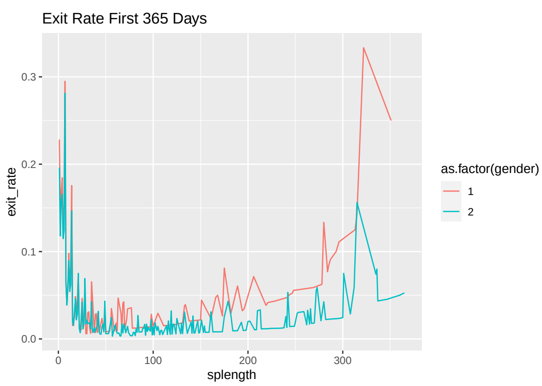
Applied Microeconometrics - Assignment 4

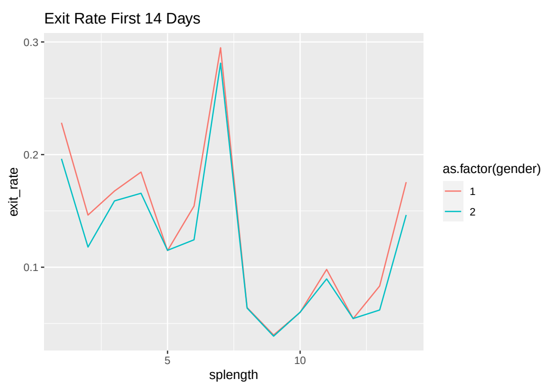
Walter Verwer (589962) & Bas Machielsen (590049)

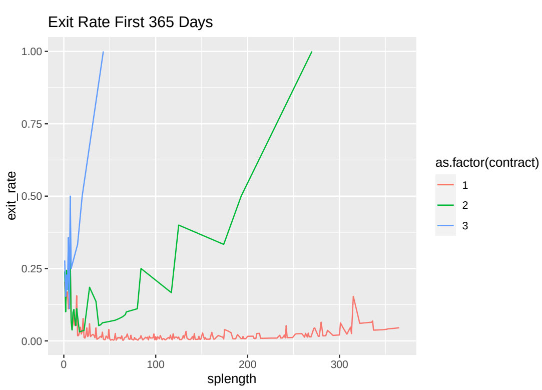
October 8, 2021

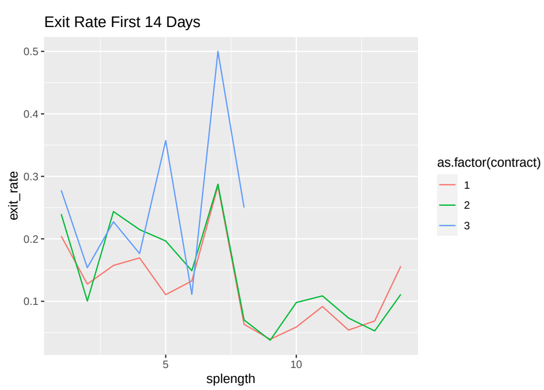
1. Describe the sickness spell data, i.e. do a simple listing of the survivor function and plot the hazard rate and the survivor function. Make separate plots for the first two weeks and for the first year. Also plot the hazard by different subgroups (for instance gender) and test whether the survival curves are the same for the different subgroups.

First, we plot the exit rates:

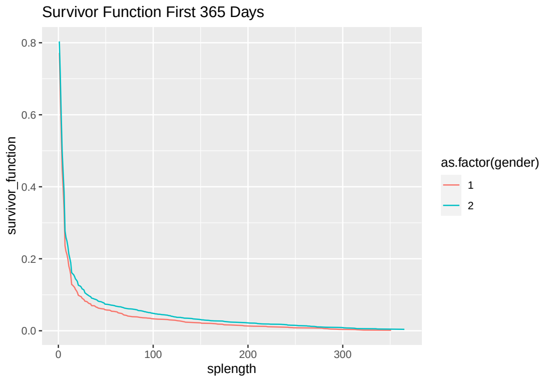


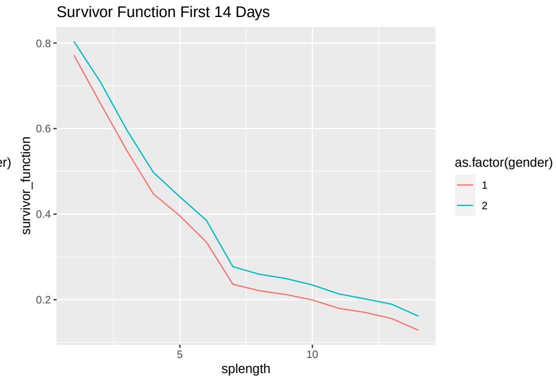


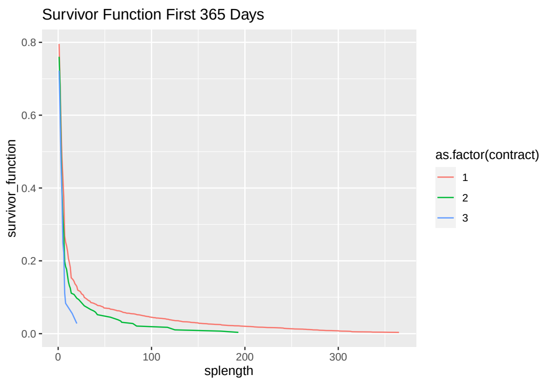


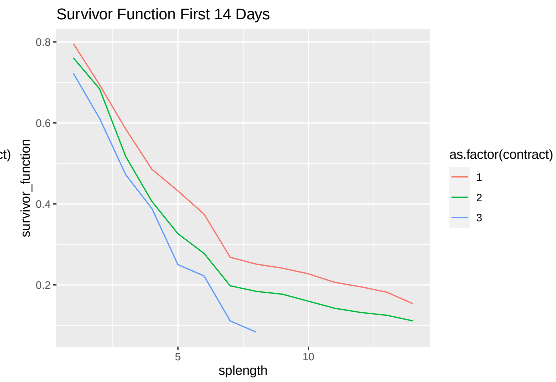


Then, we plot the survivor functions:


Survivor Function First 365 Days
survivor_function
0.8
0.6
0.4
0.2
0.0
0
100
200
300
splength
as.factor(gender)
1
2


Survivor Function First 14 Days
survivor_function
0.8
0.6
0.4
0.2
5
10
splength
as.factor(gender)
1
2


Survivor Function First 365 Days
survivor_function
0.8
0.6
0.4
0.2
0.0
0
100
200
300
splength
as.factor(contract)
1
2
3


Survivor Function First 14 Days
survivor_function
0.8
0.6
0.4
0.2
5
10
splength
as.factor(contract)
1
2
3

2.1 Estimate a Weibull and an Exponential model for sickness spells. Start with a very simple specification and you only include one regressor and subsequently add more regressors. Comment on the change in the Weibull parameters and the regression parameters when you add more variables to the model. Compare the estimates of both models.

Table 1: Weibull Models

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
(Intercept)	1.938***	1.795***	2.280***	2.279***	2.334***	0.979*
	(0.127)	(0.195)	(0.213)	(0.213)	(0.251)	(0.530)
gender	0.243***	0.254***	0.273***	0.274***	0.278***	0.319***
	(0.077)	(0.078)	(0.078)	(0.087)	(0.087)	(0.092)
Log(scale)	0.490***	0.490***	0.487***	0.487***	0.486***	0.483***
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
marstat		0.069	0.055	0.055	0.049	0.031
		(0.066)	(0.066)	(0.066)	(0.066)	(0.068)
contract			−0.467***	−0.467***	−0.467***	−0.423***
			(0.093)	(0.093)	(0.094)	(0.093)
lowgroup				−0.004	−0.005	−0.020
				(0.079)	(0.079)	(0.080)
classize					0.002	0.002
					(0.005)	(0.005)
schsize					0.000	0.000
					(0.000)	(0.000)
public					−0.056	−0.048
					(0.091)	(0.089)
protest					−0.153	−0.078
					(0.106)	(0.108)
merged						0.005
						(0.013)
avgfem						−0.206
						(0.284)
avgage						0.034***
						(0.011)
avglowgr						0.122
						(0.259)
Num.Obs.	6520	6520	6520	6520	6520	6520
AIC	44 443.9	44 442.6	44 415.1	44 417.1	44 416.8	44 383.2
Log.Lik.	−22 218.934	−22 217.295	−22 202.540	−22 202.537	−22 198.402	−22 177.593

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Most of the parameters remain stable over the models, with the exception of the intercept, which has no theoretical meaning. The significant variables are gender, the Weibull scale parameter, contract, and, in the last specification, the average age in class. People with a less stable contract are more likely to remain sick, and females are more likely to get better again than males at any point in time. The average age is also correlated positively with the hazard rate of ending a sickness spell.

Table 2: Exponential Models

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
(Intercept)	2.435***	2.247***	2.852***	2.848***	2.909***	1.601**
	(0.164)	(0.236)	(0.255)	(0.254)	(0.290)	(0.633)
gender	0.267***	0.280***	0.303***	0.324***	0.332***	0.377***
	(0.097)	(0.098)	(0.097)	(0.104)	(0.103)	(0.108)
marstat		0.091	0.077	0.076	0.077	0.058
		(0.080)	(0.079)	(0.079)	(0.078)	(0.080)
contract			−0.594***	−0.595***	−0.600***	−0.537***
			(0.112)	(0.111)	(0.112)	(0.116)
lowgroup				−0.044	−0.043	−0.063
				(0.096)	(0.097)	(0.098)
classize					0.003	0.003
					(0.005)	(0.005)
schsize					0.000	0.000
					(0.000)	(0.000)
public					−0.116	−0.093
					(0.103)	(0.104)
protest					−0.160	−0.083
					(0.124)	(0.128)
merged						−0.004
						(0.015)
avgfem						−0.237
						(0.335)
avgage						0.033***
						(0.012)
avglowgr						0.091
						(0.297)
Num.Obs.	6520	6520	6520	6520	6520	6520
AIC	49 147.6	49 133.1	49 022.5	49 022.1	48 996.4	48 893.8
Log.Lik.	−24 571.812	−24 563.554	−24 507.235	−24 506.073	−24 489.220	−24 433.900

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The parameters of the exponential model are very similar to the parameters of the Weibull model. The baseline hazard is, as in the Weibull model, sensitive to the inclusion of covariates. Intuitively, this makes sense: more covariates allows for a better isolation of the baseline (unconditional) hazard.

