

Supplementary materials for the main submission entitled -  
 Detecting distributional differences between temporal granularities  
 for exploratory time series analysis

## Contents

<b>1 Recalling notations</b>	<b>1</b>
<b>2 Raw weighted pairwise distance</b>	<b>2</b>
2.1 Tuning parameter . . . . .	2
2.2 Underlying distributions . . . . .	2
2.3 Number of comparisons . . . . .	3
<b>3 Adjusted weighted pairwise distances</b>	<b>3</b>
3.1 Permutation approach . . . . .	4
3.2 Modeling approach . . . . .	4
3.3 Combination approach . . . . .	5
<b>4 Ranking and selecting harmonies</b>	<b>5</b>
4.1 Size . . . . .	5
4.2 Power . . . . .	5

## 1 Recalling notations

Let  $v = \{v_t : t = 0, 1, 2, \dots, T - 1\}$  be a continuous measured variable observed across  $T$  time points. The number of cyclic granularities considered in the display (at once) is  $m$ . Consider cyclic granularities  $A$  and  $B$ , such that  $A = \{a_j : j = 1, 2, \dots, J\}$  and  $B = \{b_k : k = 1, 2, \dots, K\}$ . For  $m = 1$ , only one of  $A$  or  $B$  is plotted at once and a *panel* refers to a display of distributions of  $v$  across  $J$  or  $K$  levels on the x-axis. For  $m = 2$ , the distribution display of  $v$  with  $A$  placed across x-axis and  $B$  across facets is referred to as a  $(J, K)$  panel ( $J$  x-axis levels and  $K$  facet levels). The pairwise distances between pairs  $(a_j b_k, a'_j b'_k)$  could be within-facets or between-facets as seen in Figure 4 of the main paper. The tuning parameter, used to put relative weight-age to the pairwise distances within and between facets is denoted by  $\lambda$ . Let the four elementary designs be  $D_{null}$  where there is no pairwise difference in distribution of  $v$  across  $A$  or  $B$ ,  $D_{var_f}$  denotes the set of designs where there is difference in distribution of  $v$  for  $B$  and not for  $A$ . Similarly,  $D_{var_x}$  denotes the set of designs where difference is observed only across  $A$ . Finally,  $D_{var_{all}}$  denotes those designs for which difference is observed across both  $A$  and  $B$ . The following method is deployed for generating different distributions across different combinations for non-null designs - suppose the distribution of the combination of first levels of x and facet category is  $N(\mu, \sigma)$  and  $\mu_{jk}$  denotes the mean of the combination  $(a_j b_k)$ , then  $\mu_{j.} = \mu + j\omega$  (for design  $D_{var_x}$ ) and  $\mu_{.k} = \mu + k\omega$  (for design  $D_{var_f}$ ), where  $\omega$  denotes the increment in mean. Table 1 shows an example of how designs are defined for a  $(2, 3)$  panel using  $\omega = 3$ .  $nx$  and  $nfacet$  denotes the number of categories placed on x-axis and facets respectively.  $wpd_{raw}$  and  $wpd$  denote the raw and normalized weighted pairwise distances.

Table 1: Simulation setup for a panel with 3 facet levels and 2 x-axis levels for different designs starting from an initial distribution  $N(0, 1)$  for the combination  $(a_1, b_1)$  and  $\omega = 3$ .

x levels	facet levels	$D_{null}$	$D_{var_f}$	$D_{var_x}$	$D_{var_{all}}$
$a_1$	$b_1$	$N(0, 1)$	$N(0, 1)$	$N(0, 1)$	$N(0, 1)$
$a_2$	$b_1$	$N(0, 1)$	$N(0, 1)$	$N(3, 1)$	$N(3, 1)$
$a_1$	$b_2$	$N(0, 1)$	$N(3, 1)$	$N(0, 1)$	$N(6, 1)$
$a_2$	$b_2$	$N(0, 1)$	$N(3, 1)$	$N(3, 1)$	$N(9, 1)$
$a_1$	$b_3$	$N(0, 1)$	$N(6, 1)$	$N(0, 1)$	$N(12, 1)$
$a_2$	$b_3$	$N(0, 1)$	$N(6, 1)$	$N(3, 1)$	$N(15, 1)$

## 2 Raw weighted pairwise distance

### 2.1 Tuning parameter

For  $m = 1$ , pairwise distances could be defined only between different categories of the cyclic granularity considered. So no tuning parameter is applicable for this case. For  $m = 2$ ,  $\lambda$  might impact  $wpd_{raw}$  differently depending on the value of  $\omega$  and different values of  $nx$ ,  $nfacet$  and designs. The following simulation study sees the impact of  $\lambda$  for different  $\omega$ ,  $nx$ ,  $nfacet$  and designs.

#### Simulation design

Observations are generated from  $N(0,1)$  distribution for each combination of  $nx$  and  $nfacet$  from the following sets:  $nx = nfacet = \{2, 3, 5, 7\}$  and  $wpd_{raw}$  is computed for  $\lambda = 0.1, 0.2, \dots, 0.9$  under two designs  $D_{var_x}$  and  $D_{var_f}$  for  $\omega = \{1, 8\}$  to observe how the value of  $wpd_{raw}$  changes for different designs. Moreover, to observe for which value of  $\lambda$  the two designs intersect, we generate observations from  $N(0,1)$  distribution for each combination of  $nx$  and  $nfacet$  from the following sets:  $nx = nfacet = \{2, 3, 5, 7, 14, 20\}$  and  $\omega = \{1, 2, \dots, 10\}$  under the same two designs.

#### Results

Figure 1 shows how the value of  $wpd$  changes for  $\lambda = 0.1, 0.2, \dots, 0.9$  for the two different designs  $D_{var_x}$  and  $D_{var_f}$  for two values of increment in mean  $\omega = 1, 8$ . For a lower value of  $\omega$ , the two designs intersect at  $\lambda > 0.7$  and for a higher  $\omega$ , the two designs intersect at  $\lambda = 0.5$ . The value of  $wpd$  increases with  $\lambda$  for  $D_{var_x}$  and decreases with increasing  $\lambda$  for  $D_{var_f}$ . Figure 2 shows the value of  $\lambda$  for which the two designs intersect for different values of  $\omega$ . It can be observed that as the value of  $\omega$  ( $\omega > 4$ ) increase, the value of  $\lambda$  at which the two designs intersect converge is  $\lambda = 0.5$  and for  $\omega$  ( $\omega \leq 4$ ), designs interact between  $\lambda = [0.6, 0.75]$ . Hence,  $\lambda = 0.67$  for the rest of the paper, which is an average of the range of values  $\lambda$  can take.

### 2.2 Underlying distributions

The following simulation study sees the impact of different underlying distributions on  $wpd_{raw}$  for different  $nx$  and  $nfacet$  before and after performing Normal Quantile Transformation.

#### Simulation design

Observations are generated from  $N(0,1)$ ,  $N(5, 1)$ ,  $N(0, 5)$ ,  $\text{Gamma}(0.5, 1)$  and  $\text{Gamma}(2, 1)$  distributions for each combination of  $nx$  and  $nfacet$  from the following sets:  $nx = nfacet = \{2, 3, 5, 7, 14, 20, 31, 50\}$  and  $wpd_{raw}$  is computed with  $\lambda = 0.67$  with (scenario 1) and without (scenario 1) NQT.

#### Results

Figure 3 shows ridge plots of  $wpd_{raw}$  for a  $\text{Gamma}(0.5,1)$ ,  $\text{Gamma}(2,1)$  before NQT. It is observed that for the underlying distribution  $\text{Gamma}(2,1)$ , location and scale of the distribution of wpd changes from top-left panel to bottom-right panel. Figure 4 shows the the distributions of  $wpd$  under same underlying distributions after performing NQT. It is observed that within each panel, the distributions of the  $wpd$  looks same, however,

the distributions change from extreme top-left panel to bottom-right panels. Similar observations could be made in Figure 5 for different underlying normal distributions  $N(0,1)$ ,  $N(5,1)$  and  $N(0,5)$ . This implies, NQT has atleast been able to make the location and scale of the distribution of  $wpd_{raw}$  same for different underlying distributions.

## 2.3 Number of comparisons

### 2.3.1 Case: $m = 1$

#### *Simulation design*

Observations are generated from a  $N(0,1)$  distribution for each  $nx = \{2, 3, 5, 7, 9, 14, 17, 20, 24, 31, 42, 50\}$  to cover a wide range of levels from very low to moderately high.  $ntimes = 500$  observations are drawn for each combination of the categories, that is, for a panel with  $nx = 3$ , 500 observations are simulated for each of the categories. This design corresponds to  $D_{null}$  as each combination of categories in a panel are drawn from the same distribution. Furthermore, the data is simulated for each of the categories  $nsim = 200$  times, so that the distribution of  $wpd$  under  $D_{null}$  could be observed. The values of  $wpd$  is obtained for each of the panels.  $wpd_{l,s}$  denotes the value of  $wpd$  obtained for the  $l^{th}$  panel and  $s^{th}$  simulation.

#### *Results*

Figure 6 shows ridge plots of  $wpd_{raw}$  for an underlying  $N(0,1)$  distribution. For each panel, it could be seen that the location shifts to the right for increasing x levels. Across each panel, the scale of the distribution seems to change for low/moderately from lower values to higher values of  $nx$  and left tails are longer for lower facet levels.

### 2.3.2 Case: $m = 2$

#### *Simulation design*

Similarly, observations are generated from a  $N(0,1)$  distribution for each combination of  $nx$  and  $nfacet$  from the following sets:  $nx = nfacet = \{2, 3, 5, 7, 14, 20, 31, 50\}$ . That is, data is being generated for each of the panels  $(2, 2), (2, 3), (2, 5) \dots, (50, 31), (50, 50)$ . For each of the 64 panels,  $ntimes = 500$  observations are drawn for each combination of the categories. That is, if we consider a  $(2, 2)$  panel, 500 observations are generated for each of the possible subsets, namely,  $\{(1, 1), (1, 2), (2, 1), (2, 2)\}$ .

Observations are generated from  $\text{Gamma}(2,1)$ ,  $G(0.5, 1)$ ,  $N(0,1)$ ,  $N(0, 5)$  and  $N(5, 1)$  distribution for each combination of  $nx$  and  $nfacet$  from the following sets:  $nx = nfacet = \{2, 3, 5, 7, 14, 20, 31, 50\}$  to cover a wide range of levels from very low to moderately high. Each combination is being referred to as a *panel*. That is, data is being generated for each of the panels  $\{nx = 2, nfacet = 2\}, \{nx = 2, nfacet = 3\}, \{nx = 2, nfacet = 5\}, \dots, \{nx = 50, nfacet = 31\}, \{nx = 50, nfacet = 50\}$ . For each of the 64 panels,  $ntimes = 500$  observations are drawn for each combination of the categories. That is, if we consider the panel  $\{nx = 2, nfacet = 2\}$ , 500 observations are generated for each of the combination of categories from the panel, namely,  $\{(1, 1), (1, 2), (2, 1), (2, 2)\}$ . The values of  $wpd$  is obtained for each of the panels. The measurement variable for each combination of categories in a panel are drawn from the same distribution and hence the design corresponds to  $D_{null}$ . Furthermore, this entire method is repeated for each panels  $nsim = 200$  times, so that the distribution of  $wpd$  under  $D_{null}$  could be observed.

