

**RELACIÓN ENTRE INVERSIÓN EN VIVIENDAS Y
ACTIVOS ARRIESGADOS DURANTE EL CICLO DE
VIDA DE LOS HOGARES ESPAÑOLES**

Pablo Martín Murillo

Trabajo de investigación 023/016

Master en Banca y Finanzas Cuantitativas

Tutor:

Dra. Arantza Gorostiaga

Universidad Complutense de Madrid

Universidad del País Vasco

Universidad de Valencia

Universidad de Castilla-La Mancha

www.finanzasquantitativas.com

Trabajo Fin de Máster

Máster en Banca y Fianzas Cuantitativas

Relación entre inversión en vivienda y activos arriesgados durante el ciclo de vida de los hogares españoles

Pablo Manuel Martín Murillo

Dirección

Arantza Gorostiaga Alonso

Septiembre 2023

Resumen

Este trabajo contrasta un modelo teórico de asignación de cartera óptima que se ajusta a la evidencia empírica observada en la participación en mercados financieros de los hogares españoles con el ciclo de vida. En este modelo, la inversión en vivienda se trata de manera distinta a la inversión en activos con riesgo debido a las características de la misma. Por lo tanto, se considera la vivienda como una variable de estado en el problema de elección de la cartera óptima. Como resultado, la elección de la cartera óptima depende del valor de la vivienda en relación con la riqueza neta del hogar y la aversión al riesgo del mismo. Mediante regresión lineal múltiple y Tobit, los resultados obtenidos utilizando datos de la Encuesta Financiera de las Familias son compatibles con el modelo teórico propuesto por Flavin y Yamashita (2011).

Palabras clave: finanzas domésticas, comportamiento inversor, inversión inmobiliaria, ciclo de vida, regresión lineal múltiple, Tobit

Índice

1. Introducción	3
2. Revisión Bibliográfica	4
3. La Encuesta Financiera de las Familias (EFF)	6
4. Metodología	9
4.1. Modelo Teórico	9
4.2. Modelos Econométricos	11
4.2.1. Regresión Lineal Múltiple	12
4.2.2. Regresión Cuantílica	12
4.2.3. Tobit	13
4.3. Estimación con Datos de Encuestas Complejas	15
4.3.1. Pesos Muestrales	16
4.3.2. Imputación Múltiple	16
4.3.3. Pesos de Replicación	18
5. Análisis Descriptivo	18
5.1. Descripción de la Muestra	18
5.2. Activos Financieros y Ciclo Vital	23
6. Análisis Econométrico	35
6.1. Modelo de Regresión Lineal	36
6.2. Regresión Cuantílica	42
6.3. Tobit	45
7. Conclusiones	50
Referencias bibliográficas	52

1. Introducción

La inversión en vivienda constituye una pieza clave en la toma de decisiones financieras de los hogares españoles. Según datos recopilados en la Encuesta de Condiciones de Vida del Instituto Nacional de Estadística (INE) de 2017, el 76.7% de los hogares en España poseen su vivienda principal en propiedad, lo que resalta la importancia de este tipo de activo en la estructura financiera familiar.

Dada la significativa participación de la vivienda en las decisiones financieras y su rol como lugar de residencia, este trabajo tiene como objetivo principal analizar la relación entre la inversión en vivienda y la tenencia de activos arriesgados en el contexto del ciclo de vida familiar en España.

La necesidad de tal investigación se magnifica debido a las particularidades de la política de vivienda y las condiciones económicas en España, que podrían diferir de otros contextos estudiados en la literatura previa. Así, este trabajo se propone contrastar el modelo teórico desarrollado por Flavin y Yamashita (2011), utilizando datos de hogares españoles obtenidos de la Encuesta Financiera de las Familias elaborada por el Banco de España para los años 2002, 2005, 2008, 2011, 2014 y 2017.

La relevancia de este estudio radica en la alta incidencia de la propiedad de vivienda en el patrimonio de las familias españolas y en la evidencia empírica observada en la que tanto el peso del valor de la vivienda sobre la riqueza neta, como el peso de activos con riesgo en la cartera de inversión del hogar, tienen una relación cuadrática con la edad.

Contrario a los modelos estándar de asignación de carteras que no distinguen entre activos inmobiliarios y financieros, Flavin y Yamashita (2011) argumentan que estas dos clases de activos tienen propiedades distintas que afectan a las decisiones de inversión a lo largo del ciclo vital. Por este motivo, en la formulación teórica, en una primera etapa se decide la inversión en vivienda. Posteriormente, una vez seleccionada la vivienda, en una segunda etapa se determina la inversión en activos financieros.

Utilizando técnicas de regresión lineal múltiple y regresión Tobit para evaluar la relación entre la inversión en vivienda y en activos arriesgados, los resultados de este estudio muestra evidencia compatible con el modelo teórico propuesto por Flavin y Yamashita (2011). En particular, se comprueba que la proporción de la riqueza neta invertida en inmuebles es una variable significativa para explicar la proporción de activos arriesgados en la cartera financiera del hogar. Además, cuando se incluye esta proporción en los modelos econométricos propuestos, la edad pierde su capacidad explicativa, sugiriendo que actúa como una variable de confusión.

El resto del trabajo se estructura de la siguiente manera: la Sección 2 ofrece una revisión bibliográfica; la Sección 3 describe la fuente de datos y sus características; la Sección 4 expone la metodología del trabajo, incluyendo el modelo teórico y las técnicas económicas empleadas; la Sección 5 presenta un análisis descriptivo de los datos; los resultados se discuten en la Sección 6; y finalmente, la Sección 7 concluye con las implicaciones de nuestro estudio.

2. Revisión Bibliográfica

Las finanzas domésticas (Campbell, 2006) son un campo de la economía financiera que estudia cómo los hogares toman las decisiones financieras. Dentro de este campo de la economía financiera encontramos una gran cantidad de artículos de investigación. Gomes et al. (2021) hacen una revisión exhaustiva de los distintos temas que se han analizado en este campo de la economía financiera. Dentro de la revisión bibliográfica que realizan Gomes et al. (2021), la mayoría de artículos que analizan la participación en los mercados financieros obtienen evidencia de una baja participación por parte de los hogares más jóvenes, el aumento de la participación para aquellos hogares con mayor nivel de estudios/ingresos y una mayor participación por parte de los hombres frente a las mujeres. Sin embargo, puesto que es un campo tan amplio, centramos la revisión bibliográfica en aquellos trabajos que estudian la relación entre la inversión en vivienda y el resto de decisiones del hogar. Para ello, haremos un recorrido a través de los distintos artículos que estudian la vivienda, pasando por el que será nuestro artículo de referencia Flavin y Yamashita (2011).

En primer lugar, Curcuru et al. (2010) estudian cómo los *background risks* afectan a las decisiones de inversión. Así, los autores argumentan que los ingresos laborales, los ingresos empresariales, la inversión en activos inmobiliarios y los costes de transacción, entre otros, son una fuente de riesgo no diversificable ni asegurable (*background risks*) que afecta a las decisiones de inversión de los hogares. Además, las diferencias existentes entre las carteras de inversión de los distintos hogares también dependen de la edad o del momento de su ciclo de vida en que se encuentre el hogar. Los autores muestran que los hogares más jóvenes tienen menos activos financieros que hogares con cabeza de familia mayor para un mismo nivel de riqueza neta.

Cocco (2005) y Vestman (2012) contrastan modelos teóricos que endogenizan la inversión en vivienda, analizando así cómo evolucionan conjuntamente la vivienda y las carteras de inversión. De este modo, Cocco (2005) utilizando los datos del *Panel Study of Income Dynamics* (PSID) desde 1970 a 1992, concluye que, para los hogares estadounidenses, la inversión en vivienda explica en gran medida la tenencia de acciones. Así, la

inversión en vivienda que realizan los jóvenes mediante endeudamiento, justifica la baja participación en los mercados financieros. De manera similar, Vestman (2012) concluye que los propietarios de viviendas tienen el doble de probabilidad de invertir en activos arriesgados que los inquilinos. Para este último estudio, utilizan datos tanto para Estados Unidos como para Suecia.

En otro orden de ideas, Paz-Pardo (2021), observa que es menos probable que las generaciones más jóvenes sean propietarios de una vivienda de lo que lo eran generaciones anteriores a una misma edad, tanto para Estados Unidos, como para muchos países europeos (España, Italia, Reino Unido entre otros). Además, el autor utiliza un modelo de ciclo de vida con elección de cartera de inversión y vivienda para predecir que la tasa de propietarios entre los hogares más jóvenes, y la baja participación de los jóvenes en mercados financieros, se debe a la incertidumbre asociada a su participación en el mercado laboral.

Becker y Shabani (2010) y Fougere y Poulhes (2014) estudian la relación de la deuda hipotecaria sobre la elección de cartera de inversión para Estados Unidos y Francia, respectivamente. Becker y Shabani (2010) utilizan la encuesta “Survey of Consumer Finances” (SCF) para los años comprendidos entre 1989 y 2004. Los autores encuentran que los hogares con deuda hipotecaria tienen un 10% menos de probabilidad de poseer acciones, y un 37% menos de probabilidad de poseer bonos en comparación con hogares similares sin deuda hipotecaria. Fougere y Poulhes (2014) utilizando datos de la encuesta “Enquête Patrimoine” también llegan a la conclusión de que un aumento en las deudas hipotecarias reduce la inversión en acciones, aunque menos que en Estados Unidos. Además, concluyen que aumentos en el valor de la propiedad aumentan la participación en los mercados financieros, ya que el valor de la vivienda menos la deuda contraída para su compra (*home equity*) aumenta. Este efecto riqueza generado por el aumento de la vivienda afecta más a la participación en mercados financieros en Francia que en Estados Unidos.

Por otro lado, Flavin y Yamashita (2011) y Chetty et al. (2017) contrastan modelos teóricos donde la compra de vivienda se determina exógenamente. Ambos artículos, mediante modelos teóricos diferentes, analizan cómo afecta la inversión en vivienda a la cartera de inversión en activos arriesgados a lo largo del ciclo de vida.

Flavin y Yamashita (2011) proponen y contrastan un modelo teórico que se ajusta a la evidencia empírica sobre el comportamiento de los hogares en la toma de decisiones de inversión en activos financieros a lo largo del ciclo de vida. Apoyándose en el modelo teórico que diseñan, muestran que la vivienda no puede tratarse como una inversión cualquiera. Esto ocurre porque la vivienda juega el papel de variable de estado en la

modelización de las decisiones de inversión en activos con riesgo. De este modo, muestran cómo afecta la evolución en el peso del valor de los inmuebles sobre la riqueza neta a lo largo del ciclo de vida a la composición financiera de las carteras de inversión a lo largo del ciclo de vida de los hogares estadounidenses utilizando la encuesta *Survey of Consumer Finances*. Al mismo tiempo, documentan que el peso de los inmuebles sobre la riqueza neta disminuye con la edad. Por este motivo, cuando se incluye la variable edad, y se excluye el peso de los inmuebles sobre riqueza neta en los modelos econométricos a estimar, la edad actúa como *proxy* de la vivienda. En este caso, la edad muestra un *confounding effect* y no de una fuerte relación entre edad e inversión en activos con riesgo. Por el contrario, cuando se incluye la variable vinculada a la vivienda, esta muestra una relación negativa con la inversión en activos arriesgados, y la edad, por consiguiente, pierde capacidad explicativa.

Chetty et al. (2017) concluyen que la vivienda es un determinante importante de la inversión en activos arriesgados para la población estadounidense utilizando la encuesta *Survey of Income and Program Participation*. Así, los aumentos en el valor de la propiedad (manteniendo constante el valor neto de la vivienda) reducen la participación en mercados de acciones, mientras que los aumentos en la riqueza sobre el valor neto de la vivienda (manteniendo constante el valor de la propiedad) aumentan la participación.

3. La Encuesta Financiera de las Familias (EFF)

Para este trabajo hemos utilizado la Encuesta Financiera de las Familias (EFF), que elabora oficialmente el Banco de España y que está incluida en el Plan Estadístico Nacional. La encuesta proporciona información detallada sobre las rentas, los gastos, las deudas y los activos de los hogares españoles mediante datos de corte transversal y datos de panel. Esta encuesta se realiza trienalmente desde el año 2002 y su diseño se inspira en la encuesta *Indagine sui bilanci delle famiglie* (IBF) elaborada por la Banca d'Italia desde 1989 y sobre todo, en la *Survey of Consumer Finances* (SCF) elaborada desde 1983 por la Reserva Federal de Estados Unidos. Al mismo tiempo, la EFF forma parte del proyecto europeo HFCS (*Household Finance and Consumption Survey*) desde el 2010. Este proyecto utiliza un cuestionario común y obtiene encuestas armonizadas en todos los países de la zona euro. En concreto, desde el año 2011 existen tres olas disponibles de la encuesta europea armonizada. Además, existen otras encuestas financieras de los hogares ajenas al proyecto HFSC, como son las elaboradas por China, Corea o Chile, entre otras muchas.

El diseño de la EFF persigue tanto el objetivo de crear una muestra que represente la población total española, como crear una base de datos con una componente de panel.

Para la elaboración de la EFF el Banco de España recibe la colaboración del INE y de la Agencia Tributaria.

Una de las particularidades de la EFF es que realiza un sobremuestreo de hogares con mayores niveles de riqueza. A diferencia del muestreo aleatorio simple, el sobremuestreo es una herramienta útil para obtener un número suficiente de observaciones en segmentos de la población que tienen una baja frecuencia en la población total. La proporción de hogares que poseen algún tipo de activo financiero es relativamente pequeña en comparación con los hogares que no los poseen. Por lo tanto, el objetivo de este sobremuestreo en la EFF es poder analizar el comportamiento inversor de los hogares. Para ello, se requiere de una muestra lo suficientemente grande de hogares que invierten en diferentes activos financieros. Puesto que los hogares ricos participan más en mercados financieros, es necesario el sobremuestreo para los hogares con mayor nivel de riqueza. En ausencia de sobremuestreo, no sería posible analizar el comportamiento inversor de los hogares, ya que la mayor parte de los hogares no poseen activos de este tipo.

Para el sobremuestreo, en primer lugar, se definen 7 estratos de riqueza basándose en el impuesto sobre el patrimonio. Los estratos superiores están asociados a hogares con mayor impuesto sobre el patrimonio y son menos frecuentes en la población española. De esta manera, los dos estratos iniciales abarcan alrededor del 33 % de la distribución de la riqueza imponible. Los siguientes tres estratos abarcan desde el percentil 30 hasta el percentil 95, mientras que los dos estratos finales abarcan los últimos cinco percentiles casi por completo. Con los estratos definidos, el sobremuestreo consiste en seleccionar hogares a una tasa creciente en dichos estratos.

En la [Tabla 1](#) tomada de Crespo (2021) se muestra el número de observaciones y la tasa de sobre-representación presente en las olas de 2014 y 2017 para cinco percentiles de riqueza. Observamos que el percentil más alto de distribución de riqueza está representado con 777 hogares de los 6413 que participaron en la encuesta de 2017. Cabe mencionar que, en ausencia de sobremuestreo de hogares con alto nivel de riqueza, solamente se contaría con 64 hogares en este percentil de riqueza. Por otro lado, la EFF dispone únicamente de 2074 observaciones por debajo de la riqueza neta mediana, mientras que, en ausencia de sesgo en la selección de la muestra, este grupo contaría con 3191 de las 6413 observaciones totales para el año 2017.

Tabla 1: Sobre-representación de los hogares ricos en las muestras finales de la EFF2014 y la EFF2017

Decilas riqueza neta	EFF 2014		EFF 2017	
	Número de observaciones	Tasa de sobre representación	Número de observaciones	Tasa de sobre representación
50 % inferior	1981	0.65	2074	0.65
Entre 50 % y 90 %	2097	0.86	2265	0.88
Entre 90 % y 95 %	469	1.53	433	1.35
Entre 95 % y 99 %	813	3.32	864	3.37
1 % superior	741	12.11	777	12.12

Tasa de sobre-representación de los hogares encuestados en las encuestas de 2014 y 2017 por nivel de riqueza. Fuente: Crespo (2021)

La EFF posee otras dos peculiaridades. La primera de ellas es que dispone de una componente longitudinal de los datos. De este modo, además de entrevistar a nuevas familias, también se entrevista a hogares que han participado en ediciones anteriores de la encuesta. Aproximadamente un 50 % de los encuestados en cada ola participaron en la ola anterior. Por ejemplo, de las 6413 entrevistas válidas de 2017, 3634 corresponden a hogares que ya colaboraron en la EFF de 2014. La segunda peculiaridad es que la EFF es muy extensa y presenta preguntas sobre hábitos de consumo y salud financiera. Esto conlleva que la tasa de no respuesta no sea despreciable, y se requiera de un gran trabajo de campo en la recogida de datos.

Así, muchas encuestas no se aceptan como válidas cuando excedan el número de preguntas no respondidas establecido. Concretamente, la tasa de respuesta total en la EFF para el año 2017 fue del 44.97 % de los encuestados. No obstante, fijándonos en la tasa de respuesta de los diferentes estratos, la mayor tasa de respuesta la encontramos en el último estrato con un 52.63 %. Por el contrario, la menor tasa de respuesta es la correspondiente al segundo estrato, con únicamente un 28.43 %. Además, los hogares que han contestado a encuestas en años anteriores son los que mayor tasa de respuesta presentan: 70.48 % vs 50.52 %

Por último, cabe mencionar que el Banco de España publica junto con las bases de datos de la encuesta, guías para el uso de los datos de la EFF, la metodología y los principales resultados obtenidos para cada ola (Banco de España, 2017; Banco de España, 2019).

