THE FORGOTTEN MANY: YOUTH CRIME AS AN ISSUE OF INEQUALITY, EDUCATION, AND INSTITUTIONAL NEGLECT*

L. Chiappelli M. Dolfato L. Garbarino G. Nespoli

December, 2023

Abstract

This paper studies the determinants of youth crime in Italy across its regions, in the years 2011-2016. We construct a simple Ordinary Least Squares (OLS) model with a series of regressors which we then augment with control variables. We then implement an Instrumental Variables (IV) approach to account for endogeneity of inequality, using predicted inequality as an instrument, and reach a final model specification. We find that inequality, education and government effectiveness are significant in explaining youth crime.

Keywords: youth crime, inequality, institutions

JEL Codes: K14, D63, Z18

*Paper realized for the course 'Empirical Research Methods and Data Analysis' of the Bachelor's in Economic and Social Sciences (BESS) at Bocconi University.
We thank Professors Giovanna Marcolongo and Noemi Sacchetti from CLEAN for their precious insights.

"Observe, that by justice I understand nothing more than that bond, which is necessary to keep the interest of individuals united; without which, men would return to the original state of barbarity. All punishments, which exceed the necessity of preserving this bond, are in their nature unjust."
Dei delitti e delle pene, Cesare Beccaria

1 Introduction

The problem of crime is an intrinsically economic one: the reason why many take the road of crime is scarcity [Buonanno, 2003]. While this is more evidently true of minor infractions, such as theft, serious violations of the law may often arise from insufficiency. By mere deduction, if there were no scarcity, there would be (nearly) no crime.
Although most crimes have an economic basis, the study of crime has not been closely related to economics until a few decades ago. Gary Becker, among others, paved the way to the use of microeconomic theory to study aspects of human interactions outside markets [Becker, 1968]. He noted that a systematic study of criminal activity had been hindered by a link between morality and scientific enquiry. Becker suggested that the neglect arose from the idea that what was immoral did not deserve scientific attention. However, this neglect led to the unsurprising effect that unscientific interventions in the field were used and are, in part, still used today.
In his treatise of the subject, Becker used the economic idea of rationality to explain criminal activity: criminals weigh costs and benefits before engaging in crimes and are utility maximizers. Another concept that Becker introduced was that resources should not be used to erase crime or punish all offenders, but rather optimized based on marginal costs of crime prevention and marginal benefits of crime reduction. This rational framework is path-breaking in guiding the criminal justice system because it no longer considers delinquents as simply deviant individuals, but rather as people who make coherent choices.
The field of applied economics expanded immensely after the 1960s, with focus extending to macro areas [Buonanno, 2003]. In fact, the study of crime is not only pursued with problems of individual utility maximization, but also introducing agents such as society and the government.

Our analysis wishes to ride the wave of the macro approach, focusing on the interaction between unlawful acts and government interventions. We will attempt to understand the main determinants of youth crime in Italy, to suggest where government resources should be allocated. Our initial thought is driven by the idea that inequality is more relevant than net income, mainly because of the strong welfare state system that Italian residents benefit from [Bonelli, 1992]. We suspect that the relative effect of inequality is stronger than the absolute one of income in young age because of a peer comparison effect. Indeed, young individuals tend to engage more in criminal activity not because of their family economic status alone, but as they are confronted with peers that are better off, with strong negative effects arising from such comparison. Also, we consider that unemployment may be especially relevant because individuals have both more possibility and incentive to spend time in the presence of potentially deviant groups, such as mafia-like organizations. This may be due to the responsiveness of youths to price incentives and the marginal returns of an hour of crime with respect to an hour of unemployment [Grogger, 1998]. We also maintain that education plays a role in youth crime. Finally, we believe that measures of government effectiveness and social participation are important because youth crime is likely to be worsened by a perception of civic neglect.

1.1 Motivation

The issue of crime is particularly relevant because it has a large spillover effect beyond the actual victims. Entire communities, if not cities, are deeply influenced by criminal activity. It suffices to mention the infamous example of the South of Italy, whose scarce development is sometimes attributed to organized crime [Felice, 2019]. In general, the field of crime prevention and resolution has suffered from a lack of empiricism in decision-making [Becker, 1968], which resulted in worse outcomes for both criminals and civilians. Indeed, much emphasis has been placed on the varying degrees of punishment as a solution [Levitt, 1998], rather than attempting a more holistic approach.
The reason for the choice of young individuals for the sake of this study is not immediate. After all, crime is a burden to society if committed at all ages, and most offences, especially severe ones, are committed well into adulthood [Eurispes, 2023]. We deem that young individuals who start on the course of wrongdoing are more easily brought back on the right way. Consider a person of 50 years who has spent half of their life in and out of prison: such a person would be very difficult to rehabilitate. Also, even if reformation were successful and diminished the negative effects of transgression on society, the reformed person would not have many work years to look forward to. Therefore, contributions to society would be harder to make. A young person, instead, still has enough time to change and make a lasting contribution to the community. These considerations are of purely economic nature, and there is of course an important aspect that cannot be overlooked: a young person is often forced, or at least pushed, to take the way of crime [Damm and Dustmann, 2014]. It is of paramount importance to provide an alternative.

1.2 An overview of youth crime in Italy

Before delving into the data used for the sake of our study, it is worth giving an overview of youth crimes in Italy. These are divided into 13 main categories, namely crimes against: justice administration, economy and trade, family, public safety, public morality, public order, persons, property, public administration, State, religion and the deceased, animals, and counterfeiting.
We reported a graph 1 which shows the relative magnitude of such crimes in Italy within regions in the cumulative period 2011-2016. The most important spheres of crimes committed are by far against property and persons, followed by public administration. The remaining crimes make up a residual portion of the total.
Lombardia is the region with the highest absolute number of youth crimes, whereas Valle D'Aosta is the one with less juvenile criminality. However, it is more insightful to look at the

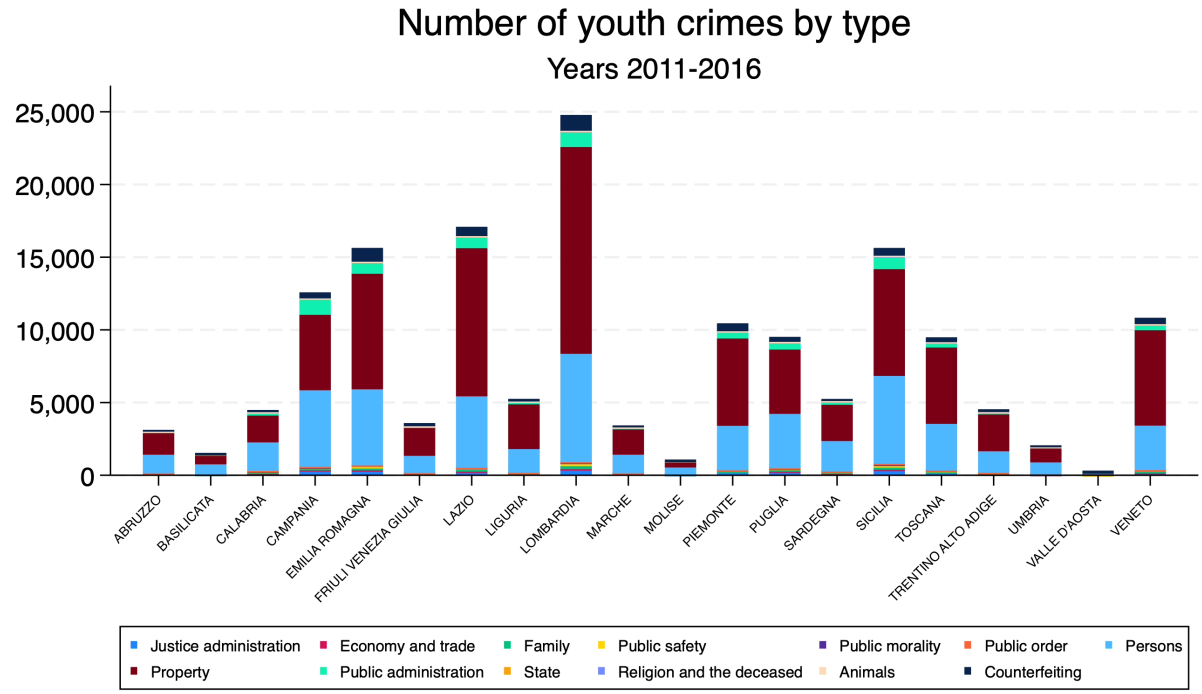

Figure 1

crime rate per 100,000 inhabitants: Trentino Alto Adige, Emilia Romagna, Liguria, Lazio and Molise have higher youth crime rates, as displayed in Figure 2. [1]

2 Data and Methodology

2.1 Data

The analysis of youth crime exploits several resources for the creation of the dataset, creating a panel of data with the twenty Italian regions over six years, from 2011 to 2016. Data on the number of criminal procedures against minors, the population, and the percentage of inhabitants for each region that have at most a middle school diploma was downloaded from the ISTAT website. The latter variable is an inverse measure of adult education,[2] that influences the transmitted importance of schooling to children, and is also a regional indicator of the population education level. The final variable for youth crimes was constructed as a share of the absolute number of crimes over the population of each region and year, multiplied by 100,000. Then, we created the variable on the shares of minors, dividing the number of minors, obtained from the population dataset, by the entire population in each region and year. The unemployment rate for people between 15 and 24 years old for each region from

[1] These figures are constructed with data from the ISTAT dataset, used also for the outcome variable of our model.

