

1. Anàlisi de tots els accidents

Anna Orteu

Contents

| | |
|---|----------|
| 1 Dades | 1 |
| 2 Estadístics | 1 |
| 3 Resultats de l'estimació Diff-in-Diff per a les reclamacions ATT | 5 |
| 3.1 Models sense covariables | 5 |
| 3.1.1 Models de regressió TWFE i OLS | 5 |
| 3.1.1.1 Variable Total | 5 |
| 3.1.1.2 Variable Night | 5 |
| 3.1.1.3 Variable Speed | 6 |
| 3.1.1.4 Variable Urban | 6 |
| 3.1.2 Estimació no paramètrica amb distribucions empíriques | 7 |
| 3.1.2.1 Variable Total | 7 |
| 3.1.2.2 Variable Night | 9 |
| 3.1.2.3 Variable Speed | 12 |
| 3.1.2.4 Variable Urban | 14 |
| 3.2 Models amb covariables canviants en el temps | 17 |
| 3.2.1 Model TWFE | 17 |
| 3.2.1.1 Variable Total amb l'efecte de covariables | 17 |
| 3.2.1.2 Variable Night amb l'efecte de covariables | 20 |
| 3.2.1.3 Variable Speed amb l'efecte de covariables | 25 |
| 3.2.1.4 Variable Urban amb l'efecte de covariables | 30 |
| 3.2.2 Model semi-paramètric - Integració de propensity scores amb models Dif-in-Dif | 33 |
| 3.2.2.1 Variable Total amb l'efecte de covariables | 34 |
| 3.2.2.2 Variable Night amb l'efecte de covariables | 34 |
| 3.2.2.3 Variable Speed amb l'efecte de covariables | 34 |
| 3.2.2.4 Variable Urban amb l'efecte de covariables | 34 |
| 3.3 Models amb covariables constants en el temps | 34 |
| 3.3.1 Model TWFE | 35 |

| | | |
|---------|--|----|
| 3.3.1.1 | Variable Total amb l'efecte de covariables | 35 |
| 3.3.1.2 | Variable Night amb l'efecte de covariables | 37 |
| 3.3.1.3 | Variable Speed amb l'efecte de covariables | 39 |
| 3.3.1.4 | Variable Urban amb l'efecte de covariables | 41 |
| 3.3.2 | Models semi-paramètrics (IWE) | 43 |
| 3.3.2.1 | Variable Total amb l'efecte de covariables | 44 |
| 3.3.2.2 | Variable Night amb l'efecte de covariables | 45 |
| 3.3.2.3 | Variable Speed amb l'efecte de covariables | 46 |
| 3.3.2.4 | Variable Urban amb l'efecte de covariables | 48 |
| 3.3.3 | Models semi-paramètrics (IWE) amb dades balancejades | 49 |
| 3.3.3.1 | Variable Total amb l'efecte de covariables | 50 |
| 3.3.3.2 | Variable Night amb l'efecte de covariables | 51 |
| 3.3.3.3 | Variable Speed amb l'efecte de covariables | 51 |
| 3.3.3.4 | Variable Urban amb l'efecte de covariables | 54 |

1 Dades

Es seleccionen només aquells individus que:

- Han conduït 100 quilòmetres o més en el període de pre-tractament
- No han declarat accidents en els períodes de pre i post tractament

Les variables d'interès són doncs:

- “Quilòmetres totals” (Total)
- “Percentatge de quilòmetres durant la nit” (Night)
- “Percentatge de distància conduïda per sobre de la velocitat” (Speed)
- “Percentatge de quilòmetres en àrees urbanes” (Urban)

I en total compte amb 6064 assegurats.

2 Estadístics

De les quals a continuació prenem els seus descriptius en el pre i post període:

Table 1: Telematics variables in pre-treatment period t=1

| | Total | Night | Speed | Urban |
|-------|-----------|--------|--------|---------|
| Means | 4694.7386 | 6.7023 | 8.4074 | 27.8237 |
| STD | 4677.8705 | 6.9342 | 8.7962 | 15.4425 |
| Min | 100.4160 | 0.0000 | 0.0723 | 0.0000 |
| Q25 | 1306.4758 | 1.6372 | 2.5918 | 16.0602 |

Table 3: Means for pre and post-treatment periods by group

| Time | D | Total | Time | D | Night |
|------|---|----------|------|---|--------|
| 1 | 0 | 4693.403 | 1 | 0 | 6.6463 |
| 2 | 0 | 3256.540 | 2 | 0 | 6.7973 |
| 1 | 1 | 4707.992 | 1 | 1 | 7.2584 |
| 2 | 1 | 4089.594 | 2 | 1 | 7.4135 |

Table 4: Means for pre and post-treatment periods by group

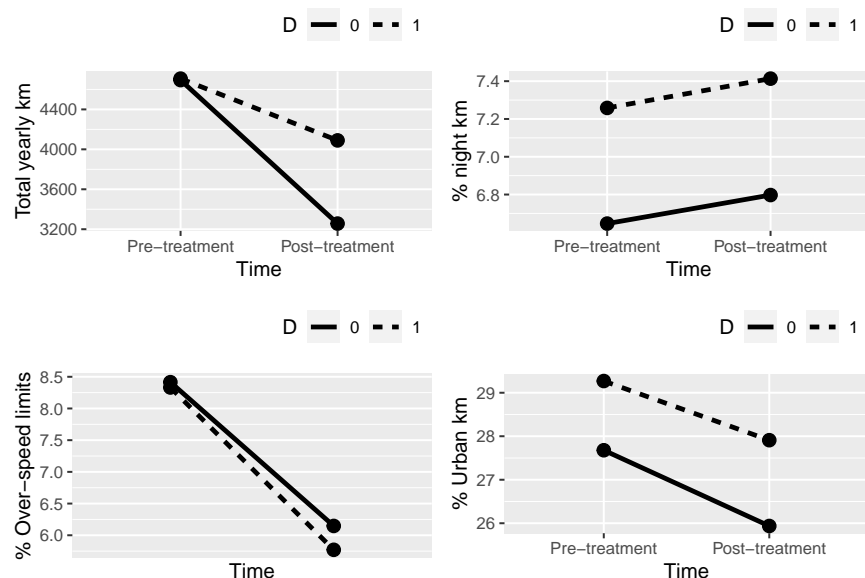
| Time | D | Speed | time | Time | D | Urban |
|------|---|--------|------|------|---|---------|
| 1 | 0 | 8.4150 | 1 | 1 | 0 | 27.6781 |
| 2 | 0 | 6.1473 | 2 | 2 | 0 | 25.9372 |
| 1 | 1 | 8.3312 | 1 | 1 | 1 | 29.2689 |
| 2 | 1 | 5.7707 | 2 | 2 | 1 | 27.9093 |

| | Total | Night | Speed | Urban |
|----------|------------|---------|---------|---------|
| Median | 3273.0655 | 4.5814 | 5.2816 | 24.5263 |
| Q75 | 6510.8150 | 9.7021 | 10.9821 | 36.6974 |
| Max | 39295.1860 | 60.3906 | 68.6549 | 95.9362 |
| Kurtosis | 1.9506 | 1.8043 | 2.1472 | 0.9486 |
| Skewness | 5.3704 | 4.5815 | 5.6431 | 0.7304 |

Table 2: Telematics variables in post-treatment period t=2

| | Total | Night | Speed | Urban |
|----------|------------|----------|---------|----------|
| Means | 3332.7845 | 6.8537 | 6.1129 | 26.1177 |
| STD | 3149.6803 | 7.9181 | 7.2510 | 16.1126 |
| Min | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Q25 | 1060.0473 | 1.2569 | 1.6409 | 14.2636 |
| Median | 2482.4205 | 4.3128 | 3.6111 | 22.4300 |
| Q75 | 4689.9253 | 9.8680 | 7.7456 | 34.5349 |
| Max | 34295.6730 | 100.0000 | 60.9904 | 100.0000 |
| Kurtosis | 1.9803 | 2.7066 | 2.5762 | 1.2323 |
| Skewness | 7.0360 | 15.6020 | 8.6906 | 1.8927 |

A continuació imprimim les mitjanes pels dos grups (control i tractament) i pels dos períodes (pre i post)



Pel que fa a la variable “Speed”, el seu gràfic té una fàcil interpretació. En tots dos cassos, tant si s’ha donat tractament com no el percentatge de quilòmetres recorreguts per sobre de la velocitat en el període de pre-tractament és major que en el post-tractament. Això pot ser degut a varis fenòmens, sent un dels principals, la introducció de varis radars en les vies urbanes. D’altra banda, la línia discontinua (amb tractament), es troba lleugerament inferior a la contínua, això indica que tal com esperat aquells individus que han patit algun accident, durant el següent any condueixen a menor velocitat, segurament degut a la por que han agafat al volant.

Pel que fa a la variable “Urban”, altre cop totes dues línies tornen a disminuir. Tanmateix, Aquest cop la línia que fa referència als individus que han tingut algun accident disminueix en menor mesura, indicant que aquests prefereixen recuperar la confiança al volant conduint per vies urbanes, segurament degut a que aquestes són més lentes.

Pel que fa a la variable “Night”, totes dues línies tenen una tendència ascendent similar. Tanmateix, sí que es pot notar que la línia discontinua es troba en tot moment per sobre, indicant que clarament aquells que condueixen amb major freqüència durant la nit, tenen una probabilitat major de tenir algun accident.

Finalment la variable “Total” és la més difícil d’explicar en tenir un comportament contrari a l’esperat. En aquest es veu clarament que el fet d’haver tingut un accident dona lloc a realitzar més quilòmetres. El fet que la unitat de mesura siguin els quilòmetres i no un percentatge, per exemple, pot ser una de les raons d’aquesta extranya conclusió.

D’altra banda, també s’afegiran les següents covariables als models:

- age= edat de l’assegurat
- age35= 1 si l’edat ≤ 35 (primer quartil aproximadament), = 0 altrament
- age_lic= edat de la llicència de conduir
- age_lic15= 1 si l’edat _lic ≤ 15 (primer quartil aproximadament), = 0 altrament
- parking_yes= 1 si s’utilitza pàrquing durant la nit, = 0 altrament
- woman= 1 si l’assegurada és una dona, = 0 altrament
- BMzones= 1 si la zona de conducció és Barcelona o Madrid, = 0 altrament

- power100= 1 si la potència del cotxe és ≤ 100 , = 0 altrament

Les variables edat i anys de llicència s’han binaritzat a nivells que s’han considerat rellevants per a poder realitzar un anàlisi en funció de valors concrets. S’han provat altres quartils i talls, però aquest ha estat el més encertat. Així doncs, durant els models no s’utilitzaran les variables numèriques com a tal, sinó les seves corresponents categòriques.

En les següents taules s’imprimeixen els descriptius estadístics de les covariables. Com que s’observen canvis en els períodes de pre i post tractament, s’assumirà que són covariables canviants en el temps durant els models.

Com a observació, fer notar que tenint en compte que sempre s’agafa l’edat a l’inscriure un primer registre i després es manté el mateix valor al llarg del temps, així com per la llicència de conduir, un es podria preguntar com és que les estadístiques d’aquestes variables poden canviar al llarg del temps. Així com la variable **woman**, ja que la única manera de canviar el sexe seria que la persona es fes pròpiament un canvi de sexe. Bé doncs, aquests cassos es donen generalment perquè es canvia el titular de la pòlissa, mantenint totes les altres característiques intactes.

Table 5: Covariates in pre-treatment period t=1

| | age | age35 | lic_age | lic_age15 | parking_yes | woman | BMzones | power100 |
|----------|---------|--------|---------|-----------|-------------|---------|---------|----------|
| Means | 37.8238 | 0.1723 | 17.5761 | 0.2010 | 0.6784 | 0.4715 | 0.1436 | 0.5935 |
| STD | 2.7727 | 0.3777 | 2.7888 | 0.4008 | 0.4671 | 0.4992 | 0.3507 | 0.4912 |
| Min | 31.1836 | 0.0000 | 13.0466 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Q25 | 35.7507 | 0.0000 | 15.3199 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Median | 37.7425 | 0.0000 | 17.1233 | 0.0000 | 1.0000 | 0.0000 | 0.0000 | 1.0000 |
| Q75 | 39.9212 | 0.0000 | 19.4096 | 0.0000 | 1.0000 | 1.0000 | 0.0000 | 1.0000 |
| Max | 45.4384 | 1.0000 | 42.4575 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| Kurtosis | 0.0495 | 1.7348 | 0.6729 | 1.4917 | -0.7638 | 0.1143 | 2.0317 | -0.3806 |
| Skewness | -0.7218 | 1.0098 | 0.5678 | 0.2251 | -1.4168 | -1.9873 | 2.1281 | -1.8554 |

Table 6: Covariates in post-treatment period t=2

| | age | age35 | lic_age | lic_age15 | parking_yes | woman | BMzones | power100 |
|----------|---------|--------|---------|-----------|-------------|---------|---------|----------|
| Means | 37.8373 | 0.1774 | 17.5746 | 0.2004 | 0.6994 | 0.4987 | 0.1405 | 0.5777 |
| STD | 2.8122 | 0.3821 | 2.8013 | 0.4003 | 0.4586 | 0.5000 | 0.3475 | 0.4940 |
| Min | 29.5288 | 0.0000 | 11.3068 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Q25 | 35.7144 | 0.0000 | 15.3151 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Median | 37.7795 | 0.0000 | 17.0795 | 0.0000 | 1.0000 | 0.0000 | 0.0000 | 1.0000 |
| Q75 | 40.0144 | 0.0000 | 19.4438 | 0.0000 | 1.0000 | 1.0000 | 0.0000 | 1.0000 |
| Max | 47.2329 | 1.0000 | 28.2767 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| Kurtosis | 0.0068 | 1.6882 | 0.5509 | 1.4968 | -0.8694 | 0.0053 | 2.0685 | -0.3144 |
| Skewness | -0.7576 | 0.8501 | -0.4617 | 0.2404 | -1.2443 | -2.0003 | 2.2791 | -1.9014 |

Tanmateix, varis dels mètodes amb que es treballa durant aquest projecte requereixen de covariables constants en el temps. Així doncs, es seleccionaran només aquells registres que tinguin aquestes covariables constants.

Table 7: Covariates in pre-treatment period t=1

| | age | age35 | lic_age | lic_age15 | parking_yes | woman | BMzones | power100 |
|----------|---------|--------|---------|-----------|-------------|---------|---------|----------|
| Means | 37.9511 | 0.1573 | 17.7626 | 0.1772 | 0.6528 | 0.4933 | 0.1458 | 0.5855 |
| STD | 2.7369 | 0.3641 | 2.7509 | 0.3819 | 0.4762 | 0.5000 | 0.3529 | 0.4927 |
| Min | 31.1836 | 0.0000 | 13.0466 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Q25 | 35.9315 | 0.0000 | 15.5699 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Median | 37.9123 | 0.0000 | 17.4411 | 0.0000 | 1.0000 | 0.0000 | 0.0000 | 1.0000 |
| Q75 | 40.0370 | 0.0000 | 19.6219 | 0.0000 | 1.0000 | 1.0000 | 0.0000 | 1.0000 |
| Max | 44.4630 | 1.0000 | 26.1507 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| Kurtosis | -0.0117 | 1.8818 | 0.4776 | 1.6903 | -0.6415 | 0.0270 | 2.0070 | -0.3471 |
| Skewness | -0.7510 | 1.5417 | -0.5418 | 0.8575 | -1.5889 | -1.9998 | 2.0284 | -1.8800 |

Table 8: Covariates in post-treatment period t=2

| | age | age35 | lic_age | lic_age15 | parking_yes | woman | BMzones | power100 |
|----------|---------|--------|---------|-----------|-------------|---------|---------|----------|
| Means | 37.9511 | 0.1573 | 17.7626 | 0.1772 | 0.6528 | 0.4933 | 0.1458 | 0.5855 |
| STD | 2.7369 | 0.3641 | 2.7509 | 0.3819 | 0.4762 | 0.5000 | 0.3529 | 0.4927 |
| Min | 31.1836 | 0.0000 | 13.0466 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Q25 | 35.9315 | 0.0000 | 15.5699 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Median | 37.9123 | 0.0000 | 17.4411 | 0.0000 | 1.0000 | 0.0000 | 0.0000 | 1.0000 |
| Q75 | 40.0370 | 0.0000 | 19.6219 | 0.0000 | 1.0000 | 1.0000 | 0.0000 | 1.0000 |
| Max | 44.4630 | 1.0000 | 26.1507 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| Kurtosis | -0.0117 | 1.8818 | 0.4776 | 1.6903 | -0.6415 | 0.0270 | 2.0070 | -0.3471 |
| Skewness | -0.7510 | 1.5417 | -0.5418 | 0.8575 | -1.5889 | -1.9998 | 2.0284 | -1.8800 |

En aquest cas queden 4075 assegurats al conjunt filtrat.

3 Resultats de l'estimació Diff-in-Diff per a les reclamacions ATT

3.1 Models sense covariables

3.1.1 Models de regressió TWFE i OLS

Estimem l' ATT_2 definit com:

$$ATT_2 = E[Y_{i2}(1) - Y_{i2}(0) | D_i = 1]. \quad (1)$$

utilitzant els models de regressió TWFE i OLS sense covariables

3.1.1.1 Variable Total

Table 9: TWFE for Total variable

| | Estimate | Std. Error | t-value | Pr(> t) | R2 |
|----|----------|------------|---------|----------|--------|
| Tr | 818.4658 | 243.7359 | 3.358 | 8e-04 | 0.5425 |

Table 10: OLS for Total variable

| | Estimate | Std. Error | t value | Pr(> t) |
|-------------|------------|------------|----------|----------|
| (Intercept) | 6130.2665 | 120.0352 | 51.0706 | 0.0000 |
| Time | -1436.8631 | 75.9169 | -18.9268 | 0.0000 |
| D | 14.5883 | 177.4421 | 0.0822 | 0.9345 |
| Tr | 818.4658 | 250.9411 | 3.2616 | 0.0011 |

3.1.1.2 Variable Night

Table 11: TWFE for Night variable

| | Estimate | Std. Error | t-value | Pr(> t) | R2 |
|----|----------|------------|---------|----------|--------|
| Tr | 0.0041 | 0.4411 | 0.0093 | 0.9926 | 0.5573 |

Table 12: OLS for Night variable

| | Estimate | Std. Error | t value | Pr(> t) |
|-------------|----------|------------|---------|----------|
| (Intercept) | 6.4953 | 0.2242 | 28.9752 | 0.0000 |
| Time | 0.1510 | 0.1418 | 1.0651 | 0.2869 |
| D | 0.6120 | 0.3314 | 1.8469 | 0.0648 |
| Tr | 0.0041 | 0.4686 | 0.0088 | 0.9930 |

3.1.1.3 Variable Speed

Table 13: TWFE for Speed variable

| | Estimate | Std. Error | t-value | Pr(> t) | R2 |
|----|----------|------------|---------|----------|--------|
| Tr | -0.2929 | 0.3854 | -0.76 | 0.4473 | 0.7177 |

Table 14: OLS for Speed variable

| | Estimate | Std. Error | t value | Pr(> t) |
|-------------|----------|------------|----------|----------|
| (Intercept) | 10.6827 | 0.2428 | 43.9889 | 0.0000 |
| Time | -2.2677 | 0.1536 | -14.7643 | 0.0000 |
| D | -0.0838 | 0.3590 | -0.2334 | 0.8154 |
| Tr | -0.2929 | 0.5077 | -0.5769 | 0.5640 |

3.1.1.4 Variable Urban

Table 15: TWFE for Urban variable

| | Estimate | Std. Error | t-value | Pr(> t) | R2 |
|----|----------|------------|---------|----------|--------|
| Tr | 0.3813 | 0.7156 | 0.5328 | 0.5942 | 0.7416 |

Table 16: OLS for Urban variable

| | Estimate | Std. Error | t value | Pr(> t) |
|-------------|----------|------------|---------|----------|
| (Intercept) | 29.4191 | 0.4752 | 61.9070 | 0.0000 |
| Time | -1.7409 | 0.3006 | -5.7924 | 0.0000 |
| D | 1.5908 | 0.7025 | 2.2645 | 0.0236 |
| Tr | 0.3813 | 0.9935 | 0.3838 | 0.7011 |

Amb aquests primers resultats notem que l'efecte del tractament només és estadísticament significatiu per la variable **Total**, amb signe positiu. Això indica que extranyament, com més accidents es tenen, més quilòmetres s'acaben duen a terme durant el següent any després d'haver tingut l'accident.

Pel que fa a la interpretació dels models, fer notar que el coeficient **Time** pel cas en que es realitza el model amb OLS (per a **Total**), és igual a -1436.8631, perquè és el resultat de fer 3256.540 (T=2, D=0) - 4693.403 (T=1, D=0), que serien els dos primers valors de les taules que mostraven les mitjanes pels períodes de pre i post tractament, per grups. D'altra banda, el coeficient de **D**, 14.5883, prové de fer 4707.992 (T=1, D=1) - 4693.403 (T=1, D=0). Finalment, el coeficient de **Tr** prové de 4089.594 (T=2, D=1) - 3256.540 (T=2, D=1) + 14.5883.

Tanmateix, en realitat se sap que el fet de tenir un accident no és completament aleatòri, sinó que és aleatòri, condicionat a unes característiques. Així doncs més endavant es realitzarà models similars però afegint covariables.

3.1.2 Estimació no paramètrica amb distribucions empíriques

Aquest mètode és una generalització del Dif-in-Dif anomenat Changes-in-Changes. Com a fets rellevants, assumeix que les característiques no són observables, però que dins de cada grup (control i tractament), la distribució d'aquestes característiques no canvia al llarg del temps. És per aquest raó, que tot i no treballar amb covariables en aquest model, s'ha treballat amb el conjunt de dades reduït, el que mostra les covariables constants al llarg del temps.

Com a exercici extra en comparació als altres mètodes, s'ha decidit calcular també els quartils. Per saber si aquests són estadísticament significatius s'ha calculat l'interval de confiança. Si conté el 0 no és estadísticament significatiu.

Els resultats mostren clarament que només la variable dependent **Total** mostra un coeficient similar al mostrar al mètode anterior i significatiu. Aquests resultats no sorprenen ja que el cap i a la fi aquest mètode està realitzant un procediment similar al TWFE: una metodologia no paramètrica amb una inferència basada en la distribució normal.

Pel que fa als quartils es nota que a mesura que augmentem el quartil, augmenta el valor del coeficient en valor absolut en la majoria dels cassos, tanmateix, els resultats només són estadísticament significatius per a la variable **Total**, excepte per al quartil 0.99 en ser aquest massa extrem. Això equival a dir que el número de quilòmetres (pel cas de la variable **Total**), augmenta pels usuaris que no han declarat cap accident durant el segon període, que és el resultat que s'esperia pensant en sentit comú, ja que són els usuaris que han tingut un accident els que agafen por a la conducció, no al contrari. Està dient doncs que els promitjos no donen informació rellevant perquè els efectes es contraresten, però que mirant els quartils es pot notar la direcció de creixement o decreixement de les variables dependents en funció del grup.

