Initial empirical results from SCOMP data

July 27, 2025

In this document I will present what we are learnign from out empirical work.

Before presenting our results it is important to clarify the sample selection criteria. We always first store the file without the sampling selection and then a second file with the sample selection.

1. We use the data in "1.solicitudes" to obtain the savings, age and other buyer characteristics. We store two files "1_solicitudes" with all the requests and then another "1_solicitudes_yytoyyRV" with initial and final years and only the requests for annuities.

2. I use the data in "2.ofertas" CURRENTLY WORKING ON BEING ABLE TO USE THE WHOLE DATA AND NOT ONLY THE REDUCED SAMPLE

   The file '2_ofertas_muestra_sol' contains a sample of offers for which requests were made, but it is not useful, because some of the requests lead to no offers or to offers that were not accepted. Hence we only use the file '2_ofertas_muestra_acep'.

   In the file '2_ofertas_muestra_acep' we do the following sample selection: 1) kept only 'cod_modalidad_pensi == 1' which are *RV inmediata*

3. **Aceptaciones** just dropped sec_beneficiario and kept one obs. per id_certificado_saldo because there was one per sec_beneficiario. NOT CLEAR WHAT SEC_BENEFICIARIO MEANS

4. **Clasificacion de riesgo** no sample selection

5. In beneficiarios we do not do any sample selection.

For our main analysis we use the following sample selection criteria:

Our sample consists on 8176 individuals and 497,000 offers, hence indiiduals receive on average 61 annuities offers. This offers differ on the number of guaranteed months and the withdrawal amount (ELD: excedente de libre disposicion). Hence we restrict our sample to offers with 0 guaranteed months and 0 ELD. Then the average individual receives 13 offers.

	stats				
	c1	c2	c3	c4	c5
r1	8176	497045	60.79318	111577	13.64689

1 IE 0

Figure 1 shows the increase of requests over time. The increase is smooth with the exception of the year 2009 to 2010 that almost doubles. Probably there was a regulatory change. On average consumers make 10.5 requests for different financial products.

CHECK WHY THE INCREASE IS SO BIG FROM 2009 TO 2010.

Important to note that individuals

Figure 1

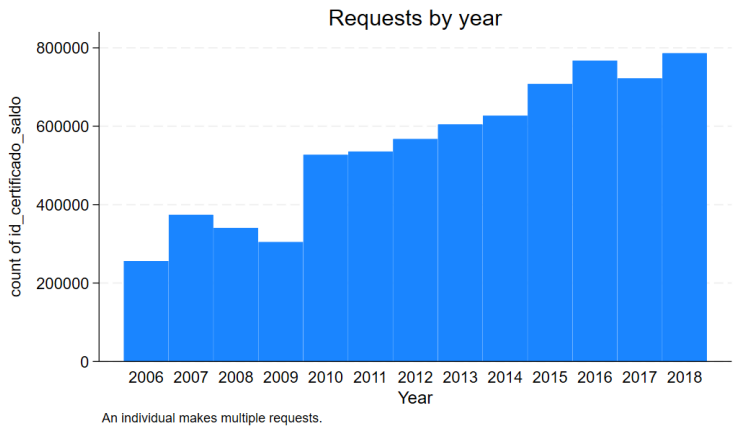


Figure 2 shows the amount of savings of individuals in the sample. Where 1UF is around 40 USD. The distribution is truncated at the 99th percentile of savings.

Figure 2

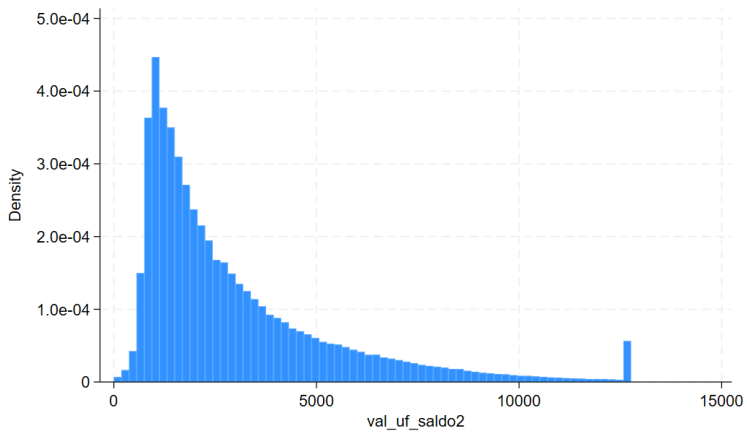


Figure **??** shows the number of requests for each type of financial product in absolute terms and then their respective shares. The changes are not particularly large, it seems like the share of each financial product is relatively stable over time.

NOTE: ANNUITIES WITH PW(GREEN) START FROM 0 IN 2006 HENCE PROBABLY THEY ARE A FINANCIAL INNOVATION.