#### *Results*

Figure 8, 9 and 10 shows the ridge plot of  $wpd_{raw}$  with  $nx$  as facets,  $nfacet$  as facets and the density plot of  $wpd_{raw}$  with  $nx$  on the x-axis and  $nfacet$  on the facets.

## 3 Adjusted weighted pairwise distances

#### *Simulation design*

Table 2: Results of generalised linear model to capture the relationship between  $wpd_{raw}$  and number of comparisons for  $m = 1$ .

distribution	term	estimate	std.error	statistic	p.value
N(0, 1)	(Intercept)	26.086352	0.5397440	48.330973	0.0e+00
N(0, 1)	log('nx * nfacet')	-1.874597	0.1894537	-9.894751	1.8e-06

Table 3: Results of generalised linear model to capture the relationship between  $wpd_{raw}$  and number of comparisons for different underlying distribution.

distribution	term	estimate	std.error	statistic	p.value
Gamma(0.5, 1)	(Intercept)	23.67	0.242	97.95	0
Gamma(0.5, 1)	log('nx * nfacet')	-1.01	0.049	-20.88	0
Gamma(2, 1)	(Intercept)	23.69	0.240	98.77	0
Gamma(2, 1)	log('nx * nfacet')	-1.02	0.048	-21.24	0
N(0, 1)	(Intercept)	23.40	0.225	104.14	0
N(0, 1)	log('nx * nfacet')	-0.96	0.044	-21.75	0
N(0, 5)	(Intercept)	23.56	0.221	106.71	0
N(0, 5)	log('nx * nfacet')	-1.00	0.044	-22.45	0
N(5, 1)	(Intercept)	23.56	0.221	106.71	0
N(5, 1)	log('nx * nfacet')	-1.00	0.044	-22.45	0
N(5, 5)	(Intercept)	23.56	0.221	106.71	0
N(5, 5)	log('nx * nfacet')	-1.00	0.044	-22.45	0

This sections shows the result from the different approaches of adjusting for the number of comparisons. Simulation design is the same as for raw weighted pairwise distances for  $m = 1$  and  $m = 2$ . In place of  $wpd_{raw}$ ,  $wpd_{perm}$ ,  $wpd_{glm}$  and  $wpd$  is computed for all panel.

### 3.1 Permutation approach

Figure 11 and 12 show the normalized  $wpd$  for  $m = 1$  and  $m = 2$  respectively. The location and scale are brought to similar values for increasing values of  $nx$  or  $nfacet$ . Due to heavy computational load,  $wpd_{perm}$  for few panels with very high value of  $nx$  or  $nfacet$  are not computed.

### 3.2 Modeling approach

#### 3.2.1 Case: $m = 1$

Figure ?? shows the scatterplot of  $wpd_{raw}$  against different values of  $nx$  and also the display of residuals from the model. Figure 14 shows the distribution of  $wpd_{glm}$  for different  $nx$ .

```
#> mapping: x = ~label_glm, y = ~as.factor(nx), label = ~label_glm
#> geom_label: parse = FALSE, label.padding = 0.25, label.r = 0.15, label.size = 0.25, na.rm = FALSE
#> stat_identity: na.rm = FALSE
#> position_identity
```

#### 3.2.2 Case: $m = 2$

Figure ?? shows the scatterplot of  $wpd_{raw}$  against different values of  $nx$  and also the display of residuals from the model. Figure 16 shows the distribution of  $wpd_{glm}$  for different  $nx$ .

```
#> [1] 0.9886192
```

### 3.3 Combination approach

#### 3.3.1 Simulation design

Observations are generated from  $\text{Gamma}(2,1)$ ,  $\text{G}(0.5, 1)$ ,  $\text{N}(0,1)$ ,  $\text{N}(0, 5)$  and  $\text{N}(5, 1)$  distribution for each combination of  $nx$  and  $nfacet$  from the following sets:  $nx = nfacet = \{2, 3, 5, 7, 14, 20, 31, 50\}$  to cover a wide range of levels from very low to moderately high. Each combination is being referred to as a *panel*. That is, data is being generated for each of the panels  $\{nx = 2, nfacet = 2\}, \{nx = 2, nfacet = 3\}, \{nx = 2, nfacet = 5\}, \dots, \{nx = 50, nfacet = 31\}, \{nx = 50, nfacet = 50\}$ . For each of the 64 panels,  $ntimes = 500$  observations are drawn for each combination of the categories. That is, if we consider the panel  $\{nx = 2, nfacet = 2\}$ , 500 observations are generated for each of the combination of categories from the panel, namely,  $\{(1, 1), (1, 2), (2, 1), (2, 2)\}$ . The values of  $wpd$  is obtained for each of the panels. The measurement variable for each combination of categories in a panel are drawn from the same distribution and hence the design corresponds to  $D_{null}$ . Furthermore, this entire method is repeated for each panels  $nsim = 200$  times, so that the distribution of  $wpd$  under  $D_{null}$  could be observed.

## 4 Ranking and selecting harmonies

### 4.1 Size

#### *Simulation design*

Observations are generated from a  $\text{N}(0,1)$  distribution for each combination of  $nx$  and  $nfacet$  from the following sets:  $nx = \{3, 7, 14\}$  and  $nfacet = \{2, 9, 10\}$ . This would result in 9 panels, viz,  $(3, 2), (3, 9), (3, 10), \dots, (14, 9), (14, 10)$ . Few experiments were conducted. In the first scenario, data for all panels are simulated using the null design  $D_{null}$ . In other scenarios, data simulated from the panel  $(14, 2)$  and  $(3, 10)$  are under  $D_{vary_{all}}$ . Moreover,  $\omega = \{0.5, 2, 5\}$  are considered to examine if the proposed test is able to capture subtle differences and non-subtle differences when we shift from the null design. In the last scenario, we consider the panel  $(3, 2), (7, 9), (14, 10)$  to be under  $D_{null}$ , the panels  $(7, 2), (14, 9)$  to be under  $D_{var_f}$ ,  $(14, 2), (3, 10)$  under  $D_{var_x}$  and the rest under  $D_{var_{null}}$ . This is done to check if the consequent ranking procedure leads to designs like  $D_{vary_{all}}$  to be chosen first followed by  $D_{vary_{all}}$ . We generate only one data set each for which these scenarios were simulated and consider this as the original data set. We generate 1000 repetitions of this experiment with different seeds.

#### *Results*

### 4.2 Power

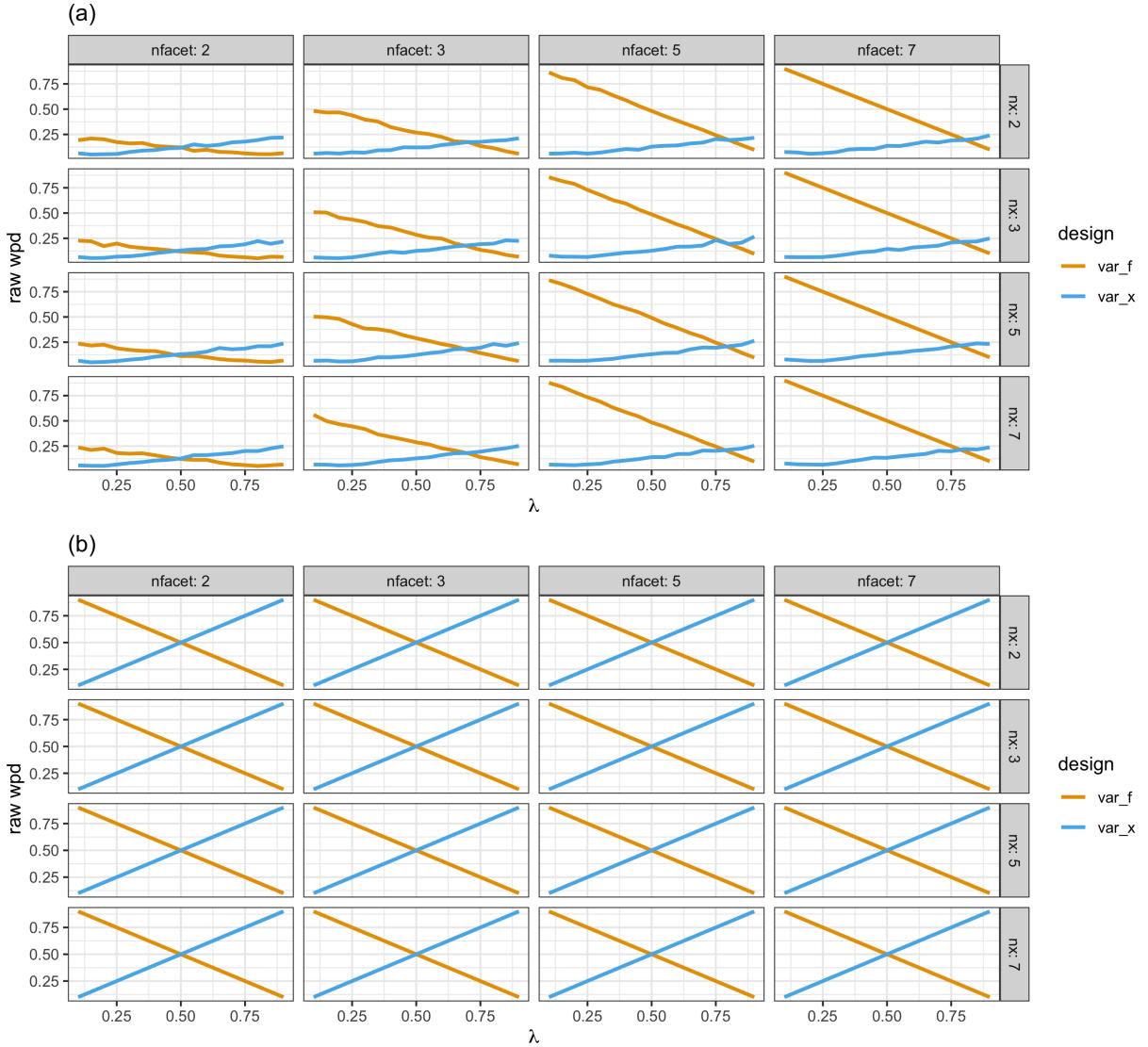


Figure 1:  $wpd_{raw}$  from two designs are plotted for different values of  $\lambda$   $nx$  and  $nfacet$ . For  $\omega = 1$ , the designs intersect for  $0.6 < \lambda \leq 0.75$ , whereas for higher omega, design intersects at  $\lambda = 0.5$ .

