INSTITUTO DE ECONOMÍA
FACULTAD DE CIENCIAS ECONÓMICAS
Y ADMINISTRATIVAS

Lecture : Econometric Theory II
Professor : Tomás Rau TA : Valle and Barrueto
Date : May 28, 2023
Student : Valentina Andrade
email : vandrade@uc.cl

Report 3

Matching and Difference-in-Difference

Abstract

The following report contains the assignment nº3 of the Econometric Theory II course directed by Professor Tomás Rau and the assistants Eduardo Barrueto and Nicolás Valle.

The report consists of 2 parts: the first section aims to review several matching techniques using historical data of Germany in order to replicate the paper Satyanath et al., 2017 . The second section seeks to replicate the article by Jayachandran et al., 2010, where the authors studies the impact of sulfa drugs in the decline of mortality rates. This exercise improved our knowledge of difference in difference approach, particularly using several specification we learn about fixed effects and triple difference in difference.

In the following web link you can find the GitHub repository that allows you to reproduce the whole homework.

Question 1

a. Summary - *Bowling for Fascism: Social Capital and the Rise of the Nazi Party*

- Principal Question of the Paper: The paper studies the relationship between social capital, specifically measured by denser social networks and the rise of the Nazi Party. The authors investigate whether higher association density, reflecting a vibrant civic society, facilitated the faster entry of individuals into the Nazi Party.
- Contribution: Relevance lies in its exploration of how social capital can undermine and contribute to destroying a democratic system. It contributes to the existing literature by providing detailed cross-sectional data demonstrating social capital's role in the rise of autocratic regimes. The findings challenge the prevailing notion that the weakness of German civic society facilitated the rise of the Nazis, suggesting instead that the opposite is closer to the truth.
- Conclusions:
    1. The presence of denser social networks, indicated by higher association density, facilitated a quicker influx of individuals into the Nazi Party.
    2. The effect of association density on Nazi Party entry rates was substantively significant, leading to a notable difference in the speed of party membership acquisition.
    3. The positive relationship between association density and Nazi Party entry encompassed various types of associations, including civic, military, "bridging," and "bonding" associations.
    4. The influence of association density on Nazi Party entry was contingent upon the political context, with higher density being less correlated with a faster entry in federal states with more stable governments.
    5. By employing panel analysis and instrumental variable exercises, the authors accounted for potential confounding factors and strengthened the validity of their findings.
    6. The study underscores the pivotal role of social capital, represented by denser social networks and vibrant civic societies, in the rise and growth of the Nazi Party.

b. Empirical strategies

The empirical analysis strategy of the paper can be visually presented through a Direct Acyclic Graph (DAG), as depicted in Figure 1. The explanation of the colors for each box can be found in the figure caption.

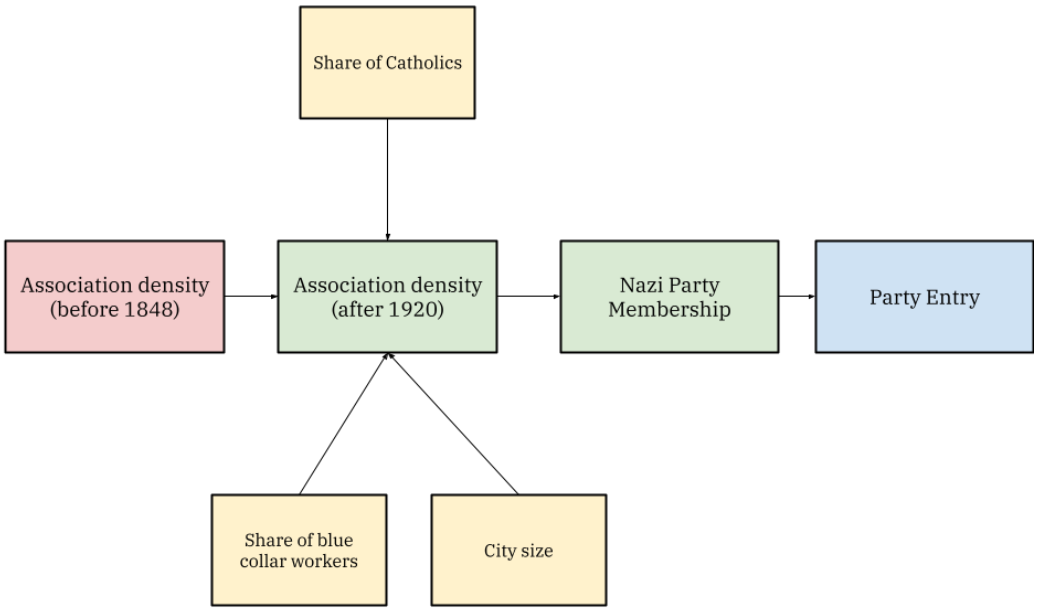

Figure 1: Direct Acyclic Graph (DAG) proposed. The red box represents an exogenous explanatory variable; the yellow boxes represent the confounders; the green boxes the explanatory variables of interest; the blue box the main dependent variable.

Now, we will analyze in particular which are the identification assumptions for the matching carried out in the article. First, it is important to note that in the matching the estimation is about the effect of being above the median in the measure of association (association density). For the estimation, we have the following assumptions:

In the matching analysis conducted in the article, several identification assumptions are made to estimate the causal effect of being above the median in association density. These assumptions include:

1. Conditional independence assumption (CIA): The potential outcomes (Y^1 and Y^0) are independent of the treatment assignment (D) given the covariates (X).

2. Common support assumption: The probability of receiving treatment ($D = 1$) given the covariates (X) falls between 0 and 1 for each stratum. These two assumptions allow for the identification of the **average treatment effect** (ATE) using the following identity equation, which considers the switching equation for each value of Y:

$$E[Y^1 - Y^0|X] = E[Y^1|X, D = 1] - E[Y^0|X, D = 0] = E[Y|X, D = 1] - E[Y|X, D = 0]$$

   By incorporating the common support assumption, the estimator for the ATE is calculated as the integral of the difference between the expected outcomes for treated and control groups, weighted by the probability distribution of the covariates.

   Additionally, to identify the average treatment effect on the treated (ATT), it is sufficient for the treatment to be conditionally independent of the potential outcome under no treatment (Y^0), and there must be units in the control group for each treatment stratum.

3. Furthermore, the assumption of independent and identically distributed (*i.i.d.*) sampling ensures that the potential outcomes and treatment statuses of individuals are unrelated to those of other individuals in the population. This assumption is part of the stable unit treatment value assumption (SUTVA).

These identification assumptions collectively form the basis for estimating the causal effects using the matching strategy.

Discussion of identification assumptions

1. CIA

One of the crucial arguments put forth by the authors is that association density (social capital) is exogenous to the rise of the Nazi Party. To support this claim, two points are addressed:

1. The density in the association comes from an exogenous source of variation unrelated to Nazi party building and hegemony.

The exogenous nature of association density is supported by the historical context. During a critical period, political and social conditions influenced the formation and strength of associations, including those related to the Nazi Party. However, the repeal of restrictions on association formation before 1948 was unrelated to the rise of the Nazis. Early associations had a liberal and nationalist agenda, devoid of a militaristic or xenophobic ideology. These associations encompassed diverse social groups, serving as "bonding" and "bridging" social capital. Therefore, the association density post-1920, specifically related to the Nazis, arises from factors independent of the Nazi party's development and hegemony, further establishing its exogenous nature.

2. There are variables that would be expected to affect both association density and Nazi party success. Adding these confounders to the analysis does not change the results at least in a direction.

In order to mitigate potential confounding factors, the analysis takes into consideration variables such as the proportion of Catholics, the proportion of blue-collar workers, and city size. Historical factors, including anti-Catholic sentiment and the Catholic Church's role, contribute to lower association density in cities with higher Catholic populations. The lower inclination of blue-collar workers to participate in formal associations and the association of club membership with the bourgeois class further necessitates controlling for the proportion of blue-collar workers. Additionally, city size is considered to account for differences in association density related to economies of scale and the concentration of workers. By addressing these confounders, the study provides a more comprehensive understanding of the relationship between association density and the success of the Nazi Party, accounting for historical and socio-political dynamics.

However, it seems to me that these criteria are not satisfied. In particular, in the matching (and throughout the analysis) it is used to capture the variables that could confound the size of the population rather than the three indicated (only because they are associated). Similarly, it seems to me that there are other confounders that may be affecting the analysis that is not considered and could violate the CIA. For example, the union density rate could be associated with higher associativity and lower adherence to Nazi values (as it is more linked to blue-collar workers). More on this I will discuss more in the last point of the question, along with the possibility that the exogeneity criterion (which is not tested in matching) could be satisfied.

Also, the CIA assumption, also referred to as unconfoundedness or selection-on-observables in the literature, is an important concept in the analysis. It implies that there is no hidden bias or confounding due to unobserved factors that could influence both the treatment (association density) and the outcome (success of the Nazi Party). By considering observable variables, such as religious, socio-political, and trade union associations, and excluding others, the analysis attempts to account for potential selection in observables. This helps ensure that the observed relationship between association density and Nazi Party success is not confounded by unobserved factors and provides a more reliable basis for drawing conclusions.

2. Common support assumption

The common support assumption states that every individual must have a positive probability of being assigned to each treatment level. It is difficult for this assumption not to hold. Let's consider a contradiction where a group has no potentially positive probability of having an association density greater than the median (treatment). However, all individuals potentially have the possibility to participate, even if the probability is very small due to different characteristics. Therefore, it is reasonable to believe that this assumption is satisfied.

3. i.i.d

The i.i.d. assumption, which stands for independent and identically distributed, is a reasonable assumption to make in this study. Random sampling techniques were likely employed to select the study sample, ensuring that each individual has an equal chance of being included. Additionally, rigorous study design and data collection methods would have aimed to minimize any systematic biases or correlations between the potential outcomes and treatment statuses of individuals. These efforts contribute to the plausibility of the i.i.d. assumption in this empirical analysis.

c. Balance

Assuming the independence assumption holds and the mean potential outcomes are the same for each type of group (represented in the tables), we would not expect the observable characteristics of the groups to be identical. This concept is known as balance, and it means that if the means of the covariates are the same for each group, we can say the covariates are balanced, and the two groups are exchangeable in terms of those characteristics.

We utilize this concept to assess the need for addressing selection based on observables. It is essential because the propensity matching score focuses only on the necessary covariates. It estimates a maximum likelihood model of the conditional probability of treatment (using methods like logit or probit to ensure common support). Subsequently, the predicted values from that estimation are used to condense those covariates into a single scalar known as the propensity matching score.

Therefore, it is likely that we may observe covariate imbalances before the matching process. However, the primary objective post-matching is to compare units that, based on observables, had very similar probabilities of being assigned to the treatment group, even though these units differed in terms of actual treatment assignment. In other words, the groups become observationally equivalent.

