Final Report-Big Data in Social Sciences 23/24

AUTHOR
Bertoni 1020372, Caggianelli 1019175, Magnaguagno 1024143, Malipiero 1024188

Childcare Services and Fertility Rates: a rational model of parenthood?

Overview of the phenomenon

Fertility rates have been decreasing in most of Europe for several decades and they are no longer sufficient to ensure the long-term replacement of the population. Italy is one of the lowest in Europe: its total fertility rate (TFR) was 1.27 in 2019 (World Bank, 2021) and 1.25 in 2021 (Eurostat). This will predictably have enormous consequences, especially in financial terms, in sectors like health care and pensions.

We have decided to investigate this phenomenon focusing on the possible relationship between higher fertility rates and the availability of childcare services. This perspective comes from the hypothesis that the determinants of below-replacement fertility might be related, among other factors, to the so-called "worker-mother conflict". Therefore, improvements in childcare availability (and so a decrease in its costs) might have a positive impact on fertility. In a similar model, childbearing would be considered in practice as a rational decision, in which parents weight the costs and benefits of having children and any reduction in the direct or indirect cost of children would produce an "increase in the demand" for children.

The aim of our research is thus to study a possible causal effect of an improvement in childcare availability on fertility rates. We have chosen to group our data by territory and use Italian Regions as units of a cross-sectional analysis.

Regional Italian Fertility Index

We could not find a unique dataset with all the information we needed about this phenomenon, so we have chosen different datasets from the ISTAT database and we have prepared them to meet our needs.

```
kable(summary(Childcare_Fertility_[,2:9]))
```

Total fertility index	Places in childcare	Women employement rate	Household Income	Religion	Poor conditions.scarce	Regional study years average	foreign_residents
Min. :1.014	Min. :19.56	Min. :30.49	Min. :26322	Min. :11.90	Min. :22.000	Min. : 9.673	Min. : 8395
1st Qu.:1.210	1st Qu.:43.12	1st Qu.:43.37	1st Qu.:27885	1st Qu.:15.28	1st Qu.:29.275	1st Qu.:10.208	1st Qu.: 90045
Median :1.252	Median :59.61	Median :57.69	Median :33665	Median :17.15	Median :31.900	Median :10.459	Median : 132451
Mean :1.255	Mean :55.27	Mean :51.79	Mean :32802	Mean :17.94	Mean :31.505	Mean :10.391	Mean : 258595
3rd Qu.:1.310	3rd Qu.:67.33	3rd Qu.:60.79	3rd Qu.:36140	3rd Qu.:20.93	3rd Qu.:34.050	3rd Qu.:10.624	3rd Qu.: 419442
Max. :1.578	Max. :82.42	Max. :66.23	Max. :41448	Max. :24.40	Max. :37.200	Max. :11.185	Max. :1190889

We firstly imported the data about reproductive behavior in Italy. The variable that indicates the average fertility rate in each Region (2021) was named "Total fertility index".

We illustrated the distribution of fertility rates on the Italian territory through a map, using the regional coordinates data from the eurostat map (nuts level = 2) and joining those data and the data from our dataset. Regions were then divided into four groups using the intervals of fertility rates given by the quartiles.

```
geo_data <-get_eurostat_geospatial(nuts_level=2)
italy_lim<- list(
xlim = c(5, 19),
ylim = c(35, 47.5)
)

geodata2 <- geo_data %>%  rename(Territory = NAME_LATN)  %>%  filter(CNTR_CODE== 'IT')

mapdata_reg <- geodata2 %>%
left_join(Childcare_Fertility_[,1:2])
```

```
mapdata_reg[10, 12 ] = 1.724 #Trento
mapdata_reg[9, 12] = 1.428 #Bolzano
mapdata_reg[12, 12] = 1.298 #Valle d'Aosta
```

```
mapdata_reg <- mapdata_reg %>% mutate(cat=cut(`Total fertility index`, breaks = c(0.966, 1.195, 1.258,1.320,1.724 )))

map_fert_reg <- ggplot(data = mapdata_reg ) +
geom_sf(aes(fill = cat), color = "black") +
scale_fill_viridis_d(option = "magma", name = "TFT", direction = -1) +
theme_bw() +
```

```
coord_sf(xlim = italy_lim$xlim,
ylim = italy_lim$ylim, expand = FALSE) +
theme(
legend.position = "top",
legend.title = element_text(size = 14),
legend.text = element_text(size = 10),
legend.background = element_rect(fill = "white")) +
labs(title = "Total Fertility Rate")
```

