

# Workshop 8 - Longitudinal Data

JPS

28/12/2021

## Introduction

We are going to explore the two longitudinal data modelling approaches presented in the lecture: growth curve models and cross-lagged models.

**Exercise 1:** In this exercise we will use sentencing data from the Czech Republic. This data is truly unique since - as far as I am aware - it is the first dataset where sentences are linked to the judges who imposed them chronologically. This allow us to explore changes in the sentencing practice across judges as they become more experienced. Dr Drápal and I have just published an article exploring these questions in the journal *Justice Quarterly*. Here, we will replicate some of the analyses we have undertaken to model changes in severity (measured as the probability of imposing a custodial sentence) throughout roughly the first 1,000 cases sentenced by different judges from the Czech Republic since they join the judiciary. This analysis will allow us to assess whether judges become harsher or more lenient through their careers, while at the same time allowing us to measure changes in between-judge disparities across time.

**Exercise 2:** In the second exercise we will use an excerpt of the 'Pathways to Desistance' survey, which we used in Workshop 3 to explore the procedural justice model using path analysis. The survey follows a panel (or prospective longitudinal) design, with young offenders repeatedly interviewed in intervals of six months. We will use the first two survey waves and cross-lagged models to explore the causal relationship between procedural justice (as self-reported perceptions of fairness in how the police and court treated the participants) and legitimacy (as individual perceptions towards those same institutions). The vast majority of the literature on the topic is based on cross-sectional data, and assumes that the effect of procedural justice on legitimacy is unidirectional. Here we will explore whether it is possible that the relationship between these two concepts is bidirectional. That is, whether it is possible that perceptions of police and courts legitimacy affect individuals' perceptions of how fair these institutions are in their interactions. This is a research question that was tested in Pina-Sánchez & Brunton-Smith (2020) using similar models to those employed here.

## Exercise 1. Modelling Judicial Trajectories

We start by importing the data, which is saved as a .csv file.

```
judges = read.csv("judgesCZ.csv")
```

The original data could be classified as sensitive so I have trimmed it extensively to ensure that it is fully anonymised. In addition, I have prepared the variables that we are going to use, so no data cleaning is required this time. Still, it is always a good idea to get started with an exploratory analysis, even if brief, to assess what is in the dataset.

```
summary(judges)
```

##	range	prevconv	female	concurrent
##	Min. :-1.16358	Min. :-0.276906	Min. :0.0000	Mode :logical
##	1st Qu.: -0.49691	1st Qu.: -0.276906	1st Qu.: 0.0000	FALSE:4731
##	Median : -0.49691	Median : -0.176906	Median : 0.0000	TRUE :17681

```
## Mean      :-0.03034   Mean      : 0.001873   Mean      :0.1542
## 3rd Qu.: 0.16976   3rd Qu.: 0.123094   3rd Qu.:0.0000
## Max.     :10.50309   Max.      : 1.723094   Max.      :1.0000
## NA's     :4
## detention      judge_sentence      judge_ID      custody
## Mode :logical   Min.      :0.0010   Min.      :4000   Min.      :0.0000
## FALSE:21386    1st Qu.:0.2410   1st Qu.:4072   1st Qu.:0.0000
## TRUE :1026     Median :0.5350   Median :4110   Median :0.0000
##              Mean      :0.6419   Mean      :4110   Mean      :0.1414
##              3rd Qu.:0.9550   3rd Qu.:4163   3rd Qu.:0.0000
##              Max.      :2.0580   Max.      :4191   Max.      :1.0000
##
```

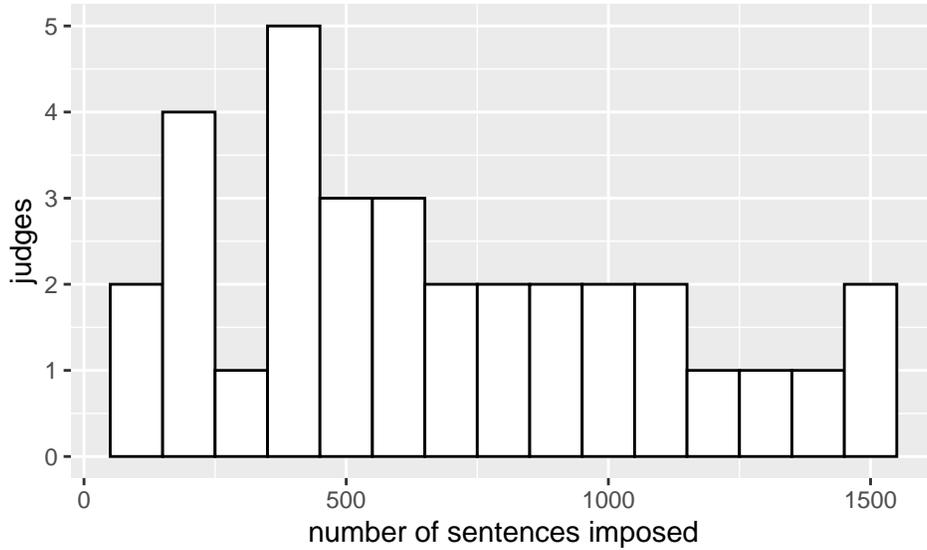
```
length(unique(judges$judge_ID))
```

```
## [1] 33
```

