

Labour Market Analysis Tasks

Shane Ellard *York University*

Sample predoc data task. Goal: using data from the U.S. Current Population survey from 1976, produce the best answer to the question: how have hourly wages and labour force participation evolved for skilled and unskilled workers since 1976?

```
# this is a large file, it may take a few minutes to load
wages <- read_csv("/Users/Shane/predoc_sample_task/cps_wages_LFP.csv")
```

Trends in the labour Force

```
## Cleaning

clean_wages <- wages %>%
  # drop if wage missing
  #filter(!is.na(wage)) %>%
  filter(!is.na(skilled)) %>%
  # ensure R treats skilled as dummy variable, not continuous
  mutate(skilled = factor(skilled),
         lfp_dummy = ifelse(lfp=="In labor force",1,0)
  )
```

Trends for wages and labour force participation

A persistent fact of labour markets is that on average skilled workers earn more than unskilled workers. This in of itself shouldn't be too surprising. One simple hypothesis is that skilled workers are more productive and as a result garner higher wages. What might be surprising, however, is that this earnings gap has been steadily increasing over time as seen in the first figure below. In 1976, a skilled worker earned about \$1.85 per week more than an unskilled worker. By 2015, a skilled worker on average earned about \$11.60 per week more than an unskilled worker.

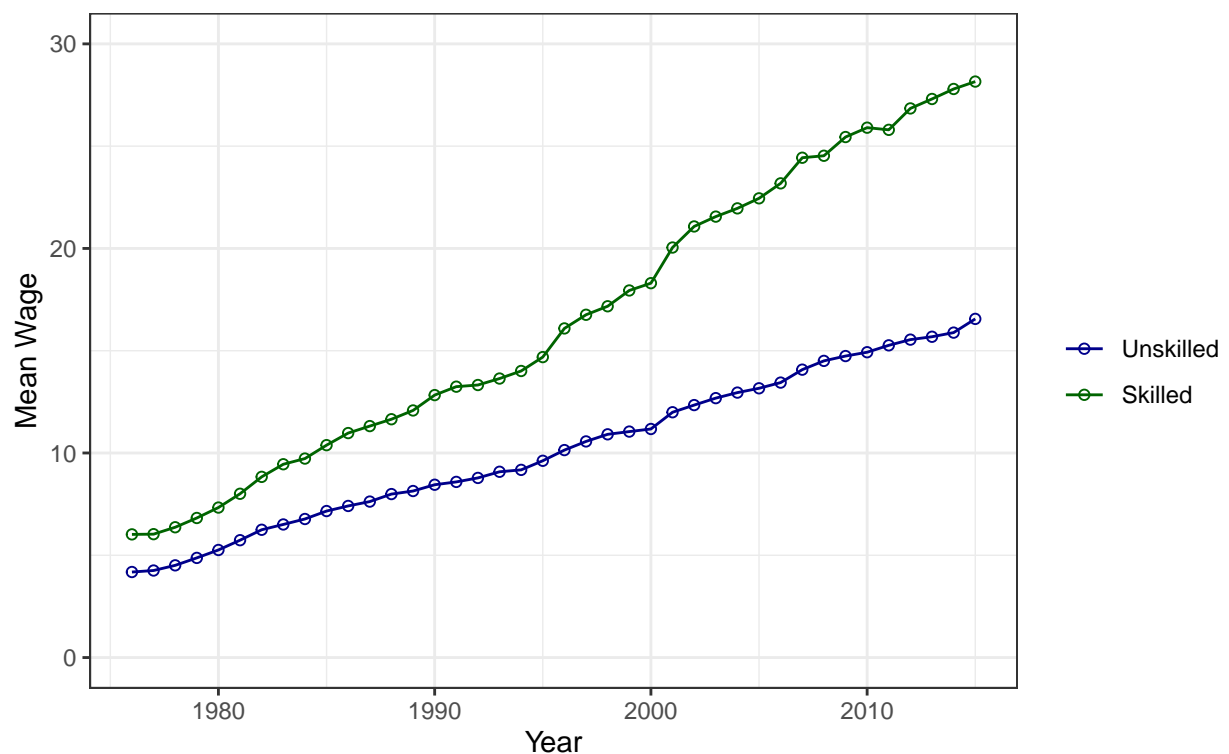
Skilled workers also have a higher labour participation rate. This also isn't too surprising. As we just uncovered, skilled workers earn higher wages and so the opportunity cost of not working is greater for skilled workers. It is noteworthy though that unlike the wage gap the difference in labour participation rates between skilled and unskilled workers is relatively stable. This pattern can be easily seen in the second figure below.

```
fig <- clean_wages %>%
  group_by(year, skilled) %>%
  # mean wage, grouped by year and skills
  summarise(mean_wage = mean(wage, na.rm=T),
            prop_lfp= mean(lfp_dummy, na.rm = T))
```

```
# graph for mean wage
ggplot(fig,
  aes(y = mean_wage , x = year , color = skilled, group=skilled)) +
  geom_line() +
  geom_point(shape = 1) +
  labs(
    title="Mean hourly wages between 1976 and 2015",
    subtitle ="Grouped by skill level",
    #legend title
    color=""
  )+
  ylim(0,30)+
  scale_color_manual(labels = c("Unskilled", "Skilled"), values = c("darkblue", "darkgreen")) +
  xlab("Year")+
  ylab("Mean Wage")+
  theme_bw()
```