Variable	Obs	Mean	Std. dev.	Min	Max
pcNSe ry$_s td$	115	-.1383384	.8080479	-1.356702	2.413341
lnpop25	115	10.88616	1.210946	8.772301	13.95746
share$_c ath$25	115	.3848245	.3086468	.0204942	.9277812
bcollar25	114	.4501207	.1081341	.2619286	.7415896

Table 1: Balance clubs$_p c_A M = 0$

Variable	Obs	Mean	Std. dev.	Min	Max
pcNSe ry$_s td$	114	.1395519	1.14885	-1.356702	4.123702
lnpop25	114	9.632706	.803283	7.632401	12.71185
share$_c ath$25	114	.27721	.3269606	.0107911	.9669098
bcollar25	113	.4283429	.1070711	.1832511	.7049965

Table 2: Balance clubs$_p c_A M = 1$

Variable	Obs	Mean	Std. dev.	Min	Max
pcNSDAP285	12	2.619885	1.181077	1.131576	5.247384
pog20s	11	0	0	0	0
clubs$_a ll_p c$	12	2.343516	1.764657	.3508168	5.570214
lnpop25	12	10.36301	1.363938	9.149741	13.42861
share$_c ath$25	12	.1073901	.1202537	.0270202	.3748851
bcollar25	12	.404666	.0883692	.1913592	.599965
latitude	12	51.80582	1.473921	48.51667	53.51207
longitude	12	10.82957	2.427314	7.155354	14.00729

Table 3: Balance $pog1349 = 0$

Variable	Obs	Mean	Std. dev.	Min	Max
pcNSDAP285	80	4.145141	4.728574	.2011494	24.73789
pog20s	80	.175	.3823644	0	1
clubs_all_pc	80	2.280236	1.285104	.0206492	7.941706
lnpop25	80	10.60346	1.272134	7.719574	13.45915
share_cath25 |	80	.4596098	.3358534	.0252205	.9570405
bcollar25	80	.3904015	.0836946	.1832511	.5776032
latitude	80	50.2649	1.336736	47.66667	53.25
longitude	80	9.227436	1.799526	6.12954	13.77259

Table 4: Balance $pog1349 = 1$

Table 1 and 2 present the balance of the variables of interest for the analysis presented in the paper in Table 10. The first table corresponds to the untreated group, while the second table corresponds to the treated group. In turn, the first variable presented is the response variable, while the others correspond to the covariates. The situation is the same for Tables 3 and 4, which correspond to the balance of the analysis presented in Table 12, on anti-Semitism. The only difference is that in this case we have three dependent variables of interest listed at the beginning of tables 3 and 4.

Just to illustrate the importance of balance, we will discuss the first tuple of tables.

As we can see by comparing the first two tables, the dependent variable has on average a very important difference. This is to be expected and gives good signals about what will be done in the analysis since we expect the treated and untreated groups to present significant differences in the response variable. Now if we look at the control variables we should expect them to be similar between the two groups. As we can see, they are close but not exactly the same, and only in the case of the control that refers to blue-collar workers is it similar in both mean and variance. However, in the case of the population size and the proportion of Catholics, this does not quite match. However, let us remember that this is a pre-matching analysis, so the empirical strategy that we will use in theory should resolve the balance of the groups.

d. Replication and interpretation -Table 10

In all the models presented in Table 5, it can be seen that in cities where the association density is above the median, the entry into Nazi parties is greater than in cities of similar size and below the median association density, all of these analyses being statistically significant.

Now it is difficult to compare the models for two reasons. The first is that not only are the nearest neighbors moving or changing in each of the models (from (1) to (4)) but also the controls that are added in each of them. The second reason is that unlike the first four models which are estimated by propensity matching score, model (5) uses another strategy which is exact matching.

Overlooking the first reason we can notice that as more nearest neighbors are added and all the controls are incorporated (indicated in the conditional independence assumption) the propensity matching score estimation shows an effect, although statistically significant, getting smaller and smaller in effect size. In particular it can be seen that in cities where there is a diversity of association above the median the entry to the Nazi party is 0.26 SD higher than in cities below the median, controlling for the base model variables and latitude.

A second interesting point to note is that the propensity score and exact matching estimates (which share the same covariates) are quite similar. Let us discuss this a bit

	m1 b	m2 b	m3 b	m4 b	m5 b
$I(ASSOC_{all} > median)$	0.452^{***}	0.339^{*}	0.288^{**}	0.257^{**}	0.275^{*}
Observations	229	229	227	227	227

Table 5: Matching estimation:dependent variable: average (standardized) NSDAP entry per capita, 1925 - January 1993

Propensity score estimation and exact matching are two different approaches used in empirical analysis to achieve comparability between treated and control groups. Although they differ in their implementation, in certain contexts, the results obtained with propensity score matching can be similar to those obtained with exact matching.

Propensity score matching consists of estimating the probability of belonging to the treatment group (treatment propensity) for each individual in the sample. Then, individuals are matched or selected from the treatment and control groups who have similar propensities. This allows selection bias to be controlled by balancing observable characteristics between the groups and comparing the treatment effect more accurately.

On the other hand, exact matching involves finding observations in the control group that are identical to observations in the treatment group in terms of the relevant variables. This implies that each individual in the treatment group has an exact "twin" in the control group. By exactly matching the characteristics of interest, perfect balance on those variables is ensured and any possible selection bias is eliminated.

However, in certain contexts, propensity score estimation can generate results similar to exact matching. This can occur when the propensity score manages to adequately capture and balance the relevant characteristics between the treated and control groups. In other words, if propensity score matching succeeds in generating groups with similar propensities and a balanced distribution of observable variables, the estimated treatment effect can be comparable to that obtained by exact matching.

Similarity in results between propensity score matching and exact matching can occur when the variables used to estimate the propensity score are highly predictive of the treatment and, at the same time, are related to the outcome of interest. In such cases, propensity score matching can generate groups that are comparable in terms of important characteristics, and the estimated effect may be similar to what would be obtained with exact matching.

e. Replicate Table 12 and discussion about OLS

One question the authors ask is whether more anti-Semitic places might have produced communities that were more closed and impenetrable to non-Germans (and thus the density of association is much higher and more likely to vote for the Nazis).

In order to prove this, the authors use the pogromo, [1] As a way of seeing if there is entrenched anti-Semitism in the sample. Thus, in Table 6 we see different models predicting the effect of this pogromo on (1) attacks against Jews; (2) voting for the Nazi party; (3) association density.

	Progroms 1920s	Progroms 1920s	NSDAP votes	NSDAP votes	ASSOC	ASSOC
Pogrom in 1349 (OLS)	.2155717***		1.969413**		.1085932	
	.0740608		.9588024		.4077375	
Pogrom in 1349 (Matching)		.175***		1.844828***		-.2248508
		.0396928		.6190096		.5536652
Observations	91	91	92	92	92	92

Table 6: Historical Anti-semitism

Interpretation

The results obtained by matching are presented. First of all we can notice that in the cities where there were persecutions of Jews in medieval times, they predict positively the persecutions and attacks on Jews during the Nazi era. In particular, in cities where there was persecution of Jews in medieval times, checks on Jews during the Nazi era are 0.175 SD higher than in cities where there was no persecution, controlling for the rest of the variables in the model. Likewise, in the cities where there was persecution of Jews in medieval times, the Nazi party shows 1.84 SD more votes than in those where there was no persecution, controlling for the variables of the model. Finally, as expected, there is no evidence of a relationship between anti-Semitism and association density. This is good because otherwise we might have some omitted variable that makes one of the main explanatory variables endogenous. It is important to note that in the case of the first two dependent variables the coefficients are not only positive but also statistically significant, while in the case of the last models related to association density these models show coefficients that are statistically significant.

On OLS vs Matching

In this case, an ordinary least squares (OLS) regression model followed by matching is performed in order to address possible endogeneity and selection bias issues in the analysis.

OLS is used as a first approximation to estimate the effect of interest without considering the endogeneity problem. However, since the variable of interest (association density) may be correlated with other unobserved variables that also influence the outcome (voting for Nazis), there is the possibility of selection bias and incorrect results.

To address this problem, matching is performed, which is a technique of matching treatment and control units with similar characteristics. By using matching, we seek to create groups that are comparable in terms of observable characteristics and more effectively control for the influence of unobserved variables.

By comparing the results of OLS with those of matching, it is possible to assess whether there is any significant difference in the estimates and whether matching helps to reduce selection bias. If the results are consistent between the two methods, this provides greater confidence in the robustness of the results and the ability to control for selection bias.

Given that the effect of pogrom has the same direction and significance on the variables of interest (association density and Nazi party entry) in both OLS and Matching, but the effect size is larger in OLS, the following can be inferred:

1. OLS results are overestimated: This means that the OLS model is capturing part of the pogrom effect that is not directly related to the variables of interest. It is possible that OLS is incorporating some bias or unintended influence in the effect estimate.

[1] A pogrom against Jews refers to an act of systematic violence and persecution directed specifically at the Jewish community. Historically, pogroms have been episodes of violent attacks, looting, murder and other forms of oppression of Jews at different times and places. These discriminatory and hostile acts are often driven by anti-Semitic prejudice and have been recorded throughout history, being especially prevalent in Eastern Europe and Russia in the 19th and 20th centuries

2. OLS assumptions may not be fully met: Although the OLS results may be significant and have the expected direction, the difference in effect size between OLS and Matching suggests that there may be unobserved factors that are influencing the results that are not controlled for in the OLS model.

3. Matching may be a more appropriate strategy: Since Matching seeks to create comparable groups and more effectively control for unobserved variables, the fact that the effect size is smaller in Matching may indicate that it is providing a more accurate and reliable estimate of the actual effect of the pogrom on the variables of interest.

Overall, the discrepancy in effect size between OLS and Matching suggests that Matching may be a more appropriate strategy in this case to control for selection bias and obtain more precise estimates of the effect of pogrom on the variables of interest.

f. Discussion about Inverse Probability Weighting (IPW)

Advantages of using IPW estimation

The use of Inverse Probability Weighting (IPW) in this case can provide several advantages:

1. Selection bias control: IPW can help to effectively control selection bias by assigning higher weights to treated or controlled units that are more similar in terms of the confounding variables. This allows for more precise estimates of the causal effect of pogrom on the variables of interest.
2. Maximizing the data: IPW uses all the information available in the data by assigning weights to each unit according to its inverse probability of receiving the observed treatment. This makes it possible to take maximum advantage of the information contained in the data and reduce potential bias.
3. Estimation flexibility: IPW can be used with different estimation models, such as generalized linear models or logistic regression models, which provides flexibility in choosing the most appropriate model for data analysis.
4. Consideration of uncertainty: IPW also allows the uncertainty associated with the estimation of inverse probabilities to be taken into account. Robust confidence intervals can be calculated to reflect this uncertainty and provide a more complete assessment of the results.
5. Treatment of outliers: From the normalization of the weights, the weights assigned to each unit are based on the inverse probabilities of receiving the observed treatment. To avoid problems of scale and numerical stability, it is common to normalize the weights so that they sum to one in the data set. This ensures that the total weights are comparable and avoids distorting the estimates. This can help to reduce the impact of outlier observations in the weighting process and improve the stability and validity of the estimates.In particular, normalization is performed by default in STATA and R (Cunningham)

In summary, the use of IPW in this case can help control selection bias, make the best use of the available data, provide more accurate and reliable estimates of the causal effect of the pogrom on the variables of interest, and avoid problems with very large or small probabilities (outliers).

Selection

The model with logit specification is selected because the information criteria (AIC/BIC) show a better fit than in the case of the probit. It is also important to point out that for model (2) now estimated by IPW we do not calculate the standard errors because variance matrix is nonsymmetric or highly singular. Therefore we cannot know the significance of the coefficients of model (2). The models are presented in Table 7.

Comparison with Table 12 of the paper

First, it is important to note that it is possible to compare the results obtained by using Inverse Probability Weighting (IPW) and Propensity Score Matching (PSM). We will now compare the results

As can be seen in the first row, in the case of the model where we want to see the effect of the persecution of Jews in medieval times on the persecution and attack during the Nazi era, we can see that in both IPW and PSM the coefficient is the same. However, we do not know if it is statistically significant in IPW (because of what was discussed above).

In the case of model (4), the results are equal in direction and significance, but the effect size is slightly lower in the case of IPW. This result means that in the case of IPW the effect of the medieval pogrom is milder on Nazi party voting. The change in the result could be due to the fact that, since the whole sample is occupied, the cities that were previously eliminated from the PSM now attenuate the effect, from which it is inferred that they should have a lower Nazi party voting rate.