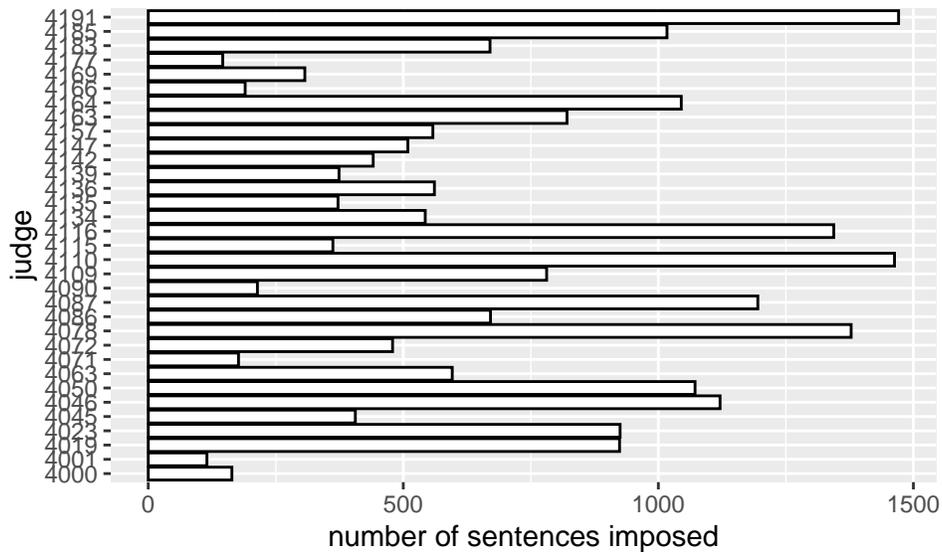
We have eight variables: ‘range’ captures the recommended sentence length for the offence type according to the Czech criminal code. This variable has been divided by ten and demeaned (the mean of the variable has been subtracted from every value so they are all centred around zero, this process helps simplify the computational process and as such helps in the estimation of complex models); ‘prevconv’, indicating the number of previous convictions, which has been demeaned in the same way; ‘female’, indicating whether the offender is a woman or a man; ‘concurrent’, a legal consideration when the offender is found guilty of multiple offences but the different penalties are not imposed on top of each other (i.e. consecutively); ‘detention’, indicating whether the offender was placed in pre-trial detention or whether given bail; ‘judge\_sentence’ indicating the order in which each judge imposed the sentences recorded (this has not been demeaned to facilitate its interpretation but it has been divided by 1,000); ‘judge\_ID’ to differentiate by judge presiding over each sentence; and ‘custody’ indicating whether a custodial sentence was imposed or not.

The longitudinal dimension of the dataset stems from the 22,412 sentences imposed by 33 judges (‘judge\_ID’). Given the different levels of experience for the judges used in the sample, and as a result of some of the trimming processes that I undertook to prepare the dataset, the distribution of the number of the sentences seen by each judge varies importantly. Growth curve models can handle this issue under the assumption that the missing segments of individual trajectories are missing at random.

```
library(dplyr)
#The following line creates a dataset with the number of sentences recorded by judge
count = judges %>% select(judge_ID) %>% table() %>% as_tibble()
count
names(count)[1] = "judge_ID"
summary(count$n)
library(ggplot2)
ggplot(count, aes(x=n)) + geom_histogram(binwidth=100, color="black", fill="white") + labs(y="judges",
```



```
ggplot(data=count, aes(x=as.character(judge_ID), y=n)) + geom_bar(stat="identity", color="black", fill=
```



We can see how the mean number of cases processed by judge is around 679, but that distribution is rather right-skewed, with more than 10 judges seeing less than 500 cases, and four judges seeing more than 1250. This uneven distribution is not ideal since those four judges will have a stronger influence in the estimation of trajectories to be explored in our models.

Let's move now to the modelling phase. I would divide this into three different steps, in order of complexity:

1. First, we could specify a standard logistic model to determine whether the number of cases imposed has an influence on sentence severity (the probability of imposing custodial sentences).
2. Then, we can use a random intercepts model to assess whether there are meaningful between-judge disparities in the use of such penalty, i.e. whether there are judges systematically harsher or more lenient than others.
3. Lastly, we can specify a random slopes model for the variable 'judge\_sentence' to explore whether those judge disparities shrink or expand as judges become more experienced, i.e. as they sentence more cases.

```
logit = glm(custody ~ range + prevconv + female + concurrent + detention +
            judge_sentence, family="binomial", data=judges)
summary(logit)
```

```
##
## Call:
## glm(formula = custody ~ range + prevconv + female + concurrent +
##      detention + judge_sentence, family = "binomial", data = judges)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -3.4698  -0.4144  -0.3014  -0.2423   2.6965
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -2.00040    0.05620 -35.597 < 2e-16 ***
## range         0.32193    0.02275  14.151 < 2e-16 ***
## prevconv      2.70232    0.05760  46.917 < 2e-16 ***
## female       -0.18171    0.07839  -2.318  0.02045 *
## concurrentTRUE -0.44929    0.05319  -8.447 < 2e-16 ***
## detentionTRUE  3.32822    0.09636  34.540 < 2e-16 ***
## judge_sentence -0.13122    0.04988  -2.631  0.00852 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 18268  on 22407  degrees of freedom
## Residual deviance: 12547  on 22401  degrees of freedom
## (4 observations deleted due to missingness)
## AIC: 12561
##
## Number of Fisher Scoring iterations: 5
```