4. Metodología

En este trabajo contrastaremos el modelo teórico propuesto por Flavin y Yamashita (2011) para los hogares españoles usando datos de la Encuesta Financiera de las Familias elaborada por el Banco de España. En esta sección, en primer lugar explicaremos brevemente el modelo propuesto por Flavin y Yamashita (2011). Una vez presentado el modelo teórico a contrastar, expondremos las técnicas de análisis de datos empleadas para obtener los resultados y conclusiones del estudio. Entre las técnicas utilizadas comenzamos exponiendo los modelos econométricos implementados para contrastar el modelo teórico de Flavin y Yamashita (2011). Por último, dado que los datos de la Encuesta Financiera de las Familias que utilizamos se enmarcan dentro del grupo de “encuestas complejas”, mostraremos las peculiaridades y la metodología a seguir en el tratamiento de los datos procedentes de este tipo de encuestas.

4.1. Modelo Teórico

El modelo teórico propuesto por Flavin y Yamashita (2011) es una versión simplificada del modelo propuesto por Flavin y Nakagawa (2008).

Así, Flavin y Yamashita (2011) proponen un modelo que incluye supuestos sobre la inversión en vivienda que hace de esta un activo diferente a los financieros. En este modelo teórico, la cartera óptima de activos con riesgo dependerá de la inversión en vivienda que se decide previamente. Es decir, la inversión en vivienda es una variable de estado del modelo de selección de cartera que se ve afectada por el grado de aversión al riesgo, pero no por la edad.

De este modo, Flavin y Yamashita (2011) suponen en su modelo que, además de la inversión en vivienda, los individuos también pueden destinar su dinero a la inversión de activos financieros arriesgados. Estos activos financieros se pueden comprar y vender sin costes de transacción, a diferencia de lo que ocurre con la vivienda. A la hora de realizar una inversión en vivienda, se compra un activo real con unas características físicas concretas. Por ello, si se quiere una vivienda con características físicas diferentes a la actual, la única alternativa posible es vender el inmueble actual para comprar otro que se ajuste a las nuevas necesidades. No obstante, al realizar este cambio se incurre en un coste de ajuste que es equivalente al coste de transacción.

De esta manera, el hogar elige la vivienda, el consumo no duradero, y la inversión en activos financieros que maximizan su utilidad. Los autores asumen que la riqueza únicamente se mantiene en forma de activos financieros y vivienda. Así, la riqueza del hogar está dada por:

$$W_t = P_t H_t + \mathbf{X}_t \mathbf{1}$$

donde P_t es el precio de la vivienda por metro cuadrado en el mercado actual, H_t son los metros cuadrados de la vivienda, \mathbf{X}_t es el vector $(1 \times n)$ de cantidades de activos arriesgados, donde el primer elemento del vector corresponde al activo hipotecario, y por último, $\mathbf{1}$ es un vector $(n \times 1)$ de unos.

Se establecen las restricciones de aval y de no negatividad. La restricción de aval afecta a los préstamos hipotecarios utilizando el primer elemento de \mathbf{X}_t . La restricción de no negatividad condiciona la inversión en activos financieros distintos de la deuda hipotecaria;

1. $-P_t H_t \leq X_{1t} \leq 0$

$X_{1t} \leq W_t - P_t H_t$. La restricción de aval postula que un hogar en el que el valor de la vivienda exceda la riqueza neta, puede tener un préstamo hipotecario mínimo de $(P_t H_t - W_t)$ y máximo de $(P_t H_t)$, ya que la cuantía de hipoteca depende del valor de la vivienda $(P_t H_t)$. En resumen, la restricción de aval establece que los hogares únicamente puede contraer deuda hipotecaria para su inversión en vivienda, por una cuantía que no exceda el precio de la vivienda a adquirir. Por el contrario, los hogares no podrán contraer deuda para la compra de activos distintos de la vivienda.

2. $0 \leq X_{it} \quad i = 2 \dots n$. Restricción de no negatividad de los activos financieros, esto es, los hogares no pueden hacer ventas de activos financieros en corto.

En cuanto al problema de optimización, es una decisión en dos etapas. Primero, el hogar decide si comprar una nueva casa o no hacerlo. En caso de decidir comprar una nueva vivienda, debe elegir las características físicas de la misma. En una segunda etapa, el hogar decide el nivel de consumo no duradero (C_t) y la proporción de riqueza invertida en activos con riesgo (\mathbf{x}_t). Reescribimos el valor de la vivienda como proporción de la riqueza mediante el escalar $h_t = \frac{P_t H_t}{W_t}$. De la misma manera, mediante el vector \mathbf{x}_t de dimensión $(n \times 1)$ denotamos la tenencia de activos financieros como proporción de la riqueza, es decir, $\mathbf{x}_t = \frac{\mathbf{X}_t}{W_t}$.

En el momento en el que el hogar decide la cartera de activos financieros (\mathbf{x}_t), h_t es una variable de estado, ya que la ha decidido en la etapa anterior. Por lo tanto, mediante la resolución del siguiente problema de maximización, el hogar elige el vector de activos financieros \mathbf{x}_t :

$$\begin{aligned} \max_{\mathbf{x}_t} & \left\{ h_t \mu_H + \mathbf{x}_t \boldsymbol{\mu} - \frac{1}{2} A_t (h_t^2 \sigma_H^2 + \mathbf{x}_t \boldsymbol{\Sigma} \mathbf{x}_t^T) \right\} \\ \text{s.a.} & \quad 1 = h_t + \mathbf{x}_t \mathbf{1}, \\ & \quad -P_t H_t \leq X_{1t} \leq 0, \\ & \quad X_{1t} \leq W_t - P_t H_t, \\ & \quad 0 \leq X_{it} \quad i = 2 \cdots n \end{aligned}$$

donde μ_H denota el rendimiento esperado de la vivienda; $\boldsymbol{\mu}$, el vector de n elementos, el cual contiene las rentabilidades esperadas de los activos financieros; σ_H^2 , la varianza de la innovación de los precios reales de la vivienda; $\boldsymbol{\Sigma}$, la matriz de covarianzas ($n \times n$) de las innovaciones de los rendimientos de los activos financieros, y por último, A_t hace referencia al grado de aversión al riesgo del hogar:

$$A_t = -\frac{\frac{\partial^2 V}{\partial W_t^2} W_t}{\frac{\partial V}{\partial W_t}} \geq 0$$

donde V es la función objetivo:

$$V(H_T, W_T) = E \left[\int_t^T e^{-\sigma s} U(H_s, C_s) ds + e^{-\sigma(T-t)} V(H_T, W_T) \right]$$

De este modo, el modelo teórico propuesto por Flavin y Yamashita caracteriza la cartera óptima de activos con riesgo en función de una variable que refleja el peso del valor actual de la vivienda sobre la riqueza neta del hogar (variable de estado de la vivienda). El modelo predice que la cartera óptima del hogar dependerá del grado de aversión al riesgo y de la variable de estado de la vivienda, pero no de la edad del hogar.

4.2. Modelos Econométricos

Para este estudio utilizaremos tres modelos econométricos, como son: regresión lineal múltiple, regresión cuantílica y Tobit. En esta sección describiremos brevemente cada uno de los tres métodos. Exposiciones detalladas de ellos pueden encontrarse en Long (1997), Deaton (1997), Koenker y Machado, (1999), Koenker y Hallock (2001), Cameron y Trivedi (2005), Winkelmann (2008), Wooldridge (2010), Stock y Watson (2012) y Wooldridge (2015).

4.2.1. Regresión Lineal Múltiple

El modelo de regresión lineal múltiple, que fue documentado por primera vez en 1805 por el matemático Adrien-Marie Legendre, modeliza la relación entre una variable de interés Y (variable respuesta, dependiente o endógena), y un conjunto de K variables explicativas, también llamadas, independientes o exógenas, X_1, X_2, \dots, X_K . Como vemos en la siguiente ecuación, en este modelo la función de regresión que relaciona la variable dependiente con las variables independientes es lineal:

$$Y = X\beta + u$$

donde Y es el vector de respuestas de dimensión $n \times 1$, X la matriz de regresores de dimensión $n \times (K + 1)$, β el vector $(K + 1) \times 1$ de parámetros a estimar y u es un vector de perturbaciones $n \times 1$. En este contexto, n representa el número de observaciones en el conjunto de datos.

El método de Mínimos Cuadrados Ordinarios (MCO) se utiliza para estimar los coeficientes β_k del modelo de regresión múltiple que minimicen la suma del cuadrado de los residuos, es decir, las estimaciones de los $K + 1$ coeficientes se obtienen a partir del siguiente problema de minimización:

$$\min_{\beta} u'u = \min_{\beta} (Y - X\beta)'(Y - X\beta)$$

Debemos mencionar que bajo el Teorema de Gauss-Markov, si se cumplen todas las hipótesis básicas del modelo de regresión lineal (supuestos Gauss-Markov), entonces el estimador MCO es un estimador eficiente. Esto es, dentro de los estimadores lineales e insesgados, es el de menor varianza.

4.2.2. Regresión Cuantílica

La regresión cuantílica (Koenker y Bassett, 1978) es un método de estimación semi-paramétrico que resulta útil cuando nos encontramos con heterocedasticidad o presencia de datos atípicos.

Así como la regresión por MCO se encuentra asociada con la media, y hace la estimación en torno a este estadístico, la regresión cuantílica modeliza los cuantiles de la variable dependiente como una función lineal de los regresores. Podemos definir el concepto de cuantil mediante la siguiente expresión:

$$\min_{b \in R} \left[\sum_{Y_i \geq b} \tau |Y_i - b| + \sum_{Y_i < b} (1 - \tau) |Y_i - b| \right]$$

donde $\tau \in (0, 1)$ es el cuantil, Y_i los distintos valores para la variable Y , y b el valor que minimiza la expresión. El cuantil τ será aquel valor b que deje una proporción τ de observaciones por debajo de b y una proporción $(1 - \tau)$ por encima.

En la regresión cuantílica, el cuantil condicional τ viene dado por una función lineal de las variables explicativas, es decir:

$$Q_\tau(Y|X) = X\beta_\tau$$

donde Y es la variable dependiente, X es la matriz de variables independientes, y β_τ es el vector de parámetros a estimar correspondiente al cuantil τ .

Partiendo de la expresión que hemos utilizado para definir el cuantil podemos expresar el problema de estimación de los parámetros de la regresión cuantílica de la siguiente manera:

$$\text{Min}_{\beta_\tau \in R} \left[\sum_{Y_i \geq X_i \beta_\tau} \tau |Y_i - X_i \beta_\tau| + \sum_{Y_i < X_i \beta_\tau} (1 - \tau) |Y_i - X_i \beta_\tau| \right]$$

Es decir, los coeficientes estimados con este método minimizan las desviaciones absolutas ponderadas con pesos asimétricos. En este caso, se penalizan los errores de forma lineal, por lo que esta estimación no se ve alterada por valores extremos. Además, en el caso particular de tener unos pesos simétricos ($\tau = 0,5$) estamos ante la llamada regresión mediana, cuyo objetivo es minimizar la suma de las desviaciones en términos absolutos sin ponderar.

4.2.3. Tobit

La regresión Tobit es un modelo econométrico para datos censurados propuesto por Tobin en 1958 (Tobin, 1958). En ocasiones nos encontramos con que la variable dependiente de un modelo de regresión lineal tiene unas características determinadas que condicionan el tipo del modelo de regresión.

A continuación definiremos en qué consisten estos tres tipos de variables y describiremos la aplicación del modelo Tobit para el caso de una variable dependiente con solución esquina. En primer lugar, una variable censurada es una variable para la que existe un determinado límite, ya sea superior y/o inferior, a partir del cual a todas las observaciones se les asigna un mismo valor. En segundo lugar, una variable truncada se da cuando en una población no existen observaciones a partir de un determinado límite, ya sea superior y/o inferior. Finalmente, una variable con solución esquina es una variable que toma el valor cero para una fracción importante de la población, mientras que el resto de

observaciones presentan una distribución continua entre valores positivos.

El modelo Tobit para una variable dependiente con solución esquina se expresa mediante las siguientes ecuaciones:

$$Y^* = X\beta + u, \quad u|X \sim Normal(0, \sigma^2) \quad (1)$$

$$Y = \max(0, Y^*) \quad (2)$$

Es decir, el modelo Tobit expresa la variable observada Y en términos de una variable latente Y^* que es lineal en los regresores con un error que presenta una distribución normal y es homocedástico. Adicionalmente, podemos ver en la ecuación (2) que la variable observable es igual a la latente en el caso de que la variable latente sea mayor o igual a cero y en el resto de casos, es igual a cero.

Además, puesto que Y^* tiene una distribución normal, Y presenta una distribución continua únicamente a través de valores positivos. Concretamente, la densidad de Y dada X coincide con la densidad de Y^* dada X para valores positivos:

$$\begin{aligned} P(Y = 0|X) &= P(Y^* < 0|X) = P(u < -X\beta|X) = \\ &= P\left(\frac{u}{\sigma} < -\frac{X\beta}{\sigma}|X\right) = \Phi\left(\frac{-X\beta}{\sigma}\right) = 1 - \Phi\left(\frac{X\beta}{\sigma}\right) \end{aligned} \quad (3)$$

Asimismo, si (X_i, Y_i) se extrae aleatoriamente, obtenemos que la densidad de Y_i dada X_i es:

$$(2\pi\sigma^2)^{-\frac{1}{2}} \exp\left(-\frac{(Y - X_i\beta)^2}{2\sigma^2}\right) = \frac{1}{\sigma} \phi\left(\frac{(Y - X_i\beta)}{\sigma}\right), Y > 0 \quad (4)$$

Los parámetros a estimar de la regresión Tobit se obtienen maximizando el logaritmo de la función de verosimilitud. Así, el logaritmo de la función de verosimilitud para cada observación i viene dada por la siguiente expresión:

$$L_i(\beta, \sigma) = 1(Y_i = 0) \log\left(1 - \Phi\left(\frac{X_i\beta}{\sigma}\right)\right) + 1(Y_i > 0) \log\left(\frac{1}{\sigma} \phi\left(\frac{Y_i - X_i\beta}{\sigma}\right)\right) \quad (5)$$

donde ϕ y Φ son la función de densidad y de distribución normal estándar respectivamente. Y cada término de la verosimilitud corresponde a un tipo de valor que puede tomar la variable dependiente observada. $Y_i = 0$ correspondiente a la ecuación (3) e $Y_i > 0$ corresponde a la ecuación (4) en la función de verosimilitud.

Podemos observar que la ecuación (5) depende tanto de la desviación estándar de u (σ) como del vector β . Por lo tanto, las estimaciones de máxima verosimilitud de β y σ se obtienen al maximizar la función de verosimilitud (5), es decir:

$$\max_{\beta, \sigma} \sum_{i=1}^n L_i(\beta, \sigma)$$

En los modelos Tobit, a diferencia de la regresión lineal múltiple, los efectos marginales no coinciden con los coeficientes estimados. Definimos los efectos marginales como la variación experimentada por la variable dependiente cuando una variable independiente se incrementa marginalmente manteniendo el resto de variables constantes.

De esta manera, en el caso del modelo Tobit, si el regresor es una variable continua, el efecto marginal será:

$$\frac{\partial E(Y|X)}{\partial X_k} = \Phi\left(\frac{X\beta}{\sigma}\right) \beta_k$$

Por otro lado, si el regresor es una variable categórica, el efecto marginal se obtiene de manera diferente a las variables continuas. A modo de ejemplo, si la variable independiente es dicotómica, el efecto marginal se obtiene como la diferencia entre $E(Y|Y > 0, X)$ con $X_1 = 1$ y $X_1 = 0$.

No obstante, debemos mencionar que si únicamente nos interesase el signo, esto es, la manera en que afecta una variable explicativa a Y , no su magnitud, bastaría con mirar el signo del coeficiente estimado, ya que este será común al arrojado por el efecto marginal.

4.3. Estimación con Datos de Encuestas Complejas

Cuando se dispone de una “encuesta compleja”, como lo es la EFF, el análisis estadístico convencional deja de ser apropiado. En los próximos subapartados explicaremos las técnicas estadísticas empleadas en el análisis de este tipo de datos, considerando la presencia de pesos muestrales, imputaciones múltiples y pesos de replicación. Explicaciones detalladas sobre estas técnicas pueden encontrarse en Rubin (1987), Deaton (1997) y Barceló (2006). Además, Lumley (2011) muestra las herramientas disponibles en el *software* estadístico “R” para implementar las técnicas estadísticas necesarias en encuestas complejas.

4.3.1. Pesos Muestrales

En primer lugar, si la frecuencia de algún estrato en la población es muy baja, puede ocurrir que con muestreo aleatorio no se recoja ni una sola observación para dicho estrato. Por este motivo, la muestra de las encuestas complejas tienen un número de observaciones para algunos estratos que no representan su frecuencia en la población total.

De este modo, para obtener magnitudes representativas de la población a partir de una muestra que tiene un sesgo en la selección de encuestados (no aleatorio), se utilizan los llamados pesos muestrales. Los pesos muestrales reflejan la representatividad de cada observación en la población total. Así, estos pesos miden la frecuencia con la que cada observación encuestada se encuentra en la población en la que se realiza el muestreo.

La entidad que elabora la encuesta es quien proporciona los pesos muestrales. Para el cálculo de los pesos muestrales se utiliza información sobre las características de la población en la que se realiza la encuesta, y se realizan ajustes para recoger la tasa de no respuesta para los diferentes estratos. Los pesos muestrales incorporan la información necesaria para obtener estimaciones poblacionales a partir de la muestra. De este modo, para obtener resultados poblacionales, los pesos muestrales deben ser aplicados en las estimaciones que se deseen obtener.

A modo de ejemplo, se obtiene un estimador consistente e insesgado de la media poblacional a partir de una muestra con sesgo en la selección, incorporando pesos muestrales para una variable x cualquiera del siguiente modo:

$$\bar{x}_w = \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i}$$

donde \bar{x}_w es la media poblacional de la variable de interés que incluye pesos muestrales, n refleja el tamaño muestral, w_i la inversa de la probabilidad de cada observación de ser incluida en la muestra (pesos muestrales) y x_i los valores que toma la variable de interés para cada observación i .