[2] As the share of adults with at most middle school education increases, the average years of education of the population are lower (in Italy the average educational attainment is higher than middle school).

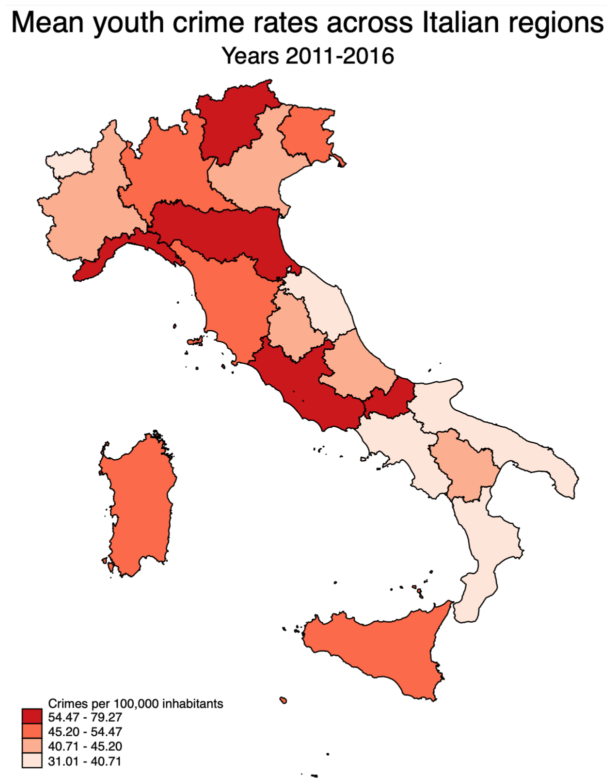

Figure 2

2011 to 2016 was taken from the ISTAT website as well. We chose the range 15 to 24, in order to capture the unemployment that affects minors that could enter the labor market. Using the ISTAT website, a variable including the shares of adolescent second-generation immigrants was also constructed. Since data was only available on the number of births of second generation infants, we considered the birth years from 1993 to 2002 and then, the numbers of people that would have been 14 to 18 years old in the span of years 2011-2016 were summed. These numbers were, subsequently, divided by the population of each region in each year. The variables government effectiveness and social participation were downloaded from the Institutional Quality Index site. In fact, these two variables are used to construct, with other three indexes, a composite indicator that assesses Institutional Quality in Italy [Nifo and Vecchione, 2014]. The first one, government effectiveness, measures the endowment of social and economic structures for each Italian province and the administrative capability of provincial and regional governments in terms of health policies, waste management and environment; the second, social participation, named in the website "voice and accountability", captures citizen's degree of participation in public elections, civic and social associations, the number of social cooperatives, INVALSI test results, and the cultural liveliness, measured in terms of books published. Both variables were taken at the regional level from the dataset.

The variable income was instead constructed from the INPS dataset, taking the median income for each region in each year. Using the income data, we created the inequality vari-

able consistently with how the instrumental variable of predicted inequality is constructed, following the process of Sara Lindgren in the Master's Thesis "Income inequality and crime: Evidence from Sweden" [Lindgren, 2019]. We calculated the income 90^{th} and 10^{th} percentiles for Italy as a whole and then we considered the shares of people in those percentiles for each region and year. The final variable of inequality is a ratio between these two shares. To create the instrumental variable of predicted inequality, it is essential to compute the national inequality growth rate with a leave-one-out approach. We computed the share of people in the bottom income bracket for region i at time t and subtracted it from the sum of the shares of each of the 20 regions. Dividing this number by 19 we obtained the national average share of people in the bottom income bracket at time t excluding region i, and, to compute the growth rate, we divided this number by the same average share (leaving region i out) at time $t-1$. By multiplying such growth rate by the share of people in the bottom income bracket at time $t-1$, we obtain the *predicted* share of people in the bottom income bracket for region i at time t.
The same procedure was applied to all regions and years, and it was also followed analogously for the top income bracket. Finally, we computed the predicted inequality variable as a ratio between the predicted share of people in the top income bracket over the predicted share of people in the bottom income bracket, for each year and region. To see this process in detail, please refer to section 6.1.
Three geographical macro-areas dummies were created, following the standard division of regions into North, South and Center[3]. Moreover, a time-trend was considered, through the variable that contains the years relevant to our analysis.

2.2 Methodology

The research paper is based on linear regressions methods and OLS analysis. In order to study the determinants of youth crime, we considered three main dimensions: poverty, inequality, and social factors. We began by analyzing one static picture of youth crime with the logarithm of income, inequality, unemployment rate, education, social participation and government effectiveness as regressors. Then, we added our controls: the Italian macro-regions, the time-trend spanning the years between 2011 and 2016, and the share of minors.

$$\begin{aligned} Crime_{i,t} = &\beta_0 + \beta_1 Lincome_{i,t} + \beta_2 Ineq_{i,t} + \beta_3 Goveff_{i,t} + \beta_4 Socpart_{i,t} \\ &+ \beta_5 Unemp_{i,t} + \beta_6 Educ_{i,t} + \beta_7 Min_{i,t} + \beta_8 North_i + \beta_9 Center_i \\ &+ \beta_{10} Year_t + \epsilon_{i,t} \end{aligned} \tag{1}$$

Since an estimate of a time-trend for all regions throughout all years is quite vague, we further investigated these results by looking at the interaction between the regional and time controls. However, almost all the variables resulted as not significant; thus, we regarded

[3]The North area includes Liguria, Lombardia, Piemonte, Valle d'Aosta, Emilia Romagna, Friuli Venezia Giulia, Trentino Alto Adige, and Veneto; the Center is formed by Lazio, Marche, Toscana, and Umbria; the South is composed of Abruzzo, Basilicata, Calabria, Campania, Molise, Puglia, Sardegna, and Sicilia.

our previous model as better specified with the controls of a general time trend and macro-regions [4].
A similar approach can be followed using a fixed effect model that captures and eliminates from the coefficients estimates of the explanatory variables biases due to the region specific, but time-invariant, characteristics on youth crime. We decided not to apply this methodology to our final model as, given the little variation in our dataset, regional fixed effects would make all other variables insignificant, capturing a great part of the variation in youth crime (see Table 7).

