Initial empirical results from SCOMP data

October 3, 2025

In this document I will present what we are learnign from out empirical work. This is the continuation of the file which presents the initial datawork.

1 IE 5

1.1 Granularity of the insurer mortality tables

Summary in general it appears as if bigger firms use more granular mortality tables. but the evidence is not super robust. For example when creating buckets based on age, gender, savings, and date of purchase, there is still variation in the ratio of flow payment to stock of savings, it is not clear whether it is because the buckets are not granular enough or there is some noise due to comissions and also because interest rates can change over time even within a month.

The logic in table 1 is that more sophisticated firms should have a statistically significant value of savings (val_uf_prima) because they include it in the mortality tables. Whereas less sophisticated firms do not include it and therefore the coefficient should be statistically insignificant. We rank firms according to sales and we show the firms with odd ranks (1,3,5,7,9,..., 15), and we assume is a proxy of sophistication[1] There does not appear to be a trend in which firms with lower ranks have a smaller value of the savings coefficient. This could be caused by the fact that the model is misspecified, we are trying to approximate the pricing formula which contains thousands of coefficients (mortality rates) with 5 coefficients.

For example if the age has a non-linear effect, let's say age squared is statistically significant, and at the same time age squared is correlated with savings, then the coefficient of savings could be significantly different from zero even if the firm is not using savings in the mortality tables.

In table 2 we repeat the excercise but without controlling for savings under the logic that the R2 of more sophisticated firms should be lower because we are missing a key variables in the pricing formula. There does not appear to be a trend in which firms with lower ranks have a higher R2. Again, this could be caused by misspecification. Also in table 3 we show that adding the savings in the regresssion increases the R2 ver little (around 0.001-0.002) for two particular firms. Hence it appears as if the big determinant of ratio is gender and age and the savings are not a first order importance determinant, which could explain that even if only some firms use the savings in the mortality tables, we do not see a difference in the amount the R2 changes when we include savings in the regression, in Table 2.

In section 5 we run regressions for the determinants of ratio for all firms, ranked by their market share (first 3 tables). It seems like bigger firms have more significant coefficients for savings. To corroborate the this hypothesis in the fourth table of the section we run a regression with an interaction

---

[1] If creating better mortality tables is a fixed cost then it is natural to expect firms with more sales to have more granular mortality tables.

between savings and firm buckets, where the Nth column has N buckets. For example the variable Size2 takes the value of 1 for the biggest 8 firms and 2 for the smallest 8 firms (there are a total of 16 firms), similarly for size3, size4 and size5. Generally biggger firms have a higher coefficient for savings, which is consistent with the hypothesis that bigger firms have more granular mortality tables.

Table I: ratio on controls by rank_sales

	(1) ratio	(2) ratio	(3) ratio	(4) ratio	(5) ratio	(6) ratio	(7) ratio	(8) ratio
val_uf_prima2	0.2***	1.4***	0.4***	0.3***	0.3***	1.0***	0.1*	0.1
	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)	(0.1)
male	-8.8***	-10.7***	-9.9***	-7.5***	-8.2***	-10.7***	-7.7***	-6.6***
	(0.3)	(0.3)	(0.3)	(0.3)	(0.3)	(0.4)	(0.4)	(0.4)
age_years	-5.7***	-5.5***	-5.8***	-6.1***	-5.9***	-5.5***	-5.6***	-5.7***
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.1)	(0.1)	(0.1)
A	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)
D	-1.0***	-1.3***	-1.5***	-0.7**	-0.7**	-0.4	-0.2	0.2
	(0.3)	(0.3)	(0.3)	(0.3)	(0.3)	(0.4)	(0.3)	(0.4)
P	-0.1	-0.3	-0.2	-1.0**	0.1	-0.2	0.1	0.0
	(0.4)	(0.3)	(0.4)	(0.4)	(0.4)	(0.5)	(0.4)	(0.5)
Constant	577.1***	563.8***	587.4***	601.6***	587.7***	559.2***	564.0***	573.0***
	(3.0)	(2.9)	(2.8)	(3.0)	(2.8)	(3.6)	(4.6)	(3.2)
N	18266.0	20151.0	21324.0	13485.0	19945.0	11870.0	10148.0	9475.0
R-sq	0.7	0.7	0.6	0.7	0.7	0.6	0.7	0.7

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table II: ratio on controls by rank_sales

	(1) ratio	(2) ratio	(3) ratio	(4) ratio	(5) ratio	(6) ratio	(7) ratio	(8) ratio
male	-8.9***	-11.3***	-10.0***	-7.6***	-8.3***	-9.6***	-7.8***	-6.6***
	(0.3)	(0.3)	(0.3)	(0.3)	(0.3)	(0.4)	(0.4)	(0.4)
age_years	-5.7***	-5.3***	-5.7***	-6.0***	-5.9***	-5.3***	-5.5***	-5.7***
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.1)	(0.1)	(0.1)
A	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	(.)	(.)	(.)	(.)	(.)	(.)	(.)	(.)
D	-1.0***	-0.9***	-1.4***	-0.6**	-0.6**	-0.2	-0.2	0.2
	(0.3)	(0.3)	(0.3)	(0.3)	(0.3)	(0.4)	(0.3)	(0.4)
P	0.0	0.1	-0.1	-0.9**	0.2	-0.0	0.1	0.0
	(0.4)	(0.4)	(0.4)	(0.4)	(0.4)	(0.5)	(0.4)	(0.5)
Constant	575.3***	552.2***	583.8***	599.7***	585.5***	552.4***	563.0***	572.5***
	(2.9)	(2.9)	(2.7)	(2.9)	(2.8)	(3.6)	(4.5)	(3.2)
N	18266.0	20151.0	21324.0	13485.0	19945.0	11870.0	10148.0	9475.0
R-sq	0.7	0.6	0.6	0.7	0.7	0.6	0.7	0.7

