

NEPA_Delays_RMD

Katie Murenbeeld

2022-10-21

Introduction

Intro stuff

Load Libraries and Packages

```
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.0 --
## v ggplot2 3.3.3      v purrr   0.3.4
## v tibble  3.1.0      v dplyr  1.0.5
## v tidyr   1.1.3      v stringr 1.4.0
## v readr   1.4.0      v forcats 0.5.1

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()

library(forcats)
library(survminer)

## Loading required package: ggpubr

library(survival)
library(lubridate)

##
## Attaching package: 'lubridate'

## The following objects are masked from 'package:base':
##
##   date, intersect, setdiff, union

library(reshape2)

##
## Attaching package: 'reshape2'

## The following object is masked from 'package:tidyr':
##
##   smiths

library(ggplot2)
library(ggpubr)
library(fastDummies)
library(chron)
```

```
##
## Attaching package: 'chron'

## The following objects are masked from 'package:lubridate':
##
##      days, hours, minutes, seconds, years

library(vtable)

## Loading required package: kableExtra

##
## Attaching package: 'kableExtra'

## The following object is masked from 'package:dplyr':
##
##      group_rows

# Set your working directory to wherever you placed your processed FS-PALS dataset.

setwd("/Users/kathrynmurenbeeld/Desktop/NEPA_DELAYS/R/NEPA_Delays")

# Set up to see more rows of output if desired

options(max.print = 10000)
```