Figure 3

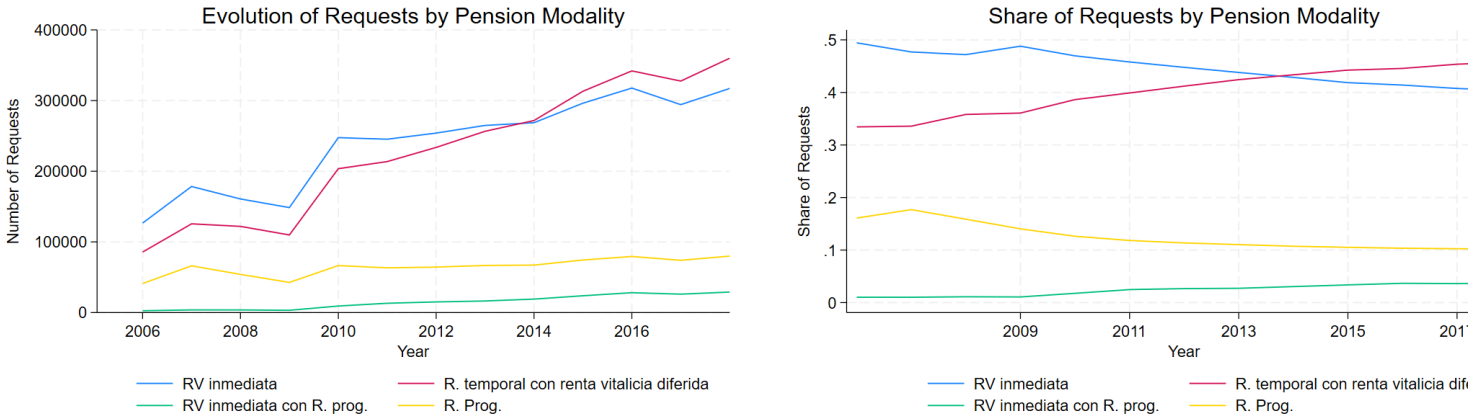


Figure 4 shows the shares for the whole sample as a function of savings. Richer individuals tend to buy more annuities with PW and less PW. This is explained by the fact that there is a minimum amount of savings required to buy an annuity and also could have to do with a higher life expectancy.

Figure 4

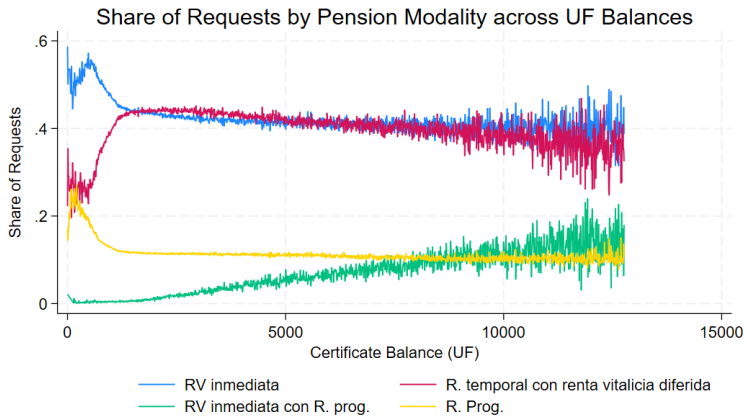
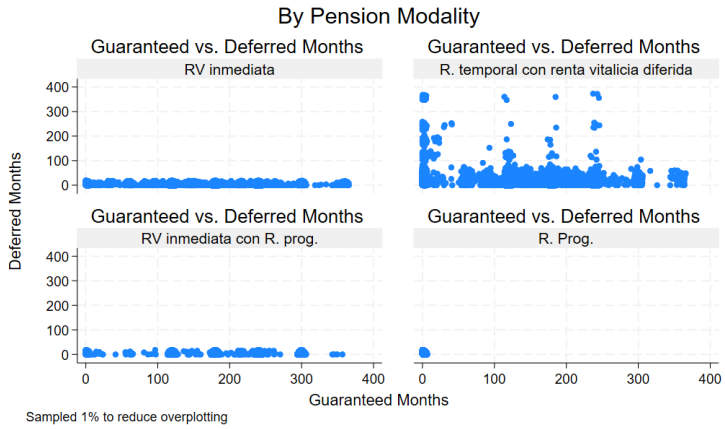


Figure 5 shows how many guaranteed months people buy.

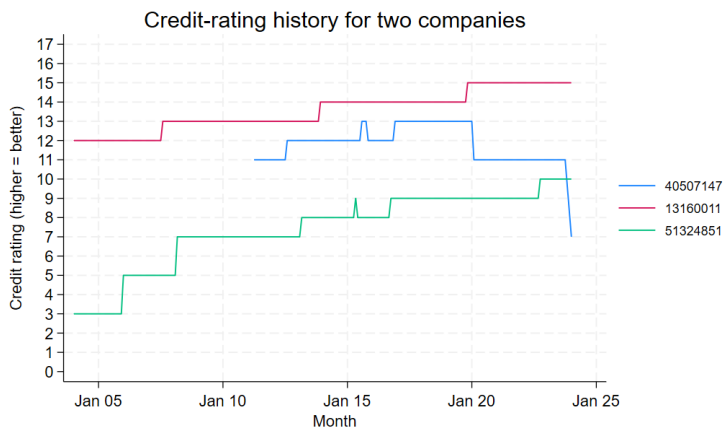
Figure 5



2 IE 1

Figure 6 the history of the credit rating for selected insurers.

Figure 6



Only 2% of the offers constitute external offers, because individuals request offers from many financial products (make many requests), then receive many offers (one for each offering company) for each request, and then in case of requesting external offers they do it for only a subset of the initial offers.

3 IE1b

This code just takes all the data of the offers (not only a sample) and creates .dta files where each of them contains a chunck of the data. Then filters the files to keep only the annuities and some years and finally joins all this filtered files into one big file with the offers.

4 IE 3

Figure 7 shows the number of external offers that buyers request and figure 8 shows the distribution disaggregated by year.