2.2 Estimate separate Weibull models for males and females. Comment on the results (is it better to estimate separate models for males and females?) Estimate the Weibull duration model for other subgroups that may differ in their behavior and where the baseline hazard may differ.

First, we estimate a model for males:

Table 3: Weibull Models - Males Only

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
(Intercept)	2.204***	2.446***	3.162***	3.149***	3.363***	1.558*
	(0.048)	(0.329)	(0.394)	(0.394)	(0.419)	(0.904)
Log(scale)	0.461***	0.460***	0.458***	0.457***	0.456***	0.449***
	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)	(0.024)
marstat		−0.125	−0.138	−0.138	−0.150	−0.149
		(0.168)	(0.165)	(0.166)	(0.161)	(0.157)
contract			−0.675***	−0.685***	−0.631***	−0.559***
			(0.169)	(0.174)	(0.179)	(0.188)
lowgroup				0.083	0.084	0.094
				(0.122)	(0.122)	(0.123)
classize					−0.009	−0.008
					(0.007)	(0.007)
schsize					0.000	0.000
					(0.000)	(0.000)
public					−0.073	−0.036
					(0.128)	(0.132)
protest					0.009	0.032
					(0.180)	(0.174)
merged						0.035
						(0.028)
avgfem						0.457
						(0.400)
avgage						0.041**
						(0.016)
avglowgr						−0.357
						(0.410)
Num.Obs.	2046	2046	2046	2046	2046	2046
AIC	13 422.3	13 422.4	13 416.2	13 417.1	13 421.4	13 405.4
Log.Lik.	−6709.152	−6708.217	−6704.101	−6703.534	−6701.689	−6689.697

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Next, we estimate a model for females:

Table 4: Weibull Models - Females Only

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
(Intercept)	2.414***	2.228***	2.715***	2.755***	2.763***	1.729***
	(0.042)	(0.138)	(0.174)	(0.197)	(0.260)	(0.643)
Log(scale)	0.503***	0.502***	0.499***	0.499***	0.498***	0.494***
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
marstat		0.106	0.093	0.092	0.076	0.058
		(0.074)	(0.074)	(0.074)	(0.073)	(0.076)
contract			−0.435***	−0.435***	−0.433***	−0.393***
			(0.103)	(0.102)	(0.106)	(0.103)
lowgroup				−0.048	−0.038	−0.081
				(0.099)	(0.099)	(0.103)
classize					0.005	0.006
					(0.006)	(0.006)
schsize					0.000	0.000
					(0.000)	(0.000)
public					−0.039	−0.030
					(0.116)	(0.114)
protest					−0.211*	−0.123
					(0.127)	(0.133)
merged						−0.008
						(0.014)
avgfem						−0.557
						(0.383)
avgage						0.030**
						(0.013)
avglowgr						0.403
						(0.334)
Num.Obs.	4474	4474	4474	4474	4474	4474
AIC	31 018.2	31 013.9	30 994.0	30 995.4	30 991.3	30 969.4
Log.Lik.	−15 507.109	−15 503.936	−15 493.017	−15 492.706	−15 486.663	−15 471.686

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The results show that the coefficients for males and females are quite similar, but it might be different to estimate separate models because the results do not suffer from a lack of efficiency (the same coefficients show significance), and the log-likelihoods are smaller (closer to zero) relative to the pooled model.

Now, given that under question 1 we observed a difference between the exit rates and the survival functions for the different contract types, we have estimated Weibull models for the contract types. However, our results are difficult to compare. The reason for this is that the models for the temporary and mixed contract have very little observations. Especially, the mixed contract has too little (36 observations). For this reason we can not make an educated comparison for this subgroup.

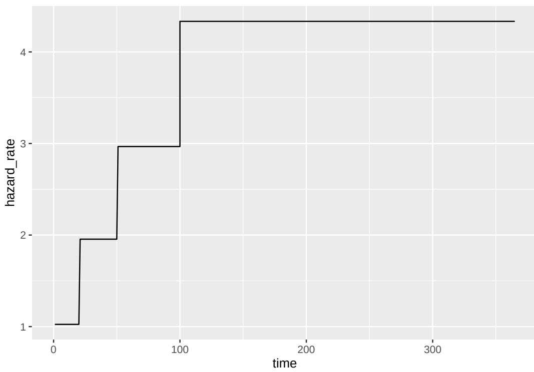
Table 5: Weibull Models - Models for the all contract types

