Conflict and Nation-Building: The Case of the 2014 Russia-Ukraine War

Serhii Abramenko, EIEF and LUISS
Valerio Sergio Castaldo, EIB
Vasily Korovkin, UPF and CEPR
Alexey Makarin, MIT and CEPR

January 6, 2024

Motivation

- Social cohesion is key for countries' security and development
- There is increased interest in nation-building processes and policies aimed at strengthening social cohesion (Rohner and Zhuravskaya, 2023)
  - Especially crucial for newly independent nations
    - 34 new countries recognized in the past 40 years
    - >100 new countries recognized in the past 80 years
- A classic theory states that nations coalesce (or even emerge) in response to an external threat (Guibernau, 2004)
- **What is the impact of armed conflict on national identity?**

Conceptual Framework

- **What is the impact of armed conflict on national identity?**
- Competing hypotheses: **Consolidation** vs. **Polarization**
- Reasons for **Consolidation**:
  - "Rally 'round the flag" effect (Mueller, 1970)
  - Reduced importance of domestic issues and corresponding (ethnic) cleavages
  - Shared animosity toward the aggressor
- Reasons for **Polarization**:
  - Negative income shocks may cause in-fighting for reduced resources
  - External conflict initiators may stoke divisions and aid separatist movements
  - External conflict initiators may be linked to a specific ethnolinguistic group in the country

Context: The 2014 Russia-Ukraine War

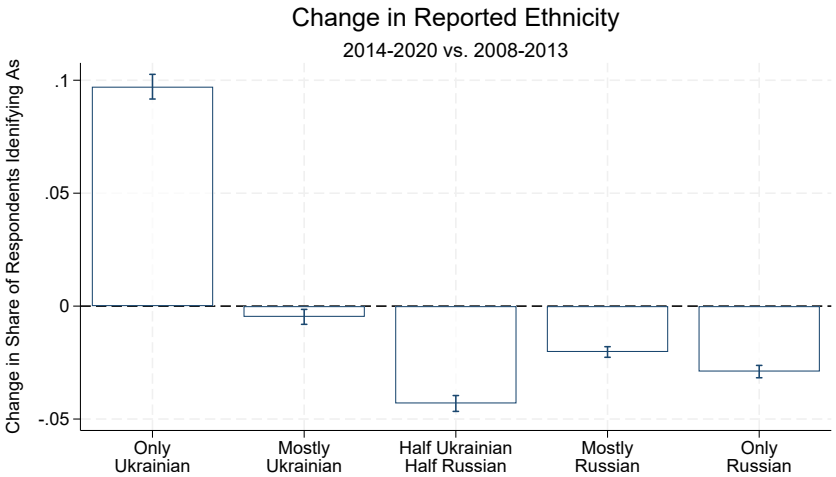
- In 2014, Russia annexed Crimea and started a war in Donbas, occupying part of it.
- Both **Consolidation** and **Polarization** effects were plausible ex ante.
- Substantial preexisting ethnolinguistic cleavages (KIIS, 2013)
  - 47% speak Ukrainian, 40% speak Russian
  - 40% are in favor of joining the EU, 35% are against it
    - Correlated with ethnicity
- Russia tried to exploit these identity differences by spreading several narratives
  - Crimea's breakaway and the Donbas war were due to pro-Russian separatists
  - Russian language and identity were threatened in Ukraine
  - The pro-European part of Ukraine had far-right views
  - Pro-European elites were trying to "sell" Ukraine to the West