Figure 7

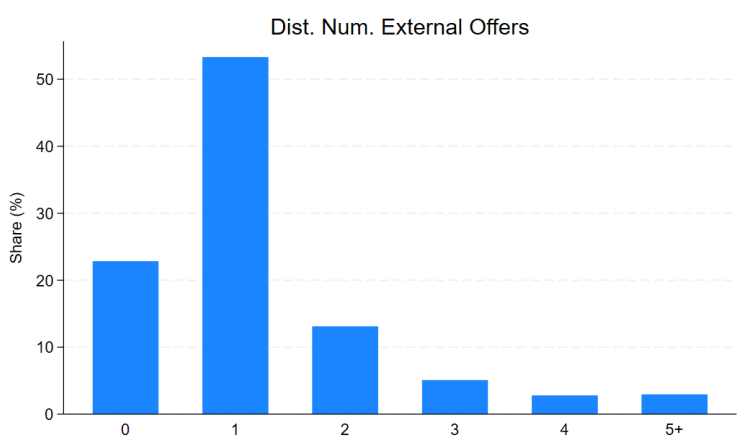


Figure 8

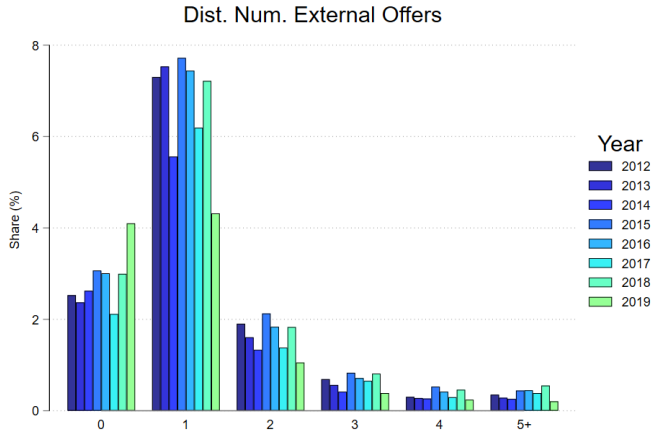


Figure 9 shows

Figure 9

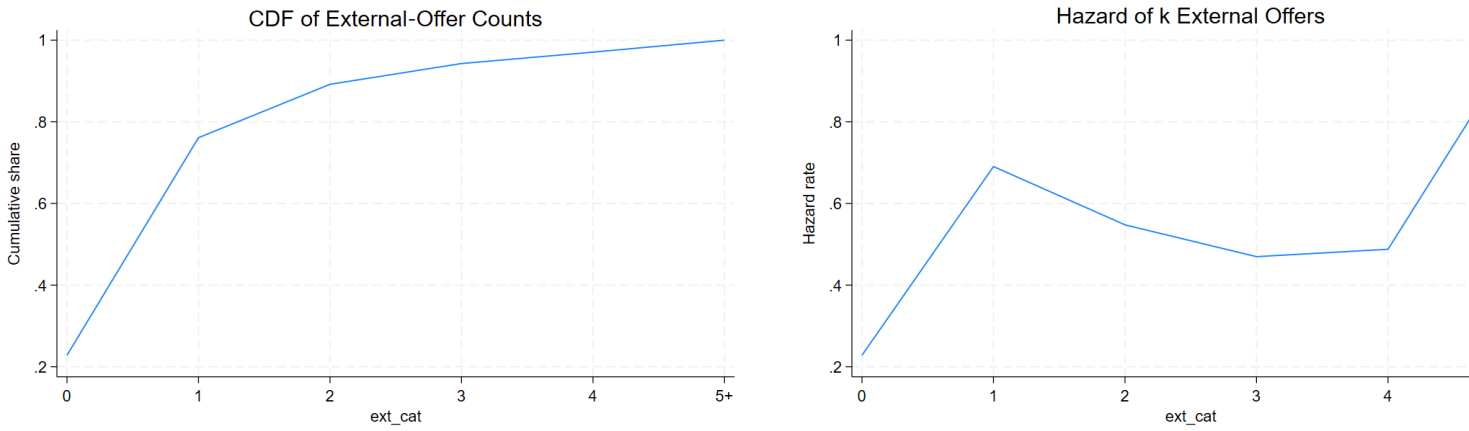


Figure 10 shows the average number of searches for individuals grouped by their quintile of savings, which is a proxy of income. Specifically an increase of savings by a standard deviation is related to 0.30 additional searches.

Figure 10

Figure 11 shows the CDF and hazard rate of searches grouped by savings quintile.

Figure 11

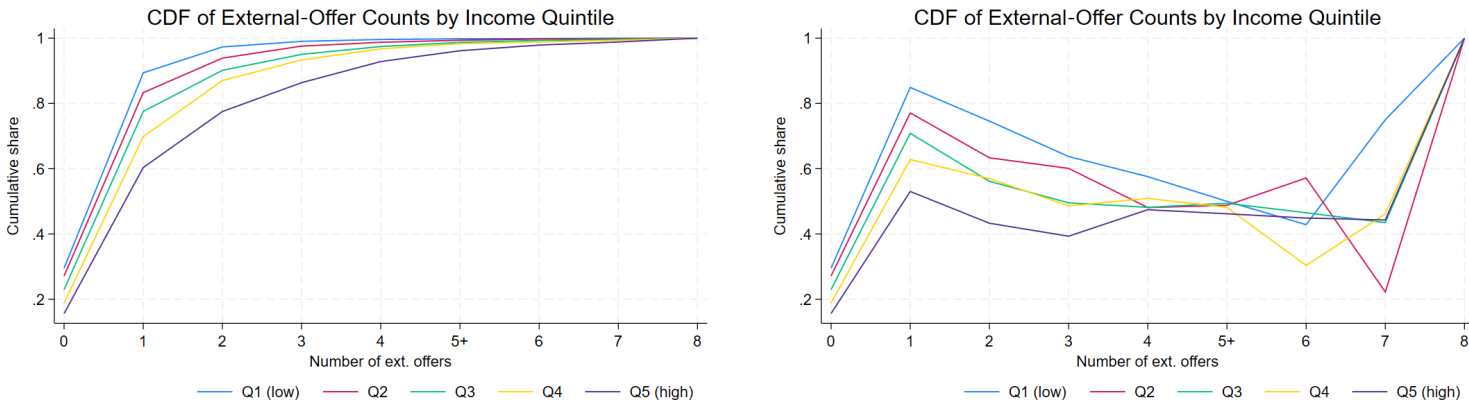
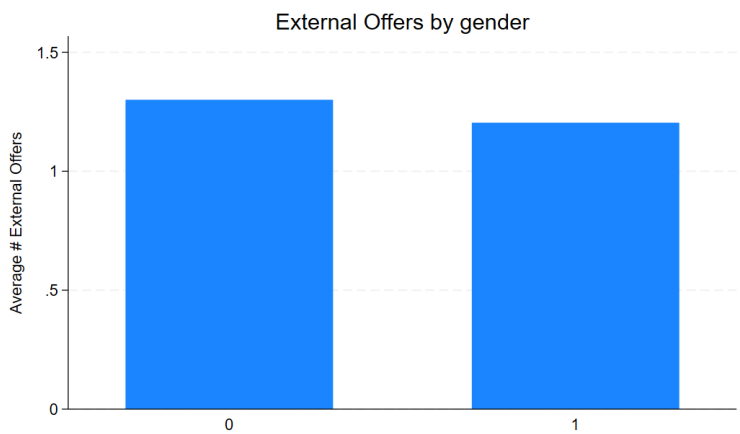


Figure 12 shows the average number of searches for individuals grouped by their gender. Men do on average -0.10 more searches, which does not seem to be economically significant, although it is statistically significant.

