

# Estate origins of democracy? Social persistence during communism in Russia

Final Paper for Social Stratification I

Stepan Polikanov

## Table of contents

<b>Introduction</b>	<b>2</b>
<b>Estate origins of democracy in Russia?</b>	<b>3</b>
<b>Mechanisms of social persistence</b>	<b>4</b>
Persistence of education . . . . .	5
Occupational choices . . . . .	7
Urbanization . . . . .	8
Great leveller? Repressions in Soviet Union . . . . .	10
<b>Empirical analysis</b>	<b>12</b>
Data . . . . .	12
Replication . . . . .	12
Robustness . . . . .	13
Statistical issues and alternative specifications . . . . .	13
Other elections . . . . .	14
<b>Results</b>	<b>17</b>
Education . . . . .	17
Urbanization . . . . .	18
Repressions . . . . .	20
Comparing mechanisms of social persistence . . . . .	21
Sequential-G estimation results . . . . .	21
Factor Analysis . . . . .	22
<b>Conclusion</b>	<b>25</b>

<b>Appendix</b>	<b>29</b>
Random effects . . . . .	29

## Introduction

Persistence of social status is a classic issue in sociology and social stratification research. Many studies link past social structures to present social inequalities (Corak, 2013; Torche, 2015). Intergenerational mobility as the main source of persistent social structures has been linked to income, wealth, class, educational and occupational persistence (Carina Mood & Mood, 2017; Erzsébet Bukodi & John H. Goldthorpe, 2013; Jani Erola et al., 2016; John Jerrim & Lindsey Macmillan, 2015; Martin Hällsten & Fabian T. Pfeffer, 2017; Sandra E. Black & Paul J. Devereux, 2010). Low intergenerational mobility implies that the “starting point” exerts a strong influence on the social stratification throughout multiple generations.

Alesina et al., 2020 show that income and education inequalities survive the Chinese Communist Revolution and Cultural Revolution, reproducing in grandchildren of the pre-revolution elites even in the context of extensive expropriation and other “homogenization” policies. Authors conclude, that since these policies were effective in equalizing the previous elites with the lower classes, the observed persistence must go through family, community, and the values transmitted within them.

In contrast, the Russian communist experience allowed for reproduction of societal legacies at levels other than family. From this point of view, Russia presents a rich case for the study of social persistence and its many possible origins (Torche, 2015). The existence of distinct estates at the turn of the century and the (comparative) limited industrialization of the empire meant that occupations, income and education were dynamically (see Wirtschafter, 1997 for a perspective on estate definitions and their subversion) dependent on the basis of the estate. Bolshevik reforms and policies aimed at “leveling” the societal structure can be regarded as inconsistent at best. As Lankina, 2021; Lankina and Libman, 2021b argue, after World War I, the Russian Civil War and the Revolution itself, Bolsheviks had to co-opt institutions and classes they would rather see rebuilt. The educational reform meant to open up the educational system to the masses introduced bottlenecks at the university level, and early decisions of attending a secondary school or a specialized vocational institution dictated the probability to attend higher education, ultimately leading to a system where social origin conditioned educational attainment (Gerber & Hout, 1995).

In the case covered in this paper, the “starting point” is the pre-revolutionary social structure, and the survivability of it is questioned in the presence of expanded equality of opportunity, policies and repressions aimed at “leveling” the social structure and over the very long run.<sup>1</sup> I exploit that variation in transition to Soviet rule to show that educational and

---

<sup>1</sup>Here I refer to expansion in equality of opportunity as the changes in *de jure* requirements for educational enrollment. Specifically, until 1905 Revolution, there was no compulsory elementary schooling and until 1918 there was no compulsory secondary education and the system was stratified by gender and class.

urban legacies of pre-revolutionary Russia persist over the very long run. I engage with the arguments of Tomila Lankina and Alexander Libman, who argue for the persistence of pre-revolutionary social structures in Soviet- and post-Soviet times (Lankina, 2021; Lankina & Libman, 2021b). Their research provides historical, theoretical and empirical base for the argument that the social structure of the Russian Empire, in particular that of the “middle (urban) class”, *meshchane*, survived the Russian Revolution and the Soviet period, and had an impact on the early democratic processes in Russian Federation.

This work entails multiple goals. My primary objective is to formulate a theoretical framework and a set of explanations pertaining to the persistence of social structures in Russia. Specifically I explore educational persistence, touch upon occupational persistence and work with Russia-specific urbanization process and Stalin’s repression legacies. I also replicate the empirical analyses of Lankina and Libman, 2021b and extend it with a broader scope, data and methodological perspectives.

I start with a brief overview of the theoretical arguments put forward by Lankina and Libman as well as alternative theories of social stratification, democracy and the effect of the Soviet rule in Russia. I then proceed to formulate my own hypotheses regarding the topic. My empirical part starts with a replication of the original study by Lankina and Libman, and subjects it to a rigorous testing under different conditions. I then focus on a narrower issue of social transmission in Russia without necessarily connecting it to democratic outcomes. The last section concludes.

## Estate origins of democracy in Russia?

Lankina and Libman, 2021b view democracy in Russia through a lens of the widely established hypothesis of middle class importance for the process. The conventional argument establishes that the middle classes, or the bourgeoisie, through “market-supporting values, entrepreneurial ethos, and demands to protect property rights” (Lankina & Libman, 2021b), education and associated tolerance and respect for autonomy or activism (Chen & Suen, 2017) are the central proponents of democracy. While the hypothesis is well-established (Acemoglu & Robinson, 2006; Ansell & Samuels, 2014), a variety of studies show that the relationship is far from universal in different contexts (Lertchoosakul, 2021; Resnick, 2015; Schotte, 2021). Of particular note is the study by Rosenfeld, 2021 that shows that across post-Soviet countries, state dependency of the middle class diminishes support for democracy.

Testing the hypothesis in Russia, Lankina and Libman, 2021b argue that the modern levels of democracy, measured by electoral competitiveness on the municipal level<sup>2</sup>, are positively associated with the incidence of pre-revolutionary bourgeoisie, *meshchane*<sup>3</sup>. The authors argue that the social structure survived the Revolution and the Soviet period at large due to

---

<sup>2</sup>The municipal level in Russia refers to districts or *rayons*, the lowest administrative unit.

<sup>3</sup>From Polish *mieszczanin* - burgher or townsman.

co-optation of the professions needed for socialist state-building, adoption of pre-revolutionary infrastructure and, ultimately, intergenerational transmission of values and norms associated with townsmen through educational, occupational and family channels.

One can immediately see the ambition of such a continuity. Authors do not only aim to explain the (rather limited) democratic experience of Russia but also the social structure persistence over the very long run. The argument, expanded upon in Lankina's book, does not merely address persistent social stratification as a "black box", with values and capital transmitted through families in a closed format. Instead, the cognitions, associated with imperial bourgeoisie, are seen as manifesting in organizational forms of the new regime. Specifically, the author argues that this class would embrace "possibilities for careers in sites that permit autonomous thought, sensibility, and action even under the most inquisitional and totalizing regimes" (Lankina, 2021). The mechanisms of transmission cover the both the "black box" of familial, cultural and communal transmission (*social closure* in author's terms) and explicit *educational* and *professional* incorporation.

The study lies at the intersection of sociology, political science and economic history. It therefore makes sense that it may be vulnerable to critiques inherent to all of these fields. For example one might take issue with the definition of the middle class, the operationalization of democratic competitiveness, or the causal chain needed to establish the relationship. I focus on the last one as the most important for the validity of the argument. One classic issue of economic history is "compression of time" or "compression of history" where by comparing variables between two points in time one might discard causal chain links in-between, which can lead to spurious correlations through unspecified mechanisms and causal oversimplification (Austin, 2008). In other words, before we can talk about pre-revolutionary social structures affecting modern democracy, we need to establish that the social structures are indeed persistent.

Therefore, to address this concern, I focus on observable aspects of those mechanisms. The logic is simple - if the original social structures have no causal effect on those, then we cannot reasonably sustain the authors' argument about the democratic outcomes of this persistence. I also bring up causal chain links that were left without proper attention - namely I look at Stalin-era repressions, urbanization trends and discuss changes in occupational structures due to industrialization, planned economy in general and transition to market economy. This is a wide project: this paper presents a "first look" into establishing long-run persistence of social structures between two "revolutions" <sup>4</sup>.

## Mechanisms of social persistence

I first turn my attention to the mechanisms that could explain the persistence of social structures in Russia. I focus on the aforementioned *education* and *occupational* transmission.

---

<sup>4</sup>The Russian Revolution of 1917 and the collapse of the Soviet Union in 1991.

## Persistence of education

The most straightforward indicator of social class persistence is in education. Classic works have long noted that a large share of educational and even professional success of individuals can be explained by their parents' education (Becker & Tomes, 1979; Blau & Duncan, 1978; Bourdieu & Passeron, 2000). In particular, Liu, 2018 finds that "up to a half of the association between parents' polygenic scores and children's education is mediated by parents' education". Other research on the topic supports the idea: mother's polygenetic scores affect their children attainment through environmental pathways in Belsky et al., 2018 and Jani Erola et al., 2016.