Load the Data

```
## Load the data for survival analysis

df_fin <- read.csv("/Users/kathrynmurenbeeld/Desktop/NEPA_DELAYS/Data/df_NEPA_2009_2018_survial_analysis")

# Need to convert DECISION TYPE, LITIGATED? and APPEALED? to factors
# Need to convert event variables (NEPA_COMP2, PROJ_COMP, PROJ_INIT, and PROJ_AWARDED) from true, false
# I think all of the duration variables (e.g ASSESSMENT_TIME) should be good, need to be numeric

df_fin$DECISION.TYPE <- as.factor(df_fin$DECISION.TYPE)
df_fin$LITIGATED. <- as.factor(df_fin$LITIGATED.)
df_fin$APPEALED.OR.OBJECTED. <- as.factor(df_fin$APPEALED.OR.OBJECTED.)

df_fin$NEPA_COMP2 <- as.integer(as.logical(df_fin$NEPA_COMP2))
df_fin$PROJ_COMP <- as.integer(as.logical(df_fin$PROJ_COMP))
df_fin$PROJ_INIT <- as.integer(as.logical(df_fin$PROJ_INIT))
df_fin$PROJ_AWARDED <- as.integer(as.logical(df_fin$PROJ_AWARDED))

# Use the survfit function from the survival package to
# calculate the KM estimate

fit_assess_all <- survfit(Surv(ASSESSMENT_TIME, NEPA_COMP2) ~ 1, data = df_fin)
#fit_assess_all
#summary(fit_assess_all)

fit_comp_all <- survfit(Surv(IMPLEMENTATION_TIME, PROJ_COMP) ~ 1, data = df_fin)
#fit_comp_all
#summary(fit_comp_all)
```

```

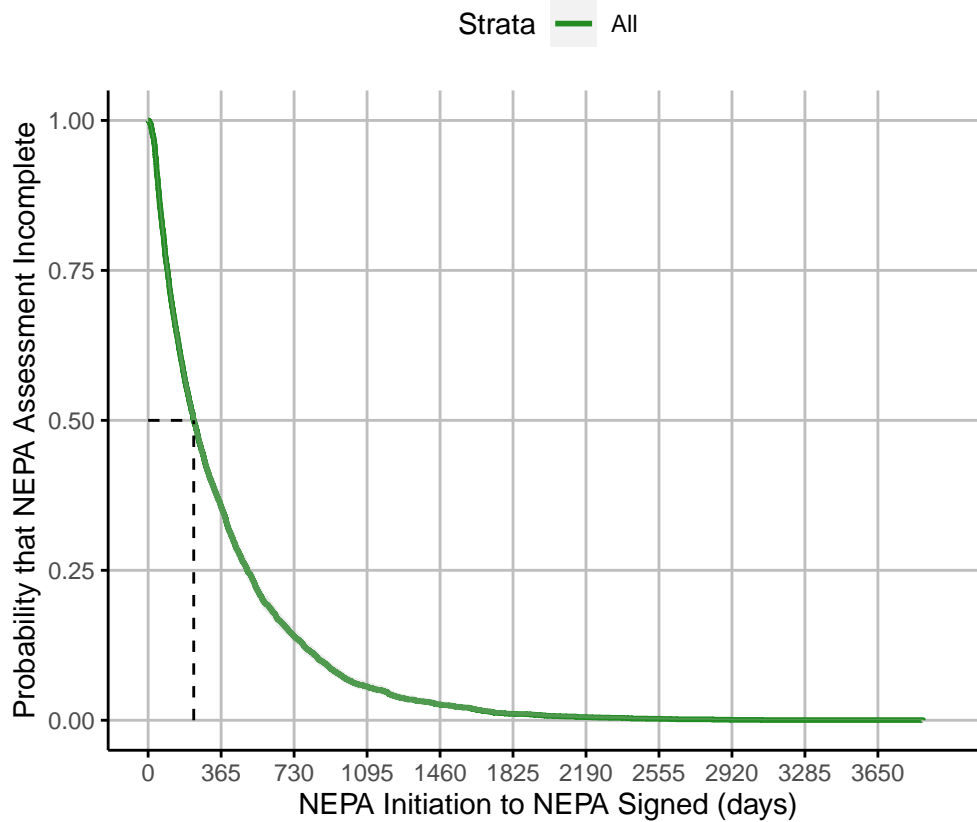
fit_init_all <- survfit(Surv(NEPA_TO_FIRST_ACT_MAX, PROJ_INIT) ~ 1, data = df_fin)
#fit_init_all
#summary(fit_init_all)

fit_award_all <- survfit(Surv(NEPA_TO_CONTRACT_MAX, PROJ_AWARDED) ~ 1, data = df_fin)
#fit_award_all
#summary(fit_award_all)

## ggsurvplot will create survival curves from the KM model created above.

km_fit_assess_all <- ggsurvplot(fit_assess_all,
  conf.int = TRUE,
  risk.table = FALSE,
  risk.table.col = "strata",
  surv.median.line = "hv",
  break.time.by = 365,
  risk.table.y.text=FALSE,
  censor = FALSE,
  ylab = "Probability that NEPA Assessment Incomplete",
  xlab = "NEPA Initiation to NEPA Signed (days)",
  palette = c("forestgreen"),
  surv.plot.height = 1,
  ggtheme = theme(aspect.ratio = 0.75,
    axis.line = element_line(colour = "black"),
    panel.grid.major = element_line(colour = "grey"),
    panel.border = element_blank(),
    panel.background = element_blank()),
  tables.theme = theme(aspect.ratio = 0.06)
)
print(km_fit_assess_all)

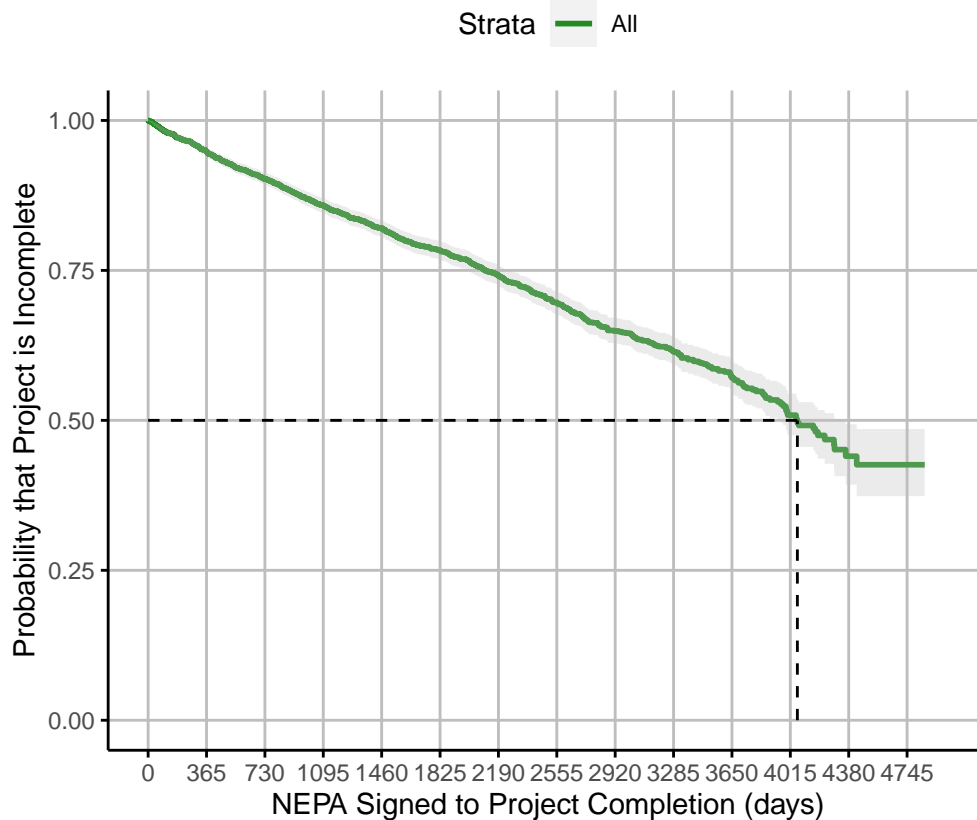
```



```
#ggsave( "/Users/kathrynmurenbeeld/Desktop/Survival_Analysis/WRITING/FIGS/km_fit_assess_all.pdf", print
```

```
## ggsurvplot will create survival curves from the KM model created above.
```