Data

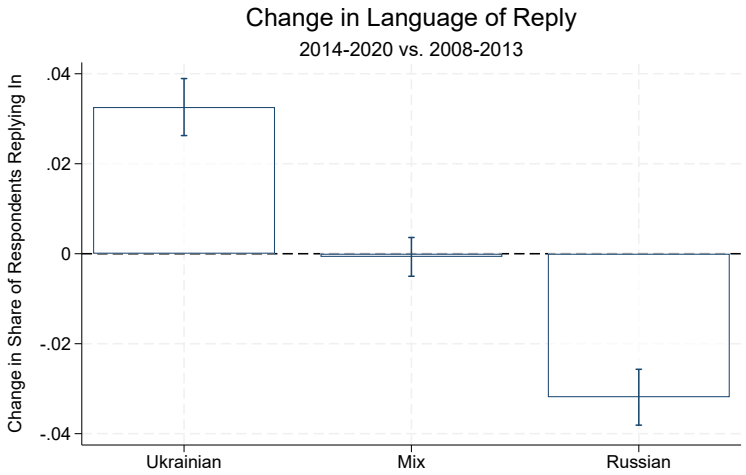
- Social media data
  - Self-reported language proficiency of VK users in Ukraine, September 2013 and March 2014
  - Language of posts on Twitter in Ukraine, 2012–2022
- OMNIBUS survey by the Kyiv International Institute of Sociology (KIIS), 2008–2020
  - Nationally representative repeated cross-section
  - 53 quarterly survey waves; more than 103,000 responses
  - Consistent questionnaire
    - Self-reported ethnicity: Likert scale (1–5) from fully Ukrainian to fully Russian
    - Language used during surveys: Ukrainian, Russian, or a mix
- 2001 Census: ethnicity and language in Ukraine

Change in Identity Measured via Surveys

Change in Reported Ethnicity in Ukraine, Before vs. After Annexation of Crimea



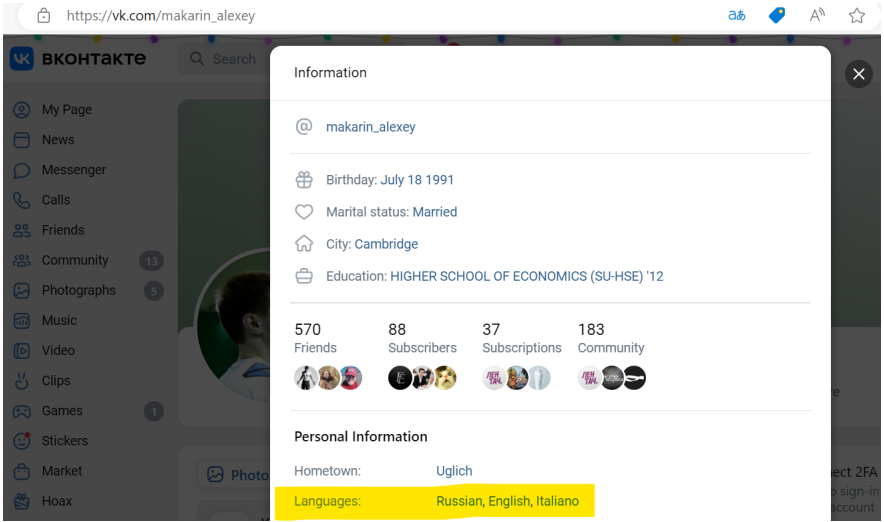
Change in Language Use in Ukraine, Before vs. After Annexation of Crimea



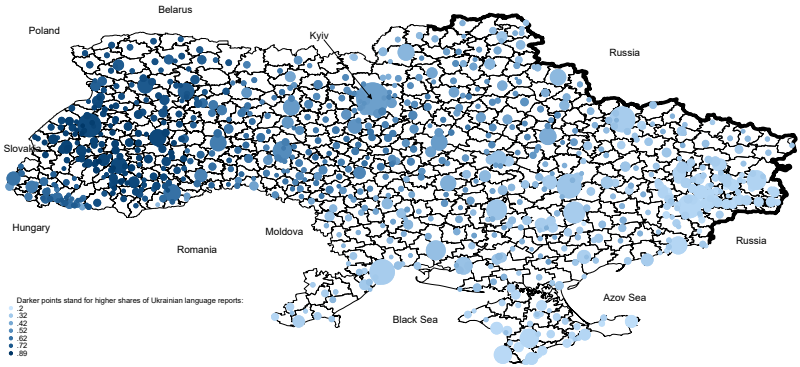
Social Media and Identity Barometer

- Social media data can be extremely useful for measuring shifts in identity
  - During conflicts or other disruptive events, surveys may not be feasible
  - Allows for granular analysis at the level of the hundreds of smaller settlements
- We apply two social media datasets to measuring shifts in Ukrainian identity after the 2014 conflict and validate them against the survey data presented earlier
  - VK: reported knowledge of specific languages
  - Twitter: actual language of posts