Figure 12



4.1 Dispersion in offers within group

The insurers in the SCOMP platform observe the age, gender and savings and same zip code? of the individual. We want to see whether individuals with similar characteristics receive similar offers.

Given the sparsity of the distribution of individuals it is difficult to find individuals with exactly the same characteristics. Hence we created three criteria to define a group:

1. Individuals within the 5-year age interval, same savings quintile and same gender who receive offers the same year by same firm. we call them group 1

2. Indivduals with the same age, gender and savings. who receive offers the same year by same firm. we call them group 2

3. Indivduals with the same age, gender and savings. we call them group 3

where the first criteria is less sparse than the second, and the second is less sparse than the third. Note that using criteria 1 we could group a man of age 60 and savigs of 954 and and an individual of age 64 and savings of 1089, hence certain dispersion is expected. To reduce the role of savings on the offer dispersion we define the ratio as the savings didivded by the offer.

Then we studied the dispersion of the offers within each group.

The following table presents the summary statistics for dispersion variables. z_offer is the percentual dispersion of the offers within groups formed by criteria 1, and then z_offer2 and z_offer3 for groups 2 and 3. Note that the variation is reduced once we define the groups using a finer criteria.

We exclude the external offers from the analysis since we expect them to be different from the internal offers.

In row 3 we show the standard deviation of the offers using criteria 2. The average deviation is .51UF which is around 20USD. This dispersion could be justified by intermediaries, ZIP code, particular day, etc.

Table I: Summary Statistics

	(1)				
	mean	sd	min	max	count
z_offer	.1991379	.1166944	0	.8129947	173923
z_offer2	.1205333	.0795119	0	.8659106	171572
z_offer3	.0638504	.0510833	0	.3931563	60530
sd_offer3	.5127416	.5968534	0	10.04091	60530
N	173964				

Finally we run a regression of the offer on the group fixed effects. The coefficients are shown below. When using the finer group definition almost all the variation in offers is explained by the group fixed effects.

In terms of modeling, the previous findings suggest that we can assume that the insurers look at savings, gender and age and send offers based on these characteristics.

	Group1	Group2	Group3
Val UF Pension	.71	.84	.99

4.2 dispersion within group for external and internal offers.

I used the 3 criteria to define groups presented previously and the within group standard deviation of the internal and external offers was not very different. I also run a regression of the offer on group fixed effects for the sample of internal and external offers and the R-squared was similar.

explain why I was expecting more diespersion for external offers. and under which conditions this could be true. I had previously written "If there is revelation of information in the aftermarket then the R2 of the regression of the initial offers would be lower than for the external offers, since there would be information that firms are considering when bidding but which we do not have in our data. but is not the case"

4.3 Others

4.3.1 Dispersion of the choice set

Figure 13 shows the distribution of standard deviations of the choice set of the consumers and the left panel shows the distribution of the range, both of them normalized by the mean of the offers

each individual receives. The table below shows the summary statistics of both variables. Offers have an average deviation from the mean of the offers that represenet a 1.7% of the mean offer and the range represent a 6% of the mean offer. Considering that this are the savings of their lifetime, this differences translate into considerable absolute differences.

Figure 13

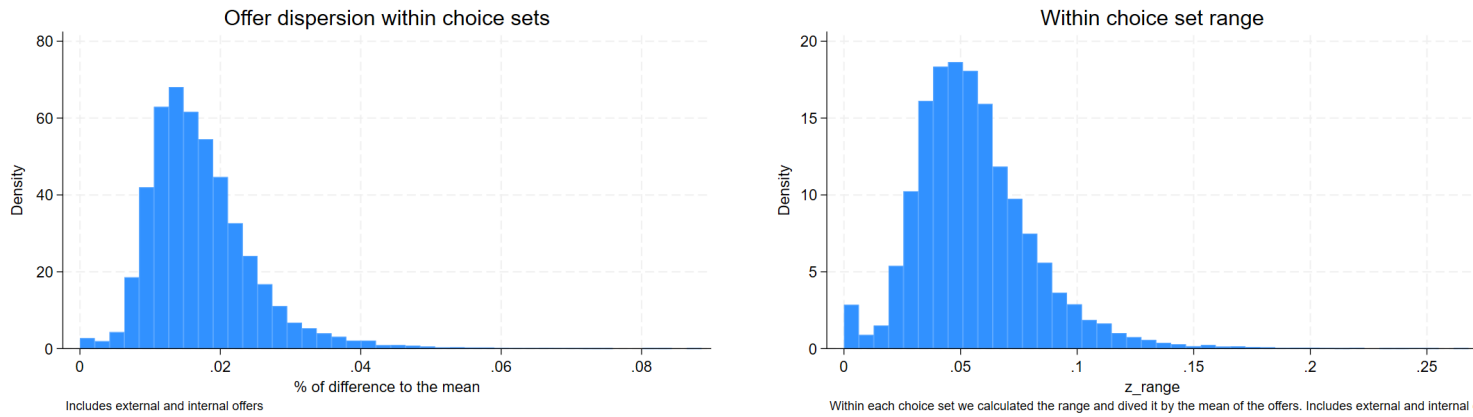


Table II: Summary Statistics

	(1)				
	mean	sd	min	max	count
diff_pct	.0175359	.0073122	0	.0885337	228942
z_range	.0599199	.0251982	0	.2678275	229102
N	229102				

4.3.2 Improvement when asking for external offers

The table below shows the improvement with respect to the initial offer made by the same company when asking for an external offer. The first row shows summary statistics for the absolute improvement, with a mean of .24 UF which is around 10 USD. The second row shows the improvement in percentage terms. The third column calculates the present value of the improvement, using the assumption that the individual lives 20 years and the monthly interest rate is 0.3% which has been the average rate of return of the pension funds in the last 20 years [1]. Finally, to understand the dimension we take the previously calculated present value and divide it by the wage of the last month of the individual, on average the improvement in the offer corresponds to the salary of almost 4 months of work.