So, as could be expected ‘range’, ‘prevconv’, and ‘detention’, all point at more serious offences, or more persistent offending, and as such they are associated with a higher probability of receiving a custodial sentence; ‘concurrent’ sentences are by definition less severe than consecutive sentences, and female offenders tend to commit less serious crimes, so those two coefficients showing a negative coefficient also makes sense. The most interesting part is that our main variable of interest ‘judge\_sentence’ has got a negative and statistically significant effect too. This means that judges tend to become more lenient with experience. Specifically, after 1,000 cases processed, compared to the moment when a judge imposes their first sentence, the odds ratio of imprisonment is...

```
exp(summary(logit)$coefficients[7,1])
```

This seems pretty substantial; using a reference category we can also calculate the probabilities of receiving a custodial sentence at those two time points in the judicial career. To define this reference category we will use an average offence type (range = 0) committed by an average male offender with no previous convictions (female and prevconv = 0), receiving bail (detention = 0) and a consecutive sentence (concurrency = 0). That is, we will just use the regression coefficients for the intercept and ‘judge\_sentence’ as setting them all to zero cancels the terms. Anything multiplied by 0 is 0. Remember, to go from an odds ratio to a probability we use the equation  $Odds/(1+Odds)$ .

```
#The probability of imposing a custodial sentence for a judge with no experience is just calculated using
exp(summary(logit)$coefficients[1,1]) / (1+exp(summary(logit)$coefficients[1,1]))
```

```
## [1] 0.1191613
```

```
#The probability of imposing a custodial sentence for a judge that has imposed 1000 sentences.  
#Remember that we divided the number of cases by 1000, so an increment of 1 represents 1000 cases.  
exp(summary(logit)$coefficients[1,1]+summary(logit)$coefficients[7,1]) /  
(1+exp(summary(logit)$coefficients[1,1]+summary(logit)$coefficients[7,1]))
```

```
## [1] 0.1060616
```

Using these comparison of reference categories we can put effect sizes in context a little more clearly, and see how the leniency effect attributed to experience is not really that dramatic. Still, it is interesting to see how the sentencing process is not entirely deterministic, there are discrepancies that can be attributed to non-legal factors such as the stage of the career of the judge imposing the sentence. To assess the extent of between-judge disparities, i.e. differences in severity between judges, we can extend this model by including a random intercepts term. This model can take some time to compute as the addition of random effects makes it considerably more complex.

```
library(lme4)  
RI = glmer(custody ~ range + prevconv + female + concurrent + detention + judge_sentence +  
          (1|judge_ID), family="binomial", data=judges)  
summary(RI)
```

```
## Generalized linear mixed model fit by maximum likelihood (Laplace  
## Approximation) [glmerMod]  
## Family: binomial ( logit )  
## Formula: custody ~ range + prevconv + female + concurrent + detention +  
## judge_sentence + (1 | judge_ID)  
## Data: judges  
##  
## AIC BIC logLik deviance df.resid  
## 12122.5 12186.7 -6053.3 12106.5 22400  
##  
## Scaled residuals:  
## Min 1Q Median 3Q Max  
## -27.1526 -0.2970 -0.2039 -0.1415 8.9859  
##  
## Random effects:  
## Groups Name Variance Std.Dev.  
## judge_ID (Intercept) 0.3025 0.55  
## Number of obs: 22408, groups: judge_ID, 33  
##  
## Fixed effects:  
## Estimate Std. Error z value Pr(>|z|)  
## (Intercept) -2.02261 0.11212 -18.040 < 2e-16 ***  
## range 0.33157 0.02321 14.283 < 2e-16 ***  
## prevconv 2.76728 0.05983 46.255 < 2e-16 ***  
## female -0.24794 0.08014 -3.094 0.001975 **  
## concurrentTRUE -0.49964 0.05435 -9.193 < 2e-16 ***  
## detentionTRUE 3.44188 0.10045 34.264 < 2e-16 ***  
## judge_sentence -0.21891 0.06093 -3.593 0.000327 ***  
## ---  
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
##  
## Correlation of Fixed Effects:  
## (Intr) range prvcnv female cnTRUE dtTRUE  
## range -0.071
```

```
## prevconv    -0.116  0.129
## female      -0.064 -0.009  0.135
## cncrrntTRUE -0.332  0.170 -0.010 -0.051
## detentnTRUE -0.085 -0.190  0.107  0.008  0.038
## judge_sntnc -0.228 -0.114 -0.074 -0.024 -0.050 -0.008
```

Notice how now all the standard errors are larger than in the previous model. That is the result of adjusting for the within cluster correlation (sentences within judges), which were left unadjusted in our previous model (we assumed sentences are independent). We can also see that some regression coefficients have also changed by a few decimal points, that is due to the different estimation method undertaken in this model, but in general all regression coefficients related to the fixed part of the model are quite similar to what we observed before.