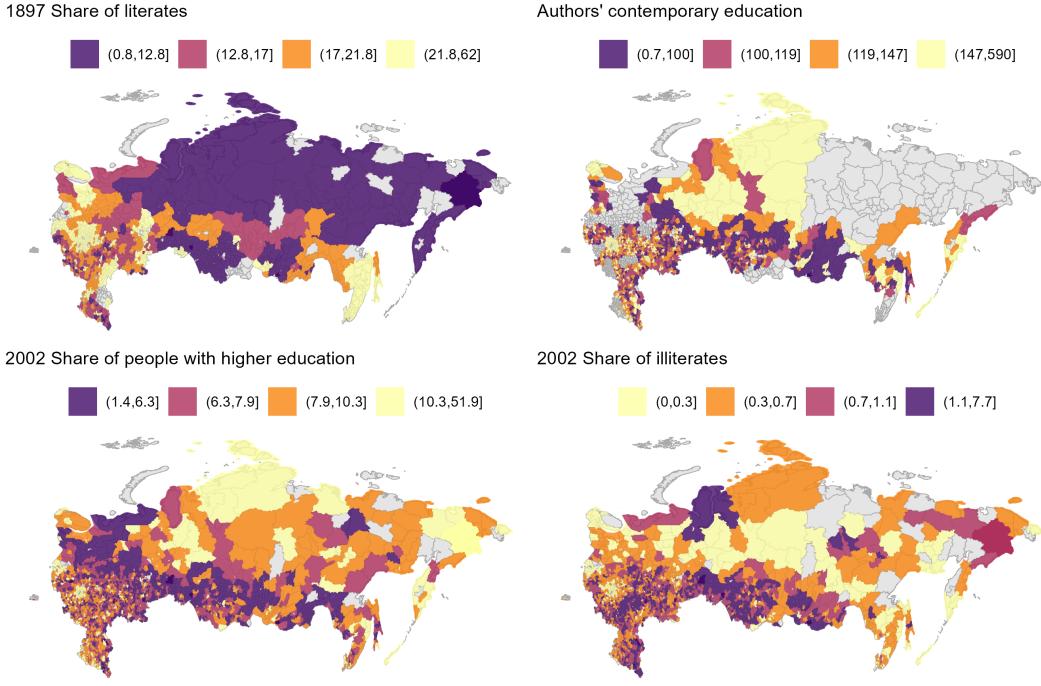
The argument in Lankina, 2021 is that the *meshchane* class is strongly associated with investment in children education, underpinned by the value structure. Therefore, one would expect meshchane's descendants to also value education more and invest in it, reproducing the pre-revolutionary pattern of social structure.

However, how do we know that persistence of education is a unique result of the class persistence of meshchane descendants? The initial condition of different places could be an alternative mechanism, linking the pre-revolutionary literacy levels to modern educational measures. Literacy rates and shares of the meshchane class are positively associated, yet it is not the case that the latter fully explains the former - and we should not assume a single effect captured by those two variables.

The idea to identify the effect is one of a placebo test - if the share of meshchane has an effect on education, it should be on the upper levels of education, particularly tertiary (higher) education. On the other hand, we can hypothesize that there is no inverse relationship - that is, of meshchane shares on the educational underperformers. This leads me to formulate *Hypothesis I*.

$H_1$ : The share of the *meshchane* class in 1897 is positively associated with the modern share of people with higher education but not with the share of illiterate people, once the effects of pre-revolutionary literacy are taken into account.

Figure 1: Distributions of education measures



In a previous study, Lankina et al., 2016 find that the positive relationship between pre-revolutionary literacy and modern democratic outcomes is mediated by Communist Party saturation, meaning that the Communist party of the Soviet Union (CPSU) recruited more educated cadre, leading to its overrepresentation in cultural and educational hubs. As Lankina, 2021 mentions, the “bureaucratic” path, which almost certainly entailed party membership, was one of the occupational choices the former meshchane undertook. This appears to be problematic for the interpretation of the democratic legacy of the meshchane class. It could be that the competitiveness measures of democracy capture the more intense struggles between the old communist regime and non-communist alternatives in 1996 rather than the democratic values of the meshchane descendants.<sup>5</sup>

In any case, it appears that while the educational persistence is the most likely candidate for the observed outcomes, there is substantial ambiguity as to how exactly we can approach the Soviet experience. In line with the arguments of Rosenfeld, 2021, we can hypothesize that the state indeed co-opted the educated, which muddies the interpretation of democratic outcomes - but does not disprove the persistence itself.

---

<sup>5</sup>The conjecture is easy enough to test, yet I do not have the data on district-level CPSU membership, so it remains unexplored empirically here.

## Occupational choices

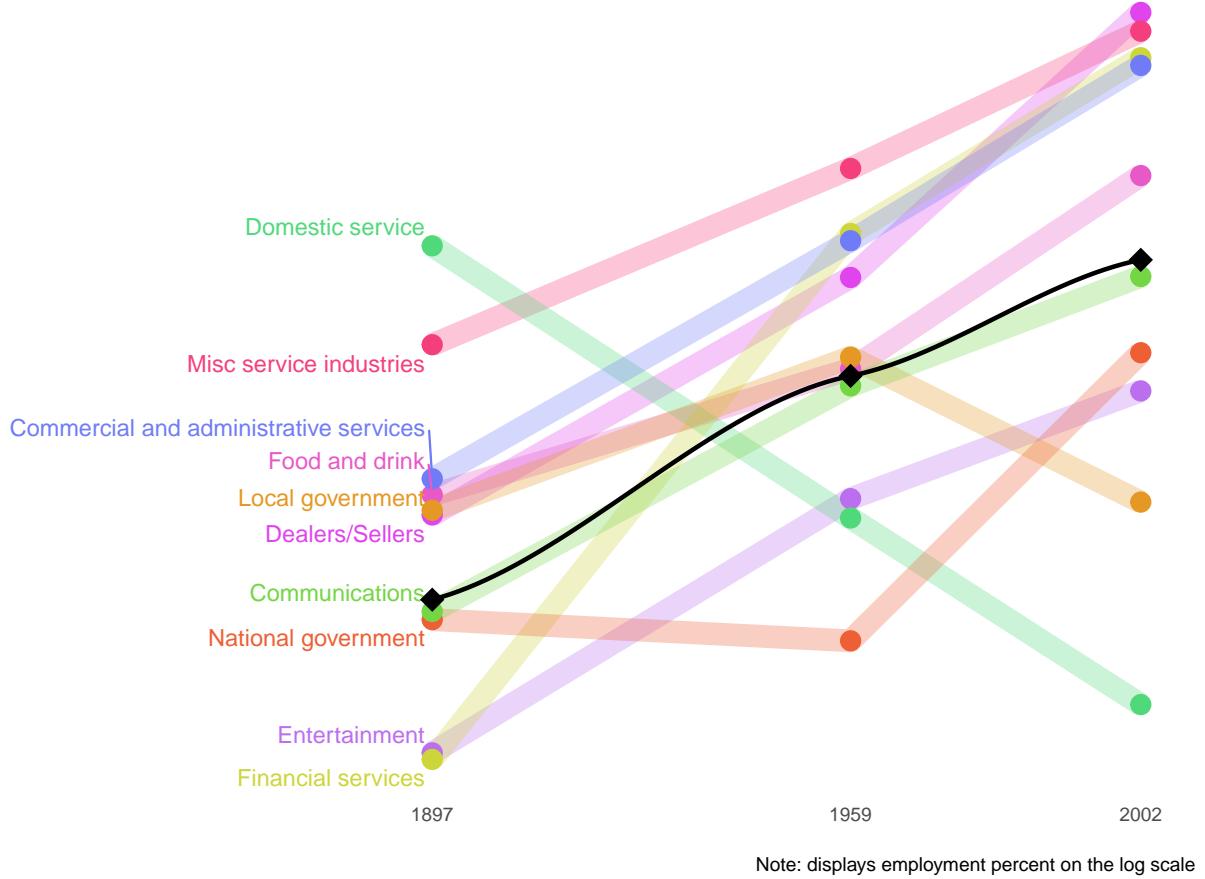
I now consider the persistence in occupational choices. As Torche, 2015 suggest, addressing the link between inequality and mobility is a harder issue when occupational measures of the latter are used. The issue is that the occupational structure of the society is not only a result of the educational system, but also of the economic structure itself, and as a consequence, are strongly related to the time. Given the very-long-run nature of the study, it is hard to establish a direct link between the pre-revolutionary occupational structure and the Soviet one, let alone the post-market transition one, riddled with pervasive inequality.

It is neither the case that a lawyer's descendants will become lawyers, nor that what it means being a lawyer will remain the same. If anything, if any persistence may emerge after two transitions - to Soviet rule and from it - it is in the educational choices, which enable high mobility across the non-menial occupation specter. Lankina, 2021 suggests that the former meshchane would be co-opted for state-building needs, which for some sectors would be unchanged - health-care, education, basic administration and the like; left in a precarious positions; or seeking shelter in isolated cultural domains of museums, libraries or other cultural institutions.

There is no late Soviet Union data on scientists or arts professions. Even if there were, these tend to concentrate near their respective institutions, which may be path dependent on imperial institutions. While that is a mechanism Lankina and Libman, 2021b mention, it is very hard to test. The issues start with general impossibility to match professions in pre-revolutionary Russia, a society only starting to industrialize, with fixed estates and relevant restrictions on occupational mobility, to Soviet-era occupational classifications, that are also peculiar as far as sociological research into occupations goes, and to modern classifications as well, especially considering Soviet emigration, WWII and economic turmoils of the 1990s.

With that in mind, I provide a a general dynamic of national employment in the different subsectors of the tertiary sector, which for all intents and purposes most closely resembles what Lankina, 2021 describes as occupations the persistent middle classes would choose.

Figure 2: Employment in tertiary sectors in 1897-2002



The inferences available from this are very limited indeed: one can note that the share of the tertiary sector grew both between 1897 and 1959 and between 1959 and 2002. The growth is not uniform across the subsectors, and the detailization does not allow to differentiate between the change associated with the changing needs of the era or the regime and the persistence of the pre-revolutionary social structure and occupations.

Moreover, lack of granular occupational data does not allow us to estimate either the effects of professional continuity (matching occupations or sectors one-to-one) or the share of meshchane class on the incidence of occupational choices. Moreover, there is once again no way to disentangle these effects from each other or from the educational persistence.