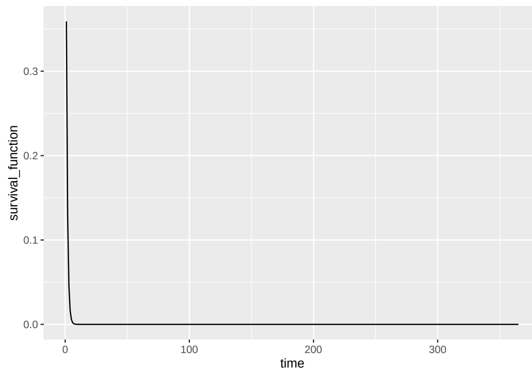
	Fixed contract	Temporary contract	Mixed contract
	Model 1	Model 2	Model 3
(Intercept)	0.526	0.290	2.377
	(0.523)	(1.742)	(2.100)
marstat	0.019	0.439**	−0.323
	(0.070)	(0.220)	(0.225)
gender	0.304***	0.781**	0.616
	(0.095)	(0.309)	(0.629)
lowgroup	−0.010	−0.083	−0.089
	(0.083)	(0.271)	(0.739)
classize	0.003	0.002	0.008
	(0.005)	(0.009)	(0.035)
schsize	0.000	0.000	0.005*
	(0.000)	(0.001)	(0.003)
public	−0.041	−0.087	−0.464**
	(0.090)	(0.264)	(0.189)
protest	−0.085	0.154	0.367
	(0.108)	(0.372)	(0.522)
merged	0.002	0.056	−0.013
	(0.014)	(0.046)	(0.052)
avgfem	−0.172	−0.796	−1.661*
	(0.285)	(0.934)	(1.003)
avgage	0.035***	−0.013	−0.045
	(0.011)	(0.038)	(0.040)
avglowgr	0.080	0.978	0.581
	(0.262)	(0.820)	(1.867)
Log(scale)	0.490***	0.317***	−0.226*
	(0.014)	(0.050)	(0.127)
Num.Obs.	6196	288	36
AIC	42 336.4	1838.7	202.8
Log.Lik.	−21 155.217	−906.333	−88.386

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

3.1 Estimate a Piece Wise Constant (PWC) model for the entire sample. Use the stsplit command to create multiple record data. You can have as many steps as the data allow you to take, but first start with only a few (3 or 4 steps). Next estimate a model with 15-20 steps, or even more. Plot de duration pattern implied by the estimates and comment on these and the regression parameters. How do the regression parameters (β) compare with those of the Weibull model?

First, we estimate a piece wise constant model.





Above, we plot the hazard rate and survival function of our estimated piecewise constant model. The piecewise-constant model coefficient estimates can be found in the table below. As the attrition from the sample is very high in the starting period, our Weibull model predicts a high hazard rate, and thus high attrition in the first period. Afterwards, piecewise, it attempts to catch up the imbalances by fitting a new Weibull distribution and hazard rate (implied by the new constant term), which significantly increases the hazard rate relative to the previous period. This proves the model does catch up with the decay in the data that is not as fastly decaying as implied by a Weibull model under the previous parameters.

One other consequence of including a dummy for the first twenty days is an underestimation of the survival probability in the early period. Even though it mildly realistic, it underestimates survival by only focusing on the first 20 days, under which attrition is very high. We repeat this exercise again with a piecewise constant model with more dummies, but omit the plot for brevity's sake.

In the table below, we show the estimates of two piece-wise constant models. For the first model, we have plotted the hazard rate and survival function, for the second, we omitted it for brevitiy's sake. We see that the hazard rate is increasing over time, in general, though not monotonically. The estimated survival curve is just the integral over all these hazard rates, and show a very fast tendency to go to zero. As before, this makes sense, as this is the case in the data, but piecewise dummies focusing on a small first period make the fitted distribution such that the survival rate is approaching zero in a very short period.

Comparing these results to the Weibull models, we observe that the inclusion of piece-wise constants greatly affects the coefficient estimates for the covariates: in particular, variation that was previously attributed to the covariates is now attributed to a more flexible baseline hazard over time: likely, we have overestimated the influence of covariates in the preceding analyses, because the covariates should only be modeled after the baseline hazard is accurately specified.

Table 6: PWC (Weibull)

	PWC 14 steps	PWC 4 steps
	Model 1	Model 2
(Intercept)	1.109***	1.024***
	(0.199)	(0.181)
gender	0.056*	0.054*
	(0.033)	(0.032)
marstat	−0.032	−0.024
	(0.024)	(0.023)
contract	−0.149***	−0.114**
	(0.044)	(0.045)
merged	0.007	0.012*
	(0.007)	(0.007)
lowgroup	0.012	0.005
	(0.032)	(0.031)
classize	−0.001	0.000
	(0.001)	(0.001)
schsize	0.000	0.000
	(0.000)	(0.000)
public	0.037	0.035
	(0.038)	(0.034)
protest	−0.005	0.001
	(0.041)	(0.039)
avgfem	0.066	0.014
	(0.113)	(0.104)
avgage	0.015***	0.014***
	(0.004)	(0.004)
avglowgr	−0.010	0.030
	(0.122)	(0.115)
timeperiod 10	4.632***	
	(0.246)	
timeperiod 11	4.397***	
	(0.193)	
timeperiod 12	4.938***	
	(0.250)	
timeperiod 13	4.137***	
	(0.038)	
timeperiod 14	5.551***	
	(0.305)	
timeperiod 2	2.058***	1.955***
	(0.032)	(0.029)
timeperiod 3	2.704***	2.967***
	(0.052)	(0.048)
timeperiod 4	3.150***	4.334***
	(0.075)	(0.075)
timeperiod 5	3.412***	
	(0.080)	
timeperiod 6	3.731***	
	(0.121)	
timeperiod 7	4.158***	
	(0.157)	
timeperiod 8	3.736***	
	(0.089)	
timeperiod 9	4.550***	
	(0.268)	
Num.Obs.	6473	6520
AIC	36 079.3	36 146.3
Log.Lik.	−18 013.641	−18 057.168

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

3.2 Estimate separate models for males and females.

First we estimate a PWC model for males.

