintos valores del PIB dentro de las opciones limitadas ofrecidas por el conjunto de datos. Se han elegido dos valores extremos, que corresponden a Alemania y a Portugal, y un valor medio que corresponde a Italia. A continuación, se presenta una tabla con los distintos valores del PIB per cápita en 2022 para cada uno de los países elegidos.

Tabla 3. Valores del PIB per cápita en el año 2022

PIB Per Cápita 2022	
Alemania	46.180 €
Italia	32.390 €
Portugal	23.290 €

En cuanto al resto de filtros, han sido elegidos para tratar de aportar una mayor generalidad al modelo dada la especificidad del resto de variables. Es decir, se he elegido el rango de edad de 25 a 65 años dado que suele ser considerado como la edad laboral activa; y la categoría laboral “employed persons” porque engloba el resto de las categorías.

La razón por la que se ha decidido prescindir de los datos recogidos entre los años 1992 y 2014 es porque pueden distorsionar las predicciones. Esto se debe a que remontarse tan hacia

atrás en el tiempo supondría tener en cuenta variables como el avance tecnológico anual o el acceso a los medios tecnológicos para las que no se dispone de información. Por lo tanto, el año que se toma como referencia para el modelo es el año 2015. A continuación, se muestra el código utilizado en R studio para crear el *data set* y el resultado del subconjunto de datos.

```
data_model <- datos [datos$Sex=="T" & datos$Frequency=="USU" & datos$Edad ==
"Y25-64" & datos$Working_Status=="EMP", ]
```

Ilustración 1. Visualización del dataset “data_model”

Country	2015	2016	2017	2018	2019	2020	2021	2022
DE	3.5	3.4	5.2	5.3	5.6	14.2	18	15.3
IT	3.4	3.4	3.6	3.7	3.7	12.6	8.5	5.3
PT	6.5	6.6	6.2	6.4	6.9	14.3	14.6	7.8

Dado que los porcentajes de teletrabajo estaban dispuestos en columnas distintas según los años, ha sido necesario reestructurarlos para tener únicamente tres columnas: una con los países, otra con los años, y otra con los porcentajes por año. A continuación, se muestra tanto la función que se ha utilizado, como la disposición final del conjunto de datos.

```
data_long <- gather(data_model, key = "año", value = "porcentaje_teletrabajo", -Country)
```

Ilustración 2. Visualización del dataset “data_long” previa inclusión de las variables independientes

Country	año	porcentaje_teletrabajo
DE	2015	3.5
IT	2015	3.4
PT	2015	6.5
DE	2016	3.4
IT	2016	3.4
PT	2016	6.6
DE	2017	5.2
IT	2017	3.6
PT	2017	6.2
DE	2018	5.3
IT	2018	3.7
PT	2018	6.4
DE	2019	5.6
IT	2019	3.7
PT	2019	6.9
DE	2020	14.2
IT	2020	12.6
PT	2020	14.3
DE	2021	18
IT	2021	8.5
PT	2021	14.6
DE	2022	15.3
IT	2022	5.3
PT	2022	7.8

Una vez explicada la variable dependiente, que es la que se quiere predecir, se explican a continuación las variables independientes.

4.3. LAS VARIABLES INDEPENDIENTES

La base de datos contiene la información recogida en variables categóricas, por tanto, para construir el modelo predictivo ha sido necesario reestructurar el conjunto de datos para crear las variables independientes.

En primer lugar, se ha incorporado como variable económica el PIB per cápita, cuyos valores para los años 2015-2022 han sido obtenidos de la plataforma Datosmacro (Datosmacro, 2023). El Producto Interior Bruto mide el valor monetario del conjunto de bienes y servicios finales producidos por un país durante un año. El PIB per cápita es el resultado de dividir el PIB entre el número de habitantes, y representa el valor de todos los bienes y servicios producidos por cada uno de ellos (Datosmacro, 2023). La razón por la que se ha decidido incorporar esta variable es porque, como se ha mencionado anteriormente, hay una relación positiva entre el nivel de renta y los trabajos que permiten el teletrabajo.

La incorporación al conjunto de datos de las cifras del PIB per cápita se ha hecho de la siguiente forma:

```
data_long$PIB <- c (37050, 27480, 17350, 38070, 28210, 18060, 39530, 28940, 19020,
40590, 29580, 19950, 41510, 30080, 20840, 40950, 27940, 19470, 43290, 30230, 20870,
46180, 32390, 23290)
```

En segundo lugar, se ha introducido la variable SELF, que recoge el porcentaje de teletrabajo de personas autónomas. Los datos correspondientes a esta variable han sido extraídos de la información proporcionada por el *data set* original. Para esto se ha seleccionado en R studio seleccionando el total de personas (T) que teletrabajan a menudo (frecuencia “USU”), la categoría laboral “SELF”, y el grupo de edad de 25 a 65 años. Este mismo proceso se ha seguido para incorporar al modelo la variable “SAL”, que corresponde al grupo de personas empleadas por cuenta ajena (employees). La diferencia con la variable “employed persons” es que esta engloba al grupo de población activa, independientemente de su categoría profesional.

En tercer lugar, se han incorporado las variables MUJER y HOMBRE para estudiar si hay diferencias por sexo en relación con el teletrabajo. Para obtener los datos de teletrabajo de los hombres y de las mujeres por separado se ha acudido al conjunto de datos original. Utilizando R studio se han filtrado los datos seleccionando por un lado los hombres (M) que teletrabajan a menudo, dentro de la categoría laboral “employed persons”, para el tramo de edad de 25-65; y por otro las mujeres (F) que teletrabajan a menudo, dentro de la categoría laboral “employed persons”, para el tramo de edad de 25-65. A continuación se muestra el *data set* definitivo al que se ha llamado “data_long” que se utilizará para en entrenamiento del modelo.

Ilustración 3. Visualización del dataset “data_long” con las variables independientes

Country	año	porcentaje_teletrabajo	PIB	teletrabajo_SELF	teletrabajo_SAL	teletrabajo_MUJER	teletrabajo_HOMBRE
DE	2015	3.5	37050	19.0	1.6	3.7	3.3
IT	2015	3.4	27480	11.6	1.0	3.1	3.7
PT	2015	6.5	17350	13.6	5.2	7.0	6.0
DE	2016	3.4	38070	18.8	1.6	3.6	3.3
IT	2016	3.4	28210	11.6	1.0	3.2	3.6
PT	2016	6.6	18060	14.7	5.2	7.0	6.2
DE	2017	5.2	39530	24.5	3.0	5.3	5.0
IT	2017	3.6	28940	12.0	1.2	3.4	3.8
PT	2017	6.2	19020	13.9	4.9	6.2	6.1
DE	2018	5.3	40590	24.7	3.2	5.5	5.2
IT	2018	3.7	29580	12.6	1.2	3.4	3.9
PT	2018	6.4	19950	15.4	4.9	6.9	5.9
DE	2019	5.6	41510	25.8	3.4	5.7	5.4
IT	2019	3.7	30080	13.0	1.2	3.4	3.9
PT	2019	6.9	20840	17.4	5.1	7.1	6.6
DE	2020	14.2	40950	32.3	12.4	13.7	14.6
IT	2020	12.6	27940	17.0	11.5	14.7	11.1
PT	2020	14.3	19470	17.4	13.8	15.0	13.7
DE	2021	18.0	43290	31.7	16.7	17.6	18.4
IT	2021	8.5	30230	9.0	8.4	10.2	7.2
PT	2021	14.6	20870	15.4	14.5	15.8	13.3
DE	2022	15.3	46180	30.0	13.9	15.1	15.5
IT	2022	5.3	32390	8.4	4.5	6.0	4.8
PT	2022	7.8	23290	14.0	6.8	7.7	7.8

4.4. NORMALIZACIÓN DE LAS VARIABLES INDEPENDIENTES

Si bien es cierto que todas las variables recogen porcentajes de teletrabajo con la excepción del PIB per cápita, es necesario normalizar las variables. De lo contrario, las diferentes escalas y rangos de valores de las variables pueden incidir en la influencia relativa de cada una en el modelo de predicción. Al normalizar las variables se eliminan también los sesgos que pueden estar causados por las diferentes unidades de medida: porcentaje y euros.

Además, la normalización aporta una mayor estabilidad al modelo durante el proceso de entrenamiento y prueba. Esto se debe a que, al reducir las diferencias en las escalas de las variables, se evita que el modelo se vea afectado por valores demasiado grandes o pequeños (Maurandi López & González Vidal, 2022).

Para normalizar las variables se ha utilizado la función “scale” en R studio. Es importante tener en cuenta que esta función solo es válida para variables numéricas por lo que si no lo fueran habría que transformarlas. Las variables normalizadas han sido recogidas en un nuevo conjunto de datos llamado “datos_long_norm”. Con este *data set* se entrenan los modelos predictivos, cuyos resultados serán analizados en el apartado titulado “Resultados de la Investigación”. A modo de adelanto, cabe mencionar que se han creado tres modelos distintos para comparar los resultados. Por un lado, se ha desarrollado un modelo de regresión lineal, que se ha validado utilizando el método Cross Validation. Por otro lado, se han utilizado las técnicas de regularización Lasso y Ridge, que penalizan los coeficientes de las variables predictoras o independientes. De esta forma, se puede reducir la magnitud de las variables y seleccionar las más relevantes o disminuir el impacto de variables redundantes en el modelo.

4.5. MODELOS

Los datos normalizados están recogidos en un *data set* llamado “datos_long_norm”, y se han utilizado para entrenar los tres modelos, siendo el primero de regresión lineal, el segundo de regresión Lasso y el tercero de regresión Ridge. Antes de entrenar y evaluar de forma detallada cada uno de los modelos, lo primero que se ha hecho, es tratar de decidir cuál es el mejor modelo utilizando una función “resamples” en R studio. Antes de llamar a la función, se han definido los predictores o variables independientes del modelo y la variable objetivo o dependiente del modelo, que en este caso es “porcentaje_teletrabajo”. El siguiente paso ha sido crear un conjunto de datos al que se ha llamado “train_control” para entrenar y evaluar los modelos.

Se ha utilizado el método de validación cruzada fijando el número de pliegues en 5, es decir que el conjunto de datos se divide en 5 partes. De esta forma se entrena y evalúa el modelo en múltiples iteraciones utilizando diferentes combinaciones de partes como conjuntos de entrenamiento y de prueba. Hecho esto, se entrenan el modelo 1 de regresión lineal, el modelo

2 de regresión Ridge y el modelo 3 de regresión Lasso con la función “train”. Por último, la función “resamples” calcula el rendimiento utilizando medidas como el RMSE (error cuadrático medio) y el R-squared (coeficiente de determinación) para cada modelo utilizando los resultados obtenidos en la validación cruzada. A continuación, se muestran los resultados de la comparación del rendimiento de los modelos hecha por R studio.

Ilustración 4. Resultados de la comparativa del rendimiento de los tres modelos hecha por R studio

```
Call:
summary.resamples(object = compare_models)

Models: modelo1, modelo2, modelo3
Number of resamples: 5

MAE
      Min.   1st Qu.   Median     Mean   3rd Qu.   Max. NA's
modelo1 0.02531656 0.04188274 0.04656459 0.05027575 0.06428285 0.07333200 0
modelo2 0.00901312 0.03266833 0.04633755 0.04316818 0.06319425 0.06462765 0
modelo3 0.21746936 0.27840618 0.30141805 0.29794509 0.32375228 0.36867959 0

RMSE
      Min.   1st Qu.   Median     Mean   3rd Qu.   Max. NA's
modelo1 0.03389015 0.04496340 0.04995501 0.06564103 0.09455170 0.10484487 0
modelo2 0.01049717 0.04228977 0.04939289 0.05510299 0.08239504 0.09094008 0
modelo3 0.27576809 0.30821518 0.36096219 0.35747290 0.37179604 0.47062301 0

Rsquared
      Min.   1st Qu.   Median     Mean   3rd Qu.   Max. NA's
modelo1 0.9998973 0.9999275 0.9999277 0.9999383 0.9999617 0.9999772 0
modelo2 0.9998725 0.9998731 0.9999508 0.9999369 0.9999921 0.9999961 0
modelo3 0.9988999 0.9998433 0.9998863 0.9997168 0.9999579 0.9999967 0
```

De este resumen se puede concluir que el rendimiento en la predicción del porcentaje de teletrabajo de los tres modelos es bueno. Por un lado, el coeficiente de determinación R-squared es alto, lo que significa que los modelos explican la mayoría de la variabilidad de la variable objetivo. Por otro lado, los valores de los errores MAE y RMSE, cuya explicación y análisis detallado se proporciona más adelante, son bajos, lo que indica que los modelos tienen una buena precisión.