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

	(1) ratio	(2) ratio	(3) ratio	(4) ratio
ym_now	-2840.6***	-2829.0***	0	0
	(-31.63)	(-31.49)	(.)	(.)
ym_now2	0	0	0	0
	(.)	(.)	(.)	(.)
ym_now3	0.0129***	0.0128***	0.000900**	0.000901**
	(32.11)	(31.97)	(3.26)	(3.27)
ym_now4	-0.0000193***	-0.0000193***	-0.00000216**	-0.00000216**
	(-32.34)	(-32.21)	(-3.28)	(-3.29)
ym_now5	8.71e-09***	8.68e-09***	1.38e-09***	1.38e-09***
	(32.57)	(32.44)	(3.30)	(3.31)
male	-8.300***	-8.014***	-5.928***	-5.885***
	(-23.47)	(-22.29)	(-11.62)	(-11.25)
val_uf_prima2		1.349***		0.593
		(4.15)		(0.87)
val_uf_prima_pol2		-0.215***		-0.166
		(-3.52)		(-1.01)
val_uf_prima_pol3		0.0120**		0.0197
		(3.22)		(1.45)
val_uf_prima_pol4		-0.000208**		-0.000519
		(-3.12)		(-1.54)
Constant	758074.4***	754647.2***	-19825.8*	-20992.5**
	(31.42)	(31.27)	(-2.50)	(-2.65)
Observations	18266	18266	6151	6151
R-squared	0.670	0.671	0.671	0.673

t statistics in parentheses

A second approach is to:

1. Generate groups based on covariates $x = (age, savings, year)$ and $\hat{x} = (x, savings)$

2. Calculate the coefficient of variation for groups within the groups based on x and $\hat{x}$.

3. Variation in groups using x should be higher for more sophisticated firms since we are not accounting for savings. Moreover if we calculate the decrease in variation when using $\hat{x}$ instead of x then we should see a decrease for sophisticated firms but not for the other firms.

2 IE 6: Differences in Hazard rates across firms.

In this section we prove that there are diffferences in hazard rates between companies and they are not only explained by differences in observables.

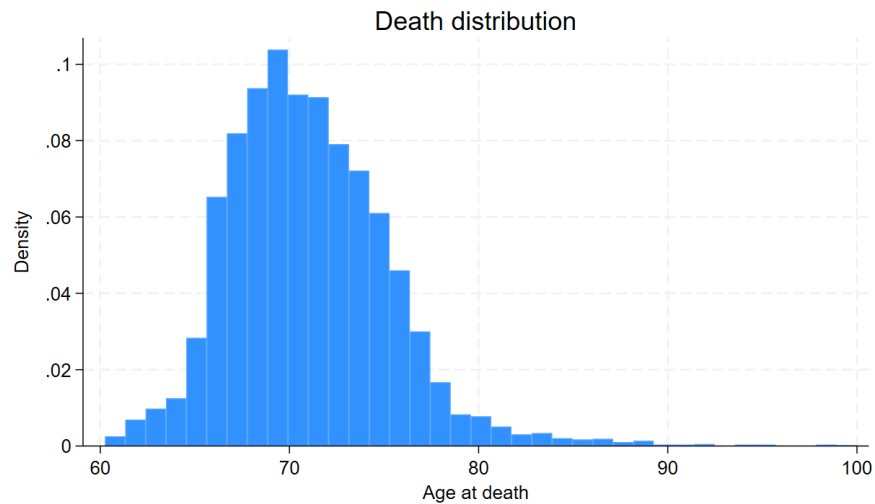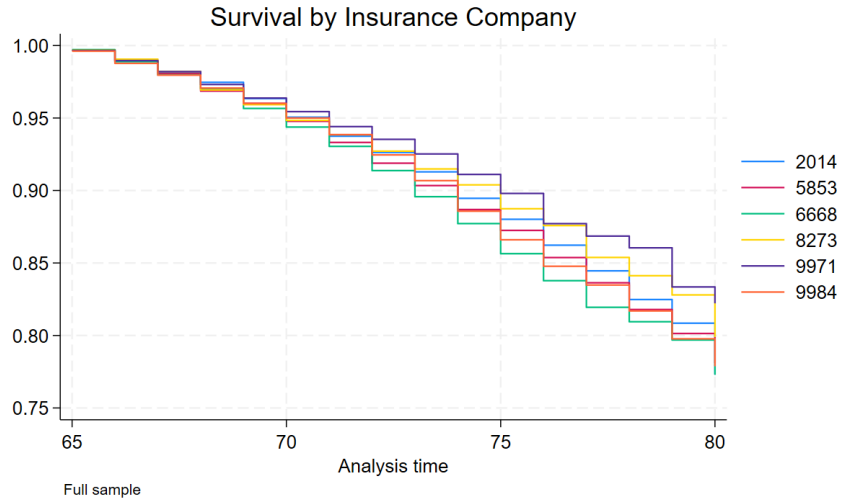Figure 1 shows the distribution of age at death.

Figure 1

Figure 4 shows the survival function by each firm. There are differences among firms.

Figure 2

Survival by Insurance Company

Full sample

One possibility is that the differences represent differences in observable characteristics, not on unobservables.