3.1.2.1 Variable Total

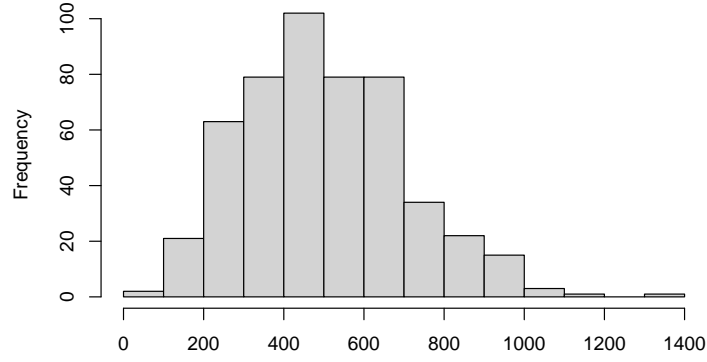
Table 17: Changes-in-changes for Total variable

| | ATT_2 | sd | z_stat | p_value | CI_lower | CI_upper |
|---------|----------|----------|--------|---------|----------|----------|
| 2-tails | 1001.715 | 275.2579 | 3.6392 | 3e-04 | 462.2198 | 1541.211 |
| 1-tail | 1001.715 | 275.2579 | 3.6392 | 1e-04 | 548.9564 | Inf |

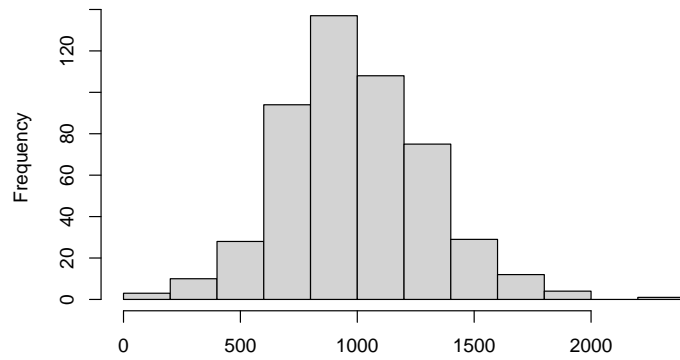
Table 18: Changes-in-changes for Total variable - quantiles

| | ATT_2 | CI_lower | CI_upper |
|------|----------|-----------|-----------|
| 0.25 | 472.530 | 203.854 | 838.577 |
| 0.5 | 990.871 | 529.478 | 1524.050 |
| 0.75 | 831.629 | 141.984 | 1753.910 |
| 0.9 | 1782.700 | 606.162 | 3105.809 |
| 0.95 | 2622.827 | 114.321 | 3945.219 |
| 0.99 | 2201.995 | -2155.281 | 10233.921 |

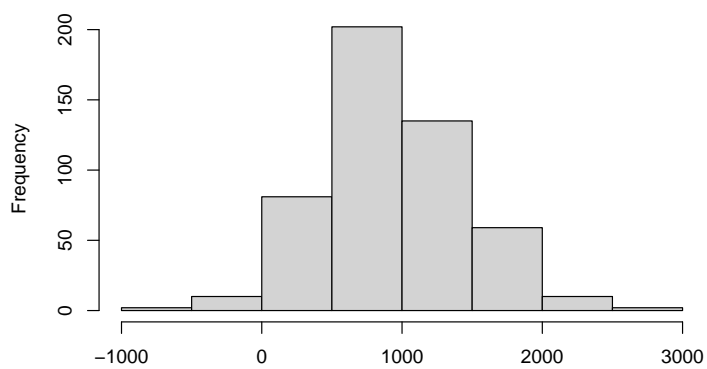
quantil 0.25



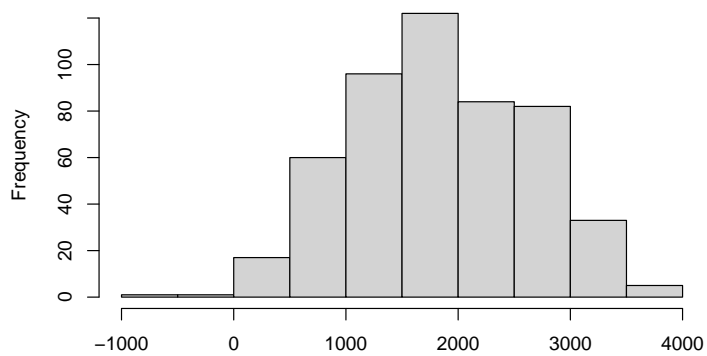
quantil 0.5



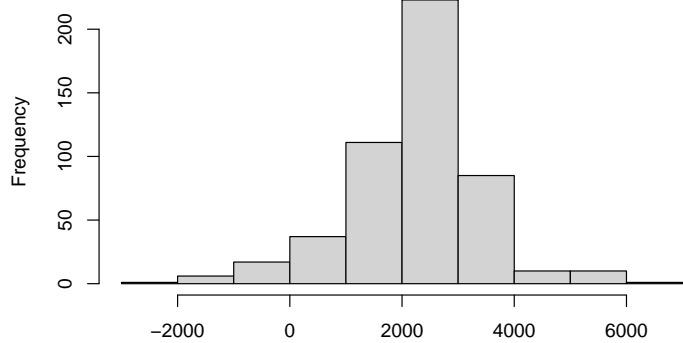
quantil 0.75

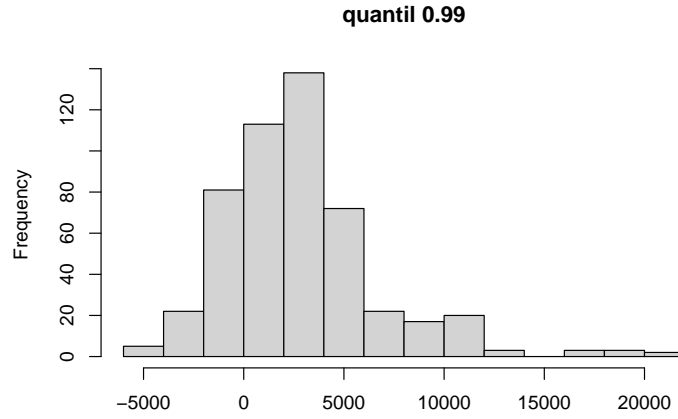


quantil 0.9



quantil 0.95





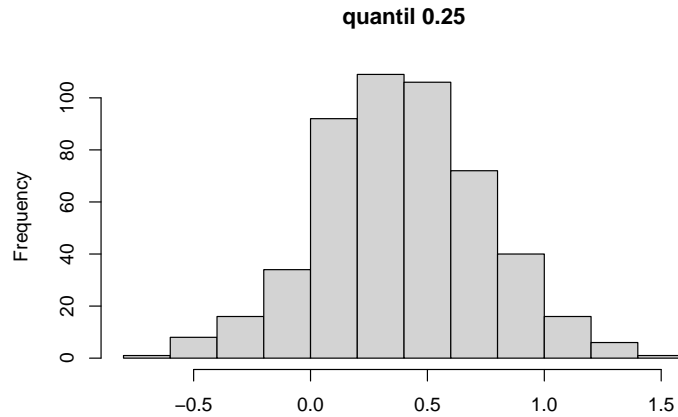
3.1.2.2 Variable Night

Table 19: Changes-in-changes for Night variable

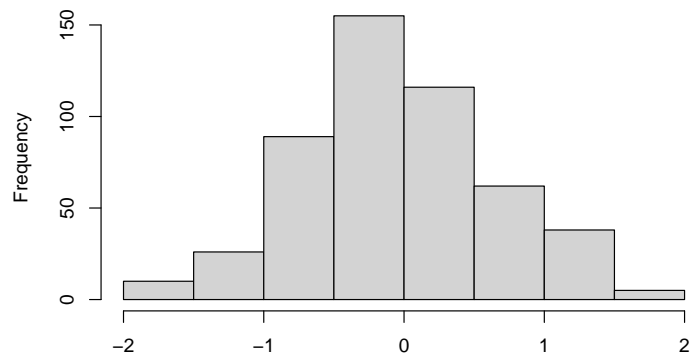
| | ATT_2 | sd | z_stat | p_value | CI_lower | CI_upper |
|---------|--------|--------|--------|---------|----------|----------|
| 2-tails | 0.1949 | 0.6296 | 0.3096 | 0.7569 | -1.0390 | 1.4288 |
| 1-tail | 0.1949 | 0.6296 | 0.3096 | 0.3784 | -0.8406 | Inf |

Table 20: Changes-in-changes for Night variable - quantiles

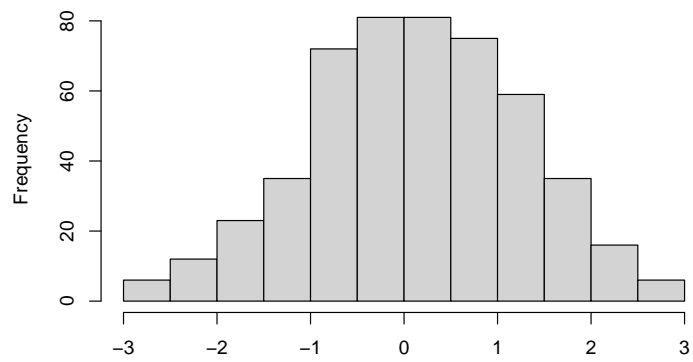
| | ATT_2 | CI_lower | CI_upper |
|------|---------|----------|----------|
| 0.25 | 0.4424 | -0.1940 | 0.9881 |
| 0.5 | -0.0381 | -1.1459 | 1.1515 |
| 0.75 | 0.6812 | -1.7821 | 1.9141 |
| 0.9 | 0.9619 | -1.2339 | 4.0262 |
| 0.95 | 0.2223 | -3.2888 | 4.6907 |
| 0.99 | -3.9011 | -11.9132 | 20.3423 |



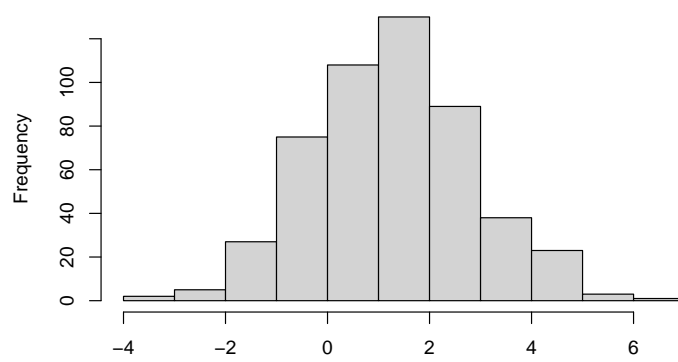
quantil 0.5

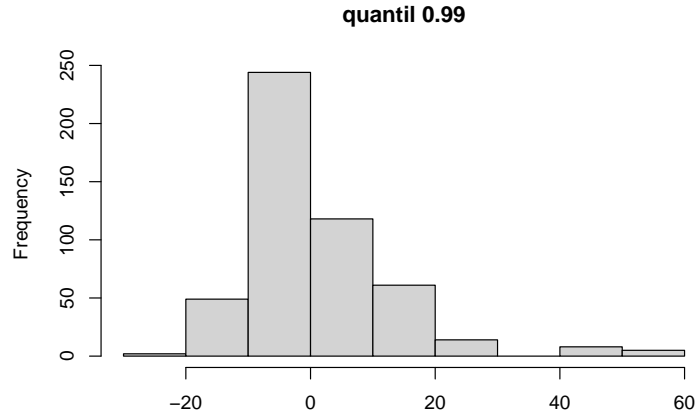
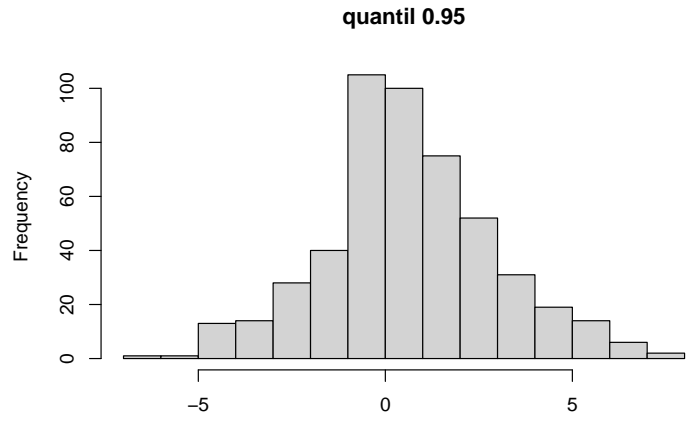


quantil 0.75



quantil 0.9





3.1.2.3 Variable Speed

Table 21: Changes-in-changes for Speed variable

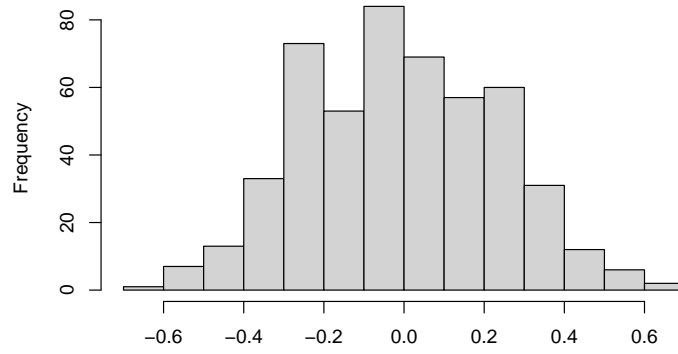
| | ATT_2 | sd | z_stat | p_value | CI_lower | CI_upper |
|---------|---------|--------|---------|---------|----------|----------|
| 2-tails | -0.0153 | 0.5395 | -0.0283 | 0.9774 | -1.0727 | 1.0421 |
| 1-tail | -0.0153 | 0.5395 | -0.0283 | 0.4887 | -Inf | 0.8721 |

Table 22: Changes-in-changes for Speed variable - quantiles

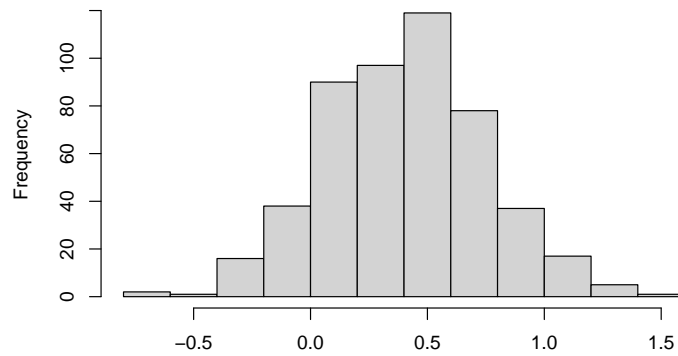
| | ATT_2 | CI_lower | CI_upper |
|------|---------|----------|----------|
| 0.25 | -0.0257 | -0.3658 | 0.3668 |
| 0.5 | 0.4094 | -0.1535 | 0.9886 |
| 0.75 | 0.6278 | -0.4226 | 1.4970 |
| 0.9 | 0.5994 | -2.7302 | 2.9519 |
| 0.95 | -1.6100 | -5.0113 | 5.3885 |

| | ATT_2 | CI_lower | CI_upper |
|------|---------|----------|----------|
| 0.99 | -8.2742 | -19.1886 | -0.1008 |

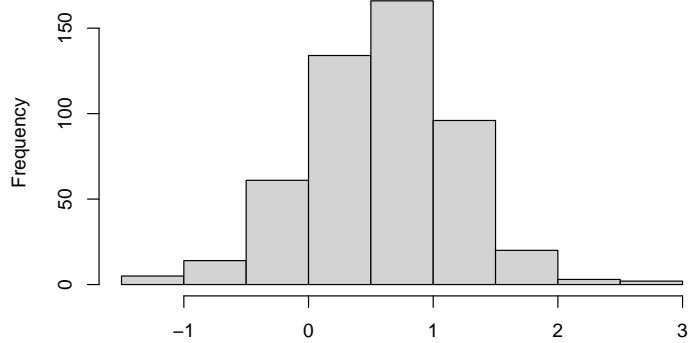
quantil 0.25



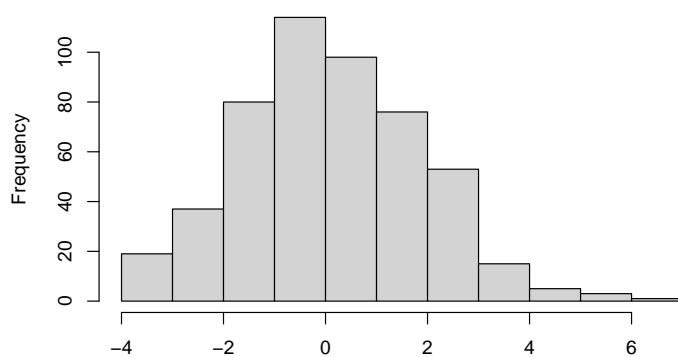
quantil 0.5



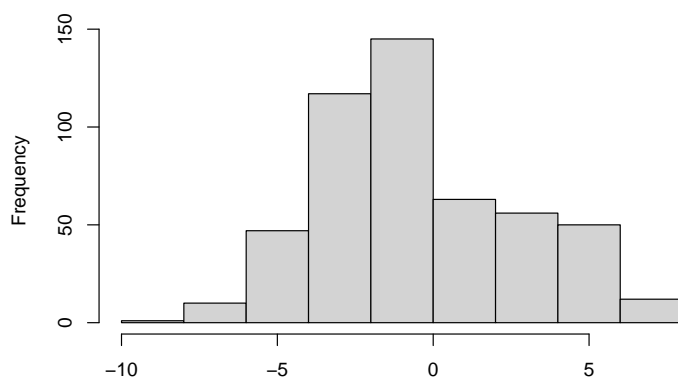
quantil 0.75

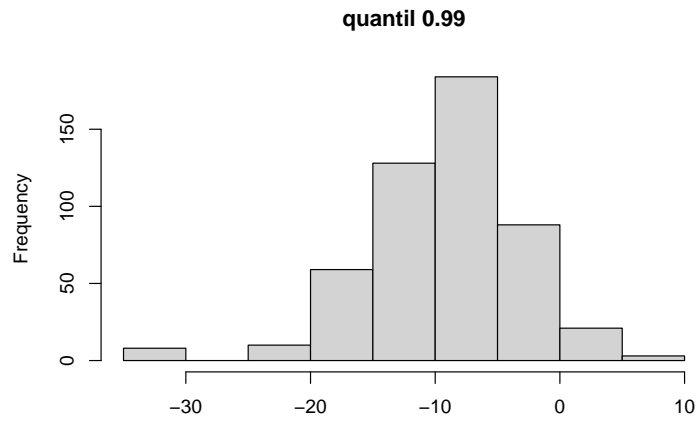


quantil 0.9



quantil 0.95





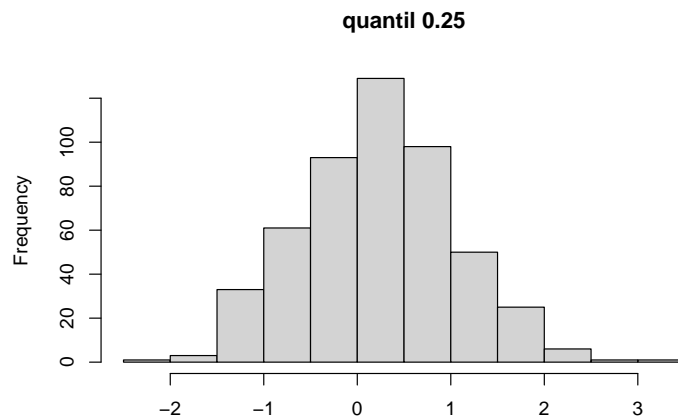
3.1.2.4 Variable Urban

Table 23: Changes-in-changes for Urban variable

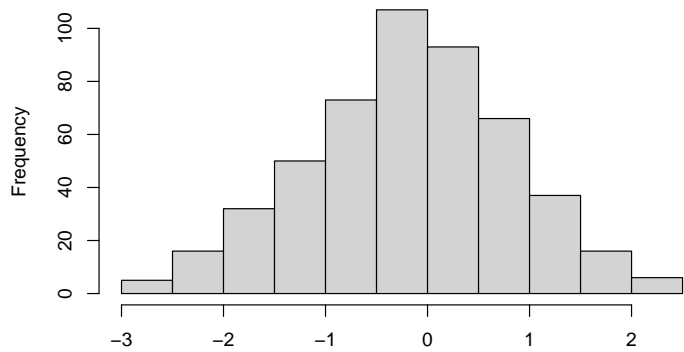
| | ATT_2 | sd | z_stat | p_value | CI_lower | CI_upper |
|---------|---------|--------|---------|---------|----------|----------|
| 2-tails | -0.2988 | 1.2079 | -0.2474 | 0.8046 | -2.6664 | 2.0687 |
| 1-tail | -0.2988 | 1.2079 | -0.2474 | 0.4023 | -Inf | 1.6881 |

Table 24: Changes-in-changes for Urban variable - quantiles

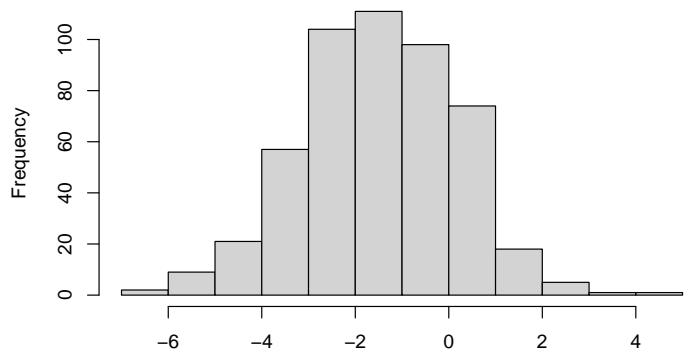
| | ATT_2 | CI_lower | CI_upper |
|------|---------|----------|----------|
| 0.25 | 0.4201 | -1.1137 | 1.5780 |
| 0.5 | 0.0553 | -1.8360 | 1.4155 |
| 0.75 | -1.2353 | -4.2262 | 0.9736 |
| 0.9 | 1.3411 | -4.6465 | 4.3217 |
| 0.95 | 2.5132 | -4.0373 | 5.4735 |
| 0.99 | -2.1839 | -18.1051 | 7.0458 |



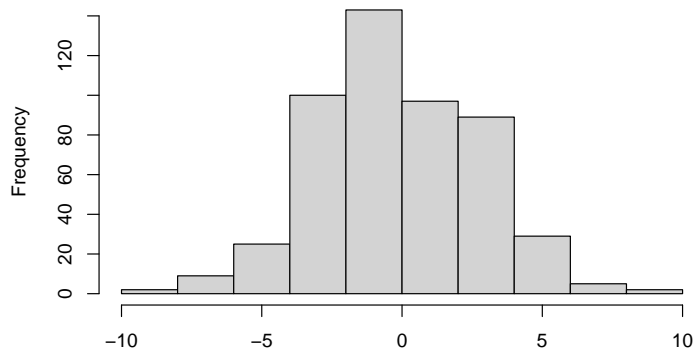
quantil 0.5

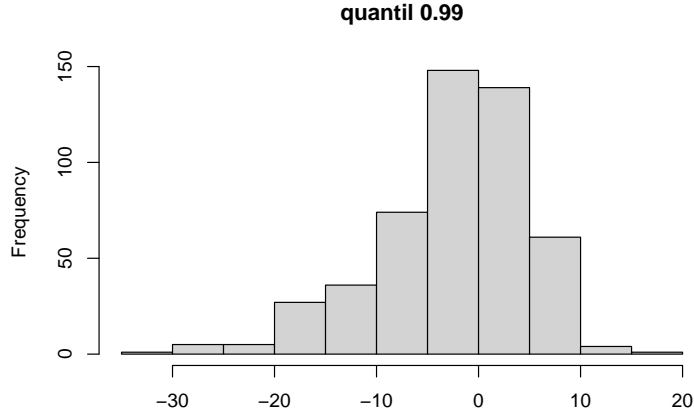
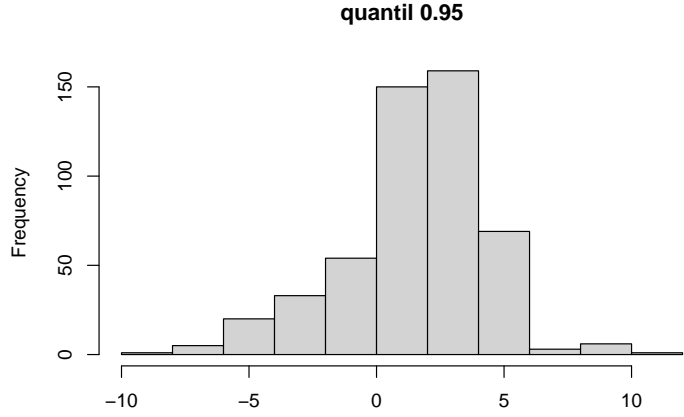


quantil 0.75



quantil 0.9





3.2 Models amb covariables canviant en el temps

3.2.1 Model TWFE

Aquest model amb totes les covariables així com aquestes al temps $t = 2$ i canviant en el temps s'està referint a la següent formulació:

$$Y_{it} = \alpha_i + \delta_t + \tau \cdot Tr_{it} + \beta \cdot X_{it} \cdot I(t = 2) + \gamma \cdot X_{it} \cdot Tr_{it} + \delta \cdot X_{it} + \epsilon_{it}, \quad (2)$$

3.2.1.1 Variable Total amb l'efecte de covariables En primer lloc creem el model amb tots els efectes additius i multiplicatius de les covariables. Es pot clarament observar que no tots els components del model són significatius en tenir pvalors molt superiors al 0.05 (agafant un nivell de confiança del 95%). Concretament, les interaccions creades entre les covariables i Tr semblen clarament ser no significatives. Així doncs, s'ha decidit seleccionar el millor model maximitzant l'R quadrat ajustat. El resultat del millor model mostra un valor estimat per ATT_2 positiu i significatiu al 0.09%. Aquesta millora s'ha hagut de fer manualment, el qual implica que sempre pot haver-hi major tendència a errors.

