# Estimating regional economic activity in Ukraine during an invasion.

Mihnea Constantinescu[1] Kalle Kappner[2] Niko Szumilo[3]

[1]National Bank of Ukraine
[2]Humboldt-Universität zu Berlin
[3]UCL - University College London

NBU results presentation - May 2022

# This project

**Estimating current economic activity (GDP) in Ukraine by region.**

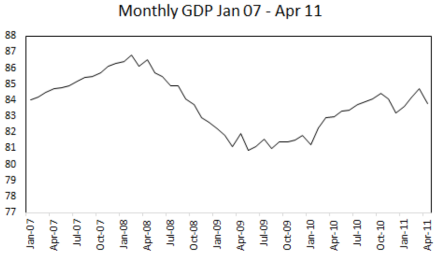
Traditionally measured annually, more recently quarterly (UK and Canada measure it monthly).

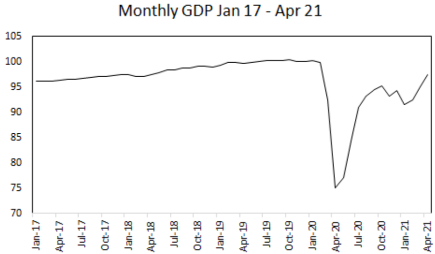
There are challenges in using/interpreting high frequency GDP as it seems to be very volatile.

Useful in a crisis:
- strong signals cut through the normal noise
- decisions need to be made faster in a crisis

# High frequency volatility.

- There is a lot of seasonal variation - seasons of the year, days of the week, national holidays, number of days in a month.
- Short term changes are not necessarily indicative of term trends.
- Strong signals are still visible immediately.




Index=100 in Dec 2019 Data for the UK from ONS

# Regional vs National.

National effects are aggregated from heterogeneous regional economies.

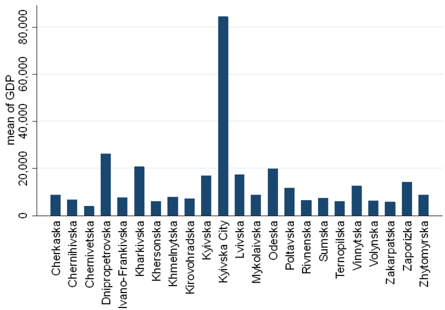
National policy affects some regions more than others → aggregate effects depend on regional effects.

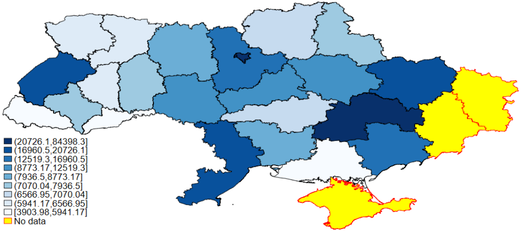
If a crisis is regional and an economy is unequal, regional breakdown is important.

A war is inherently a regional crisis where the most affected regions are not random.

# Regional heterogeneity.

- Kyiv City is around 24% of the whole economy.
- It is roughly the same as Donetsk, Zaporizhzhya, Luhansk, Mykolayiv, Sumy, Kharkiv, Kherson and Chernihiv **combined**.
- Kyiv City and Kyiv oblast are around 30%.
- Focusing on GDP will emphasize Kyiv but other policy objectives (employment, equality, growth) may suggest different focus.



# GDP with Big Data.

We focus on current GDP per region (AKA Gross Regional Product) to understand the current crisis.

The choice is:

1. Measure - popular in developed countries with strong institutions but expensive and slow. Also, often not possible during a war.
2. Model - out of sample prediction based on structure of a model:
   - 2.1 Structural theory prediction for a scenario (IMF, WB etc.)
   - 2.2 Time series prediction based on past data
   - 2.3 **Prediction using correlated data**

All indicators of GDP are biased.
Using Big Data - back-of-envelope economic performance, fast, effective and well-documented.