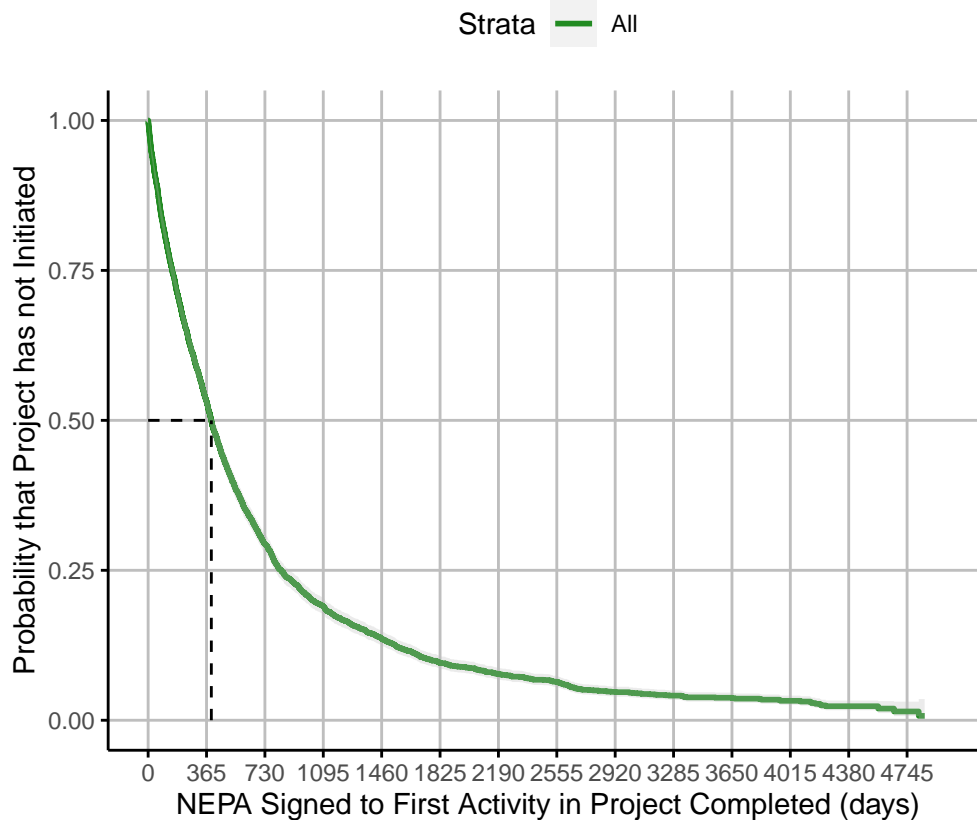
```
km_fit_comp_all <- ggsurvplot(fit_comp_all,
  conf.int = TRUE,
  risk.table = FALSE,
  risk.table.col = "strata",
  surv.median.line = "hv",
  break.time.by = 365,
  risk.table.y.text=FALSE,
  censor = FALSE,
  ylab = "Probability that Project is Incomplete",
  xlab = "NEPA Signed to Project Completion (days)",
  palette = c("forestgreen"),
  surv.plot.height = 1,
  ggtheme = theme(aspect.ratio = 0.75,
    axis.line = element_line(colour = "black"),
    panel.grid.major = element_line(colour = "grey"),
    panel.border = element_blank(),
    panel.background = element_blank(),
    tables.theme = theme(aspect.ratio = 0.06)
  )
print(km_fit_comp_all)
```



```
#ggsave( "/Users/kathrynmurenbeeld/Desktop/Survival_Analysis/WRITING/FIGS/km_fit_comp_all.pdf", print(k
```

```
## ggsurvplot will create survival curves from the KM model created above.
```