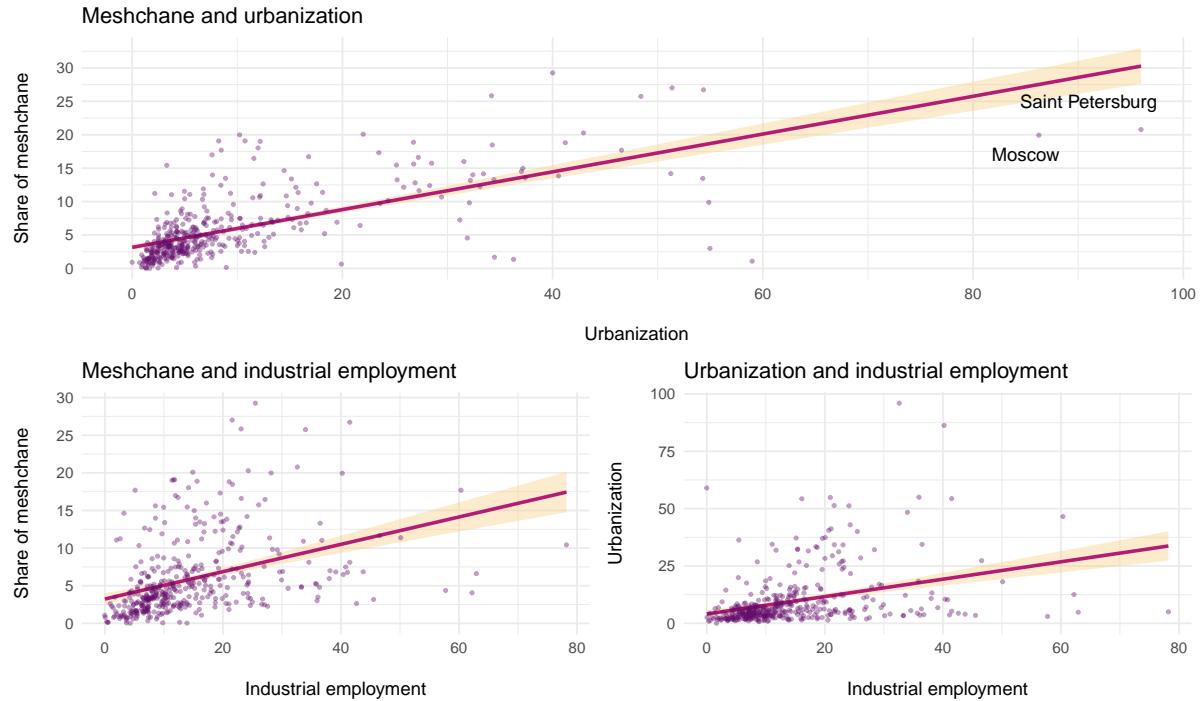
## Urbanization

One of the unduly discarded stories related to the meshchane class is its inherent connection with urbanization. The meshchane were the named urban class of the Russian Empire, created

by Catherine the Great with her “Charter to Towns” back in 1785 (Kollmann, 2017). The charter granted towns own governance and economic rights, but the reform was also aimed at creating a separate “third estate” of urban dwellers, townsmen, who would fill the gap between the nobility and the peasantry (Hittle, 1979). The meshchane were the most diverse estate, ranging from former peasants to professionals.

The urbanization process, very much complementary to the starting industrialization in late Imperial Russia, is another path linking the meshchane class to modern economic performance. It is reasonable to hypothesize that there is an alternative path to political outcomes through economic performance. The urbanization process may put towns with higher meshchane shares on a different trajectory, leading to higher economic performance and, potentially, to different democratic outcomes. Such an effect would likely be separate from social class persistence mechanisms, as it would concern the growth of the towns themselves, rather than the transmission of values and norms or indeed any characteristics of the meshchane class. Figure 3 shows the relationship between urbanization and the share of meshchane in the districts.

Figure 3: Urbanization, meshchane and industrial employment



Since both variables are correlated with industrial employment, I examine the continuity with the Soviet-era industrial configuration. Specifically, whether higher shares of meshchane, higher urbanization and higher industrial employment in 1897 are associated with higher shares of industrial (manufacturing) employment in 1989. It appears to be unlikely that the share of the bourgeoisie would dictate the distribution of industrial activity in Soviet times - however

I believe there is substantial path dependence on initial industrial sites, split among input factors and city agglomerations. I therefore formulate *Hypothesis II*:

$H_2$ : The share of the *meshchane* class in 1897 is not associated with the late Soviet Union distribution of manufacturing, once the effects of urbanization and initial industrial sites are taken into account.

Moreover, one needs to consider effects of urbanization and manufacturing on democracy itself. Natkhov and Pyle, 2023 show that the effect of the market transition put districts specializing in internationally noncompetitive sectors at a disadvantage - making their citizens abandon the transition candidates at the polls, causing a discontinuity in voting behavior between 1991 and 1995. Given this fairly significant finding, we wouldn't expect a systematic increase in competitiveness.<sup>6</sup> In addition to the main Hypothesis 2 I also test  $H_{2a}$ :

$H_{2a}$ : Late Soviet Union distribution of manufacturing does not affect the competitiveness of the 1996 election.

### Great leveller? Repressions in Soviet Union

Among the many ways in which the Soviet regime sought to level the social structure, arguably the most distinctive were the multitude of repressions undertaken against the population. While after 1917 the repressions were directed against the higher social classes of the Russian Empire, by the second half of the century Russia saw repressions against all social classes, from peasants during Industrialization to the military after the Second World War.

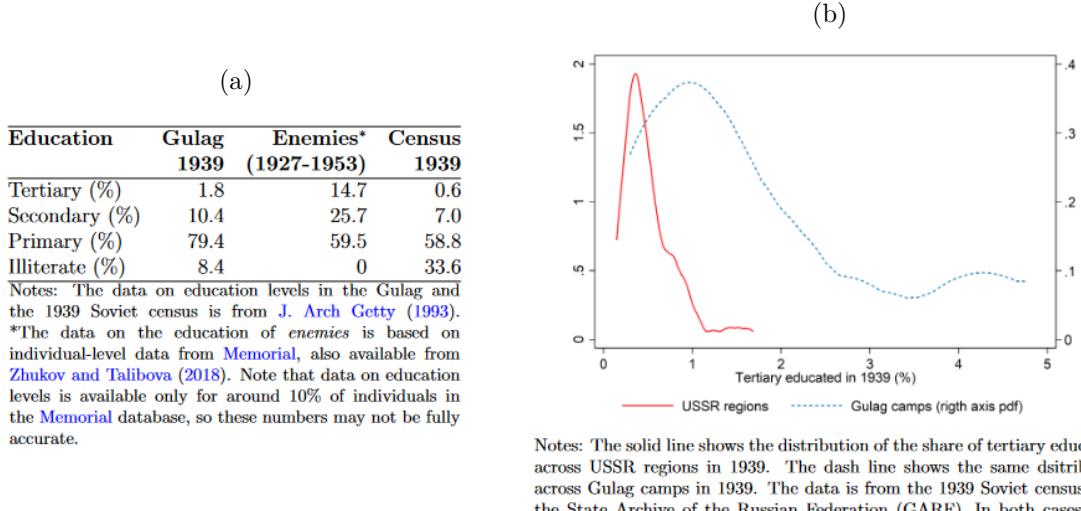
When talking about repressions in Soviet Union, one has to be wary of the classifications inherent to the process. A broad category of *enemies of the people* (“vragi naroda”) denotes all political prisoners. The notion does not include famine victims, ethnic deportations and forced resettlement or those repressed during the Russian Civil War. Within the *enemies of the people*, a variety of formal criminal charges coexist, such as “counter-revolutionary activities”, “being a dangerous social element”, “treason” and many others (Zhukov & Talibova, 2018).

As Toews and Vezina, 2020 show (Figure 4), “enemies of the people” were significantly more educated than the rest of Russia.

---

<sup>6</sup>The reason is mechanical - if, as Natkhov and Pyle, 2023 show, the relationship switched due to a non-discriminate shock to well-being, this wouldn't increase competitiveness - only relative results of candidates. Moreover, it is unlikely that voters would defect to third candidates other than the communist Zyuganov - as the nature of the issue meant that internationally noncompetitive sectors would need to be supported via subsidies.

Figure 4: Education levels in GULAGs and the rest of Soviet Union (taken from Toews and Vezina, 2020).



The implications of this for our study are profound. Firstly, the fact that repressions targeted the educated and the professional classes means that the effects those pre-revolutionary characteristics have on modern outcomes must be limited. On the other hand, due to migration restrictions, and as one of the main points in Toews and Vezina, 2020, the (surviving) repressed did not return to their original places of residence. They settled in towns and regions surrounding the GULAGs they were released from - increasing long-run educational attainment in those districts and contributing to economic activity. However those places also have lower levels of social trust and civic engagement (Nikolova et al., 2022). They were also more likely to exhibit anti-communist voting patterns after the collapse of the Soviet Union and importantly, through 1996 (Kapelko & Markevich, 2014), a goal post year for Lankina and Libman, 2021b.

Moreover, Zhukov and Talibova, 2018 show that arrest locations also matter - they depressed turnout in the variety of post-communist elections in Russia and Ukraine. This establishes that an alternative relation to democracy is possible. The logic is as follows: places where higher shares of people belonged to the higher social stratum there were more repressions (tested in Figure 8), which lead simultaneously to higher upper human capital and to a decrease in civic engagement. Therefore, Hypothesis 3 and 3a are:

- $H_3$ : The share of the *meshchane* class in 1897 is positively associated with the intensity of repressions in the Soviet Union.
- $H_{3a}$ : When accounted for in the model, the intensity of repressions in the Soviet Union decreases the association between the *meshchane* class and modern democratic competi-

tiveness.