El modelo de regresión Ridge presenta los valores más bajos de MAE y RMSE, lo que indica que se ajusta mejor a los datos y que las predicciones varían menos. Por otro lado, el modelo de regresión Lasso tiene el valor más bajo de R-squared, lo que significa que, en relación con los otros dos modelos, este explica menos de la variabilidad de la variable objetivo. El modelo de regresión lineal presenta unos valores bastante similares a los de los otros modelos. Dado que el modelo 2 tiene valores más pequeños en cuanto a MAE y RMSE, podría considerarse el modelo con mejor rendimiento.

4.6. ESTIMACIÓN DE VARIABLES

Para poder predecir el porcentaje de teletrabajo (variable dependiente) en los años 2023 y 2024, es necesario estimar primero los valores de las variables independientes para dichos años. Dada la creciente incertidumbre y la especificidad de las variables, no ha sido posible encontrar estimaciones oficiales, por lo que se ha recurrido a dos técnicas diferentes de estimación.

A la hora de estimar estas variables, es importante tener en cuenta que pueden verse afectadas por las políticas laborales, avances tecnológicos, o por eventos económicos y sociales. Sin embargo, dado que no se dispone de esta información, estos factores se han considerado exógenos, y no han sido tomados en consideración. Por lo tanto, los datos que se van a utilizar para estimar estas variables son los valores históricos recogidos en el conjunto de datos descargado de Eurostat. Es decir, se va a intentar encontrar tendencias históricas de las variables objeto de estimación.

En el presente estudio los datos se remontan a 2015 y se extienden hasta 2022, es decir que para la estimación se ha tenido en cuenta la evolución de cada una de las variables independientes en los últimos ocho años. El motivo por el que se han utilizado dos métodos diferentes para la estimación de las variables independientes es validar los resultados y verificar si son consistentes entre sí. Además, la combinación de ambos métodos ofrece una visión más completa del modelo.

4.6.1. La media del diferencial del incremento

El primer método de estimación de variables utilizado es la media del diferencial del incremento entre el último año y el primero. Este método proporciona una medida que, a pesar de ser simple, puede ser útil para evaluar la tendencia promedio de los datos a lo largo del tiempo. Este sistema se ha considerado adecuado porque los datos presentan fluctuaciones significativas que se deben al gran despunte del teletrabajo en el año 2019 provocado por la pandemia mundial del Covid-19. Al tomar la media se consigue una estimación general de la variación de los datos en el periodo de tiempo que se extiende desde 2015 hasta 2022.

Este método tiene ventajas e inconvenientes. En cuanto a ventajas, en primer lugar, se trata de un enfoque simple que no requiere conocimientos técnicos avanzados, siendo fácil de calcular y de entender. En segundo lugar, al calcular la media de los incrementos se puede medir la tendencia promedio a lo largo del ínterin de tiempo seleccionado. En tercer lugar, utilizando este método se pueden suavizar los cambios extremos en los datos de un año a otro proporcionando una perspectiva más estable. Este tercer motivo ha sido el principal de su elección, ya que en los años de pandemia hay grandes fluctuaciones en los datos.

Sin embargo, ha sido necesario complementarlo con el segundo método de estimación ya que implica la pérdida de información detallada de los valores específicos de cada año. Por otro lado, este método puede sesgar la información si hay años con cambios atípicos o fluctuaciones inusuales no reflejando adecuadamente la verdadera tendencia de los datos. Además, este método tiene limitaciones en patrones no lineales ya que asume la tendencia lineal del promedio, por lo que no tiene en cuenta el crecimiento exponencial o los ciclos estacionales. Otro factor relevante es que este enfoque asume que el crecimiento promedio observado en los años anteriores se mantendrá constante en los años futuros. Por último, para estimar utilizando este enfoque es importante elegir correctamente el periodo de tiempo específico, ya que la estimación puede verse afectada por los acontecimientos ocurridos durante esos años. Es por esto por lo que se ha elegido como punto de partida el año 2015, ya que la adopción generalizada de ordenadores ocurre en torno a este año.

La variable PIB per cápita se ha estimado utilizando las previsiones del PIB en la Eurozona para los años 2023 y 2024 publicadas en un estudio de Bankinter (Departamento de análisis

Bankinter, 2023). El incremento estimado para la Alemania e Italia en el año 2023 es de un 0,6% y para el año 2024 es de un 1,3%. Sin embargo, el incremento del PIB estimado para Portugal en 2023 es de 0,9% en 2023 y de 2% en 2024. Por lo tanto, el PIB per cápita estimado para Alemania en el año 2023 será el PIB per cápita del año 2022 multiplicado por 1,6 que, es el incremento. Esta cifra asciende a 46.464 € en 2023 y a 47.107 € en 2024. En cuanto a Italia, las estimaciones para los años 2023 y 2024 ascenderían a 32.592 y 33.004 respectivamente. Por último, en Portugal las estimaciones del PIB per cápita serían 23.511 € para 2023 y 24.001€ para 2024.

Tabla 4. Estimaciones del PIB per cápita para los años 2023 y 2024

PIB per cápita	2023	2024
Alemania	46.464 €	47.107 €
Italia	32.592 €	33.004 €
Portugal	23.511 €	24.001 €

El resto de las variables se han estimado utilizando una fórmula de R studio para calcular el diferencial del incremento. Para mayor claridad, se explica con un ejemplo la estimación de la variable SELF. El primer paso, consiste en crear un subconjunto de datos en el que se recogen los valores que corresponden al teletrabajo del total de autónomos. Este proceso se repite para cada país, de manera que habrá tres subconjuntos diferentes que recogerán los datos correspondientes a las variables para cada país.

```
SELF_alemania <- subset(data_long, Country == "DE")$teletrabajo_SELF
```

El segundo paso consiste en utilizar una función de R studio para calcular el diferencial. Supone restarle al valor del último año (2022) el valor del primer año de referencia (2015) y hacer la media dividiendo entre el total de años, que es ocho. Este valor es el incremento medio, que será diferente para cada país. Siguiendo con el ejemplo anterior, se hablaría del crecimiento medio diferencial para Alemania, y habría que multiplicarlo por el porcentaje de teletrabajo de autónomos en Alemania en el año 2022. Una vez calculado este valor, habría que multiplicar de nuevo el incremento por el valor estimado para 2023, y así se obtendría el valor estimado para 2024.

```
crecimiento_promedio_SELF_DE <- mean(diff(SELF_alemania))
```

```
SELF_alemania_2023 <- tail(SELF_alemania, 1) + crecimiento_promedio_SELF_DE
```

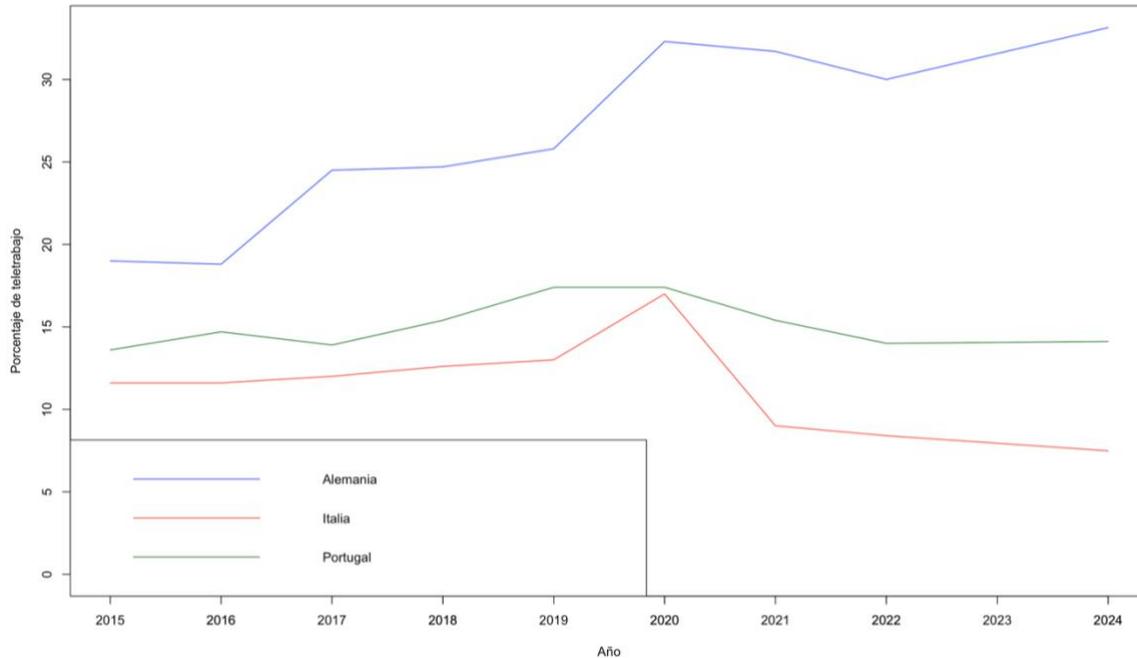
```
SELF_alemania_2024 <- SELF_alemania_2023 + crecimiento_promedio_SELF_DE
```

Tabla 5. Estimaciones para la variables SELF según el método de la media del diferencial del incremento

Estimaciones variable SELF	2023	2024
Alemania	31,57143 %	33,14286 %
Italia	7,942857 %	7,485714 %
Portugal	14,05714 %	14,11429 %

Es importante recalcar que las estimaciones se hacen con los datos sin normalizar ya que dentro de cada variable la medida es la misma. De esta forma se obtienen los porcentajes estimados para los años 2023 y 2024 para cada país y para cada variable. Por ejemplo, en virtud de estas estimaciones se puede observar que en Alemania y Portugal se espera un ligero incremento en el porcentaje total de autónomos que teletrabajan, mientras que en Italia se espera que estos valores decrezcan ligeramente con respecto al último año conocido.

Gráfico 1. Evolución del teletrabajo total de autónomos por país según las estimaciones del método de la media del diferencial del incremento. Años 2015-2024



El resultado final obtenido para la estimación de cada variable ha sido recogido en un nuevo conjunto de datos llamado “datos_nuevos”, que se presenta a continuación. Es importante tener en cuenta que para utilizar este conjunto de datos a la hora de realizar las predicciones se requiere su normalización.

Ilustración 5. Visualización del conjunto de datos “datos_nuevos” que recoge las estimaciones según el método de la media del diferencial del incremento

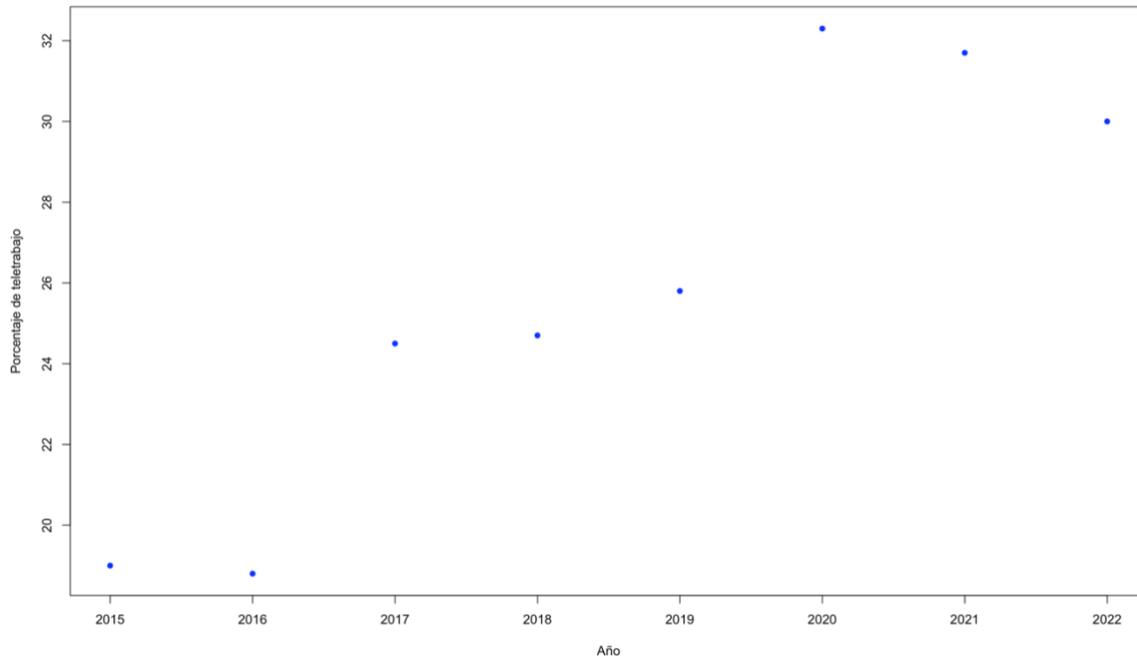
pais	año	PIB	teletrabajo_SELF	teletrabajo_SAL	teletrabajo_MUJER	teletrabajo_HOMBRE
DE	2023	46464	31.571430	15.657140	16.728570	17.242860
IT	2023	32592	7.942857	5.000000	6.414286	4.957143
PT	2023	23511	14.057140	7.028571	7.800000	8.057143
DE	2024	47107	33.142860	17.414290	18.357140	18.985710
IT	2024	33004	7.485714	5.500000	6.828571	5.114286
PT	2024	24001	14.114290	7.257143	7.900000	8.314286

4.6.2. Estimación con regresión lineal de las variables independientes

El segundo método utilizado supone entrenar un modelo predictivo con los datos históricos (2015-2022) para cada una de las variables independientes y para cada país. La estimación del PIB per cápita se ha hecho de la misma forma que en el supuesto anterior, es decir siguiendo el modelo del incremento, ya que, al ser estos datos oficiales, se considera que son más precisos. Por lo tanto, las variables estimadas con este enfoque son: el porcentaje de teletrabajo esperado para los autónomos (SELF), el porcentaje de teletrabajo esperado para los trabajadores dependientes (SAL), el porcentaje de teletrabajo esperado para las mujeres en situación laboral de ocupación (MUJER) y el porcentaje de teletrabajo esperado para los hombres en situación laboral de ocupación (HOMBRE). Este enfoque supone entrenar doce modelos distintos.