Table 7: PWC (Weibull) - Males only

	PWC 14 steps	PWC 4 steps
	Model 1	Model 2
(Intercept)	−850.507***	−1.016
	(0.322)	(0.958)
marstat	−0.017	0.260**
	(0.060)	(0.130)
contract	853.775***	5.494***
	(0.000)	(0.661)
merged	0.006	0.017
	(0.014)	(0.022)
lowgroup	0.067	−0.082
	(0.045)	(0.097)
classize	0.001	−0.008
	(0.003)	(0.007)
schsize	0.000	0.000
	(0.000)	(0.000)
public	0.075	−0.103
	(0.054)	(0.165)
protest	−0.049	−0.341**
	(0.052)	(0.145)
avgfem	−0.009	−0.445
	(0.190)	(0.457)
avgage	0.003	−0.012
	(0.006)	(0.015)
avglowgr	−0.086	0.172
	(0.180)	(0.338)
timeperiod 10	2.133***	
	(0.102)	
timeperiod 11	2.191***	
	(0.105)	
timeperiod 12	2.437***	
	(0.090)	
timeperiod 14	3.320***	
	(0.151)	
timeperiod 2	0.530***	0.256
	(0.070)	(0.222)
timeperiod 3	1.068***	0.807***
	(0.096)	(0.225)
timeperiod 4	1.217***	1.813***
	(0.087)	(0.240)
timeperiod 5	1.558***	
	(0.103)	
timeperiod 6	1.688***	
	(0.087)	
timeperiod 7	1.832***	
	(0.068)	
timeperiod 8	2.041***	
	(0.131)	
timeperiod 9	1.997***	
	(0.072)	
Log(scale)	−1.998***	−1.072***
	(0.210)	(0.114)
Num.Obs.	2029	2036
AIC	597.1	725.3
Log.Lik.	−273.545	−346.670

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Now we estimate the same model for females.

Table 8: PWC (Weibull) - Females only

	PWC 14 steps	PWC 4 steps
	Model 1	Model 2
(Intercept)	3.084***	3.780***
	(0.220)	(0.615)
marstat	0.019	0.097*
	(0.019)	(0.059)
contract	0.150	0.309
	(0.099)	(0.251)
merged	0.002	0.001
	(0.007)	(0.019)
lowgroup	−0.073*	−0.172
	(0.038)	(0.125)
classize	−0.001	−0.002
	(0.001)	(0.004)
schsize	0.000	0.000
	(0.000)	(0.000)
public	0.045	0.014
	(0.032)	(0.103)
protest	0.104**	0.137
	(0.045)	(0.143)
avgfem	−0.151	−0.160
	(0.103)	(0.347)
avgage	0.004	0.008
	(0.004)	(0.010)
avglowgr	0.198**	0.226
	(0.086)	(0.326)
timeperiod 10	2.116***	
	(0.057)	
timeperiod 11	2.302***	
	(0.064)	
timeperiod 12	2.437***	
	(0.058)	
timeperiod 13	42.067***	
	(0.000)	
timeperiod 14	3.169***	
	(0.119)	
timeperiod 2	0.591***	0.045
	(0.051)	(0.168)
timeperiod 3	1.002***	0.522***
	(0.060)	(0.174)
timeperiod 4	1.297***	1.500***
	(0.054)	(0.195)
timeperiod 5	1.509***	
	(0.053)	
timeperiod 6	1.627***	
	(0.058)	
timeperiod 7	1.749***	
	(0.053)	
timeperiod 8	2.137***	
	(0.099)	
timeperiod 9	2.069***	
	(0.070)	
Log(scale)	−1.882***	−0.874***
	(0.141)	(0.060)
Num.Obs.	4444	4456
AIC	1619.9	1922.2
Log.Lik.	−783.936	−945.106

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

4. Estimate a Cox model and compare the most elaborate specification with the results of the PWC model

Table 9: Cox Model

	Model 1	Model 2
marstat	−0.007	0.002
	(0.024)	(0.024)
gender	−0.129***	−0.147***
	(0.031)	(0.034)
contract	0.224***	0.206***
	(0.050)	(0.050)
lowgroup	−0.004	0.002
	(0.030)	(0.031)
classize	0.000	0.000
	(0.002)	(0.002)
schsize	0.000	0.000
	(0.000)	(0.000)
public	−0.007	−0.006
	(0.029)	(0.029)
protest	0.075	0.038
	(0.034)	(0.035)
merged		−0.009
		(0.006)
avgfem		0.076
		(0.099)
avgage		−0.018***
		(0.004)
avglowgr		−0.071
		(0.096)
Num.Obs.	6520	6520
R2	0.007	0.012
AIC	99 424.7	99 398.2
Log.Lik.	−49 704.363	−49 687.124

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Comparing the most elaborate specification of the Cox model with the most elaborate PWC model, we observe an interesting finding. That is, the model estimates appear to be very sensitive to our choice of model. For example, gender appears to be significant at the 1% level for the PWC model, with a positive sign, and for the Cox model it has the same significance level, but has a negative sign. This clearly illustrates the sensitivity of the parameter estimates to the parametric form imposed.