Más detalles sobre el uso de los pesos muestrales para el cálculo de diversos estadísticos pueden encontrarse en los capítulos 1 y 2 de Deaton (1997) y en Solon et al. (2015)

4.3.2. Imputación Múltiple

En todas las encuestas existe ausencia de respuesta a ciertas preguntas. No obstante, en encuestas que proporcionan información sobre las rentas, los activos, las deudas y los gastos de los hogares, la ausencia de respuestas todavía es más habitual.

Las imputaciones múltiples son una técnica útil para tratar las muestras que contienen

datos faltantes o *missing values* en inglés. Algunos procedimientos estadísticos optan por eliminar las no-respuesta de la base datos. Sin embargo, además de que eliminar toda observación que contenga una no-respuesta puede llevar a un sesgo en los resultados, diseñar y crear las bases de datos es muy costoso como para eliminar toda observación que contenga una no-respuesta.

Una posible solución a este problema es el método de imputación múltiple. El método de imputación múltiple genera tantas imputaciones como se quieran y en cada una de ellas sustituye los datos faltantes por valores estadísticamente factibles. Así, en cada imputación se sustituirán los *missing values* por un valor distinto.

De cara a obtener estimaciones de parámetros estadísticos con imputaciones, tendremos tantas estimaciones como imputaciones tengamos. Sin embargo, el objetivo es tener un único valor estimado. Por ello, debemos obtener a través de las m imputaciones y, por tanto, de las m estimaciones un único resultado final. Rubin (1987) propone las fórmulas a implementar para llegar a dicho objetivo;

Tendremos un estadístico θ_i calculado para cada imputación y debemos calcular la media aritmética de todos las estimaciones θ_i :

$$\bar{\theta} = \frac{1}{m} \sum_{i=1}^m \theta_i$$

Sin embargo, el cálculo de la desviación estándar requiere de más pasos intermedios. En primer lugar, se calcula:

$$V_W = \frac{1}{m} \sum_{i=1}^m SE_i^2$$

donde SE_i es el error estándar de la estimación para cada imputación i . A continuación se calcula:

$$V_B = \frac{1}{m-1} \sum_{i=1}^m (\theta_i - \bar{\theta})^2$$

Así, mediante V_B , que refleja la varianza entre las imputaciones, podemos calcular la varianza conjunta del siguiente modo:

$$V_{\bar{p}} = V_W + V_B \left(1 + \frac{1}{m} \right)$$

4.3.3. Pesos de Replicación

Si queremos estimar correctamente los errores estándar, por lo general requerimos de una descripción detallada del método de muestreo. Sin embargo, la información de la estratificación no está disponible para proteger el anonimato de los encuestados.

Por este motivo, para estimar el error estándar utilizando datos de encuestas complejas se crean los pesos de replicación (*replicate weights*). Estos son un conjunto de pesos muestrales que contienen la información necesaria para calcular los errores estándar análogos a los obtenidos con el método *bootstrap*, pero manteniendo el anonimato de los encuestados.

5. Análisis Descriptivo

5.1. Descripción de la Muestra

En esta sección describiremos los datos de la Encuesta Financiera de las Familias que vamos a utilizar para este estudio. En este trabajo analizamos los microdatos para seis olas: la ola de 2002, 2005, 2008, 2011, 2014 y 2017, en las que tenemos 5143, 5962, 6197, 6106, 6120 y 6413 observaciones respectivamente. Debemos mencionar que los estadísticos que se muestran en este apartado se han calculado usando los pesos muestrales que hemos explicado en el apartado 4.3.

Puesto que el objetivo del trabajo es comprobar las hipótesis planteadas en el modelo teórico propuesto por Flavin y Yamashita (2011) para los hogares españoles, únicamente seleccionamos aquellos hogares que poseen riqueza neta no negativa. Además, eliminamos de la muestra aquellos hogares donde el cabeza de familia¹ tiene menos de 24 años o más de 89. De este modo, para el año 2002 seleccionamos 4977 observaciones de las 5143, 5672 para la ola de 2005, 5925, 5846, 5779 y 6014 para las siguientes olas cronológicamente.

De la EFF hemos extraído las variables de las características de los hogares que pueden determinar la inversión de los hogares según el modelo teórico a contrastar. Para seleccionar el conjunto de variables nos hemos guiado principalmente por la selección realizada por Flavin y Yamashita (2011).

En la selección de las variables realizada, podemos distinguir entre características socio-demográficas, aquellas que nos aportan información sobre cuál es la situación y los

¹La definición de cabeza de familia es la utilizada por el Banco de España. Cuando el hogar esté formado, al menos, por una pareja compuesta por un hombre y una mujer, el hombre será el cabeza de familia. En el resto de casos, siempre que quien responda a la encuesta no sea uno de los hijos del hogar, esta será el cabeza de familia. Por último, en caso de que uno de los hijos del hogar sea quien contestó a la encuesta, el cabeza de familia será el progenitor en caso de haberlo, o la progenitora en caso de no existir progenitor en el hogar

atributos de un hogar respecto al entorno en el que vive. Y variables económicas, estas últimas nos muestran cuáles son las capacidades económicas de los hogares, así como las variables relacionadas con las inversiones financieras o inmobiliarias del hogar.

A continuación, exponemos tanto las variables seleccionadas en nuestro análisis, como los estadísticos descriptivos para las características socio-demográficas (Tabla 2) y para las variables económicas (Tabla 3).

Dentro del grupo que denominamos como características socio-demográficas, tenemos la edad, el sexo, el estado civil, el número de hijos que forman parte del hogar en el momento de contestar a la encuesta, la situación laboral del cabeza de familia, la situación laboral del cónyuge del cabeza de familia, los estudios del cabeza de familia y los estudios del cónyuge.

En primer lugar, de la Tabla 2 destacamos que la edad media y mediana de los hogares encuestados ha aumentado en los últimos años. Mientras que la edad media del cabeza de familia en 2002 es de 52.98 años, en 2017 es de 56.68 años.

En segundo lugar, aunque condicionado por la manera que hemos designado al cabeza de familia, aproximadamente el 75 % de los hogares tienen un cabeza de familia hombre para cualquier ola.

Respecto al estado civil, se observa que la mayoría de los hogares están formados por individuos casados. Sin embargo, en las últimas olas, aunque siguen siendo mayoritarios los hogares casados, se reducen respecto a 2002. En consecuencia, aumentan los hogares con cabeza de familia divorciado.

El número de hijos medio que tienen los hogares en cada ola ha disminuido sensiblemente de la primera encuesta a la última. Concretamente, los hogares en 2017 tenían en media un 25 % menos de hijos de los que tenían en 2002.

En cuanto a la situación laboral del cabeza de familia, trabajar por cuenta ajena es la situación más frecuente para todas las olas. Por otro lado, la situación laboral del cónyuge del cabeza de familia más frecuente en las encuestas de 2002 y 2005 es la inactividad. Sin embargo, a partir de 2008 trabajar por cuenta ajena se convierte en la categoría más frecuente para los cónyuges. A pesar de que en 2008 existe prácticamente el mismo porcentaje de cónyuges inactivos que trabajando por cuenta ajena, a partir de este año trabajar por cuenta ajena va cobrando peso. Hasta tal punto es así que, en 2017, había un 14 % más de cónyuges trabajando por cuenta ajena que inactivos. Mientras que en 2002 había un 16 % más de inactivos que trabajadores por cuenta ajena en este grupo.

También destacamos que el nivel de estudios tanto del cabeza de familia, como del

cónyuge, ha aumentado a lo largo de los años. Concretamente, mientras que únicamente el 15.5 % de los hogares tenían un cabeza de familia con estudios universitarios en 2002, en 2017 el 22.5 % de los hogares tenían un cabeza de familia universitario. De igual modo, el 23.1 % de los cónyuges del cabeza de familia eran universitario en 2017, mientras que en 2002, solamente el 14 % de los cónyuges eran universitarios.

Tabla 2: Estadísticos descriptivos de las características socio-demográficas

Ola		2002	2005	2008	2011	2014	2017
Edad	Media	52.98	53.23	54.17	55.34	56.22	56.68
	Mediana	52	52	53	54	55	55
	Desv. Típica	0.336	0.38	0.38	0.41	0.24	0.17
	Mínimo/Máximo	24/89	24/89	24/85	24/86	24/87	24/86
Sexo	Hombre	77.9 %	75.5 %	74.4 %	73.8 %	72.8 %	73 %
	Mujer	22.1 %	21.5 %	25.6 %	26.2 %	27.2 %	27 %
Estado Civil	Casado	67 %	64.7 %	61.6 %	60 %	57 %	55 %
	Soltero	12.87 %	13.9 %	14.2 %	14.2 %	16.8 %	19.5 %
	Divorciado	5.7 %	6.2 %	7.2 %	9 %	10.6 %	12 %
	Viudo	14.3 %	15.2 %	16.9 %	16.9 %	15.6	13.5 %
Hijos en el Hogar	Media	0.97	0.89	0.86	0.81	0.74	0.73
	Mediana	1	1	1	0	0	0
	Desv. Típica	0.02	0.03	0.03	0.03	0.10	0.01
	Mínimo/Máximo	0/7	0/7	0/7	0/7	0/7	0/7
Situación Laboral Cabeza de Familia	Trabajador Cuenta Ajena	44.4 %	44.7 %	42.1 %	37.2 %	35.6 %	38.3 %
	Jubilado	25.5 %	26.4 %	25.4 %	31.0 %	30.2 %	29.8 %
	Inactivo	14 %	15 %	16 %	10.9 %	13 %	12.7 %
	Trabajador Cuenta Propia	10.4 %	9 %	8.7 %	8.9 %	8.4 %	9.9 %
	Desempleado	5.7 %	5 %	7.8 %	11.2 %	13 %	9.4 %
Situación Laboral Cónyuge	Inactivo	46.1 %	38.2 %	35.2 %	26.1 %	25.2 %	23.4 %
	Trabajador Cuenta Ajena	29.9 %	35.6 %	35.8 %	34.1 %	34.3 %	37.1 %
	Jubilado	11.8 %	14.0 %	13.9 %	20.0 %	19.2 %	19.3 %
	Trabajador Cuenta Propia	6.3 %	6.1 %	7.2 %	6.4 %	6.5 %	7.6 %
	Desempleado	5.8 %	6.1 %	8 %	13.3 %	14.8 %	12.5 %
Nivel de Estudios Cabeza de Familia	Menor que Bachillerato	58.6 %	58 %	58.6 %	57.8 %	54.8 %	52.3 %
	Bachillerato	25.8 %	24.4 %	24.3 %	23.1 %	22.4 %	25.2 %
	Universitarios	15.5 %	17.6 %	17.1 %	19.1 %	22.8 %	22.5 %
Nivel de Estudios Cónyuge	Menor que Bachillerato	61.1 %	59.2 %	59.3 %	57.8 %	55.1 %	52.2 %
	Bachillerato	24.9 %	23.9 %	23.4 %	23.8 %	21.4 %	24.7 %
	Universitarios	14 %	16.9 %	17.4 %	18.4 %	23.5 %	23.1 %
Número de Observaciones		4977	5672	5925	5846	5779	6014

Estadísticos descriptivos de las variables socio-demográficas de interés de la EFF de 2002, 2005, 2008, 2011, 2014 y 2017. Todos los estadísticos que se muestran se han obtenido haciendo uso de los pesos muestrales, los pesos de replicación y las 5 imputaciones para cada ola.

Describimos a continuación las variables económicas seleccionadas:

En primer lugar, el “peso de activos con riesgo en la cartera de inversión” es una variable continua que refleja el peso que tiene la suma del valor de las acciones cotizadas, las acciones no cotizadas y los fondos de inversión que posee cada hogar sobre su riqueza financiera total. Definimos la riqueza financiera total como la suma del saldo de las cuentas que dispone el hogar para realizar pagos, el saldo en las cuentas con las que no puede realizar pagos, el valor de las acciones cotizadas, las acciones no cotizadas, los activos de renta fija, las carteras gestionadas, los fondos de inversión, los planes de pensiones, los seguros de vida, así como el valor de lo que se le debe al hogar.

Para la creación de la variable “peso de activos con riesgo en la cartera de inversión” Flavin y Yamashita (2011) únicamente contemplan la proporción de acciones cotizadas sobre la riqueza financiera total. Nosotros, dada la idiosincrasia de la población española y su menor inclinación hacia la inversión en acciones cotizadas en comparación con la población estadounidense, consideramos necesario incluir tanto las acciones no cotizadas como los fondos de inversión al crear la variable “peso de activos con riesgo en la cartera de inversión”. Esto se debe a que, en el año 2017, solamente un 12.8% de los hogares españoles seleccionados tenía acciones cotizadas, un 2.7% tenía acciones no cotizadas y un 7.8% poseía fondos de inversión. A su vez, el porcentaje de hogares que poseían acciones cotizadas y/o acciones no cotizadas y/o fondos de inversión era del 18.9% en la ola de 2017

Por otro lado, la variable “peso de inmuebles sobre riqueza neta” es la siguiente variable que forma parte del grupo de variables económicas. Esta variable muestra el valor que tienen las propiedades inmobiliarias sobre la riqueza neta del hogar. Definimos la riqueza neta del hogar como el valor total de los activos reales y financieros, sin incluir vehículos, y restando el importe de las deudas pendientes de amortizar por el hogar. Esta es la variable vinculada a la vivienda en la que se sustenta la hipótesis a contrastar del modelo teórico, y a la que debemos prestar especial atención.

El grado de aversión al riesgo se mide en el cuestionario como la autopercepción del riesgo financiero dispuesto a asumir para obtener un nivel de beneficios esperado.

También se ha seleccionado una variable binaria que refleja si el hogar ha ahorrado cualquier cantidad no negativa en los últimos 12 meses.

Y también una variable que refleja los planes de ahorro que tiene el hogar para el futuro, ya sea mantener, aumentar o disminuir su ahorro respecto a lo que ha ahorrado en el último año.

La última variable de este grupo es la renta del hogar. La renta se calcula como la suma de rentas laborales y no laborales de todos los miembros del hogar en el año anterior al de la encuesta.

Tabla 3: Estadísticos descriptivos de las variables económicas

Ola		2002	2005	2008	2011	2014	2017
Peso de Activos con Riesgo en Cartera de Inversión	Media	0.07	0.08	0.05	0.06	0.07	0.07
	Mediana	0	0	0	0	0	0
	Desv. Típica	0.00	0.00	0.00	0.00	0.00	0.00
Peso de Inmuebles sobre Riqueza Neta	Media	0.99	1.14	1.72	1.26	1.61	1.37
	Mediana	0.96	0.97	0.98	0.96	0.95	0.926
	Desv. Típica	0.00	0.01	0.46	0.23	0.22	0.14
Aversión al Riesgo	Mucho	80.4 %	83 %	84.3 %	87.8 %	86.3 %	84.4 %
	Bastante	16.8 %	14.5 %	13.5 %	10.7 %	12.2 %	13.3 %
	Algo	2.2 %	1.8 %	1.9 %	1 %	1.1 %	1.4 %
	Poco	0.6 %	0.8 %	0.4 %	0.5 %	0.5 %	0.8 %
Ahorro Últimos 12 Meses	¿Ha ahorrado?: No	73.2 %	66.1 %	67.1 %	63.6 %	67.9 %	65.1 %
	¿Ha ahorrado?: Sí	26.8 %	33.9 %	32.9 %	36.4 %	32.1 %	33.9 %
Planes de Ahorro Futuros	Mantener Ahorro	57 %	58.1 %	56.2 %	50.9 %	58.7 %	49.3 %
	Disminuir Ahorro	28.6 %	26.6 %	31.2 %	37 %	27.2 %	27.9 %
	Aumentar Ahorro	14.4 %	15.4 %	12.6 %	12.1 %	14.1 %	22.8 %
Renta	Media	41615.7	40116.41	38365.41	36190.49	32107.18	36158.44
	Mediana	31225.62	29040.54	29262.9	26253.24	23736.49	26518.73
	Desv. Típica	872.37	10002.06	1037.03	1054	601.74	747.49
Número de Observaciones		4977	5672	5925	5846	5779	6014

Estadísticos descriptivos de las variables económicas de interés de la EFF de 2002, 2005, 2008, 2011, 2014 y 2017. Todos los estadísticos que se muestran se han obtenido haciendo uso de los pesos muestrales, los pesos de replicación y las 5 imputaciones para cada ola. La variable “Renta” está expresada en euros de 2017.

De la [Tabla 3](#) destacamos que el peso de los activos con riesgo en la cartera de inversión tiene mediana 0, ya que la mayoría de la población española no dispone de activos con riesgo. Además, en media, tan solo el 7 % de la riqueza financiera de las familias españolas corresponde a activos con riesgo en la mayoría de las olas.

En cuanto al peso de los inmuebles sobre la riqueza neta, podemos observar que tiene una media mayor que 1 para casi todas las olas. Esto indica que la mayoría de los hogares tienen viviendas valoradas por encima de su riqueza neta, lo que solamente es posible mediante endeudamiento. Aunque, en promedio, los hogares tienen viviendas con un valor superior a su riqueza neta en cinco de las seis olas, se observan diferencias en la magnitud de esta variable en las distintas olas. Estas diferencias pueden deberse al endeudamiento previo a la crisis de 2008 y a la posterior caída en los precios de las viviendas. Mientras

que en 2002 y 2005 el valor de los inmuebles suponía el 99 % y el 114 % de la riqueza neta del hogar, en 2008 y 2014 suponía el 172 % y el 161 % respectivamente.