Childcare availability

We then imported the data about childcare services: in particular, we have chosen to use the average number of available places in childcare out of one-hundred children (0-2 years old) in that Region. We did not use public expenditure for childcare because we considered possible “economies of scale” effects and a consistent difference in wages and costs of childcare provision across Regions. The variable was called “Places in childcare”. We used data on public and private, but also employer-provided childcare. This last category might have a quite significant impact on our analysis, given the fact that “high-level” employees - that are more concentrated in the Centre-North of Italy - are also those who are more likely to access this kind of service. We excluded “recreational spaces” from the original dataset because we found it irrelevant for our purposes. We have chosen to average the variable of interest in a span of time of five years: in this way, we tried to capture the “perception” of the individual on availability of childcare services, rather than the actual current number that the common citizen may not be aware of. Purely monetary benefits were excluded from the analysis because we assumed they are similarly distributed across the Italian territory, depending on individual needs. Furthermore, for simplicity and data availability reasons, we assumed that black market for childcare (babysitting etc…) has an almost uniform distribution across Italian Regions and so it does not have a significant impact on variations in fertility rates. Similarly to what we did with fertility, we used a map to illustrate the distribution of childcare services on the Italian territory.

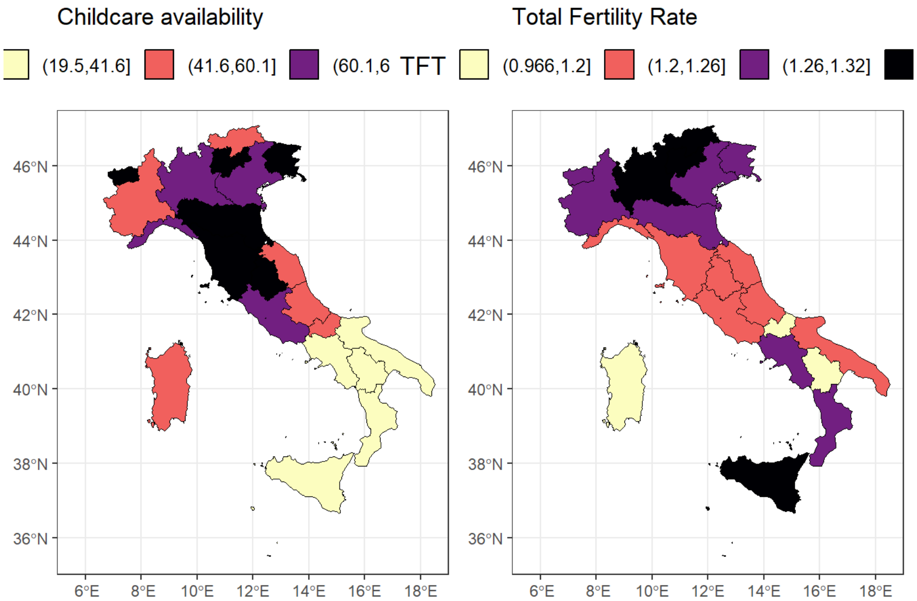
Plot regional TFR and places in childcare

```
mapdata_reg_soc <- geodata2 %>%
left_join(Childcare_Fertility_[,c(1,3)]) %>%  filter(CNTR_CODE== 'IT')
mapdata_reg_soc[10, 12 ] = 77.30 #Trento
mapdata_reg_soc[9, 12] = 51.76 #Bolzano
mapdata_reg_soc[12, 12] = 82.00 #Valle d'Aosta

mapdata_reg_soc <- mapdata_reg_soc %>% mutate(cat=cut(`Places in childcare`, breaks = c(19.5, 41.61, 60.06,67.54, 82.42) ))

map_fert_reg_soc <- ggplot(data = mapdata_reg_soc ) +
geom_sf(aes(fill = cat), color = "black") +
scale_fill_viridis_d(option = "magma", name = "TFT", direction = -1) +
theme_bw() +
coord_sf(xlim = italy_lim$xlim,
ylim = italy_lim$ylim, expand = FALSE) +
theme(
legend.position = "top",
legend.title = element_text(size = 14),
legend.text = element_text(size = 10),
legend.background = element_rect(fill = "white")) +
labs(title = "Childcare availability")

grid.arrange(map_fert_reg_soc, map_fert_reg, nrow=1 )
```