5. Repeat the procedure of question 2 for a Weibull model with (e.g., gamma) unobserved heterogeneity. Compare the estimates of the regression coefficients across the models with and without unobserved heterogeneity.

As above, but now with a Piecewise Constant (PWC) specification, where you have an elaborate specification of the baseline hazard (say, 20 dummies).

Table 10: Gamma Weibull Models

	No time dummies	4 Time Dummies	22 Time Dummeis
mu	4.481	7.907	1.541
		(1153.496)	(0.033)
sigma	1.654	0.732	0.686
		(0.094)	(0.007)
Q	1.000	1.000	1.000
gender	−1.034	−3.114	0.042**
		(576.748)	(0.019)
AIC	31 020.2	24 329.4	34 240.8
Log.Lik.	−15 507.109	−12 158.676	−17 095.402
N	4474.000	4456.000	6501.000

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

In the above table, we report the three different Gamma-Weibull models, with gender as an explanatory variable and no time dummies (model 1), 4 time dummies (model 2), and 22 time dummies (model 3). We observe that the gender coefficient is significant in the third case, indicating that an extensive and non-parametric baseline hazard helps us identify the effect of gender. When contrasting the results with the Weibull model from question 2, we find that the coefficient on gender is again much smaller, indicating that the results might have been due to spurious correlation and subject-specific effects, rather than within-subject variation.

5.2 Compare the estimates of the Cox model (question 4) with the results of the PWC model with unobserved heterogeneity.

The results from the Cox model in question four come closer to the results in this question, when focusing on the coefficient on gender. Still, if we take these results to be the true results, the Cox model overestimates the coefficient by a factor of roughly three, which is substantial. That means that also in case of duration data, a non-panel model cannot easily reflect the within-subjects effect, as we know from linear and other simpler models.

6. Multiple Spells

6.1 Estimate a standard Cox model (PL) and estimate Stratified Cox models (SPL). Concerning the latter, estimate SPL models, where the school is the stratum and estimate one where the teacher is the stratum. Comment on the teacher SPL approach. Compare the PL and the school SPL estimates. Can you think of a test to test for the relevance of using the school SPL (rather than doing the PL)?

First we estimate a model with schoolid as a strata. This gives us the following output (the package we use can not produce standard tables, hence we show its output):

```
## Call:
## survival::coxph(formula = Surv(splength, rcensor) ~ strata(schoolid) +
##     marstat + gender + contract + lowgroup + classize + schsize +
##     public + protest, data = dataset %>% filter(sptype == 2) %>%
##     mutate(rcensor = if_else(rcensor == 1, 0, 1)), cluster = schoolid)
##
##   n= 6520, number of events= 6324
##
##                coef  exp(coef)   se(coef)  robust se      z Pr(>|z|)
## marstat   0.0171112  1.0172584  0.0284297  0.0349686  0.489 0.624607
## gender   -0.1523871  0.8586559  0.0369382  0.0462039 -3.298 0.000973 ***
## contract  0.1928554  1.2127074  0.0582814  0.0550957  3.500 0.000465 ***
## lowgroup -0.0274227  0.9729499  0.0341762  0.0395282 -0.694 0.487838
## classize -0.0006756  0.9993246  0.0021278  0.0020257 -0.334 0.738734
## schsize  -0.0002969  0.9997031  0.0005075  0.0003643 -0.815 0.415017
## public   -0.1253421  0.8821950  0.1397192  0.1740631 -0.720 0.471466
## protest  -0.2016655  0.8173683  0.1689328  0.1500950 -1.344 0.179082
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##          exp(coef) exp(-coef) lower .95 upper .95
## marstat     1.0173     0.9830    0.9499     1.089
## gender      0.8587     1.1646    0.7843     0.940
## contract    1.2127     0.8246    1.0886     1.351
## lowgroup    0.9729     1.0278    0.9004     1.051
## classize    0.9993     1.0007    0.9954     1.003
## schsize     0.9997     1.0003    0.9990     1.000
## public      0.8822     1.1335    0.6272     1.241
## protest     0.8174     1.2234    0.6091     1.097
##
## Concordance= 0.516  (se = 0.012 )
## Likelihood ratio test= 37.8  on 8 df,   p=8e-06
## Wald test            = 33.26  on 8 df,   p=6e-05
## Score (logrank) test = 38.38  on 8 df,   p=6e-06,   Robust = 29.2  p=3e-04
##
##   (Note: the likelihood ratio and score tests assume independence of
##      observations within a cluster, the Wald and robust score tests do not).
```

```
## Call:
## survival::coxph(formula = Surv(splength, rcensor) ~ strata(schoolid) +
##     marstat + gender + contract + lowgroup + classize + schsize +
##     public + protest + merged + avgfem + avgage + avglowgr, data = dataset %>%
##     filter(sptype == 2) %>% mutate(rcensor = if_else(rcensor ==
##     1, 0, 1)), cluster = schoolid)
##
##   n= 6520, number of events= 6324
```

```
##
##                  coef  exp(coef)   se(coef)  robust se      z Pr(>|z|)
## marstat   0.0180778  1.0182422  0.0284262  0.0350296  0.516 0.605805
## gender   -0.1518962  0.8590774  0.0369293  0.0460497 -3.299 0.000972 ***
## contract  0.1923285  1.2120687  0.0583153  0.0549965  3.497 0.000470 ***
## lowgroup -0.0255371  0.9747863  0.0341763  0.0393892 -0.648 0.516774
## classize -0.0006425  0.9993577  0.0021221  0.0019948 -0.322 0.747373
## schsize  -0.0004039  0.9995962  0.0005133  0.0003913 -1.032 0.301978
## public   -0.1228420  0.8844034  0.1397313  0.1737639 -0.707 0.479599
## protest  -0.2045353  0.8150260  0.1689518  0.1514714 -1.350 0.176912
## merged   -0.0198450  0.9803506  0.0097281  0.0099312 -1.998 0.045689 *
## avgfem           NA         NA  0.0000000  0.0000000     NA       NA
## avgage           NA         NA  0.0000000  0.0000000     NA       NA
## avglowgr         NA         NA  0.0000000  0.0000000     NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##          exp(coef) exp(-coef) lower .95 upper .95
## marstat     1.0182     0.9821    0.9507    1.0906
## gender      0.8591     1.1640    0.7849    0.9402
## contract    1.2121     0.8250    1.0882    1.3500
## lowgroup    0.9748     1.0259    0.9024    1.0530
## classize    0.9994     1.0006    0.9955    1.0033
## schsize     0.9996     1.0004    0.9988    1.0004
## public      0.8844     1.1307    0.6291    1.2432
## protest     0.8150     1.2270    0.6057    1.0967
## merged      0.9804     1.0200    0.9615    0.9996
## avgfem          NA         NA        NA        NA
## avgage          NA         NA        NA        NA
## avglowgr        NA         NA        NA        NA
##
## Concordance= 0.523  (se = 0.012 )
## Likelihood ratio test= 42  on 9 df,   p=3e-06
## Wald test            = 35.67  on 9 df,   p=5e-05
## Score (logrank) test = 42.53  on 9 df,   p=3e-06,   Robust = 31.07  p=3e-04
##
##   (Note: the likelihood ratio and score tests assume independence of
##      observations within a cluster, the Wald and robust score tests do not).
```