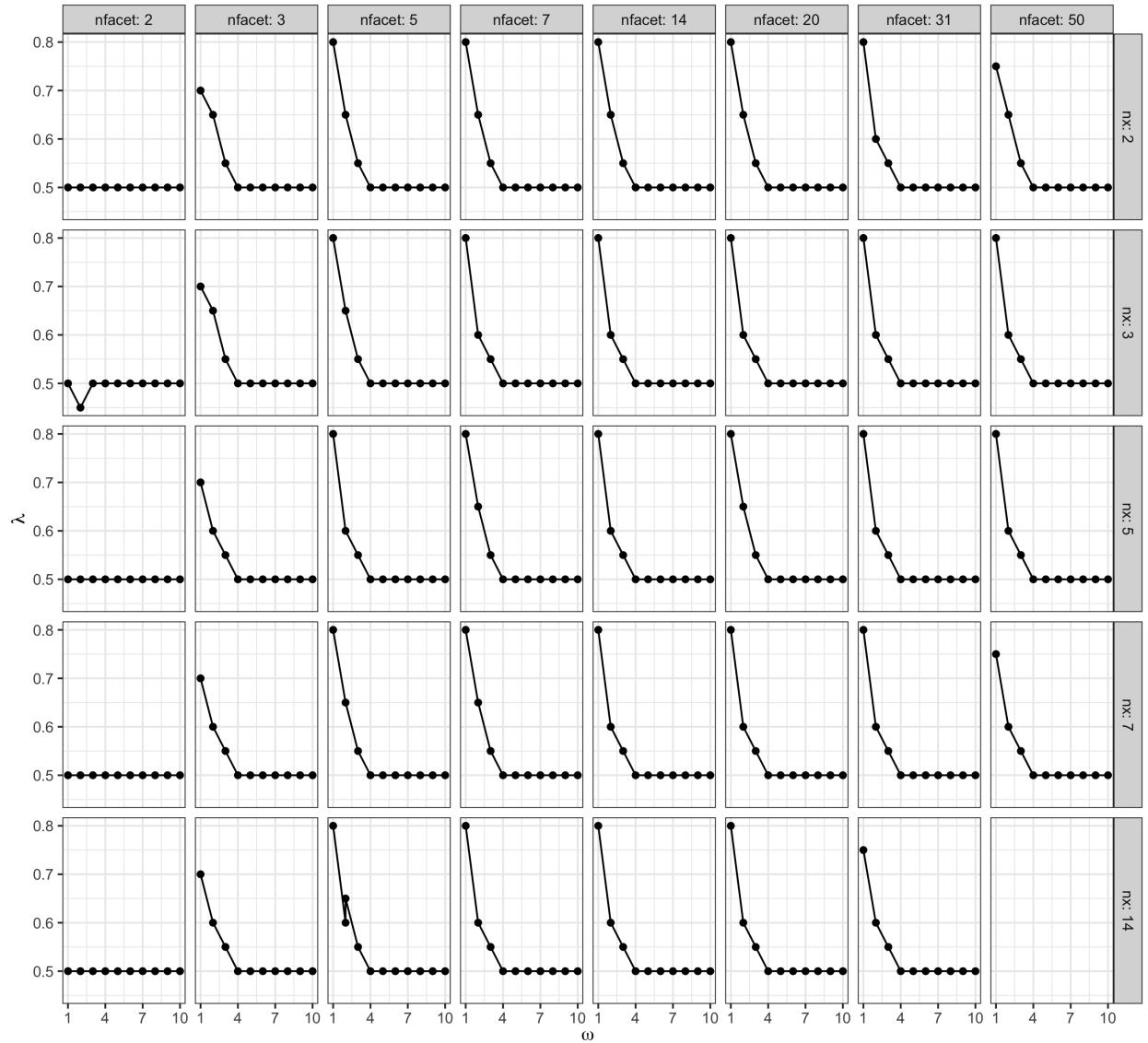


Figure 2: The point of intersection of  $wpd_{raw}$  values under the designs  $D_{var_f}$  and  $D_{var_x}$  are plotted across different  $\omega$  and  $\lambda$ . For most panels it is observed that a common value of  $\lambda$  for which the designs interact is 0.5 for  $\omega \geq 4$ , which implies any value greater than 0.5 could be chosen to up-weight the within-facet distances and down-weight the between-facet distances. The value of  $\lambda$  is higher for  $\omega < 4$  could be anywhere between 0.6 and 0.75.

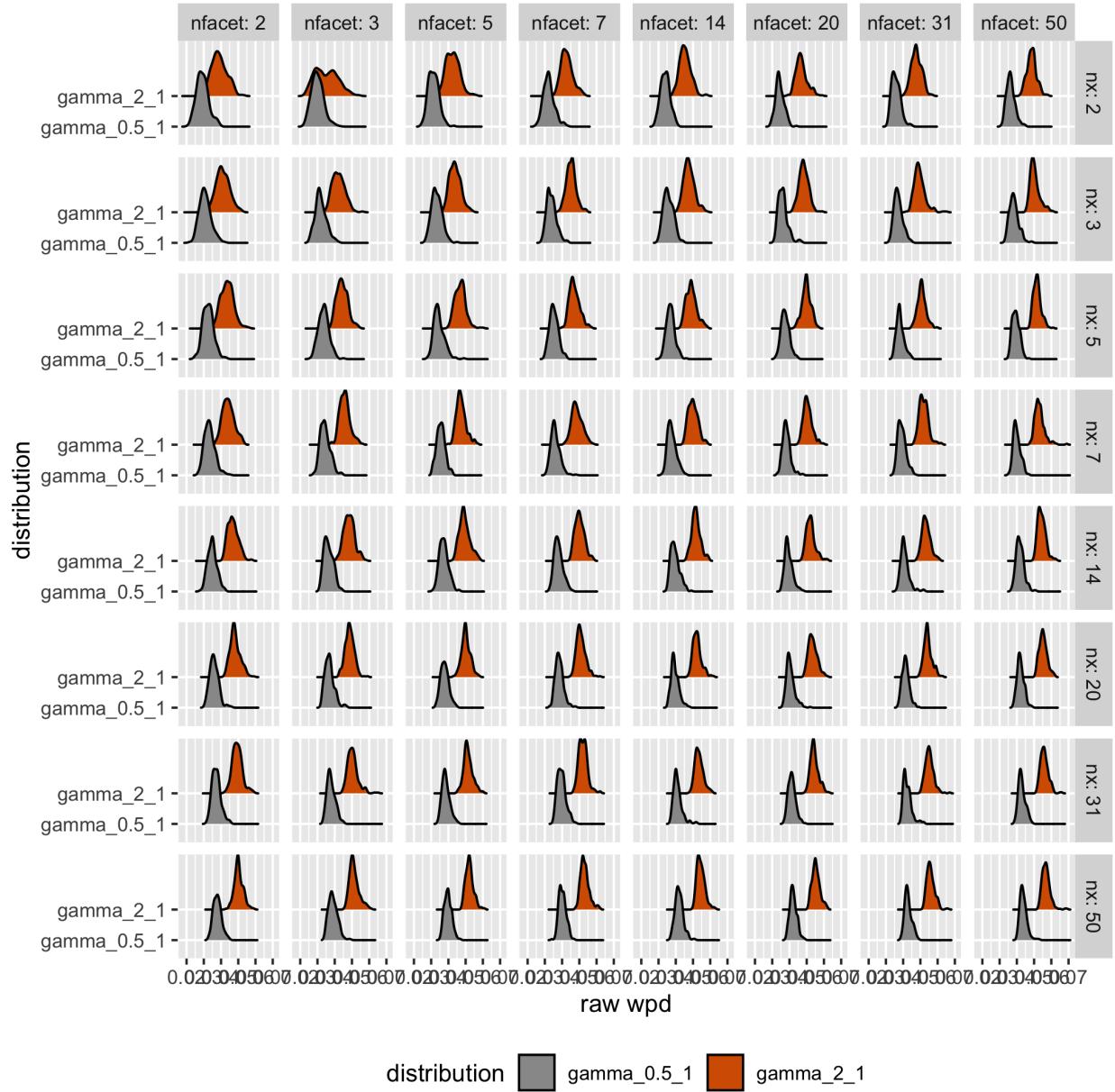


Figure 3: Ridge plots of raw wpd is shown for Gamma(0.5,1), Gamma(2,1) distribution without NQT. The densities change across different facet and x levels and also looks different for the two distributions, which implies wpd value is affected by the change in the shape parameter of the gamma distribution.

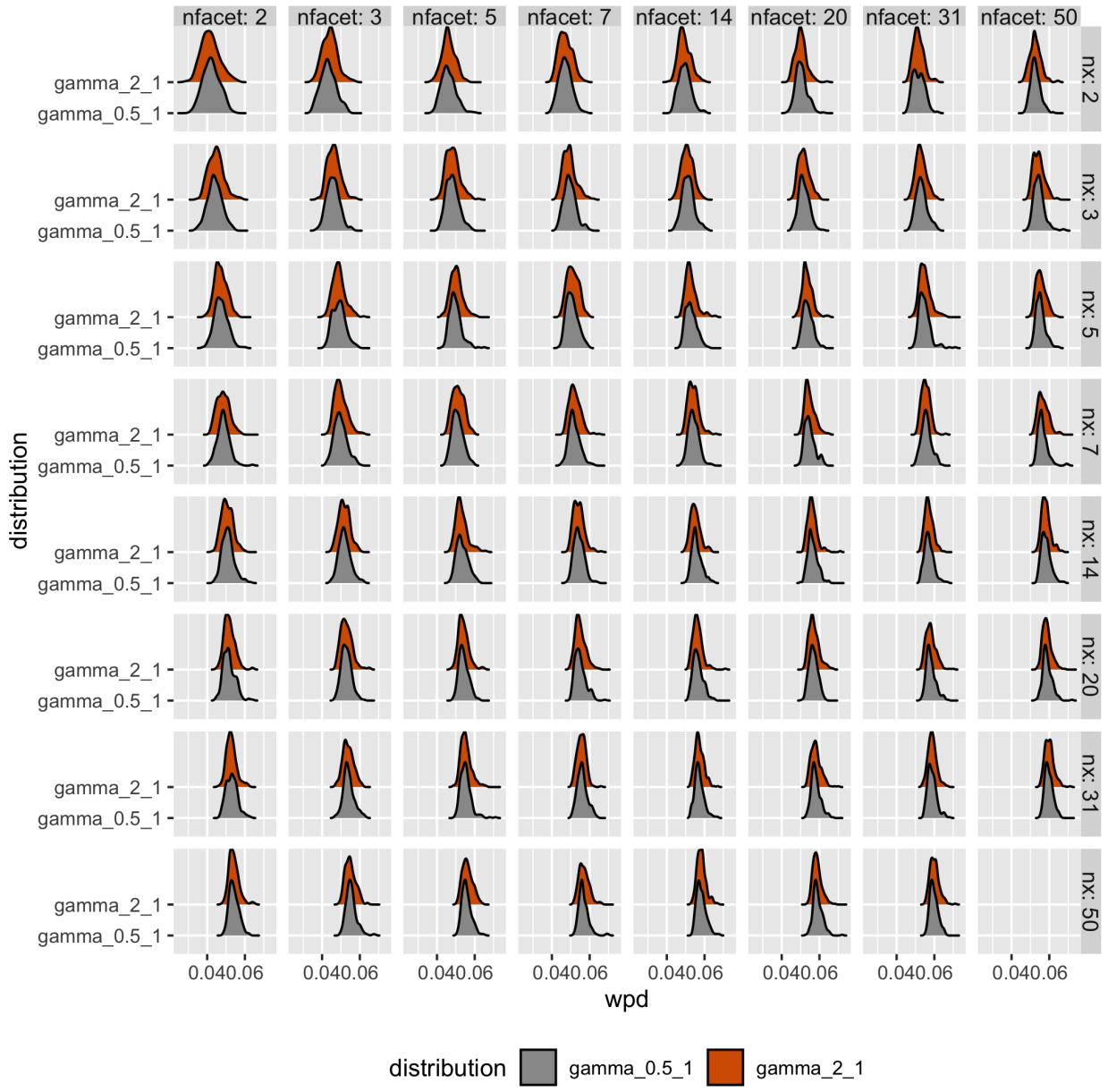


Figure 4: Ridge plots of raw wpd is shown for Gamma(0.5,1), Gamma(2,1) distribution. The densities change across different facet and x levels but look same for the two distributions, which implies wpd value is unaffected by the change in the shape parameter of the gamma distribution