In the case of model (6), we expect this variable not to be statistically significant, which is the case for both strategies. That is, both are statistical zeros, although in the case of IPW the effect size is larger. However, again, what is relevant is to verify that there is no confounder related to the association density.

	(2) Progrom 1920s	(4) NSDAP votes	(6) ASSOC
ATET			
Pogrom in 1349	.175	1.758529***	-.7504192
	.	.5794847	1.719427
POmean			
Pogrom in 1349=0	-5.06e-26	2.386612***	3.030655*
	.	.299247	1.725488
TME1			
ln(pop) in 1925	.1991545	.228726	.228726
	.	.1879304	.1879304
Share Catholics in 1925	1.843441	1.846164***	1.846164***
	.	.5783905	.5783905
Share Blue-collar workers in 1925	-.504827	-.5511318	-.5511318
	.	3.399702	3.399702
latitude	-.3005664	-.328339**	-.328339**
	.	.1450279	.1450279
longitude	-.0624926	-.0709233	-.0709233
	.	.1239309	.1239309
Constant	14.66113	15.84677**	15.84677**
	.	7.995991	7.995991
Observations	91	92	92

Table 7: Inverse Probability Weighting - Historical Anti-semitism

g. Discussion about the identification strategy of Table 10

Is it necessary to evaluate the post-matching balance for Propensity Score Matching?

Yes, it is necessary to perform a post-matching balance check after applying Propensity Score Matching (PSM). The goal of PSM is to create comparison groups that are similar in terms of the observed covariates, to reduce selection bias between treated and untreated groups.

Once the matching is done, it is essential to verify whether the observed covariates are balanced between the treated and untreated groups, as we did in classes. Post-matching balance refers to the similarity in the characteristics of the matched individuals after applying the PSM.

To assess post-matching balance, various metrics can be used, such as standardized mean difference, statistical significance test or density plots.

What if the groups are very unbalanced?

If the imbalance is small in absolute terms and the differences are not statistically significant, the imbalance may not have a significant impact on the results and may be considered acceptable (at the discretion of the researcher).

However, if the imbalance is large or the differences in the covariates are important, there may be a residual bias in the estimation of the causal effect. In such cases, additional strategies need to be considered to address the imbalance, such as the use of additional covariates, restrictions on matching, or weighting techniques.

In which case it is not necessary to check post-matching?

This occurs in the case of Exact Matching, where individuals are matched exactly on all covariates, a post-matching balance check would not be necessary. This is because Exact Matching ensures that the observed covariates are completely balanced between the treated and untreated groups, since individuals identical on all characteristics are selected.

In Exact Matching, each treated individual is exactly matched with an untreated individual who shares the same values on all covariates. As a result, there is no variability in the covariates between the matched groups and no need to assess post-matching balance.

However, it is important to note that Exact Matching can be restrictive in practice, as it may be difficult to find untreated individuals identical for each treated individual on all covariates. In such cases, more flexible matching methods can be considered, such as Propensity Score Matching, which allows a greater degree of flexibility in matching and can provide comparable results even without exact matching.

In summary, while in Exact Matching it is not necessary to check the post-matching balance due to the nature of exact matching, in other matching approaches, such as Propensity Score Matching, it is essential to perform a post-matching balance assessment to ensure the validity of the results.

h. Using *blopmatching* for Table 12

Table 8 presents the results of the BLOP matching (BLOP) estimation analysis of the effect of anti-Semitism on Nazi-era Jewish persecution, Nazi party voting, and association density.

First, as we can notice the BLOP estimates are close to those of the PMS. In particular, we can see that for model (2) the effect size is exactly the same, but the standard error is not. That is why in the case of the results presented in this table, the medieval pogrom seems not to produce significant differences in the persecution and attack of Jews during the Nazi era. This result seems strange given that on the one hand, other robustness exercises show the same result along with the fact that the standard error grows very large for this variable.

For model (4) although the effect size is not the same and instead in the BLOP it is larger, we can notice that the direction and significance of the effect are the same. In fact, the standard error in the case of the effect of the persecution of the Jews during the medieval period on the votes for the Nazi party is quite smaller, so the estimation, in this case, is more accurate.

For model (6) we do not have the results hold, this means that the association density and the persecution of Jews in medieval times have no relationship. This is seen in that controlling for the rest of the variables in the model, we can see that the effect is not statistically significant.

As a comment, it is important to point out that the estimations were made with the Euclidean metric since it is not only better known but also allowed us to calculate the standard errors. At the same time, it is important to note that the results of a balance in this case of analysis are difficult to conduct because the number of observations is already quite unbalanced. For example, the control group has only 11 observations, whereas the treated group has 80.

	(2) Progrom 1920s	(4) NSDAP votes	(6) ASSOC
atet	.175	2.107949***	-.4479266
	.133418	.3262076	.3554075
Observations	91	92	92

Table 8: BLOP matching - Historical Anti-semitism

Advantages and disadvantages of BLOP match

We will now discuss some of the advantages of BLOP matching estimation, based on the results obtained in this paper, but also on the implementation of this method in other contexts (J. Díaz et al., 2015; Satyanath et al., 2017; J. D. Díaz et al., 2021).

In the context of the BLOP matching approach proposed by J. Díaz et al., 2015, two notable advantages can be identified:

1. The BLOP matching approach does not rely on arbitrary rules to determine the counterfactual units and weighting scheme in advance. Instead, it formulates an optimization problem to determine these parameters, allowing for a more objective and data-driven approach. This feature distinguishes it from many other nonparametric methods commonly employed in the literature.

2. The BLOP matching estimator directly incorporates the optimization of post matching covariate balance. By simultaneously optimizing the weights and matches, the estimator aims to achieve an optimal balance of covariates between treated and control units. This integrated approach enhances the robustness and efficiency of the matching process, leading to more reliable treatment effect estimation.

Overall, the BLOP matching approach offers the advantages of principled parameter determination and improved post matching covariate balance optimization, contributing to its appeal and potential usefulness in empirical applications. Now the specific characteristics of the BLOP matching estimator can be highlighted:

1. Estimation: The BLOP matching estimator produces estimates that closely align with the benchmark and are comparable to those obtained using the nearest neighbor approach with a larger number of neighbors. Notably, several studies employing the BLOP estimator have demonstrated that an average of only 2 neighbors per observation is often sufficient (ranging from 1 to 8). This advantage arises from the fact that the BLOP estimator does not require a predetermined choice of the number of matches.

2. Performance: The performance of the nearest neighbor estimator tends to diminish as the number of neighbors increases, while the BLOP estimator consistently delivers slightly higher estimates than the nearest neighbor estimator. Importantly, the BLOP estimator avoids the issue of bias escalation that occurs with the nearest neighbor approach since it optimally determines the number of neighbors to consider.

3. Inference: The BLOP estimator demonstrates good asymptotic properties. In repeated simulations (500 repetitions), the authors of the paper never reject the null hypothesis of equality in the distribution of treated units. This suggests that the BLOP estimator provides reliable inferential results.

4. Specification: The BLOP estimator exhibits remarkable performance in the presence of misspecified propensity scores, such as when interactions are included (column 2). Compared to other estimators, the BLOP estimator demonstrates superior performance in capturing relevant characteristics when the propensity score is overspecified, including both linear terms and interactions (column 3). This robustness to specification issues further enhances the applicability of the BLOP estimator.

Overall, the BLOP matching estimator showcases favorable characteristics in terms of estimation accuracy, performance with varying numbers of neighbors, reliable inferential properties, and robustness to propensity score specification. These strengths position the BLOP estimator as a promising approach to empirical applications

i. Discussion about the identification strategy of the paper

Three criticisms will be presented that may be relevant when discussing whether the authors succeed in constructing a sufficiently credible identification strategy. These are: (1) the use of the propensity matching score as a source of bias; (2) the ad hoc sample selection that is constructed; (3) some unused confounders relevant to identification.

Why not use propensity matching scores (PMS) for matching? (King & Nielsen, 2019)

King and Nielsen, 2019 presents a very strong argument for evaluating the use of PMS, in particular, they indicate that its use makes the causal strategy very weak because it actually produces model dependence. Model dependence is defined as choosing what the causal effect is going to be by making an ad hoc choice of model specification. This means that if a regression is done, for example, with a quadratic term or a linear relationship, the causal relationship would change. Similarly, if you select one part of the sample or another the causal effect would also change. In particular, we know that without matching the imbalance of the variables is a source of model dependence and thus an avenue of discretion available to researchers. Now, with matching we can supposedly achieve balance and solve this problem, i.e., not relying on the discretion of the researcher but rather on statistical techniques.

This is how the authors indicate that the PMS seeks to assimilate a complete randomization that in its best case will be suboptimal because the controls and the treated will not be exact. To be more specific, "pruning at random increases imbalance because the pruning process is independent of X" (King & Nielsen, 2019). Being independent of X PMS does not attempt to make the PMS completely balanced. The authors also point out that this technique does not solve the dimensional problem because the more variables the problem worsens and therefore the standards of this technique are quite low: PMS never optimizes, it only projects a complete randomization. [2]. In conclusion, it is preferable to use techniques such as blopmatching or exact matching alone.

Sample selection

The authors construct a sample in which they exclude political and religious organizations, which could bias the effect of association density on Nazi party entry. At the same time, although they indicate that rural areas are more prone to Nazi ideas, they exclude rural areas from the analysis. In this sense, the authors generate an ad hoc sample that is useful to test a certain previously defined causal effect.

Confunders

When analyzing certain variables that might influence both association and success in joining the Nazi party, it is indicated that workers might be less likely to join the Nazi party. However, this analysis does not consider the organizations that constitute the workers: the unions. In this sense, the authors make an ad hoc choice of organizations to measure associativity but do not consider that at the time the workers' organizations are active and very strong since before 1848. This is a third example of how the causal effect is chosen by choosing the specification of the model (model dependence) and therefore makes the causal strategy not very credible.

[2] It is important because complete randomization is dominated in its capacity and power by fully blocked experiments. Exact matching and blopmatching are assimilated to fully blocked. For more discussion see the authors' paper.

Question 2

(a) Summary, hypothesis, data and results

The article "Modern Medicine and the Twentieth Century Decline in Mortality: Evidence on the Impact of Sulfa Drugs" by Jayachandran et al., 2010 examines the impact of sulfa drugs in the reduction in infectious disease mortality. [3] Using US vital statistics at national-level and state-level that include rates of mortality and incidence by treated and non-treated diseases by sulfa drugs, they estimate times series model and difference in difference analysis in order to prove the main hypothesis (H_1): when the sulfa drugs become more available, mortality falls.

More specifically, using **Difference-in-difference** the paper found support for the following four hypotheses [4]:

Hypothesis	Prediction	Result
$H_{1.1}$	The reduction of mortality was larger for treated diseases than the control diseases	Drugs led to significant mortality declines for diseases that were treatable with sulfa drugs
$H_{1.2}$	The reduction of mortality was faster for treated diseases after the introduction of sulfa drugs	Mortality started to decline faster with the introduction of sulfa drugs
$H_{2.1}$	The reduction of maternal mortality was larger for treated diseases for white people than black people (SES)	Maternal Mortality from diseases treatable with sulfa drugs fell less for black people than for whites.
$H_{2.2}$	The reduction of maternal mortality was faster for treated diseases for white people than black people areas	There is evidence of steeper declines in mortality for whites than blacks after the introduction of sulfa drugs
$H_{3.1}$	The reduction of maternal mortality was larger for treated diseases in urban areas than in rural areas	Cities experienced statistically significantly larger declines in maternal mortality after 1937 in urban than in rural areas
$H_{3.2}$	The reduction of maternal mortality was faster for treated diseases in urban areas than in rural areas	Cities experienced statistically significantly faster declines in maternal mortality after 1937 in urban than in rural areas

Table 9: Hypothesis summary, models and result

Finally, it is important to highlight that the main contribution of the paper refers to the understanding of the role of medical advances in reducing mortality and increasing life expectancy in the mid-twentieth century.