As a first step we look only at the sample of men (to control for gender differences) and we also group buyers in 3 income groups (to control for income differences).
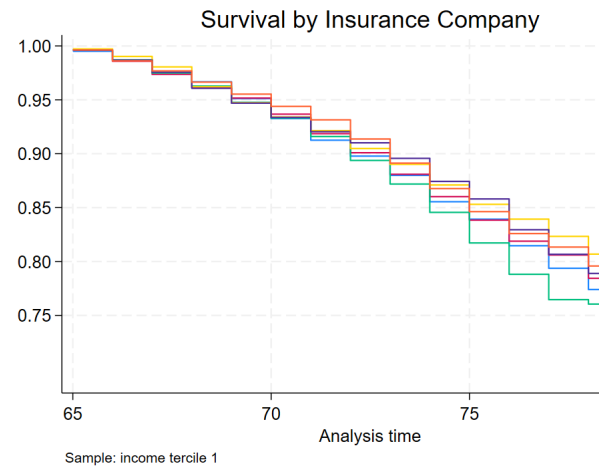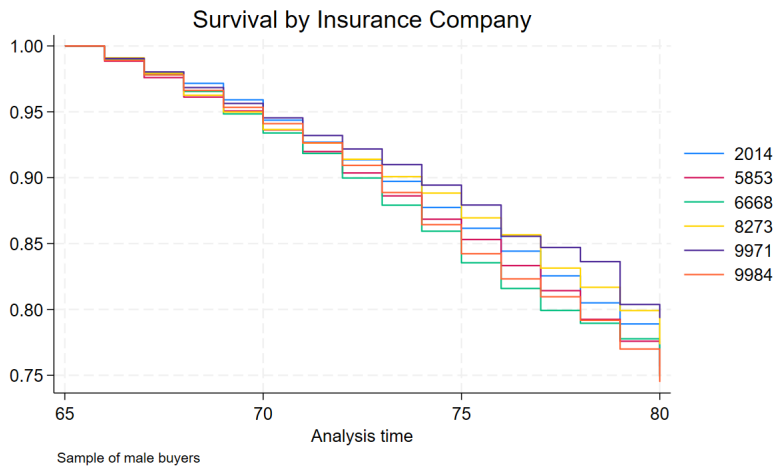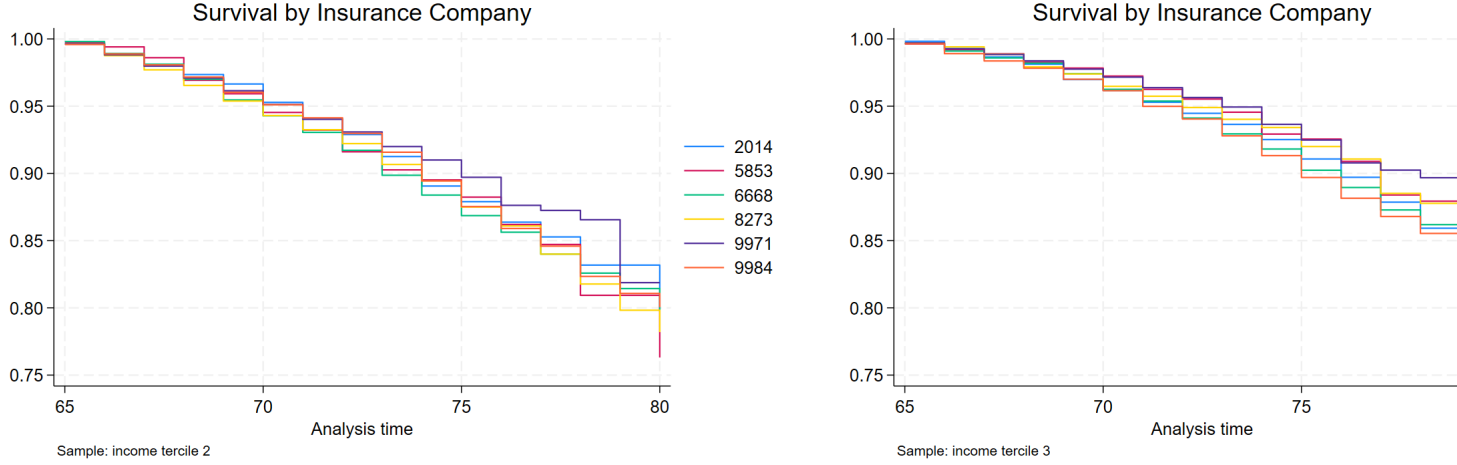
Figure 3

Survival by Insurance Company

Sample of male buyers

Survival by Insurance Company

Sample: income tercile 1

Figure 4

Sample: income tercile 2

Sample: income tercile 3

The differences once we use this subsamples continue being significant, the exception is income group 2 for which they are not significant any more (p-value 0.15).

To do a more formal test we use a Cox proportional hazard model.

We estimate Cox PH models of the form

$$h_i(t \mid X_i) \;=\; h_0(t)\, \exp\!\big(X_i^\top \beta\big),$$

where $h_0(t)$ is an unspecified baseline hazard and X collects the screened covariates (sex, premium level, intermediation type, guaranteed-months dummies). We then add company indicators to test for across-firm differences:

$$\textbf{Baseline PH:} \qquad h_i(t \mid X_i, \mathrm{Firm}_i) = h_0(t)\, \exp\!\big(X_i^\top \beta + \gamma_{\mathrm{Firm}_i}\big).$$

To relax proportionality for the *covariates* (but not for firm dummies), we also estimate a model with time-varying coefficients using $\ln t$ interactions (Stata's `tvc()` with `texp(ln(_t))`):

$$h_i(t) = h_0(t)\, \exp\!\big(X_i^\top \beta + (\ln t)\, X_i^\top \delta + \gamma_{\mathrm{Firm}_i}\big).$$

This allows the effect of X to change with time while still testing whether firms differ by a *proportional* shift.

After each Cox fit with firm dummies, we run a Wald test of

$$H_0: \;\; \gamma_2 = \cdots = \gamma_J = 0,$$

we get 0.01 and 0.01 as the p-values.