VK profile example

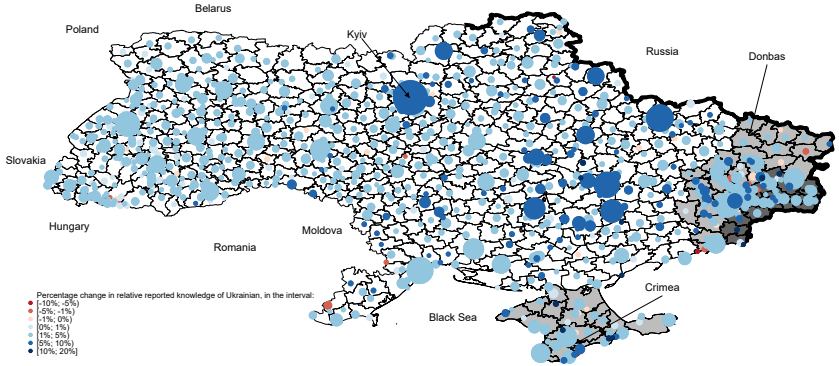


Share of VK Users Reporting Knowledge of Ukrainian, September 2013



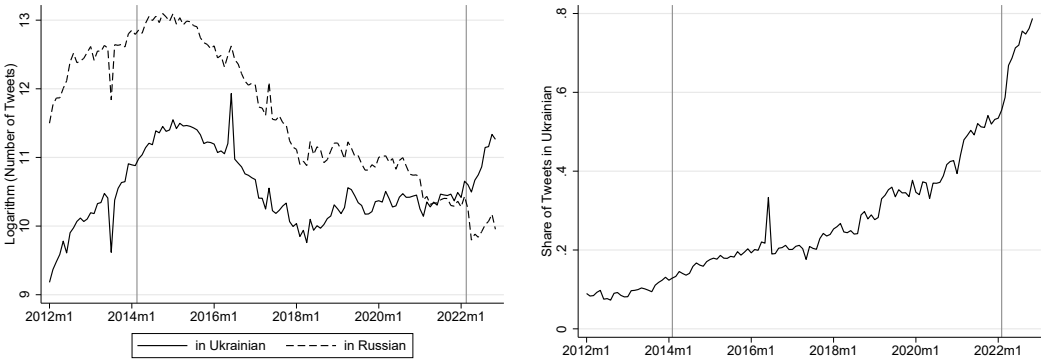
Notes: The figure maps the share of VK users reporting knowledge of the Ukrainian language. For a given population center i, it is calculated as $S_{i,t} = U_{i,t}/(U_{i,t} + R_{i,t})$ and where $U_{i,t}$ and $R_{i,t}$ are, respectively, the total numbers of VK users claiming knowledge of Ukrainian and Russian at time t. The size of each circle is weighted by the square root of the total number of VK users in September 2013.

Change in Share of VK Users Reporting Knowledge of Ukrainian, 2013–2014



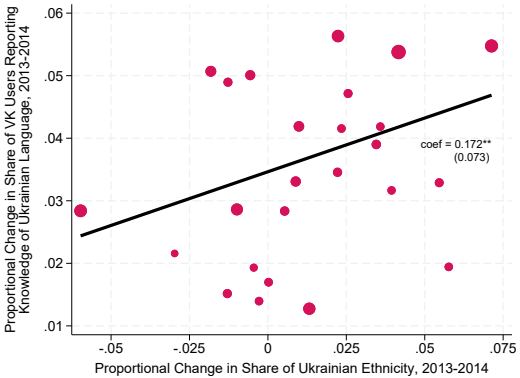
Notes: The figure maps the proportional change in the share of VK users reporting knowledge of the Ukrainian language. For a given population center i, it is calculated as $\Delta_i = 100 \times (S_{i,2014} - S_{i,2013})/S_{i,2013}$, where $S_{i,t} = U_{i,t}/(U_{i,t} + R_{i,t})$ and where $U_{i,t}$ and $R_{i,t}$ are, respectively, the total numbers of VK users claiming knowledge of Ukrainian and Russian at time t. The size of each circle is weighted by the square root of the total number of VK users in September 2013.