# Our approach.

Ukraine historic data challenges:
- Only low frequency of observations (annual)
- Only short time series (6-7 time periods)
- Conventional current data (payments, transactions, employment etc.) is unavailable

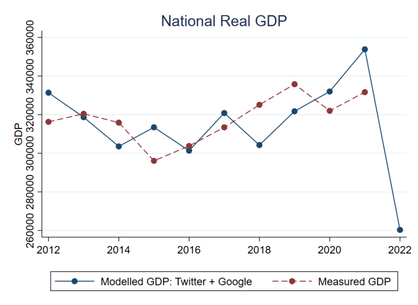
Solutions:
- Model annual GDP (for shorter periods assume that the period's level persists for a year and report the annual GDP)
- Use panel methods (current literature relies on time series)
- Use big data (Nightlights, Twitter and Google Trends)

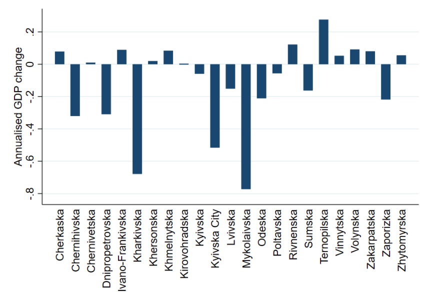
No new indicators - using we know works and combine it in a new way.

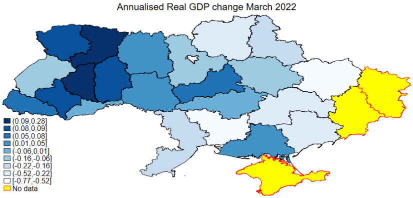
# Assumptions.

1. We ignore regions occupied by Russians before 2022: Donetsk, Luhansk and Crimea (we have the data and can look at them later).
2. We model total economic activity within areas defined by Ukrainian regions: this includes areas that are occupied or contested.
3. We are not sure what goes on in areas were there is fighting but model them nonetheless.
4. We model everything in 2004's Hryvnia.

# Results overview.




- If 2022 is the same as March 2022, economy shrinks by 26% (estimates range between 18% and 50%).
- Directly affected regions loose more than others.

# Finding data.

The GDP nowcasting literature focuses on institutional high-frequency, high-quality datasets of early economic indicators.

- Electronic payments data, sector-level indices, expert surveys
- Flexible models borrowing structure from GDP's definition
- Large scope for evaluation and continuous updating

**Mostly not an option for us.**

Instead, we focus on a set of feasible GDP correlates emphasized by recent research:

- Twitter (Indaco 2020)
- Google Trends (Woloszko 2021)
- Nightlights (Henderson et al. 2011)

# Twitter.

- Idanco (2020, RSUE): (Image) tweets are indicative of economic activity and correlate with GDP $\rightarrow$ conspicuous consumption.
- Geo-location is a user choice. Only a fraction of Tweets are geo-located, and that fraction dropped after 2019.
- We count all geo-located (Image) tweets around Ukraine's largest 1000+ settlements since 2010; then aggregate to regional level.
- Documented issues:
  1. No established micro-foundations.
  2. Twitter is not too popular in UKR.
  3. The way people use Twitter changes over time.

# Google Trends.

- Search trends → more highly correlated to consumer and firm behaviour than surveys.
- Google offers an index of the *share* of searches for a particular phrase/topic over time for an area.
- Worked very well for Covid 19 around the world (Woloszko 2020, OECD).
- Has documented issues:
    1. No consensus on which search phrases to include.
    2. Changing user base induces a non-linear downward trend in indices.
    3. Large N of topics, potentially few observations.

# Google Trends.

We use 30 topics (defined by Google) guided by the literature including:

**Firm focus:**
Investment
MBA degree
Economy
Interest rate
Inflation
Recruitment
Computer security
Bankruptcy
Export

**Consumer focus:**
Mortgage loan
Holiday
BMW
Calving Klein
Emigration
Apple Inc
Fitness
Fashion
Savings

# Nightlights.

- Henderson et Al (2012, AER) document a stable, log-linear relationship between nightlight intensity (NLI) and GDP.
- NLI useful to estimate irregular activity, deflate unreliable GDP estimates and trace effects of natural disasters and conflict.
- Many traditional technical issues solved in NASA's Black Marble suite (VIIRS data).
- Some issues remain:
  1. Disregards in-home light sources; not a good proxy for some sectors (like agriculture).
  2. Monthly data is more volatile and error-prone.
  3. NLI-to-GPD elasticity may not be stable at the sub-national level.