Table 25: TWFE for Total variable with covariates

| | Estimate | Std. Error | t-value | Pr(> t) | R2 | R2Adj |
|-----------------|-----------|------------|---------|----------|--------|-------|
| Tr | 800.4120 | 563.1053 | 1.4214 | 0.1552 | 0.5465 | 0.09 |
| age35 | 127.2525 | 275.3050 | 0.4622 | 0.6439 | 0.5465 | 0.09 |
| age35_2 | -271.2564 | 228.3853 | -1.1877 | 0.2350 | 0.5465 | 0.09 |
| lic_age15 | -412.7710 | 254.9081 | -1.6193 | 0.1054 | 0.5465 | 0.09 |
| lic_age15_2 | 162.6168 | 219.5563 | 0.7407 | 0.4589 | 0.5465 | 0.09 |
| parking_yes | 929.0184 | 257.0805 | 3.6137 | 0.0003 | 0.5465 | 0.09 |
| parking_yes2 | -686.6278 | 169.0524 | -4.0616 | 0.0000 | 0.5465 | 0.09 |
| woman | -732.9228 | 206.2766 | -3.5531 | 0.0004 | 0.5465 | 0.09 |
| woman_2 | 218.7486 | 163.8448 | 1.3351 | 0.1819 | 0.5465 | 0.09 |
| BMzones | -528.8030 | 451.1723 | -1.1721 | 0.2412 | 0.5465 | 0.09 |
| BMzones_2 | 274.4212 | 217.6779 | 1.2607 | 0.2075 | 0.5465 | 0.09 |
| power100 | -230.0002 | 216.6729 | -1.0615 | 0.2885 | 0.5465 | 0.09 |
| power100_2 | 30.9797 | 166.8933 | 0.1856 | 0.8527 | 0.5465 | 0.09 |
| Tr:age35_2 | 91.8772 | 701.0373 | 0.1311 | 0.8957 | 0.5465 | 0.09 |
| Tr:lic_age15_2 | 253.1422 | 669.3358 | 0.3782 | 0.7053 | 0.5465 | 0.09 |
| Tr:parking_yes2 | -61.3778 | 527.5085 | -0.1164 | 0.9074 | 0.5465 | 0.09 |
| Tr:woman_2 | -345.7381 | 504.1764 | -0.6857 | 0.4929 | 0.5465 | 0.09 |
| Tr:BMzones_2 | 70.2823 | 683.2903 | 0.1029 | 0.9181 | 0.5465 | 0.09 |
| Tr:power100_2 | 255.7215 | 508.8650 | 0.5025 | 0.6153 | 0.5465 | 0.09 |

Table 26: TWFE for Total variable with covariates

| | Estimate | Std. Error | t-value | Pr(> t) | R2 | R2Adj |
|--------------|-----------|------------|---------|----------|-------|-------|
| Tr | 807.3561 | 243.0341 | 3.3220 | 0.0009 | 0.546 | 0.091 |
| lic_age15 | -360.3023 | 218.4160 | -1.6496 | 0.0991 | 0.546 | 0.091 |
| parking_yes | 975.3139 | 255.3572 | 3.8194 | 0.0001 | 0.546 | 0.091 |
| parking_yes2 | -706.2360 | 159.2244 | -4.4355 | 0.0000 | 0.546 | 0.091 |
| woman | -799.2458 | 199.7551 | -4.0011 | 0.0001 | 0.546 | 0.091 |
| woman_2 | 221.5632 | 151.6239 | 1.4613 | 0.1440 | 0.546 | 0.091 |

Altrament, també hauríem pogut decidir agafar l'AIC com a criteri de selecció del millor model. Existeix una funció `stepAIC()` que selecciona el millor model donat el model complet. Tanmateix, aquesta requereix crear un model lineal (`lm`) com fet a continuació.

Abans de crear el model, centrarem els valors de les variables, encara que les covariables siguin binàries, perquè així ho requereix el model. A més a més, com que s'han de centrar a 0 totes les covariables introduïdes, creem també una variable per a les interaccions que es volen introduir en el model. S'ha decidit centrar les variables, perquè aquest procés dona lloc al mateix resultat que incloure variables fictícies (*dummy*) pels afectes fixos, però evita haver d'introduir tantes variables en el model amb les qual a vegades R no és ni capaç de tractar. A més a més, com més variables s'introdueixen en el model, major robustesa es perd.

Com podem observar, els coeficients obtinguts són exactament iguals que els obtinguts amb la funció `plm`, sent les significacions d'aquests (*p*valors) el que canvia de forma poc notòria. Així doncs, a partir d'ara es treballarà amb aquest model en ser més fàcil d'optimitzar. Així mateix, amb aquest tipus de model s'ha de tenir en compte que l'R quadrat ajustat fa referència a les variables centrades (transformades). Així doncs, s'ha de calcular altre cop aquesta mesura de forma manual per a obtenir un resultat fiable d'aquest ajust. Com a últim comentari sobre aquesta manera de construir el model, dir que el fet d'optimitzar, com es farà a continuació, de forma automàtica, pot implicar que la variable "Tr" s'elimini d'aquest. Això s'ha d'evitar en tot moment en ser el seu coeficient el que dona el valor de l' ATT_2 .

Table 27: TWFE for Total variable with covariates

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|-----------------|-----------|------------|---------|----------|--------|--------|----------|
| Tr | 800.4120 | 397.8301 | 2.0119 | 0.0442 | 0.5495 | 0.0961 | 226320.4 |
| age35 | 127.2525 | 194.5011 | 0.6543 | 0.5130 | 0.5495 | 0.0961 | 226320.4 |
| age35_2 | -271.2564 | 161.3527 | -1.6811 | 0.0928 | 0.5495 | 0.0961 | 226320.4 |
| lic_age15 | -412.7710 | 180.0908 | -2.2920 | 0.0219 | 0.5495 | 0.0961 | 226320.4 |
| lic_age15_2 | 162.6168 | 155.1150 | 1.0484 | 0.2945 | 0.5495 | 0.0961 | 226320.4 |
| parking_yes | 929.0184 | 181.6257 | 5.1150 | 0.0000 | 0.5495 | 0.0961 | 226320.4 |
| parking_yes2 | -686.6278 | 119.4344 | -5.7490 | 0.0000 | 0.5495 | 0.0961 | 226320.4 |
| woman | -732.9228 | 145.7330 | -5.0292 | 0.0000 | 0.5495 | 0.0961 | 226320.4 |
| woman_2 | 218.7486 | 115.7553 | 1.8898 | 0.0588 | 0.5495 | 0.0961 | 226320.4 |
| BMzones | -528.8030 | 318.7503 | -1.6590 | 0.0971 | 0.5495 | 0.0961 | 226320.4 |
| BMzones_2 | 274.4212 | 153.7880 | 1.7844 | 0.0744 | 0.5495 | 0.0961 | 226320.4 |
| power100 | -230.0002 | 153.0779 | -1.5025 | 0.1330 | 0.5495 | 0.0961 | 226320.4 |
| power100_2 | 30.9797 | 117.9090 | 0.2627 | 0.7928 | 0.5495 | 0.0961 | 226320.4 |
| age35_2_Tr | 91.8772 | 495.2782 | 0.1855 | 0.8528 | 0.5495 | 0.0961 | 226320.4 |
| lic_age15_2_Tr | 253.1422 | 472.8813 | 0.5353 | 0.5924 | 0.5495 | 0.0961 | 226320.4 |
| parking_yes2_Tr | -61.3778 | 372.6812 | -0.1647 | 0.8692 | 0.5495 | 0.0961 | 226320.4 |
| woman_2_Tr | -345.7381 | 356.1973 | -0.9706 | 0.3317 | 0.5495 | 0.0961 | 226320.4 |
| BMzones_2_Tr | 70.2823 | 482.7401 | 0.1456 | 0.8842 | 0.5495 | 0.0961 | 226320.4 |
| power100_2_Tr | 255.7215 | 359.5098 | 0.7113 | 0.4769 | 0.5495 | 0.0961 | 226320.4 |

Com a resultat del millor model obtingut després de l’optimització prèviament esmentada, es conclou inicialment que és homogeni perquè no hi ha efectes interactius entre les variables. Pel que fa a l’efecte del tractament, ATT_2 , aquest és positiu i significatiu, indicant que aquells usuaris que han tingut una accident passen a recórrer més quilòmetres totals, després de tenir-lo, que aquells que no n’han tingut cap.

L’efecte de les variables a $t=2$ és la suma de la variable + variable_2. Així doncs, per exemple l’efecte de la variable parking_yes a $t=2$ és significativa i positiva. Això implica que aquells usuaris que tenen pàrquing, recorren més quilòmetres totals, el qual té molt sentit ja que generalment són aquells que tenen més diners i que per tant es poden permetre més gasolina i més quilòmetres. Tanmateix, la variable parking_yes2 en realitat mostra un efecte negatiu. Això indica que l’efecte entre $t=1$ i $t=2$ s’ha reduït, és a dir, que durant el primer període es recorrien molts més quilòmetres totals si es tenia pàrquing, mentre aquest efecte tot i que continua sent positiu, ha disminuït durant el segon període.

Contràriament, les dones recòrrer menys quilòmetres totals que els homes, tot i que la diferència entre homes i dones es redueix durant el període de post-tractament. També recorren menys quilòmetres totals aquells que s’han tret el carnet fa menys de 15 anys. La variable BMzones, que recordem que és = 1 si la zona de conducció és Barcelona o Madrid i 0 altrament, també té un efecte negatiu a $t=2$ però aquest no és significatiu, tot i que millori l’explicació del model.

Table 28: TWFE for Total variable with reduced covariates

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|--------------|-----------|------------|---------|----------|--------|--------|----------|
| Tr | 815.3623 | 171.8530 | 4.7445 | 0.0000 | 0.5491 | 0.0967 | 226305.8 |
| lic_age15 | -330.2755 | 155.1279 | -2.1291 | 0.0333 | 0.5491 | 0.0967 | 226305.8 |
| parking_yes | 931.3211 | 181.3547 | 5.1354 | 0.0000 | 0.5491 | 0.0967 | 226305.8 |
| parking_yes2 | -696.8813 | 112.7775 | -6.1793 | 0.0000 | 0.5491 | 0.0967 | 226305.8 |
| woman | -742.8234 | 144.9557 | -5.1245 | 0.0000 | 0.5491 | 0.0967 | 226305.8 |
| woman_2 | 208.5263 | 107.3896 | 1.9418 | 0.0522 | 0.5491 | 0.0967 | 226305.8 |
| BMzones | -530.9050 | 317.5070 | -1.6721 | 0.0945 | 0.5491 | 0.0967 | 226305.8 |
| BMzones_2 | 281.5846 | 146.3098 | 1.9246 | 0.0543 | 0.5491 | 0.0967 | 226305.8 |

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|----------|-----------|------------|---------|----------|--------|--------|----------|
| power100 | -207.3325 | 141.4845 | -1.4654 | 0.1428 | 0.5491 | 0.0967 | 226305.8 |

Només com a comprovació, fer notar que de fet l'R quadrat ajustat que haguéssim obtingut si haguéssim posat aquestes variables en el model plm és millor (major), indicant que utilitzar funcions és més fiables, raó per la qual a partir d'ara s'utilitzaran models lm per a aquest propòsit:

Table 29: TWFE for Total variable with covariates

| | Estimate | Std. Error | t-value | Pr(> t) | R2 | R2Adj |
|--------------|-----------|------------|---------|----------|--------|--------|
| Tr | 815.3623 | 243.1473 | 3.3534 | 0.0008 | 0.5463 | 0.0911 |
| lic_age15 | -330.2755 | 219.4836 | -1.5048 | 0.1324 | 0.5463 | 0.0911 |
| parking_yes | 931.3211 | 256.5908 | 3.6296 | 0.0003 | 0.5463 | 0.0911 |
| parking_yes2 | -696.8813 | 159.5639 | -4.3674 | 0.0000 | 0.5463 | 0.0911 |
| woman | -742.8234 | 205.0914 | -3.6219 | 0.0003 | 0.5463 | 0.0911 |
| woman_2 | 208.5263 | 151.9408 | 1.3724 | 0.1700 | 0.5463 | 0.0911 |
| BMzones | -530.9050 | 449.2266 | -1.1818 | 0.2373 | 0.5463 | 0.0911 |
| BMzones_2 | 281.5846 | 207.0073 | 1.3603 | 0.1738 | 0.5463 | 0.0911 |
| power100 | -207.3325 | 200.1801 | -1.0357 | 0.3004 | 0.5463 | 0.0911 |

Com a analizi extra, s'ha volgut mirar quina variable és més important per a saber el nombre total de quilòmetres recorreguts d'un assegurat, l'edat o els anys que fa que pot conduir. El model deixant únicament la variable `lic_age15` és exactament el mateix que quan teníem totes dues variables, mentre al deixar únicament l'edat, aquesta ha "reemplaçat" el lloc de `lic_age15` en el model, en trobar exactament els mateixos coeficients i magnituds, amb significacions lleugerament diferents. No obstant això, tot i que la variable `age35` apareix en el model aquesta és clarament no significativa, mentre la variable `lic_age15` sí que presentava un coeficient significatiu. Això indica que tot i que es necessita una mesura del temps en el model, els anys de llicència permeten aproximar amb més precisió els quilòmetres totals recorreguts pels assegurats.

Table 30: TWFE for Total variable with reduced covariates without `lic_age15`

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|--------------|-----------|------------|---------|----------|--------|--------|----------|
| Tr | 819.6395 | 171.8453 | 4.7696 | 0.0000 | 0.5492 | 0.0969 | 226307.9 |
| age35_2 | -202.4363 | 130.0292 | -1.5569 | 0.1195 | 0.5492 | 0.0969 | 226307.9 |
| parking_yes | 935.7181 | 181.3462 | 5.1598 | 0.0000 | 0.5492 | 0.0969 | 226307.9 |
| parking_yes2 | -685.4251 | 113.0504 | -6.0630 | 0.0000 | 0.5492 | 0.0969 | 226307.9 |
| woman | -749.1139 | 144.9453 | -5.1683 | 0.0000 | 0.5492 | 0.0969 | 226307.9 |
| woman_2 | 211.2827 | 107.4133 | 1.9670 | 0.0492 | 0.5492 | 0.0969 | 226307.9 |
| BMzones | -557.0888 | 317.4161 | -1.7551 | 0.0793 | 0.5492 | 0.0969 | 226307.9 |
| BMzones_2 | 292.4147 | 146.4098 | 1.9972 | 0.0458 | 0.5492 | 0.0969 | 226307.9 |
| power100 | -230.0755 | 140.9145 | -1.6327 | 0.1026 | 0.5492 | 0.0969 | 226307.9 |

Table 31: TWFE for Total variable with reduced covariates without age35

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|--------------|-----------|------------|---------|----------|--------|--------|----------|
| Tr | 815.3623 | 171.8530 | 4.7445 | 0.0000 | 0.5491 | 0.0967 | 226305.8 |
| lic_age15 | -330.2755 | 155.1279 | -2.1291 | 0.0333 | 0.5491 | 0.0967 | 226305.8 |
| parking_yes | 931.3211 | 181.3547 | 5.1354 | 0.0000 | 0.5491 | 0.0967 | 226305.8 |
| parking_yes2 | -696.8813 | 112.7775 | -6.1793 | 0.0000 | 0.5491 | 0.0967 | 226305.8 |
| woman | -742.8234 | 144.9557 | -5.1245 | 0.0000 | 0.5491 | 0.0967 | 226305.8 |
| woman_2 | 208.5263 | 107.3896 | 1.9418 | 0.0522 | 0.5491 | 0.0967 | 226305.8 |
| BMzones | -530.9050 | 317.5070 | -1.6721 | 0.0945 | 0.5491 | 0.0967 | 226305.8 |
| BMzones_2 | 281.5846 | 146.3098 | 1.9246 | 0.0543 | 0.5491 | 0.0967 | 226305.8 |
| power100 | -207.3325 | 141.4845 | -1.4654 | 0.1428 | 0.5491 | 0.0967 | 226305.8 |

3.2.1.2 Variable Night amb l'efecte de covariables El model complet amb totes les covariables i interaccions possibles, així com la seva reducció dona lloc als següents resultats. En aquest cas s'ha obtingut un valor molt proper a 0 i no significatiu per l' ATT_2 , indicant doncs que no implica cap diferència en el percentatge d'hores conduïdes durant la nit el fet d'haver patit un accident durant l'any anterior o no. Pel que fa als coeficients sense interaccions comentar que la variable **age35** presenta un efecte positiu. Això indica que els usuaris amb menys de 35 anys condueixen un percentatge de quilòmetres major durant la nit que els majors de 35 anys, tot i que aquest percentatge es redueix durant el segon període. També es pot observar que tenir pàrquing i conduir per Madrid o Barcelona augmenta aquest percentatge de conducció nocturna, mentre ser dona o tenir un cotxe amb baixa potència el disminueix.

D'altra banda, es nota que s'ha inclòs varies interaccions entre la covariables i la variable **Tr**. Per aquestes s'ha decidit calcular l' ATT_2 associat dels quals només el de la variable **power100** ha resultat ser significatiu amb una significació del 3,5%, mentre **lic_age15** i **parking_yes** presenten un pvalor superior al 5% de significància. A més a més, el valor resultant ha estat rarament positiu. Això és degut, tal com veiem a la gràfica, que aquells que augmenten el percentatge de quilòmetres conduïts durant la nit del primer al segon període són els usuaris que condueixen cotxes grans que no han tingut cap accident i els usuaris amb cotxes de baixa potència que han tingut algun accident. Mentre els altres dos tipus d'usuaris disminueixen el percentatge. És a dir, la diferència entre $t=1$ i $t=2$ disminueix pels usuaris que condueixen cotxes grans, independentment de si han tingut un accident o no, mentre aquesta augmenta pels usuaris que condueixen cotxes de baixa potència.

Table 32: TWFE for Night variable with covariates

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|--------------|----------|------------|----------|----------|-------|--------|----------|
| Tr | 0.6255 | 0.7170 | 0.8723 | 0.3831 | 0.562 | 0.1212 | 73055.32 |
| age35 | 1.5221 | 0.3506 | 4.3417 | 0.0000 | 0.562 | 0.1212 | 73055.32 |
| age35_2 | -1.1314 | 0.2908 | -3.8903 | 0.0001 | 0.562 | 0.1212 | 73055.32 |
| lic_age15 | 0.1953 | 0.3246 | 0.6018 | 0.5473 | 0.562 | 0.1212 | 73055.32 |
| lic_age15_2 | -0.0116 | 0.2796 | -0.0415 | 0.9669 | 0.562 | 0.1212 | 73055.32 |
| parking_yes | 1.0956 | 0.3274 | 3.3467 | 0.0008 | 0.562 | 0.1212 | 73055.32 |
| parking_yes2 | -0.2792 | 0.2153 | -1.2972 | 0.1946 | 0.562 | 0.1212 | 73055.32 |
| woman | -2.6873 | 0.2627 | -10.2307 | 0.0000 | 0.562 | 0.1212 | 73055.32 |
| woman_2 | 0.8012 | 0.2086 | 3.8404 | 0.0001 | 0.562 | 0.1212 | 73055.32 |
| BMzones | -0.6565 | 0.5745 | -1.1427 | 0.2532 | 0.562 | 0.1212 | 73055.32 |
| BMzones_2 | 0.8627 | 0.2772 | 3.1124 | 0.0019 | 0.562 | 0.1212 | 73055.32 |
| power100 | -0.2020 | 0.2759 | -0.7320 | 0.4642 | 0.562 | 0.1212 | 73055.32 |
| power100_2 | -0.7506 | 0.2125 | -3.5321 | 0.0004 | 0.562 | 0.1212 | 73055.32 |
| age35_2_Tr | -1.1078 | 0.8927 | -1.2410 | 0.2146 | 0.562 | 0.1212 | 73055.32 |

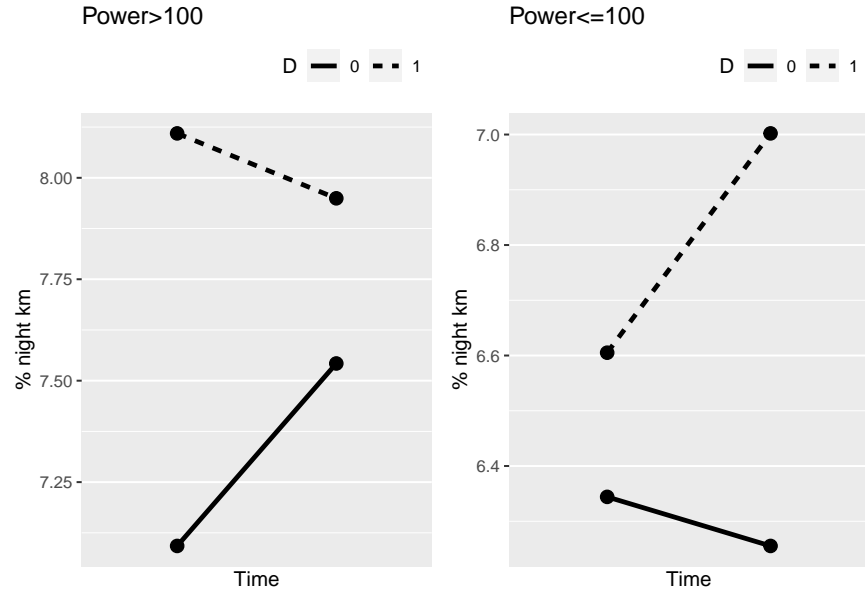
| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|-----------------|----------|------------|---------|----------|-------|--------|----------|
| lic_age15_2_Tr | -1.5725 | 0.8523 | -1.8450 | 0.0651 | 0.562 | 0.1212 | 73055.32 |
| parking_yes2_Tr | -1.1269 | 0.6717 | -1.6776 | 0.0935 | 0.562 | 0.1212 | 73055.32 |
| woman_2_Tr | 0.1639 | 0.6420 | 0.2553 | 0.7985 | 0.562 | 0.1212 | 73055.32 |
| BMzones_2_Tr | -0.5505 | 0.8701 | -0.6327 | 0.5269 | 0.562 | 0.1212 | 73055.32 |
| power100_2_Tr | 1.1836 | 0.6480 | 1.8266 | 0.0678 | 0.562 | 0.1212 | 73055.32 |

Table 33: TWFE for Night variable with reduced covariates

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|-----------------|----------|------------|----------|----------|--------|--------|----------|
| Tr | 0.6711 | 0.6572 | 1.0212 | 0.3072 | 0.5619 | 0.1222 | 73045.43 |
| age35 | 1.5971 | 0.3290 | 4.8546 | 0.0000 | 0.5619 | 0.1222 | 73045.43 |
| age35_2 | -1.2378 | 0.2567 | -4.8219 | 0.0000 | 0.5619 | 0.1222 | 73045.43 |
| parking_yes | 1.0100 | 0.3135 | 3.2213 | 0.0013 | 0.5619 | 0.1222 | 73045.43 |
| woman | -2.7319 | 0.2543 | -10.7416 | 0.0000 | 0.5619 | 0.1222 | 73045.43 |
| woman_2 | 0.8216 | 0.1990 | 4.1282 | 0.0000 | 0.5619 | 0.1222 | 73045.43 |
| BMzones_2 | 0.7639 | 0.2575 | 2.9661 | 0.0030 | 0.5619 | 0.1222 | 73045.43 |
| power100_2 | -0.7750 | 0.1957 | -3.9597 | 0.0001 | 0.5619 | 0.1222 | 73045.43 |
| lic_age15_2_Tr | -1.8673 | 0.7655 | -2.4393 | 0.0147 | 0.5619 | 0.1222 | 73045.43 |
| parking_yes2_Tr | -1.4154 | 0.6409 | -2.2085 | 0.0272 | 0.5619 | 0.1222 | 73045.43 |
| power100_2_Tr | 1.1956 | 0.6266 | 1.9079 | 0.0564 | 0.5619 | 0.1222 | 73045.43 |

Table 34: Some effects

| | lic_age15 | parking_yes | power100 |
|---------|-----------|-------------|----------|
| coef | -1.1962 | 0.2657 | 1.0917 |
| Z | -1.2608 | 0.4417 | 1.8010 |
| p_value | 0.1037 | 0.3293 | 0.0358 |



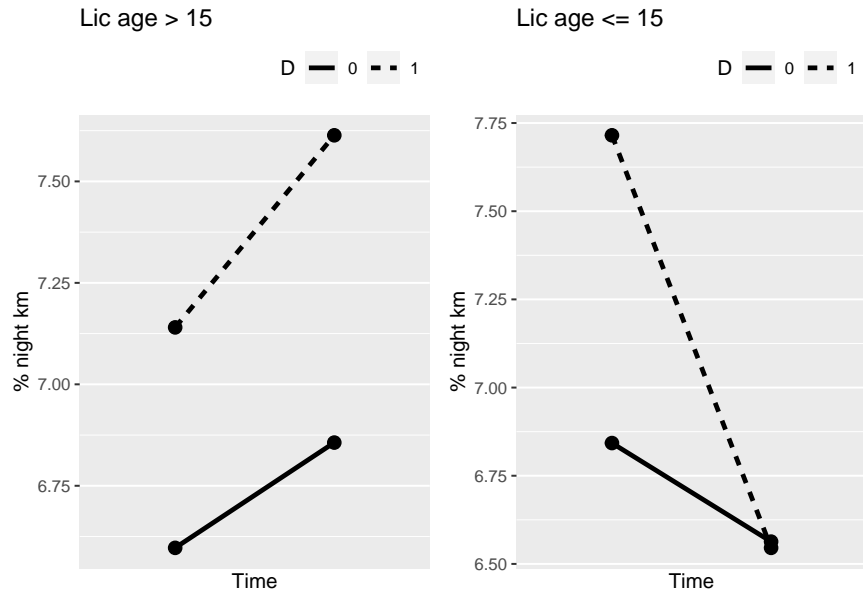
A més a més, a continuació es mirarà què passaria si traièssim de l'últim model realitzat la variable **parking_yes** en ser aquesta una variable que sabem que no és gaire rellevant per a la nostra variable dependent i que podria afectar a que altres variables importants no surtissin significatives. Com es pot notar, en aquest cas **lic_age15** ha passat a ser significativa mentre **power100** ho ha deixat de ser. A l'analitzar la gràfica notem que tots els individus amb menys de 15 anys de llicència han reduït el percentatge de quilòmetres conduïts durant la nit, mentre el comportament ha estat contrari pels usuaris amb més de 15 anys de carnet. A més a més, mentre la diferència en percentatge ha reduït del pre al post tractament pels usuaris amb menys de 15 anys de llicència, aquesta ha augmentat per la resta d'assegurats. No obstant això, l'AIC del model ha augmentat i l' R^2_{Adj} ha disminuït respecte a l'últim model creat, indicant que el prèvi és millor, com esperat perquè sinó la funció stepAIC ja hauria eliminat la variable **parking_yes** ella mateixa.