After analysing our regression, another methodology concept was introduced: the instrumental variable. In fact, we postulated a correlation between the error term and the variable inequality due to a simultaneity issue. In particular, we considered that if episodes of crime are concentrated in certain regions, richer people may be encouraged to move out, whereas poor people cannot, and this could decrease inequality in the regions with more crimes [Boustan et al., 2013]. With a robust Hausman test we confirmed that inequality is endogenous: in fact, if we use the residuals of the first stage regression (see equation 2) as explanatory variables in the second stage regression, they have a p-value of 0.002, confirming endogeneity at the 99% confidence level (see Table 7). Hence, we expect the coefficient for inequality of the second stage regression to be higher with respect to the OLS one, since the possible decrease in inequality due to an increase in crime is cancelled out using the IV methodology.
Thus, we constructed the first stage regression, regressing inequality on predicted inequality:

$$\begin{aligned} Ineq_{i,t} = \gamma_0 + \gamma_1 PredIneq_{i,t} + \gamma_2 Lincome_{i,t} + \gamma_3 Goveff_{i,t} + \gamma_4 Socpart_{i,t} \\ + \gamma_5 Unemp_{i,t} + \gamma_6 Educ_{i,t} + \gamma_7 Min_{i,t} \\ + \gamma_8 North_i + \gamma_9 Center_i + \gamma_{10} Year_t + \eta_{i,t} \end{aligned} \tag{2}$$

From the first stage regression, we see that predicted inequality is a relevant instrument at the 99% confidence level. We can be confident that also the exclusion restriction is respected, as by construction predicted inequality is directly linked to inequality and does not have additional effects on other variables, including youth crimes. Hence, the final model equation will include the fitted values of the inequality variable, related to the first stage regression. Moreover, to make our model more robust, we added the education variable squared in the model.

To further investigate macro-regional differences, we introduced the methodology of heterogeneous outcomes. In particular, we wanted to capture the differences across regions in the effect of some variables on youth crime rate.
The variable unemployment rate was taken into consideration with the idea that unemployment in the south could have a higher effect on youth crime due to the presence of organized crime. However, we saw that this is not the case, by interacting the South dummy with the unemployment rate and obtaining a non-significant coefficient. Hence, we disregarded

[4] The coefficients can be found in the replication package *(interactions1.doc)*.

heterogeneous outcomes in our final model (see Table 7).

In this context, to further investigate both regional differences and the regressors' effects on our outcome variable, we constructed a counterfactual. The main aim of this method is to empirically analyse how, *ceteris paribus*, changing the values of the worst performing region with the best performing one, with respect to a variable of our choice, would change the youth crime rates. The variables that were selected are government effectiveness and inequality.
Firstly, we predicted the baseline model from our dataset. Secondly, substituting the values of government effectiveness of Molise with the ones from Friuli Venezia Giulia, we predicted the counterfactual. Then, we repeated the same procedure, replacing the inequality values of Trentino Alto Adige with the ones from Calabria.

Finally, we introduced another variable: the share of adolescent second-generation immigrants. In particular, the interest is on the interaction between this variable and inequality. In fact, the presence of adolescent second-generation immigrants could augment the effect of inequality, due to the effects of social exclusion and lack of recognition generally spread among this category, leading to a different impact of the variable on youth crime. However, we did not consider the interaction in our model, since it is not significant (see Table 7).
The final model that we took into consideration is given by the following equation:

$$\begin{aligned} Crime_{i,t} =& \beta_0 + \beta_1 Lincome_{i,t} + \beta_2 \hat{Ineq}_{i,t} + \beta_3 Goveff_{i,t} + \beta_4 Socpart_{i,t} \\ &+ \beta_5 Unemp_{i,t} + \beta_6 Educ_{i,t} + \beta_7 Educ^2_{i,t} + \beta_8 Min_{i,t} \\ &+ \beta_9 North_i + \beta_{10} Center_i + \beta_{11} Year_t + \epsilon_{i,t} \end{aligned} \tag{3}$$

3 Results and robustness checks

3.1 First specifications and tests

From the first plain OLS regression, the significant determinants of youth crime seem to be inequality, with a *p-value* equal to 0, government effectiveness, with a slightly higher *p-value*, but still significant at the 95% confidence level, and education, only significant at the 90% confidence level.
A control for each region's share of minors is added, and we see that, contrarily to what was expected, it is not significant. This could be due to the low variation allowed by our data, as the share of minors in each region is not likely to vary much in a span of 6 years. Instead, the controls of a time trend and macro-regional characteristics all prove to be significant at least at the 95% confidence level. In particular, the macro-regional fixed effects estimates indicate the differential crime rates over the Italian territory: compared to southern regions, there are 10.64 less crimes in the center and 12.48 in the north every 100,000 inhabitants per year, keeping all other variables fixed. Such differences are all significant (as we see from the *p-values* of the north and south fixed effects, both lower than 0.05). This indicates a relevant geographical difference in juvenile crime, that is most marked in southern Italy and less in the north. The time trend estimate of -2.638 and its *p-value* lower than 0.01, instead,

indicate that there is a significant decreasing trend in crime rates from 2011 to 2016 in the whole Italian territory, that is not due to the explanatory variables included in the model, but to other factors changing at the national level.

The model is further specified by instrumenting inequality, that is endogenous, with a measure of predicted inequality, as explained in section 2.
Once instrumented in the second stage regression, inequality is still significant at the 99% confidence level and the estimated coefficient is 8.557, against the one estimated in regression (1), of 3.361 (see Table 1). This suggests that the inequality coefficient was underestimated in the previous model, as our hypothesis of simultaneity actually forecast. So, a marginal increase in the inequality ratio determines roughly 8.6 crimes more every 100,000 inhabitants yearly. Besides government effectiveness and education, that remain significant, also the regional unemployment rate seems to be a relevant determinant of youth crime (its *p-value* is lower than 0.05). The instrumented model has a lower goodness of fit, according to the *R-squared* coefficient that decreases from 0.32 to 0.19; however, given the endogeneity of inequality and the relevance of its instrument, it is convenient to maintain this specification to properly identify youth crimes determinants.

While testing the assumption of linearity of our regressors, some relevant results have been obtained. In particular, inspecting graph 8, it can be noticed that the effect of education on crime rate is non-linear, and including education squared in the regression we get that it is significant at the 99% confidence level, as displayed under column (6) of Table 1.[5] Thus, we add it to our final model.
Moreover, while inspecting linearity of income and unemployment rate (see Figures 6 and 7), we noticed that their relationship with crime rate is mainly linear, except at one of the extremes, for high values of crime rate. Thus, we examined our data and found that out of the 7 highest observations of crime rate, 6 of them are of the region Trentino Alto Adige, and the highest four are actually to be considered outliers, as seen in boxplot Figure 9 in the Appendix [6]. Trentino reports extreme values also for the other variables (see Figure 3 to visualize this). Some of them are not intuitively associated with the high crime rate, like high values of average income and low unemployment rate. Moreover, these variables are not significant according to our estimated model, so they would not explain youth crime in Trentino in any case. Instead, there are high values of inequality and a small share of inhabitants without secondary school education; both are significant explainers of a higher crime rate. Since the high level of education in Trentino would explain youth criminality with higher reporting rates, it is irrelevant policy-wise. This suggests that inequality could be the main (and possibly the only) aspect that requires intervention to reduce youth crime in Trentino. This hypothesis is tested via a counterfactual analysis whose output, displayed in Figure 4, shows how a lower inequality level decreases importantly the crime rate in Trentino in each year (light blue bars). We find that if Trentino had the lowest values of inequality of the sample, there would be on average 54.85 crimes less per 100,000 inhabitants each year

[5]Where the squared variable is denoted as *c.educ#c.educ*.

[6]However, we decided not to exclude data on Trentino from our dataset, to avoid reducing even further its variation.

(see Table 5). If we consider that crime rate in Trentino is on average 79.27, inequality determines roughly a 70% reduction in youth crime, a very notable one.