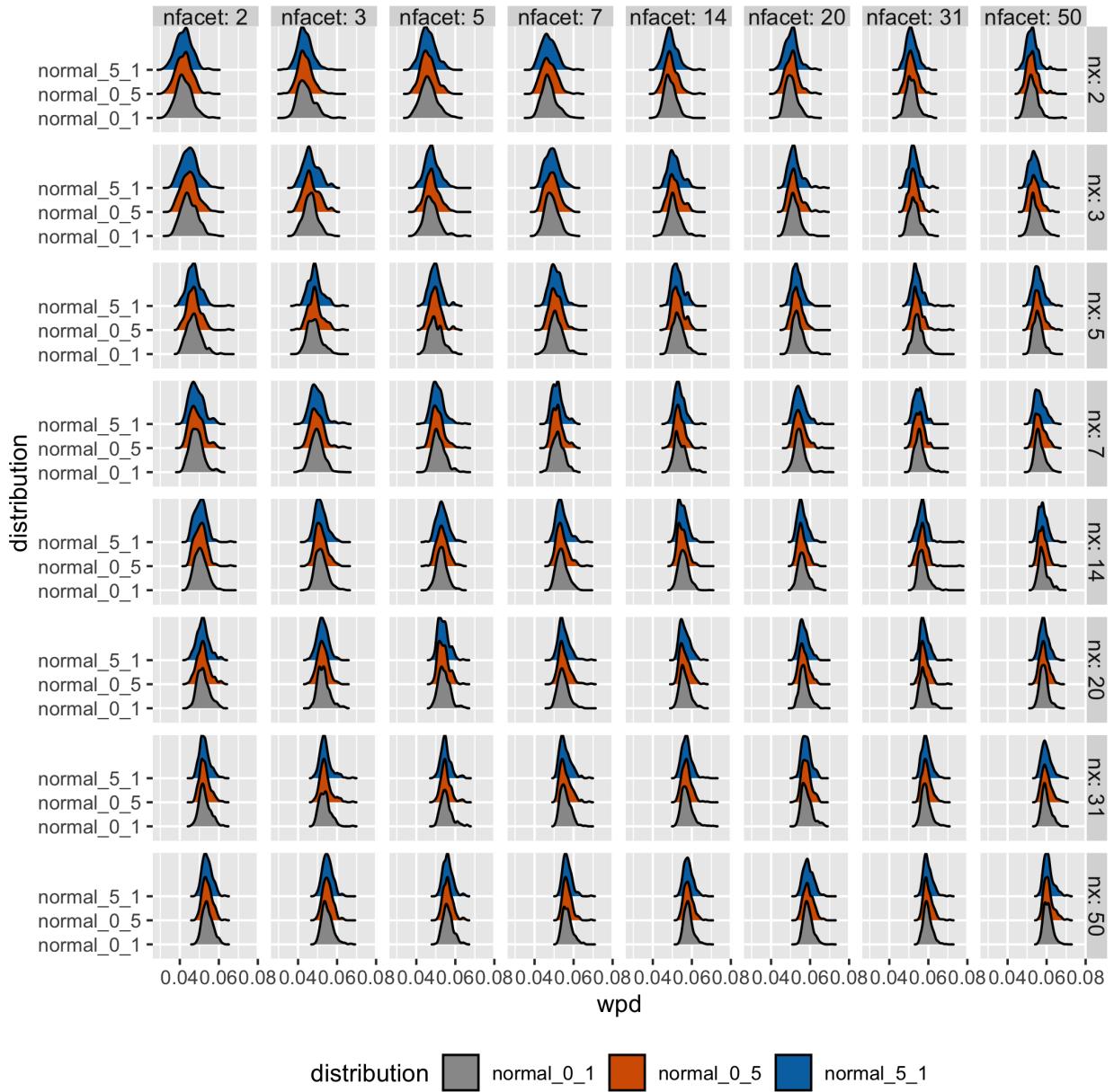


Figure 5: Ridge plots of raw wpd is shown for  $N(0,1)$ ,  $N(5,1)$  and  $N(0,5)$  distribution. The densities change across different facet and x levels but look same for each panel, which implies wpd value is unaffected by the change in mean and standard deviation of the normal distribution.

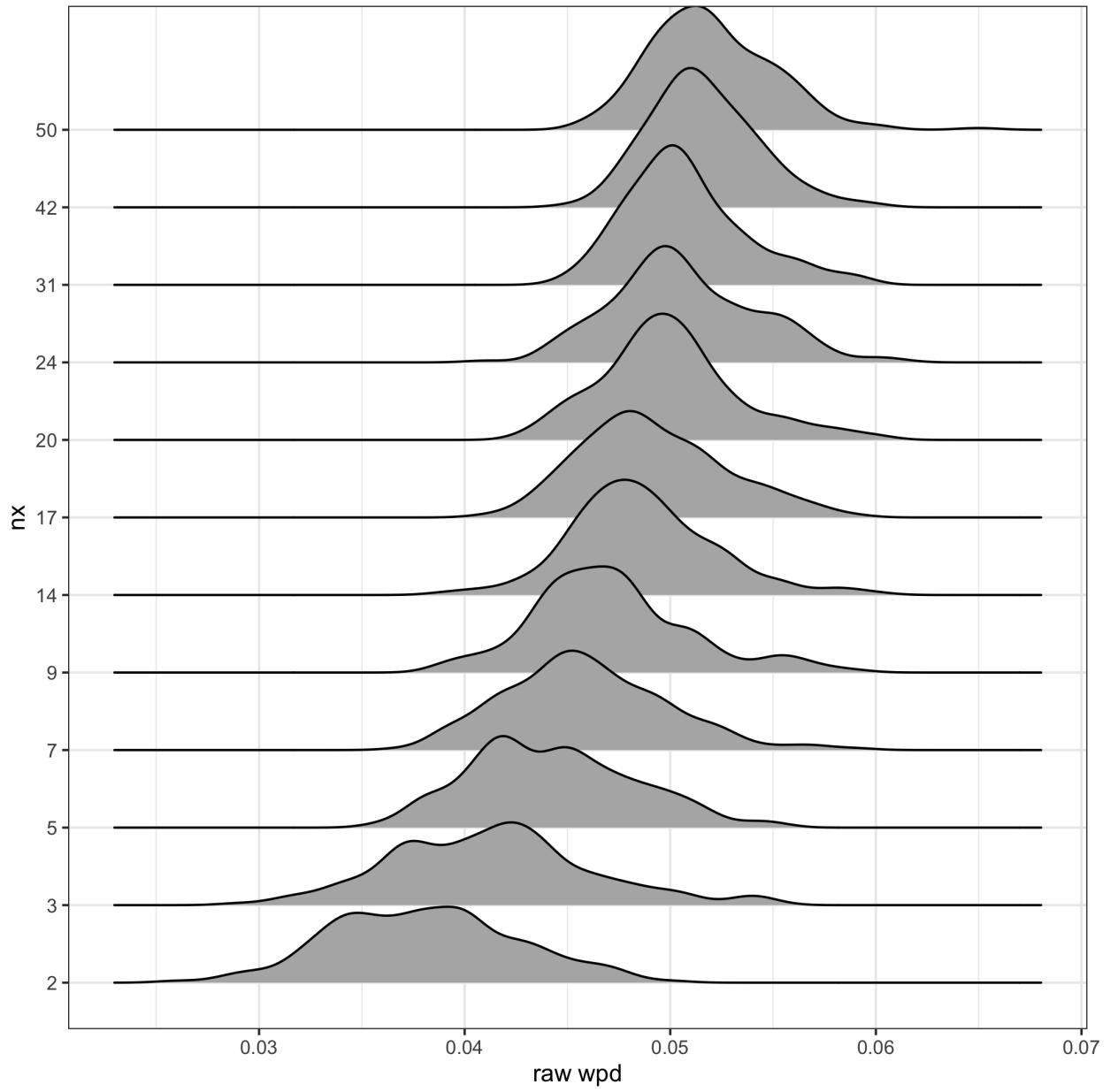


Figure 6: Ridge plots of raw wpd is shown for  $N(0,1)$  distribution. For each panel, it could be seen that the location shifts to the right for increasing  $x$  levels. Across each panel, the scale of the distribution seems to change for low/moderately lower values and higher values of nfacets and left tails are longer for lower facet levels.

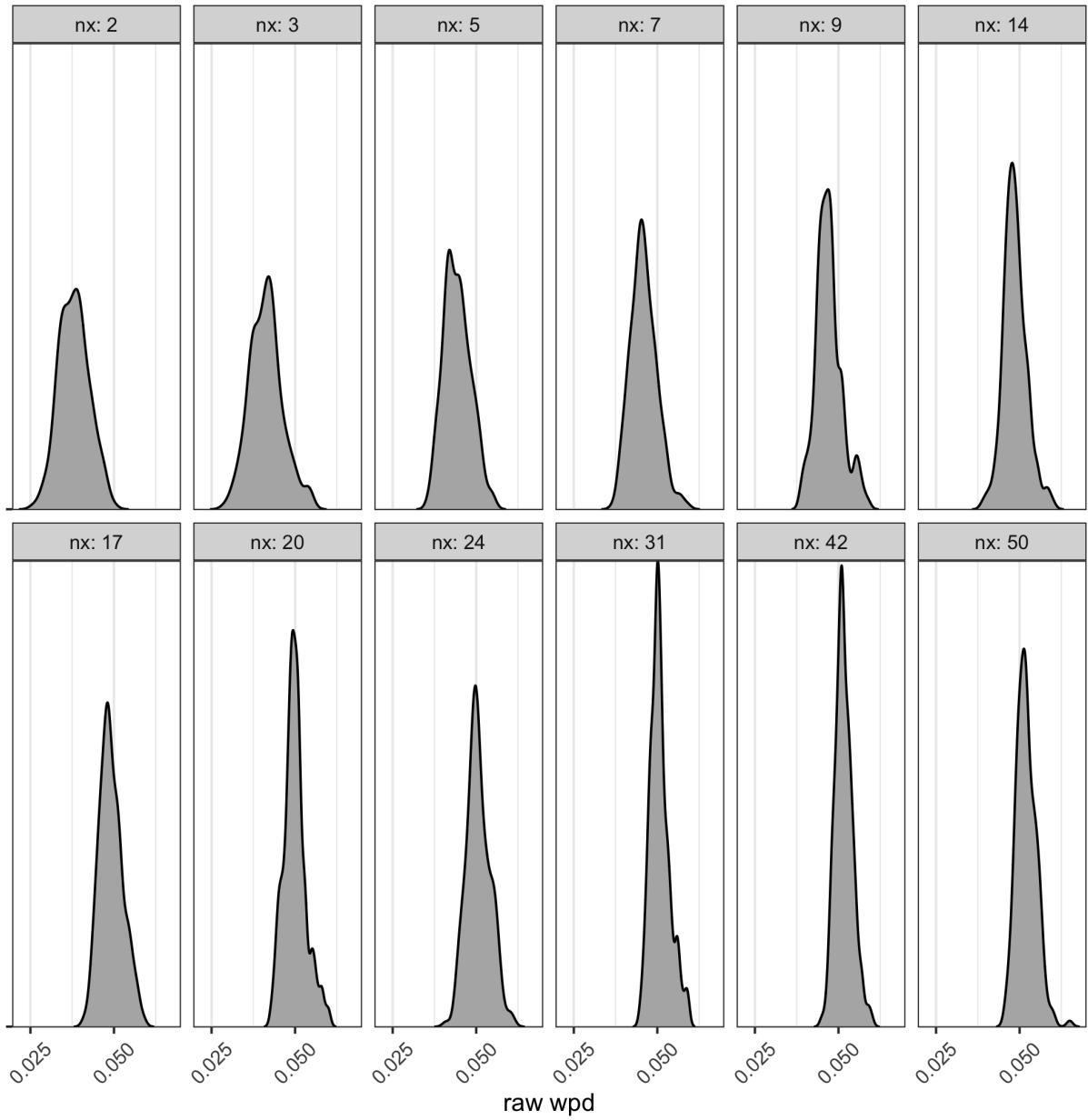


Figure 7: Ridge plots of raw wpd is shown for  $N(0,1)$  distribution. For each panel, it could be seen that the peakedness shifts for increasing  $x$  levels.