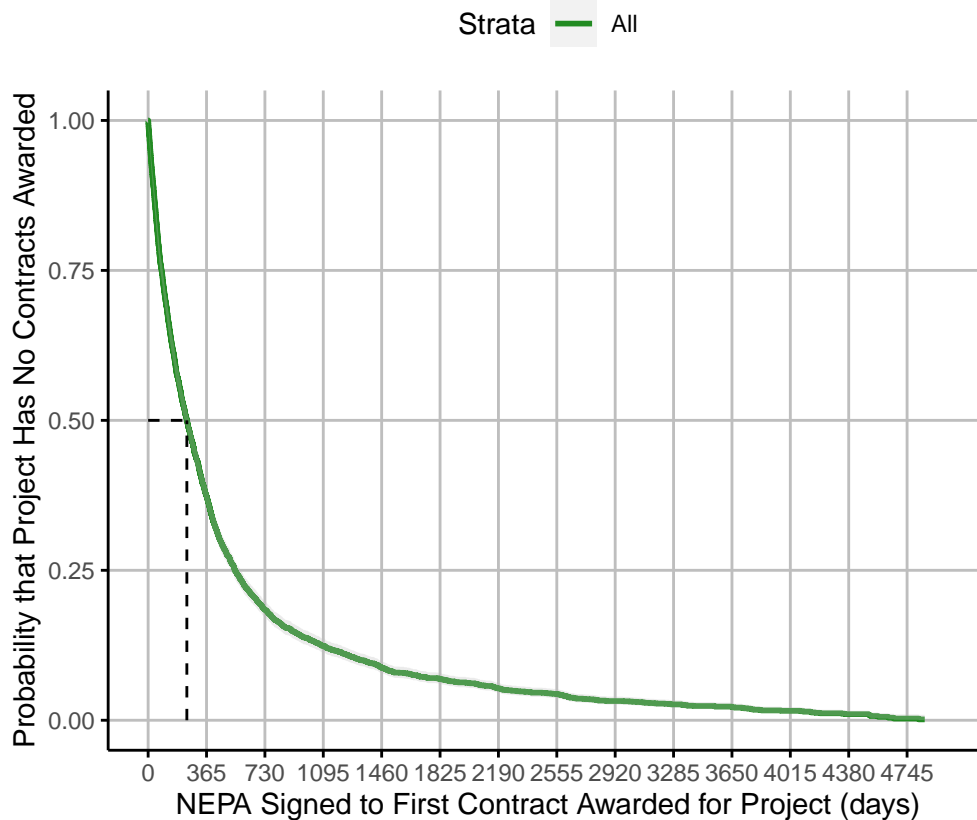
```
km_fit_init_all <- ggsurvplot(fit_init_all,
  conf.int = TRUE,
  risk.table = FALSE,
  risk.table.col = "strata",
  surv.median.line = "hv",
  break.time.by = 365,
  risk.table.y.text=FALSE,
  censor = FALSE,
  ylab = "Probability that Project has not Initiated",
  xlab = "NEPA Signed to First Activity in Project Completed (days)",
  palette = c("forestgreen"),
  surv.plot.height = 1,
  ggtheme = theme(aspect.ratio = 0.75,
    axis.line = element_line(colour = "black"),
    panel.grid.major = element_line(colour = "grey"),
    panel.border = element_blank(),
    panel.background = element_blank(),
    tables.theme = theme(aspect.ratio = 0.06)
  )
print(km_fit_init_all)
```



```
#ggsave( "/Users/kathrynmurenbeeld/Desktop/Survival_Analysis/WRITING/FIGS/km_fit_init_all.pdf", print(k
```

```
## ggsurvplot will create survival curves from the KM model created above.
```

```
km_fit_award_all <- ggsurvplot(fit_award_all,
  conf.int = TRUE,
  risk.table = FALSE,
  risk.table.col = "strata",
  surv.median.line = "hv",
  break.time.by = 365,
  risk.table.y.text=FALSE,
  censor = FALSE,
  ylab = "Probability that Project Has No Contracts Awarded",
  xlab = "NEPA Signed to First Contract Awarded for Project (days)",
  palette = c("forestgreen"),
  surv.plot.height = 1,
  ggtheme = theme(aspect.ratio = 0.75,
    axis.line = element_line(colour = "black"),
    panel.grid.major = element_line(colour = "grey"),
    panel.border = element_blank(),
    panel.background = element_blank(),
    tables.theme = theme(aspect.ratio = 0.06)
  )
print(km_fit_award_all)
```



```
#ggsave( "/Users/kathrynmurenbeeld/Desktop/Survival_Analysis/WRITING/FIGS/km_fit_award_all.pdf", print()
```

```
# Use the survfit function from the survival package to
# calculate the KM estimate for data grouped by appealed or non-appealed.
# This code chunk and the next two act as a template for the K-M estimation.
```

```
fit_assess_nepa <- survfit(Surv(ASSESSMENT_TIME, NEPA_COMP2) ~ DECISION.TYPE, data = df_fin)
fit_assess_nepa_table <- summary(fit_assess_nepa)
#fit_assess_nepa

fit_comp_nepa <- survfit(Surv(IMPLEMENTATION_TIME, PROJ_COMP) ~ DECISION.TYPE, data = df_fin)
fit_comp_nepa_table <- summary(fit_comp_nepa)
#fit_assess_nepa

fit_init_nepa <- survfit(Surv(NEPA_TO_FIRST_ACT_MAX, PROJ_INIT) ~ DECISION.TYPE, data = df_fin)
fit_init_nepa_table <- summary(fit_init_nepa)
#fit_init_nepa

fit_award_nepa <- survfit(Surv(NEPA_TO_CONTRACT_MAX, PROJ_AWARDED) ~ DECISION.TYPE, data = df_fin)
fit_award_nepa_table <- summary(fit_award_nepa)
#fit_award_nepa
```