Figure 14 shows the summary statistics of the distribution of the different measures of improvement desaggregated by gender. On average men get a higher absolute improvement, but since they also have higher initial offers (due to their higher savings) women get a higher improvement in percentage terms.

[1] source note that the rate of return is in real terms because all the flows are in UF.

Table III: Improvement when searching

	(1)				
	mean	sd	min	max	count
improvement_abs	.2456159	.2925267	-.6700001	9.050003	17331
improvement_pct	1.777408	1.313942	-10.48513	15.43956	17331
improvement_PV20	41.97763	49.99504	-114.5081	1546.714	17331
improvement_wage	3.900746	36.68072	-2.413689	2802.884	17236
N	17331				

Figure 14

sexo		e(impro~)	e(impro~)	e(imp~20)	e(impro~)
F					
	mean	.2306832	1.961764	39.42552	3.733815
	sd	.2103937	1.313165	35.95788	31.30257
	min	.0099983	.0270957	1.708789	.0215787
	max	4.280006	15.28358	731.4857	1281.804
	count	7362	7362	7362	7327
M					
	mean	.2566436	1.641264	43.86234	4.02418
	sd	.3402902	1.297877	58.15818	40.19849
	min	-.6700001	-10.48513	-114.5081	-2.413689
	max	9.050003	15.43956	1546.714	2802.884
	count	9969	9969	9969	9909
Total					
	mean	.2456159	1.777408	41.97763	3.900746
	sd	.2925267	1.313942	49.99504	36.68072
	min	-.6700001	-10.48513	-114.5081	-2.413689
	max	9.050003	15.43956	1546.714	2802.884
	count	17331	17331	17331	17236

Figure 15 shows the summary statistics of the distribution of the different measures of improvement desaggregated by income quintile. Higher quintiles get bigger improvements (col1) but in terms of the percentual improvement there does not seem to be a clear pattern (col2). Finally almost all groups get an increase which in PV corresoponds to around 5 months of work, with the exception of the first quintile, were buyers get only 2 months of work on average.

Figure 15

by categories of: income_q

income_q	e(impro~)	e(impro~)	e(imp~20)	e(impro~)
Q1(low)				
mean	.0647068	1.636535	11.05889	2.454564
sd	.0529108	1.330991	9.042865	7.808938
min	.0099998	.1814837	1.709034	.0243112
max	.29	6.779655	49.5632	85.45383
count	631	631	631	623
Q2				
mean	.0970513	1.710333	16.5868	4.077004
sd	.0897887	1.501962	15.34558	23.24705
min	-.6700001	-10.48513	-114.5081	.0215787
max	.4499998	7.269787	76.90839	358.906
count	312	312	312	307
Q3				
mean	.1575732	1.791593	26.93046	8.384937
sd	.1240831	1.361736	21.20674	84.2988
min	.0099998	.0929389	1.709034	.0448822
max	.9499998	9.443337	162.3622	1281.804
count	239	239	239	236
Q4				
mean	.2808503	2.013786	47.99944	5.575082
sd	.1941725	1.358816	33.18555	48.08287
min	.0100002	.0374821	1.709115	.0230029
max	1.440001	10.11236	246.107	751.9944
count	247	247	247	247
Q5(high)				
mean	.4817462	1.564664	82.33409	5.447399
sd	.4221773	1.307268	72.15332	21.33151
min	.0100002	.0282068	1.709115	.0215797
max	3.840004	15.28358	656.2859	194.0033
count	315	315	315	315
Total				
mean	.1891571	1.711435	32.32839	4.54436
sd	.256725	1.372678	43.87625	38.73872
min	-.6700001	-10.48513	-114.5081	.0215787
max	3.840004	15.28358	656.2859	1281.804
count	1744	1744	1744	1728

4.3.3 Improvement when asking for external offers wrt to highest offer

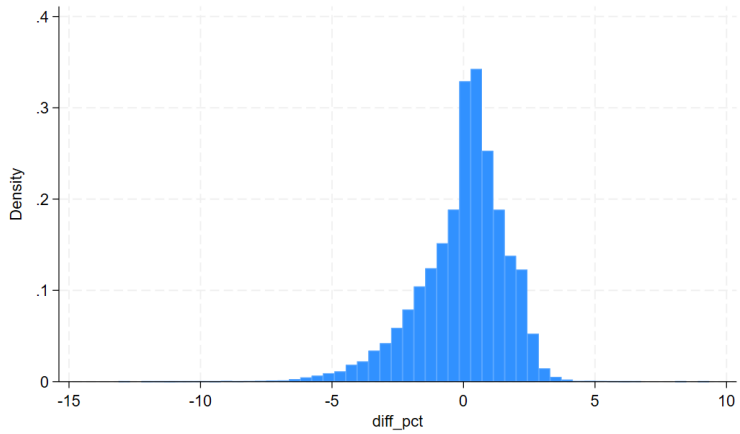
The table below shows the improvement with respect to the highest initial offer when asking for an external offer. Sometimes the external offer is higher and sometimes lower than the higher internal offer but on average they are the same (the highest initial offer is only .03% higher than the external offer), but with considerable dispersion, see figure 16.

Table IV: External offer vs. highest initial offer

	(1)				
	mean	sd	min	max	count
diff	-.0339503	.4220157	-16.45	2.610001	162667
diff_pct	-.0275759	1.684481	-13.11523	9.338141	162667
N	162667				

hola

Figure 16



4.3.4 Number of initial offers

Almost all the insurers bid in the first round. , see table below.

Table V: Dist. of number of initial offers

	(1)				
	mean	sd	min	max	count
internal_offers	10.53644	3.204842	1	31	19799
N	19799				

4.3.5 Discrete choice determinants

The share of buyers who end up buying an annuity that is not the highest offer is: 0.462 . Moreover when they do not buy the highest offer they choose offers which are significantly lower. Figure 17 shows the distribution of the highest offer minus the chosen offer (foregone amount) as a share of the accepted offer in percentage terms.