The most interesting part of this model is its random part. We can see that the standard deviation of the random intercepts term is 0.55; but what does that mean? Widespread between-judge disparities, or relatively negligible differences? As we learnt in the previous workshop, we can get the intercept for each judge using *coef*, and the specific random effect for each judge using *ranef*.

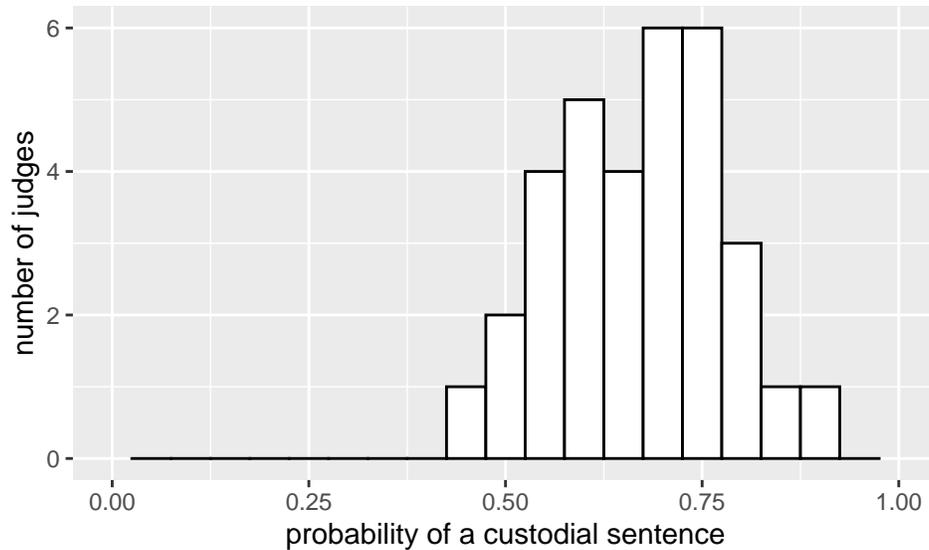
```
coef(RI)
ranef(RI)
```

Well, this looks rather substantive. For example, notice how for the second and third to last judges have got random intercepts larger and smaller than 50% the size of the model's intercept. To obtain a clearer view of the magnitude of these between-judge disparities we can calculate the probability of imposing a custodial sentence for each judge using just the intercept, that is, when 'judge\_sentence' is equal to zero. However, to somehow plot most of those probabilities around the centre of the probability range (0.5), we are also going to consider a reference case where 'prevconv' is equal to one. The reason why I suggest doing this is because otherwise we would find that the probabilities for the reference category for too many judges would be close to 0, which makes it hard to assess meaningful disparities. You can go ahead and see how that is the case by removing 'coef(RI)\$judge\_ID[,3]' from the following equation.

```
prob = exp(coef(RI)$judge_ID[,1] + coef(RI)$judge_ID[,3]) /
        (1 + exp(coef(RI)$judge_ID[,1] + coef(RI)$judge_ID[,3]))
summary(prob)
```

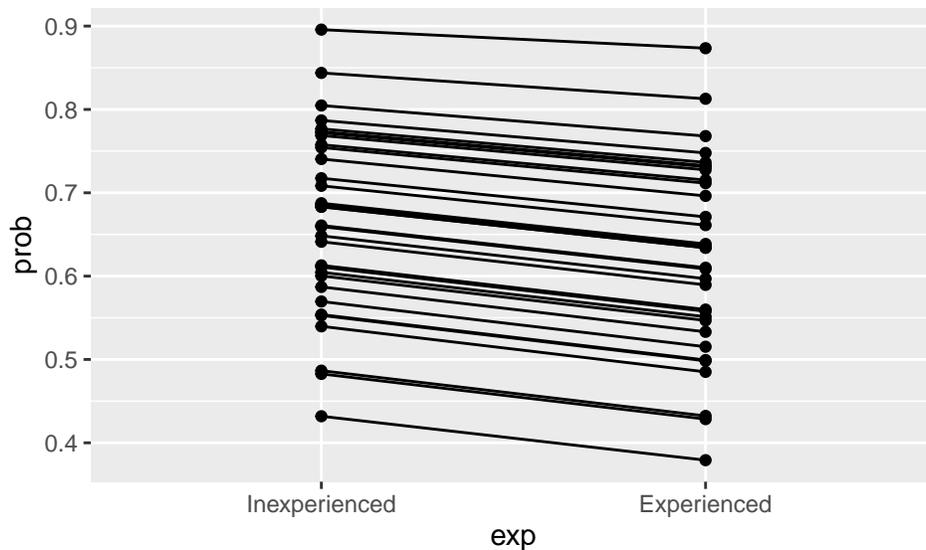
```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.4319  0.6003  0.6831  0.6692  0.7577  0.8958
```

```
prob = as.data.frame(prob)
names(prob) = "intercept"
ggplot(prob, aes(x=intercept)) + geom_histogram(binwidth=0.05, color="black", fill="white") +
  xlim(0,1) + labs(x = "probability of a custodial sentence", y = "number of judges")
```