3 IE 7: learning or discriminating

It is not clear why firms increase the amount offered when making an external offer.

Two possible explanations are that 1) firms are learning about the offers made by competitors and they want to increase their offer to match teh competitors or 2) buyers who ask for external offers are more price elastic and firms are trying to price discriminate.

Given that people in the industry say that the pricing formula equalizes the savings to the discounted value of payments, the first hypothesis does not appear to be very likely.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Increase	Increase	Increase	Increase	Increase	Increase	Has External Offer
main							
Avg. Gap	0.316***	0.155***	0.155***	0.139***	0.147***	0.071***	
	(0.006)	(0.010)	(0.010)	(0.016)	(0.019)	(0.020)	
Max. Gap		0.110***	0.110***		-0.021	-0.006	
		(0.009)	(0.009)		(0.029)	(0.028)	
gap_from_avg							-0.191***
							(0.032)
Constant	1.893***	1.375***	1.375***	1.381***	1.387***	1.511***	-2.012***
	(0.010)	(0.082)	(0.082)	(0.045)	(0.046)	(0.121)	(0.028)
Observations	14133	14133	14133	2046	2046	2046	16164

Average: is the difference between the mean of other firms' initial offers and own initial offer

Max Gap: is the difference between the highest other firm's initial offer and own initial offer.

Cols (1)-(3) use the population of initial offers that are not the highest, (4)-(6) only use the highest offer

Cols (4) and (6) include firm fixed effects

In columns 1 to 3 we can see that firms when making an external offer increase their offer more if they see that the initial offers of the competitors are higher. In columns (4) to (6) we also see that the firm that made the highest initial offer increases their offer as a function of the average of the other offers but not as a function of the maximum of the other offers.

The issue is that if one thinks about the pricing game as a sort of Bertrand with imperfect information, where firms want to just overbid the competitors, then the firm with the highest initial offer might just not submit an external offer. Hence there is an endogeneity issue. In column (7) we run a logit regression to see what is the effect of seeing the competitors offers in the probability of the highest firm making an external offer. The coefficient is negative meaning that the higher the gap (in absolute value) the lower the probability of making an external offer, which is consistent with the idea of the firm trying to just overbid the competitors.

Figures 5 and 6 show visually the positive relationship between the gap to other offers and the increase in the external offer.

Figure 5

Learning: Improvement vs Gap from Others' Average

7

Figure 6

Learning from Average

Figure 7

Distribution of Improvement (%)
Testing for Bunching at Zero

Distribution of 100*(external offer/ highest initial offer from competitor -1). Black line at zero.

Manipulation Testing Plot
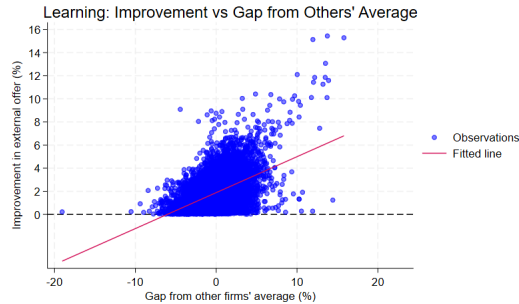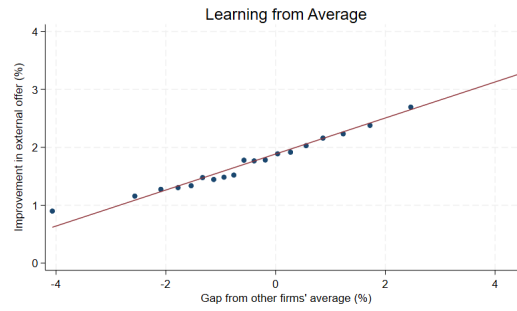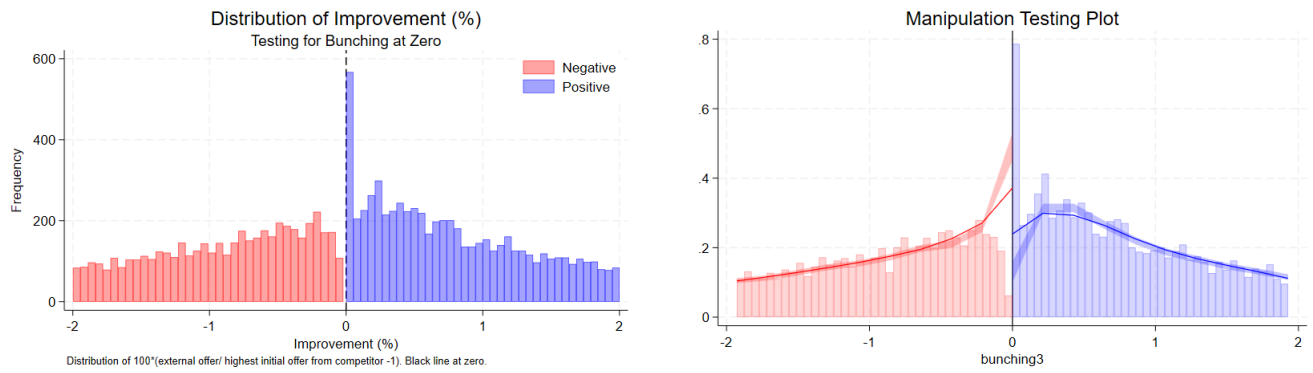
Figure 7 shows that visually there appears to be bunching around the cutoff where the cutoff is the amount the external offer has to increase to be higher than any of the initial offers.

4 IE10:

In the table below we can observe that some firms are correlated in their offers. It is not clear what to make out of it.

Some firms have consistently more similar algorithms.