And the table below shows in the first row the average foregone percentage among the buyers who did not choose the highest offer. 1.5% might not seem like a lot, but if we assume that the buyer has 20 years of life left and we use a .03% monthly interest rate [2] it means they are foregoing almost 20 UF in present value terms, which when divided by the last monthly salary of the individual is around 1.6 months of salary.

[2] previously explained why is sensible

Figure 17

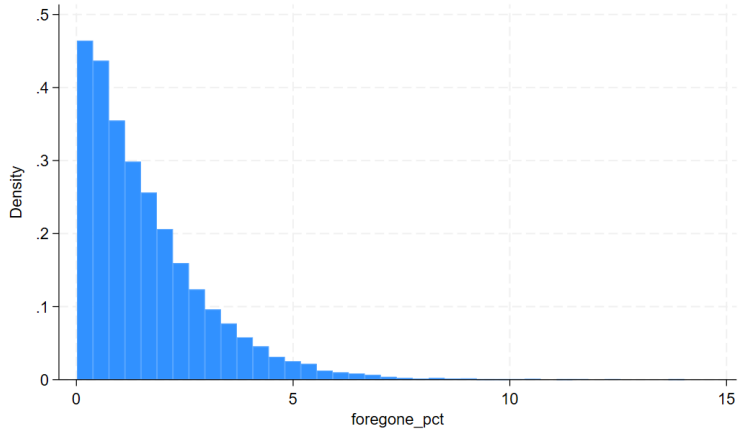


Table VI: Foregone pension

	(1)				
	mean	sd	min	max	count
foregone_pct	1.537301	1.386817	.0132041	14.029	109090
foregonePV20	19.85657	60.06566	0	2811.43	229102
foregone_wage	1.579052	16.94421	0	1452.715	227894
N	229102				

One possibility is that buyers perceive insurers as differentiated,for example in conversations with a financial advisor they mentioned the importance of risk rating of the insurer[3], the customer service or the brand appeal of the company.

Table VII: Acceptance (clogit) Results

	(1)
accepted2	
Standarizd amount	13.956***
	(0.144)
Standarized risk score	1.253***
	(0.044)
Obs.	228,287

Standard errors in parentheses
Variables are standarized at the choice set level
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

To assess the possibility that buyers perceive insurers as differentiated, we run a conditional logit regression on the choice set of the buyers, and as expected we observe that although the primary

[3] See conversation with Claudia Ayala (9 jul 25), where she says: *la clasificación en una Compañía de Seguros es muy importante porque demuestra la espalda financiera de cada una de ellas, como se encuentra en posición al resto, sucursales, spot venta, es decir un servicio completo, EN EL FONDO QUE NOS ENTREGA ... UNA TRANQUILIDAD EN EL PAGO DE PENSIONES FUTURAS*

determinant of the choice is the amount of the offer, the risk rating of the insurer does affect the choice.

4.3.6 Death and external offers

When bargaining for an external offer one can always disclose personal health information. Possibilities

- Sick individuals are more likely to ask for external offers

  The first table below shows the tabulation of individuals who die in our data (rows) with a dummy for whether they asked for external offers(cols). Then the second table shows the share of individuals who negotiate an external offer by their current status. Individuals who already died also are less likely to ask for external offers.

Table VIII: Cross-tabulation of Death Status and External Status

	0 b	1 b	Total b
0	3453	11729	15182
1	309	996	1305
Total	3762	12725	16487

  Then when running a logit model of the probability of asking for an external offer on whether the buyer is dead and year fixed effect, we find that already dead individuals asked for fewer external offers.

  to icnrease the sample size I can also use all the other annuities

Table IX

	(1)				
	mean	sd	min	max	count
has_ex_alive	.7725596	.4191931	0	1	15182
has_ex_dead	.7632184	.4252701	0	1	1305
N	16487				

Table X: Acceptance (clogit) Results

	(1)
has_external	
dead	-0.070***
	(0.021)
val_uf_saldo	0.000***
	(0.000)
Obs.	229,102

Standard errors in parentheses
Includes year fixed effects
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

- Sick individuals get higher increases when bargaining

  One could think that less healthy individuals when bargaining with firms can reveal their lower life expectancy and hence get higher increases in the offers. This appears not to be the case, conditional on bargaining dead and alive individuals get similar increases in the offers.

  To explore this possibility we first run a regression of the percentual improvement of the initial offer when the buyer approaches the insurer to bargain, we observe that individuals who already died in our data actually get lower increases in the offers.

  Then we run a regression of the improvment on the dead dummy and year fixed effects to account for possible trends in negotiation rates in our data. The coefficient is negative which suggests that actually healthier individuals get higher increases in the offers.

Table XI: T-test: Improvement Percentage by Death Status

	Mean Dead=0	Mean Dead=1	Difference	Std. Error	t-statistic	p-value
improvement_pct	1.790	1.614	0.176***	0.039	4.531	0.000
N	17331					

*** p¡0.01, ** p¡0.05, * p¡0.1

Table XII: Effect of bad health on negotiation gains

	(1)
dead	-0.154***
	(0.039)
Obs.	17,331

Standard errors in parentheses
Includes year fixed effects
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.4 Supply side

4.4.1 Market share of insurers

Figures 18, 19 and 20 show the market share of the insurers among different groups of the population. For example in figure 18 we see that the market share of insurer 827... is higher among older buyers

whereas insurer 201... has just the opposite pattern. In figure 19 we see that insurer 997... sells mostly to rich consumers whereas insurer 585... sells mostly to poor consumers.