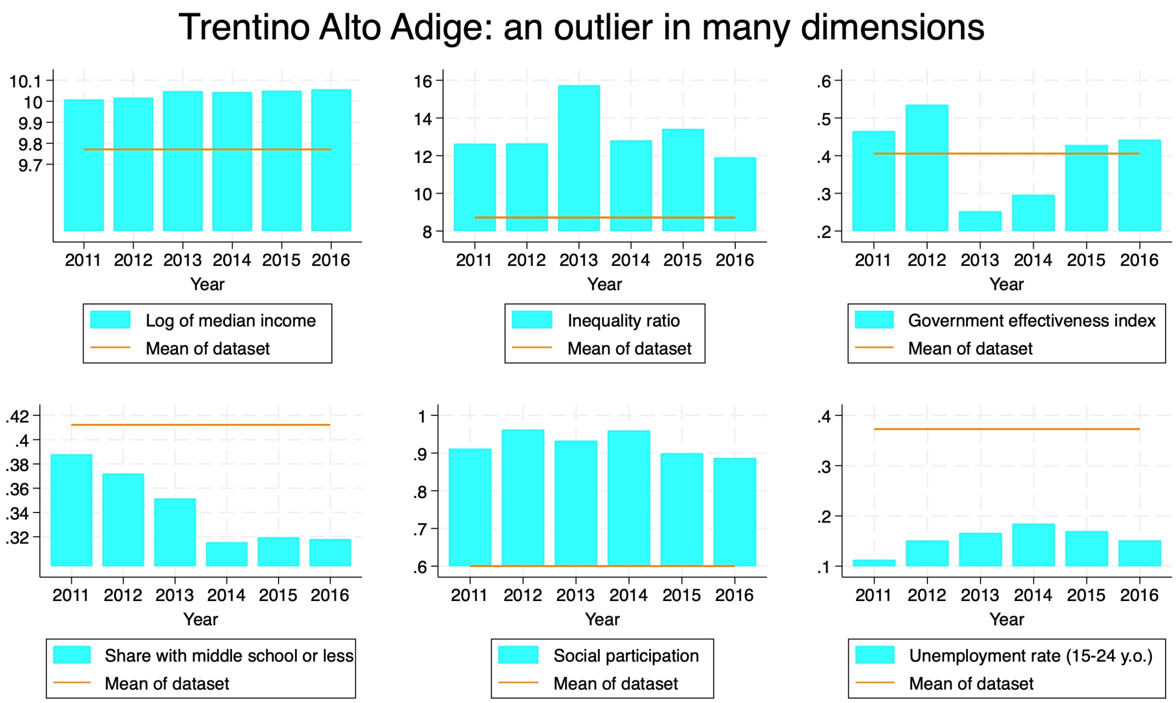

Figure 3

3.2 Final model and further analysis

The final model of choice is regression (3): according to its estimation, the significant determinants of juvenile crime in the Italian territory are inequality, the regional level of education, and government effectiveness, with the former two being significant at the 99% confidence level, and the latter at the 95%. Moreover, there is a significant time trend: on average, there seem to be 2.17 crimes less per 100,000 inhabitants every year in each Italian region.

A marginal increase in the inequality ratio determines 8.3 more crimes every year, for every 100,000 inhabitants, in each region. An interesting insight in this regard is that inequality is more relevant than income, indicating that peer comparison drives the tendency to engage in criminal activities among youngsters, as we hypothesized. The high significance of inequality is further proven by the counterfactual analysis for the region of Trentino Alto Adige explained above, where we saw that reducing inequality can decrease youth criminality by up to 70%.

The effect of education on youth crime rates is a less straight-forward one (recall that the variable is an inverse measure of the regional education level). As education has a quadratic effect on youth crime (displayed in graph 8), to inspect its effect we can compute it as the

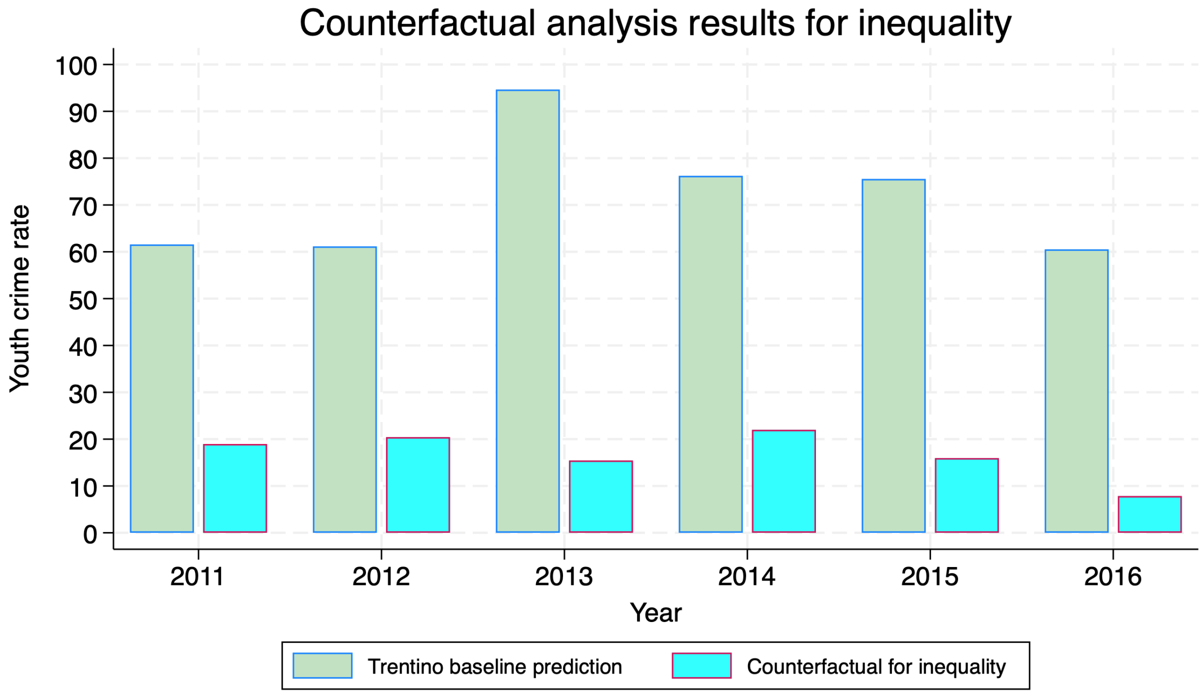

Figure 4

first derivative of youth crime rate on education:

$$\frac{\partial Crime_{i,t}}{\partial Educ_{i,t}} = \beta_6 + 2\beta_7 \ Educ_{i,t} \tag{4}$$

By substituting the estimates $\hat{\beta}_6 = -839.8$ and $\hat{\beta}_7 = 909$, we have

$$\frac{\partial Crime_{i,t}}{\partial Educ_{i,t}} = -839.8 + 1818 \ Educ_{i,t} \tag{5}$$

and we find that the effect of education is positive for all the values greater than 46.19%, and negative otherwise. The regions with a share of inhabitants with at most a middle school diploma around 46% or greater (considering an average of the 6 years) are southern regions, as displayed in Table 4. Thus, for these regions we can say that, as the share of inhabitants with at most middle school education increases, so as the population is less educated, youth criminality increases significantly. This is the kind of effect of education on crime we would have expected. However, this regards only 5 regions out of 20, which is about 1/4 of our sample, as displayed in the cumulative distribution graph in Figure 10. For all the remaining regions, the impact of education is negative. This means that as the number of inhabitants without basic schooling decreases, meaning that the education level increases, crimes increases. We hypothesize that this could be due to a dual effect of our crime rate variable: it captures both how many crimes are *committed* and how many crimes are *reported*. As education can be linked with a more civic behavior, we hypothesize that more educated people could report crimes more often, determining a higher value of our

crime rate variable. This would be the reason why regions with a high level of education - in particular a share of people without secondary school education lower than 46% - present a high crime rate. The effect of reported crimes would be relatively stronger with low values of the education variable (a more educated population), whereas the effect of committed crimes would be stronger with high values of it (a less educated population).
Lastly, a 0.1 increase in the index for government effectiveness decreases the annual crime rate by 2.04. We further investigated the effect of institutional support with a counterfactual analysis. In particular, we inspected how crime rate would change in Molise if it had the government effectiveness of Friuli Venezia Giulia, the two regions where the variable attains its lowest and highest values in the sample, respectively. Please refer to Figure 5 to visualize how crime rate in Molise would be lower in each year (light blue bars) if it had the government effectiveness of Friuli Venezia Giulia. On average over the 6-year time period, an improved government effectiveness could determine 11.32 youth crimes less for every 100,000 inhabitants each year (see Table 6). This change corresponds to a 20% decrease in crime rates (given the average of 54.76 in Molise), which would be a remarkable improvement.