Mean hourly wages between 1976 and 2015

Grouped by skill level

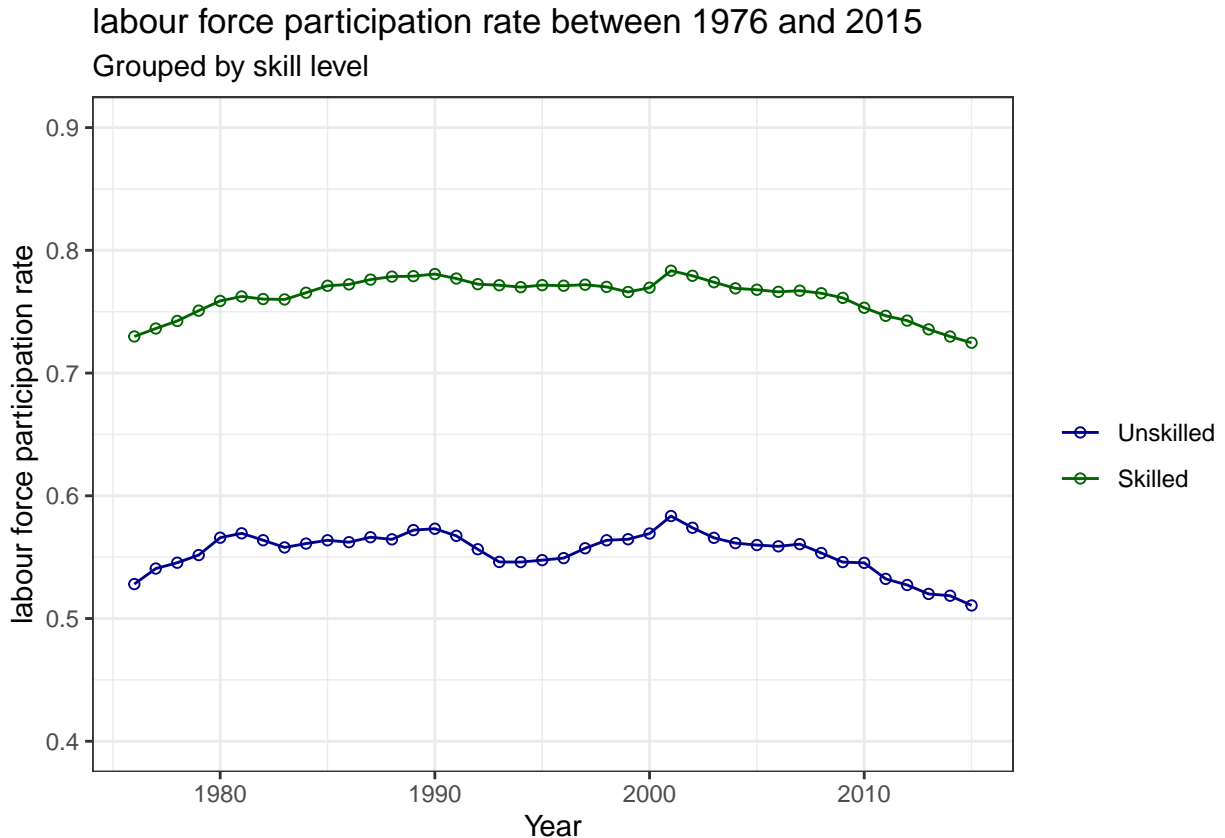


```
# graph for LFP
ggplot(fig,
  aes(y = prop_lfp , x = year , color = skilled, group=skilled)) +
  geom_line() +
  geom_point(shape = 1) +
  labs(
    title="labour force participation rate between 1976 and 2015",
```

```

    subtitle = "Grouped by skill level",
    #legend title
    color=""
)+
ylim(.4,.9)+
scale_color_manual(labels = c("Unskilled", "Skilled"), values = c("darkblue", "darkgreen")) +
xlab("Year")+
ylab("labour force participation rate")+
theme_bw()