Single variable linear regression

We started by producing a linear regression describing the causal relationship between the 'total fertility rate' dependent variable and the 'places in childcare' indipendent variable.

```
Childcare_Fertility_lm <-lm(data = Childcare_Fertility_, formula =  `Total fertility index`~ `Places in childcare`)

summary(Childcare_Fertility_lm)
```

```
Call:
lm(formula = `Total fertility index` ~ `Places in childcare`,
    data = Childcare_Fertility_)

Residuals:
     Min       1Q   Median       3Q      Max
-0.24178 -0.04522 -0.00306  0.05123  0.32140

Coefficients:
                      Estimate Std. Error t value Pr(>|t|)
(Intercept)          1.2448813  0.0782819  15.903 4.83e-12 ***
`Places in childcare` 0.0001849  0.0013401   0.138    0.892
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1133 on 18 degrees of freedom
Multiple R-squared:  0.001056,  Adjusted R-squared:  -0.05444
F-statistic: 0.01904 on 1 and 18 DF,  p-value: 0.8918
```

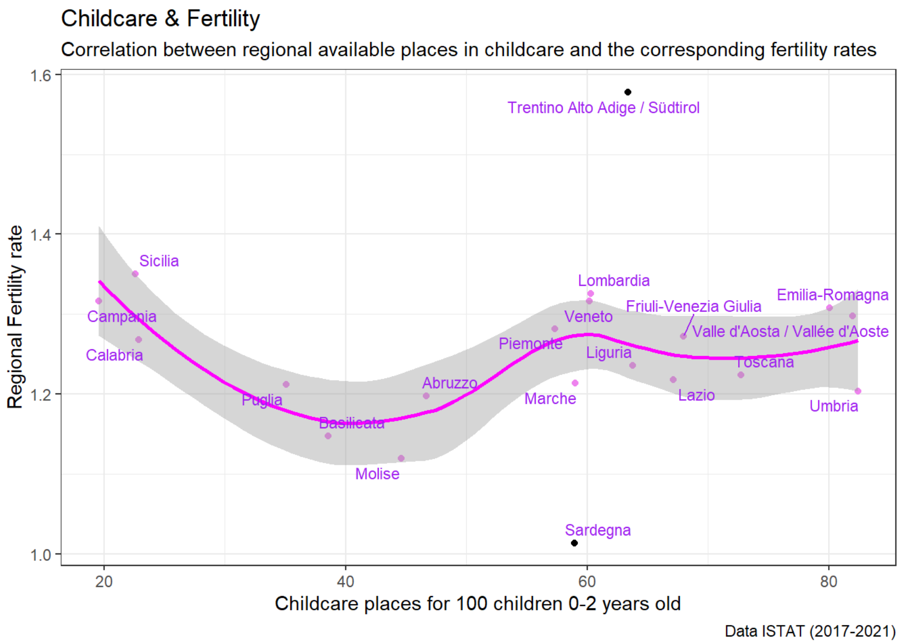
We found a regression coefficient equal to 0.0001849, which means that on average an additional available place in childcare services (for 100 children) would produce a 0.0001849 increase in the fertility rate of that Region. We found a R squared equal to 0.001056, which means that 0.001056 of the variance of the fertility rate is explained by the dependent variable. The standard error is equal to 0.0013401. The p- value associated with Places in Childcare is equal to 0.8918, which means the causal effect of the independent variable on the fertility rate is not statistically significant.

```
#plotting the regression
potential_outliers_fert <- Childcare_Fertility_ %>%  filter(`Places in childcare` > 60 & `Total fertility index` > 1.5 | `Places in childcare`> 40 &

Childcare_Fertility_plot <- Childcare_Fertility_ %>% filter(!(`Places in childcare` > 60 & `Total fertility index` > 1.5 | `Places in childcare`> 40
  ggplot(aes(x=`Places in childcare`, y=`Total fertility index`, labels(Territory))) +
  geom_point(color= 'violet') +
  theme_bw()+
    geom_smooth(formula = y ~ x , method= 'loess', color = "magenta", se = T) +
  labs(
    title = "Childcare & Fertility",
    x= "Childcare places for 100 children 0-2 years old",
    y= "Regional Fertility rate",
    caption = "Data ISTAT (2017-2021)",
```

```
    subtitle = "Correlation between regional available places in childcare and the corresponding fertility rates", color= 'violet') +
  geom_point(data=potential_outliers_fert) +
  geom_text_repel(data = Childcare_Fertility_, color= 'purple',
                  size= 3,
            aes(label = Territory))

Childcare_Fertility_plot
```