Table 35: TWFE for Night variable with reduced covariates without parking_yes

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|----------------|----------|------------|----------|----------|--------|-------|----------|
| Tr | -0.3008 | 0.4865 | -0.6182 | 0.5364 | 0.5617 | 0.122 | 73055.49 |
| age35 | 1.5955 | 0.3291 | 4.8480 | 0.0000 | 0.5617 | 0.122 | 73055.49 |
| age35_2 | -1.2442 | 0.2568 | -4.8449 | 0.0000 | 0.5617 | 0.122 | 73055.49 |
| woman | -2.7344 | 0.2544 | -10.7466 | 0.0000 | 0.5617 | 0.122 | 73055.49 |
| woman_2 | 0.8200 | 0.1991 | 4.1184 | 0.0000 | 0.5617 | 0.122 | 73055.49 |
| BMzones_2 | 0.7732 | 0.2577 | 3.0007 | 0.0027 | 0.5617 | 0.122 | 73055.49 |
| power100_2 | -0.7944 | 0.1957 | -4.0588 | 0.0000 | 0.5617 | 0.122 | 73055.49 |
| lic_age15_2_Tr | -2.0905 | 0.7616 | -2.7448 | 0.0061 | 0.5617 | 0.122 | 73055.49 |
| power100_2_Tr | 1.2828 | 0.6260 | 2.0492 | 0.0405 | 0.5617 | 0.122 | 73055.49 |

Table 36: Some effects

| | lic_age15 | power100 |
|---------|-----------|----------|
| coef | -2.3913 | 0.1877 |
| Z | -3.0147 | 0.4185 |
| p_value | 0.0013 | 0.3378 |



A més a més, ens podríem plantejar el fet que l'edat i els anys de llicència són variables que aporten informació massa similar en quan a l'actitud o característiques de la persona i que per tant només 1 d'elles s'hauria de quedar en el model. Ja que de fet, en l'últim model provat han sortit com a significatives les variables que parlen sobre l'edat per elles mateixes, mentre eren els anys de llicència els que es trobaven relacionats amb Tr. En aquest cas, els resultats dels models amb i sense la variable pàrquing i quedant-nos amb la variable age35 o lic_age15 són els que es poden trobar a continuació.

Les observacions més rellevants extretes d'aquests models són que introduir la variable **parking_yes** en el model dona lloc a que l' ATT_2 de la variable **power100** sigui significatiu (o quasi significatiu pel cas en que va acompanyat de la variable anys de llicència) i positiu, mentre l'efecte del tractament tant de l'edat com dels anys de llicència són no significatius. Altrament, al treure la variable **parking_yes**, és la variable **power100** la que passa a tenir un efecte del tractament no significatiu, mentre **age35** i **lic_age15** donen lloc a un ATT_2 significatiu i negatiu.

També es conclou que si només es pogués disposar d'una de les dues variables per a predir el percentatge de quilòmetres nocturns que recorren els assegurats, es preferiria disposar de la variable edat en resultar els seus models en un AIC menor.

Table 37: TWFE for Night variable with reduced covariates without lic_age15

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|-----------------|----------|------------|----------|----------|--------|--------|----------|
| Tr | 0.6760 | 0.6598 | 1.0246 | 0.3055 | 0.5619 | 0.1221 | 73047.39 |
| age35 | 1.5804 | 0.3290 | 4.8035 | 0.0000 | 0.5619 | 0.1221 | 73047.39 |
| age35_2 | -1.1467 | 0.2659 | -4.3126 | 0.0000 | 0.5619 | 0.1221 | 73047.39 |
| parking_yes | 1.0296 | 0.3134 | 3.2850 | 0.0010 | 0.5619 | 0.1221 | 73047.39 |
| woman | -2.7277 | 0.2543 | -10.7249 | 0.0000 | 0.5619 | 0.1221 | 73047.39 |
| woman_2 | 0.8200 | 0.1990 | 4.1201 | 0.0000 | 0.5619 | 0.1221 | 73047.39 |
| BMzones_2 | 0.7570 | 0.2575 | 2.9393 | 0.0033 | 0.5619 | 0.1221 | 73047.39 |
| power100_2 | -0.7787 | 0.1958 | -3.9775 | 0.0001 | 0.5619 | 0.1221 | 73047.39 |
| age35_2_Tr | -1.6706 | 0.8358 | -1.9987 | 0.0457 | 0.5619 | 0.1221 | 73047.39 |
| parking_yes2_Tr | -1.4872 | 0.6391 | -2.3270 | 0.0200 | 0.5619 | 0.1221 | 73047.39 |
| power100_2_Tr | 1.1408 | 0.6258 | 1.8230 | 0.0683 | 0.5619 | 0.1221 | 73047.39 |

Table 38: Some effects

| | age35 | parking_yes | power100 |
|---------|---------|-------------|----------|
| coef | -0.5608 | 0.2185 | 1.0382 |
| Z | -0.5649 | 0.3624 | 1.7136 |
| p_value | 0.2861 | 0.3585 | 0.0433 |

Table 39: TWFE for Night variable with reduced covariates without lic_age15 and parking_yes

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|-----------|----------|------------|----------|----------|--------|--------|----------|
| Tr | -0.3622 | 0.4878 | -0.7425 | 0.4578 | 0.5616 | 0.1218 | 73058.33 |
| age35 | 1.5773 | 0.3291 | 4.7922 | 0.0000 | 0.5616 | 0.1218 | 73058.33 |
| age35_2 | -1.1485 | 0.2660 | -4.3174 | 0.0000 | 0.5616 | 0.1218 | 73058.33 |
| woman | -2.7297 | 0.2545 | -10.7271 | 0.0000 | 0.5616 | 0.1218 | 73058.33 |
| woman_2 | 0.8179 | 0.1991 | 4.1077 | 0.0000 | 0.5616 | 0.1218 | 73058.33 |
| BMzones_2 | 0.7656 | 0.2577 | 2.9712 | 0.0030 | 0.5616 | 0.1218 | 73058.33 |

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|---------------|----------|------------|---------|----------|--------|--------|----------|
| power100_2 | -0.7985 | 0.1958 | -4.0790 | 0.0000 | 0.5616 | 0.1218 | 73058.33 |
| age35_2_Tr | -1.8060 | 0.8342 | -2.1649 | 0.0304 | 0.5616 | 0.1218 | 73058.33 |
| power100_2_Tr | 1.2242 | 0.6253 | 1.9576 | 0.0503 | 0.5616 | 0.1218 | 73058.33 |

Table 40: Some effects

| | age35 | power100 |
|---------|---------|----------|
| coef | -1.7394 | 0.0635 |
| Z | -2.0280 | 0.1432 |
| p_value | 0.0213 | 0.4431 |

Table 41: TWFE for Night variable with reduced covariates without age35

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|-----------------|----------|------------|----------|----------|--------|--------|----------|
| Tr | 0.4715 | 0.6751 | 0.6985 | 0.4849 | 0.5613 | 0.1208 | 73073.29 |
| lic_age15 | 0.6295 | 0.3043 | 2.0684 | 0.0386 | 0.5613 | 0.1208 | 73073.29 |
| lic_age15_2 | -0.4081 | 0.2562 | -1.5927 | 0.1113 | 0.5613 | 0.1208 | 73073.29 |
| parking_yes | 1.1490 | 0.3259 | 3.5254 | 0.0004 | 0.5613 | 0.1208 | 73073.29 |
| parking_yes2 | -0.3108 | 0.2152 | -1.4441 | 0.1487 | 0.5613 | 0.1208 | 73073.29 |
| woman | -2.7647 | 0.2549 | -10.8483 | 0.0000 | 0.5613 | 0.1208 | 73073.29 |
| woman_2 | 0.8541 | 0.1996 | 4.2780 | 0.0000 | 0.5613 | 0.1208 | 73073.29 |
| BMzones_2 | 0.7309 | 0.2589 | 2.8225 | 0.0048 | 0.5613 | 0.1208 | 73073.29 |
| power100_2 | -0.8190 | 0.1979 | -4.1376 | 0.0000 | 0.5613 | 0.1208 | 73073.29 |
| lic_age15_2_Tr | -1.8734 | 0.7981 | -2.3473 | 0.0189 | 0.5613 | 0.1208 | 73073.29 |
| parking_yes2_Tr | -1.1731 | 0.6714 | -1.7472 | 0.0806 | 0.5613 | 0.1208 | 73073.29 |
| power100_2_Tr | 1.2433 | 0.6279 | 1.9799 | 0.0477 | 0.5613 | 0.1208 | 73073.29 |

Table 42: Some effects

| | lic_age15 | parking_yes | power100 |
|---------|-----------|-------------|----------|
| coef | -1.1805 | 0.1367 | 0.8958 |
| Z | -1.2056 | 0.2247 | 1.4307 |
| p_value | 0.1140 | 0.4111 | 0.0763 |

Table 43: TWFE for Night variable with reduced covariates without age35 and parking_yes

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|-------------|----------|------------|----------|----------|--------|--------|---------|
| Tr | -0.3080 | 0.4882 | -0.6308 | 0.5282 | 0.5612 | 0.1208 | 73083.5 |
| lic_age15 | 0.6372 | 0.3043 | 2.0940 | 0.0363 | 0.5612 | 0.1208 | 73083.5 |
| lic_age15_2 | -0.4431 | 0.2553 | -1.7357 | 0.0826 | 0.5612 | 0.1208 | 73083.5 |
| woman | -2.7571 | 0.2550 | -10.8110 | 0.0000 | 0.5612 | 0.1208 | 73083.5 |
| woman_2 | 0.8375 | 0.1997 | 4.1945 | 0.0000 | 0.5612 | 0.1208 | 73083.5 |
| BMzones | -0.8904 | 0.5728 | -1.5545 | 0.1201 | 0.5612 | 0.1208 | 73083.5 |

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|----------------|----------|------------|---------|----------|--------|--------|---------|
| BMzones_2 | 0.8533 | 0.2644 | 3.2273 | 0.0013 | 0.5612 | 0.1208 | 73083.5 |
| power100_2 | -0.8031 | 0.1966 | -4.0857 | 0.0000 | 0.5612 | 0.1208 | 73083.5 |
| lic_age15_2_Tr | -2.0786 | 0.7942 | -2.6171 | 0.0089 | 0.5612 | 0.1208 | 73083.5 |
| power100_2_Tr | 1.3084 | 0.6270 | 2.0868 | 0.0369 | 0.5612 | 0.1208 | 73083.5 |

Table 44: Some effects

| | lic_age15 | power100 |
|---------|-----------|----------|
| coef | -2.1925 | 0.1973 |
| Z | -2.6823 | 0.4382 |
| p_value | 0.0037 | 0.3306 |

3.2.1.3 Variable Speed amb l'efecte de covariables Per a la variable dependent **Speed** s'observa un efecte del tractament negatiu però no significatiu. El signe es troba en consonància amb allò esperat, ja que aquells usuaris que han tingut un accident, s'espera que conduixin més lentament posteriorment a l'accident. Tanmateix, com comentat el pvalor d'aquest coeficient no és estadísticament significatiu, raó per la qual no es pot concloure res mirant aquests resultats.

Pel que fa als usuaris que tenen menys de 35 anys, aquests tenen un coeficient positiu i significatiu, indicant que conduïxen a major velocitat que els usuaris de major edat. De fet, aquest coeficient passa a ser positiu gràcies a la variable **age35_2**. Això indica que a $t=1$ els usuaris joves sí que conduïxen més lent, però a $t=2$ guanyen més confiança. A més a més, el coeficient dels usuaris que tenen el carnet des de fa menys de 15 anys és negatiu. Es pot concloure doncs que els joves són irresponsables al conduir ràpidament, però que sí que es mantenen amb major percentatge dins dels límits de velocitat, si fa poc que s'han tret el carnet.

Tal com esperat, les dones trenquen menys els límits de velocitat, així com aquells usuaris que tenen cotxes petits. Conduir per Madrid o Barcelona també mostra un signe negatiu, segurament degut a la major presència de radars en aquestes ciutats. Si s'analitzen les interaccions que han sortit significatives es pot notar que tot i que aquestes milloren el model, totes elles mostren un ATT_2 no significatiu. Això indica que tenir un accident no impacta de cap manera significativa sobre la velocitat de conducció dels usuaris, independentment de les seves característiques. Tot i així, sí que cal recalcar el fet que han tornat a sortir les mateixes interaccions que sortien per a la variable dependent **Night**, (prenent **age35** i **lic_age15** com a variable similars, indicadors del temps). Això indica clarament quines són les covariables més rellevants.

Table 45: TWFE for Speed variable with covariates

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|--------------|----------|------------|----------|----------|--------|--------|----------|
| Tr | -0.1377 | 0.6206 | -0.2218 | 0.8245 | 0.7281 | 0.4544 | 69550.86 |
| age35 | -0.7480 | 0.3034 | -2.4654 | 0.0137 | 0.7281 | 0.4544 | 69550.86 |
| age35_2 | 1.2145 | 0.2517 | 4.8252 | 0.0000 | 0.7281 | 0.4544 | 69550.86 |
| lic_age15 | -2.4838 | 0.2809 | -8.8416 | 0.0000 | 0.7281 | 0.4544 | 69550.86 |
| lic_age15_2 | 1.0852 | 0.2420 | 4.4848 | 0.0000 | 0.7281 | 0.4544 | 69550.86 |
| parking_yes | 0.5479 | 0.2833 | 1.9337 | 0.0532 | 0.7281 | 0.4544 | 69550.86 |
| parking_yes2 | 0.0885 | 0.1863 | 0.4750 | 0.6348 | 0.7281 | 0.4544 | 69550.86 |
| woman | -0.8794 | 0.2273 | -3.8683 | 0.0001 | 0.7281 | 0.4544 | 69550.86 |
| woman_2 | 0.0505 | 0.1806 | 0.2798 | 0.7796 | 0.7281 | 0.4544 | 69550.86 |
| BMzones | -2.2175 | 0.4972 | -4.4598 | 0.0000 | 0.7281 | 0.4544 | 69550.86 |
| BMzones_2 | 0.7890 | 0.2399 | 3.2888 | 0.0010 | 0.7281 | 0.4544 | 69550.86 |
| power100 | -2.8949 | 0.2388 | -12.1232 | 0.0000 | 0.7281 | 0.4544 | 69550.86 |
| power100_2 | 0.8483 | 0.1839 | 4.6121 | 0.0000 | 0.7281 | 0.4544 | 69550.86 |

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|-----------------|----------|------------|---------|----------|--------|--------|----------|
| age35_2_Tr | -1.7088 | 0.7726 | -2.2117 | 0.0270 | 0.7281 | 0.4544 | 69550.86 |
| lic_age15_2_Tr | 1.0418 | 0.7377 | 1.4124 | 0.1579 | 0.7281 | 0.4544 | 69550.86 |
| parking_yes2_Tr | -1.1890 | 0.5813 | -2.0453 | 0.0408 | 0.7281 | 0.4544 | 69550.86 |
| woman_2_Tr | 0.5833 | 0.5556 | 1.0498 | 0.2939 | 0.7281 | 0.4544 | 69550.86 |
| BMzones_2_Tr | -0.8570 | 0.7530 | -1.1381 | 0.2551 | 0.7281 | 0.4544 | 69550.86 |
| power100_2_Tr | 1.1748 | 0.5608 | 2.0948 | 0.0362 | 0.7281 | 0.4544 | 69550.86 |

Table 46: TWFE for Speed variable with reduced covariates

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|-----------------|----------|------------|----------|----------|--------|--------|----------|
| Tr | -0.1068 | 0.5710 | -0.1871 | 0.8516 | 0.7279 | 0.4545 | 69545.53 |
| age35 | -0.7367 | 0.3033 | -2.4291 | 0.0151 | 0.7279 | 0.4545 | 69545.53 |
| age35_2 | 1.1748 | 0.2496 | 4.7068 | 0.0000 | 0.7279 | 0.4545 | 69545.53 |
| lic_age15 | -2.4838 | 0.2806 | -8.8507 | 0.0000 | 0.7279 | 0.4545 | 69545.53 |
| lic_age15_2 | 1.1915 | 0.2316 | 5.1455 | 0.0000 | 0.7279 | 0.4545 | 69545.53 |
| parking_yes | 0.5672 | 0.2728 | 2.0795 | 0.0376 | 0.7279 | 0.4545 | 69545.53 |
| woman | -0.8306 | 0.2102 | -3.9515 | 0.0001 | 0.7279 | 0.4545 | 69545.53 |
| BMzones | -2.2581 | 0.4960 | -4.5523 | 0.0000 | 0.7279 | 0.4545 | 69545.53 |
| BMzones_2 | 0.7104 | 0.2281 | 3.1148 | 0.0018 | 0.7279 | 0.4545 | 69545.53 |
| power100 | -2.9048 | 0.2378 | -12.2176 | 0.0000 | 0.7279 | 0.4545 | 69545.53 |
| power100_2 | 0.8439 | 0.1771 | 4.7658 | 0.0000 | 0.7279 | 0.4545 | 69545.53 |
| age35_2_Tr | -1.3522 | 0.7237 | -1.8684 | 0.0617 | 0.7279 | 0.4545 | 69545.53 |
| parking_yes2_Tr | -1.0175 | 0.5533 | -1.8390 | 0.0659 | 0.7279 | 0.4545 | 69545.53 |
| power100_2_Tr | 1.4009 | 0.5417 | 2.5863 | 0.0097 | 0.7279 | 0.4545 | 69545.53 |

Table 47: Some effects

| | age<=35 | parking_yes | power<=100 |
|---------|---------|-------------|------------|
| coef | -1.0209 | -0.5571 | -0.7669 |
| Z | -1.1794 | -1.0664 | -1.3941 |
| p_value | 0.1191 | 0.1431 | 0.0816 |

Tot seguit, es decideix mirar altre cop què passaria si s'eliminès de l'últim model realitzat la variable **parking_yes** en ser aquesta una variable que no hauria de ser gaire rellevant per a la variable dependent estudiada i que podria afectar a que altres variables importants no surtissin significatives.

Com podem notar, aquest simple canvi de treure la variable **parking_yes** ha provocat que ara tant **age35** com **power100** mostrin ATT_2 significatius, tot i que l'AIC augmenti de forma lleugera. Amb les gràfiques es pot notar clarament que els usuaris que han canviat més el seu comportament després de patir accidents són aquells que condueixen cotxes grans i que tenen menys de 35 anys.

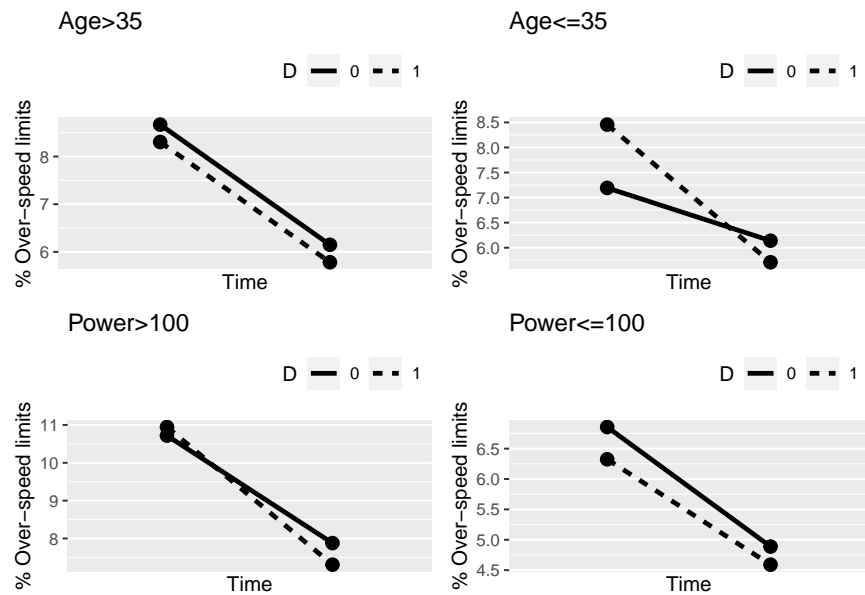
Table 48: TWFE for Speed variable with reduced covariates without parking_yes

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|---------|----------|------------|---------|----------|--------|--------|----------|
| Tr | -0.8155 | 0.4220 | -1.9322 | 0.0534 | 0.7277 | 0.4544 | 69548.61 |
| age35 | -0.7306 | 0.3033 | -2.4084 | 0.0160 | 0.7277 | 0.4544 | 69548.61 |
| age35_2 | 1.1763 | 0.2496 | 4.7120 | 0.0000 | 0.7277 | 0.4544 | 69548.61 |

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|---------------|----------|------------|----------|----------|--------|--------|----------|
| lic_age15 | -2.4930 | 0.2806 | -8.8857 | 0.0000 | 0.7277 | 0.4544 | 69548.61 |
| lic_age15_2 | 1.1827 | 0.2315 | 5.1085 | 0.0000 | 0.7277 | 0.4544 | 69548.61 |
| woman | -0.8228 | 0.2102 | -3.9144 | 0.0001 | 0.7277 | 0.4544 | 69548.61 |
| BMzones | -2.3050 | 0.4956 | -4.6507 | 0.0000 | 0.7277 | 0.4544 | 69548.61 |
| BMzones_2 | 0.7207 | 0.2281 | 3.1595 | 0.0016 | 0.7277 | 0.4544 | 69548.61 |
| power100 | -2.9462 | 0.2369 | -12.4358 | 0.0000 | 0.7277 | 0.4544 | 69548.61 |
| power100_2 | 0.8442 | 0.1771 | 4.7669 | 0.0000 | 0.7277 | 0.4544 | 69548.61 |
| age35_2_Tr | -1.4464 | 0.7220 | -2.0033 | 0.0452 | 0.7277 | 0.4544 | 69548.61 |
| power100_2_Tr | 1.4573 | 0.5411 | 2.6934 | 0.0071 | 0.7277 | 0.4544 | 69548.61 |

Table 49: Some effects

| | age35 | power100 |
|---------|---------|----------|
| coef | -1.8162 | -1.4601 |
| Z | -2.4211 | -3.5010 |
| p_value | 0.0077 | 0.0002 |



A més a més, tal com fet fins ara, es mirarà què passaria si es considerès dins del model només la variable edat o només els anys de llicència, en poder presentar aquestes dues variables multicolinealitat i no deixar entreveure altres resultats.

En aquest cas les conclusions extremes d'aquestes combinacions de models són molt diferents respecte a les que s'han obtingut anteriorment amb la variable **Night**. Concretament, considerar la variable anys de llicència per al percentatge de quilòmetres recorreguts per sobre de la velocitat permesa no ajuda en no sortir aquesta interacció com a significativa en cap cas. D'altra banda, si s'introdueix la variable edat aquesta sempre surt significativa i amb un efecte del tractament negatiu, coincidint el signe en aquest cas sí amb els resultats que s'obtenien per **Night**. Finalment, la variable **power100** ha sortit significativa (o quasi significativa) amb un signe negatiu, a diferència del que s'obtenia per **Night**.