```



Changes in labour force participation

Next, we consider labour force participation rates for males aged over 25. By grouping by race and age we can observe that participation rates are still relatively stable. It is interesting to note the variance in participation rates among skilled non-whites over the age 65 during the 80s. This can be seen in the top right panel of the third figure. Otherwise, participation rates are fairly linear. Also, it is easy to see that participation rates are fairly high for those aged 25 to 45 and aged 45 to 65 regardless of race. Unsurprisingly, participation rates are much lower for those aged 65 or older.

Every demographic group except one experienced a decline in labour force participation rates. The largest fall in labour force participation rate was experienced by skilled non-whites over the age of 65, whose participation rate fell by close to 13 percentage points. Interestingly, whites over the age of 65 (skilled or unskilled) and unskilled non-whites over the age of 65 did not experience large changes in labour force participation. It is also interesting to note that the next two demo-

graphic groups will the largest change in labour force participation rates are skilled non-whites aged 45 to 65 or aged 25 to 45. Perhaps this empirical finding will give a hint as to what is driving this pattern.

```
# males older then 25
adult_wages <- clean_wages %>%
  filter(age >= 25, sex=="male")
```

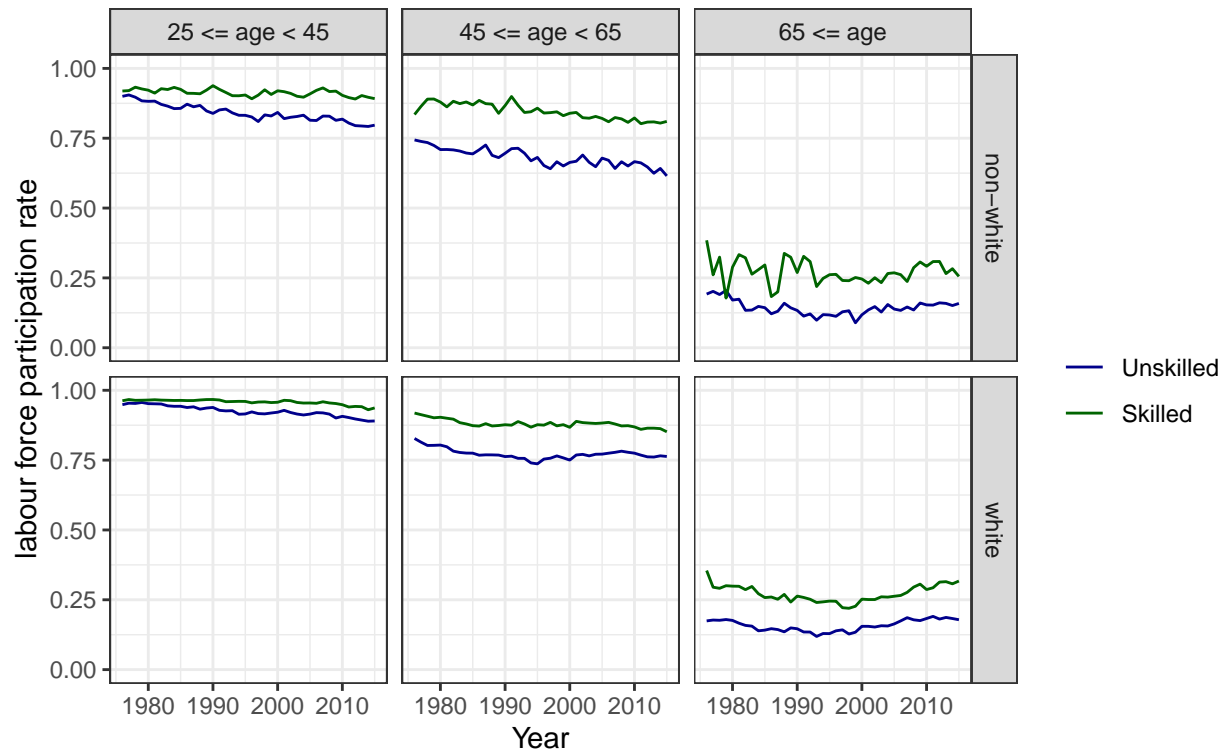
```
fig2 <- adult_wages %>%
  group_by(year, skilled,white,age_group) %>%
  # mean wage, grouped by year and skills
  summarise(mean_wage = mean(wage, na.rm=T),