[3] Some of the causes of death reviewed by the authors are maternal death as a proxy for puerperal fever death, pneumonia mortality, and scarlet fever mortality.

[4] It is important to note that the results of the two items in Table 9. The first is that H_1 can be disaggregated into two hypotheses: one hypothesis referring to the difference in intercept and the other to a difference in slope. The second is that the results can be more precise with respect to the percentage change for each cause of mortality. However, in order to make the analysis more parsimonious, we only refer to the more general hypotheses.

(b) Econometric strategy

Identification strategy

The authors indicate in their article:

"Our estimates are based on the identifying assumption that post-1937 mortality declines for treated diseases, beyond those that occurred for the control disease, are due to sulfa drugs." (Jayachandran et al., 2010, p.120)

This sentence, which in addition to being uninformative and unspecific, is tautological [5]. We will therefore be guided by, in theory, which identification assumptions the difference-in-differences (DiD) model should meet, and then look for information in the paper that can provide support and/or support for these assumptions.

The DiD model, in seeking to estimate the effect of a treatment, must meet three general assumptions (1-3) plus one specific to the DiD estimator (4)

(1) SUTVA

This assumption establishes that the treatment effect on a particular unit does not depend on the treatments received by other units in the sample. In other words, SUTVA assumes that there is no interference or external effects between treated and untreated units.

In the paper this assumption can be seen reflected in that in the period from the mid-1930s to the mid-1940s the only medical advance was the diffusion of sulfonamide-based drugs. this tells us that there is no interference from other new drugs that might have been created in the period that would affect mortality.

(2) Overlap

As we reviewed in question (1) we must ensure that each individual has a chance of receiving the treatment (with non-negative probability). In the paper, this is reflected in the timing of the sulfa drug discovery. It can be taken as exogenous and, because sulfa drugs could not be patented, diffusion was rapid.

Thus, the diffusion of sulfa drugs was very effective because of their mass production. Not only because the rapidity of the diffusion could not mark a before and after, but access to the drug could not have been probable for the entire population. [6].

Conditional Independence Assumption (CIA)

This assumption establishes that after conditioning on covariates when no unobservable variable affects both treatment assignment and the potential outcomes, the potential outcomes are conditionally independent of the treatment. That is, there are no unmeasured confounders.

This assumption is more strongly justified in the analysis. The authors indicate that "there is strong evidence from clinical trials that sulfa drugs were effective against some infectious diseases, but not others. Therefore, we are able to use infectious diseases that were impervious to sulfa compounds as a comparison group in our analysis, netting out the effects of other factors that may have coincidentally lowered mortality around 1937."

(4) Parallel trends

The parallel trends assumption is the most distinctive of the DiD estimator. This assumption establishes that, in the absence of the intervention, the trends observed in the treatment group and the control group would have been parallel over time. In other words, the two groups would have similar trends prior to the introduction of the treatment.

Unfortunately, this assumption is not directly discussed in the paper. However, it is pointed out that until 1937 (the introduction of sulfa drugs) there were no factors that differentiated the treated and untreated groups.

Are feasible the identification assumptions?

We will now discuss the feasibility of each identification assumption.

(4) Parallel trends

According to Cunningham, 2021 the parallel trends assumption cannot be tested directly but is obviously violated if in the absence of treatment the trends would have changed equally.

Indirectly we could see a test of this assumption if we were to join figure 2 (treated group trend) and figure 3 (control group trend, in this case tuberculosis).

[5] The authors indicate that their identification assumption is the very effect they want to find

[6] It is interesting to note that this drug, despite its toxicity, was sold without prescription until 1938. This element will be discussed later in question (b)

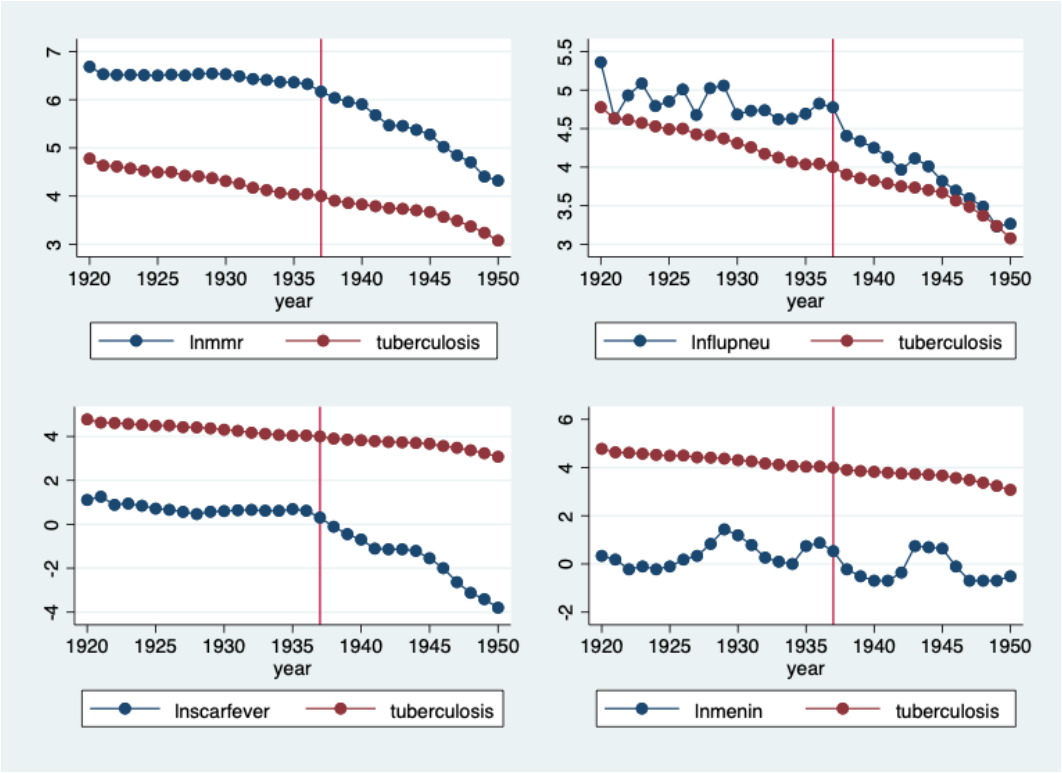

Figure 2: Parallel trends

As we can notice at least for maternal mortality and scarlet fever mortality we can notice that the trend before treatment (introduction of sulfa drugs) was the same. We cannot say exactly the same for influenza/pneumonia because of its oscillations. What the authors say is that these may be due more to influenza than pneumonia, so in the case of this research it would only be white noise. However, the robustness of this comment is never checked even though there are later data where both mortalities (influenza and pneumonia) are disaggregated.

(3) CIA

Although the authors mention them as variables that could be affecting the potential outcome, they are not considered in the regression so they could confound the effect of sulfanomide entry on the decline in mortality. These are: increased living standards, better nutrition, and public health initiatives that improved water supply, sanitation systems, and household hygiene.

In addition, these covariates are also likely to correlate with the effect that sulfonamides may have, especially those related to the advancement of the health system since their efficacy and implementation depends on it.

Another possible counfonder has to do with the increase in the price of sulfonamide. Although it is indicated that its price is low ($35 to $100), this difference could produce differences in the possibility of access to the drug. For example, if we did not condition on price, the potential outcome would not be the same between blacks and whites (thinking that measures SES). Similarly, sales of sulfa drugs (and their price) will be higher in places where the mortality rate of sulfa drug treatable diseases is higher.

Finally, the authors consider these as a comparison group but it is common in epidemiology to consider them as a control when evaluating infectious disease settings. [7]. These are the comorbidities or pre-existing conditions, or in this case, the incidence rate of chronically ill by state or nationally. For example, the treated and untreated groups could not be equal if comorbidity generates statistical differences between the two groups (which is quite likely).

(2) Overlap One element discussed by Goodman-Bacon, 2021 concerns when the pretreatment effect becomes complex, which makes it difficult to enforce that the study population has a nonnegative probability of being treated. In this paper this assumption is moot.

In the first instance, the sulfa drug was discovered in 1908, almost 30 years before where the start of treatment is indicated [8]. However, sulfonamides were introduced into the healthcare market before 1937 for two reasons

[7] Just think of the case of COVID-19

[8] In 1937, the German pharmaceutical company Bayer introduced the first sulfonamide-based drug called "Prontosil rubrum"

1. At that date, the FDA did not require human clinical trials for drug distribution. So, even as the paper states, in the absence of a patent, production began quickly.
2. The paper misrepresents which is the first major clinical trial (which was actually 2 years before the treatment). Also, connecting to the above, clinical trials in cells and mice started many years earlier.

With this in mind, it does not seem so clear that the start of treatment was in 1937 or even unique. Probably the access to sulfonamides was segregated for some part of the population at the beginning of their sale.

(1) SUTVA

This assumption is not criticized if it is considered as evidence that between 1930 and 1940 the only medical advance was the diffusion of sulfonamide-based drugs. Based on the epidemiological literature review this is true.

The only thing that could be judged is the stability of the treatment on the outcome. The criticism, although minor, is that sulfa drugs were quickly withdrawn from the market because of their toxicity and because they produced high resistance to the bacteria that produced infectious diseases. That tells us that their effect can only be captured during that period, and not after the mid-1940s. However, it also tells us that the authors' choice of evaluation periods was well done, otherwise, this basic assumption could also be put into play.

DiD equations

The following variables are included in the models of Difference in Difference (DiD)

Variable	Description	Type	Model
$\log(M)_{dt}$	Log of the mortality rate for disease d and year t.	Outcome	(5) to (8)
$\log(MMR)_{ict}$	Log of the maternal mortality ratio for geographic category c in state i in year t.	Outcome	(9) to (10)
$Treated_d$	Indicator variable for whether disease d is a treated disease	Treatment	(5) to (8)
$Post-1937_t$	Indicator variable equal to zero for t between 1925 and 1936 and equal to one for the period 1937 to 1943	Time application treatment	(5) to (10)
$Year_t$	Continuous year variable	Tendency	(5) to (10)
$Urban_c$	Indicator variable for whether the observation is a city	Group	(9) to (10)
$State$	Indicator variable for each state in US	Control	(7) to (10)

Table 10: Variables included in models (5) to (10) in \cite{} paper

The equations will now be explained. In order to make the analysis more parsimonious, we will concentrate on those coefficients that are related to the hypotheses indicated at the beginning of this question. Now we can briefly indicate that in all the models

- β_0 corresponds to the slope, the value of which tells us the expected value of the outcome in the absence of treatment, before the treatment occurs.
- The coefficients related to $Year_t$ are changes in the trend of some coefficients.
- $gamma_i t$ corresponds to the fixed effect of $State$.
- $mu_i t$ corresponds to the fixed effect of $State$ interacted with the trend.
- Coefficients that do not contain interaction (e.g., when we look only at the $Treated$ or $Post-1937$ variable) are not interesting for the analysis of the causal effect of sulfa drugs.

Models (5) and (6) seek to examine the effect over the mortality rates (ATT) at national level ($H_{1.1}$, captured by $H_{1.1}$).

In particular, the difference between models (5) and (6) is that the latest one allows us to see the change for both the intercept ($H_{1.1}$) and the slope ($H_{1.2}$, captured by β_2) after 1937. The models are looked for β_1, β_2 to be jointly statistically significant and negative.