Al utilizar datos históricos, los modelos de regresión lineal pueden reconocer patrones y tendencias a largo plazo generando una ecuación que describa la relación entre las variables de interés. Este enfoque asume que hay una relación lineal entre los porcentajes de teletrabajo de 2015 a 2022. Para evaluar si esto es o no cierto se puede hacer un gráfico de dispersión o estudiar en coeficiente de correlación de Pearson. Si los puntos en el gráfico de dispersión quedan distribuidos alrededor de una línea recta, significa que hay relación lineal. Sin embargo, es importante tener en cuenta que hay patrones no lineales que pueden tener apariencia de relación lineal sin tenerla. Por eso, se ha contrastado el resultado del gráfico con el coeficiente de correlación de Pearson calculado con la función “cor” de R studio. Este coeficiente, que oscila entre -1 y 1, indica la fuerza y la dirección de la relación lineal de las variables. Por ejemplo, en el caso de Alemania, se puede concluir que efectivamente hay relación lineal entre los porcentajes de teletrabajo de los distintos años ya que el coeficiente es de 0,922311. Cuanto más cercano a 0 sea este coeficiente, menor es la correlación; y cuanto más próximo sea el coeficiente a 1 o -1, más fuerte será la correlación (Maurandi López & González Vidal, 2022).

Gráfico 2. Gráfico de dispersión que refleja la evolución del teletrabajo de los autónomos en Alemania



```
cor(years, SELF_alemania)
```

Ilustración 6. Coeficiente de correlación Pearson

```
> cor(years, SELF_alemania)  
[1] 0.922311
```

Este modelo tiene limitaciones ya que, dado que los datos disponibles solo se extienden hasta 2022, puede ser que los datos históricos no reflejen condiciones actuales o futuras que pueden afectar a las estimaciones. Además, existe riesgo de extrapolación ya que, se asume que la relación entre las variables se mantiene constante fuera del rango de los datos conocidos. Esto puede tener cierta similitud con la suposición previamente explicada en el apartado anterior, donde el método de estimación asume que el incremento se mantiene constante en los años futuros. Sin embargo, es importante tener en cuenta que estas suposiciones pueden conducir a estimaciones menos confiables. Además, los modelos de regresión lineal tienen una alta sensibilidad a los valores atípicos, luego las observaciones inusuales correspondientes a la pandemia pueden influir y afectar a la precisión del modelo.

A continuación, se detalla un ejemplo de cómo se han realizado las estimaciones utilizando R studio para los años 2023 y 2024 de la variable SELF para Alemania. En primer lugar, se filtran los datos para crear un subconjunto que recoja toda la información relativa a un país, por ejemplo, para Alemania este subconjunto se ha denominado “Datos_DE”. Utilizando este subconjunto de datos se selecciona la columna correspondiente a la variable que se quiere estimar, en este caso “teletrabajo_SELF”, que corresponde al total de teletrabajo de personas autónomas, y se recoge en una lista. A continuación, se crea una columna en el subconjunto “Datos_DE” para incorporar los años. Por último, se ajusta el modelo de regresión lineal a los datos utilizando la función “lm” de R studio. Para terminar de estimar la variable SELF habría que repetir el proceso con los países Italia y Portugal.

```
datos_DE <- subset(data_long, Country == "DE")
```

```
SELF_DE <- datos_DE$teletrabajo_SELF
```

```
datos_DE$año_num <- as.numeric(datos_DE$año)
```

```
modelo <- lm(SELF_DE ~ año_num, data = datos_DE)
```

Repitiendo este proceso para las cuatro variables restantes se obtendrían las estimaciones para 2023 y 2024 correspondientes a cada país. Con el resultado de las estimaciones obtenidas se ha creado un nuevo conjunto de datos llamado “datos_nuevos”. Se llama igual que el que se ha creado utilizando el otro método de estimación ya que el otro no se utiliza.

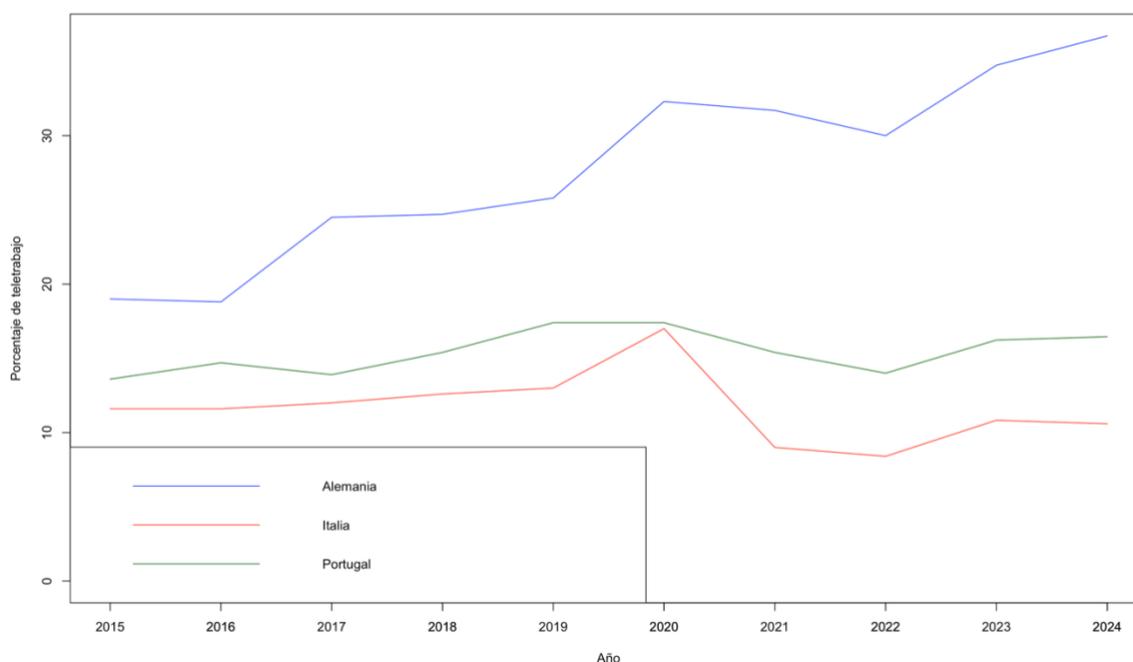
Ilustración 7. Visualización del conjunto de datos “datos_nuevos” que recoge las estimaciones realizadas por los modelos de regresión lineal

pais	año	PIB	teletrabajo_SELF	teletrabajo_SAL	teletrabajo_MUJER	teletrabajo_HOMBRE
DE	2023	46464	34.74286	17.15357	18.16071	19.010710
IT	2023	32592	10.82857	8.70000	10.70357	7.800000
PT	2023	23511	16.23214	12.08214	13.13214	12.035710
DE	2024	47107	36.71905	19.41548	20.24643	21.271430
IT	2024	33004	10.59048	9.80000	11.76548	8.366667
PT	2024	24001	16.45595	13.08929	14.03095	12.888100

Por último, a continuación se presenta un gráfico similar al del apartado anterior en el que se muestra la evolución del porcentaje de autónomos que teletrabajan con las estimaciones para

los años 2023 y 2024 según el modelo de regresión lineal. En este gráfico se puede ver que las estimaciones para todos los países están por encima del valor que tuvieron en el año 2022. Es decir, el modelo ha extraído de los datos históricos proporcionados que la tendencia del teletrabajo en la categoría laboral de los autónomos es a la alta.

Gráfico 3. Evolución del teletrabajo de los autónomos según el método de estimación con modelos de regresión lineal. Años 2015-2024



4.6.3. Comparación de los resultados según el método de estimación

A continuación se presentan una serie de gráficas que plasman la comparativa de las estimaciones para cada una de las 4 variables en 2023. En el eje X están los tres países objeto del presente estudio, y en el eje Y aparece el porcentaje de teletrabajo que varía según la variable representada. El orden de disposición de los gráficos está configurado de forma que se muestran primero las categorías laborales: autónomos y dependientes; y después el porcentaje de teletrabajo de mujeres y de hombres.

Gráfico 4. Comparación de las estimaciones de la variables “teletrabajo_SELF” según cada método de estimación. Año 2023



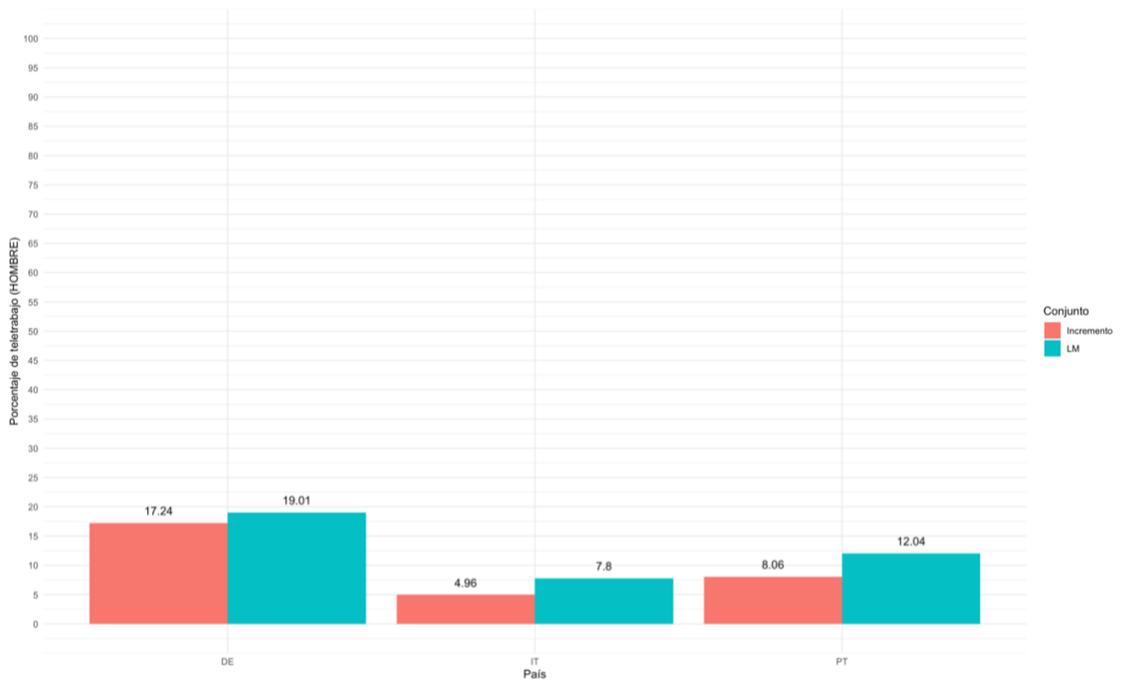
Gráfico 5. Comparación de las estimaciones de la variables “teletrabajo_SAL” según cada método de estimación. Año 2023



Gráfico 6. Comparación de las estimaciones de la variables “teletrabajo_MUJER” según cada método de estimación. Año 2023



Gráfico 7. Comparación de las estimaciones de la variables “teletrabajo_HOMBRE” según cada método de estimación. Año 2023

