Elementary Effects for the BTD Model

Setup packages.

In [1]:
```
require(data.table)
require(magrittr)
require(sensitivity)

require(ggplot2)
```

```
Loading required package: data.table
Loading required package: magrittr
Loading required package: sensitivity
Registered S3 method overwritten by 'sensitivity':
  method    from
  print.src dplyr
Loading required package: ggplot2
```

Design experiment.

Load input ranges.

In [ ]:
```
z.ranges <- fread("input-ranges.tsv")
z.ranges %>% dim
```

In [ ]:
```
z.ranges %>% summary
```

One-at a time experiment with 500 repetitions, à la Morris.

```
In [ ]: z.design <- morris(
            NULL,
            factors = z.ranges$Variable,
            r = 500,
            design = list(
                type = "oat",
                levels = mapply(function(t, x0, x1) {
                    if (t == "Integer")
                        x1 - x0 + 1
                    else if (t == "Boolean")
                        2
                    else
                        5
                }, z.ranges$Type, z.ranges$`Sensitivity Minimum`, z.ranges$`Sens
        itivity Maximum`),
                grid.jump = 1
            )
        )
        z.design$X %>% dim
```

```
In [ ]: write.table(z.design$X, file = "design.tsv", row.names = FALSE, col.name
        s = TRUE, sep = "\t", quote = FALSE)
```

Relate the design to the model's variables.

```
In [ ]: z.inputs <- cbind(
            Run = 1:(dim(z.design$X)[1]),
            data.table(
                sweep(
                    sweep(z.design$X, MARGIN = 2, z.ranges$`Sensitivity Maximum`
        - z.ranges$`Sensitivity Minimum`, `*`),
                    MARGIN = 2,
                    z.ranges$`Sensitivity Minimum`,
                    `+`
                )
            )
        )
        z.inputs %>% summary
```

```
In [ ]: write.table(z.inputs, file="inputs.tsv", row.names = FALSE, col.names =
        TRUE, sep = "\t", quote = FALSE)
```

Analyze results.

Read files.

Read design.

In [2]:
```
z.design <- fread("design.tsv")
z.design %>% dim
```

42500 84

Read inputs.

In [3]:
```
z.inputs <- fread("inputs.tsv")
z.inputs %>% dim
```

42500 85

Read outputs.

In [4]:
```
z.outputs <- fread("outputs.tsv")
z.outputs[Time == 2050] %>% dim
```