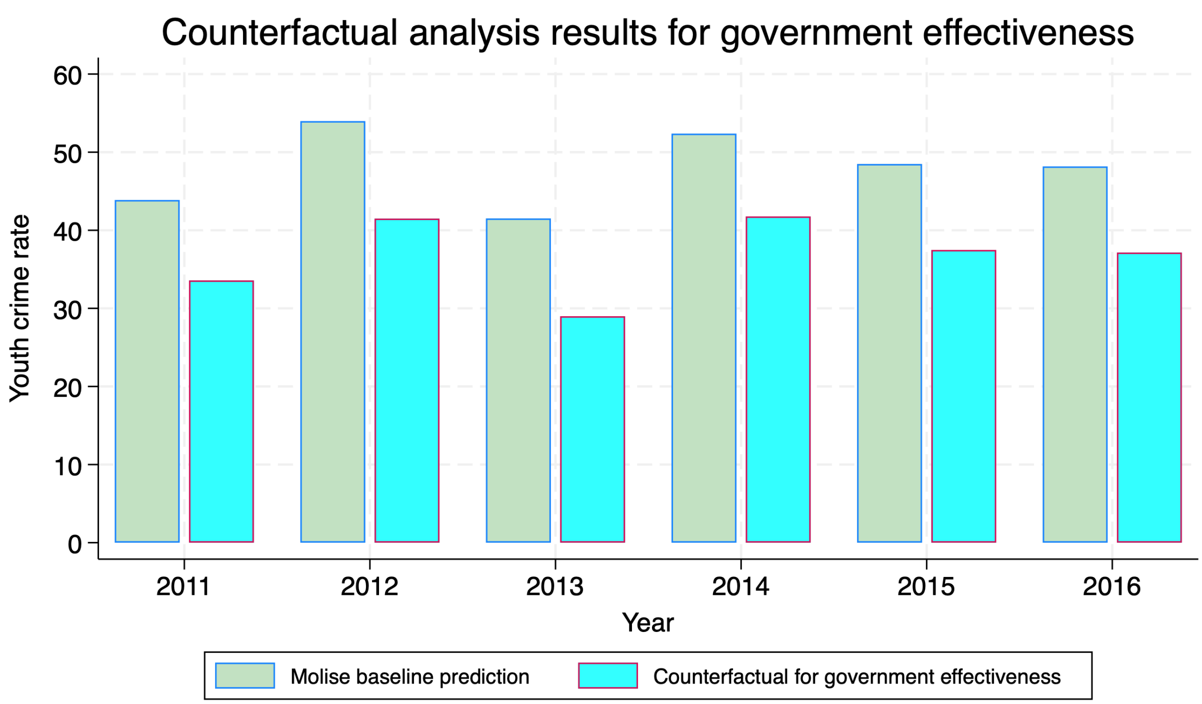

Figure 5

4 Limitations

The limitations in our study are due to both dataset characteristics and model specification. The main limitation of our data is given by the small amount of observations. In particular,

our panel data spans only over six years and twenty regions, for a total of 120 observations, generating issues on the degree of variation in our dataset. Similarly, the short span of time is not entirely adequate to study a complex social phenomenon.
The measurement of youth crime also poses challenges: for a transgression to be accurately measured, the crime must first be reported, and then the transgressor must be identified. Hence, a part of the data goes lost: some regions tend to have lower levels of reporting, which results in levels of youth crime that are possibly lower than reality. Also, some regions may have more efficient policing services to identify criminals and report them, which instead pushes crime rates upwards.
The effect of second generation immigrants on inequality depends on their origins. Indeed, those coming from poorer countries are at higher risk of being marginalized. The data used does not allow us to distinguish between immigrants from richer and poorer countries with respect to Italy. This lack of detail could have driven the insignificance of the interaction between inequality and the share of second-generation immigrants.
The *R-squared* coefficient of the model is 0.257, indicating that it explains roughly 25% of the variation in youth crime rates across Italian regions, which is quite low, probably be due to the limited size of the sample considered.

5 Conclusion

Our analysis seeks to study the determinants of youth crime and conceive potential interventions to mitigate its effects. We initially hypothesised that inequality would be particularly relevant, as well as the quality of institutions, education and unemployment. The final model of choice seems to support our hypothesis on the strong relevance of inequality in explaining youth crime. Instead, the role of income was not as significant, which we had postulated. Education and government effectiveness are both very pertinent, while unemployment and social participation are not.
We believe the role of inequality in youth crime may be due to a “comparison effect”, where minors compare themselves to their better-off peers and use crime as a means of compensation. Instead, we deem that income may not be very significant because of the attention on welfare in Italy. Perhaps, countries where absolute poverty is an issue will experience a stronger connection between income and youth crime. The variable education is significant conceivably because schooling has an effect on discipline of minors, although the estimated association with youth crime is puzzling. Unemployment is not significant in the final model, but it is likely that it would be more relevant if the data were more granular, so as to allow to link it with territorial idiosyncrasies that may lead to youth crime (such as criminal groups). Social participation is perhaps too broad to capture the involvement of minors and may be more important in adulthood.

Policies for the reduction of crime are difficult to implement in part for their scope and in part for the risk of creating popular discontent. The “zero tolerance” approach may work only in emergency situations where repression is the sole alternative, but evidence for its long term use is scarce. In this sense, interventions of broader scope should be considered. Government effectiveness offers an interesting, concrete measure of potentially effective plans of action.

For instance, improving infrastructures and augmenting social facilities can be a tangible, and sensible, way forward. In contrast, the issue of inequality, although very significant, is difficult to solve from a policy standpoint. The most immediate solution is a change in taxation, but this carries issues of desirability and feasibility. In addition, availability of education should be guaranteed in all Italian regions, and provisions on mandatory education may be better enforced in order to decrease dropout rates.
In conclusion, the problem of youth crime is a multifaceted social burden with many causes. An evidence-based approach to its mitigation may be especially useful in reducing wasted resources in meaningless interventions and focusing on the more significant determinants of the phenomenon. It is utopian to believe that youth crime can be solved entirely, but its burden can, and must, be significantly decreased.

5.1 Future approaches

The results from our study give a direction to follow, but are hampered by the low variation in our dataset. The first possibility for further analysis would be to augment our model with variation at a provincial level, or even communal if possible. Also, the inclusion of a larger number of years would allow to see trends that are slow-moving. Furthermore, data on police enforcement and reticence to report violations would give more insights on how efficiently criminals are identified and how often crimes are actually reported.
Another interesting follow-up analysis would be to seek a geographical area of interest and perform a micro-analysis to observe variations at a neighbourhood level. Milan would be a candidate of interest due to the heterogeneity of its population and the inequality within its metropolitan area. We would look at spatial inequality in particular, to better capture the dynamics of minors who compare themselves with classmates and peers living nearby.

References

[Abrams, 2013] Abrams, L. S. (2013). Juvenile justice at a crossroads: science, evidence, and twenty-first century reform. *Social Service Review*, 87(4):725–752.

[Becker, 1968] Becker, G. (1968). Crime and punishment: an economic approach. *Journal of Political Economy*, 76(2):169–217.

[Bonelli, 1992] Bonelli, F. (1992). Appunti sul 'Welfare State' in Italia. *Studi Storici*, 33(2):669–80.

[Boustan et al., 2013] Boustan, L., Ferreira, F., Winkler, H., and Zolt, E. (2013). The effect of rising income inequality on taxation and public expenditures: Evidence from U.S. municipalities and school districts, 1970–2000. *The Review of Economics and Statistics*, 95(4):1291–1302.

[Buonanno, 2003] Buonanno, P. (2003). The socioeconomic determinants of crime. a review of the literature. *University of Milano-Bicocca*, (63):1156–1185.

[Damm and Dustmann, 2014] Damm, A. P. and Dustmann, C. (2014). Does growing up in a high crime neighborhood affect youth criminal behavior? *The American Economic Review*, 104(6):1806–1832.

[Drago et al., 2009] Drago, F., Galbiati, R., and Vertova, P. (2009). The deterrent effects of prison: evidence from a natural experiment. *Journal of Political Economy*, 117(2):257–280.

[Ehrlich, 1973] Ehrlich, I. (1973). Participation in illegitimate activities: a theoretical and empirical investigation. *Journal of Political Economy*, 81(3):521–565.

[Eurispes, 2023] Eurispes (2023). La criminalità: tra realtà e percezione. Report, Ministero dell'Interno, Rome, Italy.

[Felice, 2019] Felice, E. (2019). Economia e società: Il divario nord-sud all'unità. *Meridiana*, 95:39–62.

[Grogger, 1998] Grogger, J. (1998). Market wages and youth crime. *Journal of Labor Economics*, 16(4):756–791.

[Levitt, 1998] Levitt, S. D. (1998). Juvenile crime and punishment. *Journal of Political Economy*, 106(6):1156–1185.

[Lindgren, 2019] Lindgren, S. (2019). Income inequality and crime: evidence from sweden. Master's thesis.