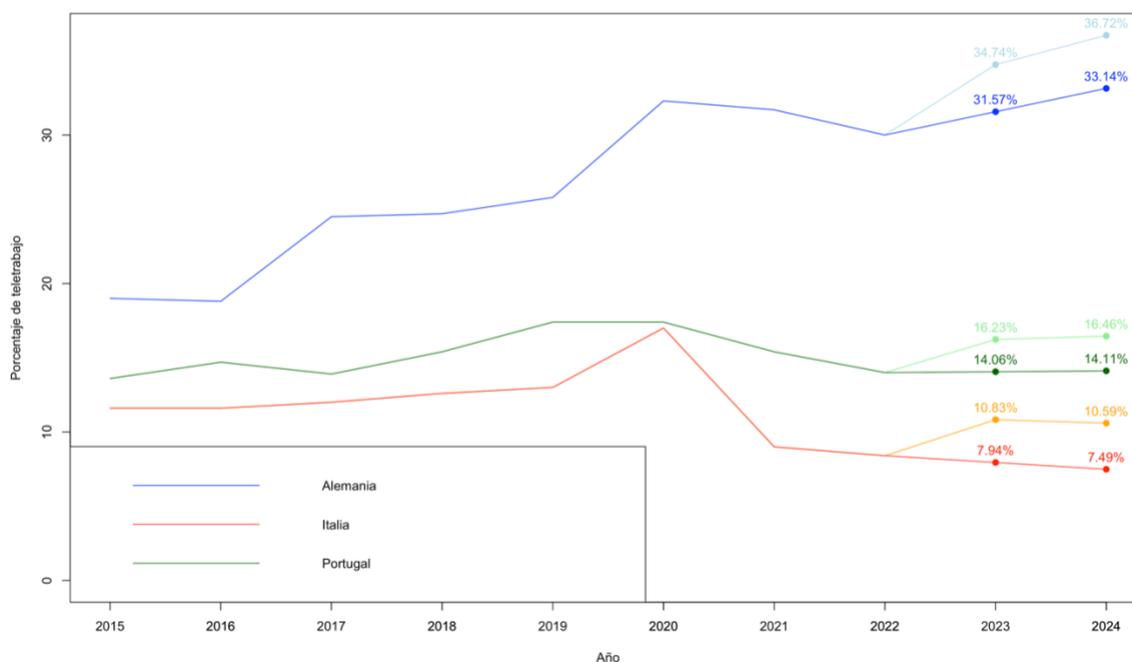


Por lo general, observando ambos conjuntos de datos se puede afirmar que el modelo de regresión lineal ha realizado estimaciones más altas para todas las variables, lo que puede indicar una tendencia a la alta del teletrabajo. Para justificar la razón por la que las estimaciones obtenidas con los modelos de regresión lineal son más altas hay que atender al funcionamiento de ambos métodos, que ya ha sido explicado en los apartados anteriores.

La estimación con la media del incremento asume que el crecimiento es constante durante los años 2023 y 2024. Sin embargo, los modelos de regresión lineal son capaces de captar patrones complejos ajustando una línea o una curva a los puntos existentes. Es decir, mientras que el primer método solo tiene en cuenta los valores del año 2015 y 2022; el segundo tiene en cuenta el valor de teletrabajo asociado a cada año, lo que le permite predecir cambios más grandes o más pequeños según los datos históricos. Además, este enfoque capta los patrones de los datos históricos, lo que es obviado por el método del incremento.

Esto último puede apreciarse de forma visual en los gráficos presentados en los apartados anteriores en los que se muestra la evolución del porcentaje de teletrabajo de las personas autónomas (SELF) incluyendo las predicciones desde ambos enfoques. A continuación se muestra un gráfico en el que se combinan las predicciones hechas con ambos enfoques, además de la evolución de los datos existentes desde 2015 hasta 2022. Las líneas que están pintadas en un color más claro representan las estimaciones según los modelos de regresión lineal; mientras que las que están pintadas de colores más oscuros son las que se han calculado utilizando el método de la media del incremento.

Gráfico 8. Evolución del teletrabajo del total de autónomos por país según cada método de estimación. Años 2015-2024



En primer lugar, en este gráfico se puede ver, que según el criterio del incremento, el crecimiento o decrecimiento de la variable es constante para los años futuros. Si se observan las líneas de color más oscuro entre los años 2022 y 2024, se ve que no hay picos y que las líneas son rectas y constantes. Es decir, este criterio asume que la variable siempre va a crecer o a decrecer. Además, en este gráfico puede observarse también la diferencia en la velocidad a la que crecen los valores, ya que los valores estimados por el modelo de regresión lineal están siempre por encima.

Además, concretamente en este ejemplo se da un supuesto muy particular, ya que los valores estimados para Italia según el criterio de la media del incremento son más bajos que el valor recogido para el año 2022. Es decir, según este criterio el teletrabajo en Italia tiende a decrecer de manera constante, sin tener en cuenta que ha habido momentos de aumento del teletrabajo incluso antes del Covid. Es decir, el incremento sale negativo porque el dato que corresponde al teletrabajo de autónomos en Italia el año 2022 es inferior al del año que se toma como referencia para comenzar el estudio (2015). Por motivos como estos se ha

decidido escoger el criterio que utiliza modelos de regresión lineal para estimar las variables independientes.

5. RESULTADOS DE LA INVESTIGACIÓN

5.1. EVALUACION DE LOS MODELOS

5.1.1. Modelo de Regresión Lineal

Con los datos normalizados recogidos en el *data set* llamado “datos_long_norm”, se ha entrenado el modelo 1, que es el de regresión lineal en el que se han incluido todas las variables independientes (*full model*). Para entrenar modelos de regresión lineal existen dos métodos que se distinguen por su implementación. Uno de ellos utiliza la función de R studio llamada “*train*” para entrenar el modelo de regresión lineal. Los argumentos de la función son los siguientes: las variables predictoras (independientes) que se utilizan para predecir la variable objetivo a las que se han llamado “predictors”, la variable objetivo a la que se ha llamado “target”, el método utilizado para predecir (lm o regresión lineal), y el “*train control*” que define el control del entrenamiento (por ejemplo, *Cross Validation* con 5 pliegues). La razón por la que no se ha utilizado esta función radica en que el otro método permite obtener medidas más precisas del rendimiento del modelo en datos que no ha visto previamente. Esto se debe a que, como se explica a continuación, se realiza una partición de los datos de forma que una parte se destina a entrenar el modelo y otra parte a probarlo.

El método aplicado en el presente trabajo para entrenar el modelo consiste en dividir el *data set* “datos_long_norm” en dos conjuntos de datos distintos: el conjunto de entrenamiento al que se ha llamado “datos_entrenamiento1” y el conjunto de prueba al que se ha llamado “datos_prueba1”. Esta división se ha hecho utilizando las funciones “*set.seed*”, que genera números aleatorios; y “*CreateDataPartition*”, que genera una partición aleatoria de los índices de los datos. El conjunto de datos de entrenamiento contiene el 70% de los datos de la muestra (que en este caso es “datos_long_norm”), y el 30% restante ha sido asignado al conjunto de datos de prueba. El siguiente paso es entrenar el modelo con los datos de entrenamiento.

```
modelo1 <- lm(porcentaje_teletrabajo ~ PIB_norm + teletrabajo_SELF_norm +
teletrabajo_SAL_norm + teletrabajo_MUJER_norm + teletrabajo_HOMBRE_norm, data =
datos_entrenamiento1)
```

A continuación, se van a interpretar las métricas del resumen del modelo 1 recogidas en la siguiente imagen.

Ilustración 8. Resumen modelo 1

Call:

```
lm(formula = porcentaje_teletrabajo ~ PIB_norm + teletrabajo_SELF_norm +
  teletrabajo_SAL_norm + teletrabajo_MUJER_norm + teletrabajo_HOMBRE_norm,
  data = datos_entrenamiento1)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.06533	-0.02315	-0.01436	0.02700	0.07787

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	7.67849	0.01125	682.767	< 2e-16 ***
PIB_norm	-0.03584	0.01618	-2.215	0.0452 *
teletrabajo_SELF_norm	0.02100	0.02351	0.894	0.3878
teletrabajo_SAL_norm	-0.09227	0.20323	-0.454	0.6573
teletrabajo_MUJER_norm	2.31921	0.19131	12.123	1.85e-08 ***
teletrabajo_HOMBRE_norm	2.35280	0.07564	31.107	1.36e-13 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.04292 on 13 degrees of freedom

Multiple R-squared: 0.9999, Adjusted R-squared: 0.9999

F-statistic: 4.439e+04 on 5 and 13 DF, p-value: < 2.2e-16

En primer lugar, los residuos son la diferencia entre los valores observados y los valores predichos por el modelo. En este caso, el residuo más negativo es de -0,06 y el más positivo es de 0,07; por lo tanto, los residuos oscilan entre -0,6 y 0,7. Se observa por tanto cierta variabilidad en los residuos, lo que puede indicar que el modelo no se ajusta perfectamente a los datos o que hay variables importantes que no se han incluido en el modelo (Maurandi López & González Vidal, 2022).

En segundo lugar, se van a analizar los coeficientes estimados del modelo, que aportan información sobre la dirección, que puede ser positiva o negativa; y magnitud de la relación entre las variables independientes y la dependiente. El hecho de que el “intercept” tenga un coeficiente estimado de 7,66 quiere decir que cuando todas las variables predictoras son iguales a cero el valor estimado es 7,66 (Maurandi López & González Vidal, 2022). En cambio, para el resto de las variables el coeficiente estimado indica cómo cambia el porcentaje de teletrabajo cuando la variable aumenta en una unidad. Por ejemplo, para el PIB, si el resto de las variables se mantienen constantes, un incremento de una unidad en la variable “PIB_norm” se asocia con una disminución de aproximadamente 0,048 en el porcentaje de teletrabajo (Cortés & Rubalcava, 1983).

En tercer lugar, el *t value* mide la significancia estadística de los coeficientes, luego un *t* valor más grande implica una mayor significancia estadística. Por otro lado, el *p* valor es la probabilidad asociada al *t* valor. Un *p* valor pequeño indica que hay evidencia para rechazar la hipótesis nula, que generalmente asume que el coeficiente es cero. Se entiende que cuando el *p* valor asociado a un *t* valor está por debajo del 0,05 la variable es estadísticamente significativa (Maurandi López & González Vidal, 2022). Por lo tanto, en este supuesto podría considerarse el PIB como estadísticamente significativa a un nivel de significancia de 0,05.

Por último, en cuanto a los errores, el RSE, o error estándar residual, indica la medida de dispersión promedio de los residuos, por lo que un valor más bajo indica un mejor ajuste a los datos (en este modelo es de 0,043) (Maurandi López & González Vidal, 2022). El *Multiple R-squared* es 0,9999, lo que indica que el modelo explica aproximadamente el 99,99% de la variabilidad en la variable dependiente (Rodríguez Sánchez et al., 2018). El *F-statistic* valora la existencia de una relación significativa entre las variables predictoras y la variable dependiente. Un valor *F-statistic* alto asociado a un *p* valor bajo indica que hay relación significativa entre las variables predictoras y la independiente, como es este caso (Maurandi López & González Vidal, 2022).

Para poder evaluar de forma más precisa el rendimiento del modelo conviene estudiar también las métricas correspondientes a los siguientes errores:

Ilustración 9. Métricas Errores modelo 1

```
"Metricas modelo 1"  
"RMSE1"  
0.1247343  
"MAE1"  
0.08987597  
"MAPE1"  
1.055863
```

En primer lugar, el Error Cuadrático Medio (RMSE) mide la diferencia entre las predicciones del modelo y los valores reales en el conjunto de prueba. Un menor valor sugiere una mejor adaptación del modelo a los datos. En este caso el error es de 0,12, lo que indica que en promedio las predicciones difieren en 0,12 unidades de los valores reales (ORACLE, s.f.).

En segundo lugar, el Error Absoluto Medio (MAE) mide la magnitud media de los errores de predicción del modelo. Cuanto menor sea el valor, mejor es la precisión del modelo; en este caso, el MAE indica que como media, las predicciones del modelo tienen una diferencia absoluta de 0,089 unidades en relación con los valores reales. Por último, el Error Porcentual Absoluto Medio (MAPE) mide el porcentaje promedio de error en las predicciones del modelo en comparación con los valores reales. Para este modelo el MAPE tiene un valor de 1,05%, lo que indica que las predicciones tienen un error del 1,05% en relación con los valores reales (ORACLE, s.f.).

Por lo general, se considera que, si los valores de RMSE y MAE son bajos, el modelo es bueno ya que estos valores sugieren que las predicciones del modelo se ajustan a los valores reales. En este supuesto, dado que las métricas muestran valores cercanos a cero se puede afirmar que el modelo 1 presenta un buen rendimiento; aunque sería necesario comparar estos valores con los de otros modelos. Para esto se van a entrenar dos modelos utilizando regresión Ridge y Lasso.

A continuación, se estudia si hay *overfitting* comparando los errores del *training test* y del *test set*, para lo que se reutiliza la partición hecha anteriormente para entrenar el modelo. Se realizan las predicciones para el conjunto de datos destinado al entrenamiento y para el conjunto de datos destinado a poner a prueba el modelo. El RMSE en el conjunto de entrenamiento es de 0,03; mientras que para el conjunto de prueba es de 0,12. El error

asociado al conjunto de prueba es ligeramente más alto al del conjunto de entrenamiento, lo que puede indicar que en el modelo 1 haya *overfitting*. Esto implicaría un sobre ajuste a los datos de entrenamiento, lo que conllevaría a que el modelo no generalice de manera adecuada para los datos nuevos. Por esta razón se van a utilizar técnicas de regresión Ridge y Lasso que penalizan la complejidad del modelo (Amat Rodrigo, 2020).

Ilustración 10. Visualización de los resultados del estudio del overfitting del modelo 1