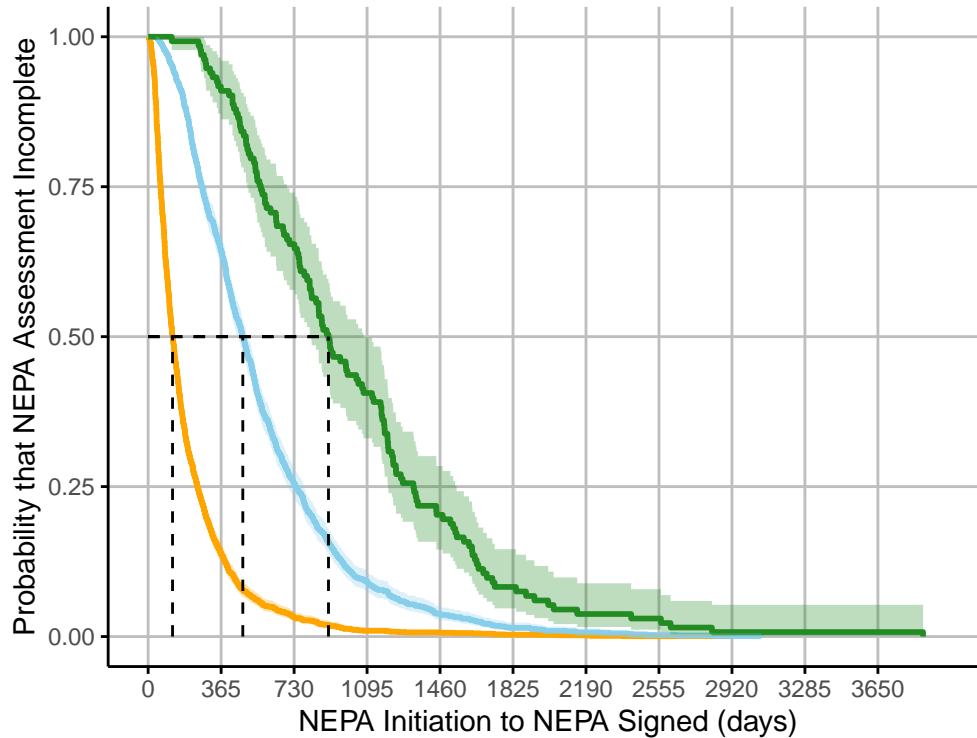
```
km_curv_assess_nepa <- ggsurvplot(fit_assess_nepa,
  conf.int = TRUE,
  risk.table = FALSE,
  risk.table.col = "strata",
  surv.median.line = "hv",
  break.time.by = 365,
```

```

risk.table.y.text=FALSE,
tables.height = 0.3,
censor = FALSE,
ylab = "Probability that NEPA Assessment Incomplete",
xlab = "NEPA Initiation to NEPA Signed (days)",
palette = c("orange", "skyblue", "forestgreen"),
surv.plot.height = 1,
ggtheme = theme(aspect.ratio = 0.75,
                 axis.line = element_line(colour = "black"),
                 panel.grid.major = element_line(colour = "grey"),
                 panel.border = element_blank(),
                 panel.background = element_blank()),
tables.theme = theme(aspect.ratio = 0.06)
)
print(km_curv_assess_nepa)