Multiple variable regression

Given the small and statistically not-significant coefficient found in the single variable regression, we tried to correct a possible omitted variable bias through a multiple variable regression. In fact, we found implausible the possibility of the variable 'Places in childcare' being a satisfactory description of such a complex mechanism, that involves several personal factors and complicated decisions. Thus we controlled for other two possible regressors, that might have a statistically relevant impact on fertility rates:

1. *The first one is women's occupational status (Women employment rate):* We have imported the dataset from ISTAT and we have called the variable "Women employment rate". In particular, the variable reflects the percentage of employed women (age 15-64) in that Region in 2022. This variable was chosen due to the fact that female employment might have a negative effect on the decision to become a mother (and vice-versa), because of the incompatibility between work and family responsibilities. For the same reason, there might also be a possible "postponement effect" after having the first child on subsequent births. However, on the other hand, employment may also have a positive impact on fertility in terms of improved economic conditions.

2. *The second one is economic status of the family (Household Income)*: Economic difficulties might be a leading cause for not wanting more children (ISTAT, 2017) and, in a rational choice model, an increase in income would be expected to result in an increase in the number of children. Therefore, we have included a variable indicating average household's income in each Region, and we have named it "Household income".

We started by computing the correlation between the omitted variables and the fertility rate, as well as the one between the two variables and childcare. A correlation with the independent variable different from zero would suggest a possible causal relationship, while a correlation with the regressor is a necessary condition for omitted variable bias (for the sake of being more concise we will include in the table also other variables that will be used later in the analysis)

```
kable(Childcare_Fertility_[,2:9] %>%
  cor(use = "pairwise.complete.obs"), caption = "Correlation Table")
```

Correlation Table

	Total fertility index	Places in childcare	Women employement rate	Household Income	Religion	Poor conditions	Regional study years average	foreign_residents
Total fertility index	1.0000000	0.0325025	0.2728453	0.5360781	0.0639748	-0.3684838	0.1280864	0.2461604
Places in childcare	0.0325025	1.0000000	0.9196262	0.7727786	-0.7325983	-0.4758205	0.7697516	0.2172149
Women employement rate	0.2728453	0.9196262	1.0000000	0.8747741	-0.6447537	-0.5907198	0.7457803	0.3068403
Household Income	0.5360781	0.7727786	0.8747741	1.0000000	-0.3568793	-0.6325038	0.6585214	0.5059070
Religion	0.0639748	-0.7325983	-0.6447537	-0.3568793	1.0000000	0.4271322	-0.6771318	0.1606619
Poor conditions	-0.3684838	-0.4758205	-0.5907198	-0.6325038	0.4271322	1.0000000	-0.5994867	-0.2019895