```
> rmse_entrenamiento1 # Error cuadrático medio en el conjunto de entrenamiento para el Modelo 1
[1] 0.03550599
> rmse_prueba1 # Error cuadrático medio en el conjunto de prueba para el Modelo 1
[1] 0.1247343
```

5.1.2. Modelo de Regresión Ridge

La regresión Ridge es una técnica que penaliza el *overfitting* del modelo ya que añade un término de penalización (L2) a la función de pérdida que se utiliza para ajustar el modelo. Este término de regularización se controla con el parámetro *Lambda* y el objetivo es penalizar los coeficientes de las variables independientes reduciendo su magnitud (Gil Martínez, 2018).

Para ajustar este modelo se ha utilizado también la partición hecha anteriormente por la que se han creado los conjuntos de datos llamados “datos_entrenamiento1” y “datos_prueba1”. En primer lugar, es necesario convertir los datos de entrenamiento en una matriz que contenga las variables predictoras a la que se ha llamado “X”. En segundo lugar, se ha creado un vector que contiene los datos de la variable objeto del conjunto de datos de entrenamiento (“porcentaje_teletrabajo”) al que se ha llamado “y”.

El siguiente paso es ajustar el modelo utilizando la función “glmnet”. Los argumentos de esta función son los siguientes: los datos utilizados para ajustar el modelo contenidos en la matriz “X”; la variable objetivo contenida en el vector “y”; el parámetro *Alpha* igual a 0, que indica que el tipo de regularización empleada es Ridge; y el valor de *Lambda* que es el factor de penalización (Pérez Planells et al., 2015).

Para determinar el valor óptimo de *Lambda* se ha utilizado la técnica de validación cruzada, que selecciona el valor del parámetro que minimiza el error de predicción. El resultado ha sido 0,4620567 (Gil Martínez, 2018).

```
ridge_model <- glmnet(X, y, alpha = 0, lambda = best_lambda_ridge)
```

Este tipo de regresión no permite hacer selección de variables, es decir no puede fijar el coeficiente de las variables en 0. A continuación se muestran los coeficientes asociados a cada una de las variables independientes.

Ilustración 11. Visualización de los coeficientes asociados a cada variable independiente según el modelo de regularización Ridge

(Intercept)	7.68486624
PIB_norm	-0.05885364
teletrabajo_SELF_norm	0.31050206
teletrabajo_SAL_norm	1.40112168
teletrabajo_MUJER_norm	1.43323319
teletrabajo_HOMBRE_norm	1.42314458

Para poder comparar este modelo con el anterior, se van a estudiar las métricas asociadas a los siguientes errores.

Ilustración 12. Métricas Errores modelo de regularización Ridge

```
"Metricas modelo Ridge"  
"RMSE"  
0.0418624  
"MAE"  
0.03102622  
"MAPE"  
0.4881706
```

El valor del RMSE de este modelo es de 0,042, lo que sugiere que el modelo tiene un buen ajuste a los datos. Además, es menor que el valor obtenido para el modelo anterior (0,12), lo que sugiere que este tiene un menor ajuste. Lo mismo pasa con el valor del MAE obtenido para este modelo, ya que en este caso es de 0,031; mientras que en el anterior era de 0,089. Esto significa que este modelo tiene una menor discrepancia promedio entre las predicciones

y los valores reales en comparación con el anterior. Por último, el valor del MAPE es de 0,488%, lo que lo hace también inferior al del modelo anterior (1,05%). Por lo tanto, parece que este modelo tiene mejor rendimiento que el modelo 1.

5.1.3. Modelo de Regresión Lasso

La regresión Lasso es un método que ajusta el modelo haciendo selección automática de variables. Este tipo de regresión también utiliza un término de regularización en la función de pérdida que se utiliza para entrenar el modelo. La diferencia es que la regresión Ridge utiliza la norma L2, que penaliza los coeficientes sumando el cuadrado de sus magnitudes (Amat Rodrigo, 2020). En cambio, la regresión Lasso utiliza la norma L1, que penaliza los coeficientes del modelo sumando el valor absoluto de sus magnitudes. Por lo tanto, la regresión Lasso hace selección de variables estableciendo el valor de algunos coeficientes en cero (Gil Martínez, 2018). Al igual que en la regresión Ridge, el parámetro que permite controlar el nivel de penalización de los coeficientes es *Lambda* (Adelheid Januavian et al., 2020).

Para ajustar este modelo, se han utilizado la matriz “X” y el vector “y”, creados para entrenar el modelo de regresión Ridge, ya explicados en el apartado anterior. Se utiliza la misma función que para la regresión Ridge, pero para indicar que se trata de regresión Lasso se especifica que *Alpha* es igual a 1. Para determinar el valor del parámetro *Lambda* (término de regularización), se ha utilizado de nuevo la técnica de validación cruzada que da como resultado 0,134692. Fijado el valor de *Lambda*, se ha entrenado el modelo de la siguiente forma:

```
lasso_model <- glmnet(X, y, alpha = 1, lambda = best_lambda_lasso) #0.134692
```

Las variables seleccionadas por el modelo de regresión Lasso son las siguientes: “teletrabajo_SAL_norm”, “teletrabajo_MUJER_norm”, y “teletrabajo_HOMBRE_norm”. Es decir, el modelo hace selección de variables reduciendo algunos de los coeficientes de las variables a cero para obtener un modelo más sencillo. A continuación se muestran los coeficientes de las variables:

Ilustración 13. Visualización de los coeficientes asociados a cada variable independiente según el modelo de regularización Lasso

(Intercept)	7.6825618
PIB_norm	.
teletrabajo_SELF_norm	.
teletrabajo_SAL_norm	0.3462167
teletrabajo_MUJER_norm	1.8907448
teletrabajo_HOMBRE_norm	2.2092351

Para terminar, se van a analizar las métricas de los errores asociados al modelo de regresión Lasso en comparación con el modelo de regresión Ridge.

Ilustración 14. Métricas Errores modelo de regularización Lasso

```
"Metricas modelo lasso"  
"RMSE"  
0.102183  
"MAE"  
0.08175293  
"MAPE"  
1.803652
```

En cuanto al RMSE, este modelo ha obtenido un valor de 0,10, lo que implica que la diferencia entre los valores predichos y los reales es de media 0,10. El modelo de regresión Ridge tenía un valor más pequeño, lo que indica que en términos de RMSE el modelo anterior se ajusta mejor a los datos. Sin embargo, el modelo de regresión Lasso tiene unos valores más grandes en cuanto a MAE (0,08 contra 0,03) y MAPE (1,8 contra 0,48). Teniendo en cuenta estas últimas dos métricas y sobre todo la considerable diferencia del valor del MAPE, se ha considerado que el modelo de regresión Ridge tiene una mejor capacidad de ajuste a los datos.

5.2. RESULTADO PREDICCIONES

En cuanto a los resultados de la estimación, hay que tener en cuenta que, como ya se ha explicado en la metodología del presente trabajo, el método que se ha utilizado para estimar las variables independientes ha sido el de regresión lineal. Los resultados de las estimaciones han sido recogidos en un conjunto de datos al que se ha llamado “datos_nuevos”. Hay que

tener en cuenta que para proceder a la predicción con los modelos es necesario estandarizar las variables, ya que de lo contrario los resultados pueden verse distorsionados por las diferentes magnitudes de los datos. El método que se ha empleado para estandarizar las variables ha sido la normalización, que consiste en restar la media y dividir entre la desviación estándar.

La normalización de las variables se ha hecho con la función “scale” de R studio, siguiendo el mismo proceso explicado para normalizar las variables del conjunto de datos que se ha utilizado para entrenar los modelos. Al conjunto de datos que recoge las estimaciones de las variables predictoras normalizadas se le ha llamado “datos_nuevos_norm”.

```
datos_nuevos_norm <- datos_nuevos%>%  
  
mutate(  
  
  PIB_norm = scale(PIB),  
  
  teletrabajo_SELF_norm = scale(teletrabajo_SELF),  
  
  teletrabajo_SAL_norm = scale(teletrabajo_SAL),  
  
  teletrabajo_MUJER_norm = scale(teletrabajo_MUJER),  
  
  teletrabajo_HOMBRE_norm = scale(teletrabajo_HOMBRE) )
```

A continuación, se van a presentar las estimaciones de las variables para cada uno de los modelos entrenados.

5.2.1. Modelo de Regresión Lineal

Para obtener las predicciones según el modelo de regresión lineal, se ha utilizado la función “predict” de R studio con las estimaciones normalizadas, es decir el conjunto de datos llamado “datos_nuevos_norm”. La función “predict” en R studio requiere de los siguientes argumentos: el modelo entrenado previamente, en este caso modelo1; y las nuevas observaciones para las que se quiere predecir, en este caso datos_nuevos_norm (Maurandi López & González Vidal, 2022).

```
predicciones_lm_2023_2024 <- predict(modelo1, newdata = datos_nuevos_norm)
```

Tabla 6. Predicciones del porcentaje de teletrabajo según el modelo 1

Predicciones porcentaje teletrabajo modelo lm	2023	2024
Alemania	12,06065 %	14,267098 %
Italia	2,855516 %	3,728247 %
Portugal	6,131574 %	7,027864 %

5.2.2. Modelo de Regresión Ridge

Igual que en el caso anterior, la función utilizada para predecir ha sido “predict” de R studio. Sin embargo, hay que tener en cuenta que para entrenar el modelo de regresión Ridge se han recogido las variables predictoras en la matriz “X”. Por lo tanto, para predecir, también hay que convertir las estimaciones normalizadas de las variables predictoras en una matriz utilizando la función “as.matrix”.

```
predicciones_ridge <- predict(ridge_model, newx = as.matrix(datos_nuevos_norm[,
c("PIB_norm", "teletrabajo_SELF_norm", "teletrabajo_SAL_norm",
"teletrabajo_MUJER_norm", "teletrabajo_HOMBRE_norm")]))
```

Tabla 7. Predicciones del porcentaje de teletrabajo según el modelo de regularización Ridge

Total datos modelo Ridge	2023	2024
Alemania	11,989845 %	14,178469 %
Italia	2,855336 %	3,768198 %
Portugal	6,206271 %	7,111079 %

5.2.3. Modelo de Regresión Lasso

El sistema para entrenar el modelo de regresión Lasso es igual al del modelo de regresión Ridge en cuanto a que también llama a la matriz “X” que recoge las variables. Por lo tanto,

a la hora de predecir, también es necesario convertir las estimaciones normalizadas de las nuevas observaciones que se desean predecir en una matriz.