# Nightlights.

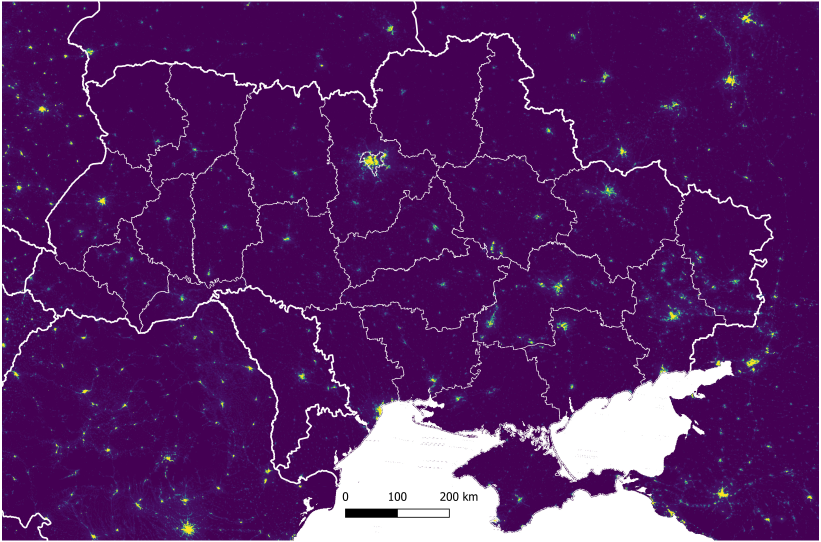



Figure: Nighttime luminosity in 2012 (NASA Black Marble)

# Nightlights.

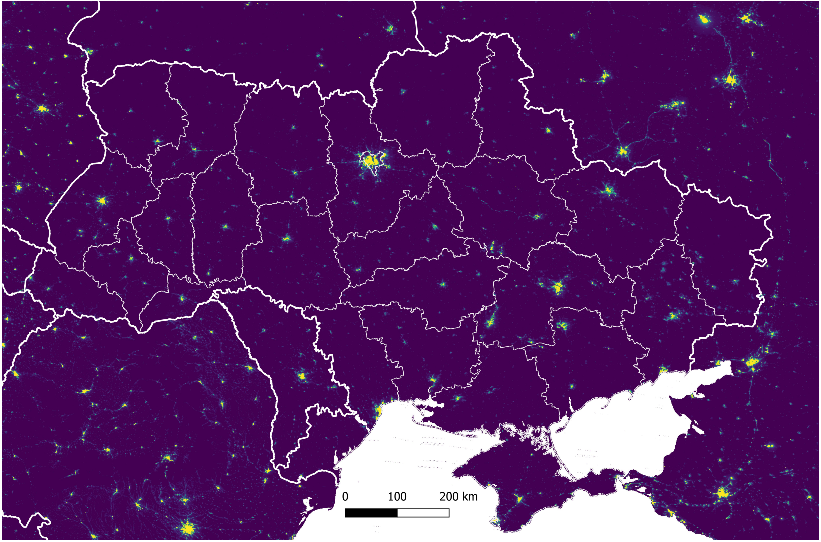



Figure: Nighttime luminosity in 2021 (NASA Black Marble)

# Nightlights.

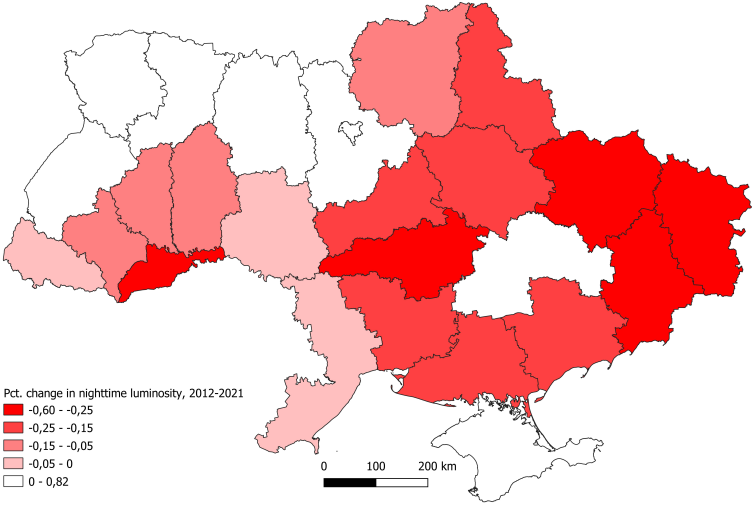



Figure: Regional change in nighttime luminosity, 2012 to 2021 (NASA Black Marble)