As anticipated, the between-judge disparities are quite substantial, with some judges being more punitive than others. To assess how those disparities look like after judges become more experienced we can compare the probabilities that we have estimated to a new set of probabilities by judge after they have sentenced their first 1,000 cases.

```
prob2 = exp(coef(RI)$judge_ID[,1] + coef(RI)$judge_ID[,3] + coef(RI)$judge_ID[,7]) /
  (1 + exp(coef(RI)$judge_ID[,1] + coef(RI)$judge_ID[,3] + coef(RI)$judge_ID[,7]))
prob2 = as.data.frame(prob2)
names(prob2) = "intercept2"
lineplot = data.frame(exp=factor(rep(c("Inexperienced", "Experienced"), each=33),
  levels=c("Inexperienced", "Experienced")),
  judge=rep(1:33, 2), prob=c(prob$intercept, prob2$intercept2))
ggplot(data=lineplot, aes(x=exp, y=prob, group=judge)) + geom_line() + geom_point()
```



We see that the disparities are somehow wider when we consider experienced judges, but that is just due to the fact that probabilities are less sensitive to change the closer they are to their minimum and maximum values (0 and 1), or similarly, the further they are from the middle point of their range (0.5). Remember that

we are still assuming that the effect of 'judge\_sentence' is constant across judges. To relax this assumption we can specify a random slopes model, just like we did in our last workshop. This will be today's unguided exercise. Specifically, you are requested to specify a random slopes model to assess whether between-judge disparities change (become larger or smaller) after judges become more experienced.

-Remember from last week, all you need to do to go from a random intercepts model (like the one we have just specified above) to a random slopes model is include the name of the explanatory variable for which you want to estimate random slopes in the random part of the model, so, in our case:  $(1 + \text{'variable\_name'}|judge\_ID)$ .

-To check that the random slopes term is statistically significant run a likelihood ratio test. Remember from last week, you will need first to load the library *lmtest*, then you will need the command, *lrtest(model1, model2)*.

-In interpreting the random slopes model take a look at the standard deviation of the random intercept, the standard deviation of the random slope (does it seem significant?) and the correlation between these two terms (is it positive, negative, large or small?). The answer to our research question (do between-judge disparities shrink or grow as judges become more experience?) can be derived from that correlation term.

-Lastly, see if you can plot the between-judge disparities for experienced and unexperienced judges using the above lineplot as a template. Specifically, you will need to estimate 'prob' and 'prob2' as we did before, with the only difference that now these need to be derived from your new random slopes model. Once you have those, you can simply put them together in the same dataset as we did before (we called it 'lineplot'), and run the same *ggplot* code.

#####Remove the code from the handout#####

```
RS = glmer(custody ~ range + prevconv + female + concurrent + detention + judge_sentence +
          (1 + judge_sentence|judge_ID), family="binomial", data=judges)
summary(RS)
coef(RS)
```

Looking at the fixed part of the model the first thing to notice is that 'judge\_sentence' is not significant anymore, this is likely due to the within cluster correlation not being perfectly accounted for with the simpler random intercepts specification. On the random part of the model we can also see that the standard deviation for the random intercepts term remains quite substantial (even bigger than before), but so is the standard deviation for the random slopes. If we want to make sure that this random slopes term is significant we can run a likelihood ratio test

```
library(lmtest)
lrtest(RI, RS)
```