The estimated model with teachers as the strata is:

```
## Call:
## survival::coxph(formula = Surv(splength, rcensor) ~ strata(teachid) +
##     marstat + gender + contract + lowgroup + classize + schsize +
##     public + protest, data = dataset %>% filter(sptype == 2) %>%
##     mutate(rcensor = if_else(rcensor == 1, 0, 1)), cluster = schoolid)
##
##   n= 6520, number of events= 6324
##
##                  coef  exp(coef)   se(coef)  robust se      z Pr(>|z|)
## marstat   3.810e-03  1.004e+00  2.436e-02  2.887e-02  0.132 0.894994
## gender   -1.357e-01  8.731e-01  3.251e-02  4.001e-02 -3.392 0.000695 ***
## contract  2.280e-01  1.256e+00  5.307e-02  6.491e-02  3.512 0.000445 ***
## lowgroup -1.182e-02  9.882e-01  3.058e-02  3.810e-02 -0.310 0.756286
## classize -2.504e-05  1.000e+00  1.781e-03  1.935e-03 -0.013 0.989673
```

```
## schsize  -1.006e-04  9.999e-01  1.270e-04  2.149e-04 -0.468 0.639710
## public   -2.719e-03  9.973e-01  2.939e-02  4.571e-02 -0.059 0.952576
## protest   7.510e-02  1.078e+00  3.423e-02  5.156e-02  1.457 0.145182
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##          exp(coef) exp(-coef) lower .95 upper .95
## marstat     1.0038     0.9962    0.9486    1.0623
## gender      0.8731     1.1454    0.8072    0.9443
## contract    1.2560     0.7962    1.1060    1.4264
## lowgroup    0.9882     1.0119    0.9171    1.0649
## classize    1.0000     1.0000    0.9962    1.0038
## schsize     0.9999     1.0001    0.9995    1.0003
## public      0.9973     1.0027    0.9118    1.0908
## protest     1.0780     0.9276    0.9744    1.1926
##
## Concordance= 0.529  (se = 0.007 )
## Likelihood ratio test= 47.01  on 8 df,   p=2e-07
## Wald test            = 30.95  on 8 df,   p=1e-04
## Score (logrank) test = 48.85  on 8 df,   p=7e-08,   Robust = 29.2  p=3e-04
##
##   (Note: the likelihood ratio and score tests assume independence of
##      observations within a cluster, the Wald and robust score tests do not).

## Call:
## survival::coxph(formula = Surv(splength, rcensor) ~ strata(teachid) +
##     marstat + gender + contract + lowgroup + classize + schsize +
##     public + protest + merged + avgfem + avgage + avglowgr, data = dataset %>%
##     filter(sptype == 2) %>% mutate(rcensor = if_else(rcensor ==
##     1, 0, 1)), cluster = schoolid)
##
##   n= 6520, number of events= 6324
##
##                coef  exp(coef)   se(coef)  robust se      z Pr(>|z|)
## marstat   0.0132498  1.0133379  0.0246037  0.0292414  0.453 0.650465
## gender   -0.1517235  0.8592258  0.0348124  0.0423751 -3.580 0.000343 ***
## contract  0.2114603  1.2354809  0.0533250  0.0637240  3.318 0.000905 ***
## lowgroup -0.0057975  0.9942192  0.0321718  0.0380892 -0.152 0.879022
## classize -0.0002979  0.9997021  0.0017902  0.0019545 -0.152 0.878852
## schsize  -0.0002714  0.9997286  0.0001325  0.0001915 -1.417 0.156416
## public   -0.0006410  0.9993592  0.0296442  0.0439327 -0.015 0.988360
## protest   0.0388644  1.0396295  0.0351635  0.0526585  0.738 0.460486
## merged   -0.0105901  0.9894657  0.0056826  0.0073454 -1.442 0.149374
## avgfem    0.0553857  1.0569482  0.1004482  0.1392552  0.398 0.690831
## avgage   -0.0186106  0.9815615  0.0035485  0.0049979 -3.724 0.000196 ***
## avglowgr -0.0726250  0.9299495  0.0976814  0.1429178 -0.508 0.611342
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##          exp(coef) exp(-coef) lower .95 upper .95
## marstat     1.0133     0.9868    0.9569    1.0731
## gender      0.8592     1.1638    0.7907    0.9336
## contract    1.2355     0.8094    1.0904    1.3998
## lowgroup    0.9942     1.0058    0.9227    1.0713
```

```
## classize    0.9997     1.0003    0.9959    1.0035
## schsize     0.9997     1.0003    0.9994    1.0001
## public      0.9994     1.0006    0.9169    1.0892
## protest     1.0396     0.9619    0.9377    1.1527
## merged      0.9895     1.0106    0.9753    1.0038
## avgfem      1.0569     0.9461    0.8045    1.3886
## avgage      0.9816     1.0188    0.9720    0.9912
## avglowgr    0.9299     1.0753    0.7028    1.2306
##
## Concordance= 0.541  (se = 0.007 )
## Likelihood ratio test= 81.48  on 12 df,   p=2e-12
## Wald test            = 45.53  on 12 df,   p=8e-06
## Score (logrank) test = 83.08  on 12 df,   p=1e-12,   Robust = 38.96  p=1e-04
##
##   (Note: the likelihood ratio and score tests assume independence of
##      observations within a cluster, the Wald and robust score tests do not).
```