# Nightlights.

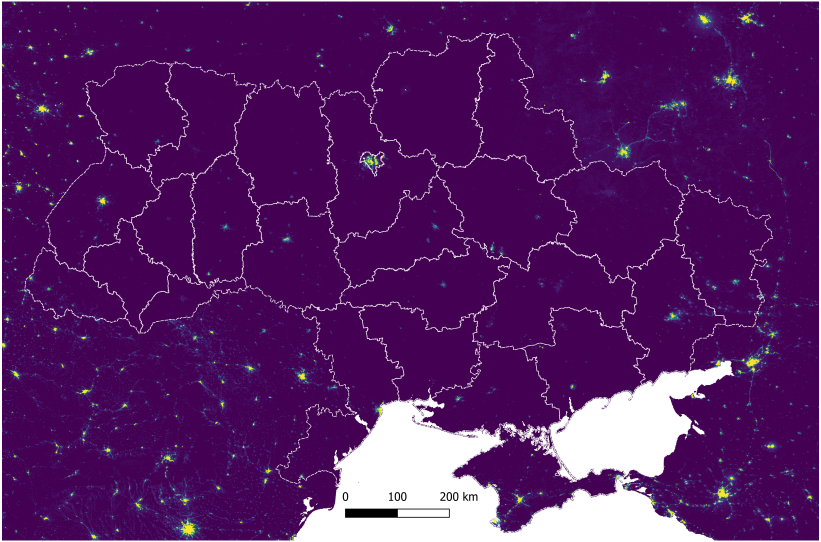



Figure: Nighttime luminosity in March 2022 (NASA Black Marble)

# Methods.

Key steps:

1. Get historic  current data at regional level: GDP, Twitter, Google Trends and Nightlights.
2. Model annual regional GDP as a function of Twitter, Nightlights and Google Trends.
3. Use the model and current data to generate out-of-sample predictions.

The key is getting the model right, so that it can be assumed to hold even during a war.
Similar approaches worked during the pandemic and natural disasters.

# Index of Nightlights.

We follow the literature → use the log of nightlights intensity per squared kilometer (NLI) per region as our main indicator.

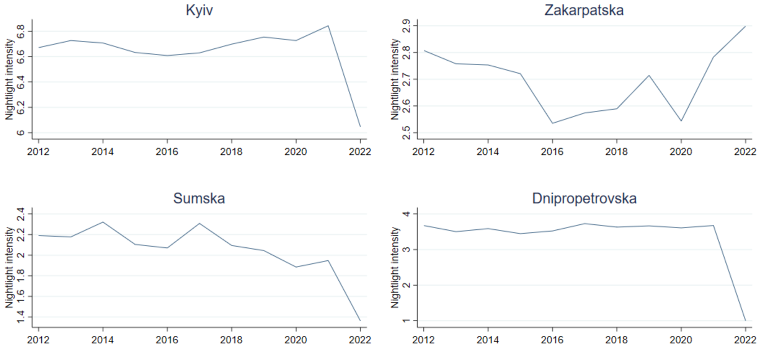
$$ln(GDP) = \alpha + \beta ln(NLI) + \epsilon \tag{1}$$

| | (1) lnGDP | (2) lnGDP |
|---|---|---|
| lnNLI | 0.0343**** | 0.330*** |
| | (2.14) | (11.97) |
| *Region FE* | Y | N |
| *Sample* | '11-'20 | '15-'20 |
| $N$ | 213 | 138 |
| $R^2$ | 0.995 | 0.7 |

$t$ statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

# Index of Nightlights.

Big changes in 2022.



# Index of Google Trends.

Fat data problem → 6-7 time periods but 30 variables.
To get an index we use a simple random effects panel.

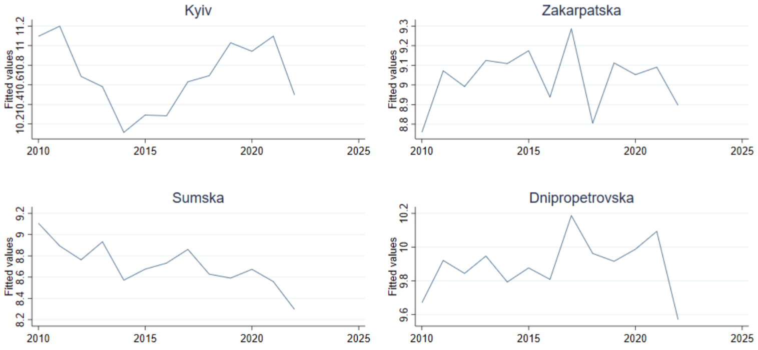
$$ln(GDP) = \alpha + \sum_{n=1}^{n=30} \beta_n Topic_n Index + \epsilon \tag{2}$$