## Empirical analysis

In my empirical analysis I follow Freese and Peterson, 2017 on replication and reproducibility in social sciences. I start with a straightforward replication of authors' tables and analyses to ensure *verifiability*. I then consider alternative tests and specifications for *robustness*. Lastly, I use additional data and methods to test competing hypotheses for *repeatability* and *generalizability*.

## Data

The first source of data for the analyses is the replication dataset of Lankina and Libman, 2021a. The dataset contains the variables needed to replicate the figures in the paper and variables used in robustness checks (most of which are unavailable due to the lack of the Statistical Appendix at Harvard Dataverse).

To these data I match by location the 2002 10% sample of the Russian Census, which contains information on education (Kopytok et al., 2021).

I also use data on electoral results from the Central Election Commission of the Russian Federation to create the “effective number of candidates/parties” measure for the 1995, 2000, 2003, 2004 and 2007 elections (Central Election Committee of Russian Federation, 1995, 2000, 2003, 2004, 2007).

The data from Natkhov and Pyle, 2023 on 1989 manufacturing employment is used for post-treatment controls employed in g-estimation and auxiliary tests that did not make it into the paper.

Lastly, I use data on the repressions in the Soviet Union from the Zhukov and Talibova, 2018, taking variables for repression intensity (number of people repressed) and distance to the nearest GULAG camp.

## Replication

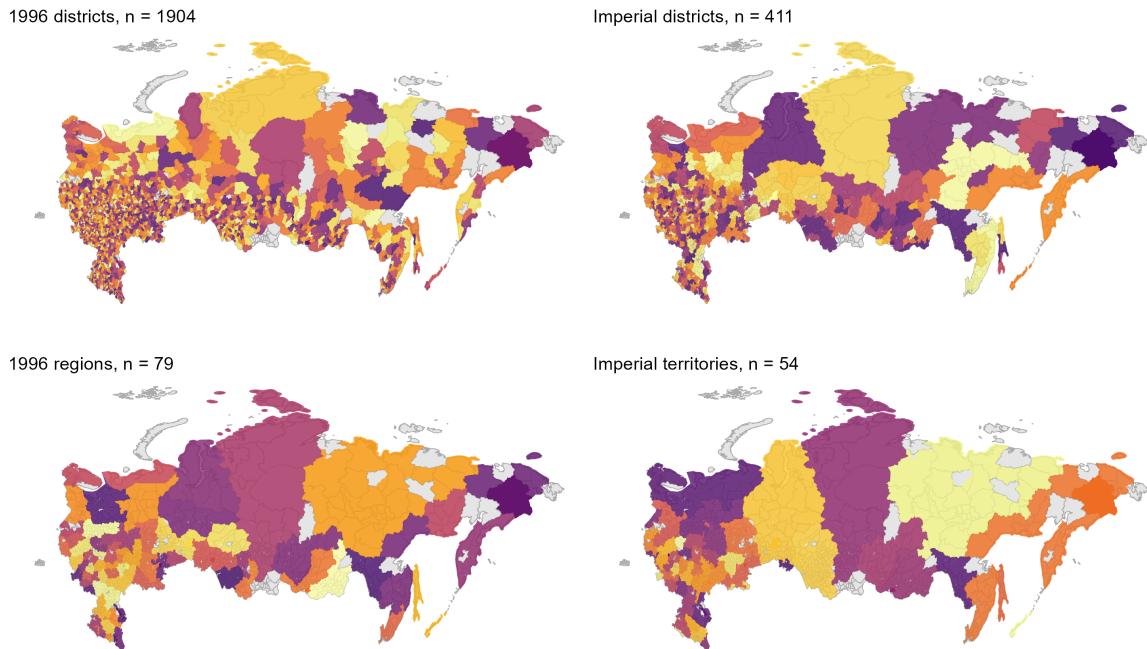
I am able to fully replicate all tables and figures from the original paper.

## Robustness

### Statistical issues and alternative specifications

I subject the models from the paper to a variety of alternative specifications. There are multiple statistical concerns with the models employed by the authors. The first and one of the more major ones is the nested structure of the data. In main regression models, the authors have 2064 to 1780 observations for modern outcomes but only 423 for pre-revolutionary variables. This reflects a natural evolution of the administrative system throughout those two points in time - higher population densities, changes in production and bureaucratic systems necessitated and allowed more fine-grained administrative divisions. However, this also means that one level of the share of the meshchane class predicts multiple results within the modern electoral districts. This is an aggregation issue: modern *rayons* (districts) are nested within historical *uezds* (Illustrated in Figure 5).<sup>7</sup> The nesting means that values of pre-revolutionary variables within the *uezd* are more alike than between *uezds* (in our case they are identical within the *uezd*). This violates the assumption of independent observations and can lead to biased estimates.

Figure 5: Data nesting



<sup>7</sup>This is implied in the paper but no issues are acknowledged in connection to this. The paper promises details in the Statistical Appendix, however, that file is unavailable from the replication data on Harvard Dataverse (Lankina & Libman, 2021a).

We cannot introduce a set of imperial district fixed effects, such as with 1996 regions or 1897 territories, as they will be collinear with the pre-revolutionary predictors. We can, however, treat imperial districts as random effects, allowing us to both account for nesting in the data and retrieve the relevant estimates. This correctly specifies pre-revolutionary predictors as level 2 variables.

Another issue concerns the post-treatment bias. Inclusion of post-treatment variables, in this case, modern doctors per capita, housing and urbanization rates biases the coefficient on the variable of interest, percent of meshchane. Moreover, the authors also use modern region fixed effects. As Homola et al., 2024 argue and show, introduction of post-treatment fixed effects also induces bias in the estimates, even if the treatment has no causal pathway to the border configurations themselves. I therefore treat models with no post-treatment variables as the baseline ones.<sup>8</sup>

Moreover, I estimate the models with imperial-era territory fixed effects instead of modern region fixed effects.<sup>9</sup> To address the issue systematically, I employ a correct procedure that allows for the inclusion of post-treatment controls, sequential-g estimation (Acharya et al., 2016). Apart from assessing whether results hold when accounting for doctors per capita, housing and urbanization, it allows me to test competing mechanisms.

It remains a question whether both random intercepts and random slopes are needed for the correct account of the clustering structure. Given that we are interested in the overall effect of the meshchane class, the random intercepts model is probably sufficient, as the intercepts themselves are a nuisance parameter and are simply needed to account for the structure of the data. I also test this formally in the Appendix. My preferred model therefore substitutes modern region FE for historical imperial territories FE and includes imperial district random effects to account for the nesting structure.

Table 1 presents the results of the alternative models.

The effect appears to be more pronounced for the effective number of candidates measure, even when accounting for the nesting structure. On the other hand, the democratic competitiveness index does not survive the robustness checks. Note that the aggregated and random-intercept models take care of the nesting in the data, albeit in different ways. While these exercises inflate standard errors and reduce coefficients, they are by no means disproving of authors' hypotheses - but suggestive of sensitivity to specifications.

## Other elections

While authors' selection of 1996 elections as the most competitive reflects a valid way of sampling, it makes sense to redo the analysis for the larger selection of post-Soviet election.

---

<sup>8</sup>Those correspond to columns (7) and (8) in Table I of the original paper.

<sup>9</sup>Due to the way imperial territories were defined, see @fig-nesting, the fixed effects are highly correlated with longitude. I drop longitude from all models employing territory fixed effects. The results are robust to this decision.

Table 1: Alternative specifications

	Original	Imperial FE	Robust LM	Clustered FE	Aggregated	Random intercepts
<i>Dependent variable: Vanhanen index of democratic competitiveness</i>						
Percent Meschane	0.196*** (0.050)	0.153** (0.054)	0.152*** (0.042)	0.196*** (0.053)	0.105 (0.068)	0.135* (0.068)
<i>Dependent variable: Effective number of candidates</i>						
Percent Meschane	0.017*** (0.003)	0.019*** (0.004)	0.020*** (0.004)	0.017*** (0.004)	0.015** (0.005)	0.017*** (0.005)
Std. Errors	Robust	Robust	Classical	~ Imperial District	Robust	Classical
Num.Obs.	2061	2061	2061	2061	415	2061

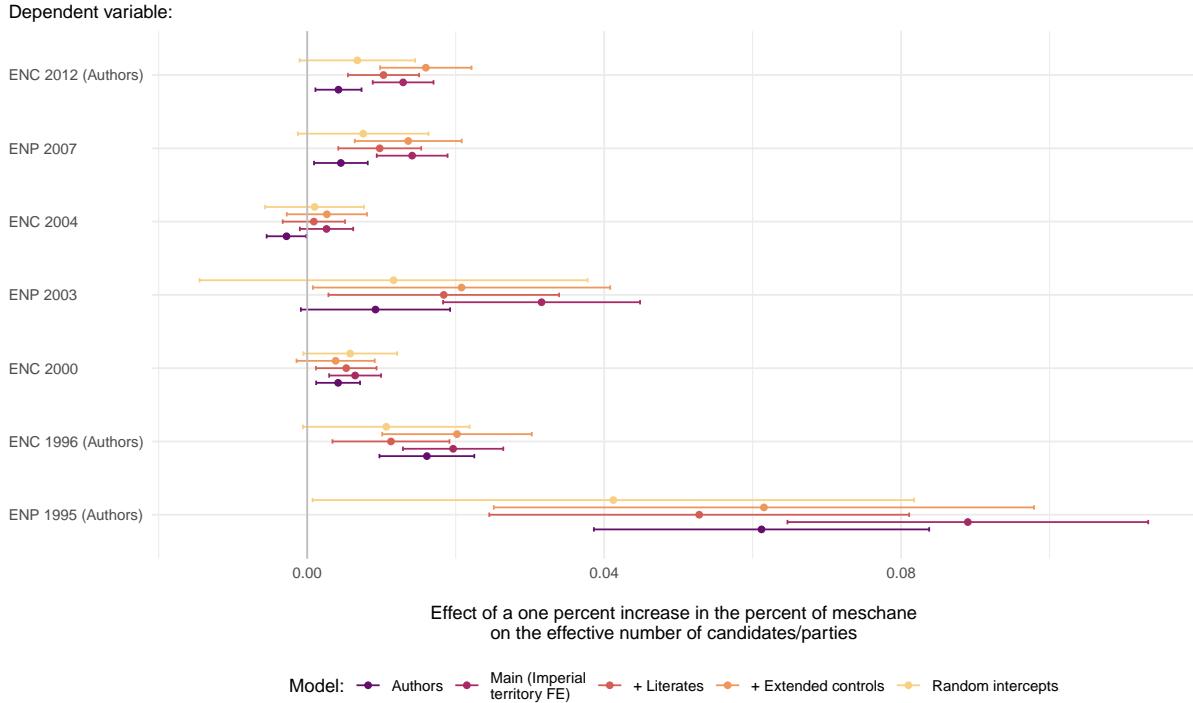
\* p < 0.1, \*\* p < 0.005, \*\*\* p < 0.001