	Total fertility index	Places in childcare	Women employement rate	Household Income	Religion	Poor conditions	Regional study years average	foreign_residents
Regional study years average	0.1280864	0.7697516	0.7457803	0.6585214	-0.6771318	-0.5994867	1.0000000	0.3422487
foreign_residents	0.2461604	0.2172149	0.3068403	0.5059070	0.1606619	-0.2019895	0.3422487	1.0000000

We then produced a multiple linear regression containing these additional two variables. The variable of income was introduced with a linear-log regression, in order to tackle the probable decreasing marginal effect of income on the fertility rate. Moreover, we introduced an interaction term between “Women employement rate” and “Household Income”, that allows to account for the interaction between the omitted variables and, therefore, to investigate their role in a possible further analysis.

```
lm_final <- lm(formula = `Total fertility index` ~ (`Places in childcare`+ `Women employement rate` + log(`Household Income`) + (`Women employement

summary(lm_final)
```

```
Call:
lm(formula = `Total fertility index` ~ (`Places in childcare` +
    `Women employement rate` + log(`Household Income`) + (`Women employement rate` *
    `Household Income`)), data = Childcare_Fertility_)

Residuals:
      Min        1Q    Median        3Q       Max
-0.071284 -0.027659 -0.001441  0.025692  0.070513

Coefficients:
                                             Estimate Std. Error t value
(Intercept)                                -4.267e+01  1.953e+01  -2.185
`Places in childcare`                      -4.418e-03  1.521e-03  -2.906
`Women employement rate`                   -7.537e-02  1.830e-02  -4.119
log(`Household Income`)                     5.082e+00  2.151e+00   2.363
`Household Income`                         -2.828e-04  9.090e-05  -3.111
`Women employement rate`:`Household Income`  2.602e-06  5.822e-07   4.469
                                           Pr(>|t|)
(Intercept)                                 0.04636 *
`Places in childcare`                       0.01151 *
`Women employement rate`                    0.00104 **
log(`Household Income`)                     0.03314 *
`Household Income`                          0.00767 **
`Women employement rate`:`Household Income`  0.00053 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.04695 on 14 degrees of freedom
Multiple R-squared:  0.8667,    Adjusted R-squared:  0.8191
F-statistic: 18.21 on 5 and 14 DF,  p-value: 1.106e-05
```

The multiple regression resized our main coefficient and it improved the fit of the model measured by the adjusted multiple R-squared up to 0.8191. However, this might be due to the fact that we are adding variables rather than to an actual improvement of the explanatory power of the model. Even if jointly statistically significant, as revealed by the p-value related to the F-statistic of 18.21 (5, 14 degrees of freedom) , all the coefficients are quite small.

In general, the analysis presented suggests that the effect of the availability of childcare services may have a very small effect on fertility in Italy. On the contrary, it seems to be more influenced by the other factors included in the multiple analysis.

Exploratory analysis

Given the poor results of our multiple variable regression, we have tried to move to an exploratory type of analysis, in order to better understand the principal factors that might explain the variation in fertility rates among regions. We have taken in consideration some other factors that might be somehow impactfull and we tested their correlation with our variable of interest.

a. *Perception of poverty*: firstly, a variable indicating the perceived economic conditions of families in the last twelve months (in 2021) was chosen. We considered the fact that, in our case, an increase in income may alternatively result in higher expenditure on fewer children (higher quality education, expensive clothes, etc…) rather than higher fertility. The Dataset included the perception of one-hundred households and it indicated the percentage of answers reflecting respectively a “absolutely insufficient”, “scarce”, “adequate” and “very good” perceived economic status. We selected the percentage of households in each Region who declared to be in an overall bad economic condition (so “scarce” and “absolutely insufficient”) and we called the variable “*Poor conditions$scarce*”.

b. *Religion*: cultural and religious differences might be considered as a component of fertility choices: many Faiths could be said to bear a predominant attitude towards marriage and maternity, that might influence individual preferences on the number of children. We therefore chose to include a variable indicating the percentage of religious individuals in a Region and we called the variable “Religion” (any religion was included).