42500 51

In [5]:
```
z.outputs[Time == 2050] %>% summary
```

```
      Run             Time      bioproduct market share mass
 Min.   :    1   Min.   :2050   Min.   :-27.65680           
 1st Qu.:10626   1st Qu.:2050   1st Qu.:  0.01149           
 Median :21250   Median :2050   Median :  0.33107           
 Mean   :21250   Mean   :2050   Mean   :  0.59533           
 3rd Qu.:31875   3rd Qu.:2050   3rd Qu.:  0.71766           
 Max.   :42500   Max.   :2050   Max.   : 54.99460           
 current market size economic current market size mass long term market
share
 Min.   :-9.881e+13          Min.   :1.750e+04        Min.   :0.0000   
 1st Qu.:-2.500e+10          1st Qu.:2.069e+07        1st Qu.:0.2126   
 Median : 7.524e+10          Median :6.609e+07        Median :0.3172   
 Mean   : 3.949e+11          Mean   :2.866e+08        Mean   :0.3959   
 3rd Qu.: 5.717e+11          3rd Qu.:2.009e+08        3rd Qu.:0.5406   
 Max.   : 5.844e+13          Max.   :7.965e+09        Max.   :0.9810   
 long term market value    Adopters           NonAdopters       
 Min.   :0.000e+00      Min.   : -11334900   Min.   :5.130e+02  
 1st Qu.:0.000e+00      1st Qu.:    416403   1st Qu.:5.613e+06  
 Median :0.000e+00      Median :   7035080   Median :2.763e+07  
 Mean   :2.098e+09      Mean   : 135755526   Mean   :1.205e+08  
 3rd Qu.:7.470e+02      3rd Qu.:  59639400   3rd Qu.:9.357e+07  
 Max.   :6.297e+11      Max.   :5976570000   Max.   :5.388e+09  
 Potential Adopters   abandoning bioproduct Cumulative Demoing Producti
on
 Min.   : -20212000   Min.   :0             Min.   :       0     
 1st Qu.:     20273   1st Qu.:0             1st Qu.:       0     
 Median :   1188780   Median :0             Median :       0     
 Mean   :  30356549   Mean   :0             Mean   :  182732     
 3rd Qu.:  19697400   3rd Qu.:0             3rd Qu.:       0     
 Max.   :2077400000   Max.   :0             Max.   :11109600     
 Cumulative Production  prepiloting     pilot plant construction
 Min.   :       0      Min.   :0.0000   Min.   :0.0000000       
 1st Qu.:       0      1st Qu.:0.0000   1st Qu.:0.0000000       
 Median :       0      Median :0.0000   Median :0.0000000       
 Mean   :  696871      Mean   :0.2168   Mean   :0.0002118       
 3rd Qu.:       0      3rd Qu.:0.0000   3rd Qu.:0.0000000       
 Max.   :67509600      Max.   :1.0000   Max.   :1.0000000       
 pilot plant is built startup piloting complete piloting ongoing
 Min.   :0.000        Min.   :0.0000           Min.   :0.0000  
 1st Qu.:1.000        1st Qu.:1.0000           1st Qu.:0.0000  
 Median :1.000        Median :1.0000           Median :1.0000  
 Mean   :0.783        Mean   :0.7767           Mean   :0.5731  
 3rd Qu.:1.000        3rd Qu.:1.0000           3rd Qu.:1.0000  
 Max.   :1.000        Max.   :1.0000           Max.   :1.0000  
 piloting progress piloting complete   predemoing      demo plant const
ruction
 Min.   :0.00      Min.   :0.0000    Min.   :0.00000   Min.   :0.000000
 1st Qu.:0.00      1st Qu.:0.0000    1st Qu.:0.00000   1st Qu.:0.000000
 Median :0.00      Median :0.0000    Median :0.00000   Median :0.000000
 Mean   :0.21      Mean   :0.2099    Mean   :0.06772   Mean   :0.003106
 3rd Qu.:0.00      3rd Qu.:0.0000    3rd Qu.:0.00000   3rd Qu.:0.000000
 Max.   :1.00      Max.   :1.0000    Max.   :1.00000   Max.   :1.000000
 demo plant is built regulatory process ongoing startup demoing complet
ed
 Min.   :0.0000      Min.   :0.00               Min.   :0.0000  
 1st Qu.:0.0000      1st Qu.:0.00               1st Qu.:0.0000  
 Median :0.0000      Median :0.00               Median :0.0000  
```

```
 Mean   :0.1372      Mean   :0.01              Mean   :0.1329  
 3rd Qu.:0.0000      3rd Qu.:0.00              3rd Qu.:0.0000  
 Max.   :1.0000      Max.   :1.00              Max.   :1.0000  
 demoing ongoing   demoing progress  demoing complete  regulatory delay
 Min.   :0.00000   Min.   :0.00000   Min.   :0.00000   Min.   :0.0000  
 1st Qu.:0.00000   1st Qu.:0.00000   1st Qu.:0.00000   1st Qu.:0.0000  
 Median :0.00000   Median :0.00000   Median :0.00000   Median :0.0000  
 Mean   :0.09219   Mean   :0.05074   Mean   :0.04504   Mean   :0.0024  
 3rd Qu.:0.00000   3rd Qu.:0.00000   3rd Qu.:0.00000   3rd Qu.:0.0000  
 Max.   :1.00000   Max.   :1.00000   Max.   :1.00000   Max.   :1.0000  
 precommercial      commercial plant construction commercial plant is b
uilt
 Min.   :0.000000   Min.   :0.000000             Min.   :0.00000  
 1st Qu.:0.000000   1st Qu.:0.000000             1st Qu.:0.00000  
 Median :0.000000   Median :0.000000             Median :0.00000  
 Mean   :0.007906   Mean   :0.003294             Mean   :0.03144  
 3rd Qu.:0.000000   3rd Qu.:0.000000             3rd Qu.:0.00000  
 Max.   :1.000000   Max.   :1.000000             Max.   :1.00000  
 commercial plant operation technology readiness level stage in progres
s
 Min.   :0.00000           Min.   :6.000             Min.   : 2.000  
 1st Qu.:0.00000           1st Qu.:6.000             1st Qu.: 5.000  
 Median :0.00000           Median :6.000             Median : 6.000  
 Mean   :0.03071           Mean   :6.311             Mean   : 8.823  
 3rd Qu.:0.00000           3rd Qu.:6.000             3rd Qu.:10.000  
 Max.   :1.00000           Max.   :9.000             Max.   :58.000  
   BS equity          payback period     NPV at required return
 Min.   :-2.075e+09   Min.   :    0.00   Min.   :-3.664e+10    
 1st Qu.:-1.073e+07   1st Qu.:    8.78   1st Qu.:-8.530e+08    
 Median :-1.996e+06   Median :   12.02   Median :-2.630e+08    
 Mean   : 2.478e+09   Mean   :   28.52   Mean   : 8.902e+08    
 3rd Qu.: 1.777e+08   3rd Qu.:   15.89   3rd Qu.: 8.030e+06    
 Max.   : 4.108e+11   Max.   :34389.80   Max.   : 2.937e+11    
 profitability indicator bioproduct favorability indicator
 Min.   :0.0000          Min.   :0.0000                   
 1st Qu.:1.0000          1st Qu.:0.0000                   
 Median :1.0000          Median :0.0000                   
 Mean   :0.9214          Mean   :0.2349                   
 3rd Qu.:1.0000          3rd Qu.:0.0000                   
 Max.   :1.0000          Max.   :1.0000                   
 long term selling price without green premium after market entry
 Min.   : 369.7                                                  
 1st Qu.:1399.2                                                  
 Median :2078.5                                                  
 Mean   :2136.7                                                  
 3rd Qu.:2768.2                                                  
 Max.   :6965.5                                                  
 total approval cost total approval time in business indicator
 Min.   :-1599210    Min.   : 0.019      Min.   :0.0000       
 1st Qu.:  199833    1st Qu.: 1.301      1st Qu.:0.0000       
 Median : 1052730    Median : 3.086      Median :0.0000       
 Mean   : 1636119    Mean   : 3.762      Mean   :0.4059       
 3rd Qu.: 2626350    3rd Qu.: 5.521      3rd Qu.:1.0000       
 Max.   :12420100    Max.   :20.166      Max.   :1.0000       
 internal project cancelled indicator   investing            granting     
 Min.   :0.0000000                    Min.   :0.000e+00   Min.   :0.000
e+00
```

```
 1st Qu.:0.0000000                     1st Qu.:0.000e+00   1st Qu.:0.000
e+00
 Median :0.0000000                     Median :0.000e+00   Median :0.000
e+00
 Mean   :0.0006588                     Mean   :2.162e+08   Mean   :5.044
e+08
 3rd Qu.:0.0000000                     3rd Qu.:0.000e+00   3rd Qu.:3.192
e+07
 Max.   :1.0000000                     Max.   :2.189e+11   Max.   :9.984
e+10
 Total Government Grants Total Investment    Working Capital
 Min.   :0.000e+00       Min.   :3.000e+06   Min.   :-2.075e+09
 1st Qu.:0.000e+00       1st Qu.:1.145e+08   1st Qu.:-1.073e+07
 Median :1.373e+08       Median :7.370e+08   Median :-1.996e+06
 Mean   :6.579e+08       Mean   :1.274e+09   Mean   : 2.478e+09
 3rd Qu.:3.661e+08       3rd Qu.:1.488e+09   3rd Qu.: 1.777e+08
 Max.   :3.754e+10       Max.   :4.533e+10   Max.   : 4.108e+11
 IS production incentive
 Min.   :0.000e+00
 1st Qu.:0.000e+00
 Median :0.000e+00
 Mean   :1.075e+08
 3rd Qu.:0.000e+00
 Max.   :1.261e+10
```

Compute elementary effects.

Just use the final year.

In [6]:
```
z.outputs.clean <- z.outputs[order(Run)][`Time` == 2050, c(-1, -2)] %>%
as.matrix
z.outputs.clean %>% dim
```

42500 49

Define functions to compute elementary effects.