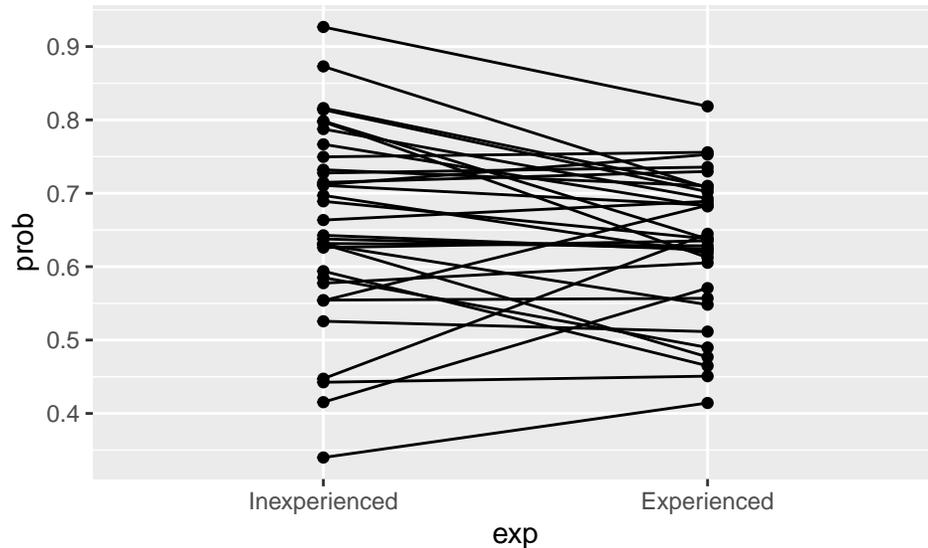
And of course, it is significant. Those random slopes disparities appear to be massive, indicating that although the average effect of 'judge\_sentence' is not statistically significant (i.e. the average effect of experience is not significantly different from zero), there is evidence pointing at very different trajectories followed by individual judges. In considering whether those different trajectories will converge (reducing the judge disparities that we observed in our previous model) or diverge (increasing those judge disparities), we can get an important clue from the correlation between the random intercepts and random slopes terms. This is negative, meaning that judges with a higher than average intercept (those who initially were harsher) will be associated with a negative experience effect (will become more lenient as they become more experienced), and the other way around, those judges with a lower than average intercept (those that were initially more lenient) will be associated with a positive experience effect (will become harsher as they become more experienced). We can see this visually as follows:

```
prob = exp(coef(RS)$judge_ID[,1] + coef(RS)$judge_ID[,3]) /
       (1 + exp(coef(RS)$judge_ID[,1] + coef(RS)$judge_ID[,3]))
summary(prob)
prob = as.data.frame(prob)
names(prob) = "intercept"
```

```

prob2 = exp(coef(RS)$judge_ID[,1] + coef(RS)$judge_ID[,3] + coef(RS)$judge_ID[,7]) /
  (1 + exp(coef(RS)$judge_ID[,1] + coef(RS)$judge_ID[,3] + coef(RS)$judge_ID[,7]))
prob2 = as.data.frame(prob2)
names(prob2) = "intercept2"
lineplot = data.frame(exp=factor(rep(c("Inexperienced", "Experienced"), each=33),
  levels=c("Inexperienced", "Experienced")),
  judge=rep(1:33, 2), prob=c(prob$intercept, prob2$intercept2))
ggplot(data=lineplot, aes(x=exp, y=prob, group=judge)) + geom_line() + geom_point()

```



This is an important finding. Even though experience does not seem to have a fixed effect on sentence severity, it does produce a really interesting effect in diminishing between court disparities. It seems that judges, as they sentenced more and more cases, become more in tune with each other. When I think of this it doesn't surprise me, I do feel the same way after marking student essays, for the first bunch I read, their marks are a little all over the place, :, until I have seen a number of them and establish what makes a good and a bad assignment, and I progressively become more consistent. More seriously, this has got important implications for sentencing policy since it seems there are procedures that could be taken into consideration to enhance consistency in sentencing without having to rely on prescriptive sentencing guidelines, which are often considered too intrusive and could undermine the principle of individualisation.

Now, putting things into perspective, our sample is huge but we have only looked at 33 judges. In addition, it is possible that our results are also affected by selection bias if those judges for whom we can see their trajectories for a much longer timespan are different from judges for which we can only see how they operate for a short timespan. Plenty of additional work needs to be done on those two fronts, but this is an interesting start.

## Exercise 2. Exploring the Causal Effect of Procedural Justice on Legitimacy

In this second exercise we are going to explore the assumed causal effect of procedural justice on legitimacy. Let's start by accessing the first two waves of data from the 'Pathways to Desistance' longitudinal survey.

```

w1 = read.csv("w1.csv")
w2 = read.csv("w2.csv")

```

We can merge the two datasets using *merge*, as we did in the second workshop. To do so we need to make sure we have a variable identifying each of the subjects that participated in the two waves, this is coded here under 'CASEID'.

```
paths = merge(w1, w2, by="CASEID")
#The following are another two ID variables generated by default after using merge that we do not need.
paths$X.x = NULL
paths$X.y = NULL
```

As always, let's run a brief exploratory analysis to take a look at our variables before moving to the modelling phase.

```
summary(paths)
```

```
##      CASEID          SOAGE          SOSGEN          SOPJCOP
## Min.   :    4   Min.   :14.00   Length:1354   Min.   :1.390
## 1st Qu.: 2764   1st Qu.:15.00   Class :character   1st Qu.:2.400
## Median : 5762   Median :16.00   Mode  :character   Median :2.730
## Mean   :26966   Mean   :16.04                      Mean   :2.767
## 3rd Qu.:53179   3rd Qu.:17.00                      3rd Qu.:3.100
## Max.   :55812   Max.   :19.00                      Max.   :4.490
##
##      SOLEGIT          S1PJCOP          S1LEGIT
## Min.   :1.000   Min.   :1.000   Min.   :1.000
## 1st Qu.:1.910   1st Qu.:2.430   1st Qu.:1.890
## Median :2.270   Median :2.790   Median :2.270
## Mean   :2.296   Mean   :2.804   Mean   :2.272
## 3rd Qu.:2.670   3rd Qu.:3.170   3rd Qu.:2.700
## Max.   :4.000   Max.   :4.540   Max.   :4.000
## NA's   :1      NA's   :93      NA's   :93
```

'SOAGE' captures the age of the offender at the time of the first interview, we can see how our sample is composed of young offenders aged 14 to 19 years old; 'SOSGEN' indicates the gender of the offender, 'SOPJCOP' and 'SOLEGIT' are aggregated indices of procedural justice and legitimacy as reported by participants in their first interview; 'S1PJCOP' and 'S1LEGIT' provide the aggregated indices of procedural justice and legitimacy obtained after participants respond to the same questions in their second interview (exactly six months after they were interviewed for the first time).