```

Strata — DECISION.TYPE=DM — DECISION.TYPE=DN — DECISION.TYPE=ROD



```
#ggsave( "/Users/kathrynmurenbeeld/Desktop/Survival_Analysis/WRITING/FIGS/km_curv_assess_nepa.pdf", pri
```

```

km_curv_comp_nepa <- ggsurvplot(fit_comp_nepa,
  conf.int = TRUE,
  risk.table = FALSE,
  risk.table.col = "strata",
  surv.median.line = "hv",
  break.time.by = 365,
  risk.table.y.text=FALSE,
  tables.height = 0.3,
  censor = FALSE,

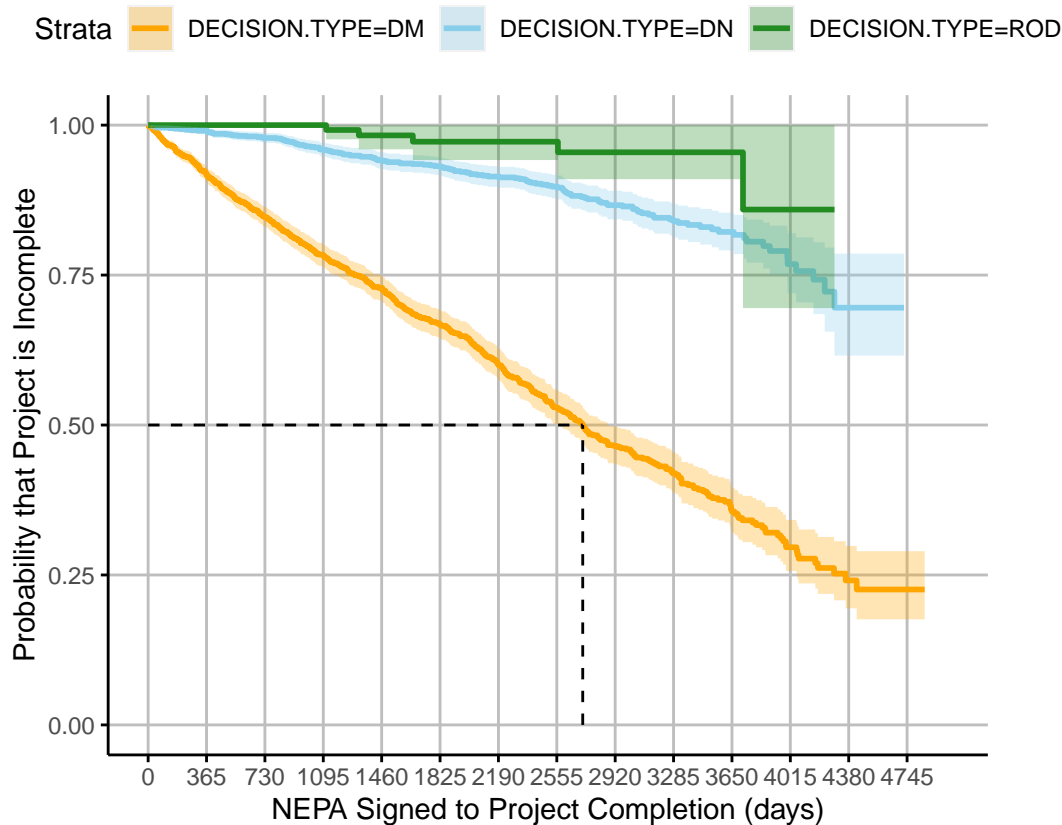
```



```

ylab = "Probability that Project is Incomplete",
xlab = "NEPA Signed to Project Completion (days)",
palette = c("orange", "skyblue", "forestgreen"),
surv.plot.height = 1,
ggtheme = theme(aspect.ratio = 0.75,
                 axis.line = element_line(colour = "black"),
                 panel.grid.major = element_line(colour = "grey"),
                 panel.border = element_blank(),
                 panel.background = element_blank()),
tables.theme = theme(aspect.ratio = 0.06)
)
print(km_curv_comp_nepa)