c. *Female educational level*: we considered the fact that, apart from women occupational status, also female education might be an important variable when analyzing reproductive behavior. More educated women might have higher level jobs that they would be reluctant to put aside when having a baby; on the other hand, more educated women might also have more intellectual (and so flexible) jobs, so they might be facilitated in conciliating the difficulties that come with having children. Furthermore, education might also shape a person's preferences in terms of the preferred number of children. We therefore imported data illustrating average qualifications among women in each Region and we called the variable "Study years".

d. *Foreign residents*: the presence of immigrants and foreign nationals is often regarded as an indicator of an enhanced reproductive behaviour, usually having higher fertility rates compared to the native population. This might be true usually due to the fact that immigrants tend to be younger and different cultures can put a different emphasis on the importance of family size. Some immigrant communities also have stronger networks which might provide a more available support system for childbearing and childhood. For this variable we decided to use data on foreign nationals who have the residence in Italy for a more relevant indicator of this effect in the stable Italian population. In particular, the variable "foreign residents" indicates the amount of non-italian nationals currently residing in each region.

```
corrmatrix <- cor(Childcare_Fertility_[,2:9])
```

Principal Component Analysis

After we improved the dataset through the addition of four variables, we used Principal Components Analysis in order to conduct a deeper analysis on the variability of fertility rates in Italy. Since the dataset's complexity increased, this tool was crucial in the reduction of our dataset dimensionality, whilst preserving as much information as possible needed for the additional analysis to be carried out. Additionally, given that PCA focuses on capturing variance in the data, it can be used as a tool for noise reduction, thus filtering out irrelevant information and identifying the crucial features in the data. This leads to an easier and more robust understanding of varying fertility rates between different regions.

```
Scaled_Dataset <- scale(Childcare_Fertility_[,2:9])
row.names(Scaled_Dataset) <- Childcare_Fertility_$Territory

PCA <- Scaled_Dataset %>% prcomp()
PCA %>%   tidy(matrix= "eigenvalues")
```

```
# A tibble: 8 × 4
     PC std.dev percent cumulative
  <dbl>   <dbl>   <dbl>      <dbl>
1     1   2.13  0.566        0.566
2     2   1.24  0.191        0.758
3     3   0.932 0.109        0.866
4     4   0.715 0.0639       0.930
5     5   0.569 0.0405       0.971
6     6   0.404 0.0204       0.991
7     7   0.227 0.00646      0.997
8     8   0.145 0.00261      1
```

In order to simplify the understanding of the results of the PCA we made a numerical and visual summary, giving insights on which variables contribute most to the principal components and their respective weight. In the graph is highlighted the strength and direction of the association between each variable and the first two principal components.

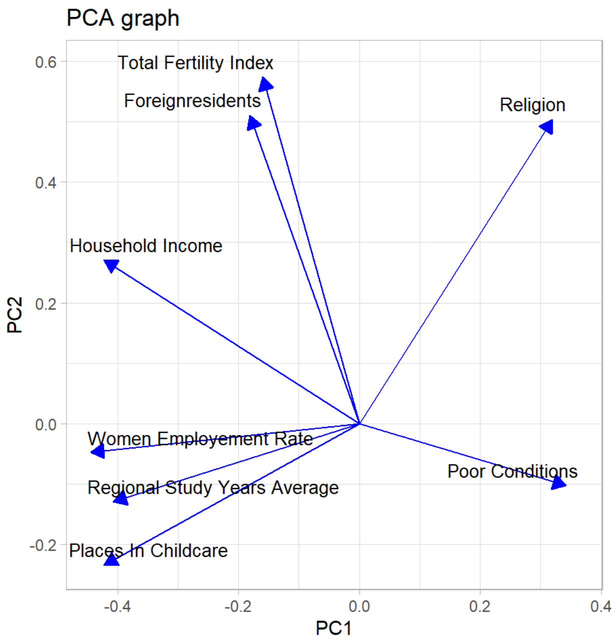
```
#PCA Var
PCA_var <- PCA|>
  tidy(matrix="rotation")|>
  pivot_wider(names_from="PC", names_prefix="PC", values_from="value")|>
  rename(Variable=column)|>
  mutate(Variable=stringr::str_to_title(Variable))|>
  mutate(Variable=stringr::str_replace_all(Variable, "_", ""))

kable(head(PCA_var))
```