In [7]:

```
ind.rep <- function(i, p) {
# indices of the points of the ith trajectory in the DoE
  (1 : (p + 1)) + (i - 1) * (p + 1)
}

ee.oat <- function(X, y) {
  # compute the elementary effects for a OAT design
  p <- ncol(X)
  r <- nrow(X) / (p + 1)

#  if(is(y,"numeric")){
  if(inherits(y, "numeric")){
    one_i_vector <- function(i){
      j <- ind.rep(i, p)
      j1 <- j[1 : p]
      j2 <- j[2 : (p + 1)]
      # return((y[j2] - y[j1]) / rowSums(X[j2,] - X[j1,]))
      return(solve(X[j2,] - X[j1,], y[j2] - y[j1]))
    }
    ee <- vapply(1:r, one_i_vector, FUN.VALUE = numeric(p))
    ee <- t(ee)
    # "ee" is now a (r times p)-matrix.
#  } else if(is(y,"matrix")){
  } else if(inherits(y, "matrix")){
    one_i_matrix <- function(i){
      j <- ind.rep(i, p)
      j1 <- j[1 : p]
      j2 <- j[2 : (p + 1)]
      return(solve(X[j2,] - X[j1,],
                   y[j2, , drop = FALSE] - y[j1, , drop = FALSE]))
    }
    ee <- vapply(1:r, one_i_matrix,
                 FUN.VALUE = matrix(0, nrow = p, ncol = dim(y)[2]))
    # Special case handling for p == 1 and ncol(y) == 1 (in this case,
 "ee" is
    # a vector of length "r"):
    if(p == 1 && dim(y)[2] == 1){
      ee <- array(ee, dim = c(r, 1, 1))
    }
    # Transpose "ee" (an array of dimensions c(p, ncol(y), r)) to an arr
ay of
    # dimensions c(r, p, ncol(y)) (for better consistency with the stand
ard
    # case that "class(y) == "numeric""):
    ee <- aperm(ee, perm = c(3, 1, 2))
#  } else if(is(y,"array")){
  } else if(inherits(y, "array")){
    one_i_array <- function(i){
      j <- ind.rep(i, p)
      j1 <- j[1 : p]
      j2 <- j[2 : (p + 1)]
      ee_per_3rd_dim <- sapply(1:(dim(y)[3]), function(idx_3rd_dim){
        y_j2_matrix <- y[j2, , idx_3rd_dim]
        y_j1_matrix <- y[j1, , idx_3rd_dim]
        # Here, the result of "solve(...)" is a (p times dim(y)[2])-matr
ix or
```

```
        # a vector of length dim(y)[2] (if p == 1):
        solve(X[j2,] - X[j1,], y_j2_matrix - y_j1_matrix)
      }, simplify = "array")
      if(dim(y)[2] == 1){
        # Correction needed if dim(y)[2] == 1, so "y_j2_matrix" and
        # "y_j1_matrix" have been dropped to matrices (or even vectors,
 if also
        # p == 1):
        ee_per_3rd_dim <- array(ee_per_3rd_dim,
                                dim = c(p, dim(y)[2], dim(y)[3]))
      } else if(p == 1){
        # Correction needed if p == 1 (and dim(y)[2] > 1), so "y_j2_matr
ix" and
        # "y_j1_matrix" have been dropped to matrices:
        ee_per_3rd_dim <- array(ee_per_3rd_dim,
                                dim = c(1, dim(y)[2], dim(y)[3]))
      }
      # "ee_per_3rd_dim" is now an array of dimensions
      # c(p, dim(y)[2], dim(y)[3]). Assign the corresponding names for t
he
      # third dimension:
      if(is.null(dimnames(ee_per_3rd_dim))){
        dimnames(ee_per_3rd_dim) <- dimnames(y)
      } else{
        dimnames(ee_per_3rd_dim)[[3]] <- dimnames(y)[[3]]
      }
      return(ee_per_3rd_dim)
    }
    ee <- sapply(1:r, one_i_array, simplify = "array")
    # Special case handling if "ee" has been dropped to a vector:
#    if(is(ee,"numeric")){
    if (inherits(ee, "numeric")){
      ee <- array(ee, dim = c(p, dim(y)[2], dim(y)[3], r))
      dimnames(ee) <- list(NULL, dimnames(y)[[2]], dimnames(y)[[3]], NUL
L)
    }
    # "ee" is an array of dimensions c(p, dim(y)[2], dim(y)[3], r), so i
t is
    # transposed to an array of dimensions c(r, p, dim(y)[2], dim(y)
[3]):
    ee <- aperm(ee, perm = c(4, 1, 2, 3))
  }
  return(ee)
}
```

Elementary effects.

In [8]:
```
z.ee <- ee.oat(z.design, z.outputs.clean)
z.ee %>% dim
```

500 84 49

mu, mu*, and sigma.

In [9]:
```
z.mu <- apply(z.ee, 3, function(M){
  apply(M, 2, mean)
})
z.mu <- melt(
    cbind(
        data.table(Input=rownames(z.mu)),
        data.table(z.mu)
    ),
    id.vars="Input",
    variable.name = "Output",
    value.name = "mu"
)
z.mu %>% head
```

A data.table: 6 x 3

Input	Output	mu
<chr>	<fct>	<dbl>
advertising budget	bioproduct market share mass	0.303273857
advertising start time	bioproduct market share mass	-0.387986934
aversion to NPV deviation	bioproduct market share mass	-0.001418284
base external investor ask rate	bioproduct market share mass	-0.023685767
bioproduct long term price	bioproduct market share mass	-0.160606092
bioproduct offtake agreement	bioproduct market share mass	0.512129831

In [10]:
```
z.mu.star <- apply(abs(z.ee), 3, function(M){
  apply(M, 2, mean)
})
z.mu.star <- melt(
    cbind(
        data.table(Input=rownames(z.mu.star)),
        data.table(z.mu.star)
    ),
    id.vars="Input",
    variable.name = "Output",
    value.name = "mu*"
)
z.mu.star %>% head
```

A data.table: 6 x 3

Input	Output	mu*
<chr>	<fct>	<dbl>
advertising budget	bioproduct market share mass	0.31027692
advertising start time	bioproduct market share mass	0.41195148
aversion to NPV deviation	bioproduct market share mass	0.01152693
base external investor ask rate	bioproduct market share mass	0.26447606
bioproduct long term price	bioproduct market share mass	0.39980623
bioproduct offtake agreement	bioproduct market share mass	0.77572582

In [11]:
```
z.sigma <- apply(z.ee, 3, function(M){
  apply(M, 2, sd)
})
z.sigma <- melt(
    cbind(
        data.table(Input=rownames(z.sigma)),
        data.table(z.sigma)
    ),
    id.vars="Input",
    variable.name = "Output",
    value.name = "sigma"
)
z.sigma %>% head
```

A data.table: 6 x 3

Input	Output	sigma
<chr>	<fct>	<dbl>
advertising budget	bioproduct market share mass	1.1085135
advertising start time	bioproduct market share mass	1.4664525
aversion to NPV deviation	bioproduct market share mass	0.1115146
base external investor ask rate	bioproduct market share mass	0.6278341
bioproduct long term price	bioproduct market share mass	1.7387415
bioproduct offtake agreement	bioproduct market share mass	4.5942430

In [12]:
```
z.results <- merge(merge(z.mu, z.mu.star, on=c("Input", "Output")), z.si
gma, on=c("Input", "Output"))
z.results %>% head
```

A data.table: 6 x 5

Input	Output	mu	mu*	sigma
<chr>	<fct>	<dbl>	<dbl>	<dbl>
advertising budget	bioproduct market share mass	3.032739e-01	3.102769e-01	1.108513e+00
advertising budget	current market size economic	-2.412707e+11	9.415770e+11	4.866741e+12
advertising budget	current market size mass	0.000000e+00	0.000000e+00	0.000000e+00
advertising budget	long term market share	-6.742848e-04	1.255821e-03	1.507683e-02
advertising budget	long term market value	-3.252294e+08	3.252294e+08	7.252179e+09
advertising budget	Adopters	4.651019e+07	5.092098e+07	1.587423e+08

In [13]:
```
z.results[, `:=`(
    `mu rank`     = frank(- `mu`    ),
    `mu* rank`    = frank(- `mu*`   ),
    `sigma rank`  = frank(- `sigma`)
), by=.(Output)]
z.results %>% head
```

A data.table: 6 x 8

Input	Output	mu	mu*	sigma	mu rank	mu* rank	sigma rank
<chr>	<fct>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
advertising budget	bioproduct market share mass	3.032739e-01	3.102769e-01	1.108513e+00	3	15	24
advertising budget	current market size economic	-2.412707e+11	9.415770e+11	4.866741e+12	80	23	25
advertising budget	current market size mass	0.000000e+00	0.000000e+00	0.000000e+00	43	44	44
advertising budget	long term market share	-6.742848e-04	1.255821e-03	1.507683e-02	64	49	51
advertising budget	long term market value	-3.252294e+08	3.252294e+08	7.252179e+09	71	30	24
advertising budget	Adopters	4.651019e+07	5.092098e+07	1.587423e+08	6	13	23

Interpret results.