Controls include shares of different religious adherents as well as latitude

All models except the original authors' use imperial district fixed effects

Using data on 1995, 1996, 2000, 2003, 2004, 2007 and 2012 elections, I can test whether the size of pre-revolutionary meshchane class has a correlation with their competitiveness (Central Election Committee of Russian Federation, [1995](#), [2000](#), [2003](#), [2004](#), [2007](#); Lankina & Libman, [2021a](#)). . Due to data limitations, I can only estimate the effective number of candidates (parties) and not the Vanhanen Index for the alternative years. Figure 6 presents the results of various estimations.

Figure 6: Effects of meschane shares on the effective number of candidates/parties in different years



As with the the 1996 models, adding imperial district random effects in general reduces the effect of the meschane class on the effective number of candidates/parties for different years. I have to conclude that if accounting for such a nesting structure is correct, one cannot conclusively confirm an effect of meshchane on the competitiveness of elections, especially given the more optimistic nature of the *effective number of candidates/parties* measure.

Accounting for pre-treatment fixed effects, compared to the models authors' fit interestingly only increases the effect, while accounting for the share of literates predictably decreases it, due to correlation between the two variables (Person's  $\rho = 0.43$ ). The 2004 Presidential election appears as a bit of an outlier - all estimates are non-significant and the authors' model is even producing a negative coefficient. This may be because of Vladimir Putin's rise to power and the "rally around the flag" effect due to the end of active phase of the Second Chechen War in 2002-2003, a feat generally attributed to the president.

# Results

## Education

I start by working with  $H_1$ , on the issue of educational persistence. Using modern representative census microdata I build multiple measures of educational attainment on the district (rayon) level. I then regress two variables of interest - pre-revolutionary share of literate people and the share of meshchane class - on those measures. Results from different model specifications are presented below.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Dependent variable: Percent with higher education in 2002</i>								
% Literate	0.130*** (0.016)		0.155*** (0.021)		0.102*** (0.022)	0.100*** (0.027)		0.064* (0.028)
Meshchane		0.183*** (0.023)		0.160*** (0.026)	0.124*** (0.027)		0.117*** (0.033)	0.098** (0.033)
<i>Dependent variable: Percent illiterate in 2002</i>								
% Literate	-0.033*** (0.002)		-0.026*** (0.004)		-0.021*** (0.004)	-0.018*** (0.004)		-0.016*** (0.005)
Meshchane		-0.025*** (0.002)		-0.023*** (0.004)	-0.012* (0.004)		-0.013* (0.005)	-0.002 (0.005)
Controls	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Region FE	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Random intercept	No	No	No	No	No	Yes	Yes	Yes
Num.Obs.	2056	2080	2055	2079	2055	2055	2079	2055

\* p < 0.1, \*\* p < 0.005, \*\*\* p < 0.001

Controls include shares of different religious adherents as well as latitude and longitude.

Models in columns (1) - (5) use stata-style robust standard errors.

Models in columns (6) - (8) are mixed effects models estimated with REML and use classical standard errors

As is evident from the table, both pre-revolutionary literacy rates and meshchane shares have effects on the share of people with higher education in 2002. The effects are robust to different model specifications and controls. An increase of 5 percent in the share of the literate in 1897 on average is associated with an increase of 0.32 percent in the share of people with higher education in 2002, controlling for pre-treatment variables.<sup>10</sup> An analogous effect of meshchane bears an average association of a 0.49 percent (0.07 of the dependent's standard deviation) increase in the share of people with higher education in 2002 for a 5 percent increase in the share of meshchane in 1897.

When using the first dependent variable, the share of meshchane has an effect independent of district's pre-revolutionary literacy rates. This results suggests that there is indeed intergenerational transmission of upper level educational preferences.

---

<sup>10</sup>When jointly considering its effects with meshchane shares, column (7).

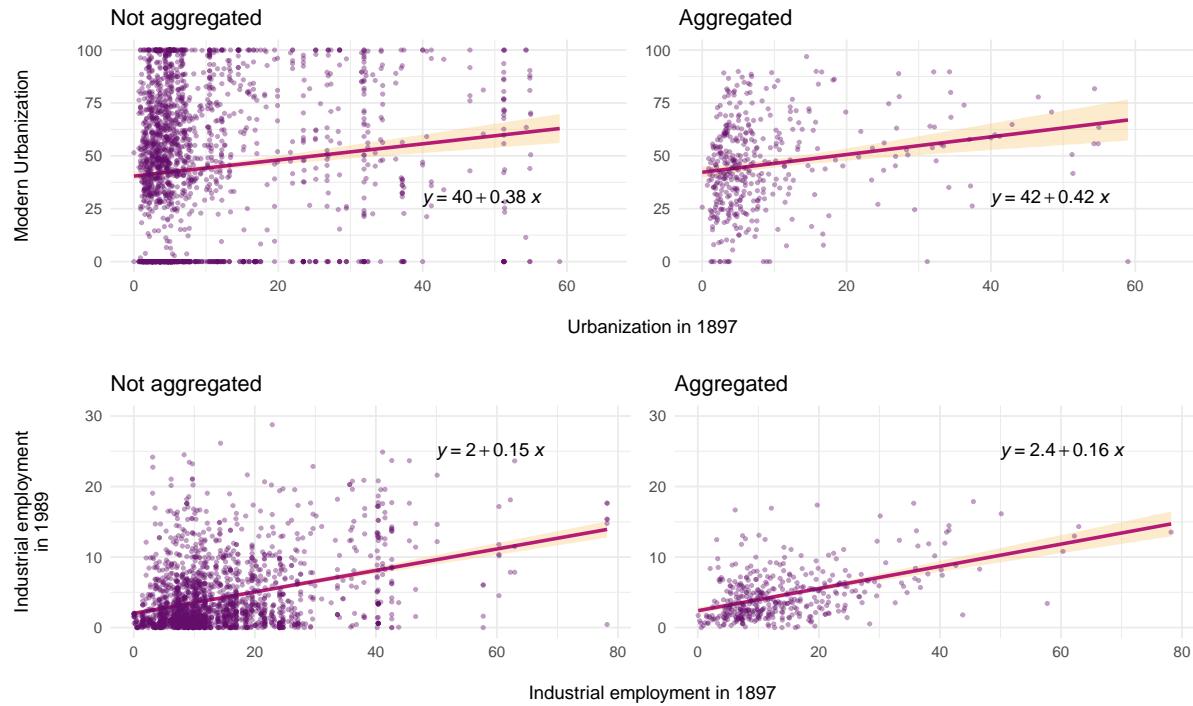
Supporting this notion, the effects of the meshchane class disappear when considering the modern shares of the illiterate. Pre-revolutionary literacy rates, on the other hand, remain important for that variable. An increase of 5 percent in the share of literate in 1897 is associated with a decrease of 0.08 (0.11 of the dependent's standard deviation) percent in the share of illiterate in 2002.

This evidence suggests that there is a persistent effect of the meshchane class on modern upper human capital educational outcomes. While this does not confirm a causal relationship, it illustrates that there is indeed significant social persistence, even over the period of more than a 100 years.

Therefore we can provisionally confirm a relationship hypothesized in  $H_1$ .

## Urbanization

Figure 7: Trends in urbanization and industrial employment by nesting



The issue with testing the hypothesis of urban persistence is that the issue of nesting in the data is exacerbated. Namely, modern urbanization rates are computed by-rayon, while the pre-revolutionary urbanization rates are computed by-uezd. This means that there is a big mismatch between urbanization rates within the uezd - a typical uezd would consist of an uezd city, its satellites and the surrounding rural parishes. Therefore, if most people lived in the city,

Table 2: Effects of estate shares on urbanization and industrial employment

	(1)	(2)	(3)	(4)	(5)	(6)
<i>Dependent variable: Modern urbanization, percent</i>						
Urbanization	0.407*** (0.107)				-0.211 (0.173)	-0.303 (0.187)
Percent Meshchane		1.304*** (0.208)			0.839* (0.385)	1.303** (0.407)
Industrial Employment			0.842*** (0.107)		0.622*** (0.120)	
Literates				1.005*** (0.218)	0.468* (0.275)	0.861** (0.267)
Num.Obs.	415	415	415	412	412	412
R2	0.417	0.439	0.478	0.429	0.492	0.454
<i>Dependent variable: 1989 Soviet manufacturing employment, percent</i>						
Urbanization	0.026 (0.018)				-0.063* (0.027)	-0.079* (0.028)
Percent Meshchane		0.155*** (0.038)			0.159* (0.059)	0.253*** (0.065)
Industrial Employment			0.142*** (0.020)		0.143*** (0.024)	
Literates				0.078* (0.035)	-0.022 (0.043)	0.071* (0.043)
Num.Obs.	411	411	411	408	408	408
R2	0.403	0.426	0.496	0.407	0.509	0.437
Std.Errors	HC1	HC1	HC1	HC1	HC1	HC1

\* p < 0.1, \*\* p < 0.005, \*\*\* p < 0.001

the district would have high urbanization rates - but if we would compute the urbanization by parish, most of them would be rural. Since in modern times the administrative division treats those rural areas as own municipalities, this is precisely what happens. Figure 7 illustrates the issue by plotting association between pre-revolutionary urbanization and modern urbanization with and without aggregation.