Let us first compare the PL model with the SPL model with schoolid as the strata. The most striking feauture of the PL model was that gender and contract were highly statistically and economically significant. This feauture is also observable in the SPL model, and the signs are the same. Apart from these findings we do not observe any other interesting findings.

Second, we can compare the teacher stratified model with the school stratified model. Again we obtain very comparable estimates. This indicates that choosing a strata does not seem to have an added value.

6.2 Estimate a model with school specific dummies and compare these estimates with those obtained from the school SPL.

Table 11: Cox Model with School Dummies

	Model 1	Model 2
schooldummies	0.001***	0.001***
	(0.000)	(0.000)
marstat	−0.092	−0.079
	(0.129)	(0.129)
gender	−0.316*	−0.392**
	(0.182)	(0.192)
contract	−0.419	−0.445
	(0.706)	(0.711)
lowgroup	0.094	0.110
	(0.172)	(0.179)
classize	0.003	0.004
	(0.008)	(0.008)
schsize	0.000	−0.001
	(0.001)	(0.001)
public	0.027	0.006
	(0.166)	(0.169)
protest	−0.002	−0.036
	(0.196)	(0.202)
merged		−0.008
		(0.036)
avgfem		0.559
		(0.571)
avgage		−0.011
		(0.020)
avglowgr		0.038
		(0.567)
Num.Obs.	6520	6520
R2	0.002	0.002
AIC	2051.4	2057.4
Log.Lik.	−1016.708	−1015.698

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

The results of the model with school specific dummies compared to the school stratas show us that the schooldummies are highly significant and including them seem to have an effect on the gender and contract variables. Gender becomes less significant (depending on the model 10% or 5% level), contract becomes insignificant and changes sign. Allowing for schooldummies implies that the base hazard rate per school can be different, but we do not study the variation of the covariates within schools.

6.3 Observed sickness patterns vary between schools. This may be due to sorting effects (bad teachers are the reason why the school scores bad in absenteeism) and/or the school effects (it is elements of the school that make some schools worse than others. Can you think of a test/procedure to shed some more light on this issue?

We decide to test the model specification of schoolid as a strata against the standard cox model based on the log partial likelihood. This test indicates whether stratas based on schoolid matter. We implement this test using the survival package in R. Our test results are shown below. The results indicate that the standard model is worse compared to the stratified model (likelihoods differ highly significantly). The associated p-value is a lot smaller than 0.01, meaning it is a highly significant result according to the test. However, given the previous comparison of the model's estimates, we cannot conclude that there might be sorting effect, even though it intuitively makes sense.

```
## Analysis of Deviance Table
##  Cox model: response is  Surv(splength, rcensor)
##  Model 1: ~ marstat + gender + contract + lowgroup + classize + schsize + public + protest + merged +
##  Model 2: ~ strata(schoolid) + marstat + gender + contract + lowgroup + classize + schsize + public +
##   loglik Chisq Df P(>|Chi|)
## 1 -49687
## 2 -14554 70267  3 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```