Això ens dur a concloure que l'edat de la persona és molt més rellevant que els anys de llicència d'aquesta pel que fa a predir la velocitat en funció de si l'usuari ha tingut un accident o no. A més a més, sabem

que els usuaris amb menys de 35 anys que pateixen un accident redueixen la velocitat després de tenir-lo, així com aquells que tenen cotxes amb baixa potència. Aquesta última afirmació era contrària pel cas del percentatge de quilòmetres recorreguts nocturnament, en mostrar la variable **power100** un signe positiu de l'efecte del tractament, indicant doncs que les persones amb cotxes de baixa potència augmenten la diferència de quilòmetres recorreguts nocturnament de forma positiva, respecte als usuaris que no n'han tingut cap.

No obstant això, cal tenir en compte que l' R^2_{Adj} és major per al model que elimina la variable edat i manté **lic_age15**. Així doncs, mentre si es vol analitzar què succeix després d'un accident disposar de la variable edat és millor. Si es vol predir amb major precisió el percentatge de quilòmetres per sobre de la velocitat dels assegurats, disposar dels anys de llicència és millor.

Table 50: TWFE for Speed variable with reduced covariates without **lic_age15**

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|-----------------|----------|------------|----------|----------|--------|-------|----------|
| Tr | -0.1099 | 0.5728 | -0.1919 | 0.8478 | 0.7266 | 0.452 | 69622.99 |
| age35 | -1.6354 | 0.2868 | -5.7026 | 0.0000 | 0.7266 | 0.452 | 69622.99 |
| age35_2 | 1.6197 | 0.2309 | 7.0133 | 0.0000 | 0.7266 | 0.452 | 69622.99 |
| parking_yes | 0.6295 | 0.2736 | 2.3013 | 0.0214 | 0.7266 | 0.452 | 69622.99 |
| woman | -0.8531 | 0.2108 | -4.0477 | 0.0001 | 0.7266 | 0.452 | 69622.99 |
| BMzones | -2.4923 | 0.4970 | -5.0148 | 0.0000 | 0.7266 | 0.452 | 69622.99 |
| BMzones_2 | 0.8154 | 0.2283 | 3.5721 | 0.0004 | 0.7266 | 0.452 | 69622.99 |
| power100 | -3.0884 | 0.2377 | -12.9950 | 0.0000 | 0.7266 | 0.452 | 69622.99 |
| power100_2 | 0.9401 | 0.1767 | 5.3194 | 0.0000 | 0.7266 | 0.452 | 69622.99 |
| age35_2_Tr | -1.4631 | 0.7258 | -2.0157 | 0.0439 | 0.7266 | 0.452 | 69622.99 |
| parking_yes2_Tr | -0.9337 | 0.5549 | -1.6826 | 0.0925 | 0.7266 | 0.452 | 69622.99 |
| power100_2_Tr | 1.3864 | 0.5434 | 2.5514 | 0.0107 | 0.7266 | 0.452 | 69622.99 |

Table 51: Some effects

| | age35 | parking_yes | power100 |
|---------|---------|-------------|----------|
| coef | -1.5887 | -0.4140 | -0.8717 |
| Z | -1.8430 | -0.7903 | -1.5799 |
| p_value | 0.0327 | 0.2147 | 0.0571 |

Table 52: TWFE for Speed variable with reduced covariates without **lic_age15** and **parking_yes**

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|---------------|----------|------------|----------|----------|--------|--------|----------|
| Tr | -0.7605 | 0.4234 | -1.7961 | 0.0725 | 0.7264 | 0.4519 | 69626.48 |
| age35 | -1.6334 | 0.2868 | -5.6946 | 0.0000 | 0.7264 | 0.4519 | 69626.48 |
| age35_2 | 1.6173 | 0.2310 | 7.0014 | 0.0000 | 0.7264 | 0.4519 | 69626.48 |
| woman | -0.8449 | 0.2108 | -4.0085 | 0.0001 | 0.7264 | 0.4519 | 69626.48 |
| BMzones | -2.5452 | 0.4965 | -5.1258 | 0.0000 | 0.7264 | 0.4519 | 69626.48 |
| BMzones_2 | 0.8261 | 0.2283 | 3.6188 | 0.0003 | 0.7264 | 0.4519 | 69626.48 |
| power100 | -3.1352 | 0.2368 | -13.2416 | 0.0000 | 0.7264 | 0.4519 | 69626.48 |
| power100_2 | 0.9400 | 0.1768 | 5.3172 | 0.0000 | 0.7264 | 0.4519 | 69626.48 |
| age35_2_Tr | -1.5492 | 0.7242 | -2.1392 | 0.0324 | 0.7264 | 0.4519 | 69626.48 |
| power100_2_Tr | 1.4399 | 0.5428 | 2.6526 | 0.0080 | 0.7264 | 0.4519 | 69626.48 |

Table 53: Some effects

| | age35 | power100 |
|---------|---------|----------|
| coef | -2.3258 | -1.5158 |
| Z | -3.1236 | -3.6240 |
| p_value | 0.0009 | 0.0001 |

Table 54: TWFE for Speed variable with reduced covariates without age35

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|-----------------|----------|------------|----------|----------|--------|--------|---------|
| Tr | -0.2780 | 0.5653 | -0.4918 | 0.6229 | 0.7272 | 0.4533 | 69562.7 |
| lic_age15 | -2.7122 | 0.2647 | -10.2464 | 0.0000 | 0.7272 | 0.4533 | 69562.7 |
| lic_age15_2 | 1.5956 | 0.2130 | 7.4918 | 0.0000 | 0.7272 | 0.4533 | 69562.7 |
| parking_yes | 0.5647 | 0.2730 | 2.0688 | 0.0386 | 0.7272 | 0.4533 | 69562.7 |
| woman | -0.8236 | 0.2103 | -3.9168 | 0.0001 | 0.7272 | 0.4533 | 69562.7 |
| BMzones | -2.2234 | 0.4959 | -4.4841 | 0.0000 | 0.7272 | 0.4533 | 69562.7 |
| BMzones_2 | 0.7091 | 0.2283 | 3.1065 | 0.0019 | 0.7272 | 0.4533 | 69562.7 |
| power100 | -2.9168 | 0.2376 | -12.2753 | 0.0000 | 0.7272 | 0.4533 | 69562.7 |
| power100_2 | 0.8591 | 0.1772 | 4.8486 | 0.0000 | 0.7272 | 0.4533 | 69562.7 |
| parking_yes2_Tr | -1.0594 | 0.5523 | -1.9180 | 0.0551 | 0.7272 | 0.4533 | 69562.7 |
| power100_2_Tr | 1.3564 | 0.5418 | 2.5034 | 0.0123 | 0.7272 | 0.4533 | 69562.7 |

Table 55: Some effects

| | parking_yes | power100 |
|---------|-------------|----------|
| coef | -0.7726 | -0.9793 |
| Z | -1.5186 | -1.8062 |
| p_value | 0.0644 | 0.0354 |

Table 56: TWFE for Speed variable with reduced covariates without age35 and parking_yes

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|---------------|----------|------------|----------|----------|-------|--------|-------|
| Tr | -1.0309 | 0.4076 | -2.5290 | 0.0115 | 0.727 | 0.4532 | 69566 |
| lic_age15 | -2.7194 | 0.2646 | -10.2765 | 0.0000 | 0.727 | 0.4532 | 69566 |
| lic_age15_2 | 1.5843 | 0.2129 | 7.4416 | 0.0000 | 0.727 | 0.4532 | 69566 |
| woman | -0.8160 | 0.2103 | -3.8805 | 0.0001 | 0.727 | 0.4532 | 69566 |
| BMzones | -2.2712 | 0.4954 | -4.5842 | 0.0000 | 0.727 | 0.4532 | 69566 |
| BMzones_2 | 0.7192 | 0.2283 | 3.1507 | 0.0016 | 0.727 | 0.4532 | 69566 |
| power100 | -2.9572 | 0.2368 | -12.4896 | 0.0000 | 0.727 | 0.4532 | 69566 |
| power100_2 | 0.8597 | 0.1772 | 4.8507 | 0.0000 | 0.727 | 0.4532 | 69566 |
| power100_2_Tr | 1.4129 | 0.5413 | 2.6101 | 0.0091 | 0.727 | 0.4532 | 69566 |

Table 57: Some effects

| | power100 |
|---------|----------|
| coef | -1.7156 |
| Z | -4.2943 |
| p_value | 0.0000 |

3.2.1.4 Variable Urban amb l'efecte de covariables Tornem a trobar que algunes interaccions són significatives. D'altra banda, l'estimador de l'efecte del tractament és negatiu però no significatiu. Pel que fa a les variables sense interaccions podem veure que el fet de tenir menys de 35 anys implica conduir més per vies urbanes, així com tenir un cotxe petit, tot i que per tots dos factors aquest percentatge disminueix durant el període de post-tractament. Ser dona i conduir per Madrid o Barcelona implica conduir un percentatge menor per vies urbanes i finalment tenir la llicència fa menys de 15 anys, conduir-ne més per aquest tipus de via.

No obstant això, si considerem les interaccions incloses en el model entre **Tr** i les covariables, podem observar com l'estimador de l' ATT_2 associat amb **lic_age15** i **power100** són significatius al 1% i 0.3%, respectivament. A més a més, l'estimador associat amb les llicències joves és negatiu, volent dir que aquells usuaris que s'han tret el carnet fa poc, després de patir un accident preferixen conduir menys per vies urbanes, mentre els usuaris amb cotxes petits es troben més segurs conduint en aquestes condicions, en ser el seu estimador de signe positiu. Finalment pel que fa a les interaccions, comentar que tot i que la interacció entre la variable **woman** i **Tr** millora el model en quant a informació captada en relació a la quantitat d'informació considerada, al treure els estadístics particulars, aquesta és no significativa.

Amb les gràfiques d'aquelles interaccions que han sortit significatives es pot veure que les persones amb llicència des de fa més de 15 anys condueixen sistemàticament menys quilòmetres per vies urbanes. Així mateix, aquells que no han tingut cap accident en condueixen cada cop menys. Pel que fa als usuaris amb menys de 15 anys de carnet el comportament és a l'inversa, en ser els que han tingut un accident els que passen a fer menys quilòmetres per vies urbanes, tal com comentat anteriorment quan s'anализava el signe de l'efecte del tractament.

Pel que fa a la potència dels cotxes, s'observa que els usuaris que han tingut un accident condueixen sempre més per vies urbanes, independentment de si el seu cotxe és gran o petit, indicant que segurament molts d'aquests accidents hauran estat en aquestes vies. Així mateix en aquest cas són els usuaris amb cotxes petits qui condueixen més per aquest tipus de vies. Finalment es nota que tenir un accident dona lloc a fer menys quilòmetres per vies urbanes en el cas de cotxes grans, mentre l'efecte no és tant significatiu sobre els petits.

Table 58: TWFE for Urban variable with covariates

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|--------------|----------|------------|---------|----------|--------|--------|----------|
| Tr | -1.7464 | 1.1674 | -1.4960 | 0.1347 | 0.7427 | 0.4837 | 84878.47 |
| age35 | 2.4984 | 0.5708 | 4.3774 | 0.0000 | 0.7427 | 0.4837 | 84878.47 |
| age35_2 | -1.1842 | 0.4735 | -2.5010 | 0.0124 | 0.7427 | 0.4837 | 84878.47 |
| lic_age15 | 1.8767 | 0.5285 | 3.5512 | 0.0004 | 0.7427 | 0.4837 | 84878.47 |
| lic_age15_2 | -0.0370 | 0.4552 | -0.0813 | 0.9352 | 0.7427 | 0.4837 | 84878.47 |
| parking_yes | -0.5348 | 0.5330 | -1.0035 | 0.3157 | 0.7427 | 0.4837 | 84878.47 |
| parking_yes2 | 0.3235 | 0.3505 | 0.9231 | 0.3560 | 0.7427 | 0.4837 | 84878.47 |
| woman | -0.1805 | 0.4277 | -0.4221 | 0.6730 | 0.7427 | 0.4837 | 84878.47 |
| woman_2 | -0.3900 | 0.3397 | -1.1481 | 0.2510 | 0.7427 | 0.4837 | 84878.47 |
| BMzones | -2.3852 | 0.9354 | -2.5500 | 0.0108 | 0.7427 | 0.4837 | 84878.47 |
| BMzones_2 | -0.6111 | 0.4513 | -1.3540 | 0.1758 | 0.7427 | 0.4837 | 84878.47 |
| power100 | 2.1348 | 0.4492 | 4.7524 | 0.0000 | 0.7427 | 0.4837 | 84878.47 |
| power100_2 | -0.5670 | 0.3460 | -1.6386 | 0.1013 | 0.7427 | 0.4837 | 84878.47 |

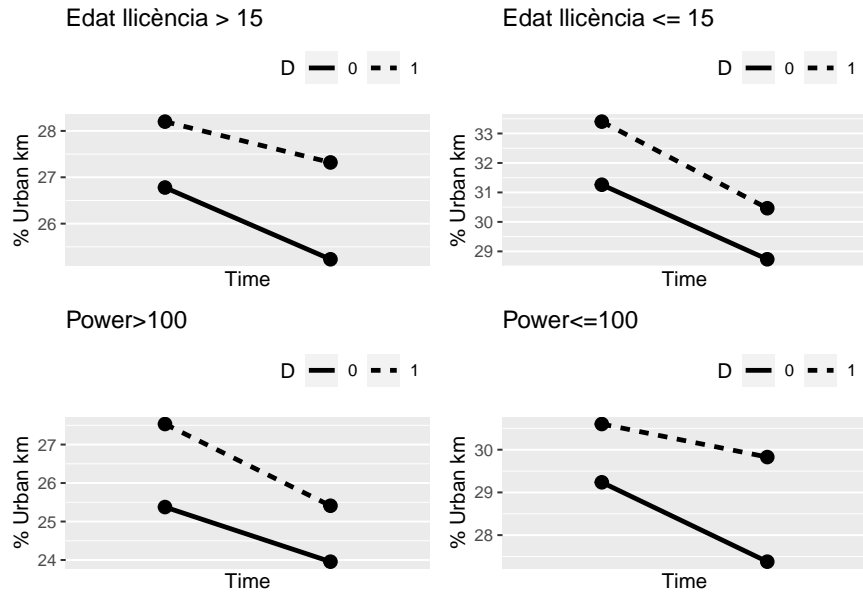
| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|-----------------|----------|------------|---------|----------|--------|--------|----------|
| age35_2_Tr | -1.8221 | 1.4534 | -1.2537 | 0.2100 | 0.7427 | 0.4837 | 84878.47 |
| lic_age15_2_Tr | -2.8606 | 1.3877 | -2.0615 | 0.0393 | 0.7427 | 0.4837 | 84878.47 |
| parking_yes2_Tr | 0.3418 | 1.0936 | 0.3126 | 0.7546 | 0.7427 | 0.4837 | 84878.47 |
| woman_2_Tr | 2.3575 | 1.0453 | 2.2554 | 0.0241 | 0.7427 | 0.4837 | 84878.47 |
| BMzones_2_Tr | 0.7338 | 1.4166 | 0.5180 | 0.6045 | 0.7427 | 0.4837 | 84878.47 |
| power100_2_Tr | 2.7862 | 1.0550 | 2.6410 | 0.0083 | 0.7427 | 0.4837 | 84878.47 |

Table 59: TWFE for Urban variable with reduced covariates

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|----------------|----------|------------|---------|----------|--------|--------|----------|
| Tr | -1.6205 | 0.8621 | -1.8797 | 0.0602 | 0.7426 | 0.4843 | 84868.04 |
| age35 | 2.5052 | 0.5653 | 4.4318 | 0.0000 | 0.7426 | 0.4843 | 84868.04 |
| age35_2 | -1.3304 | 0.4179 | -3.1834 | 0.0015 | 0.7426 | 0.4843 | 84868.04 |
| lic_age15 | 1.8876 | 0.4889 | 3.8608 | 0.0001 | 0.7426 | 0.4843 | 84868.04 |
| woman_2 | -0.4615 | 0.3148 | -1.4657 | 0.1428 | 0.7426 | 0.4843 | 84868.04 |
| BMzones | -2.5892 | 0.9109 | -2.8424 | 0.0045 | 0.7426 | 0.4843 | 84868.04 |
| power100 | 2.1178 | 0.4334 | 4.8869 | 0.0000 | 0.7426 | 0.4843 | 84868.04 |
| power100_2 | -0.5751 | 0.3406 | -1.6882 | 0.0914 | 0.7426 | 0.4843 | 84868.04 |
| lic_age15_2_Tr | -3.3355 | 1.2498 | -2.6689 | 0.0076 | 0.7426 | 0.4843 | 84868.04 |
| woman_2_Tr | 2.3824 | 1.0411 | 2.2883 | 0.0221 | 0.7426 | 0.4843 | 84868.04 |
| power100_2_Tr | 2.7532 | 1.0493 | 2.6237 | 0.0087 | 0.7426 | 0.4843 | 84868.04 |

Table 60: Some effects

| | lic_age<=15 | woman | power<=100 |
|---------|-------------|--------|------------|
| coef | -3.0684 | 0.3005 | 2.6754 |
| Z | -2.2092 | 0.2877 | 2.7531 |
| p_value | 0.0136 | 0.3868 | 0.0030 |



Tal com fet amb la resta de models, es mirarà altre cop què passaria si només considerèssim la variable **age35** o **lic_age15**, però no totes dues a la vegada, en els models. En aquest cas no s'ha hagut de crear 4 models (considerant a més a més introduir o no la variable **parking_yes**), perquè aquesta era no significant en tots els cassos, indicant que no té res a veure amb la quantitat de conducció urbana que realitzen els ciutadans.

Les conclusions són que independentment de si es posa la variable edat o anys de llicència, totes dues mostren un efecte del tractament significatiu (o quasi pel cas d'**age35**) i negatiu, mentre la potència també surt en tot cas significativa i positiva, com passava en el cas de la variable **Night**.

A més a més, si s'analitzen els R^2_{Adj} o els AIC es pot notar que introduir els anys de llicència és més informatiu que introduir l'edat. Tot i així, el millor model continua sent el creat anteriorment, que considera totes dues informacions a la vegada.

Després d'haver analitzat tots els models d'aquesta secció i de tenir en compte els AIC's es pot concloure que posseir la variable edat és el millor si es vol predir el percentatge de quilòmetres que es recorreran de forma nocturna, mentre per a les altres 3 variables dependents aporta més informació els anys de carnet. Òbviament, en tots els cassos el millor és disposar de totes dues informacions.

Table 61: TWFE for Urban variable with reduced covariates without **lic_age15**

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|---------------|----------|------------|---------|----------|-------|--------|---------|
| Tr | -1.5756 | 0.8566 | -1.8394 | 0.0659 | 0.742 | 0.4831 | 84881.4 |
| age35 | 3.1512 | 0.5378 | 5.8590 | 0.0000 | 0.742 | 0.4831 | 84881.4 |
| age35_2 | -1.1191 | 0.4332 | -2.5833 | 0.0098 | 0.742 | 0.4831 | 84881.4 |
| BMzones | -2.1447 | 0.9313 | -2.3028 | 0.0213 | 0.742 | 0.4831 | 84881.4 |
| BMzones_2 | -0.6456 | 0.4283 | -1.5073 | 0.1318 | 0.742 | 0.4831 | 84881.4 |
| power100 | 2.2734 | 0.4320 | 5.2625 | 0.0000 | 0.742 | 0.4831 | 84881.4 |
| power100_2 | -0.7103 | 0.3316 | -2.1423 | 0.0322 | 0.742 | 0.4831 | 84881.4 |
| age35_2_Tr | -2.7783 | 1.3586 | -2.0449 | 0.0409 | 0.742 | 0.4831 | 84881.4 |
| woman_2_Tr | 1.9275 | 0.9933 | 1.9405 | 0.0523 | 0.742 | 0.4831 | 84881.4 |
| power100_2_Tr | 2.7361 | 1.0461 | 2.6155 | 0.0089 | 0.742 | 0.4831 | 84881.4 |

Table 62: Some effects

| | age35 | woman | power100 |
|---------|---------|--------|----------|
| coef | -2.3218 | 0.3519 | 2.7236 |
| Z | -1.6121 | 0.3383 | 2.8324 |
| p_value | 0.0535 | 0.3676 | 0.0023 |

Table 63: TWFE for Urban variable with reduced covariates without **age35**

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|----------------|----------|------------|---------|----------|--------|--------|---------|
| Tr | -1.6240 | 0.8627 | -1.8826 | 0.0598 | 0.7423 | 0.4837 | 84886.4 |
| lic_age15 | 2.4395 | 0.4597 | 5.3061 | 0.0000 | 0.7423 | 0.4837 | 84886.4 |
| woman_2 | -0.4705 | 0.3151 | -1.4930 | 0.1355 | 0.7423 | 0.4837 | 84886.4 |
| BMzones | -2.7573 | 0.9107 | -3.0277 | 0.0025 | 0.7423 | 0.4837 | 84886.4 |
| power100 | 2.2219 | 0.4332 | 5.1296 | 0.0000 | 0.7423 | 0.4837 | 84886.4 |
| power100_2 | -0.6344 | 0.3404 | -1.8639 | 0.0624 | 0.7423 | 0.4837 | 84886.4 |
| lic_age15_2_Tr | -3.6399 | 1.2463 | -2.9205 | 0.0035 | 0.7423 | 0.4837 | 84886.4 |

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|---------------|----------|------------|---------|----------|--------|--------|---------|
| woman_2_Tr | 2.4324 | 1.0419 | 2.3345 | 0.0196 | 0.7423 | 0.4837 | 84886.4 |
| power100_2_Tr | 2.7945 | 1.0501 | 2.6612 | 0.0078 | 0.7423 | 0.4837 | 84886.4 |

Table 64: Some effects

| | lic_age15 | woman | power100 |
|---------|-----------|--------|----------|
| coef | -2.8245 | 0.3379 | 2.7579 |
| Z | -2.0522 | 0.3233 | 2.8363 |
| p_value | 0.0201 | 0.3732 | 0.0023 |

3.2.2 Model semi-paramètric - Integració de propensity scores amb models Dif-in-Dif

L'estimador utilitza puntuacions de propensió per predir les probabilitats que cada observació estigui en cadascun dels quatre grups (control pre-tractament / control post-tractament / tractament pre-tractament / tractament post-tractament). Concretament, normalment s'utilitza algun tipus de regressió multinomial. L'efecte del tractament es calcula doncs com una mitjana ponderada dels resultats observats. Destacar també que en aquest cas per a calcular la variància de l'estimador i posteriorment poder fer inferència s'ha utilitzat el mètode delta. Així doncs, l'estimador de la variància en realitat és una aproximació.

Els resultats, a l'igual que per l'anterior mètode no paramètric donen en tot cas no significatius. Concretament, la variable **Total** sempre és la que mostra un pvalor més petit, mentre les altres tres variables dependents mostren pvalors molt propers a l'1 si es mira el resultat a 2 cues, interpretant doncs que no s'està captant absolutament cap informació rellevant de les dades. Com en tot cas, efectivament fer inferència a 1 cua fa reduir el pvalor a la meitat.

3.2.2.1 Variable Total amb l'efecte de covariables

Table 65: Propensity scores + DD for Total variable

| | ATT_2 | sd | z_stat | p_value | CI_lower | CI_upper |
|---------|----------|----------|--------|---------|-----------|----------|
| 2-tails | 795.6984 | 1925.327 | 0.4133 | 0.6794 | -2977.873 | 4569.27 |
| 1-tail | 795.6984 | 1925.327 | 0.4133 | 0.3397 | -2371.183 | Inf |

3.2.2.2 Variable Night amb l'efecte de covariables

Table 66: Propensity scores + DD for Night variable

| | ATT_2 | sd | z_stat | p_value | CI_lower | CI_upper |
|---------|--------|--------|--------|---------|----------|----------|
| 2-tails | 0.0388 | 3.3174 | 0.0117 | 0.9907 | -6.4632 | 6.5408 |
| 1-tail | 0.0388 | 3.3174 | 0.0117 | 0.4953 | -5.4179 | Inf |

3.2.2.3 Variable Speed amb l'efecte de covariables

Table 67: Propensity scores + DD for Speed variable

| | ATT_2 | sd | z_stat | p_value | CI_lower | CI_upper |
|---------|---------|--------|---------|---------|----------|----------|
| 2-tails | -0.2023 | 3.2749 | -0.0618 | 0.9507 | -6.621 | 6.2163 |
| 1-tail | -0.2023 | 3.2749 | -0.0618 | 0.4754 | -Inf | 5.1844 |

3.2.2.4 Variable Urban amb l'efecte de covariables

Table 68: Propensity scores + DD for Urban variable

| | ATT_2 | sd | z_stat | p_value | CI_lower | CI_upper |
|---------|--------|--------|--------|---------|----------|----------|
| 2-tails | 0.5021 | 10.379 | 0.0484 | 0.9614 | -19.8405 | 20.8446 |
| 1-tail | 0.5021 | 10.379 | 0.0484 | 0.4807 | -16.5699 | Inf |

3.3 Models amb covariables constants en el temps

A continuació, es vol dur a terme un anàlisi similar al fet anteriorment però considerant el conjunt de dades que té les variables constants en el temps.