```



```
#ggsave( "/Users/kathrynmurenbeeld/Desktop/Survival_Analysis/WRITING/FIGS/km_curv_comp_nepa.pdf", print
```

```

km_curv_init_nepa <- ggsurvplot(fit_init_nepa,
  conf.int = TRUE,
  risk.table = FALSE,
  risk.table.col = "strata",
  surv.median.line = "hv",
  break.time.by = 365,
  risk.table.y.text=FALSE,
  tables.height = 0.3,
  censor = FALSE,
  ylab = "Probability that Project has not Initiated",
  xlab = "NEPA Signed to First Activity in Project Completed (days)",
  palette = c("orange", "skyblue", "forestgreen"),

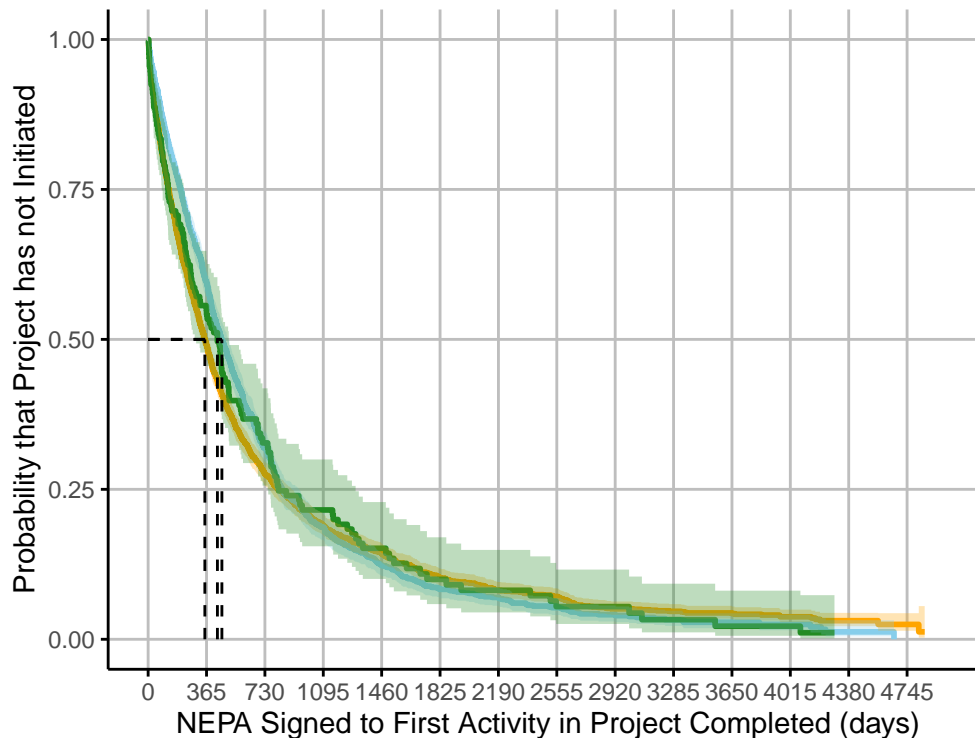
```

```

surv.plot.height = 1,
ggtheme = theme(aspect.ratio = 0.75,
                 axis.line = element_line(colour = "black"),
                 panel.grid.major = element_line(colour = "grey"),
                 panel.border = element_blank(),
                 panel.background = element_blank()),
tables.theme = theme(aspect.ratio = 0.06)
)
print(km_curv_init_nepa)