```
predicciones_lasso <- predict(lasso_model, newx = as.matrix(datos_nuevos_norm[,
c("PIB_norm", "teletrabajo_SELF_norm", "teletrabajo_SAL_norm",
"teletrabajo_MUJER_norm", "teletrabajo_HOMBRE_norm")]))
```

Tabla 8. Predicciones del porcentaje de teletrabajo según el modelo de regularización Lasso

Total datos modelo Lasso	2023	2024
Alemania	11,938006 %	14,083401 %
Italia	2,98422 %	3,838172 %
Portugal	6,186579 %	7,064992 %

5.2.4. Análisis comparativo de las predicciones de los tres modelos

Como se observa en los datos recogidos en las tablas anteriores, las diferencias entre las predicciones de cada modelo no son especialmente significativas. En el caso de Alemania, las predicciones del modelo de regresión lineal son las más altas. El modelo de regresión Ridge estima una cifra ligeramente superior para el año 2023 a la del modelo de regresión Lasso; mientras que para el año 2024 es al revés. En cuanto a Italia, las predicciones de los 3 modelos son prácticamente iguales, aunque el modelo de regresión Lasso da un valor superior por una décima. Por último, en el caso de Portugal, el modelo de regresión Ridge es el que da estimaciones más altas para los dos años, aunque la diferencia no llega a una décima.

A continuación, se incluyen dos gráficos de líneas que ofrecen una perspectiva visual de la información anteriormente expuesta. El primero incluye la evolución del porcentaje de teletrabajo para las observaciones de la base de datos original desde 2015 hasta 2022. Además, en el mismo gráfico se han incluido las predicciones hechas por cada uno de los modelos para los años 2023 y 2024. El gráfico refleja que apenas hay diferencia entre los

valores; sin embargo, si hay que destacar un país en el que los modelos presentan estimaciones de predicciones más divergentes, este sería Alemania.

El segundo gráfico muestra únicamente las predicciones de cada uno de los modelos para los años 2023 y 2024 por colores. Las predicciones del modelo de regresión lineal aparecen representadas por las líneas verdes, las del modelo de regresión Ridge por las líneas magenta, y las del modelo de regresión Lasso por las líneas azules.

Gráfico 9. Evolución del porcentaje de teletrabajo entre los años 2015-2022 con las predicciones para 2023 y 2024 según los tres modelos

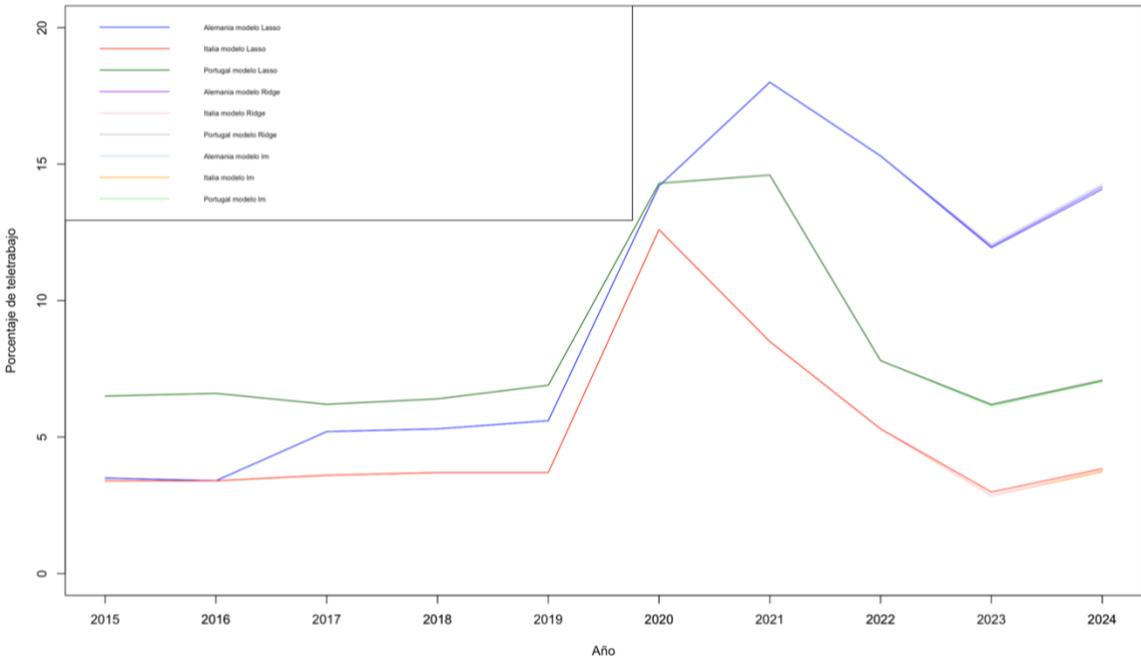
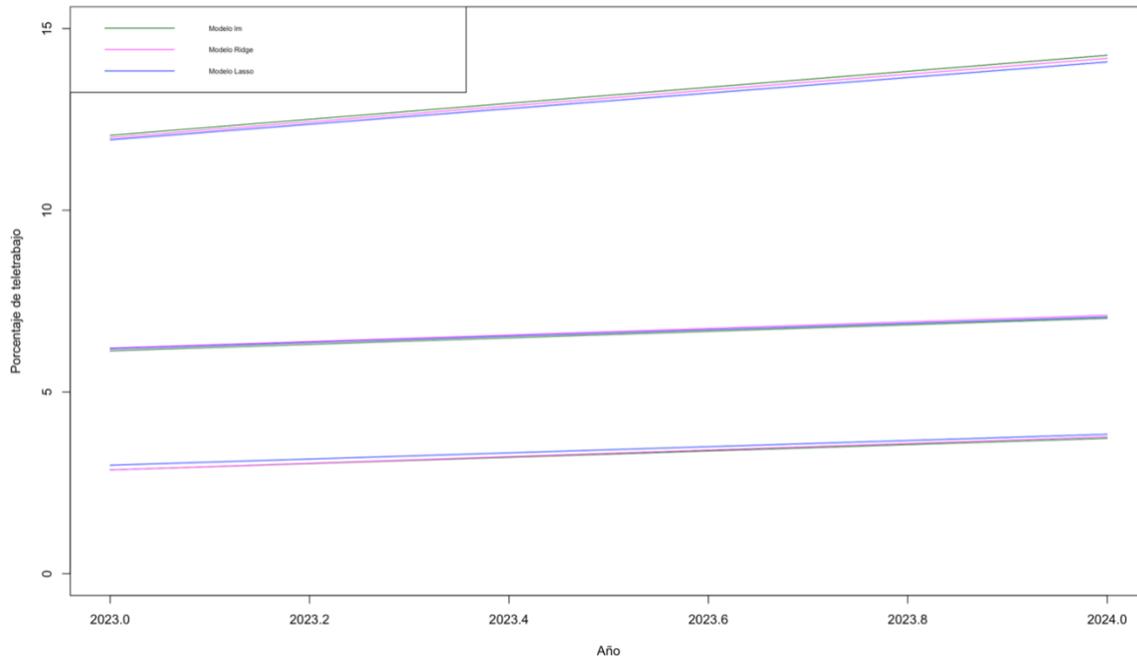


Gráfico 10. Predicciones del porcentaje de teletrabajo para 2023 y 2024 según los tres modelos



A continuación, se muestran dos tablas: una tabla con el resumen de los resultados para cada país según el modelo y otra que recoge los principales resultados para cada país.

Tabla 9. Resumen de los resultados de las predicciones para cada país según el modelo

Modelo	Año	Alemania	Italia	Portugal
Predicciones RL	2023	12,06065 %	2,855516 %	6,131574 %
	2024	14,267098 %	3,728247 %	7,027864 %
Predicciones Ridge	2023	11,989845 %	2,855336 %	6,206271 %
	2024	14,178469 %	3,768198 %	7,111079 %
Predicciones Lasso	2023	11,938006 %	2,98422 %	6,186579 %
	2024	14,083401 %	3,838172 %	7,064992 %

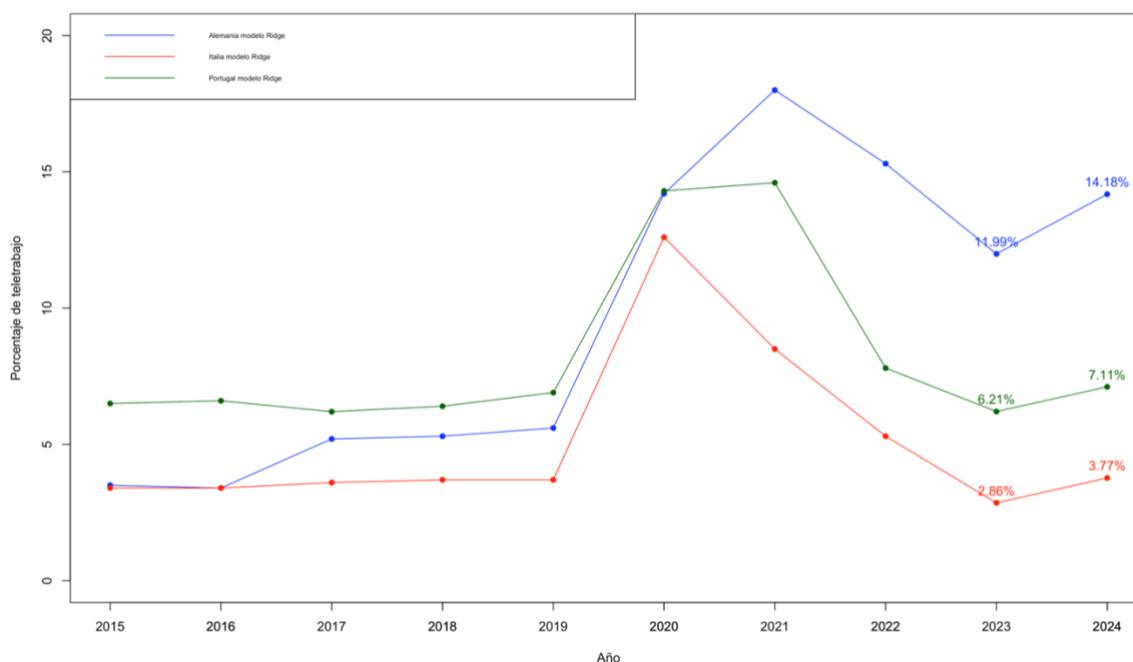
Tabla 10. Resumen de los principales resultados para cada país

Alemania	Italia	Portugal
Tendencia a crecer.	Crecimiento lento y valores muy bajos.	Crecimiento lento y cifras no muy bajas.
Estabilización en valores más altos que antes del Covid.	En 2023 se esperan valores más bajos a los de antes del Covid.	En 2023 se esperan valores más bajos a los de antes del Covid.
El Covid parece haber acelerado la implantación del teletrabajo.	El Covid no parece haber acelerado la implantación del teletrabajo.	El Covid ha consolidado la implantación del teletrabajo.

6. DISCUSIÓN

Dado que se ha determinado que el modelo de regresión Ridge exhibe las mejores métricas en términos de los errores analizados, se han representado visualmente los resultados de dicho modelo en un gráfico. El gráfico representa la evolución del porcentaje de teletrabajo para los datos conocidos hasta 2022. Además, incluye también las estimaciones hechas por el modelo y sus valores.

Gráfico 11. Evolución del porcentaje de teletrabajo entre los años 2015-2022 con las predicciones para 2023 y 2024 según el modelo de regresión Ridge



Dado que no se han encontrado cifras específicas que recojan el porcentaje de teletrabajo esperado para los años 2023 y 2024, se van a comparar las tendencias que predice el modelo de regresión Ridge con las tendencias esperadas en cada país en virtud de la literatura jurídica.

6.1. DESCENSO PREVISTO PARA EL AÑO 2023

El modelo predice una bajada de las cifras del teletrabajo para el año 2023 que coincide en todos los países. Esta bajada es de un 3,31% para Alemania, 2,45% para Italia, y 1,59% para Portugal. Lo sorprendente es que para Italia y Portugal se esperan cifras por debajo de los valores recogidos el año anterior a la pandemia.

En cuanto a Alemania, según la literatura científica, las cifras del trabajo a distancia aumentaron exponencialmente durante la pandemia. La tasa de teletrabajo llegó a alcanzar hasta un 59% en el sector de las TIC. Sin embargo, otros sectores como la gastronomía o la construcción reflejaron niveles muy bajos e incluso inexistentes de teletrabajo. Por lo tanto, el descenso en el teletrabajo previsto para el año 2023 en Alemania podría explicarse por la

vuelta a la normalidad que ya era muy notable en el año 2022. La vuelta al trabajo de los sectores que no favorecen el teletrabajo, y el hecho de que retomen sus operaciones presenciales reduciría el porcentaje general del teletrabajo del país. Este descenso es necesario para que el porcentaje de teletrabajo se estabilice tras el pico experimentado a causa de la pandemia.

En Italia, observando los datos utilizados para la implementación del modelo de regresión Ridge, se aprecia que el teletrabajo apenas era utilizado antes de la crisis sanitaria. Las cifras se mantuvieron bajas y relativamente constantes sin llegar al 3%. Esto unido a la vuelta a la normalidad tras el Covid, que ya es claramente palpable en los resultados recogidos para el año 2022, podría explicar el descenso de la cifra esperada para el año 2023. Además, también hay que tener en cuenta que la estructura industrial italiana está compuesta principalmente por empresas de reducido tamaño y que la inversión en I+D es reducida. Por lo tanto, la correcta implementación del teletrabajo podría verse dificultada por la falta de recursos y por la cultura laboral italiana tan arraigada a la modalidad de trabajo presencial. Este último factor dista mucho de cómo se ha implantado y aceptado el trabajo a distancia en otros países, como por ejemplo Alemania.

Según los datos utilizados para entrenar el modelo, parece que el porcentaje de teletrabajo recogido para el año 2022 en Portugal estaba todavía bajo la influencia de las restricciones del Covid. Por eso se espera que siga descendiendo en 2023, al igual que sucede con el resto de los países objeto del presente estudio. Según estos datos, la tasa de teletrabajo recogida para Portugal, entre los años 2015 y 2019 estaba bastante por encima en comparación con Alemania e Italia. Sin embargo, el crecimiento antes del Covid se vio bastante estancado desde 2015 con algunos descensos a lo largo de los años, a diferencia de Alemania donde el trabajo a distancia fue creciendo todos los años sin que se registrara ningún descenso. Esta bajada estimada para 2023 podría deberse también a que en Portugal el teletrabajo se concentró sobre todo en las áreas metropolitanas donde se encuentran las empresas dedicadas a actividades de las TIC, financieras o de consultoría. Si estos sectores han experimentado un descenso en la adopción del teletrabajo con la normalización tras la pandemia, quedaría justificado el descenso estimado para el año 2023.

En términos generales, este descenso en el teletrabajo puede deberse al fin de la pandemia y a la relajación de las restricciones establecidas por los gobiernos. Es decir, podría asociarse a la transición hacia la normalidad en la que las empresas retoman gradualmente la modalidad presencial de trabajo. En vistas a la vuelta progresiva a la normalidad, puede ser que las empresas hayan adoptado la modalidad de trabajo híbrido, que combina el trabajo a distancia y presencial. Además, puede ser que el teletrabajo descienda según se delimitan y aclaran las normas sobre cuándo y cómo se puede trabajar a distancia. Esta precisión en las normas relativas al teletrabajo puede haber sido necesaria por las dificultades surgidas a la hora de controlar a los trabajadores a distancia, o por la falta de compromiso con el trabajo por parte de quienes teletrabajan. Por otro lado, hay que tener en cuenta que hay determinados servicios, como la construcción o el turismo, que no permiten su prestación a distancia. Es probable que con el fin de la pandemia haya vuelto a aumentar la demanda de este tipo de trabajos que requieren de presencialidad y estuvieron paralizados o funcionando bajo mínimos durante la pandemia. Por lo tanto, el teletrabajo podría haberse visto afectado por un cambio en la demanda laboral.

6.2. ASCENSO PREVISTO PARA EL AÑO 2024

El modelo prevé un aumento general del porcentaje de teletrabajo para todos los países en el año 2024. Esta subida sería más acusada para Alemania, donde superaría el 2%; que, para Portugal e Italia, donde no llega al 1%. Sin embargo, este incremento en Italia y Portugal es lo suficientemente alto como para superar muy ligeramente las cifras recogidas en 2019.

El aumento del teletrabajo estimado para el año 2024 en Alemania según los resultados del modelo de regresión Ridge podría explicarse por la creciente aceptación del teletrabajo. Como se ha expuesto anteriormente, los resultados del estudio publicado por el IFO presentan una gran aceptación en cuanto al teletrabajo. Por lo tanto, esto explicaría la tendencia a la alza de las cifras relativas al porcentaje de teletrabajo una vez alcanzada la estabilidad. Es posible que durante el auge de la pandemia, tanto las empresas como los trabajadores hayan experimentado los beneficios del teletrabajo. Estos beneficios pueden ser de tipo económico tanto para las empresas que ahorran en costes fijos, como para los trabajadores que ahorran

en gastos de transporte. Además, los empleados probablemente hayan experimentado otros beneficios como la mayor flexibilidad y la mejor conciliación entre la vida laboral y personal.

En cuanto al incremento esperado para el año 2024 en Italia, se trata de una subida tan pequeña que no logra alcanzar niveles significativamente más elevados que los recogidos para el año 2019. Esto apunta a que la implementación del teletrabajo en Italia experimentará una recuperación gradual, lenta y limitada. Atendiendo a la literatura recopilada, si bien es cierto que el teletrabajo aumentó de forma muy acusada durante la crisis sanitaria, los sectores que concentraron las cifras más altas fueron los servicios de información y comunicación y las actividades profesionales, científicas y técnicas. Estos sectores tenían más fácil acceso a las herramientas necesarias para hacer frente al teletrabajo, y las ubicaciones de las empresas que enfocadas a las TIC o actividades profesionales suelen estar concentradas en las zonas más ricas de Italia, como Roma o Milán. Por eso, todo apunta a que el teletrabajo se adoptará muy lentamente en este tipo de regiones que son más ricas.