3.3.1 Model TWFE

3.3.1.1 Variable Total amb l'efecte de covariables En primer lloc creem el model amb tots els efectes additius i multiplicatius de les covariables, així com la seva corresponent reducció feta manualment en estar treballant amb models de panells, els quals no permeten fer una optimització automàtica. Com es pot notar en aquest cas no se li introdueixen les variables sense el `_2`.*.

Table 69: TWFE for Total variable with constant covariates

| | Estimate | Std. Error | t-value | Pr(> t) | R2 | R2Adj |
|--------------|-----------|------------|---------|----------|--------|--------|
| Tr | 807.3561 | 243.0341 | 3.3220 | 0.0009 | 0.5319 | 0.4049 |
| lic_age15 | -360.3023 | 218.4160 | -1.6496 | 0.0991 | 0.5319 | 0.4049 |
| parking_yes | 975.3139 | 255.3572 | 3.8194 | 0.0001 | 0.5319 | 0.4049 |
| parking_yes2 | -706.2360 | 159.2244 | -4.4355 | 0.0000 | 0.5319 | 0.4049 |
| woman | -799.2458 | 199.7551 | -4.0011 | 0.0001 | 0.5319 | 0.4049 |
| woman_2 | 221.5632 | 151.6239 | 1.4613 | 0.1440 | 0.5319 | 0.4049 |

Table 70: TWFE for Total variable with covariates

| | Estimate | Std. Error | t-value | Pr(> t) | R2 | R2Adj |
|--------------|-----------|------------|---------|----------|--------|--------|
| Tr | 807.3561 | 243.0341 | 3.3220 | 0.0009 | 0.5319 | 0.4049 |
| lic_age15 | -360.3023 | 218.4160 | -1.6496 | 0.0991 | 0.5319 | 0.4049 |
| parking_yes | 975.3139 | 255.3572 | 3.8194 | 0.0001 | 0.5319 | 0.4049 |
| parking_yes2 | -706.2360 | 159.2244 | -4.4355 | 0.0000 | 0.5319 | 0.4049 |
| woman | -799.2458 | 199.7551 | -4.0011 | 0.0001 | 0.5319 | 0.4049 |
| woman_2 | 221.5632 | 151.6239 | 1.4613 | 0.1440 | 0.5319 | 0.4049 |

Tanmateix, com fet anteriorment amb les dades que no eren constants en el temps, es decideix fer el model amb la funció `lm` per tal de poder optimitzar el resultat d'una forma més senzilla. Com fet anteriorment, primer s'han de centrar les variables al 0.

En aquest cas l'estimador de l'efecte del tractament és positiu i significatiu, havent augmentat el valor de l'estimador respecte al valor obtingut amb el dataset amb covariables no constants en el temps. També continuen sortint les mateixes variables com a importants excepte `lic_age15`, que ha desaparegut. Les que apareixen, ho fan amb la mateixa magnitud i signe.

Com a diferència rellevant, en aquest cas s'ha trobat 2 covariables amb interaccions significatives amb `Tr`. Tot i així, només la variable `woman` presenta un ATT_2 significatiu amb coeficient positiu. Amb l'ajuda del gràfic es pot concloure que els usuaris que han tingut un accident sempre condueixen més quilòmetres totals de mitja. Tanmateix, en el cas dels homes, haver-lo tingut implica disminuir menys la quantitat de condució que si no l'haguessin tingut respecte a les dones.

Table 71: TWFE for Total variable with cst covariates

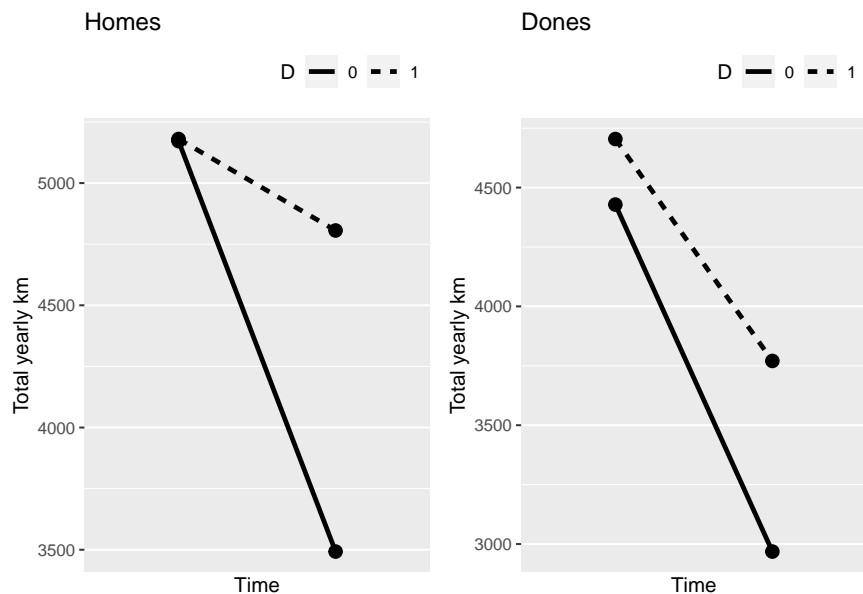
| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|-----------------|------------|------------|---------|----------|--------|--------|----------|
| Tr | 1314.9408 | 457.4835 | 2.8743 | 0.0041 | 0.5669 | 0.1309 | 151918.6 |
| age35_2 | -155.5828 | 189.1700 | -0.8224 | 0.4108 | 0.5669 | 0.1309 | 151918.6 |
| lic_age15_2 | 299.2594 | 181.8971 | 1.6452 | 0.1000 | 0.5669 | 0.1309 | 151918.6 |
| parking_yes2 | -802.4975 | 133.7476 | -6.0001 | 0.0000 | 0.5669 | 0.1309 | 151918.6 |
| woman_2 | 222.0664 | 130.0064 | 1.7081 | 0.0877 | 0.5669 | 0.1309 | 151918.6 |
| BMzones_2 | 283.2970 | 178.5069 | 1.5870 | 0.1125 | 0.5669 | 0.1309 | 151918.6 |
| power100_2 | -88.6118 | 133.3184 | -0.6647 | 0.5063 | 0.5669 | 0.1309 | 151918.6 |
| age35_2_Tr | -1185.3034 | 617.7485 | -1.9187 | 0.0551 | 0.5669 | 0.1309 | 151918.6 |
| lic_age15_2_Tr | -223.7309 | 614.4712 | -0.3641 | 0.7158 | 0.5669 | 0.1309 | 151918.6 |
| parking_yes2_Tr | -125.0706 | 442.3756 | -0.2827 | 0.7774 | 0.5669 | 0.1309 | 151918.6 |
| woman_2_Tr | -873.1574 | 440.4989 | -1.9822 | 0.0475 | 0.5669 | 0.1309 | 151918.6 |
| BMzones_2_Tr | 807.3567 | 598.7511 | 1.3484 | 0.1776 | 0.5669 | 0.1309 | 151918.6 |
| power100_2_Tr | 324.4167 | 441.7583 | 0.7344 | 0.4627 | 0.5669 | 0.1309 | 151918.6 |

Table 72: TWFE for Total variable with cst reduced covariates

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|--------------|------------|------------|---------|----------|--------|--------|----------|
| Tr | 1480.5794 | 297.5773 | 4.9754 | 0.0000 | 0.5663 | 0.1311 | 151909.7 |
| parking_yes2 | -788.1425 | 125.7845 | -6.2658 | 0.0000 | 0.5663 | 0.1311 | 151909.7 |
| woman_2 | 215.9742 | 125.3211 | 1.7234 | 0.0849 | 0.5663 | 0.1311 | 151909.7 |
| BMzones_2 | 371.8368 | 169.8095 | 2.1897 | 0.0286 | 0.5663 | 0.1311 | 151909.7 |
| age35_2_Tr | -1311.6179 | 555.9548 | -2.3592 | 0.0183 | 0.5663 | 0.1311 | 151909.7 |
| woman_2_Tr | -787.1851 | 427.0280 | -1.8434 | 0.0653 | 0.5663 | 0.1311 | 151909.7 |

Table 73: Some effects

| | age35 | woman |
|---------|----------|----------|
| coef | 168.9615 | 909.3685 |
| Z | 0.3092 | 2.7519 |
| p_value | 0.3786 | 0.0030 |



Tal com fet amb els models amb covariables canviants en el temps, es tornarà a analitzar com canvien aquests si s'introdueix únicament la variable `age35` o `lic_age15` dins d'aquests; així com també es mirarà com canvien al eliminar la variable `parking_yes` per complet si la seva interacció amb `Tr` surt significativa.

En aquest cas la interacció de `Tr` amb la variable `parking_yes` és no significativa i per tant no es considera treure-la del model. De fet, es nota que en aquest cas, considerant els estudis que es volen dur a terme, només s'ha d'estudiar si eliminar la variable `age35` del model dona lloc a que els anys de llicència "reemplacin" la informació en el model o no. A l'analitzar els resultats es pot notar que la informació és clarament no reemplaçada en no sortir el coeficient `lic_age15_2_Tr` com a significatiu. A més a més, l'AIC del model ha disminuït respecte l'anterior, tot i que l' R^2_{Adj} s'ha mantingut estable. La conclusió és clara: es necessita l'edat de l'assegurat per a predir amb major encert el nombre de quilòmetres totals que realitzarà un assegurat i els anys de llicència que tenen aquests conductors no aporta la mateixa informació. Fer notar que aquesta conclusió és contrària a la que s'obtenia quan s'analitzava el conjunt de dades amb covariables canviants en el temps. També es pot notar que eliminar l'edat fa que la interacció de la variable `woman` amb `Tr` sigui menys significativa que anteriorment, tot i que es continua trobant per sota del límit del 5%.

Table 74: TWFE for Total variable with cst reduced covariates without age35

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|--------------|-----------|------------|---------|----------|--------|--------|----------|
| Tr | 1270.3442 | 284.0007 | 4.4730 | 0.0000 | 0.5661 | 0.1311 | 151913.3 |
| parking_yes2 | -799.1166 | 125.7337 | -6.3556 | 0.0000 | 0.5661 | 0.1311 | 151913.3 |
| woman_2 | 216.3443 | 125.3561 | 1.7258 | 0.0844 | 0.5661 | 0.1311 | 151913.3 |
| BMzones_2 | 366.9327 | 169.8443 | 2.1604 | 0.0308 | 0.5661 | 0.1311 | 151913.3 |
| woman_2_Tr | -780.5223 | 427.1384 | -1.8273 | 0.0677 | 0.5661 | 0.1311 | 151913.3 |

Table 75: Some effects

| woman | |
|-------|----------|
| coef | 706.1661 |
| Z | 2.2129 |

| | woman |
|---------|--------|
| p_value | 0.0135 |

3.3.1.2 Variable Night amb l'efecte de covariables Així com passava en el cas del conjunt de dades amb les covariables canviants en el temps l'estimador de l'efecte del tractament és no significatiu i molt proper a 0. Tanmateix, en aquest cas només ha sortit 1 interacció amb la variable **Tr** com a rellevant. Al mostrar un efecte del tractament no significatiu, es conclou que cap de les covariables incloses en el conjunt de dades té un efecte rellevant en relació a si s'ha tingut un accident o no. Així doncs, els efectes anteriorment vistos a l'analitzar el conjunt de dades amb les covariables canviants en el temps sobre les variables **power100** i **lic_age15** han desaparegut.

Pel que fa a les variables sense interacció es pot notar que tenir menys de 35 anys, disposar de pàrquing i tenir un cotxe petit dona lloc a recórrer un percentatge de quilòmetres menor durant la nit durant el període de post-tractament. Contràriament, ser dona i conduir per Barcelona o Madrid, com a zones principals, augmenten aquest percentatge.

Table 76: TWFE for Night variable with cst covariates

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|-----------------|----------|------------|---------|----------|--------|--------|----------|
| Tr | 0.1125 | 0.8103 | 0.1388 | 0.8896 | 0.5858 | 0.1689 | 48641.17 |
| age35_2 | -1.3587 | 0.3351 | -4.0549 | 0.0001 | 0.5858 | 0.1689 | 48641.17 |
| lic_age15_2 | 0.4530 | 0.3222 | 1.4060 | 0.1598 | 0.5858 | 0.1689 | 48641.17 |
| parking_yes2 | -0.4272 | 0.2369 | -1.8032 | 0.0714 | 0.5858 | 0.1689 | 48641.17 |
| woman_2 | 0.7643 | 0.2303 | 3.3189 | 0.0009 | 0.5858 | 0.1689 | 48641.17 |
| BMzones_2 | 0.8543 | 0.3162 | 2.7019 | 0.0069 | 0.5858 | 0.1689 | 48641.17 |
| power100_2 | -0.7984 | 0.2361 | -3.3810 | 0.0007 | 0.5858 | 0.1689 | 48641.17 |
| age35_2_Tr | -0.5479 | 1.0942 | -0.5007 | 0.6166 | 0.5858 | 0.1689 | 48641.17 |
| lic_age15_2_Tr | -1.0566 | 1.0884 | -0.9708 | 0.3317 | 0.5858 | 0.1689 | 48641.17 |
| parking_yes2_Tr | -0.7938 | 0.7836 | -1.0130 | 0.3111 | 0.5858 | 0.1689 | 48641.17 |
| woman_2_Tr | 0.8677 | 0.7803 | 1.1120 | 0.2661 | 0.5858 | 0.1689 | 48641.17 |
| BMzones_2_Tr | -0.6225 | 1.0606 | -0.5869 | 0.5573 | 0.5858 | 0.1689 | 48641.17 |
| power100_2_Tr | 1.0581 | 0.7825 | 1.3522 | 0.1764 | 0.5858 | 0.1689 | 48641.17 |

Table 77: TWFE for Night variable with cst reduced covariates

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|---------------|----------|------------|---------|----------|--------|--------|----------|
| Tr | -0.4574 | 0.5605 | -0.8161 | 0.4144 | 0.5856 | 0.1697 | 48635.25 |
| age35_2 | -1.2535 | 0.2927 | -4.2828 | 0.0000 | 0.5856 | 0.1697 | 48635.25 |
| parking_yes2 | -0.4833 | 0.2248 | -2.1493 | 0.0316 | 0.5856 | 0.1697 | 48635.25 |
| woman_2 | 0.8613 | 0.2195 | 3.9230 | 0.0001 | 0.5856 | 0.1697 | 48635.25 |
| BMzones_2 | 0.8177 | 0.3011 | 2.7152 | 0.0066 | 0.5856 | 0.1697 | 48635.25 |
| power100_2 | -0.7935 | 0.2344 | -3.3851 | 0.0007 | 0.5856 | 0.1697 | 48635.25 |
| power100_2_Tr | 1.2558 | 0.7553 | 1.6625 | 0.0964 | 0.5856 | 0.1697 | 48635.25 |

Table 78: Some effects

| | power100 |
|------|----------|
| coef | 0.0048 |
| Z | 0.0093 |

| | power100 |
|---------|----------|
| p_value | 0.4963 |

A diferència de amb les dades amb covariables canviants en el temps, la interacció amb **Tr** de la variable **parking_yes** no surt significativa i per tant només fa falta estudiar què passa quan deixem en el model únicament **age35** o **lic_age15**. En aquest aspecte, el model altre cop es diferencia de la secció anterior, en mostrar únicament la variable **age35** com a important. Així doncs, podem mirar què passaria si únicament introduïm en el model **lic_age15**, tot i que ja sabem que els resultats empitjoraran. Es vol saber doncs si tot i no ser millor que **age35**, **lic_age15** pot ajudar a predir el percentatge de quilòmetres recorreguts nocturnament en absència de l'edat. La resposta és no, en no sortir la variable en el model creat. La significació de la interacció **power100-Tr** no canvia de forma rellevant.

Table 79: TWFE for Night variable with cst reduced covariates without age35

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|---------------|----------|------------|---------|----------|--------|--------|----------|
| Tr | -0.5020 | 0.5610 | -0.8948 | 0.3709 | 0.5849 | 0.1685 | 48651.58 |
| parking_yes2 | -0.5615 | 0.2243 | -2.5029 | 0.0123 | 0.5849 | 0.1685 | 48651.58 |
| woman_2 | 0.8644 | 0.2198 | 3.9329 | 0.0001 | 0.5849 | 0.1685 | 48651.58 |
| BMzones_2 | 0.7763 | 0.3013 | 2.5765 | 0.0100 | 0.5849 | 0.1685 | 48651.58 |
| power100_2 | -0.8741 | 0.2339 | -3.7370 | 0.0002 | 0.5849 | 0.1685 | 48651.58 |
| power100_2_Tr | 1.3244 | 0.7560 | 1.7520 | 0.0798 | 0.5849 | 0.1685 | 48651.58 |

Table 80: Some effects

| | power100 |
|---------|----------|
| coef | -0.0516 |
| Z | -0.0996 |
| p_value | 0.4603 |

3.3.1.3 Variable Speed amb l'efecte de covariables Les conclusions més rellevants aquest cop són similars al model obtingut amb les covariables canviants en el temps. D'una banda, tots dos models han trobat varies interaccions dins del millor model escollit: **age35**, **lic_age15** i **power100**. Així mateix, en tots dos s'observa un efecte del tractament negatiu però no significatiu. A l'analitzar l'efecte a $t=2$ es nota que aquest és positiu i significatiu per a totes les variables, a diferència dels resultats que es trobaven anteriorment.

A l'analitzar les característiques per separat notem que totes les variables analitzades són significatives en quant a l'efecte del tractament. D'una banda l'edat presenta un ATT_2 negatiu indicant que aquells usuaris joves que tenen un accident, generalment decreixen la velocitat després de patir-lo. D'altra banda, els anys de llicència i la potència del cotxe presenten un valor positiu indicant que els usuaris que tenen cotxes petits i pocs anys de llicència, incrementen la velocitat després de tenir un accident, contràriament a l'esperat. Amb l'ajuda de les gràfiques podem veure que de fet, pel que fa als usuaris amb pocs anys de carnet, no és que aquells que tinguin un accident augmentin la seva velocitat, sinó que els que no el tenen la disminueixen encara més respecte al seu comportament anterior. Un comportament similar però menys marcat succeeix pels cotxes petits.

Table 81: TWFE for Speed variable with cst covariates

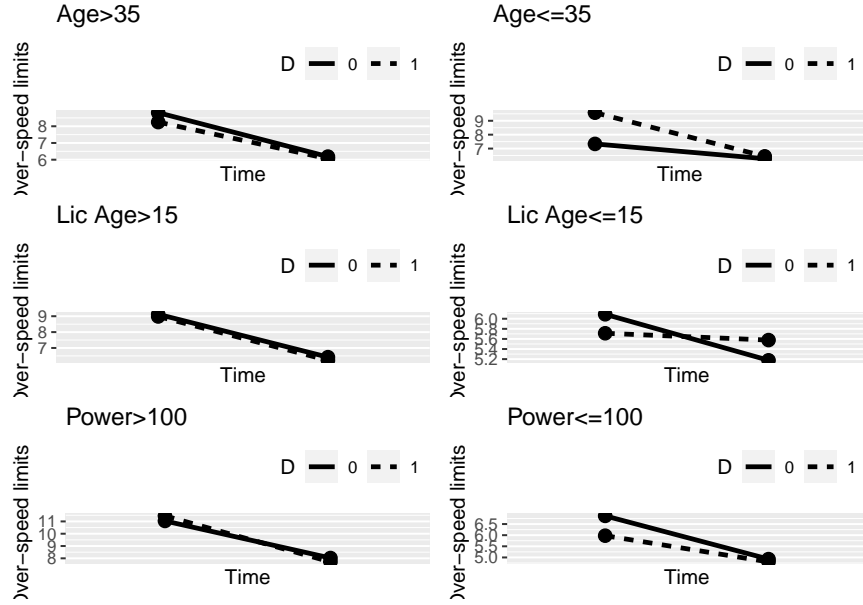
| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|-----------------|----------|------------|---------|----------|--------|--------|----------|
| Tr | -0.6282 | 0.6748 | -0.9309 | 0.3519 | 0.7708 | 0.5401 | 45657.81 |
| age35_2 | 0.9046 | 0.2790 | 3.2419 | 0.0012 | 0.7708 | 0.5401 | 45657.81 |
| lic_age15_2 | 1.2739 | 0.2683 | 4.7480 | 0.0000 | 0.7708 | 0.5401 | 45657.81 |
| parking_yes2 | -0.1447 | 0.1973 | -0.7336 | 0.4632 | 0.7708 | 0.5401 | 45657.81 |
| woman_2 | -0.1912 | 0.1918 | -0.9970 | 0.3188 | 0.7708 | 0.5401 | 45657.81 |
| BMzones_2 | 0.8940 | 0.2633 | 3.3954 | 0.0007 | 0.7708 | 0.5401 | 45657.81 |
| power100_2 | 0.9417 | 0.1967 | 4.7888 | 0.0000 | 0.7708 | 0.5401 | 45657.81 |
| age35_2_Tr | -2.7658 | 0.9112 | -3.0353 | 0.0024 | 0.7708 | 0.5401 | 45657.81 |
| lic_age15_2_Tr | 1.7177 | 0.9064 | 1.8952 | 0.0581 | 0.7708 | 0.5401 | 45657.81 |
| parking_yes2_Tr | -0.2043 | 0.6525 | -0.3131 | 0.7542 | 0.7708 | 0.5401 | 45657.81 |
| woman_2_Tr | 1.0097 | 0.6498 | 1.5540 | 0.1202 | 0.7708 | 0.5401 | 45657.81 |
| BMzones_2_Tr | -0.8850 | 0.8832 | -1.0020 | 0.3164 | 0.7708 | 0.5401 | 45657.81 |
| power100_2_Tr | 1.2600 | 0.6516 | 1.9337 | 0.0532 | 0.7708 | 0.5401 | 45657.81 |

Table 82: TWFE for Speed variable with cst reduced covariates

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|----------------|----------|------------|---------|----------|--------|--------|----------|
| Tr | -0.5644 | 0.4898 | -1.1522 | 0.2493 | 0.7702 | 0.5395 | 45652.28 |
| age35_2 | 0.9056 | 0.2788 | 3.2486 | 0.0012 | 0.7702 | 0.5395 | 45652.28 |
| lic_age15_2 | 1.2413 | 0.2665 | 4.6584 | 0.0000 | 0.7702 | 0.5395 | 45652.28 |
| BMzones_2 | 0.8177 | 0.2505 | 3.2637 | 0.0011 | 0.7702 | 0.5395 | 45652.28 |
| power100_2 | 0.9103 | 0.1890 | 4.8164 | 0.0000 | 0.7702 | 0.5395 | 45652.28 |
| age35_2_Tr | -2.8461 | 0.9078 | -3.1353 | 0.0017 | 0.7702 | 0.5395 | 45652.28 |
| lic_age15_2_Tr | 1.7032 | 0.9003 | 1.8919 | 0.0585 | 0.7702 | 0.5395 | 45652.28 |
| power100_2_Tr | 1.5360 | 0.6308 | 2.4348 | 0.0149 | 0.7702 | 0.5395 | 45652.28 |

Table 83: Some effects

| | age<=35 | lic_age<=15 | power<=100 |
|---------|---------|-------------|------------|
| coef | -2.5049 | 2.3802 | 1.8820 |
| Z | -2.7733 | 2.5666 | 4.0495 |
| p_value | 0.0028 | 0.0051 | 0.0000 |



Altres cops la variable `parking_yes` no surt en cap lloc del model i per tant no ens hem de preocupar per ella. No obstant això, en aquest cas sí que apareixen com a rellevants tant `age35` com `lic_age15`, inclús presentant totes dues una interacció amb `Tr` en el model. Anteriorment s'ha vist que totes dues variable, així com `power100` presentaven un ATT_2 significatiu. A continuació es mirarà si la significació canvia a l'eliminar una de les dues variables de model. Eliminar els anys de llicència dona lloc a un pvalor de l' ATT_2 de la variable edat significatiu al 4.5%, mentre la significació de `power100` es continua mantenint intacte. D'altra banda, eliminar l'edat dona lloc a una eliminació completa de la interacció dels anys de llicència amb `Tr`, mentre altre cop els resultats de `power100` no canvien.

Tenint en compte el mencionat anteriorment i el fet que l'AIC del model sense `age35` és menor (i l' R^2_{Adj} major) en comparació al model sense `lic_age15`, s'hauria de concloure que si es vol analitzar com canvia el comportament d'un assegurat en front un accident seria millor disposar de la variable edat, mentre si es vol donar més importància a predir correctament el percentatge de quilòmetres que es durant a terme per sobre de la velocitat els anys de llicència d'un assegurat són més importants. Òbviament, en tot cas la millor opció seria disposar de totes dues variables, com en aquest treball.