```

Strata — DECISION.TYPE=DM — DECISION.TYPE=DN — DECISION.TYPE=ROD



```
#ggsave( "/Users/kathrynmurenbeeld/Desktop/Survival_Analysis/WRITING/FIGS/km_curv_init_nepa.pdf", print
```

```

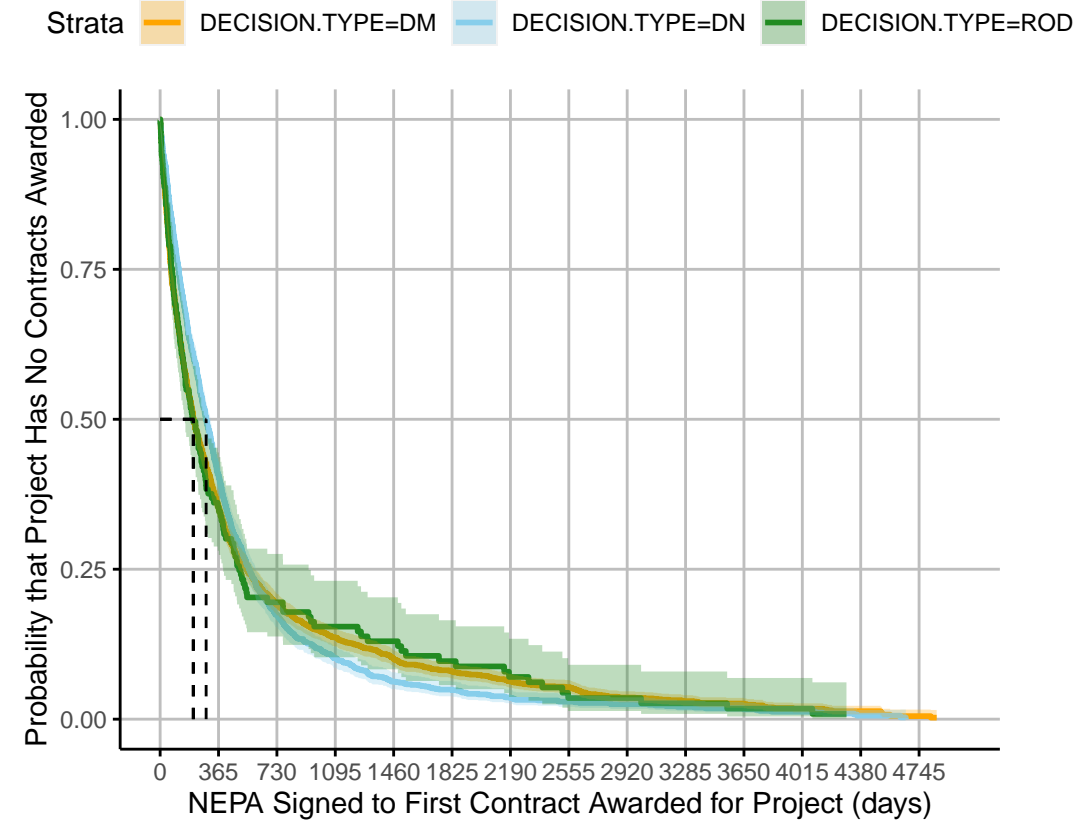
km_curv_award_nepa <- ggsurvplot(fit_award_nepa,
  conf.int = TRUE,
  risk.table = FALSE,
  risk.table.col = "strata",
  surv.median.line = "hv",
  break.time.by = 365,
  risk.table.y.text=FALSE,
  tables.height = 0.3,
  censor = FALSE,
  ylab = "Probability that Project Has No Contracts Awarded",
  xlab = "NEPA Signed to First Contract Awarded for Project (days)",
  palette = c("orange", "skyblue", "forestgreen"),
  surv.plot.height = 1,
  ggtheme = theme(aspect.ratio = 0.75,
                  axis.line = element_line(colour = "black"),

```

```

    panel.grid.major = element_line(colour = "grey"),
    panel.border = element_blank(),
    panel.background = element_blank(),
    tables.theme = theme(aspect.ratio = 0.06)
  )
print(km_curv_award_nepa)

```



```
#ggsave( "/Users/kathrynmurenbeeld/Desktop/Survival_Analysis/WRITING/FIGS/km_curv_award_nepa.pdf", prin
```

```
#summary(fit_award_nepa)
```