Variable	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8
Total Fertility Index	-0.1598557	0.5732078	-0.5725319	0.2793611	0.4120728	-0.0299847	0.0033589	-0.2633043
Places In Childcare	-0.4226867	-0.2336944	0.1731194	0.3044651	-0.2195880	0.0455591	0.4152814	-0.6490766
Women Employement Rate	-0.4449776	-0.0472918	0.0351058	0.2957671	-0.2456806	-0.0795192	-0.8024991	0.0189947
Household Income	-0.4228509	0.2710748	0.0041747	0.2499522	-0.2791699	0.1417829	0.3977166	0.6549850
Religion	0.3202953	0.5036717	0.1679486	-0.0070491	-0.4472196	0.5846049	-0.1291990	-0.2387649
Poor Conditions	0.3422968	-0.1024160	0.3742746	0.7400948	0.3844993	0.1424520	-0.0234732	0.1258520

```
PCA_variances <-ggplot(data = PCA_var, aes(PC1, PC2))+
  geom_segment(xend=0, yend=0, arrow=arrow(length = unit(0.3, "cm"), type="closed",ends="first"), color="blue")+geom_text_repel(aes(label=Variable),
```

```
PCA_variances
```

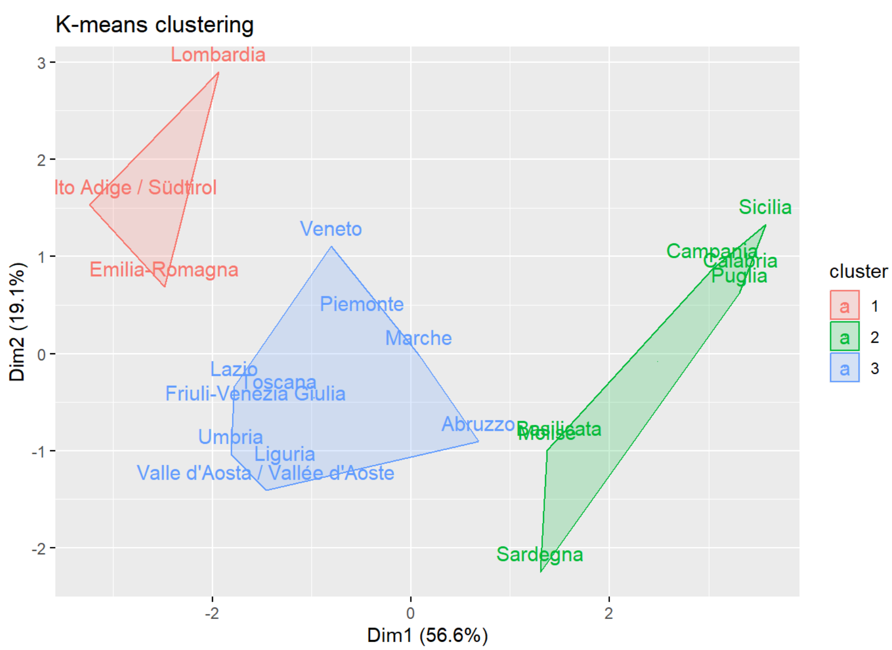


Clustering

On the context of our exploratory data analysis we deemed helpful grouping the observations in order to observe underlying connections. We chose a cluster analysis based on a k-means algorithm, which unsupervisedly grouped the data based on the number of clusters we chose, dividing them into groups with the smallest intra-cluster variations regarding the two main dimensions of the PCA. We chose a number of clusters equal to 3 by observing the graph plotting the within cluster variation for different values of k.

```
set.seed(12345)
#Rule of the elbow: search for the 'knee' in the within sum of square
cluster_plot<-fviz_nbclust(Scaled_Dataset, kmeans, method = "wss")

#the optimal number of clusters seems to be k=3
set.seed(1234)
k3 <- kmeans(Scaled_Dataset, centers = 3)
fviz_cluster(k3, data= Scaled_Dataset,
             geom = c("points", "text"),
             show.clust.cent = T,
             main = "K-means clustering")
```



The three clusters obtained draw an almost trivial picture: similarities across regions as explained by the principal components regarding the variables of interest result in groups of southern (and the island of Sardinia) and northern regions, with a third cluster formed by Lombardia, Emilia Romagna and Südtirol, regions which exhibit high levels of income, women empowerment and regional childcare as well as low poverty and above average fertility indexes.