We can see a few missing cases, which stem mainly from wave-2 variables, this is probably due to the effect of attrition, i.e. participants dropping off throughout the study. To simplify things and stay focused on today's content, we will proceed to delete these cases, however we know we could do better, e.g. using regression or multiple imputation.

```
paths = paths[complete.cases(paths), ]
```

As in the previous exercise, we should undertake the modelling strategy in different stages, building up its complexity progressively. The simplest and most commonly used approach in the procedural justice literature, when longitudinal data is available, involves exploring the effect of past perceptions of procedural justice on future perceptions of legitimacy. This approach has been used by multiple researchers. Their rationale is as follows, most of the evidence of the effect of procedural justice on legitimacy in the literature stems from cross-sectional studies, but as we know we cannot safely infer a causal connection from those studies, amongst other reasons because they do not allow exploring the temporal dimension of the assumed causal effect. To look for more robust evidence of the causal link between procedural justice on legitimacy researchers first step has been to look at the effect of the former in time period t0, on the latter at a later time period, say t1.

```
naive = lm(S1LEGIT ~ SOPJCOP, data=paths)
summary(naive)
```

```
##
## Call:
## lm(formula = S1LEGIT ~ SOPJCOP, data = paths)
```

```
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.59025 -0.35196  0.00593  0.34740  1.73804
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.16805    0.08382   13.94 <2e-16 ***
## SOPJCOP      0.39838    0.02974   13.40 <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5446 on 1259 degrees of freedom
## Multiple R-squared:  0.1248, Adjusted R-squared:  0.1241
## F-statistic: 179.5 on 1 and 1259 DF,  p-value: < 2.2e-16
```

And yes, it seems that there could be a causal effect here, with the regression coefficient of procedural justice being positive and statistically significant. But hang on a minute, could it be that the observed relationship is due to some time-invariant factors that are confounding that relationship, i.e. factors that are simultaneously correlated with ‘SOPJCOP’ and ‘S1LEGIT’? E.g. gender, ethnicity, personality traits, etc. We can do better including a lagged term of the outcome variable, essentially turning the previous model into a first order autoregressive model.

```
causal = lm(S1LEGIT ~ SOPJCOP + SOLEGIT, data=paths)
summary(causal)
```

```
##
## Call:
## lm(formula = S1LEGIT ~ SOPJCOP + SOLEGIT, data = paths)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.76162 -0.29971  0.01334  0.30544  1.69843
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.79533    0.07640   10.410 < 2e-16 ***
## SOPJCOP      0.09197    0.03074    2.992  0.00283 **
## SOLEGIT      0.53072    0.02784   19.063 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4799 on 1258 degrees of freedom
## Multiple R-squared:  0.3209, Adjusted R-squared:  0.3198
## F-statistic: 297.3 on 2 and 1258 DF,  p-value: < 2.2e-16
```

We see a rather different picture. Procedural justice still has a significant effect on legitimacy but it is now less than a quarter of the size it was before. That is a big change. It seems that procedural justice is a significant factor in determining perceptions of legitimacy but not nearly as important as it is reported to be in the literature relying on cross-sectional designs or simple longitudinal models.

Let’s now consider another important assumption. The relationship between procedural justice and legitimacy has almost always been assumed to follow one particular direction, from procedural justice (fair treatment) to legitimacy (trust, alignment and cooperation with authorities). However, it is worth asking whether that relationship might be working in reverse too. Could it be that those people who perceive institutions to be legitimate also report that they are being treated more fairly? After all, we are just using subjective

perceptions reported in a survey, what someone perceives as ‘fair treatment’ might be very different from what other people do, and such difference might be stemming from multiple factors, one of them, potentially, how aligned the person is with the goals of the institution with which he/she is interacting. It is, hence, a hypothesis worth exploring, and we are going to see how we can do that using cross-lagged models.

In essence this involves the specification of a similar autoregressive model to the above (‘causal’) but substituting legitimacy for procedural justice in the left and right-hand side of the equation. Still, we need to do better than that, since the residuals from the two models are likely to be correlated we need to use structural equation modelling. To do so we will use the package *lavaan*. Here the syntax is slightly different. We need to specify the different models as an object within ‘’, with each model to be estimated in a different line. This will automatically provide a covariance term for the wave 2 residuals, but we will also request another covariance term for wave 1. Once the model is specified into an object we can estimate it using the command *sem*.

```
library(lavaan)
## The models have to be in different lines
c12w = 'S1PJCOP ~ SOPJCOP + SOLEGIT
        S1LEGIT ~ SOLEGIT + SOPJCOP
        #Covariances
        SOPJCOP~~SOLEGIT'
c12w = sem(c12w, data=paths)
summary(c12w)
```