Zoffer3	1.000	-0.040	-0.184	-0.141	-0.187	-0.050	0.079	-0.137
p-val	.	0.017	0.000	0.000	0.000	0.000	0.000	0.000
Zoffer6	-0.040	1.000	-0.238	-0.045	-0.218	-0.103	0.102	-0.071
p-val	0.017	.	0.000	0.007	0.000	0.000	0.000	0.000
Zoffer7	-0.184	-0.238	1.000	-0.035	0.011	-0.182	-0.104	0.090
p-val	0.000	0.000	.	0.000	0.244	0.000	0.000	0.000
Zoffer9	-0.141	-0.045	-0.035	1.000	0.109	0.050	-0.094	-0.070
p-val	0.000	0.007	0.000	.	0.000	0.000	0.000	0.000
Zoffer11	-0.187	-0.218	0.011	0.109	1.000	0.143	-0.236	0.141
p-val	0.000	0.000	0.244	0.000	.	0.000	0.000	0.000
Zoffer13	-0.050	-0.103	-0.182	0.050	0.143	1.000	-0.288	-0.086
p-val	0.000	0.000	0.000	0.000	0.000	.	0.000	0.000
Zoffer15	0.079	0.102	-0.104	-0.094	-0.236	-0.288	1.000	0.047
p-val	0.000	0.000	0.000	0.000	0.000	0.000	.	0.000
Zoffer17	-0.137	-0.071	0.090	-0.070	0.141	-0.086	0.047	1.000
p-val	0.000	0.000	0.000	0.000	0.000	0.000	0.000	.

Table III: Correlation Matrix with P-values

In the table below we estimate a logit model of the form:

$$d_{itj} = \text{logit}(\beta_j + \gamma_j Z_{itj} + \epsilon_{itj}) \tag{1}$$

where i indexes individual, t indexes time and j indexes firms. d_{itj} is a dummy that takes the value of if the individual dies, β_j aims to account for certain firms consistently bidding more and z_{itj} is the deviation. In some cases we add a set of individual level controls X_i.

Discussion of specification

In our logit model we need to include a constant to account for the fact that in our sample people die between 5 and 10% of the years. If we do not include a constant the model will use the coefficients to match the average death rate and even in the case where the z_{itj} are not predictive we will end up with positive coefficients.

The next decision to make is whether the constant is firm-specific or not.

If firms make offers to groups of people who vary systematically in their mortality rates then we need to include firm-specific constants[2].

But the problem of including firm-specific constants is that the offering probability could be endogenous, for example if a firm predicts a low mortality and makes offers only to individuals with high predicted mortality then we would be in a problem. Hence we assume that the offer probability is exogenous, which allows us to use a common constant. In this case we would have censorizing, e.g. if the z would be too low then the firm does not offer.

when including a firm-specific constant most of the constants (11 out of 17) are significant. What does this mean? shoul we include firm-specific constants or not?

One possibility is to use only observations where many firms made an offer.

---

[2] If a firm only makes offers to women (healthier) that would already be included in the mean price hence would not be a problem.

	(1)	(2)	(3)	(4)
	Died	Died	Died	Died
Died				
rank_sales=1 × Z-score	-0.300***	-0.0523		
	(-7.04)	(-1.27)		
rank_sales=2 × Z-score	-0.146***	0.0317		
	(-3.92)	(0.83)		
rank_sales=3 × Z-score	0.0897*	0.0198		
	(2.09)	(0.48)		
rank_sales=4 × Z-score	0.191***	-0.0382		
	(5.80)	(-1.13)		
rank_sales=5 × Z-score	0.0108	-0.0266		
	(0.35)	(-0.85)		
rank_sales=6 × Z-score	0.184***	0.0617		
	(5.53)	(1.81)		
rank_sales=7 × Z-score	-0.0756	-0.0287		
	(-1.13)	(-0.42)		
rank_sales=8 × Z-score	0.447***	0.101		
	(5.28)	(1.22)		
rank_sales=9 × Z-score	0.0991	0.0358		
	(1.80)	(0.69)		
rank_sales=10 × Z-score	-0.116**	-0.0666		
	(-2.78)	(-1.58)		
rank_sales=11 × Z-score	0.0852*	0.0263		
	(2.03)	(0.61)		
rank_sales=12 × Z-score	0.0713	0.0541		
	(1.32)	(1.07)		
rank_sales=13 × Z-score	-0.130	0.0215		
	(-1.33)	(0.22)		
rank_sales=14 × Z-score	-0.175***	-0.0569		
	(-3.46)	(-1.17)		
rank_sales=15 × Z-score	-0.0824*	-0.00309		
	(-1.98)	(-0.08)		
rank_sales=16 × Z-score	-0.0570	-0.127*		
	(-0.90)	(-1.97)		
rank_sales=17 × Z-score	-0.533***	0.0235		
	(-4.71)	(0.20)		
male		0.836***		
		(39.39)		
log_savings		-0.465***		
		(-34.30)		
current_age		0.1000***		
		(54.52)		
Z-score × market_share			-0.215	
			(-1.73)	
Z-score × rank_sales				-0.000496
				(-0.38)
Observations	1536283	1536283	1536283	1536283