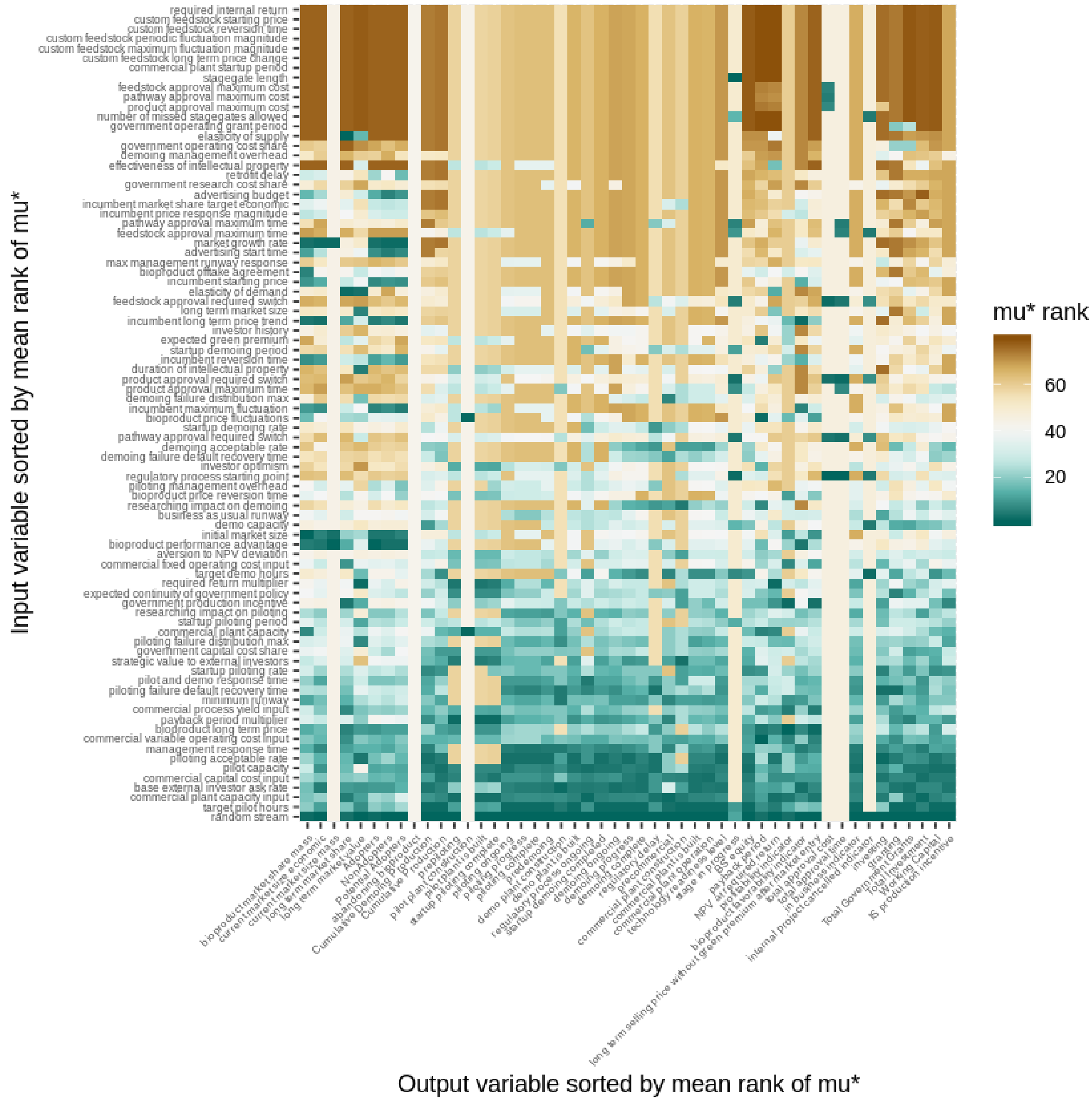
Interpretations:

- mu: influence of variable
- mustar: influence of variable, accounting for non-monoticity
- sigma: non-linear and interaction effects for variable

Heat map for mu*.

In [14]:

```
ggplot(
    z.results,
    aes(
        x=factor(
            Output,
            levels=z.results[, .(`sort` = mean(`mu* rank`)), by=.(Output
)][order(sort), `Output`]
        ),
        y=factor(
            Input,
            levels=z.results[, .(`sort` = mean(`mu* rank`)), by=.(Input)
][order(sort), `Input`]
        ),
        fill=`mu* rank`
    )
) +
    geom_tile() +
    scale_fill_distiller(type="div") +
    xlab("Output variable sorted by mean rank of mu*") +
    ylab("Input variable sorted by mean rank of mu*") +
    theme(
        axis.text=element_text(size=5),
        axis.text.x = element_text(angle = 45, hjust=1)
    )
```



Heat map for sigma.

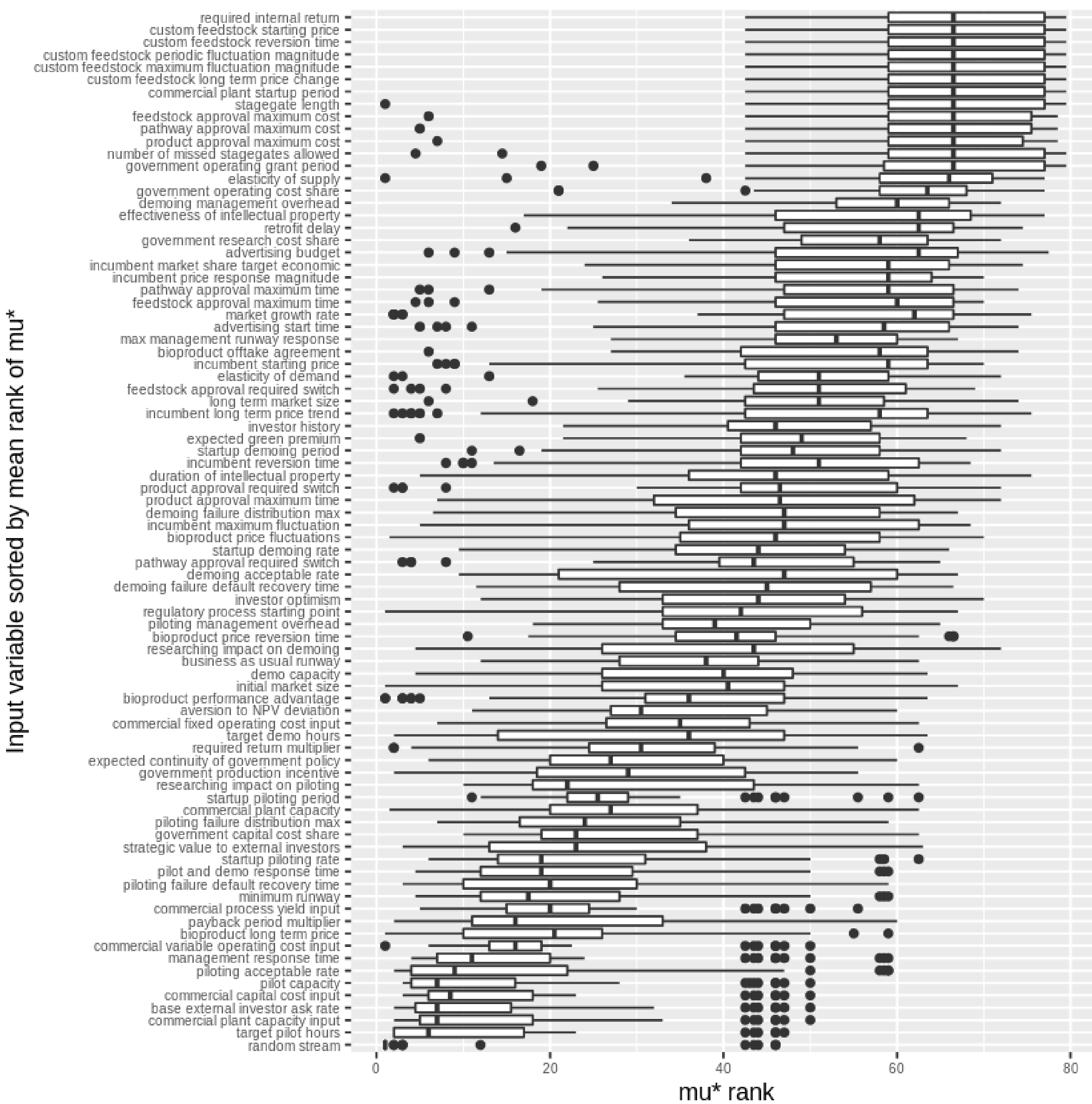
In [15]:

```
ggplot(
    z.results,
    aes(
        x=factor(
            Output,
            levels=z.results[, .(`sort` = mean(`sigma rank`)), by=.(Outp
ut)][order(sort), `Output`]
        ),
        y=factor(
            Input,
            levels=z.results[, .(`sort` = mean(`sigma rank`)), by=.(Inpu
t)][order(sort), `Input`]
        ),
        fill=`sigma rank`
    )
) +
    geom_tile() +
    scale_fill_distiller(type="div") +
    xlab("Output variable sorted by mean rank of sigma") +
    ylab("Input variable sorted by mean rank of sigma") +
    theme(
        axis.text=element_text(size=5),
        axis.text.x = element_text(angle = 45, hjust=1)
    )
```