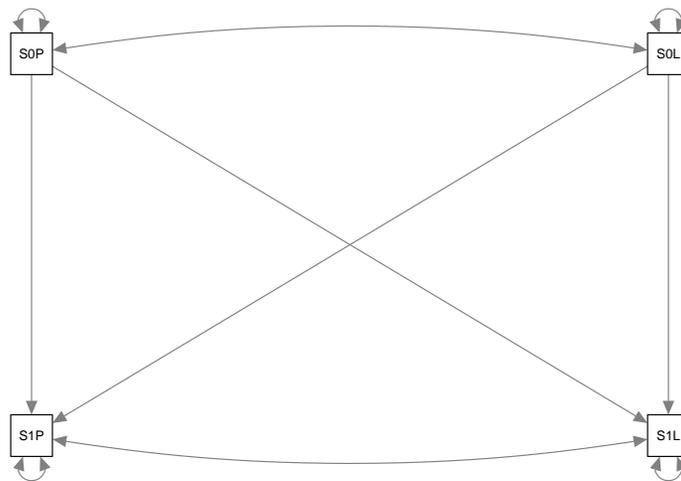
```
## lavaan 0.6-9 ended normally after 23 iterations
##
## Estimator ML
## Optimization method NLMINB
## Number of model parameters 10
##
## Number of observations 1261
##
## Model Test User Model:
##
## Test statistic 0.000
## Degrees of freedom 0
##
## Parameter Estimates:
##
## Standard errors Standard
## Information Expected
## Information saturated (h1) model Structured
##
## Regressions:
## Estimate Std.Err z-value P(>|z|)
## S1PJCOP ~
## SOPJCOP 0.325 0.032 10.220 0.000
## SOLEGIT 0.111 0.029 3.869 0.000
## S1LEGIT ~
## SOLEGIT 0.531 0.028 19.086 0.000
## SOPJCOP 0.092 0.031 2.995 0.003
##
## Covariances:
## Estimate Std.Err z-value P(>|z|)
## SOPJCOP ~~
## SOLEGIT 0.154 0.009 16.453 0.000
```

```
## .S1PJCOP ~~
## .S1LEGIT      0.063    0.007    9.083    0.000
##
## Variances:
##           Estimate Std.Err z-value P(>|z|)
## .S1PJCOP      0.246   0.010  25.110   0.000
## .S1LEGIT      0.230   0.009  25.110   0.000
## SOPJCOP       0.266   0.011  25.110   0.000
## SOLEGIT       0.324   0.013  25.110   0.000
```

The output is slightly different; you can find the main information under *Regressions*. So, what do you make of these results? Is there evidence of a bidirectional causal effect in the relationship between procedural justice and legitimacy? And if so, is it important? Which one seems stronger? And what do you make of the hundreds of studies on the topic that have used cross-sectional data and interpreted relationships between procedural justice on legitimacy to be the result of a causal effect of the former on the latter?

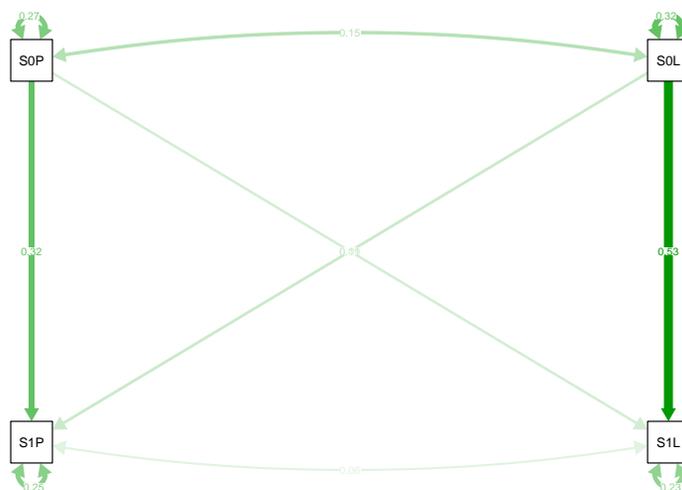
If you want to obtain measures of fit to compare different models using lavaan you can include the option `fit.measures=TRUE`. And if you want to draw a little graph representing your cross-lagged model you can use the package `semPlot` and the command `semPaths`.

```
summary(c12w, fit.measures=TRUE)
library(semPlot)
semPaths(c12w)
```



`semPlot` has lots of additional options that can be used to customize the plot; e.g. you can add the coefficients and give weight to the arrows (based on the coefficients' size).

```
semPaths(c12w, "Estimates")
```



That's it for today. We have three more workshops ahead of us where we will explore time-series, spatial modelling, and agent based modelling. But at this point you are equipped with a great range of new techniques that you could be applying to your assignment. I would recommend you to start working on it right now rather than waiting until the end of the teaching. You can access data you are already using for a research question of your interest, or you can also use the ESS to explore the optional question: "To what extent does the time spent using internet affect individual perceptions of trust in their co-citizens? And to what extent is such relationship uniform across countries?". We have already seen a potentially useful modelling strategy to explore those questions.