Por otro lado, los hogares en España presentan una alta aversión al riesgo. En general, el 80 % los hogares son muy aversos al riesgo para cualquier ola, mientras que el porcentaje de hogares que son poco aversos no llegan al 1 % en ninguna ola. Además, se observa que en los años posteriores a la crisis financiera de 2008, aumenta el porcentaje de hogares que son muy aversos al riesgo.

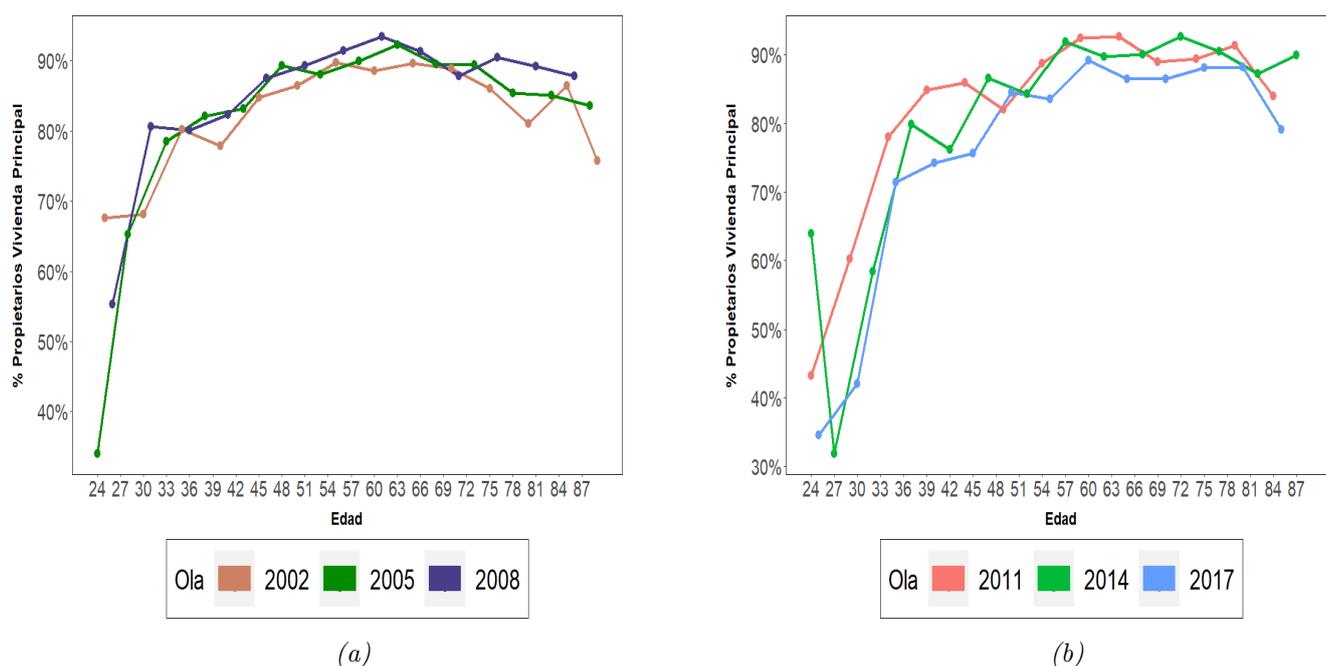
En cuanto a las variables vinculadas al ahorro, a pesar de que más de la mitad de las familias declaran no haber ahorrado en los últimos 12 meses, el 50 % de los hogares predice mantener para el próximo año el mismo nivel de ahorro que en el pasado.

Para terminar con las variables económicas, destacamos que la renta media ajustada por la inflación en 2017 ha disminuido sensiblemente respecto a los años previos a la crisis financiera.

5.2. Activos Financieros y Ciclo Vital

A continuación vamos a describir gráficamente la evolución de diversas variables con una perspectiva de ciclo vital para los hogares españoles. En las figuras que siguen hay dos tipos de análisis. El primero de ellos es el análisis por edad, donde cada línea representa una ola. El segundo es el análisis de cohorte por edad, donde cada línea representa una cohorte.

Figura 1: Tasa de propietarios de vivienda principal



Porcentaje de hogares propietarios de vivienda principal por edad y ola. Análisis por edad y olas.

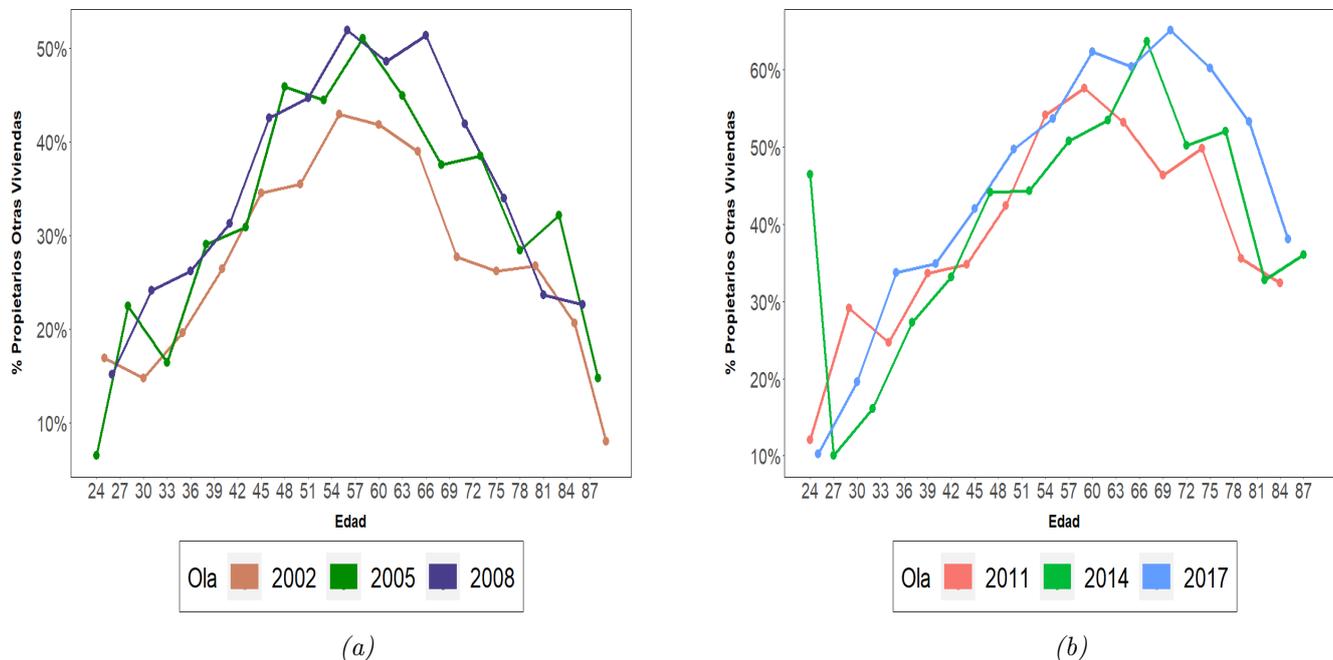
La Figura 1 representa la tasa de propietarios de vivienda principal por edad. Los datos se muestran en dos subfiguras *a* y *b* con el fin de evitar que se solapen, lo que facilita el análisis de las diversas líneas representadas. Cada una de las líneas dibujadas en las subfiguras *a* y *b* muestra una ola de la encuesta. Por otro lado, se clasifica a los hogares por la edad de su cabeza de familia en intervalos de cinco años y se calcula para cada grupo de edad el porcentaje de hogares que son propietarios de la vivienda principal.

En primer lugar, observamos que el porcentaje de propietarios con cabeza de familia más jóvenes ha disminuido en las olas más recientes. Concretamente en las olas de 2002 y 2005, más del 60% de los hogares con cabeza de familia entre 25-35 años eran propietarios de una vivienda. Por el contrario, en las olas de 2014 o 2017 menos del 40% de los hogares más jóvenes son propietarios. En cualquier caso, la tasa de propietarios de vivienda principal aumenta rápidamente con la edad para cualquier año. En todas las olas, el porcentaje máximo de propietarios de vivienda principal se alcanza cuando el cabeza de familia tiene entre 55 y 70 años, y es aproximadamente al 90%. No obstante, a partir de los 70 años aproximadamente la tasa de propietarios se mantiene prácticamente estable, aunque comienza a disminuir muy levemente, dibujando una J invertida en la relación entre la tenencia de vivienda principal y la edad.

En general, las distintas olas tienen una evolución similar, siendo la principal diferencia entre ellas un retraso en el porcentaje de propietarios de vivienda principal por edad para las olas más recientes. Los jóvenes en 2017 se convierten en propietarios más tarde de lo

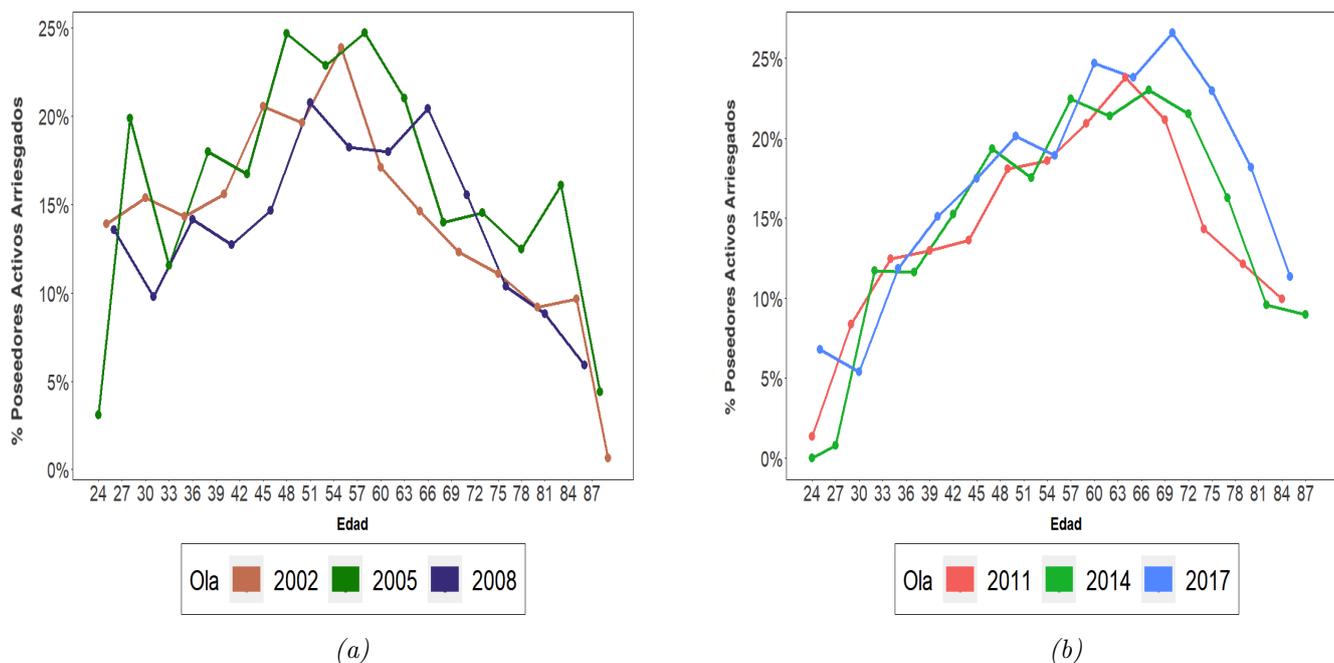
que lo hacían en 2002. De igual modo, el porcentaje máximo de propietarios se alcanza a una edad mayor en 2017 (70 años) de lo que lo hacían los hogares en 2002 (55 años).

Figura 2: Tasa propietarios otras viviendas distintas de la vivienda principal



Porcentaje de hogares propietarios de viviendas distintas a la principal por edad. Análisis por edad y olas.

Figura 3: Tasa de propietarios de acciones y fondos de inversión



Porcentaje de hogares que poseen activos con riesgo (acciones y fondos de inversión). Análisis por edad y olas.

En las Figuras 2 y 3, se analiza la evolución del porcentaje de propietarios de viviendas distintas a la principal con la edad, y la evolución del porcentaje de inversores en activos con riesgo con la edad respectivamente. A pesar de que las figuras representan distintas variables, la evolución de ellas con la edad es similar.

En los gráficos (a) y (b) de la Figura 2, se observa que la proporción de hogares españoles propietarios de viviendas distintas de la vivienda principal crece aproximadamente hasta los 55-70 años.

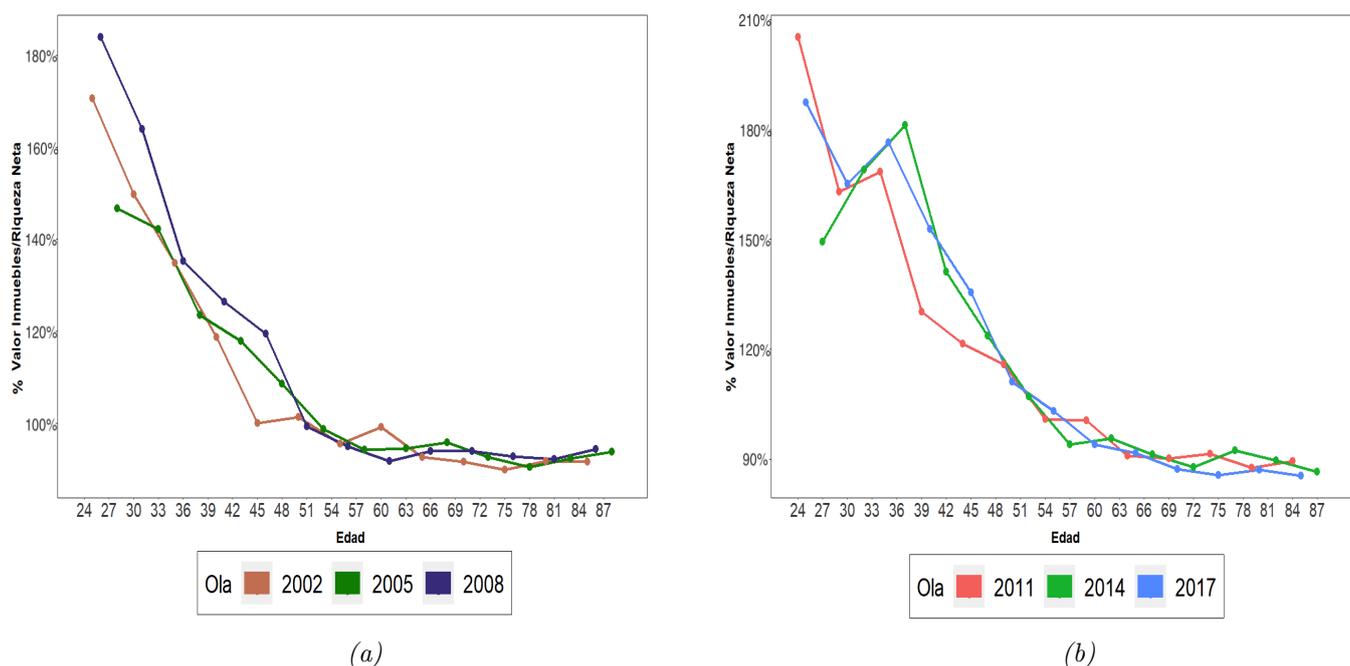
En las primeras olas, el porcentaje máximo de propietarios de segundas propiedades se alcanza cuando el cabeza de familia del hogar tiene 55 años aproximadamente y es igual a 40 %-50 %. En las olas más recientes, no solamente ha aumentado la edad a la que se alcanza el máximo, también ha aumentado el porcentaje de propietarios de segundas viviendas, superando el 60 % en 2014 y 2017.

A diferencia de lo que ocurría con la vivienda principal, una vez alcanzado el máximo, el porcentaje de propietarios de segundas residencias disminuye sensiblemente con la edad, en especial en 2002, 2005 y 2008. Esta Figura muestra una forma de U invertida en la relación entre la tenencia de segundas propiedades y la edad.

Por otro lado, las subfiguras a y b de la Figura 3 también dibujan una forma de U invertida en la relación entre la tenencia de activos arriesgados y la edad. Esta forma de U invertida es más pronunciada en las olas más recientes porque la participación en mercados financieros de los hogares más jóvenes disminuye. En 2002, 2005 y 2008, en torno al 15 % de los hogares más jóvenes poseía activos arriesgados. Por el contrario, en las olas posteriores a la crisis financiera, menos del 5-10 % de los hogares más jóvenes era propietario de este tipo de activos. En todos los años, el punto máximo se alcanza cuando aproximadamente el 25 % de los hogares son propietario de algún tipo de activo arriesgado, aunque la edad en la que esto ocurre varía en las diferentes olas. Nuevamente, al igual que en la Figura 2, se desplaza el máximo hacia la derecha en las olas más recientes. Mientras que en 2002 se alcanza el máximo cuando el cabeza de familia tiene 55 años, en 2017 lo hace a los 70 años.

Es importante recordar que la tasa de propietarios de la vivienda principal (Figura 1) no muestra la forma de U invertida que comparten la tasa de propietarios de viviendas secundarias y la tasa de poseedores de activos con riesgo. Debido a ello, podemos pensar que la inversión en propiedades inmobiliarias distintas de la vivienda principal tiene una evolución similar a la inversión en activos arriesgados con la edad.

Figura 4: Valor de los inmuebles sobre la riqueza neta (solamente propietarios de inmuebles)



Peso de los inmuebles sobre la riqueza neta (valor total de los activos reales y financieros, sin incluir vehículos, y restando el importe de las deudas pendientes de amortizar por el hogar) para los hogares que poseen inmuebles. Análisis por edad y olas.

La Figura 4 muestra el valor de las viviendas sobre la riqueza neta del hogar, restringimos la muestra a los propietarios de inmuebles y limitamos² el valor de los inmuebles sobre la riqueza neta para mantener una escala interpretable. Recordamos que la riqueza neta está compuesta por el valor total de los activos reales y financieros, sin incluir vehículos, y restando el importe de las deudas pendientes de amortizar por el hogar.