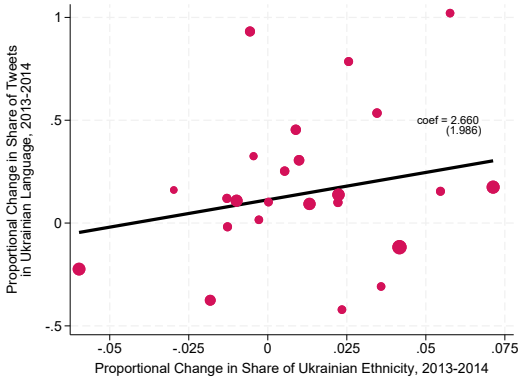
Change in Language of Twitter Posts in Ukraine, 2012–2022



Notes: The figure on the left displays the evolution of the total number of tweets in Russian and Ukrainian in the territory of Ukraine. The figure on the right displays the evolution of the share of tweets in Ukrainian as opposed to Russian in the territory of Ukraine. The vertical lines represent the months of two major acts of Russian aggression against Ukraine: the annexation of Crimea (February 2014) and the full-fledged invasion (February 2022).

Correlations Between Identity Shifts on Social Media and in Surveys





Notes: The figure on the left displays the province-level correlation between the proportional changes in the share of survey respondents claiming Ukrainian ethnicity (x-axis) and the share of VK users reporting knowledge of the Ukrainian language (y-axis). The figure on the right displays the province-level correlation between the proportional changes in the share of survey respondents claiming Ukrainian ethnicity (x-axis) and the share of Twitter posts written in Ukrainian as opposed to in Russian (y-axis).

Consolidation vs. Polarization

Empirical Strategy

- Before-after comparisons may suffer from secular trends.
  - Younger cohorts might be less likely to speak Russian or identify as ethnically Russian.
  - Generational shift could create similar patterns, even without any major political events.
- **Test:** Do respondents in "more Russian" areas of Ukraine move closer to Ukrainian identity (**Consolidation**) or away from it (**Polarization**) after the start of the war?

Empirical Strategy

- **Difference-in-Differences:** Do respondents in "more Russian" areas of Ukraine move closer to Ukrainian identity (**Consolidation**) or away from it (**Polarization**) after the start of the war?

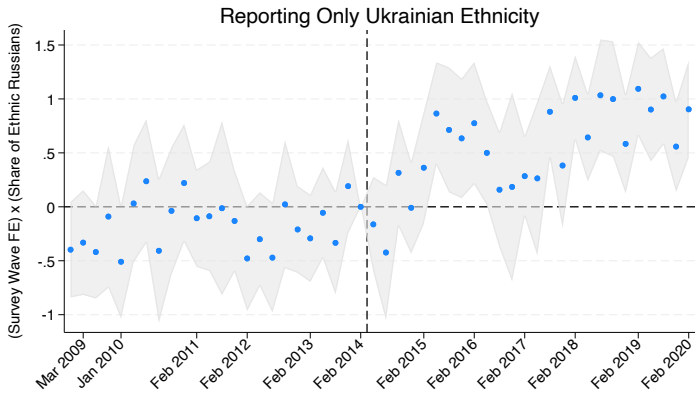
$$I_{irt} = \beta ER_{r,2001} \times PostFeb2014_t + X_{irt}\gamma + \alpha_r + \xi_t + u_{irt}$$

  - I_{irt} — a measure of identity of individual i from rayon r in period t
  - $ER_{r,2001}$ — the share of ethnic Russians in rayon r
  - $PostFeb2014_t$ — post-start-of-the-conflict indicator
  - α_r and ξ_t — rayon-level and time fixed effects
  - X_{irt} — individual i's and rayon r's characteristics, such as province-level GDP per capita and CPI, share of displaced individuals, rural indicator, and i's sex, age, and education

- **Identifying assumption:** Absent the conflict, respondent identity in Ukraine's different rayons would have evolved along parallel trends, conditional on common rayon and time shocks.

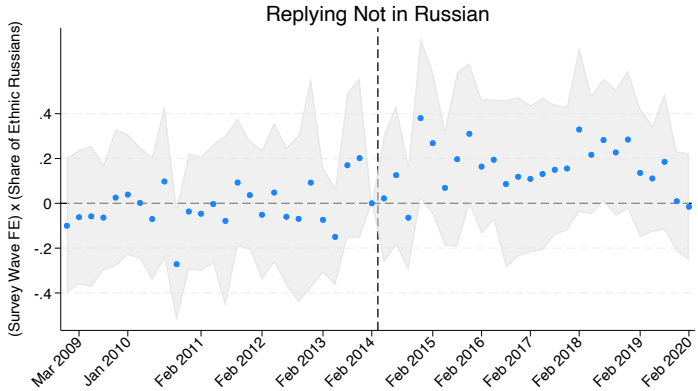
Results

Event Study: Self-Reported Ethnicity



- **Accelerated consolidation** after the start of the 2014 Russia-Ukraine conflict