(5)

$$\begin{aligned}\log(M)_{dt} =& \beta_0 + \beta_1 \text{ Treated }_d \times \text{ Post-1937 }_t + \beta_2 \text{ Treated }_d \times \text{ Year }_t \\ &+ \beta_3 \text{ Treated }_d + \beta_4 \text{ Year }_t + \beta_5 \text{ Post-1937 }_t + \varepsilon_{dt}\end{aligned}$$

(6)

$$\begin{aligned}&\log(M)_{dt} = \beta_0 + \beta_1 \text{ Treated }_d \times \text{ Year }_t \times \text{ Post } - 1937_t + \beta_2 \text{ Treated }_d \times \text{ Post-1937 }_t \\ &+ \beta_3 \text{ Treated }_d \times \text{ Year }_t + \beta_4 \text{ Treated }_d + \beta_5 \text{ Year }_t + \beta_6 \text{ Post-1937 }_t \\ &+ \varepsilon_{dt}.\end{aligned}$$

Models (7) and (8) do something similar but with state-level data. This allows us to include within-state comparisons and make the regression coefficients more precise. The hypothesis analyses are maintained for both models, only that for model (8) the roles of $beta_1$ and $beta_2$ are reversed.

(7)

$$\begin{aligned}\log(M)_{idt} =& \beta_0 + \beta_1 \text{ Treated }_d \times \text{ Post-1937 }_t + \beta_2 \text{ Treated }_d \times \text{ Year }_t + \beta_3 \text{ Treated }_d \\ &+ \beta_4 \text{ Year }_t + \gamma_{it} + \varepsilon_{idt}\end{aligned}$$

(8)

$$\begin{aligned}\log(M)_{idt} =& \beta_0 + \beta_1 \text{ Treated }_d \times \text{ Year }_t \times \text{ Post-1937 }_t + \beta_2 \text{ Treated }_d \times \text{ Post-1937 }_t \\ &+ \beta_3 \text{ Treated }_d \times \text{ Year }_t + \beta_4 \text{ Treated }_d + \beta_5 \text{ Year }_t + \gamma_{it} \\ &+ \mu_{it} \times \text{ Year }_t + \varepsilon_{idt}.\end{aligned}$$

Models (9) and (10) are quite interesting as the treated group changes: now it is no longer the diseases that are treatable by sulfonamides (*Treated*) but *Urban*, i.e. the indicator of whether the observation is a city or not. This is in order to assess whether cities faced a statistically significant decline in mortality (in this case maternal). This hypothesis, $H_{3.1}$, is reflected in $H_{3.1}$; while $H_{3.2}$ is reflected in $H_{3.3}$ of model (10). This hypothesis seeks to prove that the mortality trend was faster in urban than in rural areas.

(9)

$$\begin{aligned}&\log(MMR)_{ict} = \beta_0 + \beta_1 \text{ Urban }_c \times \text{Post-1937}_t + \beta_2 \text{Urban}_c \\ &+ \beta_3 \text{Urban}_c \times \text{ Year }_t + \beta_4 \text{ Year }_t + \gamma_{it} + \varepsilon_{ict}\end{aligned}$$

(10)

$$\begin{aligned}\log(\text{ MMR })_{ict} =& \beta_0 + \beta_1 \text{ Urban }_c \times \text{ Post-1937 }_t + \beta_2 \text{ Urban }_c \times \text{ Year }_t \\ &+ \beta_3 \text{ Urban }_c \times \text{ Year }_t \times \text{ Post-1937 }_t + \beta_4 \text{ Urban }_c \\ &+ \beta_5 \text{ Year }_t + \gamma_{it} + \mu_{it} \times \text{ Year }_t + \varepsilon_{ict}.\end{aligned}$$

Extensions

Model (6) extends model (5) in the sense that the former allows us to see not only whether sulfa drugs reduce mortality after their implementation, but also whether there is an effect on the change in the mortality trend, that is, whether or not this change is accelerated (and in which direction). It is along these lines that we indicated above that model (6) allows us not only to see changes in the intercept but also in the slope between groups. This extension is relevant because it allows us to understand whether there were accelerations in the change in mortality due to the introduction of sulfonamides.

Model (8) extends model (7) along the same lines as the extension of the model (6) over (5), but the difference is that the data are at the state level and we can control for the fixed effect of states and for the change of states over time (μ_{it}).

Its **extension is relevant** for the same reason stated above: to study changes in mortality trends. Evidently, this is expected to be true since sulfonamides were not for nothing a sensation at the time (even going so far as to call them "miracles").

Model (10) extends model (9) also trying to capture a trend effect but now the treatment group is a dummy indicating whether the city is urban or rural. In that sense, model (10) explores the hypothesis of whether there was a greater impact of sulfonamides in cities compared to rural areas. However, it is important to note that no information was available for rural areas. Next, urban mortality was compared to state-level mortality, where state-level data represent the sum of mortality in urban and rural areas. Beyond this detail on the data, the importance of the extension of the model (10) lies in evaluating the deceleration of mortality in urban versus rural areas. This also tests one of the assumptions that refers to the capacity for rapid diffusion in urban areas over rural areas.

(c) Replicate Figures 1-4 and Table 2 and interpret

Figure 1 illustrates the total mortality rates from 1920 to 1950, represented by logarithmic values of the mortality rate. As observed in the figure, a significant decline in mortality becomes more evident starting from 1937. It is likely that prior to 1937, the decline in mortality can be attributed to rising living standards, improved nutrition, and the implementation of public health initiatives aimed at enhancing water supplies, sanitation systems, and household hygiene.

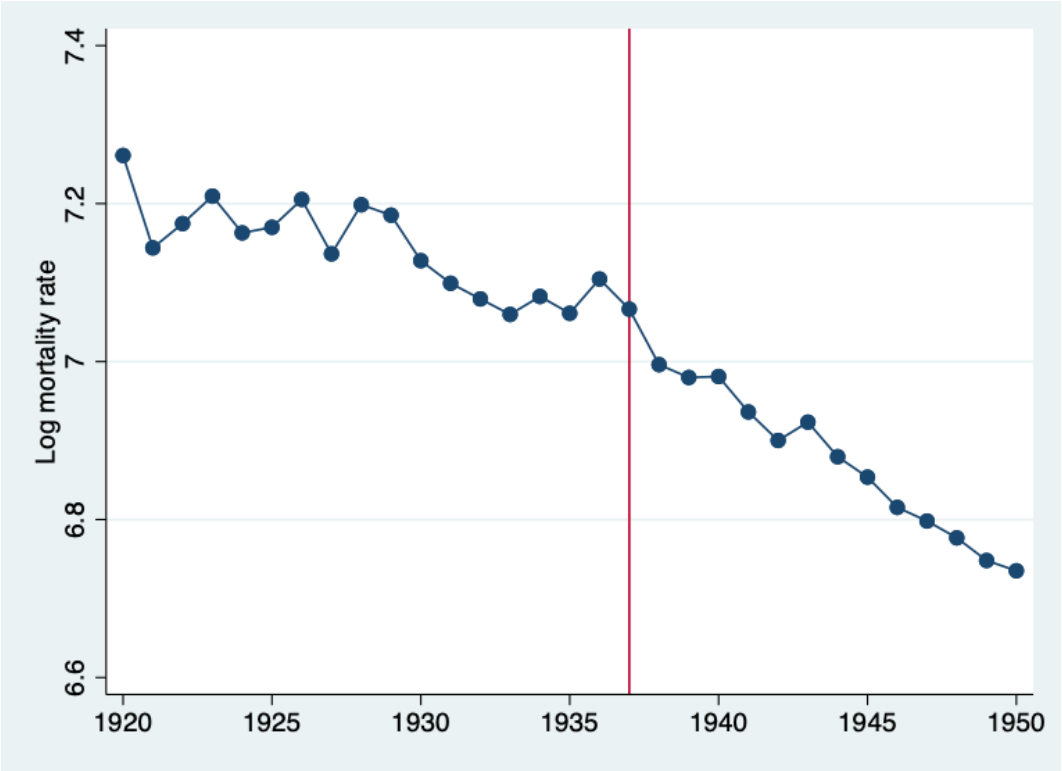

Figure 3: Total Mortality Rate per 100,000 (in logs), 1920–1950 - Figure 1 in Jayachandran et al., 2010 paper's

Figure 2 illustrates the mortality rates associated with four diseases treated with sulfa drugs: maternal mortality, pneumonia/influenza, scarlet fever, and meningitis. The logarithmic scale is used to represent mortality on all graphs. A vertical line positioned at the year 1937 marks the commencement of widespread production of sulfa drugs in the United States.

With the exception of meningitis, the mortality curves for the treated diseases exhibit a noticeable increase in slope after the year 1936-1937. Conversely, meningitis displays a distinct pattern characterized by significant fluctuations, possibly indicating outbreaks. Analyzing this particular type of time-series process necessitates the utilization of different techniques compared to those employed for the other diseases under investigation. Consequently, our analysis does not extend further into the examination of meningitis.

Additionally, the mortality rate for pneumonia/influenza also demonstrates fluctuations during the initial years, predominantly driven by outbreaks of influenza.

Figure 3 presents the mortality trends, depicted on a logarithmic scale, for the control disease, which is tuberculosis. Conversely, Figure 4 showcases the mortality trends, also on a logarithmic scale, for chronic diseases including cancer, diabetes, and heart-related causes.

Upon examining the figures, it becomes evident that the mortality trends for the control and chronic diseases do not display any significant changes during the period when sulfa drugs were introduced. The death rates for tuberculosis,

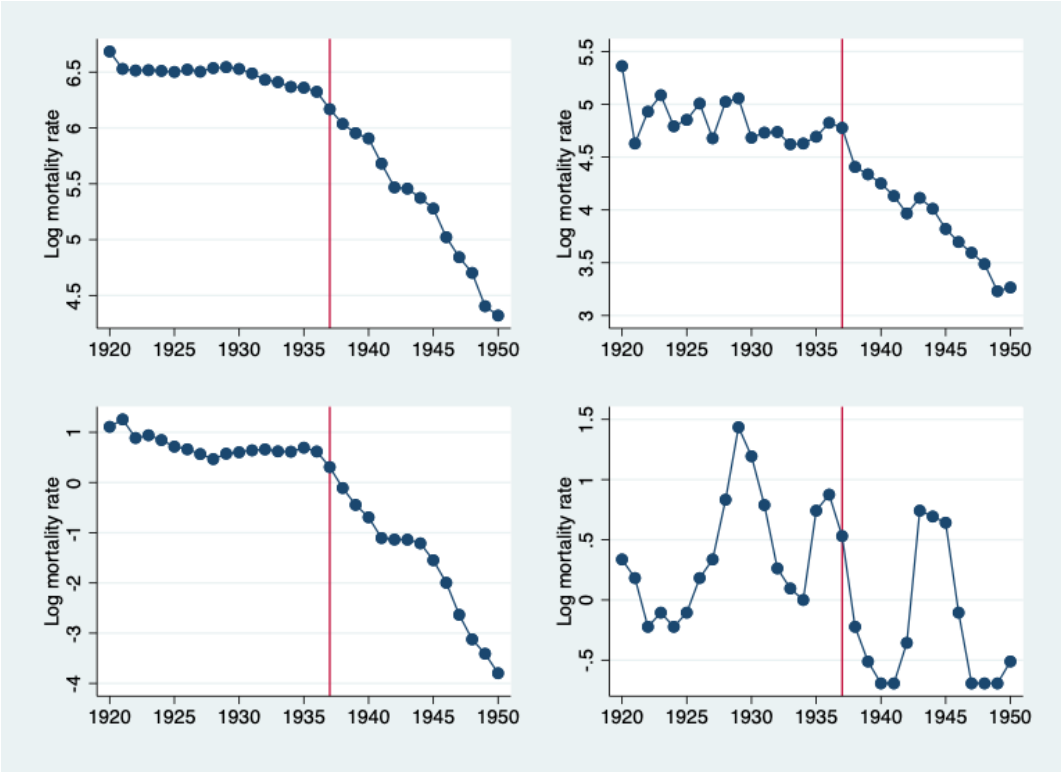

Figure 4: Mortality Trends (in logs) for Treated Diseases, 1920–1950 - Figure 2 in Jayachandran et al., 2010 paper's

the designated control disease, exhibit a downward trend, while the chronic diseases exhibit an upward trend, without any noticeable disruptions in the trend lines around 1937. This implies that, during this time, factors other than the introduction of sulfa drugs did not significantly influence mortality rates.