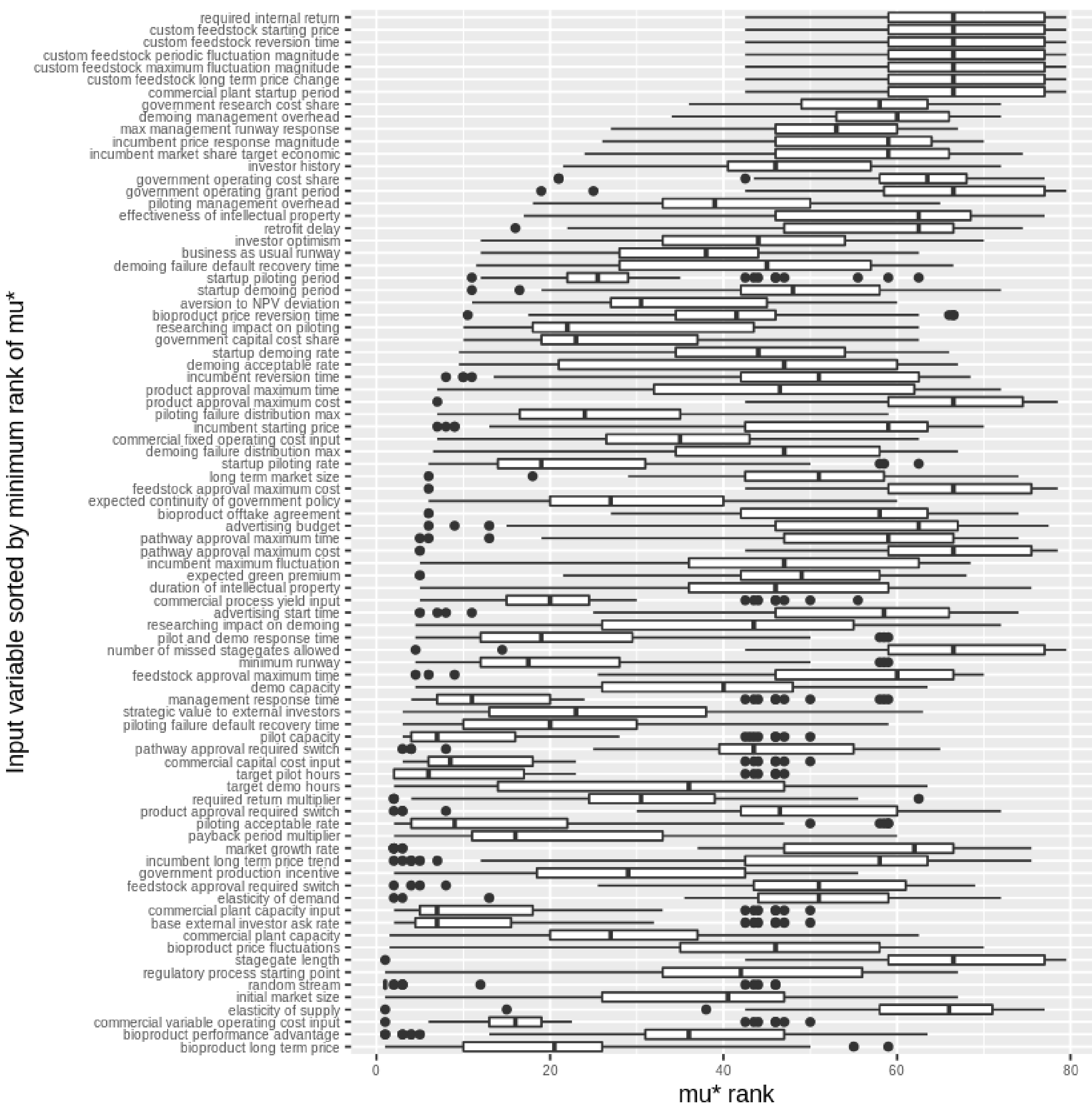
Box plots of ranks for mu*.

In [16]:

```
ggplot(
    z.results,
    aes(
        x=factor(
            Input,
            levels=z.results[, .(`sort` = mean(`mu* rank`)), by=.(Input)
][order(sort), `Input`]
        ),
        y=`mu* rank`
    )
) +
    geom_boxplot() +
    coord_flip() +
    xlab("Input variable sorted by mean rank of mu*") +
    theme(axis.text=element_text(size=6))
```