Table 84: TWFE for Speed variable with cst reduced covariates without `lic_age15`

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|---------------|----------|------------|---------|----------|--------|--------|----------|
| Tr | -0.4159 | 0.4854 | -0.8569 | 0.3915 | 0.7687 | 0.5366 | 45681.64 |
| age35_2 | 1.4369 | 0.2548 | 5.6391 | 0.0000 | 0.7687 | 0.5366 | 45681.64 |
| BMzones_2 | 0.9040 | 0.2505 | 3.6088 | 0.0003 | 0.7687 | 0.5366 | 45681.64 |
| power100_2 | 0.9952 | 0.1885 | 5.2796 | 0.0000 | 0.7687 | 0.5366 | 45681.64 |
| age35_2_Tr | -2.4448 | 0.8597 | -2.8438 | 0.0045 | 0.7687 | 0.5366 | 45681.64 |
| power100_2_Tr | 1.6132 | 0.6301 | 2.5602 | 0.0105 | 0.7687 | 0.5366 | 45681.64 |

Table 85: Some effects

| | age35 | power100 |
|------|---------|----------|
| coef | -1.4239 | 2.1924 |
| Z | -1.7012 | 4.8666 |

| | age35 | power100 |
|---------|--------|----------|
| p_value | 0.0445 | 0.0000 |

Table 86: TWFE for Speed variable with cst reduced covariates without age35

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|---------------|----------|------------|---------|----------|--------|--------|----------|
| Tr | -0.7580 | 0.4668 | -1.6238 | 0.1044 | 0.7706 | 0.5405 | 45662.63 |
| lic_age15_2 | 1.6559 | 0.2333 | 7.0989 | 0.0000 | 0.7706 | 0.5405 | 45662.63 |
| BMzones_2 | 0.8126 | 0.2507 | 3.2407 | 0.0012 | 0.7706 | 0.5405 | 45662.63 |
| power100_2 | 0.9236 | 0.1890 | 4.8877 | 0.0000 | 0.7706 | 0.5405 | 45662.63 |
| power100_2_Tr | 1.5848 | 0.6293 | 2.5183 | 0.0118 | 0.7706 | 0.5405 | 45662.63 |

Table 87: Some effects

| | power100 |
|---------|----------|
| coef | 1.7504 |
| Z | 4.0747 |
| p_value | 0.0000 |

3.3.1.4 Variable Urban amb l'efecte de covariables Tornem a trobar que algunes interaccions són significatives. D'altra banda, l'estimador de l'efecte del tractament és negatiu i significatiu, indicant que tenir un accident disminueix el percentatge de quilòmetres recorreguts per via urbana. Totes les variables que han sortit com a significatives mostren un efecte negatiu a $t=2$, indicant que tenir menys de 35 anys, tenir la llicència fa menys de 15 anys, conduir principalment per Barcelona o Madrid i tenir un cotxe petit fa disminuir el percentatge de quilòmetres recorreguts per vies urbanes.

Finalment, comentar que tot i que la interacció entre les variables `woman` i `lic_age15` i `Tr` milloren el model en quant a informació captada en relació a la quantitat d'informació considerada, al treure els estadístics particulars, els ATT_2 són no significatius.

Table 88: TWFE for Urban variable with cst covariates

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|-----------------|----------|------------|---------|----------|--------|--------|----------|
| Tr | -2.1855 | 1.1488 | -1.9024 | 0.0571 | 0.8128 | 0.6243 | 54329.76 |
| age35_2 | -0.6695 | 0.4750 | -1.4094 | 0.1587 | 0.8128 | 0.6243 | 54329.76 |
| lic_age15_2 | -1.3688 | 0.4568 | -2.9969 | 0.0027 | 0.8128 | 0.6243 | 54329.76 |
| parking_yes2 | 0.3585 | 0.3359 | 1.0676 | 0.2857 | 0.8128 | 0.6243 | 54329.76 |
| woman_2 | 0.2167 | 0.3265 | 0.6639 | 0.5068 | 0.8128 | 0.6243 | 54329.76 |
| BMzones_2 | -1.0116 | 0.4482 | -2.2567 | 0.0241 | 0.8128 | 0.6243 | 54329.76 |
| power100_2 | -0.6521 | 0.3348 | -1.9480 | 0.0514 | 0.8128 | 0.6243 | 54329.76 |
| age35_2_Tr | 0.1554 | 1.5512 | 0.1002 | 0.9202 | 0.8128 | 0.6243 | 54329.76 |
| lic_age15_2_Tr | 2.6015 | 1.5430 | 1.6860 | 0.0918 | 0.8128 | 0.6243 | 54329.76 |
| parking_yes2_Tr | 0.1759 | 1.1108 | 0.1583 | 0.8742 | 0.8128 | 0.6243 | 54329.76 |
| woman_2_Tr | 2.4314 | 1.1061 | 2.1981 | 0.0280 | 0.8128 | 0.6243 | 54329.76 |
| BMzones_2_Tr | -1.5878 | 1.5035 | -1.0561 | 0.2910 | 0.8128 | 0.6243 | 54329.76 |
| power100_2_Tr | 0.8521 | 1.1093 | 0.7682 | 0.4424 | 0.8128 | 0.6243 | 54329.76 |

Table 89: TWFE for Urban variable with cst reduced covariates

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|----------------|----------|------------|---------|----------|--------|--------|----------|
| Tr | -2.0101 | 0.7288 | -2.7581 | 0.0058 | 0.8125 | 0.6244 | 54321.44 |
| age35_2 | -0.6494 | 0.4515 | -1.4381 | 0.1504 | 0.8125 | 0.6244 | 54321.44 |
| lic_age15_2 | -1.3087 | 0.4498 | -2.9093 | 0.0036 | 0.8125 | 0.6244 | 54321.44 |
| BMzones_2 | -1.1665 | 0.4264 | -2.7355 | 0.0062 | 0.8125 | 0.6244 | 54321.44 |
| power100_2 | -0.5663 | 0.3077 | -1.8406 | 0.0657 | 0.8125 | 0.6244 | 54321.44 |
| lic_age15_2_Tr | 2.6676 | 1.4465 | 1.8442 | 0.0652 | 0.8125 | 0.6244 | 54321.44 |
| woman_2_Tr | 2.7791 | 1.0289 | 2.7009 | 0.0069 | 0.8125 | 0.6244 | 54321.44 |

Table 90: Some effects

| | lic_age<=15 | woman |
|---------|-------------|--------|
| coef | -0.6511 | 0.7690 |
| Z | -0.4700 | 0.9285 |
| p_value | 0.3192 | 0.1766 |

Altres cop la variable **parking_yes** no és en absolut rellevant en el model. No obstant això, la variable edat sí que ajuda a predir el percentatge de quilòmetres que els assegurats realitzaran per vies urbanes, mentre els anys de carnet, ajuden tant a l'anterior propòsit com a saber qui tindrà un accident amb major probabilitat, segons el model. Es realitzarà doncs altre cop el mateix procés que en la secció anterior. Destacar dels resultats que deixar únicament la variable edat fa que l' ATT_2 de **woman** quasi bé sigui significatiu, indicant que **lic_age15** és un distractor per a aquesta variable, tot i que el model que conté **lic_age15** explica més informació que el que únicament conté **age35**.

La conclusió doncs és similar a la realitzada per a la variable **Speed** d'aquesta mateixa secció. Disposar de totes dues variables seria el millor. Tanmateix, si es vol analitzar com canvia el comportament d'un assegurat en front un accident seria millor disposar de la variable edat, mentre si es vol donar més importància a predir correctament el percentatge de quilòmetres que es durant a terme per sobre de la velocitat els anys de llicència d'un assegurat són més importants.

Així com pel cas del dataset amb covariables canviants en el temps s'ha conclòs que disposar dels anys de carnet era més informatiu (menor AIC) per a 3 de les 4 variables dependents, en aquest cas no es pot concloure res perquè l'edat és millor per a predir els quilòmetre totals que realitzaran els assegurats així com el percentatge de quilòmetres recorreguts durant la nit, mentre els anys de llicència són millor per a les altres dues variables "Urban" i "Speed". Així doncs, la variable "Total" és la única que ha canviat en aquest sentit al canviar de conjunt de dades.

Table 91: TWFE for Urban variable with cst reduced covariates without lic_age15

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|------------|----------|------------|---------|----------|--------|--------|----------|
| Tr | -1.5902 | 0.6971 | -2.2810 | 0.0226 | 0.8116 | 0.6227 | 54327.12 |
| age35_2 | -1.1238 | 0.4136 | -2.7170 | 0.0066 | 0.8116 | 0.6227 | 54327.12 |
| BMzones_2 | -1.2390 | 0.4257 | -2.9102 | 0.0036 | 0.8116 | 0.6227 | 54327.12 |
| power100_2 | -0.6470 | 0.3064 | -2.1114 | 0.0348 | 0.8116 | 0.6227 | 54327.12 |
| woman_2_Tr | 2.8462 | 1.0283 | 2.7680 | 0.0057 | 0.8116 | 0.6227 | 54327.12 |

Table 92: Some effects

| | woman |
|---------|--------|
| coef | 1.2560 |
| Z | 1.5974 |
| p_value | 0.0551 |

Table 93: TWFE for Urban variable with cst reduced covariates without age35

| | Estimate | Std. Error | t value | Pr(> t) | R2 | R2Adj | AIC |
|----------------|----------|------------|---------|----------|--------|--------|----------|
| Tr | -2.0296 | 0.7287 | -2.7851 | 0.0054 | 0.8126 | 0.6246 | 54321.51 |
| lic_age15_2 | -1.5630 | 0.4136 | -3.7788 | 0.0002 | 0.8126 | 0.6246 | 54321.51 |
| BMzones_2 | -1.1668 | 0.4265 | -2.7359 | 0.0062 | 0.8126 | 0.6246 | 54321.51 |
| power100_2 | -0.5782 | 0.3076 | -1.8797 | 0.0602 | 0.8126 | 0.6246 | 54321.51 |
| lic_age15_2_Tr | 2.7186 | 1.4462 | 1.8799 | 0.0602 | 0.8126 | 0.6246 | 54321.51 |
| woman_2_Tr | 2.7924 | 1.0290 | 2.7138 | 0.0067 | 0.8126 | 0.6246 | 54321.51 |

Table 94: Some effects

| | lic_age15 | woman |
|---------|-----------|--------|
| coef | -0.8739 | 0.7629 |
| Z | -0.6348 | 0.9210 |
| p_value | 0.2628 | 0.1785 |

3.3.2 Models semi-paramètrics (IWE)

A la pràctica l' \widehat{ATT}_2 que calcula aquest model, és la mitjana ponderada de cada conjunt de dades amb els mateixos valors per les covariables X_i . Són doncs bàsicament una mitjana mostral ponderada.

Només té sentit aplicar aquest tipus de models amb les solucions que en la secció anterior (TWFE amb covariables constants en el temps) han retornat interaccions dins del millor model seleccionat. Concretament, per a aplicar el mètode s'han de seleccionar els individus que compleixen les interaccions trobades com a rellevants. Amb rellevants ens referim a interaccions que han sortit en el model final, sense la necessitat que aquestes hagin hagut de ser significatives o que el seu efecte de tractament hagi resultat significatiu. Per a aquests usuaris s'ha d'aplicar el logit o altres mètodes com un tree o un random forest. S'ha de tenir en compte que a vegades ens podem quedar sense usuaris amb $Tr = 1$ al filtrar amb les variables d'interès (interaccions). En aquests cassos no s'hi pot fer res, simplement no es podrà aplicar el mètode, és precisament una de les debilitats d'aquest.

També s'ha de tenir en compte que a l'aplicar el mètode s'ha d'assumir que els períodes $t = 1$ i $t = 2$ són independents. Tanmateix, aquesta assumpció podria ser criticable perquè s'està tractant en tot cas amb els mateixos individus a $t = 1$ i $t = 2$, encara que existeixi 1 any pel mig.

A continuació es descriuran els resultats que s'obtenen en cada cas, tanmateix, la conclusió general és que els models (logit, tree i random forest) no donen lloc a cap resultat significatiu. De fet, si es realitzen els summary d'aquests es nota que tots els coeficients mostren pvalors no significatius. Altre cop, això és degut a que aquesta tipologia d'estimadors, els semi-paramètrics i els no paramètrics no serveixen per dades desbalancejades, com amb les que s'està tractant.

3.3.2.1 Variable Total amb l'efecte de covariables La variable `Total` ha mostrat en la secció anterior que tant l'edat com el fet de ser dona eren interaccions rellevants en el model 72. Així doncs, es filtrarà per totes dues. En primer lloc, es farà l'anàlisi per cadascuna d'elles de forma separada perquè es recorda que si es filtra massa, ens podem arribar a quedar sense representació.

Com a recordatori de la última secció cal comentar que la variable `age35` i `woman` havien sortit rellevants en el model, però només `woman` havia resultat en tenir un ATT_2 significatiu i positiu, mentre l'efecte del tractament del model era 1480, sent aquest significatiu. Tanmateix, en aquest cas cap de les combinacions de filtratges ha resultat en un efecte del tractament significatiu, sent la filtració per la variable `woman` la que s'ha trobat més aprop d'estar-ho.

3.3.2.1.1 Filtratge per “age35”

```
## Assegurats amb accidents que queden a la mostra: 56
```

```
## Millor ntree: 50
```

Table 95: IWE for Total variable filtering by age35

| | ATT_2 | sd | t_stat | p_value | CI_lower | CI_upper |
|-----------------------|------------|----------|---------|---------|-----------|----------|
| Logit 2-tails | -28606.98 | 418185.9 | -0.0684 | 0.9455 | -849789.3 | 792575.3 |
| Tree 2-tails | -108565.59 | 415732.7 | -0.2611 | 0.7941 | -924930.6 | 707799.5 |
| Random Forest 2-tails | -71539.46 | 418998.0 | -0.1707 | 0.8645 | -894316.5 | 751237.6 |
| Logit 1-tail | -28606.98 | 418185.9 | -0.0684 | 0.4727 | -Inf | 660244.8 |
| Tree 1-tail | -108565.59 | 415732.7 | -0.2611 | 0.3970 | -Inf | 576245.2 |
| Random Forest 1-tail | -71539.46 | 418998.0 | -0.1707 | 0.4322 | -Inf | 618650.0 |

3.3.2.1.2 Filtratge per “woman”

```
## Assegurats amb accidents que queden a la mostra: 155
```

```
## Millor ntree: 100
```

Table 96: IWE for Total variable filtering by woman

| | ATT_2 | sd | t_stat | p_value | CI_lower | CI_upper |
|-----------------------|-----------|----------|--------|---------|-----------|----------|
| Logit 2-tails | 1019650.2 | 822555.3 | 1.2396 | 0.2153 | -593500.3 | 2632801 |
| Tree 2-tails | 1057751.4 | 821880.8 | 1.2870 | 0.1982 | -554076.5 | 2669579 |
| Random Forest 2-tails | 951941.9 | 823289.7 | 1.1563 | 0.2477 | -662648.9 | 2566533 |
| Logit 1-tail | 1019650.2 | 822555.3 | 1.2396 | 0.1076 | -333956.9 | Inf |
| Tree 1-tail | 1057751.4 | 821880.8 | 1.2870 | 0.0991 | -294745.9 | Inf |
| Random Forest 1-tail | 951941.9 | 823289.7 | 1.1563 | 0.1239 | -402873.8 | Inf |

3.3.2.1.3 Filtratge per “age35” i “woman”

```
## Assegurats amb accidents que queden a la mostra: 24
```

```
## Millor ntree: 50
```

Table 97: IWE for Total variable filtering by age35 and woman

| | ATT_2 | sd | t_stat | p_value | CI_lower | CI_upper |
|-----------------------|----------|----------|--------|---------|-----------|----------|
| Logit 2-tails | 286665.4 | 276699.3 | 1.0360 | 0.3010 | -257651.5 | 830982.4 |
| Tree 2-tails | 306679.4 | 275576.4 | 1.1129 | 0.2666 | -235428.6 | 848787.5 |
| Random Forest 2-tails | 288311.5 | 272962.4 | 1.0562 | 0.2916 | -248654.2 | 825277.3 |
| Logit 1-tail | 286665.4 | 276699.3 | 1.0360 | 0.1505 | -169745.6 | Inf |
| Tree 1-tail | 306679.4 | 275576.4 | 1.1129 | 0.1333 | -147879.5 | Inf |
| Random Forest 1-tail | 288311.5 | 272962.4 | 1.0562 | 0.1458 | -161935.5 | Inf |

3.3.2.2 Variable Night amb l'efecte de covariables Per a aquesta variable dependent els resultats obtinguts amb el model TWFE mostraven que l' ATT_2 del model era no significatiu i molt proper a 0, mentre només 1 variable, **power100**, surtia com a rellevant en el model, mostrant aquesta un efecte del tractament no significatiu. Els resultats per a aquest nou model han canviat de forma significativa en obtenir un ATT_2 positiu i significatiu al filtrar per la variable **power100**.

3.3.2.2.1 Filtratge per “power100”

```
## Assegurats amb accidents que queden a la mostra: 194
```

```
## Millor ntree: 150
```

Table 98: IWE for Night variable filtering by power100

| | ATT_2 | sd | t_stat | p_value | CI_lower | CI_upper |
|-----------------------|---------|----------|--------|---------|----------|----------|
| Logit 2-tails | 2296911 | 905241.7 | 2.5373 | 0.0112 | 521768.6 | 4072053 |
| Tree 2-tails | 2286666 | 904378.4 | 2.5284 | 0.0115 | 513216.7 | 4060115 |
| Random Forest 2-tails | 2276111 | 905505.9 | 2.5136 | 0.0120 | 500451.1 | 4051771 |
| Logit 1-tail | 2296911 | 905241.7 | 2.5373 | 0.0056 | 807341.9 | Inf |
| Tree 1-tail | 2286666 | 904378.4 | 2.5284 | 0.0058 | 798517.7 | Inf |
| Random Forest 1-tail | 2276111 | 905505.9 | 2.5136 | 0.0060 | 786107.7 | Inf |

3.3.2.3 Variable Speed amb l'efecte de covariables En aquest cas el model TWFE amb variable constants en el temps ha deixat veure que l'efecte del tractament era no significatiu però que 3 de les seves variables sí que presentaven una interacció amb un ATT_2 significatiu: **age35**, **lic_age15** i **power100**. El primer efecte era negatiu mentre els dos segons positius. Amb els models IWE realitzats a continuació es pot notar que filtrar únicament per 1 variable dona lloc a obtenir pvalors més propers a ser significatius. Concretament, filtrar per la variable **lic_age15** és el que ha donat lloc a resultats més llunyants del 0.05%, seguit per **power100** i finalment **age35**, la qual de fet presenta un efecte del tractament significatiu i negatiu, com anteriorment. Aquest efecte és únicament significatiu amb els resultats obtinguts a 1 cua. De fet, en tots els cassos podem notar com la significació (pvalor) es divideix aproximadament per 2 en passar dels resultats de 2 cues a 1 cua. La última sentència afirma que és necessari treballar amb aquest tipus d'inferència.

3.3.2.3.1 Filtratge per “age35”

```
## Assegurats amb accidents que queden a la mostra: 56
```

```
## Millor ntree: 50
```

Table 99: IWE for Speed variable filtering by age35

| | ATT_2 | sd | t_stat | p_value | CI_lower | CI_upper |
|-----------------------|-----------|----------|---------|---------|-----------|-----------|
| Logit 2-tails | -1193.041 | 687.7481 | -1.7347 | 0.0833 | -2543.556 | 157.4745 |
| Tree 2-tails | -1333.620 | 683.7136 | -1.9506 | 0.0515 | -2676.213 | 8.9727 |
| Random Forest 2-tails | -1288.942 | 689.0837 | -1.8705 | 0.0619 | -2642.080 | 64.1959 |
| Logit 1-tail | -1193.041 | 687.7481 | -1.7347 | 0.0416 | -Inf | -60.1562 |
| Tree 1-tail | -1333.620 | 683.7136 | -1.9506 | 0.0258 | -Inf | -207.3813 |
| Random Forest 1-tail | -1288.942 | 689.0837 | -1.8705 | 0.0309 | -Inf | -153.8574 |

3.3.2.3.2 Filtratge per “lic_age15”

Assegurats amb accidents que queden a la mostra: 57

Millor ntree: 400

Table 100: IWE for Speed variable filtering by lic_age15

| | ATT_2 | sd | t_stat | p_value | CI_lower | CI_upper |
|-----------------------|----------|----------|--------|---------|-----------|----------|
| Logit 2-tails | 531.9211 | 656.0526 | 0.8108 | 0.4178 | -756.0804 | 1819.923 |
| Tree 2-tails | 560.4917 | 653.7891 | 0.8573 | 0.3916 | -723.0660 | 1844.049 |
| Random Forest 2-tails | 593.5365 | 656.3586 | 0.9043 | 0.3661 | -695.0658 | 1882.139 |
| Logit 1-tail | 531.9211 | 656.0526 | 0.8108 | 0.2089 | -548.5776 | Inf |
| Tree 1-tail | 560.4917 | 653.7891 | 0.8573 | 0.1958 | -516.2792 | Inf |
| Random Forest 1-tail | 593.5365 | 656.3586 | 0.9043 | 0.1831 | -487.4662 | Inf |

3.3.2.3.3 Filtratge per “power100”

Assegurats amb accidents que queden a la mostra: 194

Millor ntree: 150

Table 101: IWE for Speed variable filtering by power100

| | ATT_2 | sd | t_stat | p_value | CI_lower | CI_upper |
|-----------------------|----------|----------|--------|---------|-----------|----------|
| Logit 2-tails | 1871.395 | 1254.430 | 1.4918 | 0.1359 | -588.4914 | 4331.281 |
| Tree 2-tails | 1821.932 | 1253.234 | 1.4538 | 0.1461 | -635.6082 | 4279.472 |
| Random Forest 2-tails | 1884.082 | 1254.796 | 1.5015 | 0.1334 | -576.5221 | 4344.686 |
| Logit 1-tail | 1871.395 | 1254.430 | 1.4918 | 0.0679 | -192.7609 | Inf |
| Tree 1-tail | 1821.932 | 1253.234 | 1.4538 | 0.0731 | -240.2551 | Inf |
| Random Forest 1-tail | 1884.082 | 1254.796 | 1.5015 | 0.0667 | -180.6761 | Inf |

3.3.2.3.4 Filtratge per “age35” i “lic_age15”

Assegurats amb accidents que queden a la mostra: 24

Millor ntree: 350

Table 102: IWE for Speed variable filtering by age35 and lic_age15

| | ATT_2 | sd | t_stat | p_value | CI_lower | CI_upper |
|-----------------------|---------|----------|--------|---------|-----------|----------|
| Logit 2-tails | 34.2923 | 499.9276 | 0.0686 | 0.9454 | -949.0175 | 1017.602 |
| Tree 2-tails | 49.2177 | 499.0116 | 0.0986 | 0.9215 | -932.2904 | 1030.726 |
| Random Forest 2-tails | 38.2617 | 496.8069 | 0.0770 | 0.9387 | -938.9100 | 1015.433 |
| Logit 1-tail | 34.2923 | 499.9276 | 0.0686 | 0.4727 | -790.2424 | Inf |
| Tree 1-tail | 49.2177 | 499.0116 | 0.0986 | 0.4607 | -773.8062 | Inf |
| Random Forest 1-tail | 38.2617 | 496.8069 | 0.0770 | 0.4693 | -781.1260 | Inf |

3.3.2.3.5 Filtratge per “age35” i “power100”

Assegurats amb accidents que queden a la mostra: 31

Millor ntree: 400

Table 103: IWE for Speed variable filtering by age35 and power100

| | ATT_2 | sd | t_stat | p_value | CI_lower | CI_upper |
|-----------------------|----------|----------|---------|---------|-----------|-----------|
| Logit 2-tails | -47.4329 | 539.0439 | -0.0880 | 0.9299 | -1106.950 | 1012.0840 |
| Tree 2-tails | -2.6267 | 532.0388 | -0.0049 | 0.9961 | -1048.375 | 1043.1213 |
| Random Forest 2-tails | -51.9238 | 534.8453 | -0.0971 | 0.9227 | -1103.188 | 999.3406 |
| Logit 1-tail | -47.4329 | 539.0439 | -0.0880 | 0.4650 | -Inf | 841.1478 |
| Tree 1-tail | -2.6267 | 532.0388 | -0.0049 | 0.4980 | -Inf | 874.4065 |
| Random Forest 1-tail | -51.9238 | 534.8453 | -0.0971 | 0.4614 | -Inf | 829.7358 |

3.3.2.3.6 Filtratge per “lic_age15” i “power100”

Assegurats amb accidents que queden a la mostra: 36

Millor ntree: 150

Table 104: IWE for Speed variable filtering by lic_age15 and power100

| | ATT_2 | sd | t_stat | p_value | CI_lower | CI_upper |
|-----------------------|----------|----------|--------|---------|-----------|----------|
| Logit 2-tails | 417.3783 | 526.8968 | 0.7921 | 0.4286 | -617.7770 | 1452.534 |
| Tree 2-tails | 344.1550 | 523.8800 | 0.6569 | 0.5115 | -685.0735 | 1373.383 |
| Random Forest 2-tails | 411.9911 | 523.9102 | 0.7864 | 0.4320 | -617.2968 | 1441.279 |
| Logit 1-tail | 417.3783 | 526.8968 | 0.7921 | 0.2143 | -450.8669 | Inf |
| Tree 1-tail | 344.1550 | 523.8800 | 0.6569 | 0.2558 | -519.1191 | Inf |
| Random Forest 1-tail | 411.9911 | 523.9102 | 0.7864 | 0.2160 | -451.3328 | Inf |

3.3.2.3.7 Filtratge per “age35”, “lic_age15” i “power100”

Assegurats amb accidents que queden a la mostra: 13

Millor ntree: 250

Table 105: IWE for Speed variable filtering by age35, lic_age15 and power100

| | ATT_2 | sd | t_stat | p_value | CI_lower | CI_upper |
|-----------------------|----------|----------|--------|---------|-----------|----------|
| Logit 2-tails | 258.9982 | 441.5572 | 0.5866 | 0.5580 | -610.6479 | 1128.644 |
| Tree 2-tails | 245.7872 | 438.5182 | 0.5605 | 0.5756 | -617.8737 | 1109.448 |
| Random Forest 2-tails | 247.7198 | 438.7857 | 0.5646 | 0.5729 | -616.4679 | 1111.907 |
| Logit 1-tail | 258.9982 | 441.5572 | 0.5866 | 0.2790 | -470.0001 | Inf |
| Tree 1-tail | 245.7872 | 438.5182 | 0.5605 | 0.2878 | -478.1938 | Inf |
| Random Forest 1-tail | 247.7198 | 438.7857 | 0.5646 | 0.2864 | -476.7029 | Inf |

3.3.2.4 Variable Urban amb l'efecte de covariables Finalment per a la variable Urban anteriorment s'havia obtingut un efecte del tractament de -2 i significatiu mentre les variables `lic_age15` i `woman` són les que es presentaven com a rellevants en el model, totes dues amb un ATT_2 no significatiu. Amb els models semiparamètrics l'efecte del tractament passa a ser no significatiu en tots els cassos, sent el filtratge de la variable `woman` el que es troba més proper de ser-ho.