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

	(1)	(2)	(3)	(4)
	Died	Died	Died	Died
Died				
rank_sales=1 × Offer: monthly payment	-0.00776**	0.00777**		
	(-2.88)	(3.07)		
rank_sales=2 × Offer: monthly payment	-0.0107***	0.00818**		
	(-3.61)	(3.01)		
rank_sales=3 × Offer: monthly payment	-0.0140***	0.00725**		
	(-5.00)	(2.78)		
rank_sales=4 × Offer: monthly payment	-0.0154***	0.0102***		
	(-4.94)	(3.31)		
rank_sales=5 × Offer: monthly payment	-0.0128***	0.00727**		
	(-4.75)	(2.90)		
rank_sales=6 × Offer: monthly payment	-0.0155***	0.00680*		
	(-5.39)	(2.53)		
rank_sales=7 × Offer: monthly payment	-0.00113	0.00637		
	(-0.31)	(1.78)		
rank_sales=8 × Offer: monthly payment	-0.0171***	0.00539		
	(-3.80)	(1.36)		
rank_sales=9 × Offer: monthly payment	0.00290	0.00528		
	(0.97)	(1.59)		
rank_sales=10 × Offer: monthly payment	-0.0131***	0.00750**		
	(-4.67)	(2.86)		
rank_sales=11 × Offer: monthly payment	-0.0113*	0.0107*		
	(-2.24)	(2.32)		
rank_sales=12 × Offer: monthly payment	-0.0152**	0.0107*		
	(-3.06)	(2.22)		
rank_sales=13 × Offer: monthly payment	-0.00665	0.00783*		
	(-1.63)	(2.25)		
rank_sales=14 × Offer: monthly payment	0.0000487	0.00340		
	(0.01)	(1.02)		
rank_sales=15 × Offer: monthly payment	-0.00755*	0.00860**		
	(-2.37)	(2.90)		
rank_sales=16 × Offer: monthly payment	0.00710	0.00985*		
	(1.80)	(2.45)		
rank_sales=17 × Offer: monthly payment	-0.0261**	0.00363		
	(-2.87)	(0.47)		
male		0.825***		
		(38.70)		
log_savings		-0.583***		
		(-21.94)		
current_age		0.0972***		
		(51.26)		
Offer: monthly payment × market_share			-0.0925***	
			(-10.14)	
Offer: monthly payment × rank_sales				-0.000811***
				(-8.52)
Observations	1536359	1536359	1536359	1536359

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

	(1)	(2)	(3)	(4)
	Died	Died	Died	Died
Died				
rank_sales=1 × Z-score2	-9.986***	-0.793		
	(-4.49)	(-0.42)		
rank_sales=2 × Z-score2	-9.658***	2.723		
	(-4.12)	(1.14)		
rank_sales=3 × Z-score2	0.415	0.263		
	(0.16)	(0.12)		
rank_sales=4 × Z-score2	7.045***	-3.216		
	(4.01)	(-1.88)		
rank_sales=5 × Z-score2	2.880	-0.316		
	(1.77)	(-0.19)		
rank_sales=6 × Z-score2	8.033***	2.575		
	(4.32)	(1.41)		
rank_sales=7 × Z-score2	4.603	1.846		
	(1.22)	(0.55)		
rank_sales=8 × Z-score2	22.18***	2.070		
	(5.46)	(0.52)		
rank_sales=9 × Z-score2	-1.477	-0.340		
	(-0.51)	(-0.13)		
rank_sales=10 × Z-score2	-5.316*	-2.161		
	(-2.22)	(-0.95)		
rank_sales=11 × Z-score2	5.907*	2.049		
	(2.43)	(0.83)		
rank_sales=12 × Z-score2	1.976	4.385		
	(0.58)	(1.32)		
rank_sales=13 × Z-score2	1.731	3.829		
	(0.40)	(0.95)		
rank_sales=14 × Z-score2	-9.114***	-1.758		
	(-3.45)	(-0.74)		
rank_sales=15 × Z-score2	-2.516	0.268		
	(-1.12)	(0.13)		
rank_sales=16 × Z-score2	-5.896*	-6.621**		
	(-2.25)	(-2.62)		
rank_sales=17 × Z-score2	-33.86***	1.669		
	(-5.31)	(0.28)		
male		0.840***		
		(39.55)		
log_savings		-0.472***		
		(-34.58)		
current_age		0.0997***		
		(53.77)		
Z-score2 × market_share			-8.334	
			(-1.23)	
Z-score2 × rank_sales				-0.0550
				(-0.81)
Observations	1536359	1536359	1536359	1536359

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

	(1)	(2)	(3)	(4)
	Died	Died	Died	Died
Died				
rank_sales=1 × Z-score	-0.168***	-0.0382		
	(-5.09)	(-1.21)		
rank_sales=2 × Z-score	-0.106**	0.0392		
	(-3.21)	(1.14)		
rank_sales=3 × Z-score	0.103*	0.0198		
	(2.41)	(0.48)		
rank_sales=4 × Z-score	0.0936***	-0.0296		
	(3.40)	(-1.08)		
rank_sales=5 × Z-score	0.0386	-0.0130		
	(1.46)	(-0.49)		
rank_sales=6 × Z-score	0.0534*	0.0299		
	(2.18)	(1.20)		
rank_sales=7 × Z-score	-0.124**	-0.00610		
	(-2.78)	(-0.13)		
rank_sales=8 × Z-score	0.280***	0.0595		
	(6.70)	(1.43)		
rank_sales=9 × Z-score	0.0179	0.0278		
	(0.38)	(0.60)		
rank_sales=10 × Z-score	-0.114**	-0.0661		
	(-2.67)	(-1.57)		
rank_sales=11 × Z-score	0.0824**	0.0345		
	(2.69)	(1.11)		
rank_sales=12 × Z-score	0.0521	0.0475		
	(1.08)	(1.03)		
rank_sales=13 × Z-score	-0.165**	-0.0286		
	(-3.21)	(-0.56)		
rank_sales=14 × Z-score	-0.180***	-0.0559		
	(-3.96)	(-1.15)		
rank_sales=15 × Z-score	-0.0412	0.0121		
	(-1.16)	(0.35)		
rank_sales=16 × Z-score	0.0795	-0.108		
	(1.52)	(-1.85)		
rank_sales=17 × Z-score	-0.432***	-0.0586		
	(-6.15)	(-0.82)		
male		0.836***		
		(39.54)		
log_savings		-0.468***		
		(-36.09)		
current_age		0.100***		
		(54.72)		
Z-score × market_share			-0.184	
			(-1.94)	
Z-score × rank_sales				-0.000582
				(-0.57)
Constant	-4.624***	-8.680***	-4.635***	-4.636***
	(-463.17)	(-55.10)	(-559.21)	(-559.98)
Observations	1536283	1536283	1536283	1536283

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Interpretation

In the second table col 1 there are coefficients different from zero when we do not include the controls. But this coefficients become no different from zero once we include the controls. What is our interpretation?