Event Study: Language Used During Survey



- **Accelerated consolidation** after the start of the 2014 Russia-Ukraine conflict

Baseline Estimates

VARIABLES	(1) Self-identifying as only Ukrainian	(2) Self-identifying as only Ukrainian	(3) Not replying in Russian	(4) Not replying in Russian
Post Feb 2014 X	0.710***	0.558***	0.162***	0.138**
Share of Ethnic Russians	(0.0909)	(0.116)	(0.0452)	(0.0580)
Observations	87,751	87,272	95,666	95,117
R-squared	0.245	0.251	0.572	0.577
Controls	No	YES	No	YES
Rayon FE	YES	YES	YES	YES
Wave FE	YES	YES	YES	YES

Notes: The controls include rayon/city fixed effects, province-level real GDP per capita, and CPI relative to 2007, rayon-level ratio of internally displaced persons (IDP) to population, rurality of settlements, and individual characteristics, including sex, age, and education level. Outcomes include self-reporting "only Ukrainian" ethnicity on a five-point scale from "only Russian" to "only Ukrainian," and de facto replying in Russian (vs. in Ukrainian or a mix) during the surveys. Standard errors in parentheses are clustered at the rayon level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

- 0% $\rightarrow$ 25% ethnic Russians $\implies$ $\uparrow$ 13.8 pp likelihood of identifying as 'only Ukrainian' & $\uparrow$ 4 pp likelihood of not replying in Russian post-Feb 2014

Baseline Estimates: Likert Scale of Ethnicity

	(1)	(2)	(3)	(4)	(5)
	Respondent is self-identifying as:				
VARIABLES	Only Ukrainian	More Ukrainian	Half Ukrainian	More Russian	Only Russian
Post Feb 2014 X	0.558***	0.00934	-0.275***	-0.0862**	-0.207***
Share of Ethnic Russians	(0.116)	(0.0450)	(0.0938)	(0.0353)	(0.0371)
Observations	87,272	87,272	87,272	87,272	87,272
R-squared	0.251	0.058	0.118	0.060	0.081
Controls	YES	YES	YES	YES	YES
Rayon FE	YES	YES	YES	YES	YES
Wave FE	YES	YES	YES	YES	YES

Notes: The controls include rayon/city fixed effects, province-level real GDP per capita and CPI relative to 2007, rayon-level ratio of internally displaced persons (IDP) to population, rurality of settlements, and individual characteristics, including sex, age, and education level. Outcomes are indicators for self-reporting ethnicity options on a five-point scale from "only Russian" to "only Ukrainian." Standard errors in parentheses are clustered at the rayon level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Potential Concerns

- Migration
  - More than 1 million internally displaced persons (IDPs) due to conflict.
  - Pro-Ukraine individuals could have been more likely to migrate within Ukraine (vs. to Russia).
  - **Response:** We control for the number of IDPs per population.
- Selection
  - Regardless of migration, perhaps only pro-Ukraine individuals agreed to interviews.
  - **Response:** (i) We control for individual-level characteristics; (ii) results are similar in a cohort-based panel; (iii) no effect on an indicator of 'no response' to the question of ethnicity.
- Ceiling effect
  - There was naturally less space for any increase in Ukrainian ethnicity in the West of Ukraine.
  - **Response:** (i) Our strategy would still be able to capture polarization if it was happening; (ii) results hold after omitting areas in Western Ukraine or areas within the top quartile of Ukrainian ethnicity; (iii) results hold using a Tobit regression.

Conclusion

- What is the role of conflict in nation-building?
- Ex ante, many forces predict either consolidation or polarization.
- We study this question in the context of the 2014 Russia-Ukraine conflict.
- We show how one could apply social media to measure shifts in identity in such contexts
- **Result**: Russian aggression has led to **accelerated consolidation** of Ukrainian identity—both in ethnicity and language use, both in survey and social media data.
- Implication: External conflict aimed at destabilizing a country led to its consolidation and stronger resistance in a future war