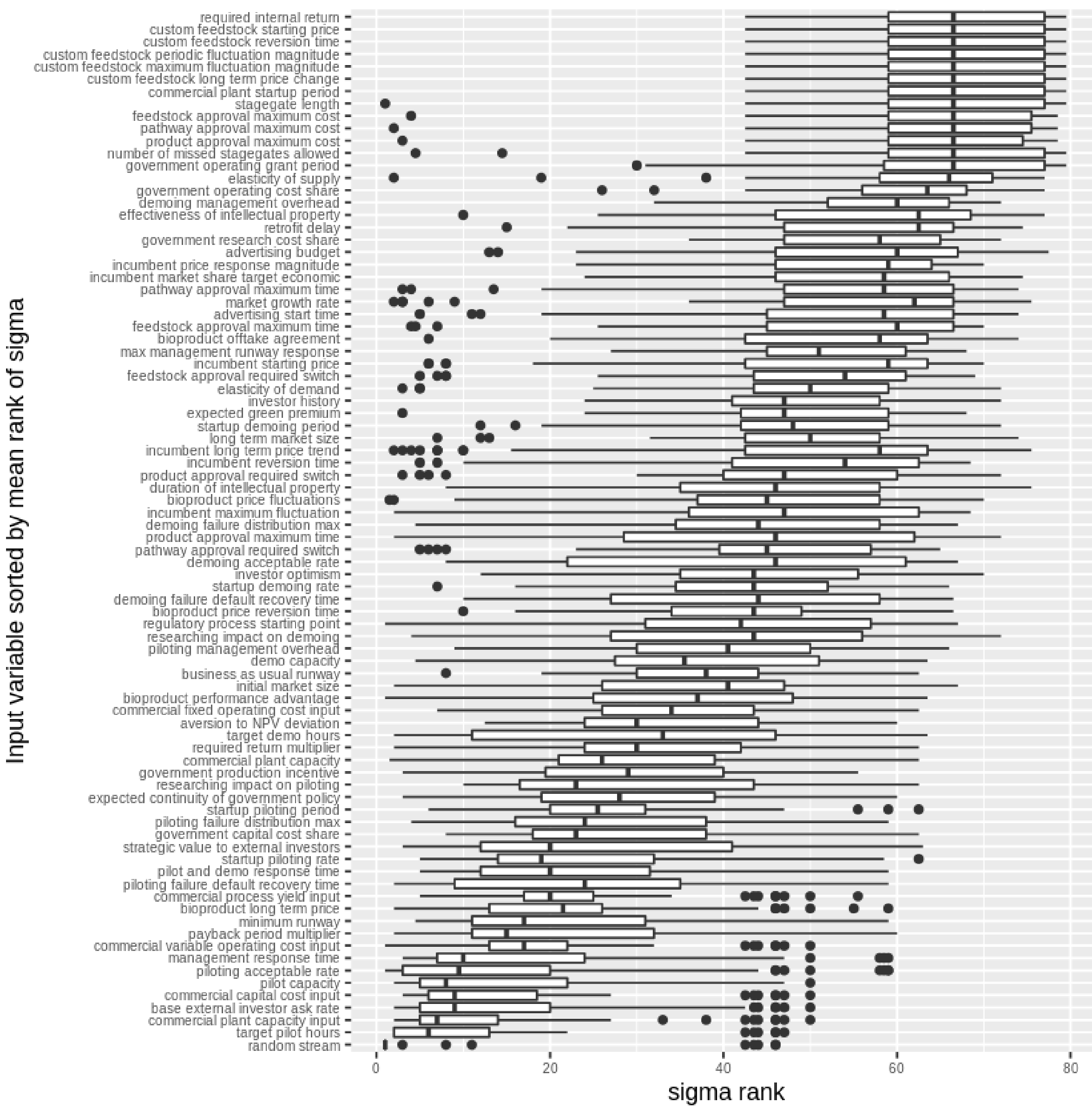


In [17]:
```
ggplot(
    z.results,
    aes(
        x=factor(
            Input,
            levels=z.results[, .(`sort` = min(`mu* rank`)), by=.(Input)]
[order(sort), `Input`]
        ),
        y=`mu* rank`
    )
) +
    geom_boxplot() +
    coord_flip() +
    xlab("Input variable sorted by minimum rank of mu*") +
    theme(axis.text=element_text(size=6))
```

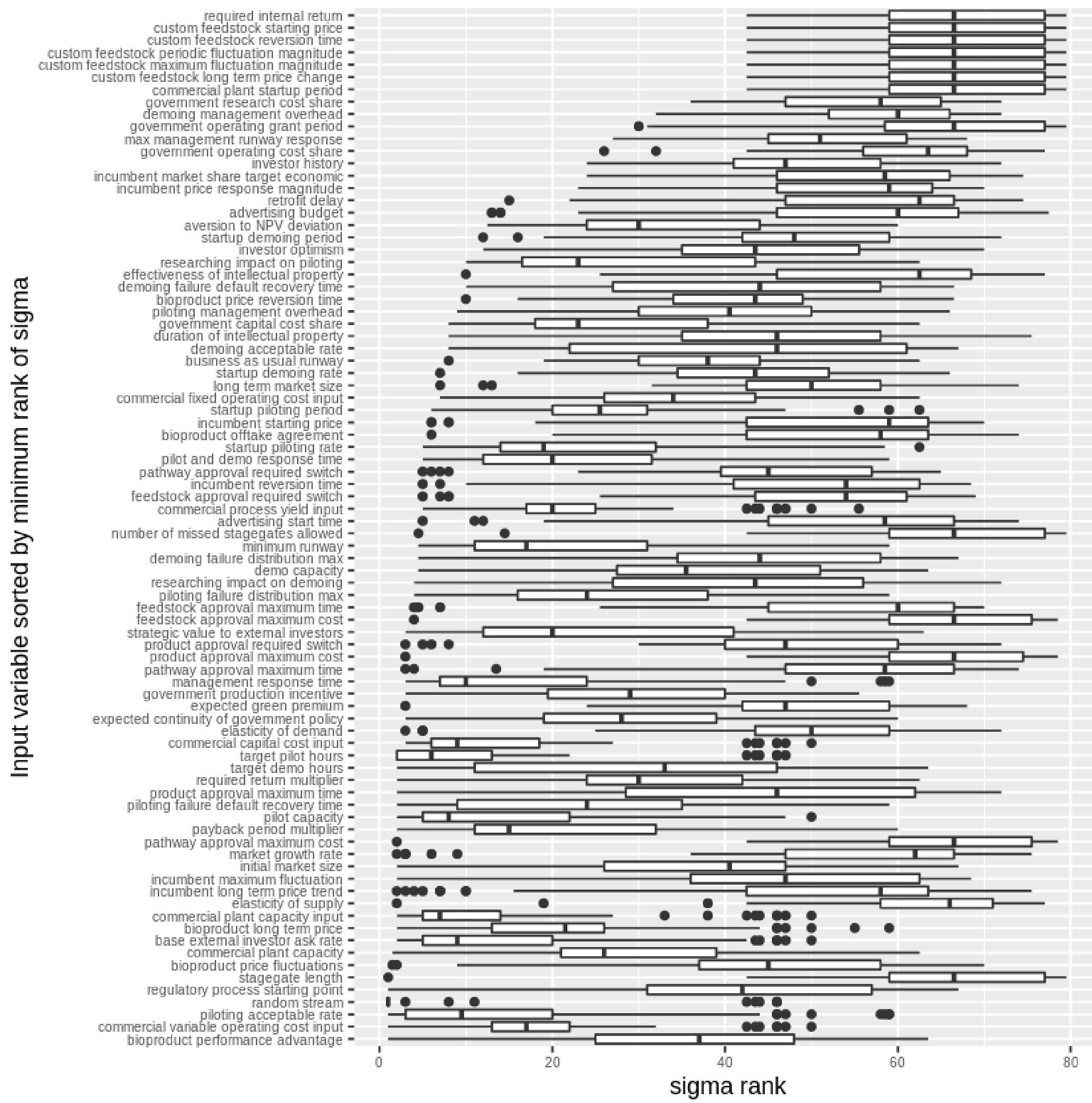


Box plots of ranks for sigma.