It means that the

5 Appendix: IE5 granularity

	(1)	(2)	(3)	(4)	(5)
	ratio	ratio	ratio	ratio	ratio
age_months	-14.36	-7.942	-8.906	-18.42	-10.83
	(-1.71)	(-0.92)	(-0.82)	(-1.94)	(-1.09)
age_m2	0.0232	0.0127	0.0143	0.0311	0.0174
	(1.57)	(0.85)	(0.74)	(1.85)	(0.99)
age_m3	-0.0000171	-0.00000965	-0.0000106	-0.0000238	-0.0000129
	(-1.49)	(-0.83)	(-0.70)	(-1.82)	(-0.94)
age_m4	4.69e-09	2.73e-09	2.93e-09	6.81e-09	3.58e-09
	(1.41)	(0.81)	(0.67)	(1.79)	(0.89)
ym_now	-2829.0***	-2139.5***	-1904.1***	-1949.7***	-1848.4***
	(-31.49)	(-35.55)	(-33.33)	(-34.00)	(-31.29)
ym_now2	0	0	0	0	0
	(.)	(.)	(.)	(.)	(.)
ym_now3	0.0128***	0.00982***	0.00868***	0.00890***	0.00843***
	(31.97)	(35.99)	(33.46)	(34.24)	(31.45)
ym_now4	-0.0000193***	-0.0000148***	-0.0000131***	-0.0000134***	-0.0000127***
	(-32.21)	(-36.21)	(-33.52)	(-34.34)	(-31.51)
ym_now5	8.68e-09***	6.73e-09***	5.91e-09***	6.08e-09***	5.74e-09***
	(32.44)	(36.42)	(33.57)	(34.44)	(31.58)
male	-8.014***	-8.071***	-10.55***	-9.016***	-9.047***
	(-22.29)	(-22.72)	(-30.40)	(-25.98)	(-25.28)
val_uf_prima2	1.349***	0.993**	-0.373	0.207	1.988***
	(4.15)	(2.98)	(-0.77)	(0.67)	(6.35)
val_uf_prima_pol2	-0.215***	-0.169**	0.514***	0.0275	-0.275***
	(-3.52)	(-2.72)	(4.27)	(0.46)	(-4.56)
val_uf_prima_pol3	0.0120**	0.0108**	-0.0486***	-0.00114	0.0139***
	(3.22)	(2.86)	(-4.71)	(-0.31)	(3.70)
val_uf_prima_pol4	-0.000208**	-0.000205**	0.00127***	-0.00000526	-0.000225***
	(-3.12)	(-3.03)	(4.76)	(-0.08)	(-3.32)
Constant	754647.2***	566418.8***	506230.3***	519747.5***	491891.0***
	(31.27)	(35.14)	(32.97)	(33.79)	(31.05)
Observations	18266	18963	20151	20498	21324
R-squared	0.671	0.677	0.655	0.667	0.649

t statistics in parentheses

Regressions for the 5 firms with highest market share

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

	(1)	(2)	(3)	(4)	(5)
	ratio	ratio	ratio	ratio	ratio
age_months	-15.89*	-32.01***	-17.00	-10.35	-17.64
	(-2.29)	(-3.81)	(-0.42)	(-1.35)	(-1.54)
age_m2	0.0255*	0.0535***	0.0334	0.0168	0.0296
	(2.11)	(3.64)	(0.45)	(1.26)	(1.47)
age_m3	-0.0000186*	-0.0000402***	-0.0000300	-0.0000127	-0.0000226
	(-1.99)	(-3.54)	(-0.49)	(-1.24)	(-1.44)
age_m4	5.04e-09	1.12e-08***	1.01e-08	3.59e-09	6.42e-09
	(1.88)	(3.44)	(0.54)	(1.22)	(1.40)
ym_now	-902.5***	0	-1615.7***	-2185.5***	-1891.4***
	(-14.10)	(.)	(-14.80)	(-35.13)	(-26.77)
ym_now2	0	0	0	0	0
	(.)	(.)	(.)	(.)	(.)
ym_now3	0.00409***	-0.000129***	0.00737***	0.0100***	0.00866***
	(14.05)	(-6.92)	(14.82)	(35.49)	(27.03)
ym_now4	-0.00000615***	0.000000295***	-0.0000111***	-0.0000151***	-0.0000131***
	(-14.01)	(7.02)	(-14.83)	(-35.65)	(-27.15)
ym_now5	2.77e-09***	-1.80e-10***	5.02e-09***	6.86e-09***	5.92e-09***
	(13.97)	(-7.11)	(14.83)	(35.82)	(27.26)
male	-7.530***	-6.858***	-7.004***	-7.655***	-7.588***
	(-22.12)	(-17.13)	(-14.67)	(-20.54)	(-18.54)
val_uf_prima2	-0.441	-3.602***	2.572***	2.206***	0.669
	(-1.45)	(-7.24)	(3.44)	(6.69)	(1.39)
val_uf_prima_pol2	0.139*	0.824***	-0.554**	-0.367***	-0.111
	(2.39)	(7.41)	(-2.89)	(-5.84)	(-1.26)
val_uf_prima_pol3	-0.00770*	-0.0576***	0.0376*	0.0212***	0.00983
	(-2.14)	(-6.70)	(2.16)	(5.47)	(1.89)
val_uf_prima_pol4	0.000107	0.00123***	-0.000746	-0.000366***	-0.000213*
	(1.65)	(6.02)	(-1.63)	(-5.23)	(-2.38)
Constant	243322.8***	11065.3***	430668.2***	579272.4***	503646.6***
	(14.28)	(5.87)	(14.25)	(34.89)	(26.52)
Observations	20049	13485	8526	19945	11249
R-squared	0.678	0.707	0.645	0.655	0.696