I therefore aggregate to the uezd level all models with urbanization as the dependent variable. I first work with continuity in urbanization itself, focusing on three variables of interest - imperial urbanization, share of meshchane and imperial industrial employment. The results are presented in Table 2.

The results show some interesting patterns. Firstly, while previous urbanization rates matter

for modern urbanization, once literacy and industrial employment are accounted for, the effect disappears. Secondly, there appears to be no effect of the meshchane share on modern urbanization, with only industrial employment having an effect when accounting for literacy, urbanization and the meshchane share. In particular, a 10 percent increase in the share of industrial employment in 1897 (around one standard deviation of that variable) is on average associated with a 8.42 percent increase in modern urbanization (0.4 of the dependent's sd), all other variables held equal.

The Soviet manufacturing configuration is also strongly associated with pre-revolutionary industrial employment. A 10 percent increase in the share of industrial employment in 1897 is associated with a 1.43 percent increase in the share of people employed in manufacturing in 1989 (also 0.4 standard deviation of the dependent). This is also an effect separate from the share of meshchane, literacy and urbanization.

Therefore, I accept  $H_2$  - the share of the meshchane class in 1897 is not associated with the late Soviet Union distribution of manufacturing, once the effects of urbanization and initial industrial sites are taken into account. I test  $H_{2a}$  in the section "Comparing mechanisms of social persistence".

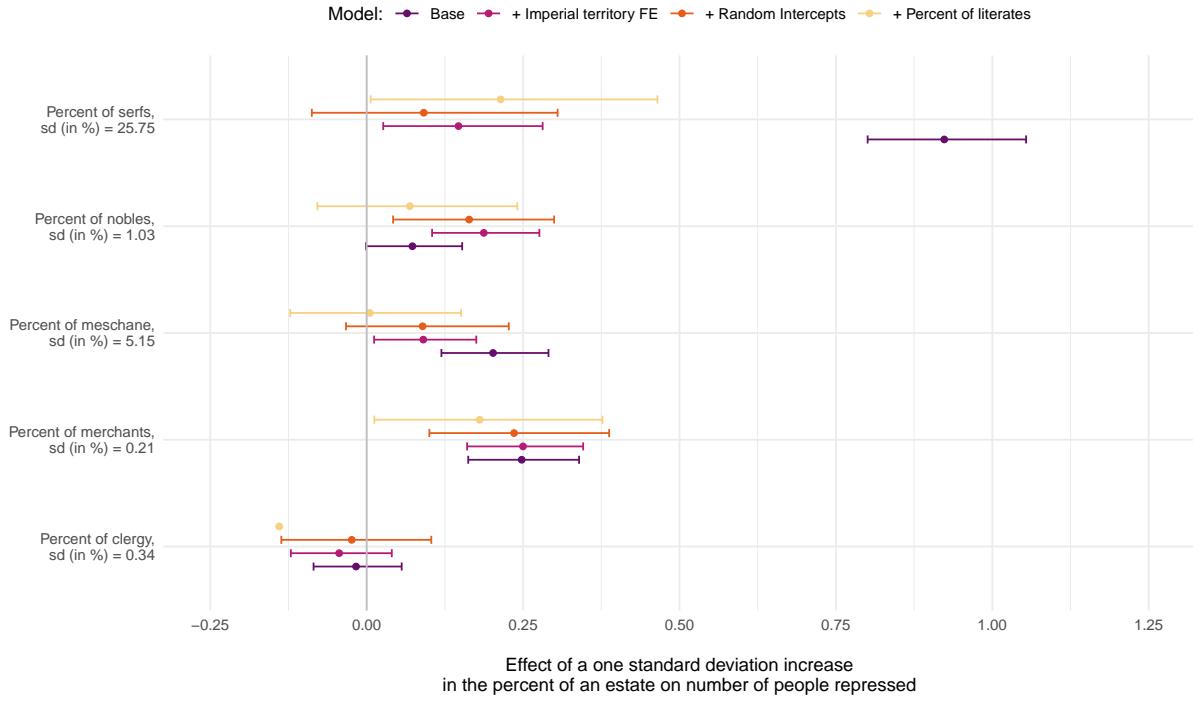
## Repressions

I now turn to repressions, which I hypothesize suppress the effects. Figure 8 plots the estimates in standard deviations of the estate predictor, to make them comparable between each other.

The two preferred specifications: random intercepts and random intercepts controlling also for literacy rates reveal that we can reject  $H_3$ . The share of the meshchane class in 1897 is not associated with the intensity of repressions in the Soviet Union. Importantly, this grants credence to the conceptual argument in Lankina, 2021; Lankina and Libman, 2021b that the meshchane class was more numerous and more needed for co-optation than the nobles and the merchants, two clearer targets for "social homogenization" policies.

Given this result, the question posited in  $H_{3a}$  is moot. I test it anyways when considering different mechanisms of social persistence in the section "Comparing mechanisms of social persistence" formally.

Figure 8: Effects of estate shares on repression intensity



## Comparing mechanisms of social persistence

### Sequential-G estimation results

I now turn to g-estimation as a way to test relative strength of different mechanisms in explaining 1996 electoral competitiveness. For the educational persistence channel, I focus on authors' measure of education (shown in Figure 1). For occupations, I assess whether authors' measure of doctors per capita (as a profession associated with extensive training and not necessarily affected by political factors). I also work with urbanization, soviet manufacturing and repressions.

Sequential g-estimation, described in Acharya et al., 2016, allows one to control for post-treatment variables and a mediator, the latter being a variable that is the mechanism, that is, affected by the treatment and affects the outcome. The coefficient on *percent of meshchane* is its demediated effect (without the specified mechanism). For all models, I control for post-treatment distance to GULAG, Soviet-era military plants per capita, housing and modern "closed city" status. In models that do not specify repressions or soviet manufacturing as mediators, I control for them as post-treatment variables. The results are presented in Table 3.

When we specify education as the mechanism, the coefficient is reduced by significantly higher margins than for any other mechanisms, indicating that when we factor in education, percent

Table 3: Effects of estate shares through different mechanisms

	No Post-treatment variables	Controlling for post-treatment	No mediators	Education	Urbanization	Doctors	Soviet manufacturing	Repressions	All
<i>Dependent variable: Vanhanen index of democratic competitiveness</i>									
Percent Meshchane	0.153** (0.048)	0.153** (0.048)	0.104* (0.054)	0.016 (0.063)	0.088 (0.054)	0.092* (0.055)	0.100* (0.054)	0.105* (0.055)	0.021 (0.064)
Num.Obs.	2061	2061	1837	1467	1837	1788	1837	1837	1458
<i>Dependent variable: Effective number of candidates</i>									
Percent Meshchane	0.019*** FE	0.012*** Imperial	0.016*** Imperial	0.007* Imperial	0.013*** Imperial	0.013** Imperial	0.015*** Imperial	0.016*** Imperial	0.007* Imperial
Estimation	LM (0.004)	LM (0.003)	Sq G (0.004)	Sq G (0.004)	Sq G (0.004)	Sq G (0.004)	Sq G (0.004)	Sq G (0.004)	Sq G (0.004)
Num.Obs.	2061	1781	1837	1467	1837	1788	1837	1837	1458

\* p < 0.1, \*\* p < 0.005, \*\*\* p < 0.001

of meshchane fails to explain any of the variation in electoral competitiveness. This leads me to conclude that it is the main channel through which observed effects on democratic competitiveness are determined.

Here we also observe that the Vanhanen Index is more responsive to competing explanations than the effective number of candidates. Results for the latter are also significantly reduced, with percent of meshchane significant at only 10% level, which is not usually considered as a robust result.

Urbanization also manages to reduce the effect of the meshchane class to non-significant, albeit only for the Vanhanen Index. Out of other potential mediators, controlling for repressions has the smallest effect on the coefficient, indicating that the repressions are not a significant mediator of the effect of the meshchane class on democratic competitiveness and disproving  $H_{3a}$ . The same can be said about the Soviet manufacturing configuration, which confirms  $H_{2a}$ .