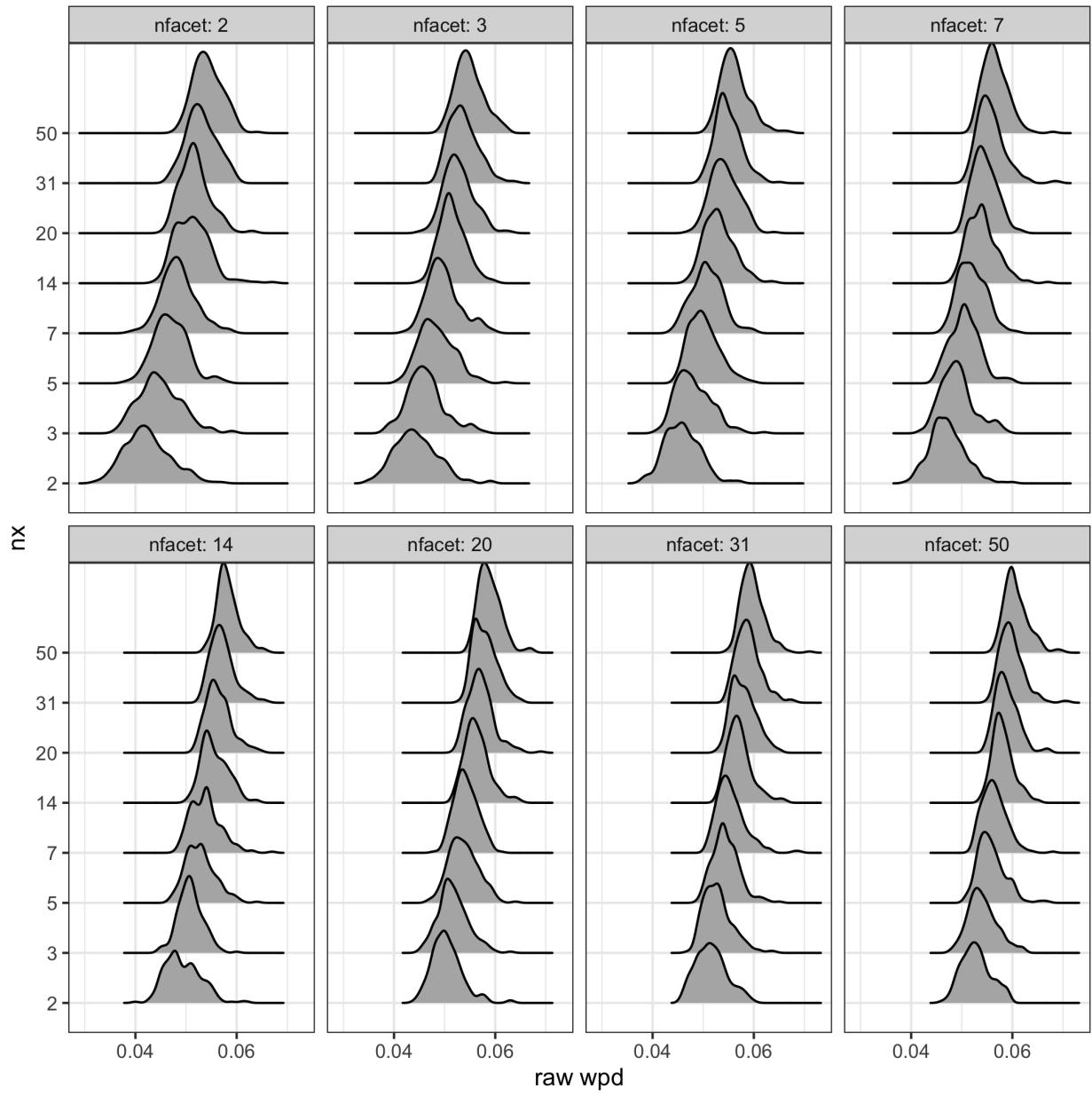


Figure 8: Ridge plots of raw wpd is shown for  $N(0,5)$  distribution. For each panel, it could be seen that the location shifts to the right for increasing  $x$  levels. Across each panel, the scale of the distribution seems to change for low/moderately lower values and higher values of nfacets and left tails are longer for lower facet levels.