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

	(1)	(2)	(3)	(4)	(5)
	ratio	ratio	ratio	ratio	ratio
age_months	-2.314	-137.4**	10.88	-5.991	-40.89
	(-0.12)	(-2.73)	(0.94)	(-0.35)	(-1.95)
age_m2	0.00249	0.250**	-0.0196	0.00822	0.0721
	(0.07)	(2.71)	(-0.96)	(0.27)	(1.91)
age_m3	-0.00000129	-0.000203**	0.0000148	-0.00000531	-0.0000570
	(-0.05)	(-2.70)	(0.94)	(-0.22)	(-1.90)
age_m4	1.72e-10	6.13e-08**	-4.19e-09	1.25e-09	1.69e-08
	(0.02)	(2.70)	(-0.92)	(0.17)	(1.89)
ym_now	0	-1746.0***	0	-128.7	-352.5***
	(.)	(-23.62)	(.)	(-1.38)	(-3.77)
ym_now2	0	0	0	0	0
	(.)	(.)	(.)	(.)	(.)
ym_now3	0.000901**	0.00794***	0.00199***	0.000493	0.00153***
	(3.27)	(23.49)	(38.17)	(1.16)	(3.61)
ym_now4	-0.00000216**	-0.0000120***	-0.00000434***	-0.000000669	-0.00000225***
	(-3.29)	(-23.41)	(-38.28)	(-1.04)	(-3.52)
ym_now5	1.38e-09***	5.40e-09***	2.52e-09***	2.68e-10	9.90e-10***
	(3.31)	(23.33)	(38.39)	(0.92)	(3.42)
male	-5.885***	-6.682***	-10.43***	-5.831***	-10.23***
	(-11.25)	(-14.77)	(-20.75)	(-14.14)	(-23.76)
val_uf_prima2	0.593	20.71***	1.078*	-0.853	1.210
	(0.87)	(13.46)	(2.40)	(-1.55)	(1.54)
val_uf_prima_pol2	-0.166	-6.569***	-0.0838	0.134	0.391*
	(-1.01)	(-12.07)	(-1.02)	(1.57)	(2.39)
val_uf_prima_pol3	0.0197	0.783***	0.00343	-0.00600	-0.0472***
	(1.45)	(10.52)	(0.70)	(-1.28)	(-3.74)
val_uf_prima_pol4	-0.000519	-0.0310***	-0.0000647	0.0000702	0.00131***
	(-1.54)	(-9.00)	(-0.76)	(0.91)	(4.29)
Constant	-20992.5**	490933.8***	-66797.0***	38547.8	104253.4***
	(-2.65)	(22.21)	(-22.43)	(1.55)	(4.13)
Observations	6151	10148	10376	9475	8714
R-squared	0.673	0.654	0.677	0.711	0.676

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

	(1)	(2)	(3)	(4)
	ratio	ratio	ratio	ratio
age_months	-10.32***	-10.34***	-10.29***	-10.41***
	(-4.01)	(-4.01)	(-4.00)	(-4.04)
age_m2	0.0165***	0.0165***	0.0164***	0.0166***
	(3.64)	(3.64)	(3.62)	(3.67)
age_m3	-0.0000121***	-0.0000122***	-0.0000121***	-0.0000122***
	(-3.45)	(-3.46)	(-3.43)	(-3.48)
age_m4	3.33e-09**	3.34e-09**	3.31e-09**	3.36e-09***
	(3.27)	(3.27)	(3.25)	(3.29)
ym_now	54.15***	54.35***	55.39***	55.11***
	(23.86)	(23.93)	(24.39)	(24.25)
ym_now2	-0.0800***	-0.0803***	-0.0819***	-0.0815***
	(-23.43)	(-23.49)	(-23.96)	(-23.82)
ym_now3	0.0000396***	0.0000397***	0.0000405***	0.0000403***
	(23.12)	(23.16)	(23.66)	(23.51)
male	-8.243***	-8.264***	-8.227***	-8.236***
	(-80.09)	(-80.23)	(-79.96)	(-79.97)
Size2=1 × val_uf_prima2	0.654***			
	(37.48)			
Size2=2 × val_uf_prima2	0.153***			
	(8.22)			
Size3=1 × val_uf_prima2		0.668***		
		(34.96)		
Size3=2 × val_uf_prima2		0.281***		
		(13.32)		
Size3=3 × val_uf_prima2		0.278***		
		(13.41)		
Size4=1 × val_uf_prima2			0.547***	
			(26.73)	
Size4=2 × val_uf_prima2			0.782***	
			(35.25)	
Size4=3 × val_uf_prima2			0.0258	
			(1.08)	
Size4=4 × val_uf_prima2			0.272***	
			(11.76)	
Size5=1 × val_uf_prima2				0.578***
				(25.51)
Size5=2 × val_uf_prima2				0.647***
				(29.05)
Size5=3 × val_uf_prima2				0.450***
				(17.66)
Size5=4 × val_uf_prima2				0.0145
				(0.50)
Size5=5 × val_uf_prima2				0.269***
				(11.64)
Observations	229190	229190	229190	229190
R-squared	0.644	0.644	0.645	0.644

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$