En relación con el aumento esperado para el año 2024 en Portugal, podría explicarse por la adaptación de los empleados y las empresas a las nuevas formas de trabajo. Es posible que, tras la vuelta a la presencialidad, tanto trabajadores como empleadores valoren los beneficios que aporta el teletrabajo y decidan optar por mantener un mínimo de trabajo a distancia. La experiencia adquirida durante la pandemia probablemente haya contribuido a su aceptación, que, si bien ya estaba relativamente generalizada, se ha visto potenciada por la pandemia, lo que conlleva a un aumento de las cifras en comparación con los niveles asociados a antes de la pandemia. Otro punto que ha de tenerse en cuenta es la aprobación de un marco legal que garantiza una mayor seguridad tanto para las empresas como para los trabajadores. Esto probablemente contribuya a impulsar los niveles de trabajo a distancia, ya que brinda garantías adicionales para los empleados, lo que posiblemente contribuya positivamente a su aceptación.

En general, la recuperación de las cifras previstas para 2024 puede deberse a motivos como la mayor aceptación del teletrabajo por parte de las empresas, que se han adaptado a esta nueva modalidad de trabajo de forma paulatina; y de los empleados que se han acostumbrado a trabajar a distancia. Además, podría ser que en virtud de las ventajas que ofrece el teletrabajo, como son la flexibilidad y la reducción de costes, se prevea un aumento a futuro.

En base a esto, puede ser que las empresas hayan implementado políticas laborales más flexibles que fomenten el trabajo a distancia. Además, quizás haya habido de nuevo un cambio en la demanda laboral, revirtiéndose así el incremento previo en la demanda de empleos que requieren presencialidad. Esto tal vez sea se deba a que el aumento anterior era consecuencia de una reactivación tras un periodo de inactividad, y ahora se está volviendo a la estabilidad del mercado laboral.

6.3. ANÁLISIS DEL NIVEL DE ESTABILIZACIÓN DEL PORCENTAJE DE TELETRABAJO DESPUÉS DEL COVID

El modelo de regresión Ridge estima que el porcentaje de trabajo a distancia en Alemania se estabilizará en un valor que se encuentra por encima a los valores recogidos para el año 2019. Esto puede deberse a la rápida evolución y adopción de tecnologías digitales en Alemania, que han permitido a los trabajadores y a las empresas explotar los beneficios del trabajo a distancia de forma efectiva y productiva. Además, tal y como se ha mencionado, el estudio del IAO explica que en Alemania ha habido un cambio en la forma de organización del trabajo. Durante la pandemia tanto los trabajadores como los empleadores tuvieron que adaptarse al teletrabajo rápidamente, y la adaptación en Alemania fue buena ya que la experiencia en este país probó que el trabajo a distancia puede ser productivo y efectivo. Esto ha traído consigo una mayor predisposición para implementarlo de forma permanente.

El motivo por el que el crecimiento no llega a alcanzar niveles tan altos como los alcanzados durante la pandemia, es precisamente porque no en todas las regiones alemanas cuentan con el mismo número de puestos de trabajo que permiten la posibilidad de implar el trabajo a distancia. Las regiones más ricas y con más fácil acceso a las herramientas necesarias para el teletrabajo, como el acceso a internet de alta velocidad o herramientas de videoconferencia, experimentarán un aumento más acusado de las cifras de teletrabajo. Sin embargo, las regiones que concentren un número más alto de puestos de trabajo que no permiten el trabajo a distancia, verán estancado o aumentado en una cifra muy pequeña el porcentaje de teletrabajo. Esto hará que se vuelvan cada vez más acusadas las diferencias entre Oriente y Occidente en Alemania, tal y como prevé el estudio realizado por el IZA.

El modelo apunta hacia la estabilización del porcentaje de trabajo a distancia por encima de los valores recogidos antes del Covid para Italia. Esto puede deberse a que los trabajadores y los empleadores hayan experimentado los beneficios económicos vinculados al teletrabajo, como el ahorro de costes de transporte o energéticos. Sin embargo, dadas las dificultades que ha supuesto la implantación del teletrabajo en Italia, lo más probable es que estas ventajas se hallan visto opacadas por los problemas con relación a este método de trabajo. La estructura industrial y empresarial italiana, construida por pequeñas empresas, unida a la escasa inversión en I+D, hace que sea más difícil facilitar las herramientas necesarias a los trabajadores para que puedan desarrollar el trabajo a distancia de forma eficaz y productiva. Quizás esto haya influido en su aceptación y acogida, lo que explicaría que esta sea más baja que en países como Alemania en los que las facilidades han sido mayores. Por lo tanto, a medida que se vayan solucionando los problemas que los empleados tuvieron con el teletrabajo cuando era obligatorio, la aceptación irá creciendo, lo que conllevará un aumento en el porcentaje de teletrabajo.

Otro factor importante para tener en consideración es la regulación legal, ya que en Italia se reguló esta modalidad de trabajo por primera vez en 1998. En 2017, es decir antes del Covid se promulgó la Ley del trabajo ágil, que, aunque no es exactamente lo mismo puede haber aportado algunas pautas para la implementación del trabajo a distancia en el país. De hecho, la prueba de ello es que en los datos que se han utilizado para implementar el modelo se puede apreciar un ligero aumento del teletrabajo el año 2018. Por lo tanto, la paulatina aceptación de los beneficios ligados al teletrabajo y la existencia de un marco legal que aporte seguridad al empresario y al trabajador apuntan hacia una mayor conciencia y adopción del teletrabajo. Sin embargo, hay que tener en cuenta factores como la concentración de empresas y recursos en las zonas metropolitanas; y las diferencias en cuanto a estructura tecnológica y cultura laboral que influirán en el crecimiento de las regiones. Esto dará lugar a una aceleración en las áreas urbanas y una ralentización de su implantación en las zonas menos desarrolladas.

Por último, tanto la literatura recopilada como las estimaciones del modelo apuntan a que el teletrabajo en Portugal crecerá en los años futuros. Esto se debe a que los valores se han estabilizado en una cifra que está por encima del valor registrado en el año 2019. Después de

un primer periodo de adaptación a cambios bruscos en los patrones laborales, es necesario que los niveles de teletrabajo se establezcan y ajusten a las necesidades de las empresas y de los trabajadores. Una vez alcanzada esta estabilidad en 2023, se ajustará la adopción del teletrabajo conduciendo este ajuste a un valor ligeramente superior a los valores pre-Covid. Además, la cultura laboral en Portugal no parece estar tan arraigada al trabajo presencial en vista a las cifras recogidas para los años anteriores a la pandemia. Esta diferencia es importante ya que en comparación con Italia el crecimiento probablemente sea más rápido en Portugal que en Italia.

En términos generales, podría decirse que el aumento del teletrabajo durante la crisis sanitaria fue debido a la necesidad de implementar medidas de distanciamiento social para evitar la expansión del Covid. Sin embargo, a medida que la situación ha vuelto a la normalidad y las restricciones han sido suprimidas, es normal que se espere una disminución del trabajo a distancia. Esto puede encontrar explicación en un fenómeno matemático que es conocido como la regresión a la media. Este concepto estadístico sugiere que los valores extremos tienden a regresar hacia la media con el paso del tiempo. Es decir, durante la pandemia hubo una demanda excepcional de trabajo a distancia; sin embargo, con la relajación de las restricciones, esta demanda de teletrabajo ha disminuido. Por lo tanto, es probable que las empresas opten por modalidades que combinen el trabajo presencial y el teletrabajo de manera equilibrada.

7. CONCLUSIONES

Para comenzar este apartado, se va a hacer mención a las limitaciones a las que se ha visto sujeta la presente investigación. En primer lugar, la base de datos contaba con muchos valores perdidos por lo que aunque aparentemente parece que hay mucha información sobre muchos países, la realidad es diferente. Esto ha sido una importante limitación a la hora de elegir los países en los que se ha focalizado el estudio. De haber tenido más información, se hubiera elegido un país con tasas más elevadas de teletrabajo antes de la pandemia, como por ejemplo Países Bajos.

En segundo lugar, otra de las limitaciones del trabajo ha sido la estructura de la base de datos utilizada, ya que no tenía variables dependientes e independientes como tal. Esto se ha

solucionado creando las variables en base a la información proporcionada por el conjunto de datos. En tercer lugar, a la hora de estimar las variables independientes, no se han encontrado estimaciones oficiales dada la especificidad de las variables creadas. Por lo tanto se han creado varios modelos de regresión lineal utilizando los datos históricos disponibles para estimar cada una de las variables independientes. La última de las limitaciones a la que ha quedado supeditada la investigación ha sido la imposibilidad de incluir algunas variables que pueden incidir en la implantación del teletrabajo, como el desarrollo tecnológico de los países o la composición de los hogares.

Basándome en los resultados de la presente investigación se puede concluir que a raíz del Covid en Alemania se dio un aumento significativo del teletrabajo, especialmente en los sectores relacionados con las TIC. En cambio, para el año 2023 se prevé un descenso en las cifras del teletrabajo que se justifica por la vuelta a la normalidad y a la presencialidad en los sectores que no favorecen el teletrabajo. Para el año 2024 se espera un incremento del trabajo a distancia gracias al alto nivel de acogida que esta modalidad ha tenido en Alemania. En términos generales, se espera que el porcentaje de teletrabajo se estabilice en un nivel superior a los valores pre-pandemia. Esto se debe principalmente a la rápida adopción de tecnologías digitales y a una mayor predisposición a implementarlo de forma permanente.

En cuanto a Italia, antes de la pandemia la tasa de teletrabajo era relativamente baja, lo que en parte puede deberse a que la estructura industrial del país está principalmente compuesta por pequeñas empresas y a las escasas inversiones en I+D. A pesar del aumento experimentado durante la pandemia, en 2023 se espera que la tasa de trabajo a distancia continúe en descenso hasta llegar a niveles por debajo de los recogidos antes del Covid. Esto puede encontrar justificación en la estructura industrial y en la cultura laboral italiana, que dificultan la implementación del teletrabajo en el país. Para 2024 se espera un ligero ascenso que no será significativo, lo que indica que la adopción del teletrabajo en Italia se hará de forma lenta y con limitaciones.