[Nifo and Vecchione, 2014] Nifo, A. and Vecchione, G. (2014). Do institutions play a role in skilled migration? The case of Italy. *Regional Studies*, 48(10):1628–1649.

[Pfeiffer, 1998] Pfeiffer, C. (1998). Juvenile crime and violence in europe. *Crime and Justice*, 23:255–328.

6 Appendix

6.1 Mathematical Appendix

To better understand the construction of our instrument, the equations used are reported.

$$S_{b,t} = \sum_{i=1}^{m} W_{bi,t} \tag{6}$$

$$\overline{W}_{bi,t} = \frac{S_{b,t} - W_{bi,t}}{m-1} \tag{7}$$

$$g_{bi,t} = \frac{\overline{W}_{bi,t}}{\overline{W}_{bi,t-1}} \tag{8}$$

$$\hat{W}_{bit} = W_{bi,t-1} * g_{bi,t} \tag{9}$$

$$PredictedRatio_{i,t} = \frac{\hat{W}_{ti,t}}{\hat{W}_{bi,t}} \tag{10}$$

6.2 Tables and Figures

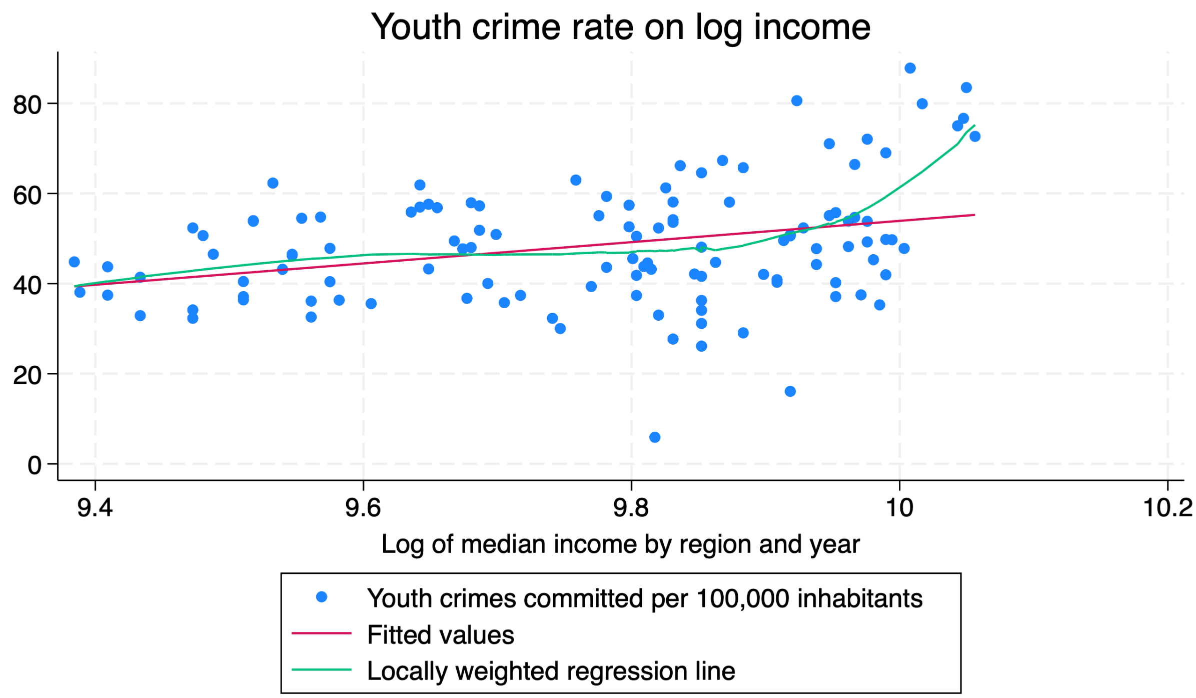

Figure 6

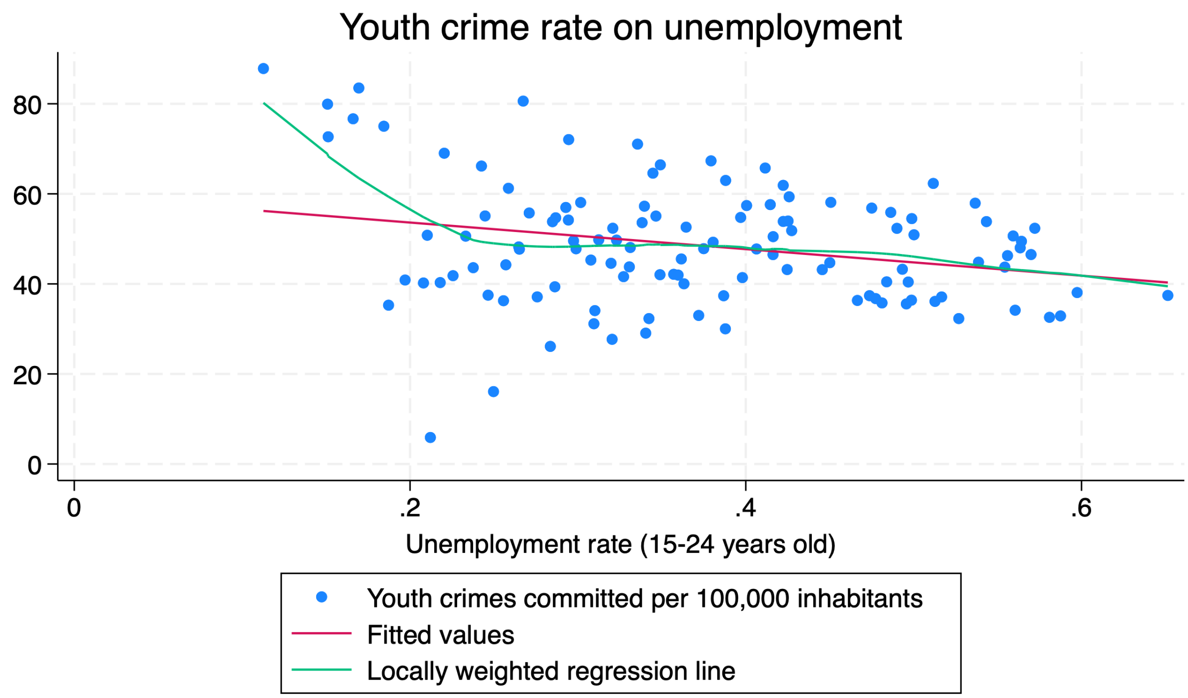

Figure 7

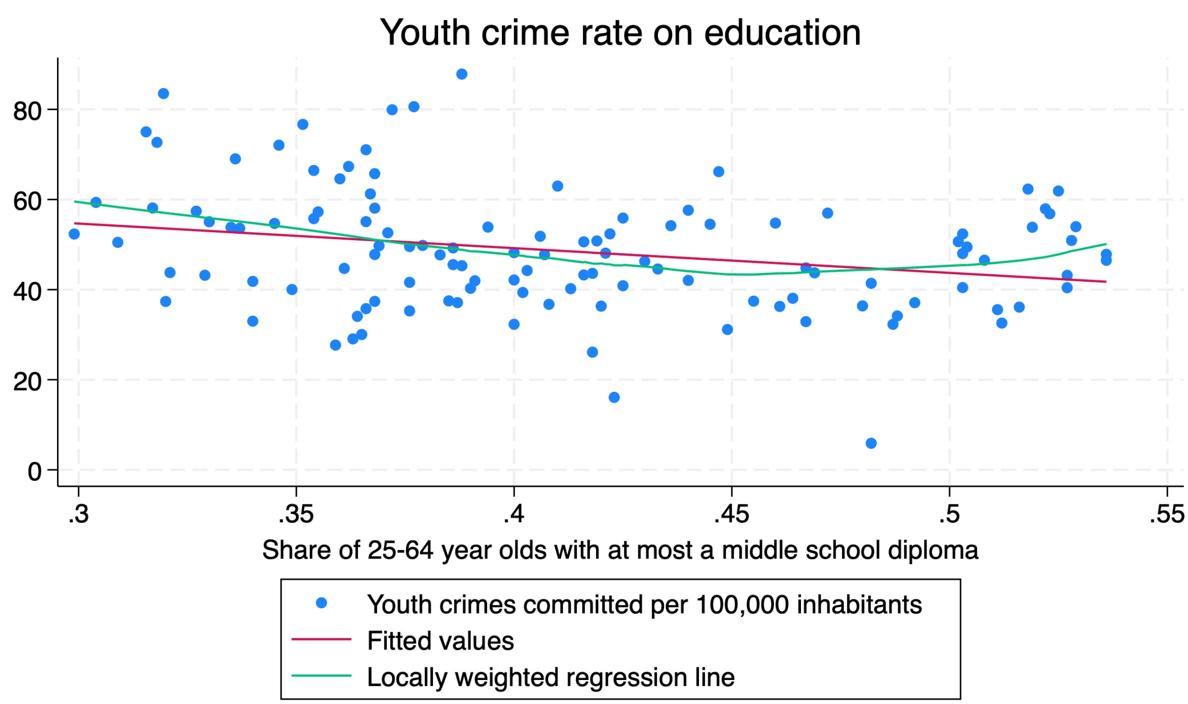

Figure 8

VARIABLES	(1) OLS	(2) Controls 1	(3) Controls 2	(4) First Stage	(5) IV	(6) Final
lincome	0.319	1.852	51.95**	5.205***	2.944	-13.66
	(18.76)	(20.16)	(24.53)	(1.644)	(30.63)	(29.95)
ineq	3.906***	3.822***	3.361***		8.557***	8.300***
	(1.067)	(1.142)	(1.151)		(2.162)	(2.076)
gov_eff	-18.37**	-18.27**	-16.47*	-0.258	-21.05**	-20.40**
	(8.303)	(8.351)	(8.398)	(0.601)	(8.863)	(8.505)
soc_part	-10.55	-10.28	-17.80	-0.572	-14.84	-9.221
	(11.55)	(11.67)	(11.40)	(0.790)	(11.89)	(11.57)
educ	-50.78*	-51.21*	-81.78***	2.332	-77.60***	-839.8***
	(25.71)	(25.90)	(27.75)	(1.977)	(28.85)	(264.4)
unemp_rate	7.306	8.217	34.00*	-1.290	41.35**	29.40
	(17.56)	(18.15)	(19.77)	(1.367)	(20.68)	(20.26)
min_shr		19.52	68.06	20.70***	-67.11	-160.8
		(91.28)	(88.05)	(5.794)	(102.7)	(103.8)
anno			-2.638***	0.0435	-2.305***	-2.166***
			(0.787)	(0.0565)	(0.825)	(0.793)
north			-12.48**	0.700*	-12.93**	-9.175
			(5.816)	(0.413)	(6.041)	(5.938)
center			-10.64**	0.0744	-6.177	-6.598
			(4.968)	(0.364)	(5.386)	(5.168)
pred_ineq				0.684***		
				(0.0988)		
c.educ#c.educ						909.0***
						(313.6)
Constant	43.37	25.67	4,859***	-139.1	4,637***	4,692***
	(179.0)	(197.9)	(1,513)	(108.2)	(1,573)	(1,509)
Observations	120	120	120	120	120	120
R-squared	0.230	0.231	0.320	0.851	0.192	0.257

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 1: Models

	Mean	Std. Dev.	Min	Max
Youth crimes committed per 100,000 inhabitants				
ABRUZZO	41.37	9.70	30.02986	57.26181
BASILICATA	43.49	9.78	32.30051	54.76673
CALABRIA	39.73	4.46	32.89363	44.82604
CAMPANIA	37.26	4.01	32.30368	43.17383
EMILIA ROMAGNA	66.58	13.74	40.29721	80.60388
FRIULI VENEZIA GIULIA	53.29	2.52	49.5564	55.75286
LAZIO	55.22	3.56	50.49609	59.35764
LIGURIA	60.28	8.32	44.7074	67.33006
LOMBARDIA	46.85	3.64	40.20592	49.80684
MARCHE	40.48	8.81	27.69831	52.60659
MOLISE	54.76	6.68	43.25317	62.98305
PIEMONTE	43.55	14.08	16.08999	53.86988