| | (1) | (2) |
|---|---|---|
| | ln_gdp | ln_gdp |
| Investment_index | 0.00235 | 0.0316*** |
| | (0.80) | (6.25) |
| Economy_index | -0.00208 | -0.0256*** |
| | (-0.81) | (-5.77) |
| Bankruptcy_index | 0.00141 | 0.00678 |
| | (0.63) | (1.41) |
| apple_index | 0.00897*** | 0.00935** |
| | (4.09) | (2.84) |
| MBA_degree_index | 0.00390* | 0.0115** |
| | (2.36) | (3.08) |
| *Region FE* | Y | N |
| $N$ | 265 | 265 |
| $R^2$ | 0.996 | 0.765 |

$t$ statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

# Index of Google Trends.

More volatility than nightlights and a smaller change in 2022.



# Tweets and regional GDP.

Not very useful after controlling for time/year FEs.
Useful in a random panel and in a cross section.

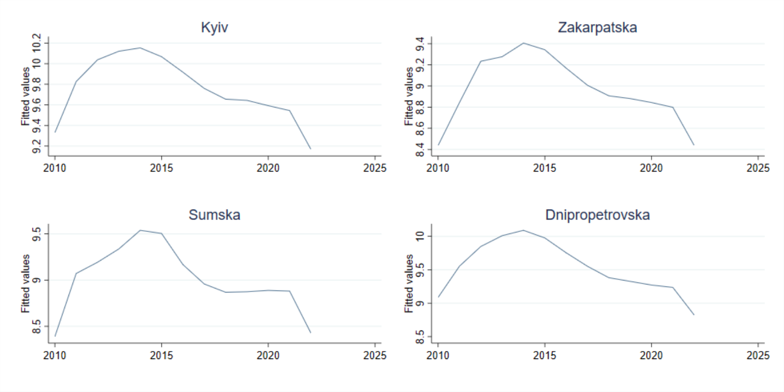
$$ln(GDP) = \alpha + \beta ln(Tweets) + \epsilon \tag{3}$$

| | (1) | (2) | (3) |
|---|---|---|---|
| | ln_gdp | ln_gdp | ln_gdp |
| ln(Tweets) | 0.00377 | 0.157*** | 0.227*** |
| | (1.79) | (9.38) | (8.02) |
| *Region FE* | Y | N | N |
| *Sample* | '11-'20 | '11-'20 | '20 |
| $N$ | 265 | 265 | 25 |
| $R^2$ | 0.993 | 0.251 | 0.46 |

$t$ statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

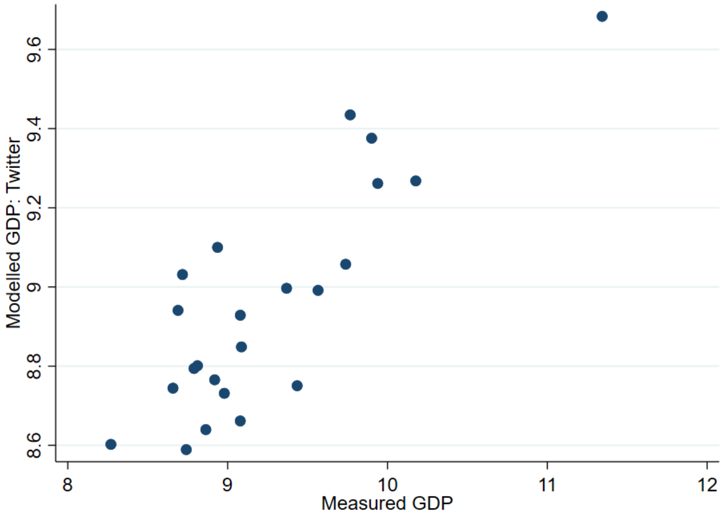
# Twitter Trends.

Clear national pattern but also local effects.



# Twitter Trends.

Good performance in a cross section.
Below data is for 2020.



# Three indicators.

Combining indicators helps is they have **uncorrelated prediction errors**.

It looks like Google is a good complement to Twitter and Nightlights. Twitter and Nightlights seem to be correlated.