In [18]:
```
ggplot(
    z.results,
    aes(
        x=factor(
            Input,
            levels=z.results[, .(`sort` = mean(`sigma rank`)), by=.(Inpu
t)][order(sort), `Input`]
        ),
        y=`sigma rank`
    )
) +
    geom_boxplot() +
    coord_flip() +
    xlab("Input variable sorted by mean rank of sigma") +
    theme(axis.text=element_text(size=6))
```



In [19]:
```
ggplot(
    z.results,
    aes(
        x=factor(
            Input,
            levels=z.results[, .(`sort` = min(`sigma rank`)), by=.(Input
)][order(sort), `Input`]
        ),
        y=`sigma rank`
    )
) +
    geom_boxplot() +
    coord_flip() +
    xlab("Input variable sorted by minimum rank of sigma") +
    theme(axis.text=element_text(size=6))
```



Select inputs variables for variance-based sensitivity analysis.

Select the variables whose median rank is less than 30.

In [20]:

```
z.results[,
    .(
        `mu* minimum rank`    = min   (`mu* rank`),
        `mu* mean rank`       = mean  (`mu* rank`),
        `mu* median rank`     = median(`mu* rank`),
        `sigma minimum rank` = min   (`sigma rank`),
        `sigma mean rank`     = mean  (`sigma rank`),
        `sigma median rank`   = median(`sigma rank`)
    ),
    by=.(Input)
][order(`mu* median rank`)][
    `mu* median rank` <= 30 | `sigma median rank` <= 30
]
```

A data.table: 26 x 7

Input	mu* minimum rank	mu* mean rank	mu* median rank	sigma minimum rank	sigma mean rank	sigma median rank
<chr>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
random stream	1.0	5.816327	1.0	1.0	5.816327	1.0
target pilot hours	2.0	11.948980	6.0	2.0	11.775510	6.0
base external investor ask rate	2.0	13.744898	7.0	2.0	14.755102	9.0
commercial plant capacity input	2.0	13.673469	7.0	2.0	14.428571	7.0
pilot capacity	3.0	14.632653	7.0	2.0	15.285714	8.0
commercial capital cost input	3.0	14.326531	8.5	3.0	15.173469	9.0
piloting acceptable rate	2.0	17.908163	9.0	1.0	17.826531	9.5
management response time	4.0	18.091837	11.0	3.0	18.897959	10.0
commercial variable operating cost input	1.0	19.183673	16.0	1.0	20.000000	17.0
payback period multiplier	2.0	22.020408	16.0	2.0	21.285714	15.0
minimum runway	4.5	22.683673	17.5	4.5	22.785714	17.0
pilot and demo response time	4.5	23.581633	19.0	5.0	24.030612	20.0
startup piloting rate	6.0	24.122449	19.0	5.0	24.377551	19.0
commercial process yield input	5.0	22.500000	20.0	5.0	22.938776	20.0
piloting failure default recovery time	3.0	23.000000	20.0	2.0	23.857143	24.0
bioproduct long term price	1.0	21.989796	20.5	2.0	22.938776	21.5
researching impact on piloting	10.0	28.571429	22.0	10.0	28.755102	23.0
government capital cost share	10.0	27.520408	23.0	8.0	27.357143	23.0
strategic value to external investors	3.0	26.132653	23.0	3.0	26.265306	20.0
piloting failure distribution max	7.0	27.571429	24.0	4.0	27.397959	24.0
startup piloting period	11.0	28.459184	25.5	6.0	28.020408	25.5
commercial plant capacity	1.5	28.193878	27.0	1.5	29.418367	26.0
expected continuity of government policy	6.0	29.071429	27.0	3.0	28.653061	28.0
government production incentive	2.0	28.897959	29.0	3.0	29.295918	29.0

Input	mu* minimum rank	mu* mean rank	mu* median rank	sigma minimum rank	sigma mean rank	sigma median rank
<chr>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
aversion to NPV deviation	11.0	34.142857	30.5	12.5	33.285714	30.0
required return multiplier	2.0	30.326531	30.5	2.0	30.285714	30.0

Since production is the most important output, check the ranks for that.

In [21]:

```
z.results[Output == "Cumulative Production"][order(`mu* rank`)]
```

A data.table: 84 x 8

Input	Output	mu	mu*	sigma	mu rank	mu* rank	sigma rank
<chr>	<fct>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
random stream	Cumulative Production	-940470.13	15393209.2	100494557	83	1	1
target pilot hours	Cumulative Production	-3423993.41	3902968.9	21377483	84	2	4
pilot capacity	Cumulative Production	-170879.34	3598234.2	23100902	72	3	2
commercial capital cost input	Cumulative Production	1125500.14	3150768.6	21928963	4	4	3
piloting acceptable rate	Cumulative Production	78101.40	2604742.6	16174465	17	5	6
base external investor ask rate	Cumulative Production	691010.42	2506264.6	16827901	7	6	5
commercial plant capacity input	Cumulative Production	1636623.64	2101824.9	15772556	1	7	7
target demo hours	Cumulative Production	70952.60	1868923.5	15558456	18	8	8
startup piloting rate	Cumulative Production	1545903.95	1793046.6	12786463	2	9	13
government capital cost share	Cumulative Production	1077553.23	1541084.7	14019066	5	10	9
bioproduct long term price	Cumulative Production	1242954.43	1450861.2	13159696	3	11	10
minimum runway	Cumulative Production	-220764.47	1448032.3	11081354	74	12	16
strategic value to external investors	Cumulative Production	-704196.98	1355163.5	12944755	82	13	12
commercial variable operating cost input	Cumulative Production	-280251.17	1307569.7	11520019	75	14	14
commercial plant capacity	Cumulative Production	880855.86	1251252.1	7678975	6	15	27
payback period multiplier	Cumulative Production	202565.29	1171182.4	8917840	11	16	23
management response time	Cumulative Production	-322141.54	1161898.0	9025454	78	17	22
demoing failure default recovery time	Cumulative Production	-346149.85	1134374.0	12946849	79	18	11
demoing acceptable rate	Cumulative Production	-692497.94	1046048.0	9645279	81	19	20
commercial process yield input	Cumulative Production	-160919.86	1019512.1	11138562	70	20	15
pilot and demo response time	Cumulative Production	343226.60	937524.6	9170210	9	21	21