            prop_lfp= mean(lfp_dummy, na.rm = T))
```

```
labels <- c("0" = "non-white", "1" = "white")

ggplot(fig2,
  aes(y = prop_lfp , x = year , color = skilled, group=skilled)) +
  geom_line() +
  labs(
    title="labour force participation rate between 1976 and 2015",
    subtitle="Grouped by age, race, and skill level",
    #legend title
    color=""
  )+
  facet_grid(white ~ age_group, labeller=labeller(white = labels))+
  ylim(0,1)+
  scale_color_manual(labels = c("Unskilled", "Skilled"), values = c("darkblue", "darkgreen")) +
  xlab("Year")+
  ylab("labour force participation rate")+
  theme_bw()
```

labour force participation rate between 1976 and 2015

Grouped by age, race, and skill level



```
old <- fig2 %>%
  filter(year == 1976)

new <- fig2 %>%
  filter(year == 2015)

table1 <- data.frame(new, old$prop_lfp, old$mean_wage) %>%
  mutate(change_lfp = prop_lfp - old.prop_lfp,
         change_wage = mean_wage - old.mean_wage) %>%
  select(year, skilled, white, age_group, change_lfp, change_wage) %>%
  arrange(change_lfp)

#can format this table to look much better, rename dummies, etc.
kable(table1)
```

year	skilled	white	age_group	change_lfp	change_wage
2015	1	0	65 <= age	-0.1292746	30.91730
2015	0	0	45 <= age < 65	-0.1287055	13.42193
2015	0	0	25 <= age < 45	-0.1031399	11.56488
2015	1	1	45 <= age < 65	-0.0670128	30.00383
2015	0	1	45 <= age < 65	-0.0644412	15.17992
2015	0	1	25 <= age < 45	-0.0583081	12.81443
2015	1	1	65 <= age	-0.0375579	30.20522
2015	0	0	65 <= age	-0.0336459	16.37606
2015	1	0	25 <= age < 45	-0.0272442	23.67246
2015	1	1	25 <= age < 45	-0.0255314	23.99867
2015	1	0	45 <= age < 65	-0.0246407	26.10177
2015	0	1	65 <= age	0.0042257	15.79110