Correlation of GDP prediction errors (post 2014)

| | Tw error | GTI Error |
|---|---|---|
| Tw error | 1 | – |
| GTI Error | 0.24 | 1 |
| NL error | 0.51 | -0.03 |

# Favourite specifications.

No.1 - Intuitive but naive specification:

$$\ln(\mathrm{GDP})_{Rt} = \alpha + \beta_1 ln(Tw) + \beta_R ln(GTI) \times R + \beta_2 ln(NLI) + \beta_3 NLI + \epsilon_{RT}$$

No.2 - naive with region fixed effects:

$$\ln(\mathrm{GDP})_{Rt} = \alpha_R + \beta_1 ln(Tw) + \beta_R ln(GTI) \times R + \beta_2 ln(NLI) + \beta_3 NLI + \epsilon_{RT}$$

No.3 - No nightlights but region-specific $\beta$:

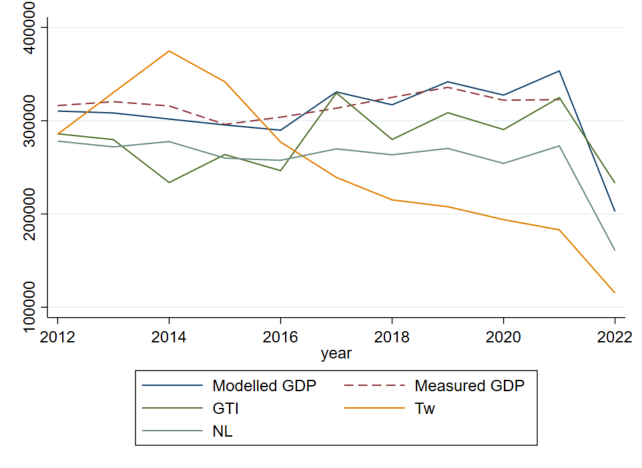
$$\ln(\mathrm{GDP})_{Rt} = \alpha + \beta_R ln(Tw) \times R + \gamma_R ln(GTI) \times R + \epsilon_{RT}$$

Where $R$ denotes a region and $t$ time.

In our estimation we leave 2021 out to test an out-of-sample prediction.

# Results national.

Measured and modelled GDP over time.



# Results national.

Favourite specifications: growth in 2022 from -50% to -22%.

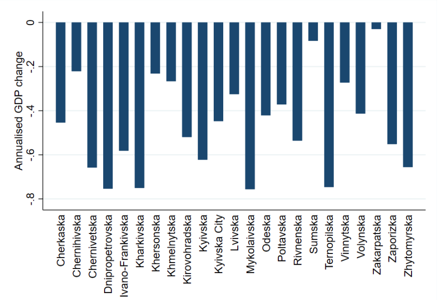
| | (1) lng | (2) lng | (3) lng |
|---|---|---|---|
| lnNLI | 0.646*** | -0.09*** | |
| | (3.67) | (3.08) | |
| NLI | 0.00000973 | 0.0007* | |
| | (1.25) | (1.66) | |
| lnTw | 0.0186 | -0.015*** | [...] |
| | (0.59) | (6.86) | [...] |
| lnGTI×*region* | [...] | [...] | [...] |
| | [...] | [...] | [...] |
| *Spec* | GTI+NL+Tw | Tw+NL+GTI | GTI+Tw |
| *Region FE* | No | Yes | No |
| ′21 | 0.14 | 0.024 | 0.065 |
| ′22 | -0.50 | -0.22 | -0.26 |
| $N$ | 213 | 213 | 213 |
| $R^2$ | 0.976 | 0.997 | 0.993 |

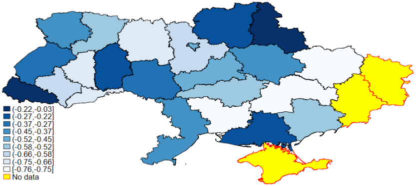
$t$ statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

# Results No.1

National growth 2021: 14%
National growth based on March 2022: -50%

Key features:
pretty dramatic,
occupied areas seem lightly affected,
west strongly affected.