Conclusion

The overall analysis suggests that improved childcare services would probably represent an almost insignificant encouragement for childbearing.

This might be due to many reasons. First of all, other not-analyzed factors may be operating, thus making it very difficult to isolate the impact of policies from other determinants.

Secondly, the structure of the chosen model itself might have some problems. With regard to this, the first critical point concerns information and, in particular, imperfect information. A rational choice model relies on the assumption that individuals make decisions based on complete information about costs and benefits. However, for such complex issues it is more likely to observe a situation of imperfect information, where people act according to perceptions more than accurate information, so a small increase in childcare service might not even be perceived by the average person. Another limit is that our analysis considered only the presence of childcare services on the various territories, but not other individual differences in terms for example of quality or access to services. Furthermore, a particularly relevant issue might be the fact that we looked at regional averages (for data availability reasons), while in this case it would be more appropriate to study the phenomenon at the municipal level. Finally, also individual non-economic preferences might play a decisive role in determining fertility outcomes and these may vary across Regions. However, even if we tried to tackle these effect through other related variables, they clearly do not represent an exhaustive answer. This adds to the fact that, in general, surveys on preferences for the number of children bear many difficulties, as they might be volatile or they might reflect social expectations rather than real preferences.

Bibliography

Bauernschuster, Hener, Rainer, “Children of a (Policy) Revolution: the Introduction of Universal Child Care and its Effect on Fertility”, CESifo, 2014, https://www.cesifo.org/en/publications/2014/working-paper/children-policy-revolution-introduction-universal-child-care-and .

Brini, Pavolini, Scherer, “Culle piene solo con misure strutturali”, Lavoce.info, 2023, https://lavoce.info/archives/102647/culle-piene-solo-con-misure-strutturali/ .

Dimai, “Shall we have another? Effects of daycare benefits on fertility, a case study in a region in Northeastern Italy”, Springer, 2023, https://genus.springeropen.com/articles/10.1186/s41118-023-00194-w#:~:text=The%20increase%20in%20public%20childcare,reviewed%20in%20Gauthier%20(2007) .

Gauthier, “The impact of family policies on fertility in industrialized countries: a review of the literature”, Springer, 2007, https://link.springer.com/article/10.1007/s11113-007-9033-x .

Imai, “Quantitative Social Sciences: An Introduction”, Princeton University Press, Princeton, 2017 .

Koka, Rapallini, “Italy’s demographic trap: Voting for childcare subsidies and fertility outcomes”, European Journal of Political Economy, 2023, https://www.sciencedirect.com/science/article/abs/pii/S0176268022000672 .

Links to datasets used

Fertility index: http://dati.istat.it/Index.aspx?DataSetCode=DCIS_FECONDITA1&Lang=en

Childcare availability: http://dati.istat.it/index.aspx?queryid=23229

Women employment: https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,Z0500LAB,1.0/LAB_OFFER/LAB_OFF_EMPLOY/DCCV_TAXOCCU1/IT1,150_915_DF_DCCV_TAXOCCU1_4,1.0

Household income: http://dati.istat.it/index.aspx?queryid=22919

Religion: http://dati.istat.it/index.aspx?queryid=24349

Economic conditions: http://dati.istat.it/index.aspx?queryid=22957

Average regional study years : https://esploradati.istat.it/databrowser/#/it/dw/categories/IT1,Z0820EDU,1.0/DCCV_POPTIT1_UNT2020/IT1,52_1194_DF_DCCV_POPTIT1_UNT2020_2,1.0

Foreign residents: http://dati.istat.it/Index.aspx?DataSetCode=DCIS_POPSTRRES1#