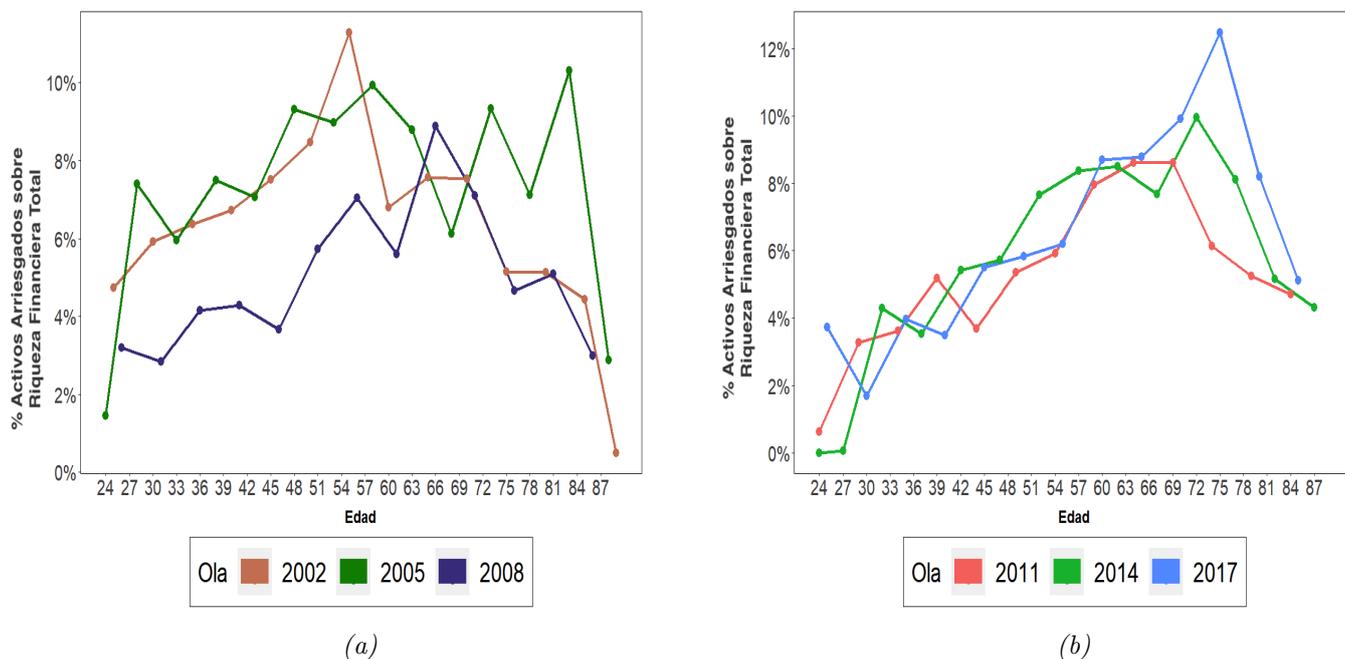
Podemos distinguir, en la Figura 4, que para los hogares con cabeza de familia más joven el valor de la vivienda es más del 100 % del valor de la riqueza neta. El ratio valor de los inmuebles sobre la riqueza neta disminuye a medida que el hogar acumula riqueza. No obstante, el valor de la vivienda sobre la riqueza neta decrece con la edad más rápidamente en las olas de 2002 y 2005, de lo que lo hace en las olas más recientes. Así, cuando el cabeza de familia tiene menos de 70 años, el valor de los inmuebles sobre la riqueza neta es mayor para una misma edad en las olas más recientes.

Por otro lado, cuando el cabeza de familia supera los 60-70 años aproximadamente, el ratio del valor de los inmuebles sobre riqueza neta es el mínimo durante el ciclo de

²Para la representación de la Figura 4 se ha limitado la variable de estado de la vivienda conforme a Flavin y Yamashita, (2011). Así, toma el valor 1 % cuando $\frac{ValorInmuebles}{RiquezaNeta} < 5\%$; el valor 10 % cuando el ratio está entre 5 % y 15 %; el 20 % entre 15 % y 25 %; el 112,5 % entre 105 % y 120 %; el 132,5 % cuando el ratio está entre 120 y 145 %; 175 % entre 145 % y 205 %; 250 % entre 205 % y 295 %; y por último, toma el valor 300 % cuando el ratio es superior a 295 %

vida. Una vez alcanzado el mínimo no parece haber grandes diferencias en el peso de los inmuebles sobre la riqueza neta en las distintas olas. Por lo que concluimos que hay un retraso con la edad en la acumulación de riqueza de los hogares en las olas más recientes (desplazamiento hacia la derecha) y que los hogares más jóvenes en las olas más recientes acceden a viviendas con un mayor valor respecto a la riqueza neta de las que accedían los jóvenes en 2002 o 2005.

Figura 5: *Peso de los activos con riesgo en la cartera de inversión*



Peso de los activos arriesgados (acciones y fondos de inversión) sobre la riqueza financiera total del hogar (suma del saldo de las cuentas que dispone el hogar para realizar pagos, el saldo de cuentas con las que no puede realizar pagos, el valor de las acciones cotizadas, las acciones no cotizadas, los activos de renta fija, las carteras gestionadas, los fondos de inversión, los planes de pensiones, los seguros de vida y el valor de lo que se le debe al hogar). Análisis por edad y olas.

En la Figura 5 se representa la variable “peso de los activos arriesgados en la cartera de inversión”. Esta variable muestra el peso que tiene el valor de las acciones y de los fondos de inversión sobre la riqueza financiera del hogar. Recordamos que la riqueza financiera total está compuesta por el saldo de las cuentas que dispone el hogar para realizar pagos, el saldo en las cuentas con las que no puede realizar pagos, el valor de las acciones cotizadas, las acciones no cotizadas, los activos de renta fija, las carteras gestionadas, los fondos de inversión, los planes de pensiones, los seguros de vida y de los derechos de cobro del hogar.

En las subfiguras (a) y (b), podemos destacar una forma de U invertida en la relación entre la tenencia de activos arriesgados y la edad. Observamos, como ocurría en la

Figura 3, que para los años 2011, 2014 y 2017, la forma de U invertida es muy evidente y pronunciada. Para las olas de 2002, 2005 y 2008, a pesar de que sigue existiendo esa forma de U invertida al dibujar la proporción de acciones y fondos de inversión sobre la riqueza financiera respecto a la edad del cabeza de familia, esta es algo más tenue.

Los hogares jóvenes tenían un mayor peso de acciones y fondos de inversión sobre su riqueza financiera en 2002 y 2005 que en el resto de años. Sin embargo, el peso máximo no ha variado en los distintos años, sino que simplemente se alcanza a una edad mayor para los hogares de las olas más recientes. Tras alcanzar el máximo, los hogares deshacen sus posiciones en activos con riesgo, lo que dibuja la U invertida. En las primeras olas, al liquidar activos con riesgo, el peso de estos llegaba a ser prácticamente cero cuando el cabeza de familia superaba los 85 años de edad. Por el contrario, en 2011, 2014 y 2017, una vez alcanzado el máximo, el ratio desciende más lentamente sin ser inferior al 4% durante la vejez.

Concluimos que en las últimas olas los hogares invierten en activos con riesgo más tarde de lo que lo hacían en las primeras olas. Además, los hogares de las olas más recientes mantienen en mayor medida las posiciones en activos arriesgados superados los 80 años.

A pesar de que el análisis de corte transversal nos ayuda a estudiar el comportamiento de los hogares durante el ciclo de vida, el que se estén comparando hogares con edades diferentes, pero que pertenecen a distintas generaciones limita la investigación. Para reforzar la investigación del ciclo vital, se podría realizar un estudio con datos de panel. Sin embargo, con los datos con los que estamos trabajando, el análisis de panel dispone de pocas observaciones que hayan participado en muchas olas y, por ello, se pierde variabilidad en el estudio a lo largo del ciclo de vida. Concretamente, de los 6413 hogares que contestaron a la encuesta en 2017, 3634 habían contestado también a la de 2014, únicamente 1647 hogares contestaron también a la encuesta de 2011 y solamente 475 contestaron a las encuestas de 2008 a 2017. A partir de aquí, ningún hogar contestó a las encuestas desde 2005 hasta 2017. De este modo, se ha optado por no realizar el análisis de panel y hacer, en su lugar, el de cohortes sintéticas.

Una “cohorte” es un grupo de personas que comparten la característica común de haber nacido en un mismo momento del tiempo. Las cohortes sintéticas son grupos de hogares entrevistados en diferentes olas cuyos cabezas de familia han nacido en un mismo periodo de tiempo (año, quinquenio, década, etc.) y representan adecuadamente a esa generación en la población total. Por tanto, el análisis de cohortes sintéticas es una herramienta útil, ya que permite, sin disponer de datos de panel, identificar relaciones o predecir el comportamiento de grupos por generaciones. Esta técnica consiste en crear pseudo-

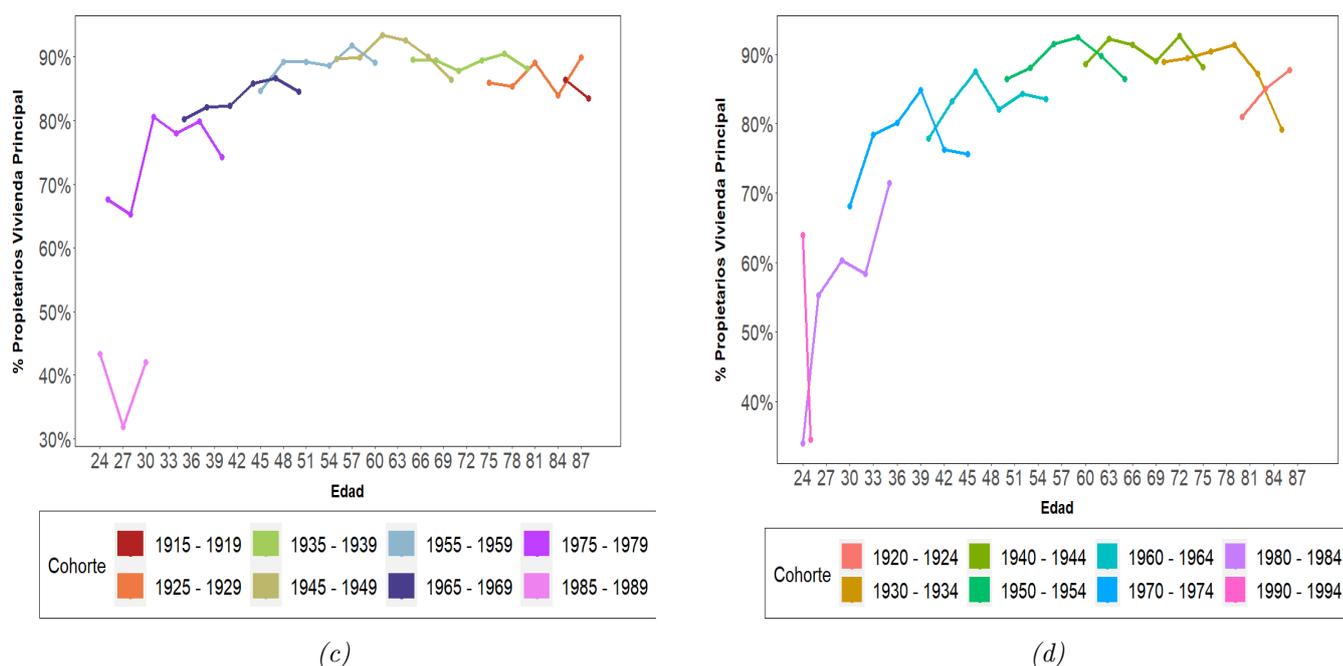
paneles utilizando datos de cortes transversales para crear las “cohortes sintéticas”. En Attanasio (1998) se encuentra una descripción detallada del estudio de cohorte.

Una vez explicada la utilidad del análisis de cohortes para nuestro trabajo, vamos a crear las cohortes sintéticas por años de nacimiento. Para ello, agrupamos los hogares por quinquenios desde 1915 hasta 1994. De este modo, los individuos nacidos entre 1915 y 1919 pertenecen a la primera cohorte, los nacidos entre 1920 y 1924 pertenecerán a la segunda cohorte y así sucesivamente hasta los nacidos entre 1990 y 1994 que pertenecen a la decimosexta cohorte.

En las subfiguras *c* y *d* de las Figuras 6-10 se representa el análisis por cohortes sintéticas. Representamos las cohortes en dos subfiguras para facilitar la interpretación. Cada línea coloreada representa una cohorte. Cada punto, representado en la mitad del intervalo de edad, muestra la media de la variable, que ha sido calculada con los datos de los hogares de esa cohorte y en ese rango de edad. En general, se observa que en todas las figuras por cohorte sintéticas, el patrón a lo largo del ciclo vital tiene la misma forma que en las subfiguras (a) y (b) de las Figuras 1-5. Sin embargo, además del patrón por ciclo de vida, este análisis nos permite comparar las variables para generaciones distintas a una misma edad, como mostramos a continuación.

Para facilitar el análisis, adelantamos que cuando dos puntos de dos líneas (cohortes) distintas a la misma edad tengan una separación vertical, diremos que a una misma edad esas generaciones se comportan de manera distinta. Cuanto mayor sea la separación vertical entre puntos, mayor será la diferencia en el comportamiento entre generaciones. Por el contrario, si los puntos de las distintas líneas se solapan, no existiría diferencia en el comportamiento de diferentes generaciones a una misma edad. Aplicaremos esta metodología a la interpretación de las siguientes figuras:

Figura 6: Tasa de propietarios de vivienda principal



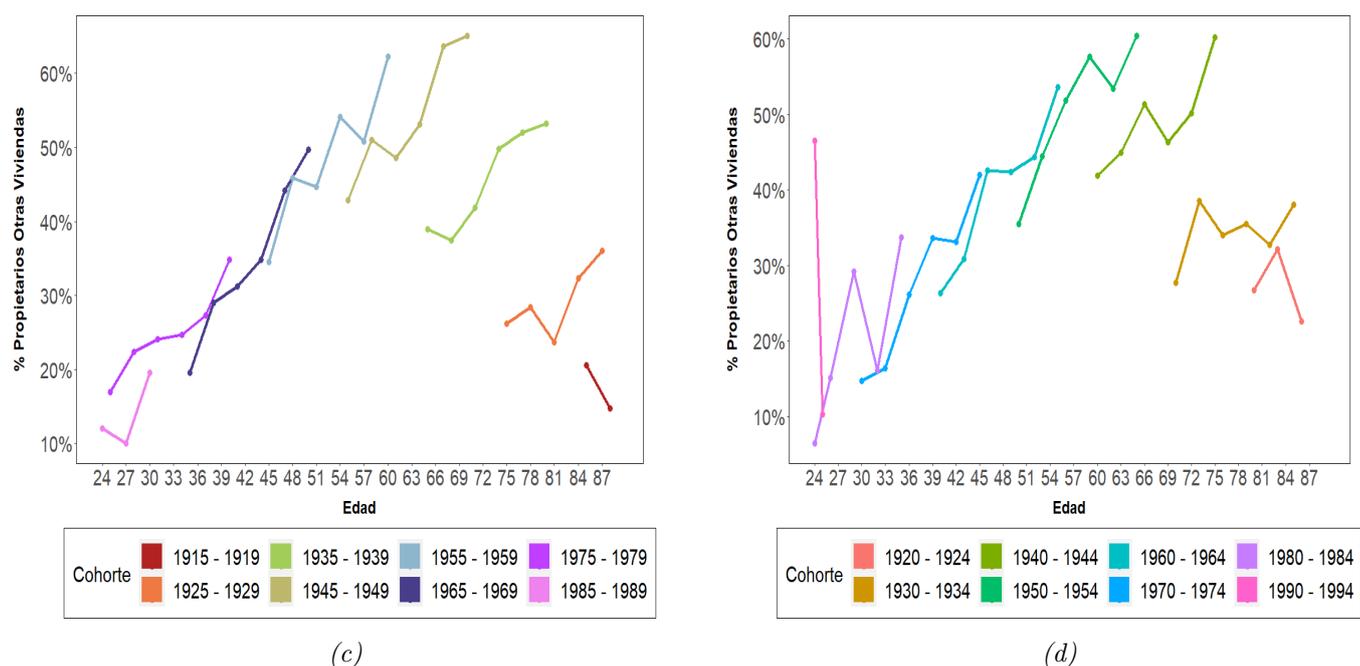
Porcentaje de hogares propietarios de vivienda principal por edad y ola. Análisis de cohorte.

En primer lugar, el análisis por cohortes de la Figura 6 muestra la proporción de propietarios de vivienda por cohortes sintéticas. Observamos, nuevamente, que la proporción de propietarios de la vivienda principal crece rápidamente con la edad en las generaciones más jóvenes. En cuanto al comportamiento de las diferentes cohortes: las mayores diferencias las encontramos en las generaciones nacidas después de 1950, donde para una misma edad, había más propietarios de vivienda principal en aquellas generaciones que nacieron antes.

Concretamente, mientras que el 75% de los hogares nacidos entre 1960 y 1964 eran propietarios de la vivienda principal a los 54 años. Más del 85% de los hogares nacidos entre 1950 y 1954 eran propietarios a los 54 años.

A pesar de que se observa cómo cada cohorte posterior tiene una menor proporción de propietarios de vivienda principal a una misma edad que las generaciones previas, este efecto se atenúa para las generaciones nacidas antes de 1950, en estas generaciones no podemos distinguir un patrón tan evidente entre generaciones.