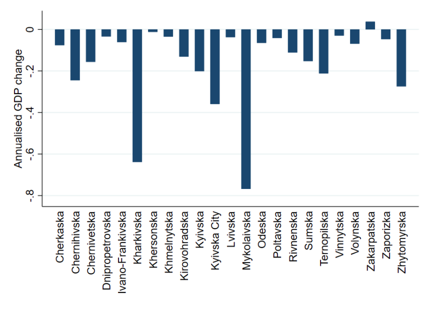


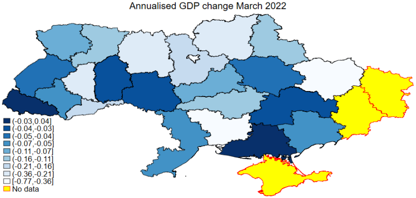



# Results No.2

National growth 2021: 2.4%
National growth based on March 2022: -22%

Key features:
optimistic in the West and around Kyiv.

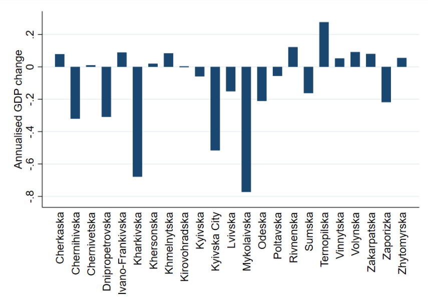


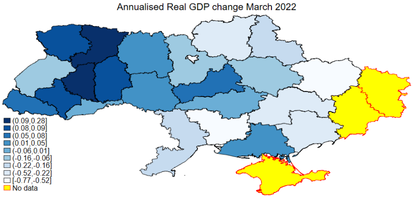



# Results No.3

National growth 2021: 6.5%
National growth based on March 2022: -26%

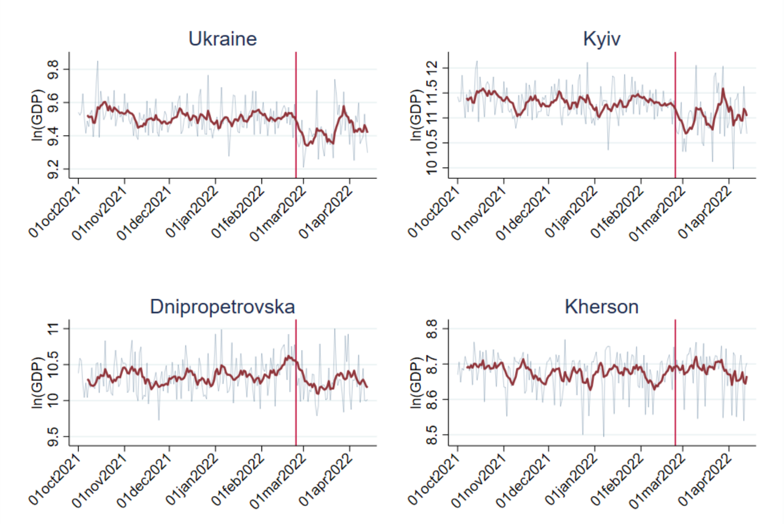
Key features:
optimistic in the West.



# Daily GTI and war

Daily GDP predicted by GTI suggests a drop at the start of the war to 70% of the pre-war mean but also a recovery to around 80% of that value in late March.

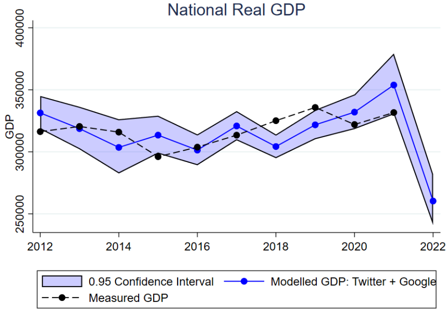


# Confidence.

**Data based:**
Favourite: -26% (CI: -32%;-20%)
Mean of 3 models: -33%

**Model based:**
IMF: -35% (plausible)
EBRD: -30% (plausible)
World Bank: -45% (seems unlikely)

1. Assumptions and methods matter.
2. Transparency is important.
3. Definitions are not obvious.



# Conclusions.

- GDP indicators offer intuitive results.
- It is unclear how reliable the results are as they cannot be tested.

- At the start of the war economic activity reduced by 20-40%.
- The first month of the war was a dramatic fall but activity increased afterwards.
- Winning the battle of Kyiv is a major economic victory.
- The World Bank's forecast of a 45% drop in GDP due to the war seems unlikely.

# Thank you.

Niko Szumilo - UCL - n.szumilo@ucl.ac.uk

Kalle Kappner - Humboldt-Universität - kallekappner@googlemail.com