PUGLIA	39.84	6.22	32.58153	47.82856
SARDEGNA	54.18	5.51	48.01468	61.89353
SICILIA	53.27	5.21	46.51875	62.31219
TOSCANA	46.88	12.61	29.06108	66.17529
TRENTINO ALTO ADIGE	79.27	5.66	72.67886	87.83557
UMBRIA	42.37	7.44	32.99522	55.06905
VALLE D'AOSTA	31.01	14.06	5.894737	44.57903
VENETO	40.94	5.71	35.27071	50.61167
Total	48.53	13.47	5.894737	87.83557

Table 2: Youth crime rate

	Mean	Std. Dev.	Min	Max
Inequality ratio				
ABRUZZO	7.10	0.12	6.926923	7.236434
BASILICATA	7.21	1.12	6.384615	9.402778
CALABRIA	6.58	0.85	5.559767	7.73444
CAMPANIA	6.64	0.36	6.174452	7.073479
EMILIA ROMAGNA	10.46	0.73	9.510586	11.19306
FRIULI VENEZIA GIULIA	11.82	0.68	10.86034	12.64103
LAZIO	7.96	0.25	7.638258	8.385946
LIGURIA	8.68	0.54	8.003676	9.353983
LOMBARDIA	10.56	0.22	10.35938	10.95276
MARCHE	8.27	0.32	7.800614	8.661074
MOLISE	7.35	0.77	6.046154	8.061225
PIEMONTE	10.08	0.69	9.272966	11.1035
PUGLIA	6.67	0.20	6.328592	6.891339
SARDEGNA	7.96	0.33	7.432314	8.315271
SICILIA	6.71	0.52	6.166013	7.614458
TOSCANA	8.77	0.32	8.506023	9.379768
TRENTINO ALTO ADIGE	13.19	1.33	11.90625	15.72917
UMBRIA	8.53	0.81	7.611702	9.905405
VALLE D'AOSTA	8.51	1.96	5.633333	10.86667
VENETO	11.24	0.48	10.58171	11.80027
Total	8.71	2.00	5.559767	15.72917
Government effectiveness index				
ABRUZZO	0.42	0.06	.3223627	.4917095
BASILICATA	0.22	0.06	.111615	.3066944
CALABRIA	0.20	0.03	.1563006	.2325911
CAMPANIA	0.44	0.08	.3210737	.5268851
EMILIA ROMAGNA	0.57	0.04	.4946007	.61145
FRIULI VENEZIA GIULIA	0.65	0.04	.6042599	.6895086
LAZIO	0.44	0.08	.3362358	.5106663
LIGURIA	0.54	0.05	.459399	.6016791
LOMBARDIA	0.57	0.06	.488154	.6370807
MARCHE	0.56	0.07	.4351813	.6330267
MOLISE	0.09	0.04	.0583513	.1699044
PIEMONTE	0.47	0.04	.417915	.5158777
PUGLIA	0.31	0.05	.260984	.4010988
SARDEGNA	0.25	0.05	.1587988	.3062001
SICILIA	0.19	0.05	.1398944	.2837094
TOSCANA	0.55	0.05	.48136	.609624
TRENTINO ALTO ADIGE	0.40	0.11	.2522286	.5348751
UMBRIA	0.49	0.05	.4147685	.5538524
VALLE D'AOSTA	0.15	0.15	0	.3638859
VENETO	0.60	0.08	.4797638	.6772476
Total	0.41	0.18	0	.6895086

Table 3: Inequality and government effectiveness

	Mean	Std. Dev.	Min	Max
Share of 25-64 year olds with at most a middle school diploma				
ABRUZZO	0.36	0.01	.349	.383
BASILICATA	0.43	0.02	.4	.46
CALABRIA	0.47	0.01	.455	.482
CAMPANIA	0.50	0.02	.48	.527
EMILIA ROMAGNA	0.36	0.02	.336	.39
FRIULI VENEZIA GIULIA	0.37	0.03	.335	.419
LAZIO	0.32	0.01	.299	.337
LIGURIA	0.36	0.00	.36	.368
LOMBARDIA	0.39	0.02	.368	.413
MARCHE	0.38	0.02	.359	.418
MOLISE	0.43	0.02	.406	.472
PIEMONTE	0.40	0.02	.386	.423
PUGLIA	0.52	0.01	.511	.536
SARDEGNA	0.52	0.01	.503	.528
SICILIA	0.51	0.01	.502	.529
TOSCANA	0.41	0.03	.363	.447
TRENTINO ALTO ADIGE	0.34	0.03	.3155	.388
UMBRIA	0.33	0.01	.32	.34
VALLE D'AOSTA	0.45	0.02	.418	.482
VENETO	0.40	0.02	.376	.425
Total	0.41	0.06	.299	.536