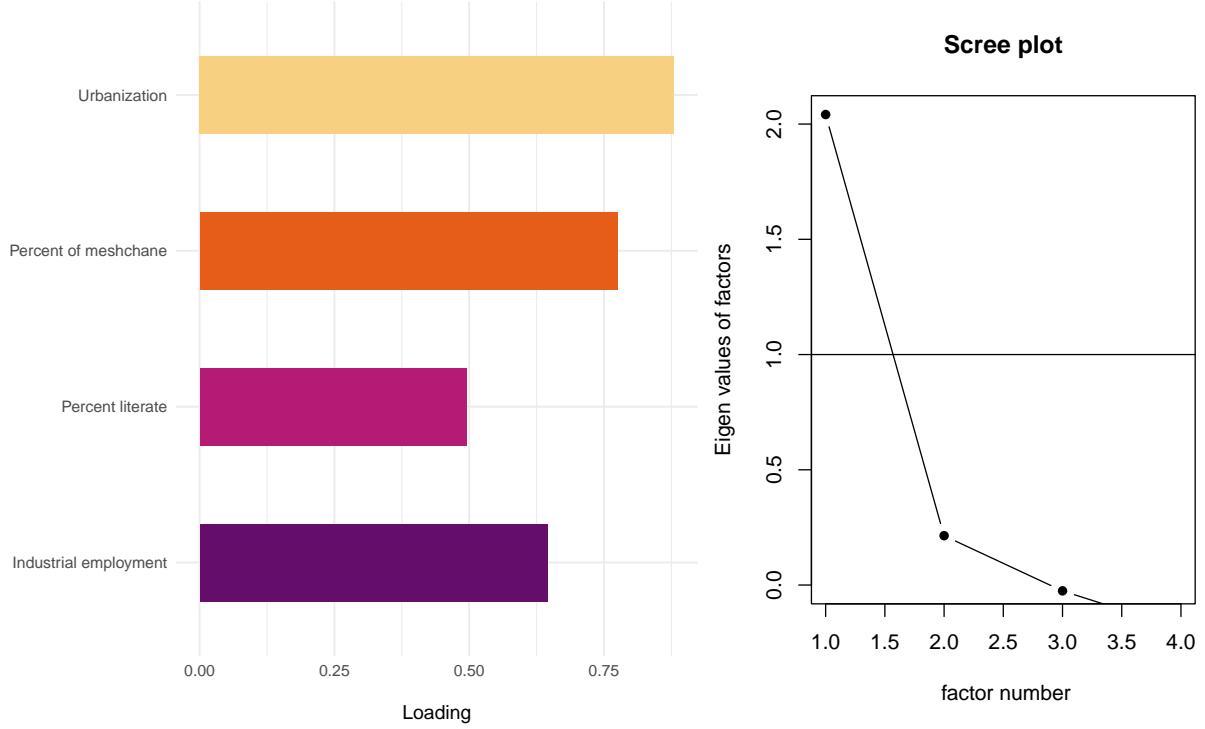
Therefore, I conclude that the main effect of the meshchane on democratic competitiveness is mediated through educational persistence. Urbanization and number of doctors per capita (as a rough occupational persistence proxy), but they are secondary to the educational channel.

## Factor Analysis

Since percent of meshchane, percent of literates, industrial employment and urbanization are correlated to the point where it is hard to distinguish their long-run effect, I perform a factor analysis to see whether they can be reduced to a single factor. Test of the hypothesis that 1 factor is sufficient confirms that there are no other latent variables.

The factor loadings and the scree plot are presented in Figure 9.

Figure 9: Factor analysis scree plot and factor loadings



I then use the factor scores as a predictor in the preferred specification model. The results are presented in Table 4 and indicate that the factor score is a significant predictor and actually has a larger effect on the dependent variables than each of its components. This suggests that there is a single process linking pre-revolutionary urban social and economic structures to modern electoral competitiveness in Russia.

Table 4: Using factor analysis score to estimate a compound effect of urbanization, meshchane, industrial employment and literacy.

	(1)	(2)	(3)	(4)	(5)
<i>Dependent variable: Vanhanen index of democratic competitiveness</i>					
Percent Meshchane	0.023*** (0.005)				
Industrial Employment		0.011*** (0.002)			
Urbanization			0.006** (0.002)		
Literates				0.020*** (0.004)	
Factor 1 Score					0.121*** (0.027)
<i>Dependent variable: Effective number of candidates</i>					
Percent Meshchane	0.131* (0.062)				
Industrial Employment		0.043 (0.032)			
Urbanization			0.042 (0.026)		
Literates				0.165** (0.051)	
Factor 1 Score					0.773* (0.355)
Num.Obs.	2028	2028	2028	2028	2028

\* p < 0.1, \*\* p < 0.005, \*\*\* p < 0.001

## Conclusion

In this paper, I have explored persistence of education and occupations in Russia, engaging with the works of Lankina, 2021; Lankina and Libman, 2021b. I argue that the multifaceted process of social transmission in Russia, specifically through the 1917 Revolution is a valid and exciting research agenda, but the explicit link to early modern democratic outcomes is a bit tricky to establish. I deconstruct some of the potential channels of such transmission, focusing on educational persistence, urbanization, industrial employment and repressions.

I find that there is persistence in higher upper human capital transmission (through higher education) that is explained by the share of the pre-revolutionary bourgeoisie even when accounting for pre-revolutionary literacy rates. When comparing relative importance of different mechanisms through which the meshchane class could affect democratic competitiveness, I find that the educational channel is the most important one. This suggests that the meshchane class is associated with higher human capital that likely increases democratic competitiveness even across multiple generations and institutional settings.

Urbanization and professional persistence (on the example of doctors) are secondary mechanisms to the educational channel. As an urban class, the meshchane are associated with higher urbanization rates, but the effect is hard to separate from previous urbanization, literacy and imperial industrial employment. When we consider the controlled effects of these factors, imperial manufacturing configurations have a strong independent effect on both modern urbanization and late Soviet Union manufacturing configurations. I document that while the former has an effect on democratic outcomes, the former doesn't. I hypothesize that the relationship between late Soviet Union manufacturing and modern electoral competitiveness is dependent on the market transition process and associated shocks, not long-run trends.

Stalin-era repressions are considered as a “great leveler” process, equalizing meshchane with the former peasant classes as a way to “homogenize” the social structure within the bolshevik project. In that regard, I validate the logic of Lankina, 2021; Lankina and Libman, 2021b, who theorized that the meshchane due to their numbers were more likely to be co-opted than the nobles or the merchants, as they were needed. I confirm this empirically: places with higher shares of meshchane did not suffer higher intensity of repressions, but places with more nobles and merchants did. I find that the meshchane class is not associated with the intensity of repressions in the Soviet Union, and subsequently, that controlling for repressions does not change the effect of the meshchane class on democratic competitiveness.

When trying to ascertain long-run covariation between pre-revolutionary predictors and modern outcomes, I find that several correlated variables are individually important. There appears to be a single latent factor that serves as pre-revolutionary determinant of modern electoral competitiveness. By using factor analysis, I show that the meshchane class, urbanization, industrial employment and literacy rates can be reduced to a single factor that is a very important predictor of modern electoral competitiveness.

## References

- Acemoglu, D., & Robinson, J. A. (2006). *Economic origins of dictatorship and democracy*. Cambridge University press.
- Acharya, A., Blackwell, M., & Sen, M. (2016). Explaining causal findings without bias: Detecting and assessing direct effects. *American Political Science Review*, 110(3), 512–529. <https://doi.org/10.1017/S0003055416000216>
- Alesina, A., Seror, M., Yang, D. Y., You, Y., & Zeng, W. (2020). *Persistence through revolutions*. National Bureau of Economic Research.
- Ansell, B. W., & Samuels, D. (2014). *Inequality and democratization: An elite-competition approach*. Cambridge University Press.
- Austin, G. (2008). The ‘reversal of fortune’ thesis and the compression of history: Perspectives from African and comparative economic history. *Journal of International Development*, 20(8), 996–1027. <https://doi.org/10.1002/jid.1510>
- Becker, G. S., & Tomes, N. (1979). An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of Political Economy*, 87(6), 1153–1189. <https://doi.org/10.1086/260831>
- Belsky, D. W., Domingue, B. W., Wedow, R., Arseneault, L., Boardman, J. D., Caspi, A., Conley, D., Fletcher, J. M., Freese, J., Herd, P., Moffitt, T. E., Poulton, R., Sicinski, K., Wertz, J., & Harris, K. M. (2018). Genetic analysis of social-class mobility in five longitudinal studies. *Proceedings of the National Academy of Sciences*, 115(31). <https://doi.org/10.1073/pnas.1801238115>
- Blau, P. M., & Duncan, O. D. (1978). *The american occupational structure*. Free press.
- Bourdieu, P., & Passeron, J.-C. (2000). *Reproduction in education, society and culture* (2. ed., reprinted). Sage Publ.
- Carina Mood & Mood, C. (2017). More than money: Social class, income, and the intergenerational persistence of advantage [MAG ID: 2601387044]. *Sociological Science*, 4(12), 263–287. <https://doi.org/10.15195/v4.a12>
- Central Election Committee of Russian Federation. (1995). Results of 1995 parliamentary election.
- Central Election Committee of Russian Federation. (2000). Results of 2000 presidential election.
- Central Election Committee of Russian Federation. (2003). Results of 2003 parliamentary election.
- Central Election Committee of Russian Federation. (2004). Results of 2004 presidential election.
- Central Election Committee of Russian Federation. (2007). Results of 2007 parliamentary election.
- Chen, H., & Suen, W. (2017). Aspiring for change: A theory of middle class activism. *The Economic Journal*, 127(603), 1318–1347. <https://doi.org/10.1111/ecoj.12346>
- Corak, M. (2013). Income inequality, equality of opportunity, and intergenerational mobility. *Journal of Economic Perspectives*, 27(3), 79–102. <https://doi.org/10.1257/jep.27.3.79>