Figure 18

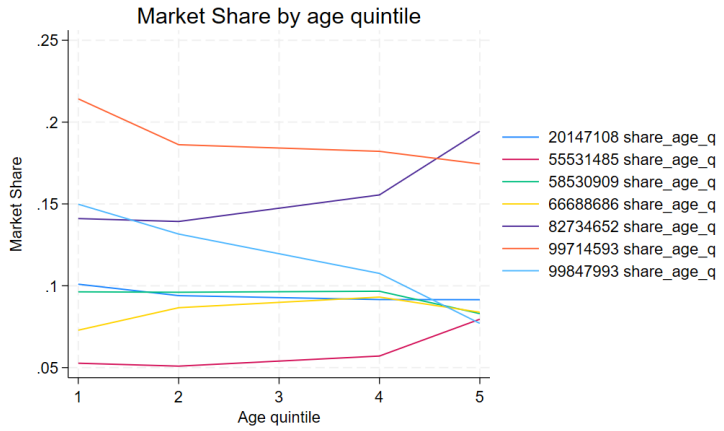


Figure 19

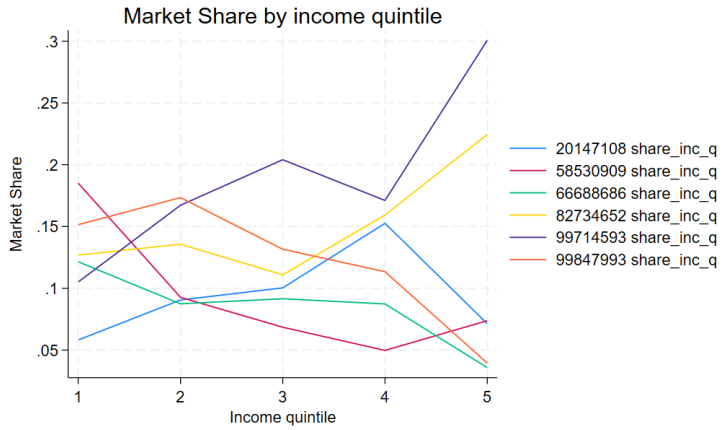
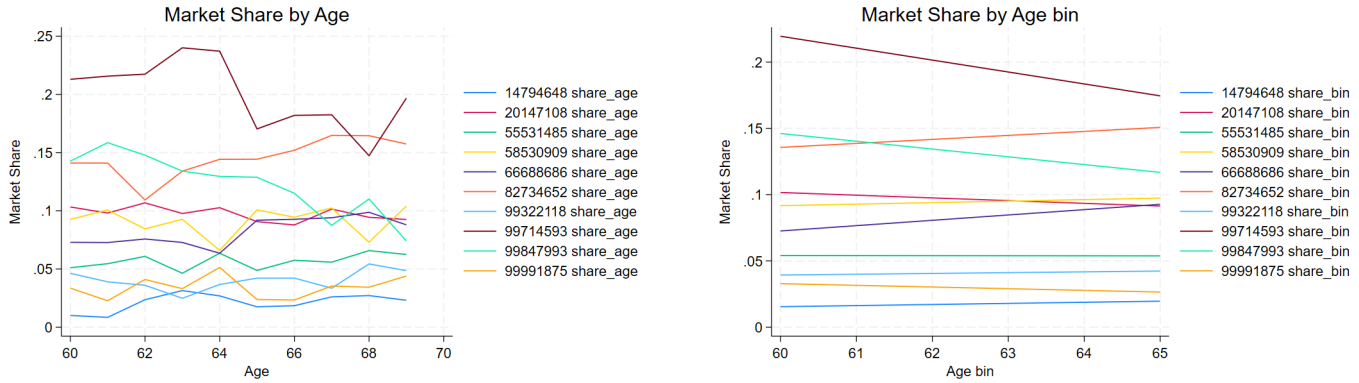


Figure 20



4.4.2 Probability of offer

Figures 21, 22 and 23 show the probability an insurer sends an offer to each of the different segments of buyers.

For example in figure 21 we see that the insurer 1479... almost always makes an offer, whereas insurer 997... targets the income quintiles 2,3 and 4, and the probability of making an offer to quintile 1 is essentially zero.

There does not appear to be so much targeting by age, see figure 22 and 23 where the lines are more horizontal than in the case of income. But what is striking is that there is high heterogeneity in the probability of making an offer at all, in our sample firm 345... makes an offer in around 20% of the cases, whereas firm 1479... makes an offer in almost all the cases. One possibility is that the previous pattern is caused by firms not being present all the years in our data, but when they are active they always make an offer. To check this we replicate figure 23 but by year for the years 2012-2015. We see that this pattern holds for the individual years.

THE BIG QUESTION IS WHY DO SOME FIRMS MAKE OFFERS ONLY TO SOME SEGMENTS OF THE POPULATION? AND EXACTLY TO WHAT SEGMENTS? IT COULD BE DEMAND DRIVEN (SOME SEGMENTS PREFER INSURER A, HENCE INSURER B DOES NOT EVEN TRY TO SELL THEM) OR SUPPLY DRIVEN. BUT THERE SEEMS TO BE SOMETHING INTERESTING HERE.