Table 4: Education

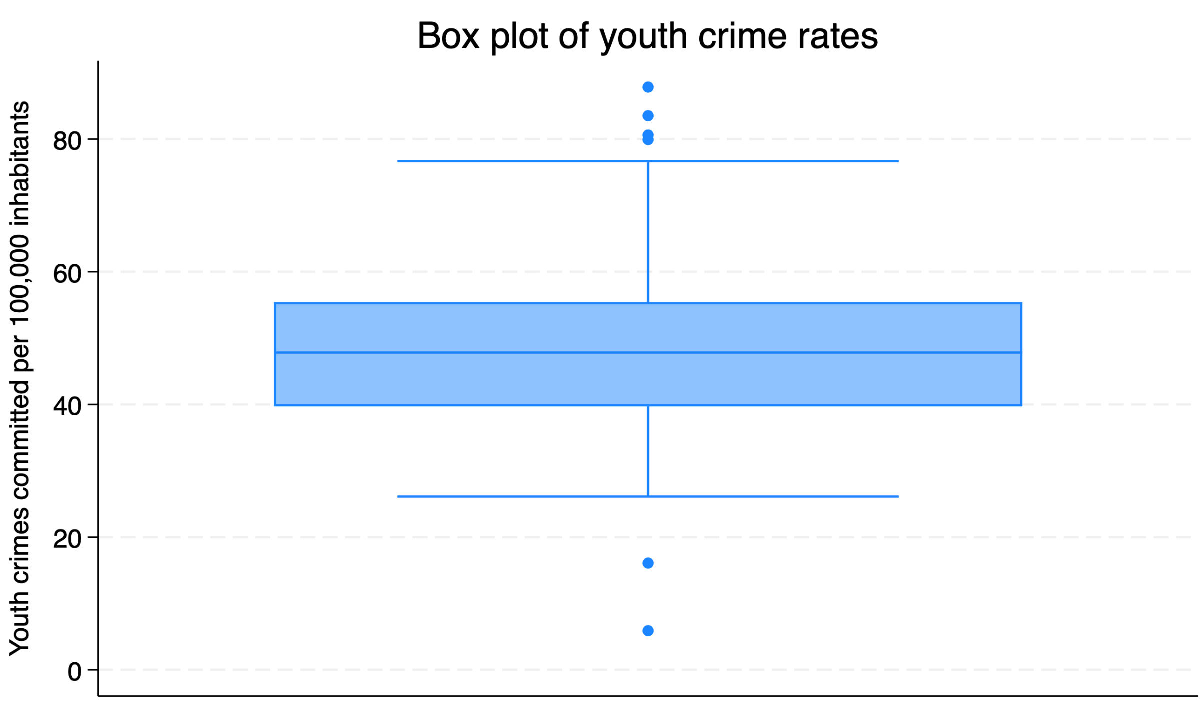

Figure 9

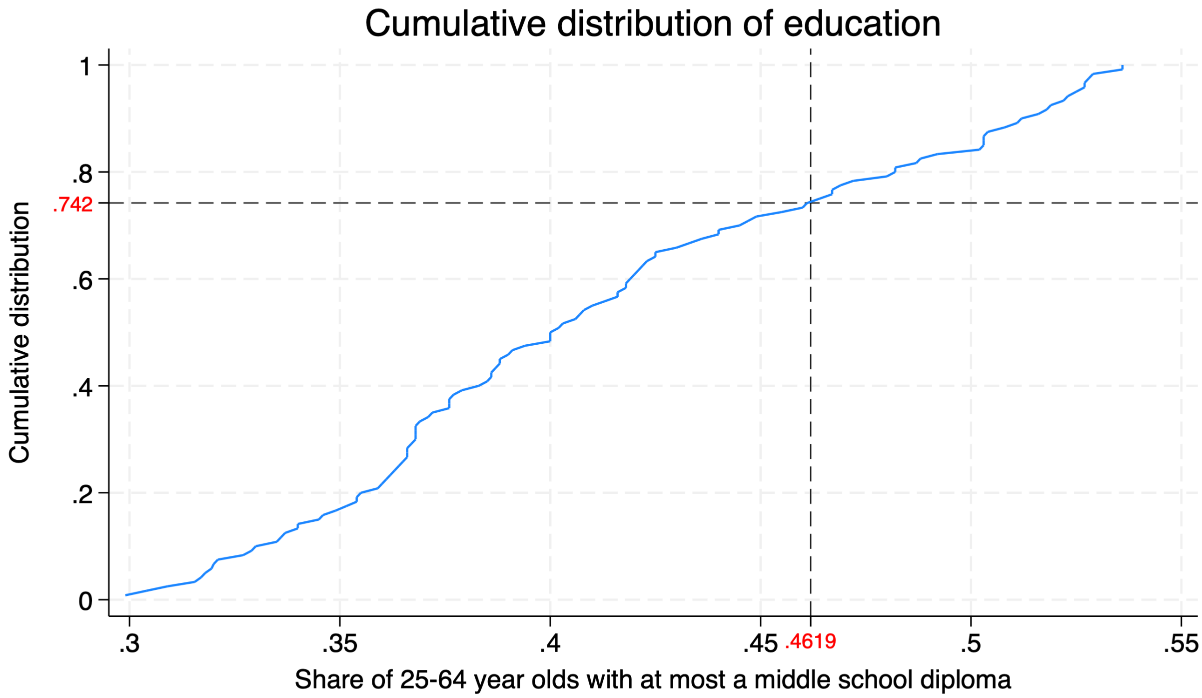

Figure 10

	Ineq. counterfactual analysis results
N	120
Trentino baseline prediction	71.621 (13.422)
Trentino counterfactual prediction	16.773 (5.050)
Difference between baseline and counterfactual	54.847 (13.933)

Table 5: Inequality counterfactual

	Gov. eff. counterfactual analysis results
N	120
Molise baseline prediction	48.067 (4.779)
Molise counterfactual prediction	36.745 (4.871)
Difference between baseline and counterfactual	11.322 (0.951)

Table 6: Government effectiveness counterfactual

VARIABLES	(1) Fixed Effects	(2) Robust Hausman test	(3) Heterogeneous outcome	(4) Interactions 2
ineq	-0.0585	8.557***	10.41***	8.291***
	(1.162)	(2.004)	(2.945)	(2.054)
lincome	0.353	2.944	-22.15	-7.108
	(34.69)	(28.38)	(39.86)	(30.26)
gov_eff	-17.58	-21.05**	-20.78**	-19.54**
	(13.36)	(8.215)	(9.295)	(8.463)
educ	36.42	-77.60***	-60.73*	-860.4***
	(77.56)	(26.74)	(32.95)	(262.8)
soc_part	-0.0226	-14.84	-11.31	-11.02
	(12.00)	(11.02)	(12.82)	(11.63)
unemp_rate	27.77	41.35**	69.04**	27.12
	(18.71)	(19.17)	(30.44)	(20.21)
min_shr	1,256***	-67.11	-69.00	-132.5
	(462.4)	(95.23)	(108.1)	(106.8)
0b.south#co.unemp_rate			0	
			(0)	
1.south#c.unemp_rate			-62.34	
			(46.16)	
north		-12.93**	-34.70**	-7.946
		(5.599)	(17.37)	(6.052)
center		-6.177	-28.26*	-5.342
		(4.992)	(16.81)	(5.369)
anno	-0.0221	-2.305***	-1.845*	-1.687*
	(0.976)	(0.765)	(0.945)	(0.976)
res_ineq_fs		-7.481***		
		(2.405)		
c.sec_gen_shr#c.ineq				-0.205
				(0.244)
c.educ#c.educ				931.1***
				(311.4)
Constant	-127.3	4,637***	3,946**	3,667*
	(1,943)	(1,458)	(1,737)	(1,933)
Observations	120	120	120	120
R-squared	0.155	0.376	0.110	0.272
Number of rgn	20			

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 7: Other regressions