Figura 7: Tasa propietarios otras viviendas distintas de la vivienda principal



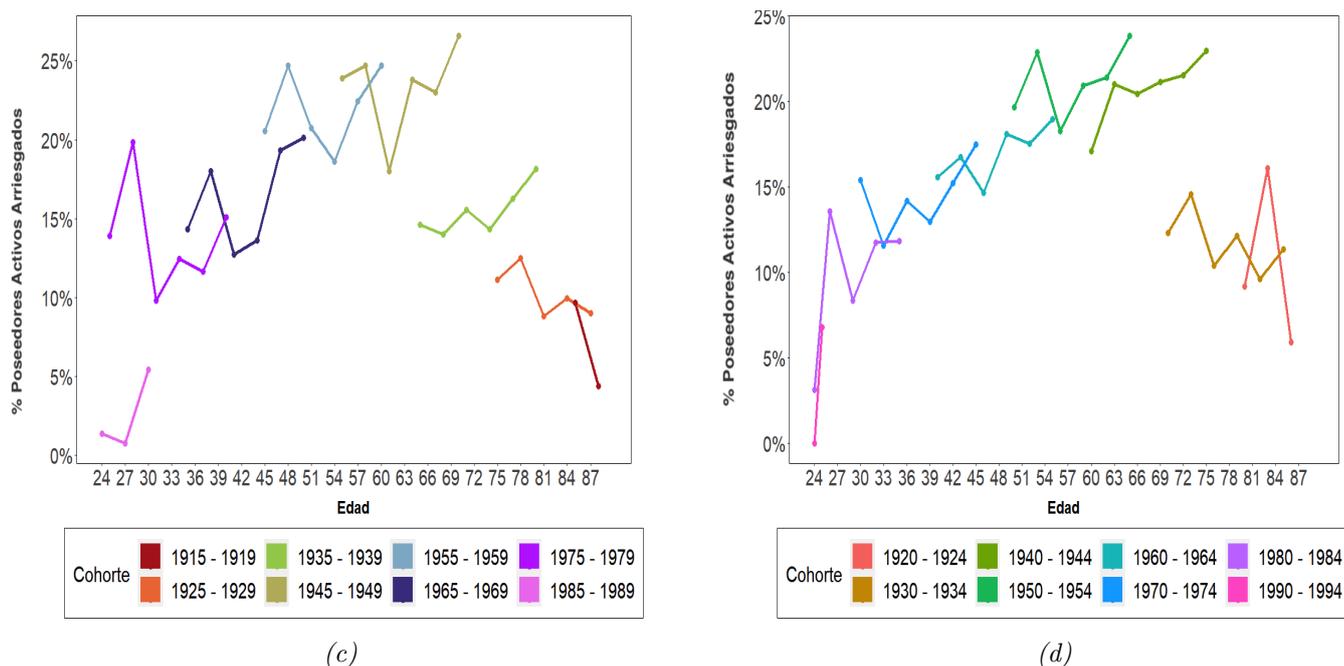
Porcentaje de hogares propietarios de viviendas distintas a la principal por edad. Análisis de cohorte.

En la Figura 7 se muestra el porcentaje de hogares que poseen segundas propiedades. En estas figuras se observan diferencias en el comportamiento de algunas cohortes.

En general, observamos que en las generaciones nacidas entre 1920 y 1960 hay más hogares con segundas propiedades en generaciones posteriores a una misma edad. La tasa de propietarios de viviendas distintas de la vivienda principal crece con las generaciones que nacieron más tarde entre 1920 y 1960. Por el contrario, en las generaciones nacidas después de 1960 no se observan grandes diferencias.

A modo de ejemplo, si se comparan generaciones que entre 2002 y 2017 han tenido de 60 a 87 años. A una misma edad, había la mitad de hogares propietarios de segundas viviendas nacidos entre 1925 y 1934, de lo que lo eran los nacidos entre 1935 y 1944. Este efecto se diluye al comparar generaciones más recientes, donde se ve un solapamiento entre líneas que no indica una diferencia en comportamiento.

Figura 8: Tasa de propietarios de acciones y fondos de inversión

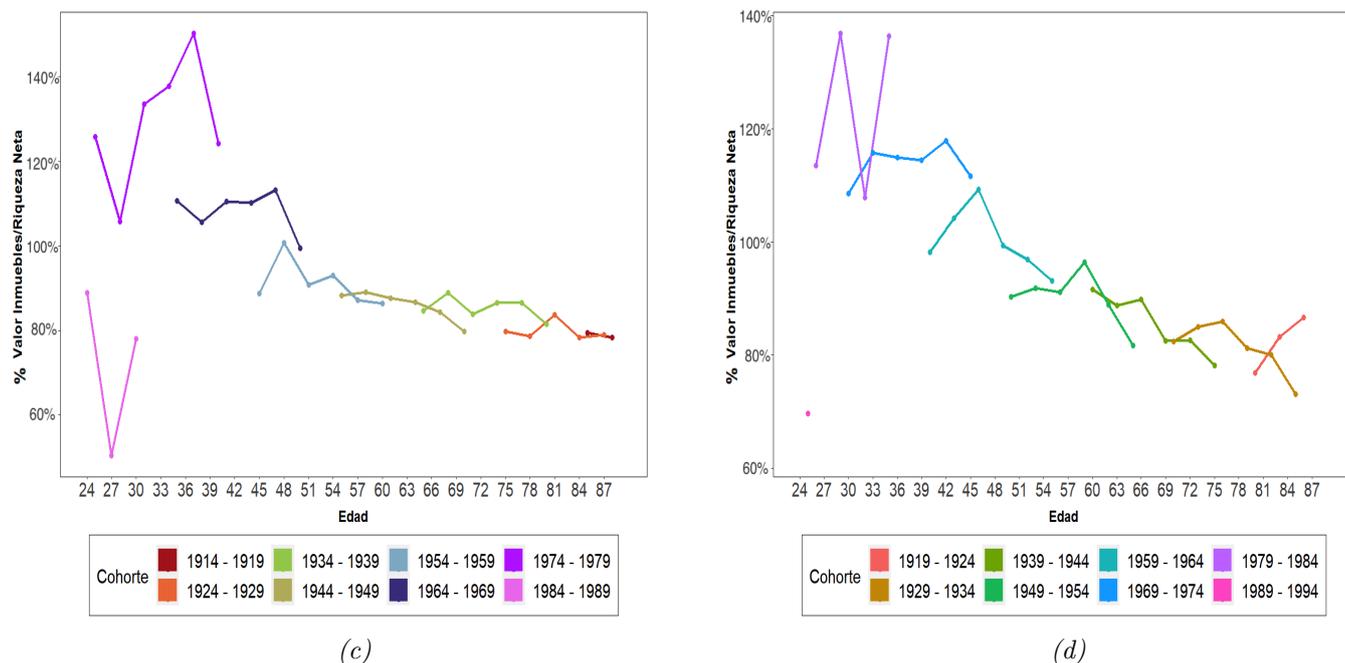


Porcentaje de hogares que poseen activos con riesgo (acciones y fondos de inversión). Análisis de cohorte.

En la [Figura 8](#) se muestra el porcentaje de hogares que poseen activos con riesgo. Observamos que los hogares nacidos en generaciones más jóvenes poseen, a una misma edad, menos activos con riesgo en sus carteras de inversiones que las generaciones que nacieron antes. Sin embargo, la evolución cambia para las generaciones nacidas antes de 1950. Podemos distinguir, por tanto, dos tendencias. La primera de ellas corresponde a los hogares que pertenecen a generaciones anteriores a 1950. Los hogares de estas generaciones invierten a una misma edad más que la generación anterior. Concretamente, en torno a un 22% de los hogares nacidos entre 1945 y 1949 poseían algún activo con riesgo a los 70 años. Por el contrario, únicamente el 12% de los hogares posee activos con riesgo en la generación anterior a esa misma edad.

La segunda tendencia corresponde a los hogares que pertenecen a generaciones posteriores a 1950. Estas generaciones a una misma edad invierten un menor peso de sus carteras de inversión en activos con riesgo que la generación anterior. Así, mientras que un 25% de los hogares nacidos entre 1955 y 1959 con cabeza de familia de 50 años, son poseedores de algún tipo de activo con riesgo, únicamente el 20% de los hogares nacidos entre 1965 y 1969 poseen activos con riesgo a los 50 años.

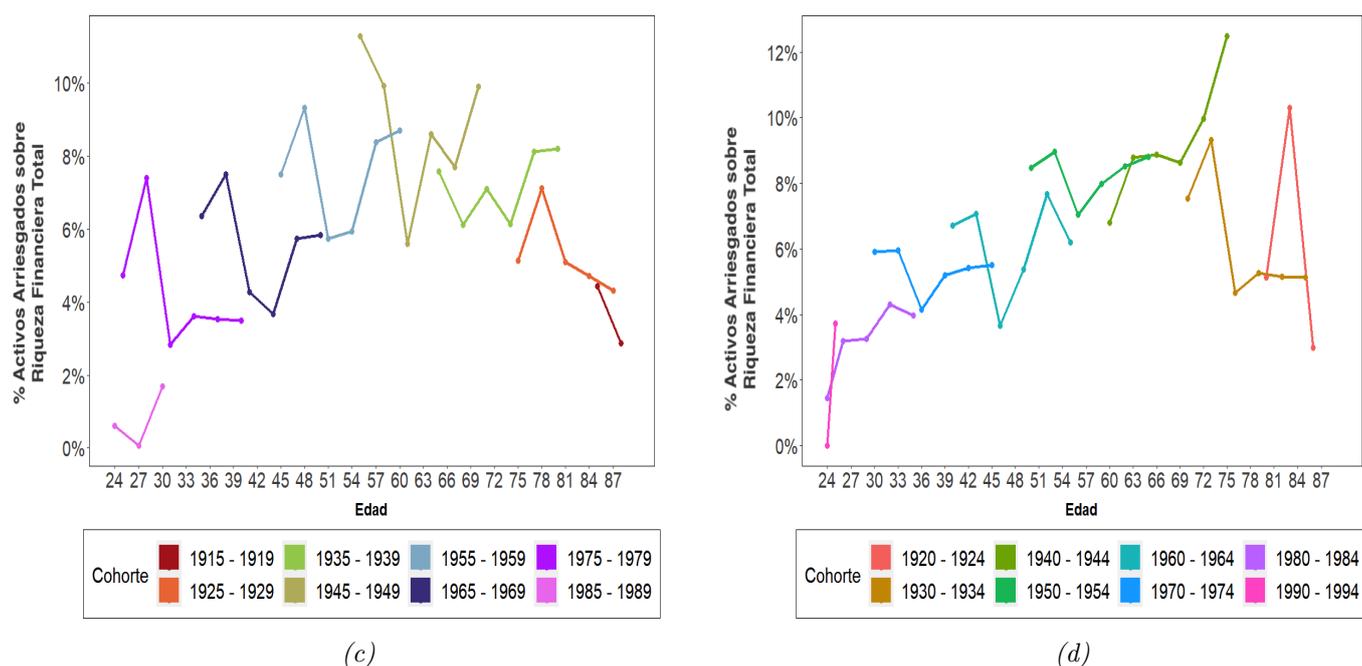
Figura 9: Valor de los inmuebles sobre la riqueza neta (solamente propietarios de inmuebles)



Peso de los inmuebles sobre la riqueza neta (valor total de los activos reales y financieros, sin incluir vehículos, y restando el importe de las deudas pendientes de amortizar por el hogar) para los hogares que poseen inmuebles. Análisis de cohorte.

La Figura 9 muestra el peso de los inmuebles sobre la riqueza neta de los hogares. Por un lado, destaca que las cohortes más jóvenes que estudiamos, esto es, las cohortes nacidas entre 1960 y 1994, presentan un mayor ratio del valor de los inmuebles sobre la riqueza neta que cohortes anteriores a una misma edad. Por otro lado, se observa que cuando el cabeza del hogar cumple los 50 años, se atenúa la diferencia en el comportamiento entre generaciones. A pesar de que, en general, las generaciones posteriores tienen un ratio del valor de los inmuebles sobre la riqueza neta mayor que las generaciones anteriores a una misma edad, esta diferencia se suaviza a medida que el hogar envejece. Por ello, se ve una forma escalonada en las generaciones más jóvenes de las subfiguras (c) y (d) de la Figura 9 que se aplana con la edad en las generaciones más mayores. Puede que esta relación entre el valor de los inmuebles y la riqueza neta sea resultado del aumento del precio relativo de la vivienda, o de que la edad de adquisición de una vivienda es cada vez mayor, como se observa en la Figura 1.

Figura 10: *Peso de los activos con riesgo en la cartera de inversión*



Peso de los activos arriesgados (acciones y fondos de inversión) sobre la riqueza financiera total del hogar (suma del saldo de las cuentas que dispone el hogar para realizar pagos, el saldo de cuentas con las que no puede realizar pagos, el valor de las acciones cotizadas, las acciones no cotizadas, los activos de renta fija, las carteras gestionadas, los fondos de inversión, los planes de pensiones, los seguros de vida y el valor de lo que se le debe al hogar). Análisis de cohorte.

Por último, la [Figura 10](#) muestra la evolución del peso de los activos con riesgo en la cartera de inversión con la edad. En estas subfiguras podemos observar relaciones distintas entre las generaciones previas y posteriores a 1950. Así, para las cohortes nacidas antes de 1950, en generaciones consecutivas, los hogares que nacieron antes tenían menos peso en activos con riesgo en sus carteras que generaciones posteriores a una misma edad. Sin embargo, para los hogares nacidos después de 1950 esta se revierte. A una misma edad, las cohortes consecutivas que nacieron antes tenían un mayor peso de activos con riesgo sobre su riqueza financiera que generaciones posteriores a una misma edad.

6. Análisis Econométrico

En esta sección, tras haber detallado las características de la muestra en lo que se refiere a las características de los hogares y su comportamiento inversor en vivienda y activos financieros arriesgados a lo largo del ciclo vital, procedemos a realizar el análisis econométrico. Nuestro objetivo principal es evaluar si el modelo teórico se alinea de manera coherente con la evidencia derivada de los datos recopilados en España por la EFF. En este proceso, emplearemos varios modelos econométricos para investigar la relación entre el porcentaje de la riqueza invertida en activos con riesgo y la edad del cabeza de

familia. Además, iniciaremos analizando esta relación sin considerar el valor de la vivienda sobre la riqueza neta como un factor relevante. Luego, introduciremos un nuevo regresor que incorporará este componente en nuestro análisis.

Para llevar a cabo la estimación de los modelos econométricos, implementamos algunas restricciones en nuestra muestra. Así, además de excluir de la muestra aquellos hogares que poseen renta neta no negativa y los hogares donde el cabeza de familia tienen menos de 24 años o más de 89, eliminamos también a todos los hogares que no posean activos inmobiliarios. Con estas restricciones disponemos de 30.048 observaciones.

La variable dependiente para las regresiones es el peso de activos con riesgo en la cartera de inversión, que como sabemos, se mide como la proporción de la riqueza total de activos financieros del hogar invertida en acciones y fondos de inversión.

Por otro lado, las variables explicativas son las enumeradas en la sección anterior. Esto es, los regresores seleccionados son la edad, el sexo, el estado civil, el número de hijos del hogar, la situación laboral del cabeza de familia, la situación laboral de cónyuge, el nivel de estudios del cabeza de familia, el nivel de estudios del cónyuge, la aversión al riesgo del hogar, haber ahorrado en los últimos 12 meses, los planes de ahorro futuro, la renta y la ola de la encuesta. Sin embargo, transformamos la edad en una variable categórica por intervalos, la renta como el logaritmo de la misma y adicionalmente incluimos la ola de la encuesta como regresor.

Es importante destacar que, debido a las categorías de las variables cualitativas, consideraremos como hogar de referencia aquel que contestó a la encuesta en 2017, con un cabeza de familia varón de entre 24 y 29 años, casado, que trabaja por cuenta ajena, muy averso al riesgo y con estudios menores a bachillerato que no ha ahorrado en el año anterior y tiene expectativas de mantener su nivel de ahorro en el futuro. Además, el cónyuge del cabeza de familia se encuentra inactivo y con estudios menores a bachillerato.

6.1. Modelo de Regresión Lineal

En esta subsección mostramos e interpretamos los resultados obtenidos en la regresión lineal múltiple. Para ello, exponemos la [Tabla 4](#) y la [Figura 11](#). En la tabla mostramos los resultados de la estimación por MCO para dos modelos. El primer modelo (Modelo sin controlar por variable vivienda) excluye como regresor el peso de los inmuebles sobre la riqueza neta. El segundo modelo (Modelo controlado por variable vivienda) incluye la variable vinculada a la vivienda como regresor. La figura, a su vez, muestra los coeficientes estimados y sus intervalos de confianza para todas las categorías de la variable edad. A continuación detallamos los resultados obtenidos.

En cuanto a las estimaciones por mínimos cuadrados ordinarios, encontramos que a un nivel de significación del 5 %, la edad, la situación laboral del cabeza de familia, la situación laboral del cónyuge, el nivel de estudios del cabeza del hogar, la aversión al riesgo, el ahorro en el último año, la renta y la ola de la encuesta resultan variables estadísticamente significativas. El signo de los coeficientes de estas variables es el mismo tanto cuando se incluye la variable que refleja el peso de los inmuebles sobre la riqueza neta (variable de estado de la vivienda), como en el modelo que no la incluye. Además, en el modelo que incluimos la variable de estado de la vivienda, el coeficiente estimado de esta también es estadísticamente significativamente diferente de cero para un nivel de significación del 5 %.