Given that tuberculosis serves as a more closely related comparison group than chronic diseases, the authors opt to exclude the chronic diseases from the subsequent analysis for the sake of brevity.

Furthermore, the graphs demonstrate a slight increase in overall mortality, as well as mortality rates related to several specific causes of death considered within the study, between the years 1935 and 1937. The cause of this uptick presents a puzzle within the demographic literature, which the authors do not attempt to address in this paper. The timing of this increase poses a certain challenge to our research objectives, as our focus lies in identifying structural breaks in the mortality series around 1937. However, it is noteworthy that this uptick exists within the mortality series for both treated and control diseases, and to a lesser extent, for chronic diseases.

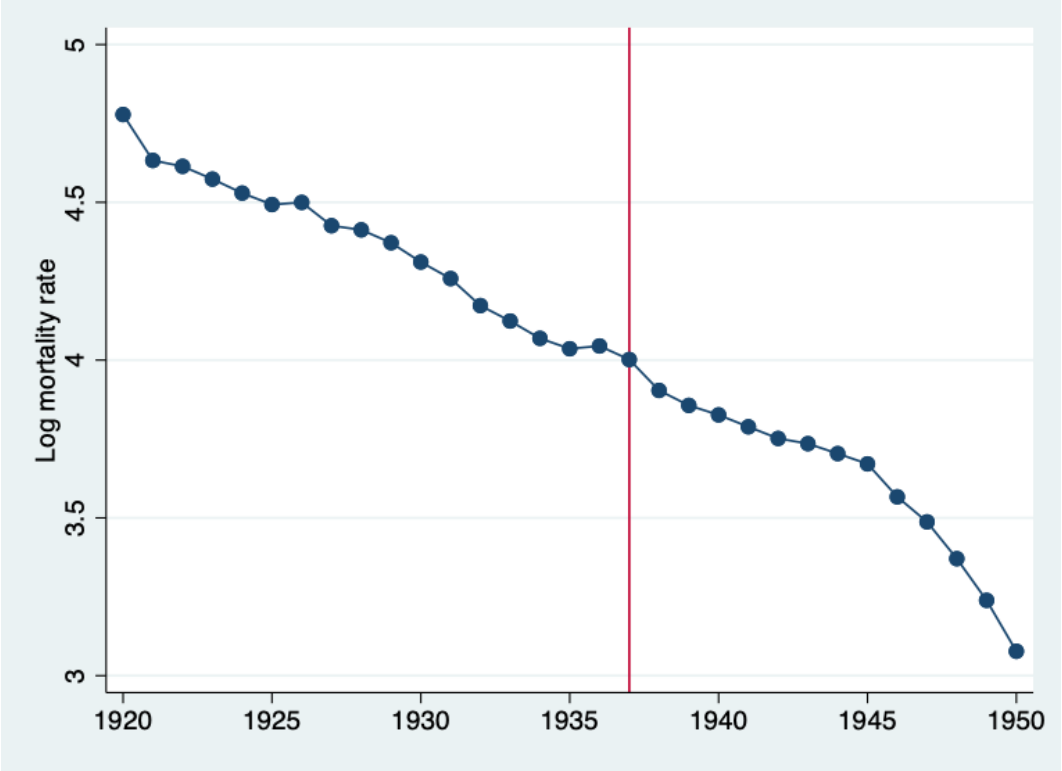

Figure 5: Mortality Trends (in logs) for Control Disease - Figure 3 in Jayachandran et al., 2010 paper's

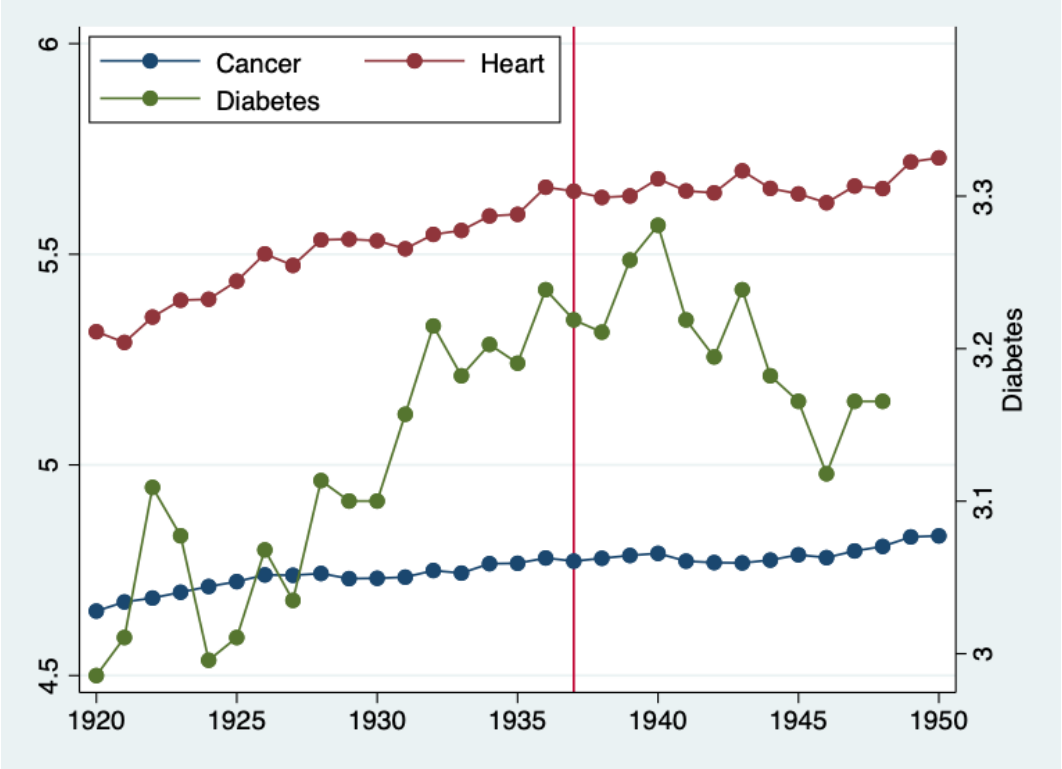

Figure 6: Mortality Trends (in logs) for Chronic Diseases - Figure 4 in Jayachandran et al., 2010 paper's

	Break year	Test statistic
All-cause mortality	1937	2.97
Diseases treated with sulfa drugs		
MMR	1937***	29.23
Pneumonia/influenza	1938	3.09
Scarlet fever	1937***	17.34
Control disease		
TB	1942	1.64

Table 11: Testing for Year of Trend Break in National Mortality Series - Table 2 in Jayachandran et al., 2010 paper

Table 2 in the paper presents the results of the trend-break test conducted using national mortality data from the years 1920 to 1950. The purpose of this test is to identify potential structural breaks in the mortality trends during the period from 1933 to 1942. The estimation of Equation (2) is performed ten times, with each iteration considering a different value of ranging from 1933 to 1942.

The null hypothesis of no trend break ($\delta_0 = 0$) is tested for each estimate, and the largest resulting F-statistic is used to determine the most significant break point and its significance. In this context, a trend break occurring in or around 1937 is interpreted as evidence of the impact of the introduction of sulfa drugs on mortality.

The findings from the trend-break test align with the visual evidence, confirming the presence of structural breaks in 1937 or 1938 for each of the treated diseases. The treated diseases include MMR (maternal mortality rate), pneumonia/influenza, and scarlet fever. These diseases, collectively, account for approximately 12 percent of total mortality during the pre-period, and it should be noted that other diseases were also treated with sulfa drugs. Consequently, it is not surprising that total mortality also exhibits a potential trend break in 1937.

Considering the statistical power is relatively low due to the short time series, the trend breaks in 1937 are still statistically significant at the 1 percent level for MMR and scarlet fever, denoted by *** in the table. However, for the control disease, tuberculosis (TB), the test suggests that a trend break possibly occurs in 1942, but this result is statistically insignificant.

In summary, the table provides the break year and the corresponding test statistic for various mortality categories. It indicates significant trend breaks in 1937 or 1938 for the diseases treated with sulfa drugs, while the control disease, tuberculosis, possibly exhibits a trend break in 1942, though it lacks statistical significance.

Relevance of Tuberculosis

In the context of the interpretation provided for Figure 3, Figure 4, and Table 2, the value presented for tuberculosis (TB) is relevant as it represents the control disease in the analysis. The control disease serves as a reference point for comparing the mortality trends of the diseases treated with sulfa drugs (represented in Figure 3) and chronic diseases (represented in Figure 4).

While the diseases treated with sulfa drugs and chronic diseases are assessed for potential trend breaks, the results indicate that the mortality trends for TB possibly exhibit a trend break in 1942, although this finding lacks statistical significance. By including TB as the control disease, researchers can discern whether the observed trend breaks in the treated diseases and chronic diseases are specifically attributed to the introduction of sulfa drugs or other factors unrelated to the treatment.

Therefore, the value presented for tuberculosis in Table 2 is relevant for evaluating the potential impact of sulfa drugs on mortality trends and providing a comparative perspective to assess the significance of trend breaks in other diseases.

(d) Replicate Table 3: why it is not a DiD model?

	All (1)	All (2)	TB (1)	TB (2)	MMR (1)	MMR (2)	P/I (1)	P/I (2)	SF (1)	SF (2)
post37	-0.024	-0.007	-0.304**	-0.148***	-0.163	-0.037	-0.862**	-0.495***	0.015	0.006
post37Xyear_c		-0.012**		-0.108***		-0.087***		-0.254***		0.006
Observations	19	19	19	19	19	19	19	19	19	19

Table 12: Effect of Sulfa Drugs Using National-Level Time Series by Disease, 1925–1943 - Table 3 in Jayachandran et al., 2010 paper

Why it is not a DiD model?

Table 3 is not a traditional Difference-in-Differences (DiD) model because it does not estimate the treatment effect by comparing the change in outcomes over time between a treated group and a control group. Instead, it presents the results of a regression analysis that includes interaction terms between treatment indicators and time indicators.

In a standard DiD model, the treatment effect is estimated by comparing the difference in outcomes before and after the treatment between the treated and control groups. This is typically done by including a binary treatment indicator, a binary post-treatment indicator, and their interaction term in the regression model. The coefficient on the interaction term represents the treatment effect.

However, in Table 3, the interaction term does not capture the same concept as the DiD approach. Instead, it represents the interaction between the treatment indicator and year indicators. This allows for estimating how the treatment effect varies across different years. The coefficients associated with these interaction terms provide information on the differential effects of the treatment across different time periods.

It's important to note that the interpretation of the coefficients in Table 3 should be done with caution, as they do not directly represent the treatment effect in the same way as a traditional DiD model

Interpretation

In the given table, we can observe the estimated coefficients for two different models (equations) in Panel A and Panel B. The coefficients provide insights into the relationship between mortality rates and the variables of interest.