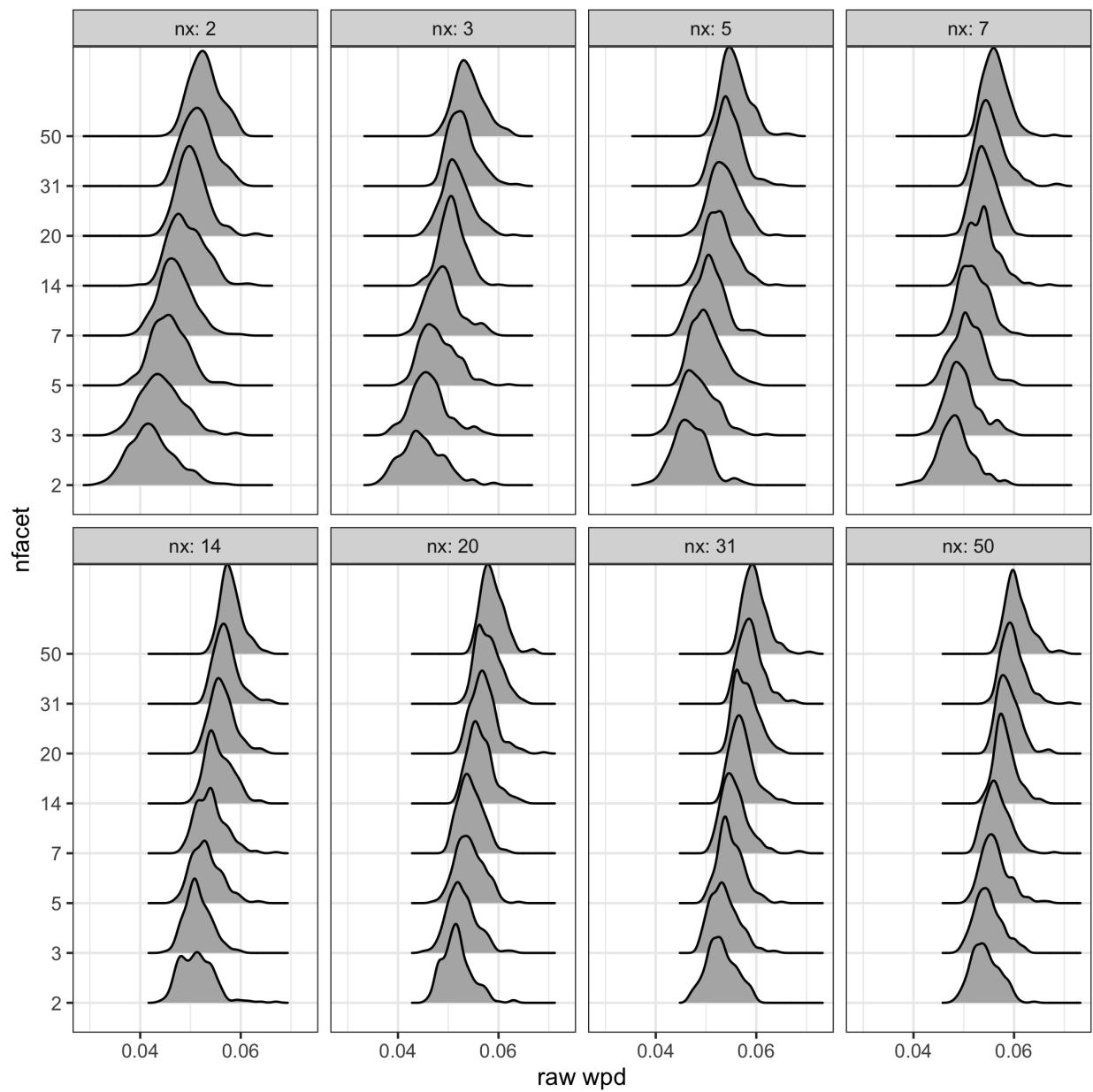


Figure 9: something1

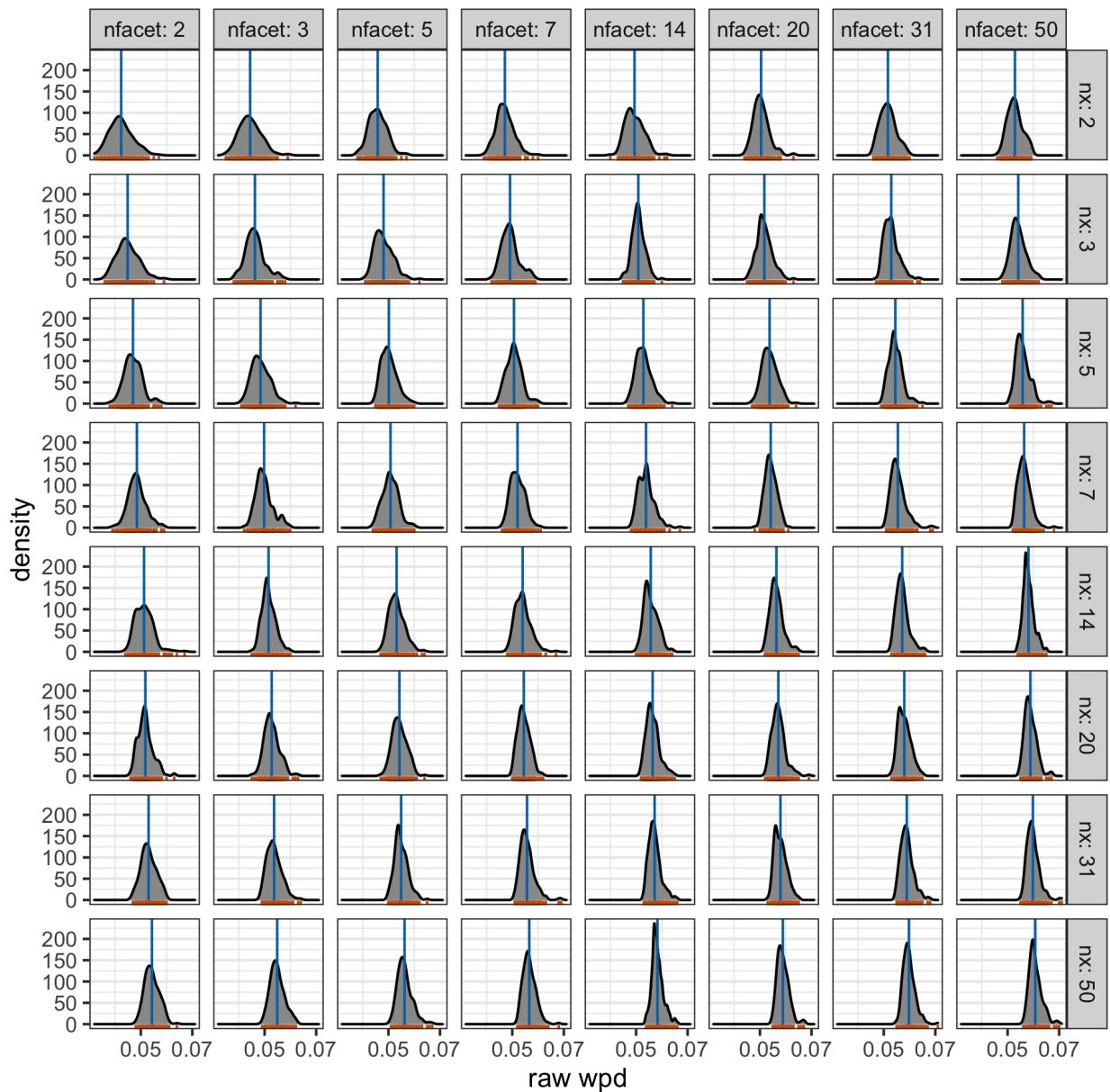


Figure 10: something2

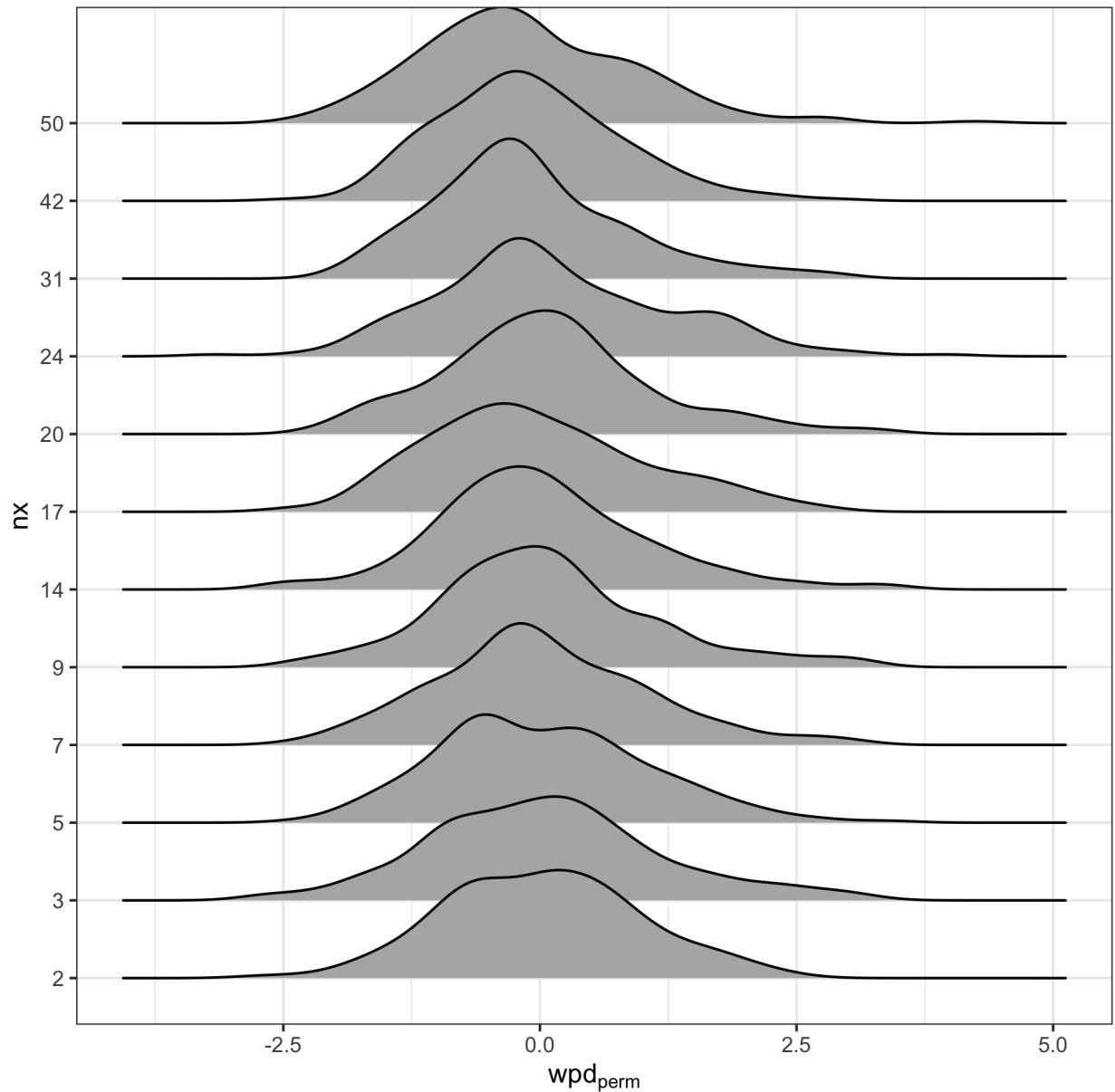


Figure 11: Distribution of  $wpd_{perm}$  are shown for different levels for  $m = 1$  under the null design. The distribution for different panels have similar location and scale.

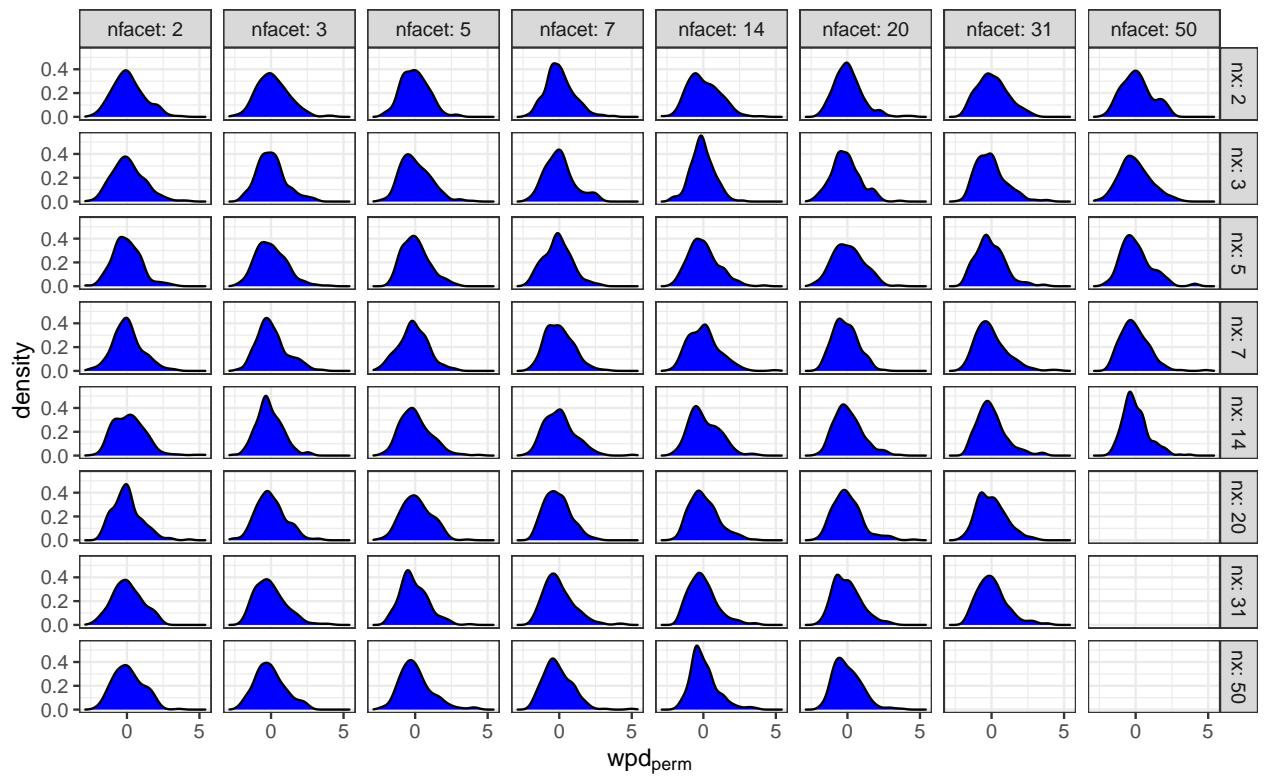


Figure 12: Distribution of  $wpd_{perm}$  are shown for different levels for  $m = 2$  under the null design. The distribution for different panels have similar location and scale.