Input	Output	mu	mu*	sigma	mu rank	mu* rank	sigma rank
<chr>	<fct>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
piloting failure distribution max	Cumulative Production	192157.40	907394.2	10032682	13	22	17
demo capacity	Cumulative Production	-104939.45	860641.1	6611611	69	23	29
startup piloting period	Cumulative Production	-303812.34	795238.1	9769802	77	24	19
required return multiplier	Cumulative Production	424134.36	767225.5	9821152	8	25	18
piloting failure default recovery time	Cumulative Production	-166458.05	742145.4	5001318	71	26	33
researching impact on piloting	Cumulative Production	-421149.29	716403.4	7262513	80	27	28
aversion to NPV deviation	Cumulative Production	-95297.74	640858.4	8586875	67	28	24
researching impact on demoing	Cumulative Production	116157.92	569312.9	5919843	16	29	30
expected continuity of government policy	Cumulative Production	-219606.51	506331.9	5234267	73	30	32
...	...	...	...	...	...	...	...
incumbent reversion time	Cumulative Production	7276.5328	7406.2928	163707.13360	25.0	55.0	55.0
government research cost share	Cumulative Production	5132.7600	5132.7600	113327.62077	26.0	56.0	56.0
duration of intellectual property	Cumulative Production	-1860.1360	4547.9440	73166.19353	55.0	57.0	57.0
advertising start time	Cumulative Production	3185.8400	3185.8400	71237.54805	27.0	58.0	58.0
long term market size	Cumulative Production	2953.9368	2953.9368	56815.66402	28.0	59.0	59.0
bioproduct offtake agreement	Cumulative Production	-1284.4000	1284.4000	28720.05710	54.0	60.0	60.0
pathway approval maximum time	Cumulative Production	-516.5520	516.5520	10631.56392	53.0	61.0	61.0
incumbent long term price trend	Cumulative Production	111.5600	111.5600	2494.55744	30.0	62.0	62.0
incumbent starting price	Cumulative Production	27.0544	27.0544	604.95477	31.0	63.0	63.0
incumbent price response magnitude	Cumulative Production	-0.5192	0.5192	11.60966	52.0	64.0	64.0
advertising budget	Cumulative Production	0.0000	0.0000	0.00000	41.5	74.5	74.5
commercial plant startup period	Cumulative Production	0.0000	0.0000	0.00000	41.5	74.5	74.5
custom feedstock long term price change	Cumulative Production	0.0000	0.0000	0.00000	41.5	74.5	74.5

Input	Output	mu	mu*	sigma	mu rank	mu* rank	sigma rank
<chr>	<fct>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
custom feedstock maximum fluctuation magnitude	Cumulative Production	0.0000	0.0000	0.00000	41.5	74.5	74.5
custom feedstock periodic fluctuation magnitude	Cumulative Production	0.0000	0.0000	0.00000	41.5	74.5	74.5
custom feedstock reversion time	Cumulative Production	0.0000	0.0000	0.00000	41.5	74.5	74.5
custom feedstock starting price	Cumulative Production	0.0000	0.0000	0.00000	41.5	74.5	74.5
effectiveness of intellectual property	Cumulative Production	0.0000	0.0000	0.00000	41.5	74.5	74.5
elasticity of supply	Cumulative Production	0.0000	0.0000	0.00000	41.5	74.5	74.5
feedstock approval maximum cost	Cumulative Production	0.0000	0.0000	0.00000	41.5	74.5	74.5
government operating cost share	Cumulative Production	0.0000	0.0000	0.00000	41.5	74.5	74.5
government operating grant period	Cumulative Production	0.0000	0.0000	0.00000	41.5	74.5	74.5
incumbent market share target economic	Cumulative Production	0.0000	0.0000	0.00000	41.5	74.5	74.5
market growth rate	Cumulative Production	0.0000	0.0000	0.00000	41.5	74.5	74.5
number of missed stagegates allowed	Cumulative Production	0.0000	0.0000	0.00000	41.5	74.5	74.5
pathway approval maximum cost	Cumulative Production	0.0000	0.0000	0.00000	41.5	74.5	74.5
product approval maximum cost	Cumulative Production	0.0000	0.0000	0.00000	41.5	74.5	74.5
required internal return	Cumulative Production	0.0000	0.0000	0.00000	41.5	74.5	74.5
retrofit delay	Cumulative Production	0.0000	0.0000	0.00000	41.5	74.5	74.5
stagegate length	Cumulative Production	0.0000	0.0000	0.00000	41.5	74.5	74.5

Format results for a github issue.

In [22]:

```
options(repr.matrix.max.rows=100)
z.results[,
    .(
        `mu* minimum rank`    = min   (`mu* rank`),
        `mu* mean rank`       = mean  (`mu* rank`),
        `mu* median rank`     = median(`mu* rank`),
        `sigma minimum rank`  = min   (`sigma rank`),
        `sigma mean rank`     = mean  (`sigma rank`),
        `sigma median rank`   = median(`sigma rank`)
    ),
    by=.(Input)
][order(`mu* median rank`)][,
    .(
        text=mapply(
            function(x, y) paste(ifelse(x, "- [x]", "- [ ]"), y),
            `mu* median rank` <= 30 | `sigma median rank` <= 30,
            Input
        )

    )
]
```

A data.table: 84 x 1

text
<chr>
- [x] random stream
- [x] target pilot hours
- [x] base external investor ask rate
- [x] commercial plant capacity input
- [x] pilot capacity
- [x] commercial capital cost input
- [x] piloting acceptable rate
- [x] management response time
- [x] commercial variable operating cost input
- [x] payback period multiplier
- [x] minimum runway
- [x] pilot and demo response time
- [x] startup piloting rate
- [x] commercial process yield input
- [x] piloting failure default recovery time
- [x] bioproduct long term price
- [x] researching impact on piloting
- [x] government capital cost share
- [x] strategic value to external investors
- [x] piloting failure distribution max
- [x] startup piloting period
- [x] commercial plant capacity
- [x] expected continuity of government policy
- [x] government production incentive
- [x] aversion to NPV deviation
- [x] required return multiplier
- [ ] commercial fixed operating cost input
- [ ] bioproduct performance advantage
- [ ] target demo hours
- [ ] business as usual runway
- [ ] piloting management overhead
- [ ] demo capacity
- [ ] initial market size
- [ ] bioproduct price reversion time

text
<chr>
- [ ] regulatory process starting point
- [ ] pathway approval required switch
- [ ] researching impact on demoing
- [ ] investor optimism
- [ ] startup demoing rate
- [ ] demoing failure default recovery time
- [ ] bioproduct price fluctuations
- [ ] duration of intellectual property
- [ ] investor history
- [ ] product approval maximum time
- [ ] product approval required switch
- [ ] demoing acceptable rate
- [ ] demoing failure distribution max
- [ ] incumbent maximum fluctuation
- [ ] startup demoing period
- [ ] expected green premium
- [ ] elasticity of demand
- [ ] feedstock approval required switch
- [ ] incumbent reversion time
- [ ] long term market size
- [ ] max management runway response
- [ ] bioproduct offtake agreement
- [ ] government research cost share
- [ ] incumbent long term price trend
- [ ] advertising start time
- [ ] incumbent market share target economic
- [ ] incumbent price response magnitude
- [ ] incumbent starting price
- [ ] pathway approval maximum time
- [ ] demoing management overhead
- [ ] feedstock approval maximum time
- [ ] market growth rate
- [ ] advertising budget
- [ ] effectiveness of intellectual property
- [ ] retrofit delay

text
<chr>
- [ ] government operating cost share
- [ ] elasticity of supply
- [ ] commercial plant startup period
- [ ] custom feedstock long term price change
- [ ] custom feedstock maximum fluctuation magnitude
- [ ] custom feedstock periodic fluctuation magnitude
- [ ] custom feedstock reversion time
- [ ] custom feedstock starting price
- [ ] feedstock approval maximum cost
- [ ] government operating grant period
- [ ] number of missed stagegates allowed
- [ ] pathway approval maximum cost
- [ ] product approval maximum cost
- [ ] required internal return
- [ ] stagegate length