Tabla 4: *Resultados de estimación MCO*

	Modelo sin controlar por variable vivienda		Modelo controlado por variable vivienda	
	<i>Coficiente</i>	<i>Error estándar</i>	<i>Coficiente</i>	<i>Error estándar</i>
Constante	-0.20	0.02	-0.12	0.02
Peso Inmuebles sobre Riqueza Neta	—	—	-0.04***	0.00
Edad				
(29-34)	-0.00	0.012	-0.01	0.01
(34-39)	0.02	0.01	0.01	0.01
(39-44)	0.02	0.01	0.00	0.01
(44-49)	0.03**	0.01	0.00	0.01
(49-54)	0.04***	0.01	0.01	0.01
(54-59)	0.05***	0.01	0.02*	0.01
(59-64)	0.05***	0.01	0.02*	0.01
(64-69)	0.06***	0.014	0.04**	0.01
(69-74)	0.09***	0.01	0.06***	0.01
(74-79)	0.08***	0.01	0.05***	0.02
(79-84)	0.07***	0.02	0.04***	0.02
(84-89)	0.06***	0.02	0.03*	0.02
Sexo				
(Mujer)	-0.01	0.01	-0.00	0.01
Estado Civil				
Soltero	0.01*	0.01	0.01	0.01
Divorciado	-0.01	0.01	-0.01	0.01
Viudo	-0.01	0.01	-0.01	0.01
Hijos Hogar	-0.00	0.00	-0.00	0.00
Situación Laboral Cabeza Familia				
Jubilado	0.01*	0.01	0.01	0.01
Trabajador Cuenta Propia	0.03***	0.01	0.02*	0.01
Desempleado	0.02***	0.01	0.02**	0.01
Inactivo	0.00	0.01	-0.00	0.01

Continúa en la próxima página

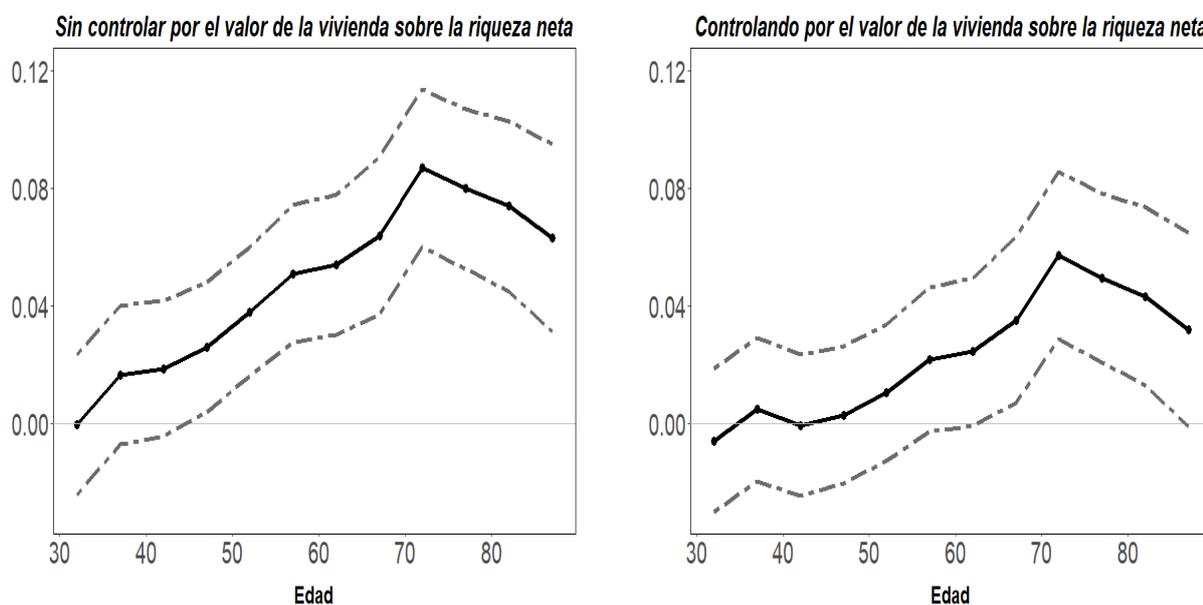
<i>Continuación tabla anterior</i>				
	Modelo sin controlar por variable vivienda		Modelo controlado por variable vivienda	
	<i>Coefficiente</i>	<i>Error estándar</i>	<i>Coefficiente</i>	<i>Error estándar</i>
Situación Laboral Cónyuge				
Jubilado	-0.01***	0.01	-0.01*	0.01
Trabajador Cuenta Ajena	-0.02***	0.01	-0.03***	0.01
Trabajador Cuenta Propia	-0.00	0.01	-0.00	0.01
Desempleado	-0.00	0.01	-0.00	0.01
Nivel Estudios Cabeza Familia				
Bachillerato	0.03***	0.01	0.03***	0.01
Universitarios	0.09***	0.01	0.09***	0.01
Nivel Estudios Cónyuge				
Bachillerato	0.01	0.01	0.01	0.01
Universitarios	0.01	0.01	0.01	0.01
Aversión al Riesgo				
Poco	0.11***	0.04	0.11***	0.04
Algo	0.17***	0.02	0.17***	0.02
Bastante	0.11***	0.01	0.11***	0.01
Ahorro Último Año				
Sí (ahorro positivo)	0.02***	0.00	0.01***	0.00
Planes Ahorro Futuros				
Aumentar Ahorro	0.00	0.01	0.00	0.01
Disminuir Ahorro	0.00	0.00	0.00	0.00
log (Renta)	0.02***	0.00	0.02***	0.00
Ola				
2002	0.02***	0.01	0.01*	0.01
2005	0.03***	0.01	0.02***	0.01
2008	-0.01	0.01	-0.01*	0.01
2011	-0.00	0.01	-0.00	0.01
2014	0.00	0.01	0.00	0.01

*Coefficientes estimados por MCO con pesos muestrales mediante imputación múltiple para la EFF de 2002, 2005, 2008, 2011, 2014 y 2017, disponemos de 30.048 observaciones para cada imputación. La variable dependiente es el peso que tienen los activos con riesgo en la cartera de inversión del hogar. Los errores estándar mostrados son errores estándar bootstrap usando pesos de replicación. Adicionalmente, junto a los coeficientes estimados se indica con '***', '**' y '*' los coeficientes estimados que son estadísticamente significativamente diferentes de 0 a un nivel de confianza del 1 %, 5 % y 10 % respectivamente.*

La variable peso de los inmuebles sobre la riqueza neta solamente se incluye, como su nombre indica, en el modelo controlado por la variable vivienda. Esta variable presenta un coeficiente estimado negativo y significativamente diferente de cero al 1%. Así, valores altos en el peso de inmuebles sobre riqueza neta, están asociados a valores bajos en el peso de la inversión en activos con riesgo sobre la riqueza financiera del hogar.

La edad, que tiene como referencia la categoría (24-29), se presenta como una variable relevante en ambos modelos. Como vemos en la [Tabla 4](#), en el modelo sin controlar por la variable vivienda, los coeficientes estimados para todas las categorías que comprenden edades superiores a los 44 años son significativamente diferentes de 0 para un nivel de significación del 5%. Por el contrario, en el modelo controlado por la variable vivienda, solamente son estadísticamente significativamente diferentes de 0 al 5% los coeficientes estimados para las categorías (64-69), (69-74), (74-79) y (79-84). Con el objetivo de estudiar más en detalle el comportamiento de la edad entre los dos modelos estimados mostramos la [Figura 11](#).

Figura 11: Relación entre edad e inversión en activos con riesgo: coeficientes estimados por MCO para las categorías de la variable “edad”



Intervalos de confianza y coeficientes estimados por MCO en el modelo sin controlar por la variable peso de inmuebles sobre la riqueza neta (variable vivienda), y en el modelo controlado por el peso de los inmuebles sobre la riqueza neta.

La [Figura 11](#) representa con las líneas continuas los valores que toman los coeficientes estimados para las distintas categorías de la variable “edad”. Cada punto dibujado refleja el valor que toma el coeficiente estimado en la regresión MCO para cada categoría de la

variable edad, tanto para el modelo sin controlar por variable vivienda (figura izquierda), como para el modelo controlado por la variable vivienda (figura derecha). Adicionalmente, se dibujan (líneas de color gris) los intervalos de confianza al 95 % para ambos modelos. Podemos observar una relación cuadrática entre el peso de los activos arriesgados sobre la riqueza financiera y la edad en ambos modelos. Además, encontramos diferencias entre el modelo que incluye la variable vinculada a la vivienda y el que no la incluye.

Las principales diferencias entre ambos modelos radican en que, en el modelo que incorpora la variable vivienda, la magnitud de los coeficientes estimados para las categorías de la variable edad disminuye considerablemente. Además, en el modelo controlado por la variable vivienda, se observa que los intervalos de confianza son ligeramente más amplios. Combinando ambos efectos, encontramos que, en el modelo controlado por la variable vivienda, muchos de los coeficientes estimados para las categorías de la variable edad incluyen el valor 0 dentro de su intervalo de confianza. Este fenómeno ilustra una pérdida de capacidad explicativa de la edad en el segundo modelo.

Respecto a la aversión al riesgo del hogar, los coeficientes estimados para las categorías de esta variable son estadísticamente significativamente distintos de cero a un nivel de significación del 1 % en ambos modelos. Como cabía esperar, ya que se trata de una variable que refleja el riesgo financiero tolerado autopercebido por el hogar, los coeficientes estimados son positivos para las categorías que menos aversión al riesgo presentan respecto a los que son muy aversos (categoría de referencia).

A pesar de que la regresión por Mínimos Cuadrados Ordinarios puede presentar buenas propiedades, bajo heterocedasticidad los errores estándar de los estimadores MCO están sesgados y la inferencia no es válida. Puesto que la heterocedasticidad es habitual en modelos con microdatos y variables financieras como las que consideramos, estimaremos, igual que hacen Flavin y Yamashita (2011) en su análisis, la regresión cuantílica.

La regresión cuantílica es una herramienta útil para estimar modelos donde los datos presentan heterocedasticidad o valores atípicos. Cada cuantil representa una línea de regresión única y su pendiente puede variar considerablemente según la distribución de los datos. Así mismo, el método de estimación de la regresión cuantílica no requiere suposiciones sobre la perturbación aleatoria y puede resolver los problemas asociados a MCO.

Además, independientemente de la presencia de heterocedasticidad, la regresión cuantílica mediana presenta ventajas respecto a la regresión lineal. Cuando la distribución de los residuos no es normal, la regresión cuantílica suele ser un estimador más eficiente que MCO.

6.2. Regresión Cuantílica

Se presenta la regresión cuantílica como una alternativa para obtener estimaciones que resuelvan las debilidades que presenta la estimación por Mínimos Cuadrados Ordinarios. No obstante, pronto nos encontramos con inconvenientes a la hora de estimar estos modelos que nos llevan a descartar la regresión cuantílica para obtener estimaciones con los datos de la EFF.

En la [Tabla 5](#) presentamos los coeficientes estimados para los cuantiles 0.5, 0.6, 0.75 y 0.9. Los cuatro modelos estimados incluyen las mismas variables dependiente e independientes que el modelo MCO controlado por la variable vivienda. Vemos que la mayoría de los coeficientes estimados son 0 para los primeros dos modelos de regresión. Por el contrario, obtenemos coeficientes distintos a 0 para $\tau = 0.75$ y $\tau = 0.9$.

Tabla 5: Resultados de estimación Regresión Cuantílica

	$\tau = 0.5$	$\tau = 0.6$	$\tau = 0.75$	$\tau = 0.9$
	<i>Coefficientes</i>			
Constante	0.00	0.00	-0.27	-0.11
Peso Inmuebles sobre Riqueza Neta	0.00	0.00	-0.05	-0.11
Edad				
(29-34)	0.00	0.00	-0.00	-0.04
(34-39)	0.00	0.00	0.00	0.04
(39-44)	0.00	0.00	-0.00	-0.04
(44-49)	0.00	0.00	-0.00	-0.02
(49-54)	0.00	0.00	-0.00	-0.02
(54-59)	0.00	0.00	-0.00	-0.01
(59-64)	0.00	0.00	0.01	0.03
(64-69)	0.00	0.00	0.03	0.14
(69-74)	0.00	0.00	0.07	0.21
(74-79)	0.00	0.00	0.05	0.23
(79-84)	0.00	0.00	0.04	0.27
(84-89)	0.00	0.00	0.04	0.28
Sexo				
Mujer	0.00	0.00	-0.00	-0.01
Estado Civil				
Soltero	0.00	0.00	0.02	0.07
Divorciado	0.00	0.00	0.02	0.03
Viudo	0.00	0.00	-0.01	-0.06
Hijos Hogar	0.00	0.00	-0.00	0.01
Situación Laboral Cabeza Familia				
Jubilado	-0.00	0.00	0.01	0.06
Trabajador Cuenta Propia	0.00	0.00	0.03	0.13
Desempleado	0.00	0.00	0.02	0.02
Inactivo	0.00	0.00	0.01	0.04

Continúa en la próxima página

<i>Continuación tabla anterior</i>				
	$\tau = 0.5$	$\tau = 0.6$	$\tau = 0.75$	$\tau = 0.9$
	<i>Coefficientes</i>			
Situación Laboral Cónyuge				
Jubilado	0.00	0.00	-0.03	-0.06
Trabajador Cuenta Ajena	0.00	0.00	-0.02	-0.04
Trabajador Cuenta Propia	0.00	0.00	-0.02	-0.04
Desempleado	0.00	0.00	-0.01	-0.06
Nivel Estudios Cabeza Familia				
Bachillerato	0.00	0.00	0.06	0.19
Universitarios	0.09	0.20	0.34	0.27
Nivel Estudios Cónyuge				
Bachillerato	0.00	0.00	0.03	0.08
Universitarios	0.00	0.00	0.01	0.09
Aversión al Riesgo				
Poco	0.29	0.46	0.44	0.26
Algo	0.44	0.42	0.43	0.24
Bastante	0.21	0.31	0.33	0.20
Ahorro Último Año				
Sí (ahorro positivo)	0.00	0.00	0.01	0.02
Planes Ahorro Futuro				
Aumentar Ahorro	0.00	0.00	0.01	0.02
Disminuir Ahorro	0.00	0.00	0.01	0.02
log (Renta)	0.00	0.00	0.04	0.04
Ola				
2002	0.00	0.00	0.01	0.06
2005	0.00	0.00	0.01	0.07
2008	0.00	0.00	-0.01	-0.01
2011	0.00	0.00	0.00	0.00
2014	0.00	0.00	0.01	0.01

Coefficientes estimados por Regresión Cuantílica con pesos muestrales mediante imputación múltiple para la EFF de 2002, 2005, 2008, 2011, 2014 y 2017, disponemos de 30.048 observaciones para cada imputación. La variable dependiente es el peso que tienen los activos con riesgo en la cartera de inversión del hogar.

En la [Tabla 5](#) se observa que los coeficientes estimados para algunos cuantiles toman el valor 0 para la mayoría de las variables independientes. La razón de que los coeficientes estimados sean 0, es la no participación en los mercados de activos con riesgo es prevalente en los hogares españoles, independientemente de cuáles sean sus características. Un análisis de la distribución de la variable dependiente condicional a los valores que toman cada uno de los regresores, devuelve valores cero para los percentiles por debajo de 70-75. La única excepción es la aversión al riesgo (véase que en ese caso los coeficientes estimados no son cero). Esta peculiaridad de los datos españoles hace la regresión cuantílica no sea apropiada para estimar y contrastar el modelo teórico de Flavin y Yamashita para España.

La baja participación en mercados de activos con riesgo lleva a una acumulación de observaciones que toman el valor 0 en la variable dependiente de nuestro estudio. Cuando la variable dependiente presenta acumulación de observaciones en 0, mientras que el resto de observaciones presentan una distribución continua entre valores positivos, es a lo que Wooldridge (2010) llama solución esquina y para la cual el modelo Tobit es una herramienta apropiada. El modelo Tobit además de ser una buena metodología para tratar datos con solución esquina, también permite manejar la heteroscedasticidad que podría limitar los resultados del modelo lineal.

6.3. Tobit

Puesto que nuestros datos presentan solución esquina en la variable dependiente peso de activos con riesgo sobre la riqueza financiera total (véase [Tabla 3](#)), aplicaremos un modelo Tobit con censura inferior e igual a cero en la variable dependiente. En esta subsección, presentamos los resultados obtenidos en la regresión Tobit, los cuales se detallan en la [Tabla 6](#) y en la [Figura 12](#).

En primer lugar, es importante señalar, como hemos explicado en la Sección 4, que en los modelos Tobit los efectos marginales no coinciden con los coeficientes estimados, y por ello la interpretación de los mismos no es directa. En la [Tabla 6](#) además de los coeficientes estimados y sus desviaciones estándar, mostramos los efectos marginales medios calculados tanto para el modelo sin controlar por la variable vivienda, como para el modelo controlado por la variable vivienda.

En la [Figura 12](#), representamos los coeficientes estimados y sus intervalos de confianza para todas las categorías de la variable edad.

En relación con las estimaciones Tobit, se observa que, a un nivel de significación del 5%, la edad, la situación laboral del cónyuge, el nivel de estudios del cabeza del hogar, el nivel de estudios del cónyuge, la aversión al riesgo, el ahorro en el último año, los

planes de ahorro futuros, la renta y la ola de la encuesta resultan variables estadísticamente significativas tanto en el modelo sin controlar por la variable vivienda, como en el modelo controlado por la variable vivienda. Adicionalmente, el coeficiente estimado para la variable “peso de inmuebles sobre la riqueza neta” también es estadísticamente significativamente diferente de cero al 5% en el modelo que incluye dicha variable (Modelo controlado por variable vivienda).

En vista de los resultados de las estimaciones Tobit podemos decir que, en general, las variables independientes mantienen la misma relación con la variable dependiente que mostraban en el modelo MCO y que la magnitud de los efectos marginales entre Tobit y MCO son similares.