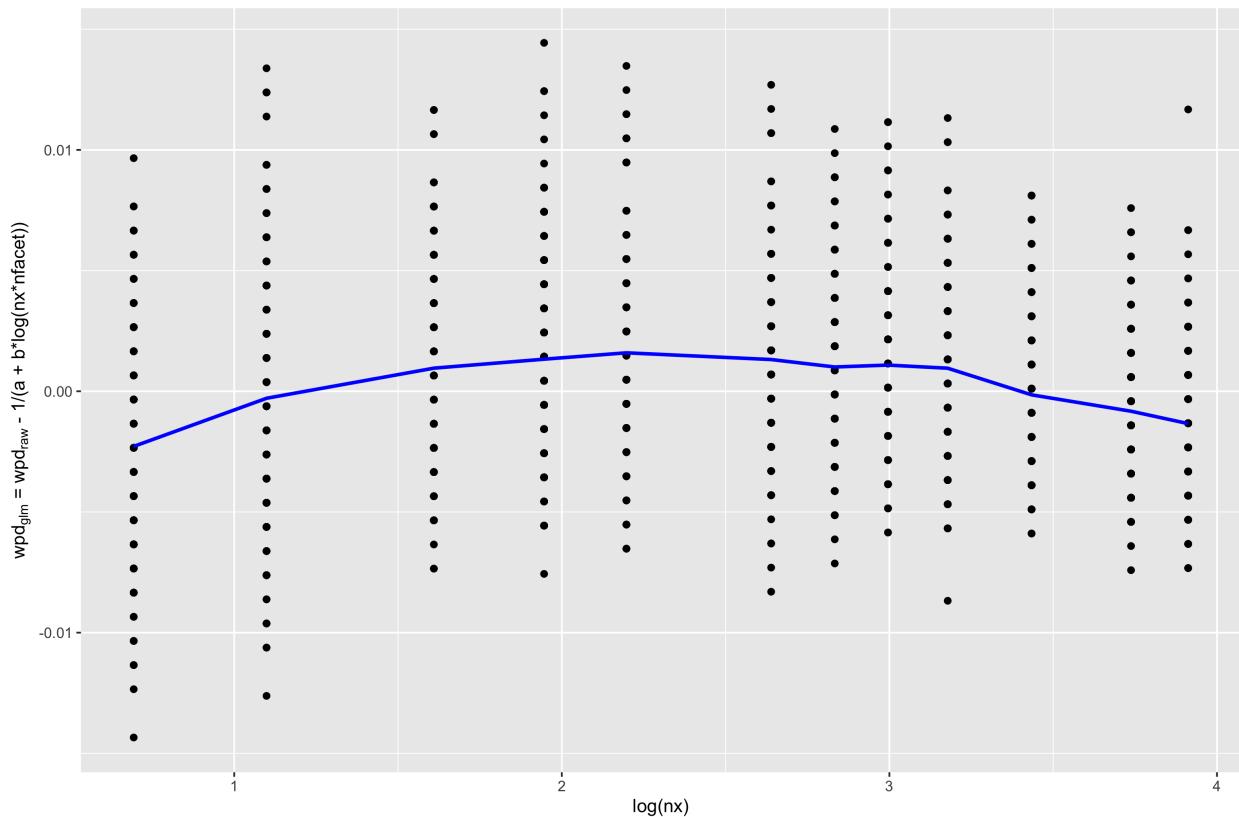


Figure 13: Residuals from the model is plotted across different number of comparisons for  $m = 1$ . Residuals seem to be independent of  $nx$  and have been defined as  $wpd_{glm}$ .

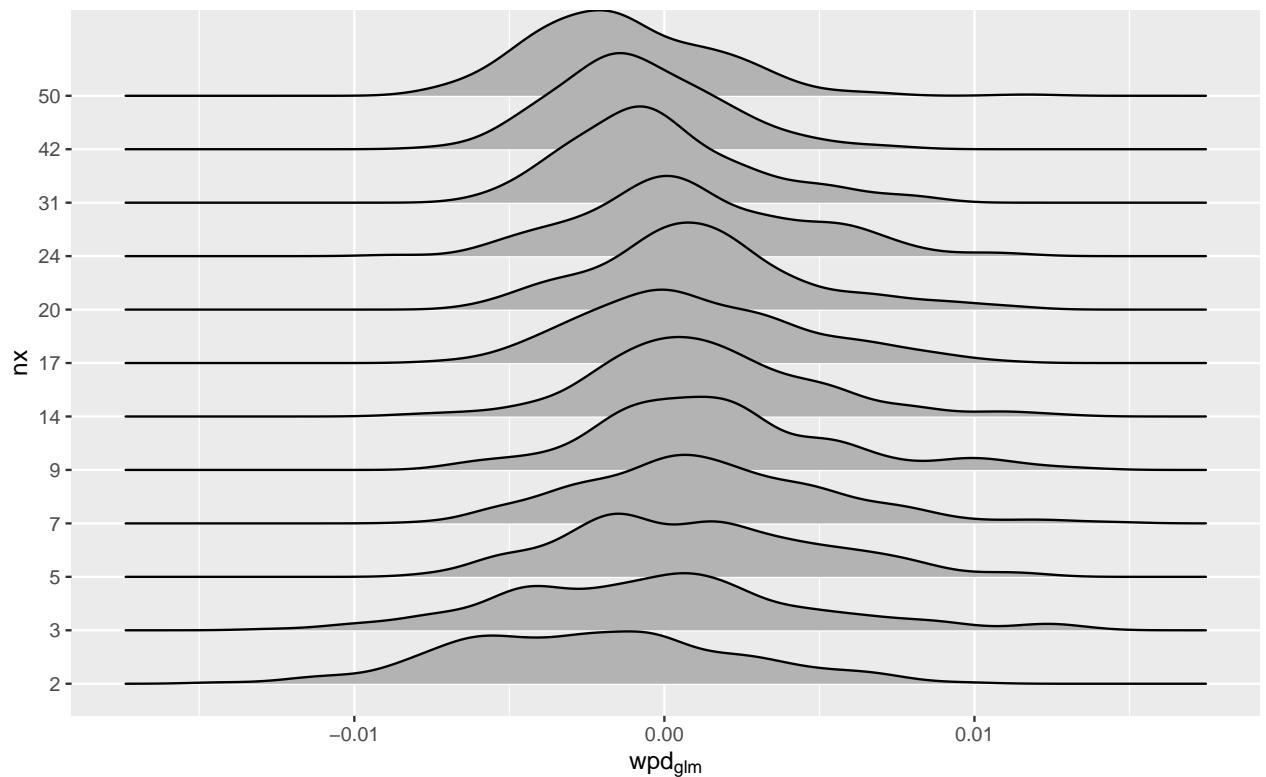


Figure 14: Distribution of  $wpd_{glm}$  is plotted across different  $nx$  for  $m = 1$ . Both shape and scale of the distributions are different for lower  $nx$  than for higher  $nx$ .

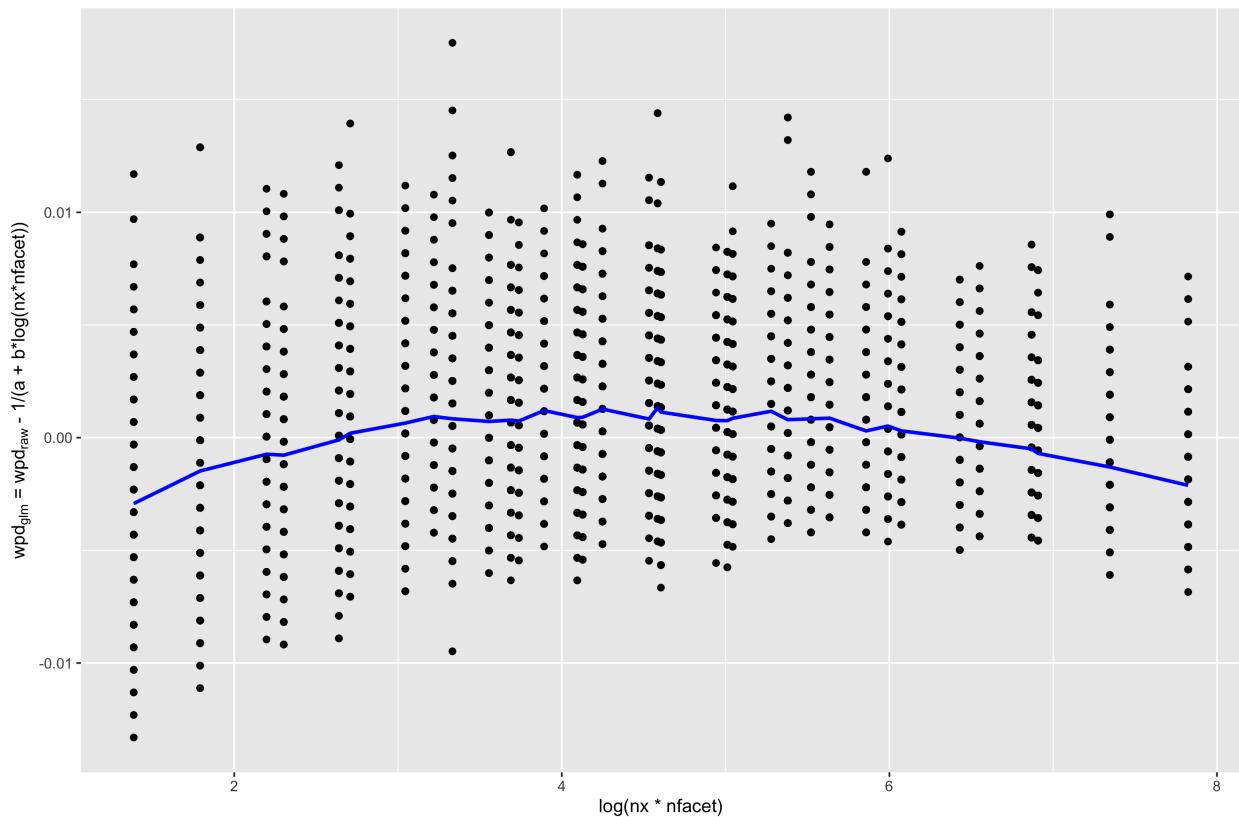


Figure 15: Residuals from the model is plotted across different number of comparisons for  $m = 2$ . Residuals seem to be independent of  $nx$  and have been defined as  $wpd_{glm}$ .

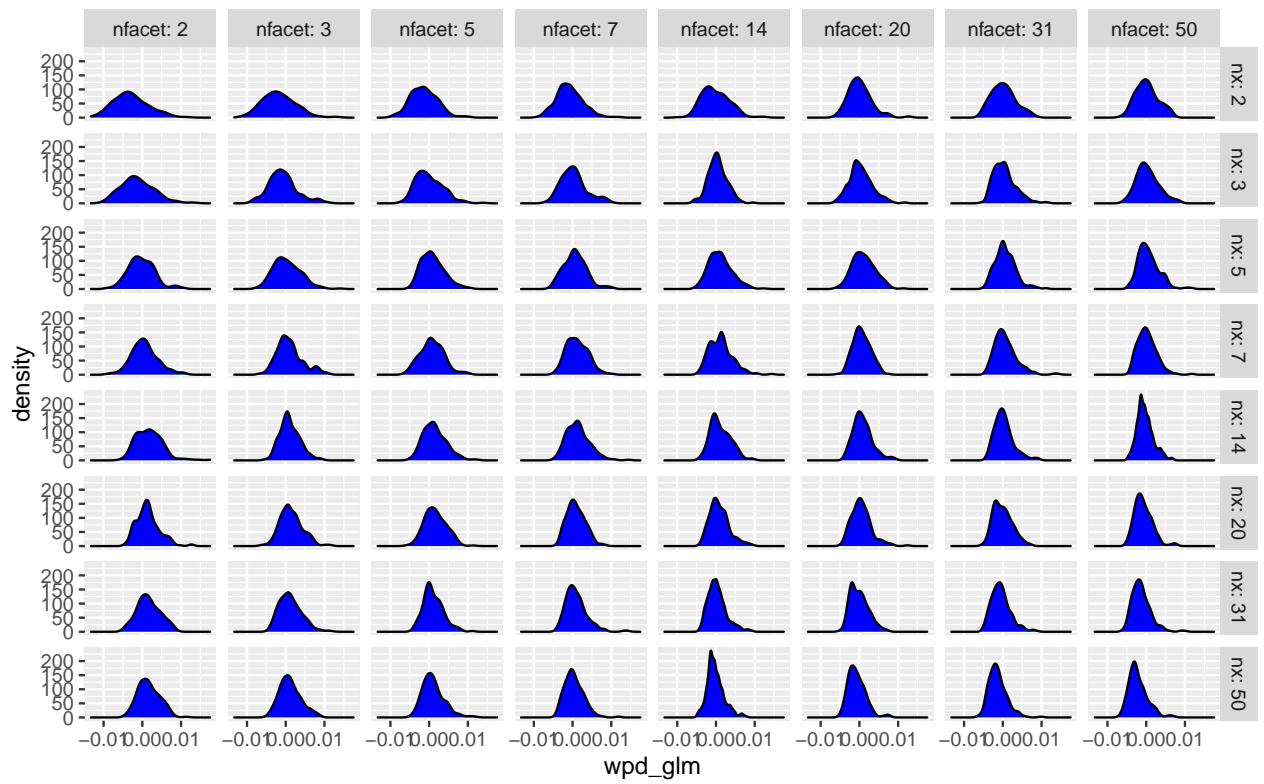


Figure 16: Distribution of  $wpd_{glm}$  is plotted across different  $nx$  for  $m = 2$ . Both shape and scale of the distributions are different for lower  $nx$  than for higher  $nx$ .