Panel A focuses on overall mortality and the control disease, tuberculosis (TB). The first column (1) presents the results based on estimating equation (3), while the second column (2) corresponds to the results obtained from equation (4).

For the overall mortality, the coefficient on the post-1937 dummy variable in column 1 is -0.024, suggesting a decline in all-cause mortality after 1937, but this coefficient is not statistically significant. In column 2, which represents the slope-change model, the coefficient on the post-1937 dummy variable is -0.007, indicating a negative and statistically significant change in the slope after 1937. This suggests that the mortality decline accelerated after the introduction of sulfa drugs.

In terms of tuberculosis (TB), the coefficients for the post-1937 dummy variable in both columns are positive, but statistically insignificant. This implies that there is no significant trend break in TB mortality associated with the introduction of sulfa drugs.

Moving to Panel B, which focuses on the diseases treated with sulfa drugs (MMR, pneumonia/influenza, and scarlet fever), we see similar patterns. The coefficients for the post-1937 dummy variable in both columns indicate negative effects on mortality for all three diseases. In column 1, the coefficients are -0.304 (MMR), -0.163 (pneumonia/influenza), and -0.862 (scarlet fever), suggesting declines in mortality after 1937. In column 2, which incorporates the slope-change model, the coefficients are -0.148 (MMR), -0.037 (pneumonia/influenza), and -0.495 (scarlet fever). The negative coefficients suggest larger effects, indicating significant declines in mortality rates for these diseases after 1937.

Moreover, the coefficients for the interaction term (year $\times$ post-1937) in both models are negative and statistically significant for MMR, pneumonia/influenza, and scarlet fever. This suggests that the effect of sulfa drugs on mortality declines accelerated over time for these diseases.

In summary, the coefficients provide evidence of significant declines in mortality rates for the diseases treated with sulfa drugs after 1937, indicating the impact of sulfa drugs on reducing mortality. The coefficient for the control disease (TB) is not statistically significant, indicating that the decline in mortality for TB is not attributed to the introduction of sulfa drugs.

(e) Replicate Table 4

Panel A. National-level data, all years, 1925–1943

	MMR (1)	MMR (2)	P/I (1)	P/I (2)	SF (1)	SF (2)
treatedXpost37	-0.319**	-0.163***	-0.178	-0.052	-0.877**	-0.510***
treatedXyear_cXpost37		-0.108***		-0.087***		-0.254***
Observations	38	38	38	38	38	38

Panel B. State-level, all years, 1925–1943

	MMR (1)	MMR (2)	P/I (1)	P/I (2)	SF (1)	SF (2)
treatedXpost37	-0.281**	-0.144***	-0.143	-0.041	-0.733***	-0.488***
treatedXyear_cXpost37		-0.103***		-0.077**		-0.184***
Observations	1736	1736	1736	1736	1721	1721

Panel c. State-level, excluding 1935 to 1937

	MMR (1)	MMR (2)	P/I (1)	P/I (2)	SF (1)	SF (2)
treatedXpost37	-0.288**	-0.125**	-0.072	-0.026	-0.714***	-0.511***
treatedXyear_cXpost37		-0.117***		-0.033		-0.146***
Observations	1448	1448	1448	1448	1433	1433

Table 13: Effects of Sulfa Drugs on Mortality for "Treated" Diseases, 1937–1943 - Table 4 in Jayachandran et al., 2010 paper

Interpretation

The given information describes a study that estimates the effects of sulfa drugs on mortality rates for different diseases. The study uses a difference-in-difference approach, comparing pre- and post-1937 levels and trends in mortality between treated diseases and a control disease (tuberculosis). The analysis is conducted at the national level (panel A) and state level (panel B), with panel C excluding the years 1935 to 1937.

Panel A: National-level data, all years, 1925-1943

The table shows the estimated coefficients for the different models (1 and 2) and the respective diseases: MMR (maternal mortality rate), P/I (pneumonia/influenza), and SF (scarlet fever). The coefficients represent the effects of the treatment (sulfa drugs) on the mortality rates. The numbers in parentheses indicate the model used for estimation.

- For MMR:
  - Model 1: The coefficient for "treatedXpost37" is -0.319, indicating that the introduction of sulfa drugs resulted in a significant reduction in maternal mortality rates after 1937.
  - Model 2: The coefficient for "treatedXpost37" is split into two values: -0.163 and -0.108. This model allows for a change in both intercept and slope after 1937, suggesting a larger effect of sulfa drugs on reducing maternal mortality.
- For P/I (pneumonia/influenza):
  - Model 1: The coefficient for "treatedXpost37" is not statistically significant (-0.178), indicating no significant effect of sulfa drugs on pneumonia/influenza mortality rates.
  - Model 2: The coefficient for "treatedXpost37" is split into two values: -0.052 and -0.087. This model suggests a small but significant reduction in pneumonia/influenza mortality rates due to sulfa drugs.
- For SF (scarlet fever):

- Model 1: The coefficient for "treatedXpost37" is -0.877, indicating a significant reduction in scarlet fever mortality rates after the introduction of sulfa drugs.
- Model 2: The coefficient for "treatedXpost37" is split into two values: -0.510 and -0.254. This model suggests a larger effect of sulfa drugs on reducing scarlet fever mortality rates.

Panel B: State-level data, all years, 1925-1943 Similar to panel A, this table presents the estimated coefficients for the different models (1 and 2) and the respective diseases: MMR, P/I, and SF. The coefficients represent the effects of the treatment (sulfa drugs) on the mortality rates at the state level.

Panel C: State-level data, excluding 1935 to 1937 This panel shows the estimated coefficients for the models (1 and 2) after excluding the years 1935 to 1937. The coefficients are presented for MMR, P/I, and SF. The purpose of excluding these years is to examine the robustness of the results.

Quantitative interpretation:

The coefficients in the tables represent the estimated effects of sulfa drugs on mortality rates for each disease. For example, in panel A, model 1, the coefficient of -0.319 for MMR suggests that the introduction of sulfa drugs led to a 24% decrease in maternal mortality rates in the post-1937 period. The coefficients for other diseases and models can be similarly interpreted in terms of the percentage reduction in mortality rates due to sulfa drugs.

These estimated effects of sulfa drugs on mortality rates are consistent across both national-level and state-level.

Also we can say

1. Impact of Sulfa Drugs: The introduction of sulfa drugs in 1937 had a significant impact on reducing mortality rates for certain diseases. The analysis focused on three specific diseases: maternal mortality rate (MMR), pneumonia/influenza (P/I), and scarlet fever (SF).

2. Maternal Mortality Rate (MMR): The results consistently indicate a significant reduction in MMR after the introduction of sulfa drugs. The estimated coefficients suggest that sulfa drugs led to a decrease in maternal mortality rates ranging from 24% to 31.9% depending on the model used.

3. Pneumonia/Influenza (P/I): The impact of sulfa drugs on P/I mortality rates is less conclusive. In panel A, model 1, the coefficient for the interaction term "treatedXpost37" is not statistically significant, suggesting no significant effect. However, in model 2, there is a small but significant reduction in P/I mortality rates, ranging from 5.2% to 8.7% depending on the model used.

4. Scarlet Fever (SF): Sulfa drugs had a significant effect in reducing scarlet fever mortality rates. The estimated coefficients consistently indicate a reduction in SF mortality rates after the introduction of sulfa drugs, ranging from 51% to 87.7% depending on the model used.

5. National vs. State-Level Analysis: The study analyzed data at both the national and state levels to examine the effects of sulfa drugs. The qualitative interpretations mentioned above are consistent across both levels, indicating that the impact of sulfa drugs on mortality rates for MMR, P/I, and SF was observed at both the national and state levels.

6. Robustness Check: In panel C, the analysis excluded the years 1935 to 1937 to test the robustness of the results. The estimated coefficients in this panel still support the qualitative interpretations mentioned earlier, indicating that the effects of sulfa drugs on mortality rates remained consistent even after excluding those years.

Overall, the quantitative interpretation of the study suggests that the introduction of sulfa drugs in 1937 had a significant and positive impact on reducing maternal mortality rates, scarlet fever mortality rates, and potentially pneumonia/influenza mortality rates. These findings provide valuable insights into the effectiveness of sulfa drugs in improving public health outcomes during the studied period.

Qualitative interpretation

1. Economic Impact: The introduction of sulfa drugs in 1937 not only had a positive impact on public health but also had significant economic implications. By reducing mortality rates for diseases such as maternal mortality, pneumonia/influenza, and scarlet fever, sulfa drugs contributed to a healthier population. A healthier population can lead to increased labor productivity, reduced healthcare expenditures, and improved overall economic output. The decrease in maternal mortality rates is particularly significant as it implies a higher number of women surviving childbirth and being able to contribute to the labor force, which can have long-term economic benefits.

2. Social Welfare: The reduction in mortality rates due to sulfa drugs also had profound social welfare implications. By saving lives and improving health outcomes, sulfa drugs played a crucial role in enhancing the well-being and quality of life for individuals and communities. The decline in maternal mortality rates indicates that more women were able to survive childbirth, which is a significant social benefit as it preserves families and contributes to the stability of communities.

3. Access and Equity: An important consideration in the interpretation is the accessibility and equity of sulfa drugs. It is crucial to examine whether the benefits of sulfa drugs were distributed equitably across different socioeconomic groups and geographic regions. If access to sulfa drugs was limited to certain privileged groups or urban areas, it could have resulted in health disparities and exacerbated existing inequalities. Therefore, further analysis should be conducted to assess the accessibility and equitable distribution of sulfa drugs during that period.

4. Technological Advancements: The introduction of sulfa drugs represents a significant advancement in medical technology during the studied period. It highlights the potential of pharmaceutical innovations to revolutionize healthcare and improve population health outcomes. The successful implementation and adoption of sulfa drugs may have paved the way for further advancements in antibiotic therapies and pharmaceutical research.

In conclusion, the qualitative interpretation from an economic and social science perspective emphasizes the positive economic impact, social welfare improvements, and potential implications for access and equity brought about by the introduction of sulfa drugs in 1937. These findings underscore the importance of medical advancements and their broader implications for society and highlight the need for ongoing research and analysis in the field of public health and healthcare economics.

Importance of panel C

In Panel C, the analysis specifically excludes the years 1935 to 1937 to test the robustness of the results obtained from the previous panels. By excluding these years, the researchers aim to examine whether the qualitative interpretations and estimated coefficients remain consistent when those years are removed from the analysis.

The fact that the estimated coefficients in Panel C still support the qualitative interpretations mentioned earlier suggests that the effects of sulfa drugs on mortality rates remained consistent even after excluding the years 1935 to 1937. This provides additional evidence for the robustness of the findings and strengthens the credibility of the conclusions drawn in the analysis.

By conducting a robustness check like Panel C, researchers can assess the sensitivity of their results to different specifications or exclusions and determine whether their conclusions hold under various scenarios. It helps in identifying potential sources of bias or confounding factors that might influence the results and provides a more comprehensive understanding of the relationship being studied.