Tabla 6: Resultados de estimación Tobit

	Modelo sin controlar por variable vivienda			Modelo controlado por variable vivienda		
	<i>Coficiente</i>	<i>Error estándar</i>	<i>Efectos Marginales Medios</i>	<i>Coficiente</i>	<i>Error estándar</i>	<i>Efectos Marginales Medios</i>
Constante	-2.71	0.21		-2.06	0.21	
Peso Inmuebles sobre Riqueza Neta	—	—	—	-0.28***	0.03	-0.05
Edad						
(29-34)	0.01	0.07	0.00	-0.01	0.07	-0.00
(34-39)	0.11	0.07	0.01	0.05	0.07	0.01
(39-44)	0.14**	0.06	0.02	0.03	0.07	0.01
(44-49)	0.21***	0.06	0.03	0.08	0.07	0.01
(49-54)	0.25***	0.06	0.03	0.09	0.06	0.01
(54-59)	0.30***	0.06	0.04	0.13*	0.07	0.02
(59-64)	0.32***	0.06	0.05	0.15**	0.07	0.02
(64-69)	0.36***	0.07	0.06	0.19***	0.07	0.03
(69-74)	0.45***	0.07	0.08	0.28***	0.07	0.05
(74-79)	0.42***	0.07	0.07	0.25***	0.08	0.04
(79-84)	0.38***	0.08	0.06	0.21**	0.08	0.04
(84-89)	0.32***	0.05	0.05	0.14	0.09	0.02
Sexo						
Mujer	-0.02	0.03	-0.00	-0.01	0.03	-0.00
Estado Civil						
Soltero	0.06*	0.03	0.01	0.04	0.03	0.01
Divorciado	-0.06	0.04	-0.01	-0.06	0.04	-0.01
Viudo	0.07	0.04	0.01	-0.07	0.04	-0.01
Hijos Hogar	-0.02	0.01	-0.00	-0.02	0.01	-0.00
Situación Laboral Cabeza Familia						
Jubilado	0.07*	0.03	0.02	0.04	0.03	0.01
Trabajador Cuenta Propia	0.13***	0.03	0.03	0.06*	0.03	0.02
Desempleado	0.04	0.04	-0.01	0.02	0.04	0.01
Inactivo	-0.02	0.04	-0.01	-0.06*	0.03	-0.02

Continúa en la próxima página

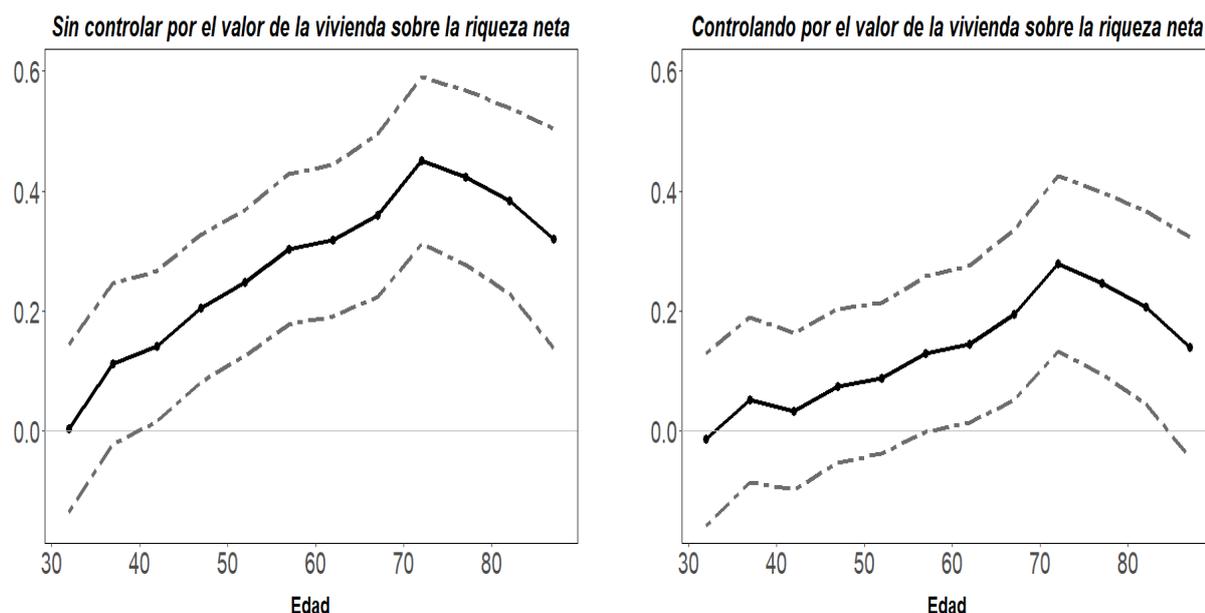
<i>Continuación tabla anterior</i>						
	Modelo sin controlar por variable vivienda			Modelo controlado por variable vivienda		
	<i>Coficiente</i>	<i>Error estándar</i>	<i>Efectos Marginales Medios</i>	<i>Coficiente</i>	<i>Error estándar</i>	<i>Efectos Marginales Medios</i>
Situación Laboral Cónyuge						
Jubilado	-0.06**	0.03	-0.01	-0.03	0.02	-0.01
Trabajador Cuenta Ajena	-0.10***	0.04	-0.02	-0.10**	0.04	-0.02
Trabajador Cuenta Propia	-0.02	0.03	-0.00	-0.02	0.03	-0.00
Desempleado	-0.00	0.04	-0.00	0.01	0.04	0.00
Nivel Estudios Cabeza Familia						
Bachillerato	0.23***	0.03	0.04	0.23***	0.03	0.04
Universitarios	0.35***	0.03	0.07	0.35***	0.03	0.07
Nivel Estudios Cónyuge						
Bachillerato	0.05*	0.04	0.01	0.04	0.03	0.01
Universitarios	0.09***	0.03	0.02	0.08***	0.03	0.02
Aversión al Riesgo						
Poco	0.38***	0.12	0.08	0.38***	0.13	0.08
Algo	0.52***	0.05	0.13	0.53***	0.05	0.14
Bastante	0.41***	0.02	0.10	0.39***	0.02	0.09
Ahorro Último Año						
Sí (ahorro positivo)	0.12***	0.02	0.02	0.09***	0.02	0.02
Plan Ahorro Futuro						
Aumentar Ahorro	0.04	0.04	0.01	0.05**	0.02	0.01
Disminuir Ahorro	0.05***	0.02	0.01	0.05***	0.02	0.01
log (Renta)	0.15***	0.02	0.03	0.14***	0.02	0.02
Ola						
2002	0.06***	0.03	0.01	0.03	0.03	0.01
2005	0.11***	0.03	0.02	0.09***	0.03	0.02
2008	-0.03	0.03	-0.00	-0.05*	0.03	-0.01
2011	-0.01	0.03	-0.00	-0.02	0.03	-0.00
2014	-0.01	0.03	-0.00	-0.02	0.03	-0.01

*Coficientes estimados por MCO con pesos muestrales mediante imputación múltiple para la EFF de 2002, 2005, 2008, 2011, 2014 y 2017, disponemos de 30.048 observaciones para cada imputación. La variable dependiente es el peso que tienen los activos con riesgo en la cartera de inversión del hogar. Los errores estándar mostrados son errores estándar bootstrap usando pesos de replicación. Adicionalmente, junto a los coeficientes estimados se indica con ‘***’, ‘**’ y ‘*’ los coeficientes estimados que son estadísticamente significativamente diferentes de 0 a un nivel de confianza del 1 %, 5 % y 10 % respectivamente.*

Entre los resultados de las estimaciones obtenidos (Tabla 6), destacamos que el coeficiente estimado para la variable peso de los inmuebles sobre la riqueza neta, que incluimos en el modelo controlado por la variable vivienda, es estadísticamente significativamente diferente de cero a un nivel de significación del 1%, y se relaciona negativamente con el peso de los activos con riesgo en la cartera de inversión.

En cuanto a la edad, los coeficientes estimados para todas las categorías que comprenden edades superiores a los 39 años son estadísticamente significativamente diferentes de 0 al 5% en el modelo sin controlar por la variable vivienda. Por el contrario, igual que ocurría en la regresión MCO algunas categorías dejan de ser estadísticamente significativas en el modelo controlado por la variable vivienda. Concretamente, solamente son estadísticamente significativamente diferente de cero a un nivel de significación del 5% los coeficientes estimados para las categorías (59-64), (64-69), (69-74), (74-79) y (79-84) en el segundo modelo. Mostramos la Figura 12 con el objetivo de analizar detalladamente la relación entre la edad y la variable dependiente entre ambos modelos:

Figura 12: Relación entre edad e inversión en activos con riesgo: coeficientes estimados por Tobit para las categorías de la variable “edad”



Intervalos de confianza y coeficientes estimados por Tobit en el modelo sin controlar por la variable peso de inmuebles sobre la riqueza neta (variable vivienda), y en el modelo controlado por el peso de los inmuebles sobre la riqueza neta.

La Figura 12 representa, con las líneas de color negro, los valores que toman los coeficientes estimados para las distintas categorías de la variable “edad”, siendo la referencia de dicha categoría el intervalo de edades entre los 24 y 29 años. Además, también se dibujan (líneas de color gris) los intervalos de confianza al 95% tanto para el modelo que incluyen la variable vivienda, como para los que no la incluyen.

Como podemos observar, los resultados obtenidos son similares a los presentados en el modelo lineal: existe una relación cuadrática entre el peso de los activos con riesgo sobre la riqueza financiera total y la edad que toca su máximo con el coeficiente estimado para la categoría (69-74), los coeficientes más bajos están estimados con menos precisión, y en su mayoría, no podemos rechazar que sea significativamente iguales a cero a un nivel de confianza del 95 %.

Por otro lado, la aversión al riesgo del hogar, además de presentar coeficientes estimados estadísticamente significativamente distintos de cero a un nivel de significación del 1 % en ambos modelos, tiene los efectos marginales más grandes de entre todas las variables.

Al igual que sucedía para el caso de los hogares de Estados Unidos en el trabajo de Flavin y Yamashita (2011), con datos españoles recogidos por la EFF también se encuentra que el peso de la vivienda sobre la riqueza neta es una variable estadísticamente significativa para explicar el peso de los activos con riesgo sobre la riqueza financiera. Por lo que para un mismo nivel de aversión al riesgo y del resto de variables de control, el peso de los inmuebles sobre la riqueza neta se relaciona negativamente con el peso de los activos con riesgo sobre la riqueza financiera total. Por el contrario, la edad del cabeza de familia deja de ser estadísticamente significativa para explicar la variable dependiente cuando se incluye el peso de los inmuebles sobre la riqueza neta en los modelos econométricos estimados. Sin embargo, en los modelos econométricos sin controlar por la variable vivienda, la edad hace de *proxy* del peso de los inmuebles sobre la riqueza neta, actuando como variable de confusión.

7. Conclusiones

Este trabajo se enmarca en la línea de estudio sobre las decisiones de inversión que realizan los hogares con el ciclo de vida. En concreto, utilizamos la Encuesta Financiera de las Familias para analizar la composición de las carteras de inversión de los hogares españoles a lo largo del ciclo de vida, prestando especial atención a la relación existente entre inversión en vivienda y activos con riesgo.

En el análisis descriptivo encontramos diferencias en el comportamiento de los hogares españoles con la edad. En primer lugar, observamos que la inversión en vivienda principal es baja para los hogares más jóvenes, crece rápidamente con la edad y se mantiene estable después de alcanzar el máximo cuando el cabeza de familia tiene entre 60 y 70 años. En contraste, la inversión en vivienda distinta de la vivienda principal, al igual que el peso de los activos con riesgo en las carteras de inversión, no se mantiene estable una vez alcanzado el máximo, sino que decrece, dibujando una forma de U invertida con la edad.

A pesar de que la evidencia empírica muestra una relación cuadrática entre inversión en activos arriesgados y la edad, los modelos estándar de asignación de carteras sostienen que la inversión en activos con riesgo que realizan las familias depende del rendimiento esperado de estos activos por encima del tipo de interés libre de riesgo, de su volatilidad y de la aversión al riesgo, pero no de la edad. Por ello, Flavin y Yamashita (2011) proponen un modelo teórico que se ajusta a la evidencia empírica observada. En este modelo, el valor de la inversión en vivienda es una variable de estado en el problema de elección de la inversión en otros activos del hogar.

Del mismo modo que Flavin y Yamashita (2011) para Estados Unidos, encontramos evidencia compatible con el modelo teórico que proponen utilizando datos de la población española. En este contexto, mediante regresiones lineales y modelos Tobit, encontramos que la edad y la inversión en activos arriesgados dejan de tener una relación estadísticamente significativa cuando se controla por el valor de la inversión en vivienda sobre la riqueza neta del hogar. Por el contrario, en los modelos que no incluyen la variable valor de la inversión en vivienda sobre la riqueza neta del hogar, la edad actúa como *proxy* de esta, generando un efecto de confusión negativo.

Con el objetivo de motivar futuros estudios sobre este tema, puesto que la Encuesta Financiera de las Familias se elabora a nivel comunitario. Resulta interesante llevar a cabo una comparación y contrastación del modelo teórico propuesto por Flavin y Yamashita (2011) en diferentes países, considerando sus diferencias y similitudes con España. Estas investigaciones podrían proporcionar una perspectiva más amplia sobre los factores, en especial la vivienda, que influyen en las decisiones de inversión en activos con riesgo a nivel global.

Referencias bibliográficas

- [1] Attanasio, OP: *Cohort Analysis of Saving Behavior by US Households. The Journal of Human Resources*. 33 (3). Summer: 575-609. 1998.
- [2] Banco de España. *Encuesta Financiera de las Familias (EFF) 2014: métodos, resultados y cambios desde 2011*. Boletín Económico, (MAR):1-38, 2017.
- [3] Banco de España. *Encuesta Financiera de las Familias (EFF) 2017: Métodos, resultados y cambios desde 2014*. Boletín económico-Banco de España, 2019.
- [4] Barceló, Cristina: *Imputation of the 2002 Wave of the Spanish Survey of Household Finances (EFF)*. Banco de España Research Paper No. OP-0603, 2006.
- [5] Becker, Thomas A y Reza Shabani: *Outstanding Debt and the Household Portfolio*. The Review of financial studies, 23(7):2900-2934, 2010.
- [6] Cameron, A Colin y Pravin K Trivedi: *Microeconometrics: Methods and Applications*. Cambridge university press, 2005.
- [7] Campbell, John Y: *Household Finance*. The journal of finance, 61(4):1553-1604, 2006.
- [8] Chetty, Raj, László Sándor y Adam Szeidl: *The Effect of Housing on Portfolio Choice*. The Journal of Finance, 72(3):1171-1212, 2017.
- [9] Cocco, Joao F: *Portfolio Choice in the Presence of Housing*. The Review of Financial Studies, 18(2):535-567, 2005.
- [10] Crespo, Laura: *La Encuesta Financiera de las Familias: Métodos y Usos. Seminario sobre Uso de Encuestas por los Bancos Centrales-CEMLA Banco de España. Mimeo*. 2021.
- [11] Curcuru, Stephanie, John Heaton, Deborah Lucas y Damien Moore: *Heterogeneity and Portfolio Choice: Theory and evidence*. En *Handbook of financial econometrics: Tools and techniques*, páginas 337-382. Elsevier, 2010.
- [12] Deaton, Angus: *The Analysis of Household Surveys: a Microeconomic Approach to Development Policy*. World Bank Publications, 1997.
- [13] Flavin, Marjorie y Shinobu Nakagawa: *A Model of Housing in the Presence of Adjustment Costs: A Structural Interpretation of Habit Persistence*. American Economic Review, 98(1):474-95, 2008.

- [14] Flavin, Marjorie y Takashi Yamashita: *Owner-occupied Housing: Life-Cycle Implications for the Household Portfolio*. American Economic Review, 101(3):609–14, 2011.
- [15] Fougere, D y M Poulhes: *The Effect of Housing on Portfolio Choice: A Reappraisal Using French Data*. CREST Working Paper, 2014.
- [16] Gomes, Francisco, Michael Haliassos y Tarun Ramadorai: *Household Finance*. Journal of Economic Literature, 59(3):919–1000, 2021.
- [17] Koenker, Roger y Gilbert Bassett Jr: *Regression Quantiles*. Econometrica: journal of the Econometric Society, páginas 33–50, 1978.
- [18] Koenker, Roger y Kevin F Hallock: *Quantile Regression*. Journal of economic perspectives, 15(4):143–156, 2001.
- [19] Koenker, Roger y Jose AF Machado: *Goodness of Fit and Related Inference Processes for Quantile Regression*. Journal of the american statistical association, 94(448):1296–1310, 1999.
- [20] Long, J. S.: *Regression Models for Categorical and Limited Dependent Variables*. *Advanced quantitative techniques in the social sciences*, 7, 1997.
- [21] Lumley, Thomas: *Complex Surveys: a Guide to Analysis Using R*. John Wiley & Sons, 2011.
- [22] Paz-Pardo, Gonzalo: *Homeownership and Portfolio Choice over the Generations*. ECB Working Paper No. 2021/2522, 2021.
- [23] Rubin, Donald B: *Multiple Imputation for Nonresponse in Surveys*. Hoboken. John Wiley & Sons, 1987.
- [24] Solon, Gary, Steven J Haider y Jeffrey M Wooldridge: *What are we weighting for?* Journal of Human resources, 50(2):301–316, 2015.
- [25] Stock, James H y Mark W Watson: *Introducción a la Econometría*. Número 330.1543 S8. Pearson, 2012.
- [26] Tobin, James: *Estimation of Relationships for Limited Dependent Variables*. Econometrica: journal of the Econometric Society, páginas 24–36, 1958.
- [27] Vestman, Roine: *Limited Stock Market Participation and Home Ownership*. Informe técnico, Citeseer, 2012.
- [28] Winkelmann, Rainer: *Econometric analysis of count data*. Springer Science & Business Media, 2008.

[29] Wooldridge, Jeffrey M: *Econometric Analysis of Cross Section and Panel Data*. MIT Press, 2010.

[30] Wooldridge, Jeffrey M: *Introductory Econometrics: A Modern Approach*. Cengage learning, 2015.