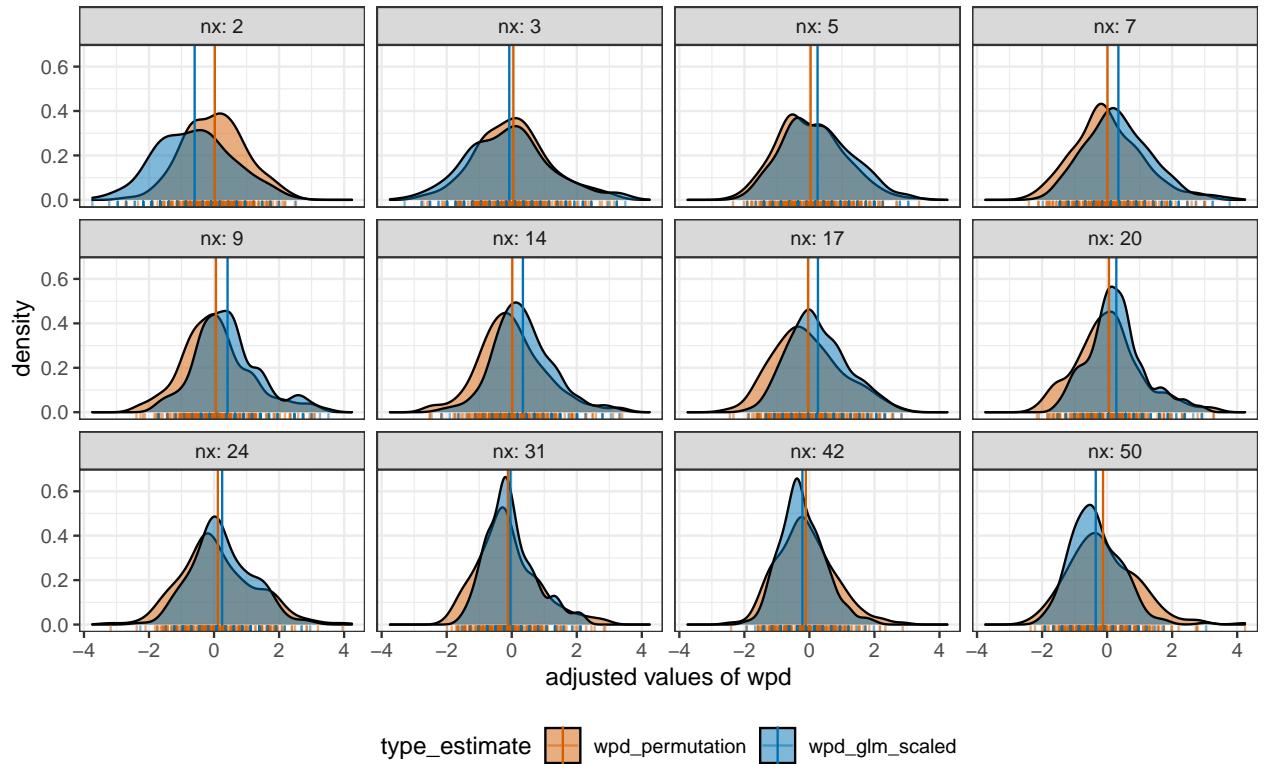


Figure 17: The distribution of  $wpd_{perm}$  and  $wpd_{glm-scaled}$  are overlaid to compare the location and scale across different  $nx$  for  $m = 1$ . The distribution of the adjusted measure looks similar for both approaches for higher levels.

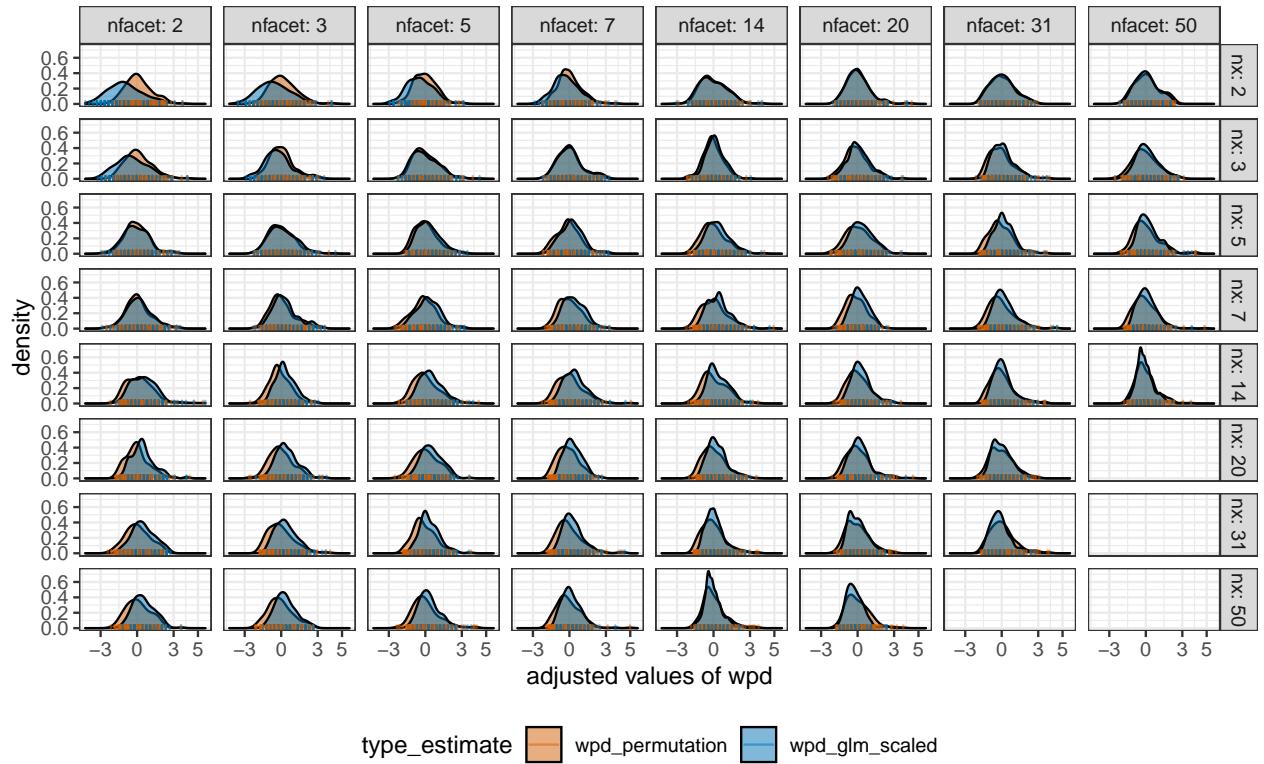


Figure 18: The distribution of  $wpd_{perm}$  and  $wpd_{glm-scaled}$  are overlaid to compare the location and scale across different  $nx$  and  $nfacet$  for  $m = 2$ .  $wpd_{norm}$  takes the value of  $wpd_{perm}$  for lower levels, and  $wpd_{glm-scaled}$  for higher levels to alleviate the problem of computational time in permutation approaches. This is possible as the distribution of the adjusted measure looks similar for both approaches for higher levels.

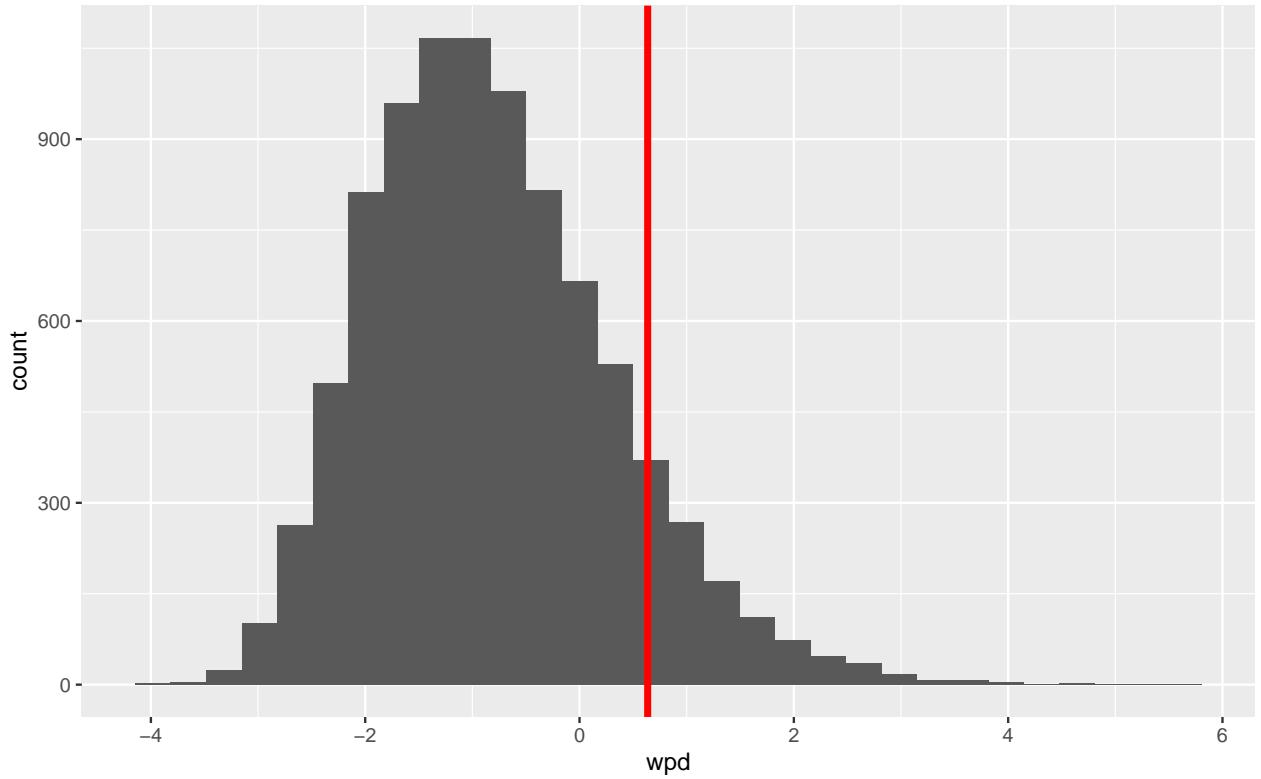


Figure 19: The distribution of  $wpd$  for 9 null designs of different levels with each harmony being rejected if the corresponding  $wpd$  value is higher than  $wpd_{threshold99}$ . Although the level of significance for each test is 1%, the level of significance for multiple tests together is 10% as displayed through the red line. This means 10% of the times the null designs will be tagged as interesting even when there is actually nothing interesting going on.