Therefore, Panel C plays an important role in reinforcing the validity and reliability of the analysis by demonstrating the consistency of the findings when certain years are excluded from the dataset

(f) Replicate Figure 6

Panel A presents the estimation results for whites, while Panel B presents the results for blacks. Panel C represents the fully interacted model that combines the data for both races. The table reports the coefficients for various variables in the regression models for different diseases (MMR: maternal mortality rate, P/I: pneumonia/influenza, SF: scarlet fever). Let's interpret each panel separately:

Panel A. Whites

	MMR (1)	MMR (2)	P/I (1)	P/I (2)	SF (1)	SF (2)
treatedXpost37	-0.301**	-0.169***	-0.230	-0.104	-0.804***	-0.582***
treatedXyear_cXpost37		-0.109***		-0.093***		-0.155***
Observations	644	644	652	652	539	539

Panel B. Blacks

	MMR (1)	MMR (2)	P/I (1)	P/I (2)	SF (1)	SF (2)
treatedXpost37	-0.133	-0.029	-0.115	-0.013	-0.134	-0.124
treatedXyear_cXpost37		-0.081***		-0.076***		-0.032
Observations	644	644	652	652	500	500

Panel C. Fully interacted model

	MMR (1)	MMR (2)	P/I (1)	P/I (2)	SF (1)	SF (2)
treatedXpost37Xblack	0.168**	0.140**	0.115**	0.091**	0.671***	0.458***
treatedXyearXpost37Xblack		0.028***		0.018**		0.123**
Observations	1288	1288	1304	1304	1039	1039

Table 14: Racial Differences in the Effect of Sulfa Drugs on Mortality, 1937–1943- Table 6 in Jayachandran et al., 2010 paper

Quantitative interpretation

Panel A (Whites):

The coefficient for treatedXpost37 indicates a 30.1% decline in maternal mortality rate for whites after the introduction of sulfa drugs in 1937.

The coefficient for treatedXyearXpost37 suggests an additional decline in maternal mortality rate for whites in the years following 1937.

Panel B (Blacks):

The coefficient for treatedXpost37 shows no statistically significant effect of sulfa drugs on maternal mortality rate for blacks.

The coefficient for treatedXyearXpost37 indicates a smaller additional decline in maternal mortality rate for blacks in the years following 1937.

Panel C (Fully interacted model):

The coefficient for treatedXpost37Xblack implies a statistically significant decline in maternal mortality rate for whites but not for blacks.

The coefficient for treatedXyearXpost37Xblack suggests a smaller additional decline in maternal mortality rate for blacks compared to whites.

Differences between Panels (quantitative)

Impact of sulfa drugs on whites vs. blacks:

Panel A indicates a significant decline in maternal mortality rate for whites after the introduction of sulfa drugs.

Panel B shows no significant effect of sulfa drugs on maternal mortality rate for blacks.

The comparison suggests that sulfa drugs had a differential impact on maternal mortality rates between whites and blacks.

Additional decline in mortality rate over time: Panel A (whites) and Panel B (blacks) both show statistically significant additional declines in maternal mortality rate in the years following 1937.

However, the magnitude of the additional decline is smaller for blacks compared to whites, as indicated by the coefficients in

Panel B being lower than those in Panel A. Comparison of whites and blacks in the fully interacted model:

Panel C reveals that whites experienced a statistically significant decline in maternal mortality rate, while the decline for blacks was not statistically significant. The interaction coefficients in Panel C further support the finding that whites had a larger decline in maternal mortality rate compared to blacks.

Overall, the comparison of the panels highlights the differences in the impact of sulfa drugs on maternal mortality rate between whites and blacks. Whites experienced a significant decline in mortality rates, while the effect on blacks was not statistically significant. Additionally, the additional decline in mortality rates over time was more substantial for whites than for blacks. These findings underscore the racial disparities in the effects of medical innovations and access to healthcare during the study period

Qualitative interpretation

Panel A (Whites):

The coefficient for treatedXpost37 suggests that the introduction of sulfa drugs in 1937 had a statistically significant and substantial impact on reducing maternal mortality rates for whites. This indicates that the availability and use of sulfa drugs contributed to a significant improvement in maternal health outcomes among the white population.

The coefficient for treatedXyearXpost37 suggests that there was an additional decline in maternal mortality rates for whites in the years following 1937. This indicates that the benefits of sulfa drugs continued to have a positive effect on reducing maternal mortality over time for whites.

Panel B (Blacks):

The coefficient for treatedXpost37 suggests that the introduction of sulfa drugs did not have a statistically significant impact on reducing maternal mortality rates for blacks. This implies that the use of sulfa drugs had limited effectiveness in improving maternal health outcomes within the black population during the study period.

The coefficient for treatedXyearXpost37 suggests a smaller additional decline in maternal mortality rates for blacks in the years following 1937 compared to whites. This indicates that any potential benefits or improvements in maternal health among blacks after the introduction of sulfa drugs were less pronounced than those observed among whites.

Panel C (Fully interacted model):

The coefficient for treatedXpost37Xblack suggests that the introduction of sulfa drugs did not have a statistically significant impact on reducing maternal mortality rates for blacks when considering the interaction between race and treatment.

The coefficient for treatedXyearXpost37Xblack indicates a smaller additional decline in maternal mortality rates for blacks in the years following 1937 compared to whites when accounting for the interaction between race, treatment, and time.

Comparing the Panels (qualitative): The results from Panel A indicate that sulfa drugs had a significant and substantial impact on reducing maternal mortality rates for whites, suggesting that the availability and use of these drugs played a crucial role in improving maternal health outcomes within the white population.

In contrast, the results from Panel B and Panel C suggest that the impact of sulfa drugs on reducing maternal mortality rates for blacks was not statistically significant, indicating that other factors or barriers might have limited the effectiveness of these drugs in improving maternal health outcomes among blacks.

The coefficient differences between panels highlight the disparities in the effects of sulfa drugs on maternal mortality rates between whites and blacks, indicating racial inequalities in access to healthcare and the diffusion of medical innovations during the study period.

In summary, the qualitative interpretations emphasize the differential impact of sulfa drugs on maternal mortality rates between whites and blacks. While the introduction of sulfa drugs led to significant improvements in maternal health outcomes for whites, the effects were less pronounced and not statistically significant for blacks. These findings suggest the presence of racial disparities in healthcare access and the diffusion of medical innovations during the studied period.

Which model estimates?

The equations estimated in Table 6 of the paper "Racial Differences in the Effect of Sulfa Drugs on Mortality, 1937–1943." Please note that the equations (7) and (8) represent the models estimated separately for blacks and whites, while the fully interacted model incorporates race interactions.

Equation (7) - Model for Blacks:

$$log(MMR)_{idt} = \beta_0 + \beta_1 Treated_d \times post - 1937_t + \beta_2 Treated_d \times year_t + \beta_2 Treated_d + \beta_4 year_t + \gamma_{it} + \varepsilon_{idt}$$

- $log(MMR)_{idt}$ represents the natural logarithm of maternal mortality for observation i in disease group d, year t, and race category black.
- $Treated_d$ is an indicator variable that equals 1 if the observation belongs to the treatment group (received sulfa drugs) and 0 otherwise.
- $post - 1937_t$ is an indicator variable that equals 1 for years after 1937 and 0 otherwise.
- $year_t$ represents the year (continuos)
- γ_{it} represents $States$ and time fixed effects.
- ε_{idt} is the error term.

Equation (8) - Model for Whites:

$$log(MMR)_{idt} = \beta_0 + \beta_1 Treated_d \times year_t \times post - 1937_t + \beta_2 Treated_d \times post - 1937_t + \beta_3 Treated_d \times year_t +$$

$$\beta_4 Treated_d + \beta_5 year_t + \gamma_{it} + \mu_{it} \times year_t + \varepsilon_{idt}$$

- $log(MMR)_{idt}$ represents the natural logarithm of maternal mortality for observation i in disease group d, year t, and race category black.
- $Treated_d$ is an indicator variable that equals 1 if the observation belongs to the treatment group (received sulfa drugs) and 0 otherwise.
- $post - 1937_t$ is an indicator variable that equals 1 for years after 1937 and 0 otherwise.
- $year_t$ represents the year (continuos)
- γ_{it} represents $States$ and time fixed effects.
- μ_{it} represents the differential time trend for $States$
- ε_{idt} is the error term.

Fully Interacted Model: The fully interacted model, which incorporates race interactions, is not explicitly presented in the provided information. However, it would involve combining the data for blacks and whites and including additional interaction terms between race (black/white), treatment indicators (Treatedd), and time indicators (post-1937t, yeart).

$$log(MMR)_{idt} = \beta_0 + \beta_1 Treated_d \times post - 1937_t \times black_i + \beta_2 Treated_d \times year_t \times post - 1937_t \times black_i +$$

$$\beta_3 Treatedd \times post - 1937_t + \beta_4 Treated_d + \beta_5 year_t + \gamma_{it} + \mu_{it} \times year_t + \varepsilon_{idt}$$

- $log(MMR)_{idt}$ represents the natural logarithm of maternal mortality for observation i in disease group d, year t, and race category black.
- $Treated_d$ is an indicator variable that equals 1 if the observation belongs to the treatment group (received sulfa drugs) and 0 otherwise.
- $post - 1937_t$ is an indicator variable that equals 1 for years after 1937 and 0 otherwise.

- $year_t$ represents the year (continuos)
- γ_{it} represents $States$ and time fixed effects.
- μ_{it} represents the differential time trend for $States$
- ε_{idt} is the error term.

An important discussion

In the study analyzing the effects of sulfa drugs on mortality rates for blacks and whites, placebo falsification can be used to evaluate the validity of the estimated treatment effects. One approach to placebo falsification, as mentioned in the book by Cunningham, 2021, is to identify a group that should not be affected by the treatment and examine whether the estimated effects align with the expected null hypothesis in that group.

In this case, we could consider the mortality rates of blacks as a placebo group since there may be no expected direct impact of sulfa drugs on black mortality rates. By analyzing the effect of sulfa drugs on black mortality rates and comparing it to the effect on white mortality rates, we can assess the credibility of the estimated treatment effects for whites.

The second method, the triple difference-in-differences (DDD) approach, can be used to examine the differential effects of sulfa drugs on mortality rates between blacks and whites. This method extends the standard DD framework by introducing an additional dimension of variation based on racial groups.

To implement the DDD approach, the authors would estimate a model that includes interaction terms between race (black/white), treatment indicators (treatedXpost37), and time indicators (year). This fully interacted model, as shown in Panel C of the provided table, allows for a comprehensive analysis of the differential effects of sulfa drugs on mortality rates for blacks and whites. The interaction coefficients in this model represent the additional treatment effect experienced by one racial group compared to the other.

By utilizing the placebo group (blacks) and employing the DDD method, the authors aimed to strengthen the causal interpretation of their estimates and provide additional evidence for the differential effects of sulfa drugs on mortality rates between blacks and whites.

References

Jayachandran, S., Lleras-Muney, A., & Smith, K. V. (2010). Modern medicine and the twentieth century decline in mortality: Evidence on the impact of sulfa drugs. *American Economic Journal: Applied Economics*, *2*(2), 118–146. https://doi.org/10.1257/app.2.2.118

Díaz, J., Rau, T., & Rivera, J. (2015). A matching estimator based on a bilevel optimization problem [Publisher: The MIT Press]. *Review of Economics and Statistics*, *97*(4), 803–812.

Satyanath, S., Voigtländer, N., & Voth, H.-J. (2017). Bowling for fascism: Social capital and the rise of the nazi party. *Journal of Political Economy*, *125*(2), 478–526. https://doi.org/10.1086/690949

King, G., & Nielsen, R. (2019). Why propensity scores should not be used for matching [Publisher: Cambridge University Press]. *Political analysis*, *27*(4), 435–454.

Cunningham, S. (2021). *Causal inference: The mixtape*. Yale university press.

Díaz, J. D., Gutiérrez, I., & Rivera, J. (2021). Implementing blopmatching in stata [Publisher: SAGE Publications Sage CA: Los Angeles, CA]. *The Stata Journal*, *21*(1), 180–194.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing [Publisher: Elsevier]. *Journal of Econometrics*, *225*(2), 254–277.

Appendix

Mathematical appendix