En general notem que el model semiparamètric dona lloc a resultats generalment no significatius. Pel que a fa a les diferents maneres de construir el model, no es pot concloure quina és la millor, en mostrar pvalors majors o inferiors respecte als altres mètodes en funció del filtratge. Pel cas de la variable `Urban`, el model construït a partir d'un random forest és el que dona lloc a pvalors menors, però això no conclou res ja que per exemple per a la variable `Total` era el tree el que donava lloc a menors pvalors, mentre pel cas de `Night` els resultats varien de forma significativa en funció del filtratge.

3.3.2.4.1 Filtratge per “lic_age15”

Assegurats amb accidents que queden a la mostra: 57

Millor ntree: 400

Table 106: IWE for Urban variable filtering by lic_age15

| | ATT_2 | sd | t_stat | p_value | CI_lower | CI_upper |
|-----------------------|----------|----------|--------|---------|-----------|----------|
| Logit 2-tails | 531.9211 | 656.0526 | 0.8108 | 0.4178 | -756.0804 | 1819.923 |
| Tree 2-tails | 560.4917 | 653.7891 | 0.8573 | 0.3916 | -723.0660 | 1844.049 |
| Random Forest 2-tails | 593.5365 | 656.3586 | 0.9043 | 0.3661 | -695.0658 | 1882.139 |
| Logit 1-tail | 531.9211 | 656.0526 | 0.8108 | 0.2089 | -548.5776 | Inf |
| Tree 1-tail | 560.4917 | 653.7891 | 0.8573 | 0.1958 | -516.2792 | Inf |
| Random Forest 1-tail | 593.5365 | 656.3586 | 0.9043 | 0.1831 | -487.4662 | Inf |

3.3.2.4.2 Filtratge per “woman”

Assegurats amb accidents que queden a la mostra: 155

Millor ntree: 100

Table 107: IWE for Urban variable filtering by woman

| | ATT_2 | sd | t_stat | p_value | CI_lower | CI_upper |
|-----------------------|----------|----------|--------|---------|-----------|----------|
| Logit 2-tails | 1602.281 | 1223.633 | 1.3094 | 0.1905 | -797.4421 | 4002.004 |
| Tree 2-tails | 1529.694 | 1222.630 | 1.2512 | 0.2110 | -868.0616 | 3927.449 |
| Random Forest 2-tails | 1584.549 | 1224.726 | 1.2938 | 0.1959 | -817.3166 | 3986.414 |
| Logit 1-tail | 1602.281 | 1223.633 | 1.3094 | 0.0953 | -411.3454 | Inf |
| Tree 1-tail | 1529.694 | 1222.630 | 1.2512 | 0.1055 | -482.2814 | Inf |
| Random Forest 1-tail | 1584.549 | 1224.726 | 1.2938 | 0.0979 | -430.8751 | Inf |

3.3.2.4.3 Filtratge per “lic_age15” i “woman”

Assegurats amb accidents que queden a la mostra: 28

Millor ntree: 400

Table 108: IWE for Urban variable filtering by lic_age15 and woman

| | ATT_2 | sd | t_stat | p_value | CI_lower | CI_upper |
|-----------------------|----------|----------|--------|---------|-----------|----------|
| Logit 2-tails | 177.0829 | 520.0664 | 0.3405 | 0.7336 | -845.1192 | 1199.285 |
| Tree 2-tails | 182.1251 | 519.0999 | 0.3508 | 0.7259 | -838.1772 | 1202.428 |
| Random Forest 2-tails | 213.8945 | 520.0784 | 0.4113 | 0.6811 | -808.3311 | 1236.120 |
| Logit 1-tail | 177.0829 | 520.0664 | 0.3405 | 0.3668 | -680.2058 | Inf |
| Tree 1-tail | 182.1251 | 519.0999 | 0.3508 | 0.3629 | -673.5703 | Inf |
| Random Forest 1-tail | 213.8945 | 520.0784 | 0.4113 | 0.3405 | -643.4140 | Inf |

3.3.3 Models semi-paramètrics (IWE) amb dades balancejades

La conclusió de la secció anterior ha estat que no es poden treure resultats significatius principalment perquè les dades es troben molt desbalancejades. Així doncs, a continuació es mirarà què passaria si es treballés amb dades balancejades. S'utilitzarà una tècnica de submostreig per a obtenir el nou conjunt de dades. És a dir, s'ha reduït el nombre de mostres de la classe majoritària (el grup control) per igualar-ho al nombre de mostres de la classe minoritària (el grup tractament). Durant aquest submostreig s'ha de tenir en compte que si s'elimina una persona, s'ha d'eliminar tant la seva fila de $t=1$ com la de $t=2$.

Durant el procediment, ja no es mirarà si queden registres amb $Tr=1$ a la mostra, perquè anteriorment s'ha vist que sí i no s'ha eliminat cap registre d'aquesta tipologia durant el submostreig.

3.3.3.1 Variable Total amb l'efecte de covariables Anteriorment s'havia vist que la filtració per la variable **woman** era la que es trobava més propera a mostrar resultats significatius. En aquest cas, aquesta filtració ha donat lloc a un warning, perquè el model random forest únicament dona lloc a 2 possibles respostes per a la variables dependent. Indicant clarament que els models continuen sense diferenciar de forma significativa. Contràriament, en aquest cas la combinació de la filtració de totes dues variables **age35** i **woman** ha estat la que s'ha trobat més propera a mostrar resultats significatius.

Tot i les explicacions anteriors, es pot confirmar que els resultats milloren en comparació als models amb dades desbalancejades, sobretot pel que fa al tree. Anteriorment, al filtrar per **age35** s'observava que el model únicament donava lloc a 1 nodes (l'arrel). En aquest cas, s'han creat fins a 29 fulles, indicant que com a mínim en aquest cas la diferenciació és major. Per a les altres filtracions, tot i que no de forma tant significativa, les fulles de l'arbre també han augmentat.

3.3.3.1.1 Filtratge per “age35”

Millor ntree: 500

Table 109: IWE for Total variable filtering by age35

| | ATT_2 | sd | t_stat | p_value | CI_lower | CI_upper |
|-----------------------|----------|----------|--------|---------|-----------|----------|
| Logit 2-tails | 42382.61 | 120049.3 | 0.3530 | 0.7247 | -195456.8 | 280222.0 |
| Tree 2-tails | 45600.88 | 125944.8 | 0.3621 | 0.7180 | -203918.5 | 295120.3 |
| Random Forest 2-tails | 52672.00 | 114408.2 | 0.4604 | 0.6461 | -173991.2 | 279335.2 |
| Logit 1-tail | 42382.61 | 120049.3 | 0.3530 | 0.3624 | -156713.2 | Inf |
| Tree 1-tail | 45600.88 | 125944.8 | 0.3621 | 0.3590 | -163272.3 | Inf |
| Random Forest 1-tail | 52672.00 | 114408.2 | 0.4604 | 0.3231 | -137068.3 | Inf |

3.3.3.1.2 Filtratge per “woman”

Millor ntree: 150

Table 110: IWE for Total variable filtering by woman

| | ATT_2 | sd | t_stat | p_value | CI_lower | CI_upper |
|-----------------------|----------|----------|--------|---------|------------|----------|
| Logit 2-tails | 278107.6 | 195977.6 | 1.4191 | 0.1568 | -107411.09 | 663626.3 |
| Tree 2-tails | 287783.2 | 194995.5 | 1.4758 | 0.1409 | -95803.49 | 671369.8 |
| Random Forest 2-tails | 243927.4 | 194373.6 | 1.2549 | 0.2104 | -138436.02 | 626290.8 |
| Logit 1-tail | 278107.6 | 195977.6 | 1.4191 | 0.0784 | -45151.62 | Inf |
| Tree 1-tail | 287783.2 | 194995.5 | 1.4758 | 0.0705 | -33856.03 | Inf |
| Random Forest 1-tail | 243927.4 | 194373.6 | 1.2549 | 0.1052 | -76686.11 | Inf |

3.3.3.1.3 Filtratge per “age35” i “woman”

Millor ntree: 350

Table 111: IWE for Total variable filtering by age35 and woman

| | ATT_2 | sd | t_stat | p_value | CI_lower | CI_upper |
|-----------------------|-----------|----------|--------|---------|------------|----------|
| Logit 2-tails | 103003.90 | 67174.66 | 1.5334 | 0.1308 | -31563.115 | 237570.9 |
| Tree 2-tails | 109117.85 | 67066.16 | 1.6270 | 0.1093 | -25231.807 | 243467.5 |
| Random Forest 2-tails | 93191.35 | 62418.60 | 1.4930 | 0.1410 | -31848.123 | 218230.8 |
| Logit 1-tail | 103003.90 | 67174.66 | 1.5334 | 0.0654 | -9347.219 | Inf |
| Tree 1-tail | 109117.85 | 67066.16 | 1.6270 | 0.0547 | -3051.794 | Inf |
| Random Forest 1-tail | 93191.35 | 62418.60 | 1.4930 | 0.0705 | -11205.143 | Inf |

3.3.3.2 Variable Night amb l'efecte de covariables En aquest cas, els resultats mostrats s'assemblen més als que s'obtenen amb el model TWFE als que s'han obtingut amb l'IWE amb dades desbalancejades, ja que els pvalors indiquen que els resultats són no significatius. Això dona lloc a pensar que els resultats anteriorment obtinguts amb l'IWE no són fiables. Finalment, comentar que altre cop el model random forest ha tornat a donar com a resposta menys de 5 valors, dificultant la regressió que es requereix fer.

3.3.3.2.1 Filtratge per “power100”

Millor ntree: 50

Table 112: IWE for Night variable filtering by power100

| | ATT_2 | sd | t_stat | p_value | CI_lower | CI_upper |
|-----------------------|----------|----------|--------|---------|----------|----------|
| Logit 2-tails | 521106.7 | 211687.2 | 2.4617 | 0.0143 | 104912.4 | 937301.0 |
| Tree 2-tails | 558447.9 | 209003.6 | 2.6720 | 0.0079 | 147530.0 | 969365.9 |
| Random Forest 2-tails | 526354.3 | 208085.5 | 2.5295 | 0.0118 | 117241.3 | 935467.2 |
| Logit 1-tail | 521106.7 | 211687.2 | 2.4617 | 0.0071 | 172080.9 | Inf |
| Tree 1-tail | 558447.9 | 209003.6 | 2.6720 | 0.0039 | 213847.0 | Inf |
| Random Forest 1-tail | 526354.3 | 208085.5 | 2.5295 | 0.0059 | 183267.1 | Inf |

3.3.3.3 Variable Speed amb l’efecte de covariables En aquest cas els resultats tornen a ser similars als obtinguts anteriorment, en notar que la variable **age35** mostra un efecte del tractament significatiu i negatiu, com també passava amb el model TWFE. A més a més, la variable **power100** mostra un ATT_2 amb un pvalor també proper al 0.05. Es confirma doncs, que tot i que aquests models semi-paramètrics en general no funcionin molt bé, la relació negativa de l’edat amb la velocitat després de tenir un accident és molt clarament significativa, així com la relació que mostra el fet de tenir un cotxe petit, en aquest cas de caràcter positiu,

3.3.3.3.1 Filtratge per “age35”

Millor ntree: 500

Table 113: IWE for Speed variable filtering by age35

| | ATT_2 | sd | t_stat | p_value | CI_lower | CI_upper |
|-----------------------|-----------|----------|---------|---------|-----------|----------|
| Logit 2-tails | -37.3991 | 197.4276 | -0.1894 | 0.8501 | -428.5388 | 353.7405 |
| Tree 2-tails | -45.2565 | 207.1230 | -0.2185 | 0.8274 | -455.6046 | 365.0916 |
| Random Forest 2-tails | -108.6458 | 188.1504 | -0.5774 | 0.5648 | -481.4057 | 264.1140 |
| Logit 1-tail | -37.3991 | 197.4276 | -0.1894 | 0.4250 | -Inf | 290.0247 |
| Tree 1-tail | -45.2565 | 207.1230 | -0.2185 | 0.4137 | -Inf | 298.2467 |
| Random Forest 1-tail | -108.6458 | 188.1504 | -0.5774 | 0.2824 | -Inf | 203.3922 |

3.3.3.3.2 Filtratge per “lic_age15”

Millor ntree: 50

Table 114: IWE for Speed variable filtering by lic_age15

| | ATT_2 | sd | t_stat | p_value | CI_lower | CI_upper |
|-----------------------|----------|----------|--------|---------|-----------|----------|
| Logit 2-tails | 225.9339 | 165.8384 | 1.3624 | 0.1759 | -102.8213 | 554.6891 |
| Tree 2-tails | 266.8858 | 166.2610 | 1.6052 | 0.1114 | -62.7073 | 596.4789 |
| Random Forest 2-tails | 241.5665 | 149.3672 | 1.6173 | 0.1088 | -54.5366 | 537.6696 |
| Logit 1-tail | 225.9339 | 165.8384 | 1.3624 | 0.0880 | -49.2283 | Inf |

| | ATT_2 | sd | t_stat | p_value | CI_lower | CI_upper |
|----------------------|----------|----------|--------|---------|----------|----------|
| Tree 1-tail | 266.8858 | 166.2610 | 1.6052 | 0.0557 | -8.9777 | Inf |
| Random Forest 1-tail | 241.5665 | 149.3672 | 1.6173 | 0.0544 | -6.2665 | Inf |

3.3.3.3.3 Filtratge per “power100”

Millor ntree: 50

Table 115: IWE for Speed variable filtering by power100

| | ATT_2 | sd | t_stat | p_value | CI_lower | CI_upper |
|-----------------------|----------|----------|--------|---------|-----------|----------|
| Logit 2-tails | 393.7753 | 277.6274 | 1.4184 | 0.1569 | -152.0627 | 939.6133 |
| Tree 2-tails | 424.7615 | 274.1078 | 1.5496 | 0.1220 | -114.1566 | 963.6796 |
| Random Forest 2-tails | 387.7393 | 272.9037 | 1.4208 | 0.1562 | -148.8115 | 924.2901 |
| Logit 1-tail | 393.7753 | 277.6274 | 1.4184 | 0.0784 | -63.9712 | Inf |
| Tree 1-tail | 424.7615 | 274.1078 | 1.5496 | 0.0610 | -27.1820 | Inf |
| Random Forest 1-tail | 387.7393 | 272.9037 | 1.4208 | 0.0781 | -62.2189 | Inf |

3.3.3.3.4 Filtratge per “age35” i “lic_age15”

Millor ntree: 450

Table 116: IWE for Speed variable filtering by age35 and lic_age15

| | ATT_2 | sd | t_stat | p_value | CI_lower | CI_upper |
|-----------------------|----------|----------|--------|---------|-----------|----------|
| Logit 2-tails | 92.2057 | 147.0927 | 0.6269 | 0.5337 | -203.3881 | 387.7994 |
| Tree 2-tails | 159.1784 | 139.9971 | 1.1370 | 0.2611 | -122.1563 | 440.5130 |
| Random Forest 2-tails | 108.9513 | 127.0237 | 0.8577 | 0.3952 | -146.3123 | 364.2149 |
| Logit 1-tail | 92.2057 | 147.0927 | 0.6269 | 0.2668 | -154.4027 | Inf |
| Tree 1-tail | 159.1784 | 139.9971 | 1.1370 | 0.1305 | -75.5338 | Inf |
| Random Forest 1-tail | 108.9513 | 127.0237 | 0.8577 | 0.1976 | -104.0103 | Inf |

3.3.3.3.5 Filtratge per “age35” i “power100”

Millor ntree: 250

Table 117: IWE for Speed variable filtering by age35 and power100

| | ATT_2 | sd | t_stat | p_value | CI_lower | CI_upper |
|-----------------------|---------|----------|--------|---------|-----------|----------|
| Logit 2-tails | 32.5741 | 131.5043 | 0.2477 | 0.8051 | -229.9826 | 295.1308 |
| Tree 2-tails | 55.6357 | 126.3000 | 0.4405 | 0.6610 | -196.5303 | 307.8018 |
| Random Forest 2-tails | 44.9311 | 116.3179 | 0.3863 | 0.7005 | -187.3051 | 277.1673 |
| Logit 1-tail | 32.5741 | 131.5043 | 0.2477 | 0.4026 | -186.8106 | Inf |
| Tree 1-tail | 55.6357 | 126.3000 | 0.4405 | 0.3305 | -155.0668 | Inf |
| Random Forest 1-tail | 44.9311 | 116.3179 | 0.3863 | 0.3503 | -149.1186 | Inf |

3.3.3.3.6 Filtratge per “lic_age15” i “power100”

Millor ntree: 500

Table 118: IWE for Speed variable filtering by lic_age15 and power100

| | ATT_2 | sd | t_stat | p_value | CI_lower | CI_upper |
|-----------------------|----------|----------|--------|---------|-----------|----------|
| Logit 2-tails | 167.5269 | 143.6692 | 1.1661 | 0.2475 | -119.0123 | 454.0662 |
| Tree 2-tails | 184.5400 | 119.4147 | 1.5454 | 0.1268 | -53.6250 | 422.7050 |
| Random Forest 2-tails | 151.0480 | 107.2324 | 1.4086 | 0.1634 | -62.8204 | 364.9163 |
| Logit 1-tail | 167.5269 | 143.6692 | 1.1661 | 0.1238 | -71.9574 | Inf |
| Tree 1-tail | 184.5400 | 119.4147 | 1.5454 | 0.0634 | -14.5140 | Inf |
| Random Forest 1-tail | 151.0480 | 107.2324 | 1.4086 | 0.0817 | -27.6993 | Inf |

3.3.3.3.7 Filtratge per “age35”, “lic_age15” i “power100”

Millor ntree: 50

Table 119: IWE for Speed variable filtering by age35, lic_age15 and power100

| | ATT_2 | sd | t_stat | p_value | CI_lower | CI_upper |
|-----------------------|---------|----------|--------|---------|-----------|----------|
| Logit 2-tails | 43.6936 | 103.6236 | 0.4217 | 0.6761 | -167.3807 | 254.7679 |
| Tree 2-tails | 64.6542 | 96.9308 | 0.6670 | 0.5095 | -132.7875 | 262.0959 |
| Random Forest 2-tails | 59.8138 | 88.0189 | 0.6796 | 0.5017 | -119.4747 | 239.1023 |
| Logit 1-tail | 43.6936 | 103.6236 | 0.4217 | 0.3380 | -131.8332 | Inf |
| Tree 1-tail | 64.6542 | 96.9308 | 0.6670 | 0.2548 | -99.5359 | Inf |
| Random Forest 1-tail | 59.8138 | 88.0189 | 0.6796 | 0.2508 | -89.2803 | Inf |

3.3.3.4 Variable Urban amb l’efecte de covariables Altre cop els resultats per a tots els filtratges són no significatius per a les dues variables provades, així com per a la seva significació. En general, es pot notar que filtrar per més d’1 variable dona lloc a obtenir una significació menor. Això es podria explicar perquè com més filtratge es realitza, menys mostres queden al conjunt de dades i més difícil és que els models captin el comportament darrere les dades.

3.3.3.4.1 Filtratge per “lic_age15”

Millor ntree: 50

Table 120: IWE for Urban variable filtering by lic_age15

| | ATT_2 | sd | t_stat | p_value | CI_lower | CI_upper |
|-----------------------|----------|----------|--------|---------|-----------|----------|
| Logit 2-tails | 225.9339 | 165.8384 | 1.3624 | 0.1759 | -102.8213 | 554.6891 |
| Tree 2-tails | 266.8858 | 166.2610 | 1.6052 | 0.1114 | -62.7073 | 596.4789 |
| Random Forest 2-tails | 241.5665 | 149.3672 | 1.6173 | 0.1088 | -54.5366 | 537.6696 |
| Logit 1-tail | 225.9339 | 165.8384 | 1.3624 | 0.0880 | -49.2283 | Inf |

| | ATT_2 | sd | t_stat | p_value | CI_lower | CI_upper |
|----------------------|----------|----------|--------|---------|----------|----------|
| Tree 1-tail | 266.8858 | 166.2610 | 1.6052 | 0.0557 | -8.9777 | Inf |
| Random Forest 1-tail | 241.5665 | 149.3672 | 1.6173 | 0.0544 | -6.2665 | Inf |

3.3.3.4.2 Filtratge per “woman”

Millor ntree: 150

Table 121: IWE for Urban variable filtering by woman

| | ATT_2 | sd | t_stat | p_value | CI_lower | CI_upper |
|-----------------------|----------|----------|--------|---------|-----------|-----------|
| Logit 2-tails | 442.0868 | 298.6881 | 1.4801 | 0.1398 | -145.4795 | 1029.6530 |
| Tree 2-tails | 419.3769 | 297.1912 | 1.4111 | 0.1591 | -165.2447 | 1003.9985 |
| Random Forest 2-tails | 408.5174 | 296.2434 | 1.3790 | 0.1688 | -174.2398 | 991.2747 |
| Logit 1-tail | 442.0868 | 298.6881 | 1.4801 | 0.0699 | -50.5903 | Inf |
| Tree 1-tail | 419.3769 | 297.1912 | 1.4111 | 0.0796 | -70.8310 | Inf |
| Random Forest 1-tail | 408.5174 | 296.2434 | 1.3790 | 0.0844 | -80.1273 | Inf |

3.3.3.4.3 Filtratge per “lic_age15” i “woman”

Millor ntree: 400

Table 122: IWE for Urban variable filtering by lic_age15 and woman

| | ATT_2 | sd | t_stat | p_value | CI_lower | CI_upper |
|-----------------------|----------|----------|--------|---------|-----------|----------|
| Logit 2-tails | 86.7743 | 135.2057 | 0.6418 | 0.5234 | -183.5862 | 357.1348 |
| Tree 2-tails | 143.5892 | 138.0187 | 1.0404 | 0.3023 | -132.3962 | 419.5747 |
| Random Forest 2-tails | 130.9522 | 126.1088 | 1.0384 | 0.3032 | -121.2180 | 383.1224 |
| Logit 1-tail | 86.7743 | 135.2057 | 0.6418 | 0.2617 | -139.0488 | Inf |
| Tree 1-tail | 143.5892 | 138.0187 | 1.0404 | 0.1511 | -86.9323 | Inf |
| Random Forest 1-tail | 130.9522 | 126.1088 | 1.0384 | 0.1516 | -79.6772 | Inf |