- Erzsébet Bukodi & John H. Goldthorpe. (2013). Decomposing ‘social origins’: The effects of parents’ class, status, and education on the educational attainment of their children [MAG ID: 2139068022]. *European Sociological Review*, 29(5), 1024–1039. <https://doi.org/10.1093/esr/jcs079>
- Freese, J., & Peterson, D. (2017). Replication in social science. *Annual Review of Sociology*, 43(1), 147–165. <https://doi.org/10.1146/annurev-soc-060116-053450>
- Gerber, T. P., & Hout, M. (1995). Educational stratification in russia during the soviet period. *American Journal of Sociology*, 101(3), 611–660. <https://doi.org/10.1086/230755>
- Greven, S., & Kneib, T. (2010). On the behaviour of marginal and conditional AIC in linear mixed models. *Biometrika*, 97(4), 773–789. <https://doi.org/10.1093/biomet/asq042>
- Hittle, J. M. (1979). *The service city: State and townsmen in russia, 1600-1800*. Harvard University Press.
- Homola, J., Pereira, M. M., & Tavits, M. (2024). Fixed effects and post-treatment bias in legacy studies. *American Political Science Review*, 118(1), 537–544. <https://doi.org/10.1017/S0003055423001351>
- Jani Erola, Sanni Jalonens, & Hannu Lehti. (2016). Parental education, class and income over early life course and children’s achievement [MAG ID: 2250131611]. *Research in Social Stratification and Mobility*, 44, 33–43. <https://doi.org/10.1016/j.rssm.2016.01.003>
- John Jerrim & Lindsey Macmillan. (2015). Income inequality, intergenerational mobility, and the great gatsby curve: Is education the key? [MAG ID: 1792190074]. *Social Forces*, 94(2), 505–533. <https://doi.org/10.1093/sf/sov075>
- Kapelko, N., & Markevich, A. (2014). The political legacy of the gulag archipelago. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2516635>
- Kollmann, N. S. (2017). *The russian empire 1450-1801* (First edition) [OCLC: ocn969962873]. Oxford University Press.
- Kopytok, V., Kuzmina, Y., & Bobylev, S. (2021). 2002 census microdata: Ten percent sample. Retrieved June 13, 2024, from [https://data-rsci-science.translate.goog/data-catalog/datasets/189/?\\_x\\_tr\\_sl=ru&\\_x\\_tr\\_tl=en&\\_x\\_tr\\_hl=en&\\_x\\_tr\\_pto=sc#dataset-custom\\_tab\\_59](https://data-rsci-science.translate.goog/data-catalog/datasets/189/?_x_tr_sl=ru&_x_tr_tl=en&_x_tr_hl=en&_x_tr_pto=sc#dataset-custom_tab_59)
- Lankina, T. (2021, November 30). *The estate origins of democracy in russia: From imperial bourgeoisie to post-communist middle class* (1st ed.). Cambridge University Press. <https://doi.org/10.1017/9781009071017>
- Lankina, T., & Libman, A. (2021a). Replication data for: The two-pronged middle class: The old bourgeoisie, new state-engineered middle class and democratic development [Version Number: V1]. <https://doi.org/10.7910/DVN/JO8C7A>
- Lankina, T., & Libman, A. (2021b). The two-pronged middle class: The old bourgeoisie, new state-engineered middle class, and democratic development. *American Political Science Review*, 115(3), 948–966. <https://doi.org/10.1017/S000305542100023X>
- Lankina, T., Libman, A., & Obydenkova, A. (2016). Appropriation and subversion: Precommunist literacy, communist party saturation, and postcommunist democratic outcomes [MAG ID: 2233648642]. *World Politics*, 68(2), 229–274. <https://doi.org/10.1017/s0043887115000428>

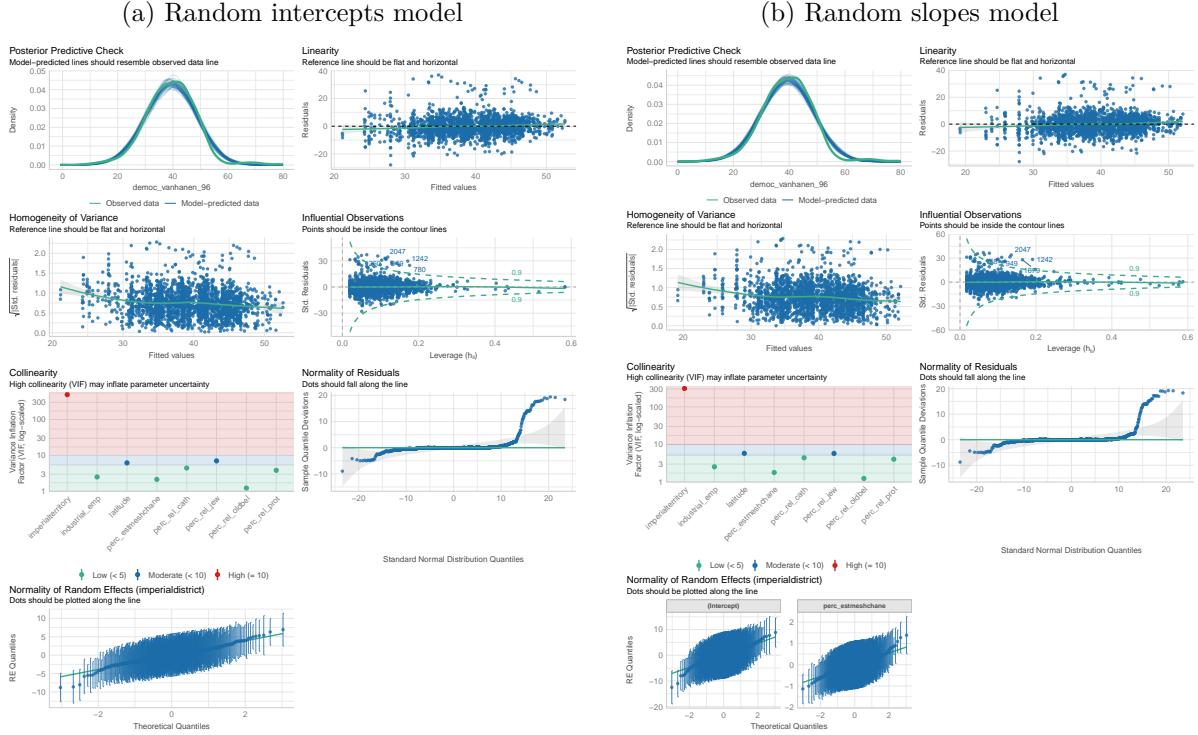
- Lertchoosakul, K. (2021). The paradox of the thai middle class in democratisation. *TRaNS: Trans -Regional and -National Studies of Southeast Asia*, 9(1), 65–79. <https://doi.org/10.1017/trn.2020.16>
- Liu, H. (2018). Social and genetic pathways in multigenerational transmission of educational attainment. *American Sociological Review*, 83(2), 278–304. <https://doi.org/10.1177/0003122418759651>
- Martin Hällsten & Fabian T. Pfeffer. (2017). Grand advantage: Family wealth and grandchildren's educational achievement in sweden. [MAG ID: 2593160432]. *American Sociological Review*, 82(2), 328–360. <https://doi.org/10.1177/0003122417695791>
- Natkhow, T., & Pyle, W. (2023). Revealed in transition: The political effect of planning's legacy. *European Economic Review*, 159, 104567. <https://doi.org/10.1016/j.eurocorev.2023.104567>
- Nikolova, M., Popova, O., & Otrachshenko, V. (2022). Stalin and the origins of mistrust. *Journal of Public Economics*, 208, 104629. <https://doi.org/10.1016/j.jpubeco.2022.104629>
- Resnick, D. (2015). The middle class and democratic consolidation in zambia. *Journal of International Development*, 27(5), 693–715. <https://doi.org/10.1002/jid.3101>
- Rosenfeld, B. (2021). State dependency and the limits of middle class support for democracy. *Comparative Political Studies*, 54(3), 411–444. <https://doi.org/10.1177/0010414020938085>
- Sandra E. Black & Paul J. Devereux. (2010). Recent developments in intergenerational mobility [MAG ID: 3124094248]. *Handbook of Labor Economics*, 4, 1487–1541. [https://doi.org/10.1016/s0169-7218\(11\)02414-2](https://doi.org/10.1016/s0169-7218(11)02414-2)
- Schotte, S. (2021). The anxious and the climbers: Ambivalent attitudes towards democracy among south africa's middle class. *The Journal of Development Studies*, 57(3), 465–483. <https://doi.org/10.1080/00220388.2020.1802007>
- Toews, G., & Vezina, P.-L. (2020, December 15). Enemies of the people. <https://doi.org/10.31235/osf.io/gnypr>
- Torche, F. (2015). Analyses of intergenerational mobility: An interdisciplinary review. *The ANNALS of the American Academy of Political and Social Science*, 657(1), 37–62. <https://doi.org/10.1177/0002716214547476>
- Wirtschafter, E. K. (1997). *Social identity in imperial russia* [OCLC: 1241099236]. Cornell University Press.
- Zhukov, Y. M., & Talibova, R. (2018). Stalin's terror and the long-term political effects of mass repression. *Journal of Peace Research*, 55(2), 267–283. <https://doi.org/10.1177/0022343317751261>

# Appendix

## Random effects

The main choice with regards to random effects is whether random slopes should be included in the models. I start with some basic model diagnostics, shown in Figure 10.

Figure 10: Random effects model diagnostics



Both models appear to be a good fit to the data. I now compare the models with and without random slopes across a variety of goodness-of-fit measures.

Table 5: Comparison of models with and without random slopes

Name	Model	R2_conditional	R2_marginal	ICC	RMSE	cAIC
Random intercepts	lmerMod	0.410	0.288	0.171	6.720	14137.45
Random slopes	lmerMod	0.428	0.295	0.189	6.739	14124.67

The more robust metric by which to compare the modes is cAIC, developed by Greven and Kneib, 2010 specifically to compare models with the same fixed effects but different random

effects. The model with random slopes is preferred by this metric, however ICC should be taken into account as well. Given that the ICC is quite low (it is easy to show that this is due to the inclusion of imperial territory fixed effects, the analogous models without them produce ICC equal to 0.327 and 0.333 respectively) and that there are no first-level predictors, there is no way to substantively include varying slopes into the model.<sup>11</sup>

---

<sup>11</sup>I've come to this obvious conclusion after too much time was spent on the issue, hence this section.