Por último, en relación a Portugal, el país experimentó un crecimiento estancado en la tasa de teletrabajo antes de la crisis sanitaria. Los porcentajes más altos estaban concentrados en las áreas metropolitanas y en sectores relacionados con las TIC. Se espera que el trabajo a distancia en Portugal experimente un descenso en 2023 ligado a la vuelta a la normalidad

asociada al fin de la pandemia. En cambio, para el año 2024 se espera un aumento gradual del trabajo a distancia por la toma de conciencia de empleados y empresas que se adaptan a esta nueva modalidad. La acogida del teletrabajo en Portugal ha sido buena ya que ha tenido una buena aceptación, por lo que se puede concluir que la tendencia esperada en este país es a la alta.

Por lo tanto, en base a las tendencias analizadas, parece que en general el trabajo a distancia no ha alcanzado todavía la estabilidad post-pandemia, lo que puede ser interesante como futura línea de investigación. Cabe esperar que una vez alcanzada la estabilidad, el trabajo a distancia no desaparezca ya que el Covid ha acelerado su implementación. Los resultados de la investigación podrían extrapolarse a países con circunstancias y características similares, lo que quizás sea interesante como futura línea de investigación. Por lo tanto, a medida que se superen las barreras y dificultades vinculadas a la implantación del teletrabajo, como por ejemplo la infraestructura tecnológica o la cultura laboral arraigada, es probable que el teletrabajo se consolide como una opción más frecuente y permanente en el mundo laboral. Aunque los escenarios sean diferentes según las particularidades de cada país y región, en términos generales se espera que el teletrabajo marque una transformación duradera en la forma de organización y realización del trabajo.

8. ÍNDICE DE GRÁFICOS, ILUSTRACIONES Y TABLAS

Gráfico 1. Evolución del teletrabajo total de autónomos por país según las estimaciones del método de la media del diferencial del incremento. Años 2015-2024	24
Gráfico 2. Gráfico de dispersión que refleja la evolución del teletrabajo de los autónomos en Alemania.....	26
Gráfico 3. Evolución del teletrabajo de los autónomos según el método de estimación con modelos de regresión lineal. Años 2015-2024	28
Gráfico 4. Comparación de las estimaciones de la variables “teletrabajo_SELF” según cada método de estimación. Año 2023	29
Gráfico 5. Comparación de las estimaciones de la variables “teletrabajo_SAL” según cada método de estimación. Año 2023	29
Gráfico 6. Comparación de las estimaciones de la variables “teletrabajo_MUJER” según cada método de estimación. Año 2023	30
Gráfico 7. Comparación de las estimaciones de la variables “teletrabajo_HOMBRE” según cada método de estimación. Año 2023	30
Gráfico 8. Evolución del teletrabajo del total de autónomos por país según cada método de estimación. Años 2015-2024	32
Gráfico 9. Evolución del porcentaje de teletrabajo entre los años 2015-2022 con las predicciones para 2023 y 2024 según los tres modelos	44
Gráfico 10. Predicciones del porcentaje de teletrabajo para 2023 y 2024 según los tres modelos.....	45
Gráfico 11. Evolución del porcentaje de teletrabajo entre los años 2015-2022 con las predicciones para 2023 y 2024 según el modelo de regresión Ridge.....	47

Ilustración 1. Visualización del dataset “data_model”	15
Ilustración 2. Visualización del dataset “data_long” previa inclusión de las variables independientes	15
Ilustración 3. Visualización del dataset “data_long” con las variables independientes	17
Ilustración 4. Resultados de la comparativa del rendimiento de los tres modelos hecha por R studio	19
Ilustración 5. Visualización del conjunto de datos “datos_nuevos” que recoge las estimaciones según el método de la media del diferencial del incremento	24
Ilustración 6. Coeficiente de correlación Pearson	26
Ilustración 7. Visualización del conjunto de datos “datos_nuevos” que recoge las estimaciones realizadas por los modelos de regresión lineal	27
Ilustración 8. Resumen modelo 1	34
Ilustración 9. Métricas Errores modelo 1	36
Ilustración 10. Visualización de los resultados del estudio del overfitting del modelo 1	37
Ilustración 11. Visualización de los coeficientes asociados a cada variable independiente según el modelo de regularización Ridge	38
Ilustración 12. Métricas Errores modelo de regularización Ridge	38
Ilustración 13. Visualización de los coeficientes asociados a cada variable independiente según el modelo de regularización Lasso	40
Ilustración 14. Métricas Errores modelo de regularización Lasso	40

Tabla 1. Porcentaje de teletrabajo en la UE según categoría profesional.....	6
Tabla 2. Tabla resumen	13
Tabla 3. Valores del PIB per cápita en el año 2022	14
Tabla 4. Estimaciones del PIB per cápita para los años 2023 y 2024	22
Tabla 5. Estimaciones para la variables SELF según el método de la media del diferencial del incremento	23
Tabla 6. Predicciones del porcentaje de teletrabajo según el modelo 1	42
Tabla 7. Predicciones del porcentaje de teletrabajo según el modelo de regularización Ridge	42
Tabla 8. Predicciones del porcentaje de teletrabajo según el modelo de regularización Lasso	43
Tabla 9. Resumen de los resultados de las predicciones para cada país según el modelo ...	45
Tabla 10. Resumen de los principales resultados para cada país	46

9. ANEXOS

Para más detalles sobre el código ver: [Código modelo predictivo teletrabajo Reyes Ruiz-Labourdette.R](#)

10. REFERENCIAS

- Adelheid Januavian, T., Sukono, S., & Lesmana, E. (2020). Modeling of money supply using LASSO regression with Cross-Validation. *Revista de Ciencias Humanas y Sociales*(91), 213-231.
- Agencia Informativa Latinoamericana. (2020). *El servicio de información del Instituto Económico Alemán*. Recuperado el Mayo de 2023, de Crece el entusiasmo por la oficina en casa: <https://www.iwd.de/artikel/begeisterung-fuer-das-homeoffice-waechst-470071/>
- Alipour, J., Falck, O., & Schüller, S. (2020). *Germany's Capacities to Work from Home*”, *IZA – Institute of Labor Economics*, n. 13152. IZA – Institute of Labor Economics, Alemania.
- Amat Rodrigo, J. (2020). *cienciadedatos.net*. Obtenido de Selección de predictores, regularización ridge, lasso, elastic net y reducción de dimensionalidad: https://www.cienciadedatos.net/documentos/31_seleccion_de_predictores_subset_selection_ridge_lasso_dimension_reduction#Introducci%C3%B3n
- Belzuz Abogados. (2021). *Belzuz Abogados*. Obtenido de Nuevo régimen de teletrabajo en Portugal: <https://www.belzuz.net/es/publicaciones/en-espanol/item/11438-nuevo-regimen-del-teletrabajo-en-portugal.html>
- Cortés, F., & Rubalcava, R. (1983). Notas sobre colinealidad en el modelo de regresión lineal de tres variables. *Estudios demográficos y urbanos*, 17(2), 181-198.
- Datosmacro*. (2023). Obtenido de Diccionario de economía de Datosmacro: [https://datosmacro.expansion.com/diccionario/pib#:~:text=PIB%20per%20c%C3%A1pita%20\(o%20PIB,por%20cada%20uno%20de%20ellos](https://datosmacro.expansion.com/diccionario/pib#:~:text=PIB%20per%20c%C3%A1pita%20(o%20PIB,por%20cada%20uno%20de%20ellos).
- Datosmacro*. (2023). Obtenido de PIB de Italia: <https://datosmacro.expansion.com/pib/italia>

- Departamento de análisis Bankinter. (2023). *Blog de Economía y Finanzas Bankinter*. Obtenido de Previsión del PIB Eurozona para 2023 y 2024 (actualizado): <https://www.bankinter.com/blog/economia/previsiones-pib-eurozona>
- Dingel, J. I., & Neiman, B. (2020). *How many jobs can be done at home?* Cambridge: National Bureau of Economic Research.
- Eurostat. (2023). *Eurostat Data Browser*. Obtenido de Employed persons working from home as a percentage of the total employment, by sex, age and professional status (%): https://ec.europa.eu/eurostat/databrowser/view/lfsa_ehomp/default/table?lang=en
- Faramondi, A., & Costa, S. (2020). *Situazione e prospettive delle imprese nell'emergenza sanitaria Covid-19*. Istituto Nazionale di Statistica (ISTAT).
- Gil Martínez, C. (2018). *Técnicas de regularización y selección del mejor modelo*. Obtenido de https://rstudio-pubs-static.s3.amazonaws.com/392216_5adc98e84ff246e3b6418f3426c8fe38.html#m%C3%A9todos_de_regularizaci%C3%B3n
- Grunau, F., Steffes, S., & Wolters, S. (2020). Ministerio del Interior en tiempos de Corona: todavía hay potencial sin explotar en muchas profesiones. *IAB-Forum*.
- Hermann, J. (2020). Crece el entusiasmo por la oficina en casa. *El servicio de información del Instituto Económico Alemán*.
- INE. (2020). El teletrabajo en España y la UE antes de la COVID-19. *Boletín Informativo del Instituto Nacional de Estadística*.
- Instituto Vasco de Estadística. (2022). *Eustat*. Obtenido de Gasto en I+D interna (% PIB) por país. 2010-2021: https://www.eustat.eus/elementos/ele0003200/ti_Gasto_en_ID__PIB_por_pais_19972012/tbl0003292_c.html

- Irlacher, M., & Koch, M. (2020). Working from Home, Wages, and Regional Inequality in the Light of COVID-19. *CESifo Working Paper*(8232).
- Maurandi López, A., & González Vidal, A. (2022). *Análisis de datos y métodos estadísticos con R*. España: Universidad de Murcia.
- Milasi, S., González-Vázquez, I., & Fernández-Macías, E. (2021). Telework before the Covid-19 pandemic: trends and drivers of differences across the EU. *OECD Productivity Working Papers*, n. 21, Paris, 2021, p. 5.(21), 5.
- Millán, V. (2022). El teletrabajo ya es solo para privilegiados: este estudio en Italia lo demuestra. *El Economista*.
- Ministerio de Empleo y Seguridad Social. (2017). Debate acerca del impacto de la digitalización en el mercado de trabajo. *Actualidad Internacional Sociolaboral*(217), 94-105.
- Ministerio de Empleo y Seguridad Social. (2017). El teletrabajo en Alemania. *Actualidad Internacional Sociolaboral*(247), 104-115.
- OIT. (2020). *El teletrabajo durante la pandemia de COVID-19 y después de ella - Guía práctica*. Ginebra: Oficina Internacional de Trabajo.
- ORACLE. (s.f.). Obtenido de EPM Cloud Planning Translated Books: https://docs.oracle.com/cloud/help/es/pbcs_common/PFUSU/insights_metrics_RMSE.htm#PFUSU-GUID-FD9381A1-81E1-4F6D-8EC4-82A6CE2A6E74
- Pérez Planells, L., Delegido, J., Rivera, J., & Verrelst, J. (2015). Análisis de métodos de validación cruzada para la obtención robusta de parámetros biofísicos. *Revista de teledetección: Revista de la Asociación Española de Teledetección*(44), 55-65.
- Rodríguez Sánchez, A., García García, C., Salmerón Gómez, R., & García García, C. B. (2018). The Coefficient of Determination in the Ridge Regression. *Anales de ASEPUMA*(26).

Static Portugal. (2021). *No contexto da pandemia COVID-19 a Área Metropolitana de Lisboa concentrou quase metade do emprego em teletrabalho*. Instituto Nacional de Estatística Statistics Portugal.

Torniamo; Volvemos.org; Helping Networks. (2022). *RemotEU*. Obtenido de Regulación para el teletrabajo en Italia: <https://remotEU.info/regulacion-teletrabajo-italia#:~:text=La%20figura%20del%20teletrabajador%20lleva,smart%20working%20o%20trabajo%20%20C3%A1gil>

Velasco, L. (2022). *El teletrabajo en España. Antes, durante y después de la pandemia*. Observatorio Nacional de Tecnología y Sociedad (ONTISI).