Figure 21

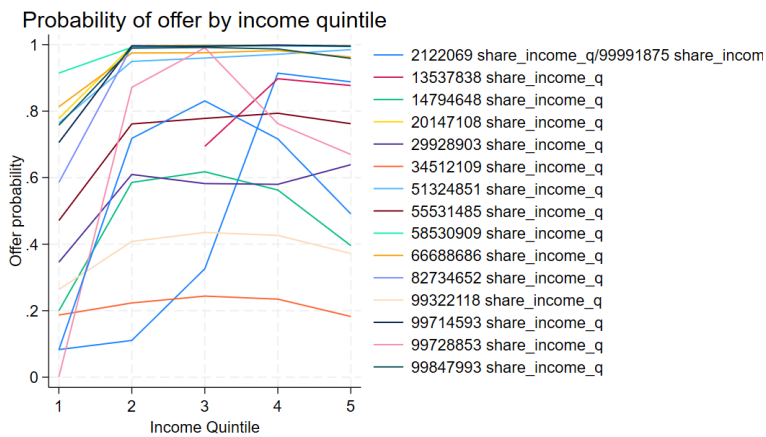


Figure 22

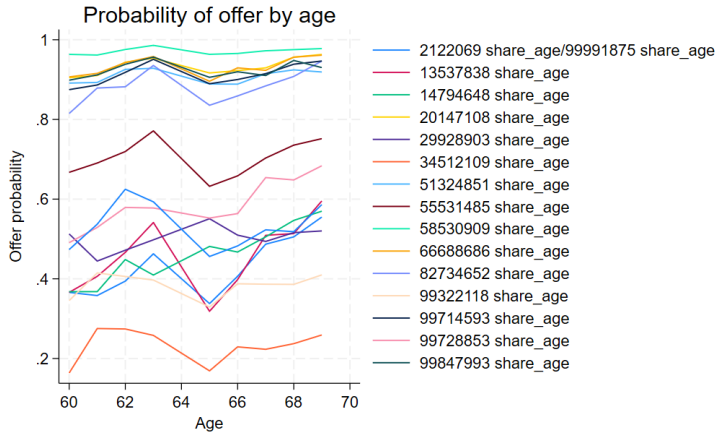


Figure 23

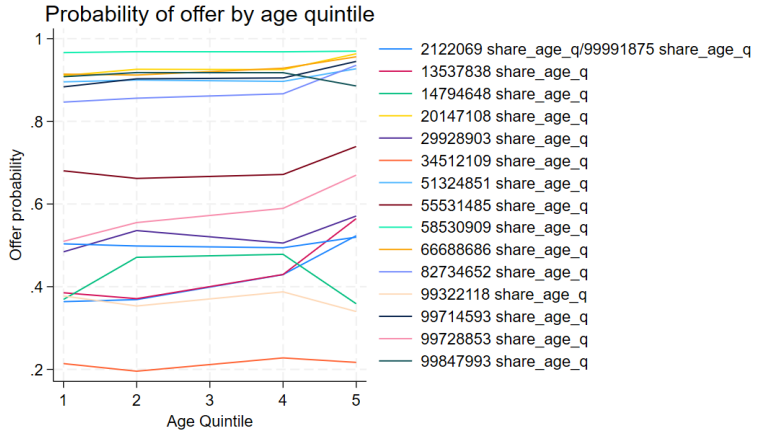
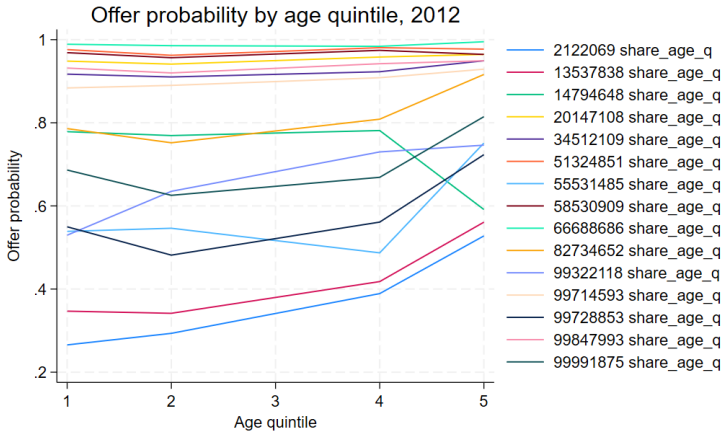


Figure 24



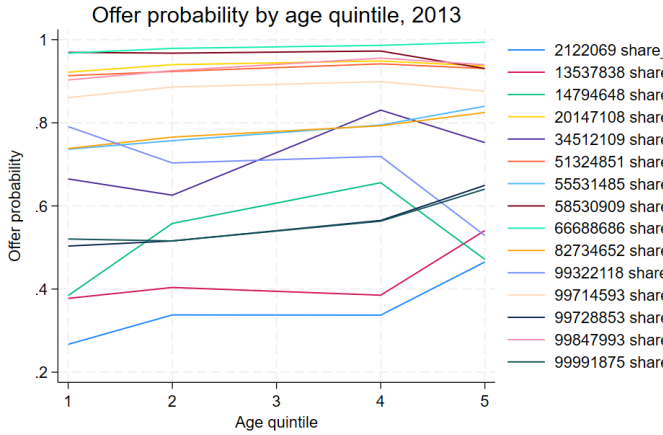
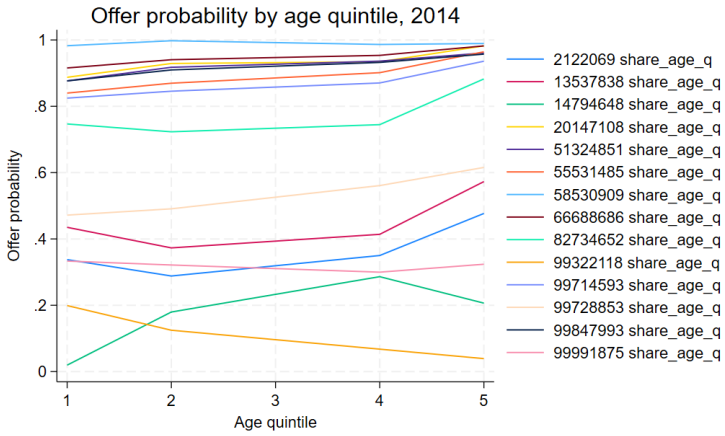
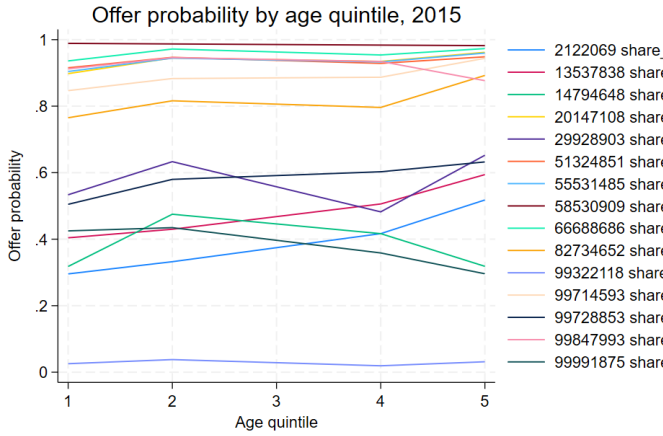



Figure 25