

This registration is a frozen, non-editable version of [this project \(/js4md/\)](#)

This registration is currently embargoed. It will remain private until its embargo end date, Thursday, Apr 25, 2019.

## Register

### Study Information

Title

Authors

Description

Hypotheses

### Design Plan

Study type

Blinding

Blinding

Study design

Randomization

### Sampling Plan

Existing Data

Explanation

Data collection procedures

Sample size

Sample size rationale

Stopping rule

### Variables

Manipulated

Measured

Indices

### Analysis Plan

Statistical models

Transformations

Inference criteria

Data exclusion

Missing data

Exploratory analysis

### Other

Other

# Study Information

## Title

*Provide the working title of your study. It may be the same title that you submit for publication of your final manuscript, but it is not a requirement.*

Can land intensification and abandonment in Latvia be linked to key socio-economic events?

## Authors

Isabelle Rich

## Description

*Please give a brief description of your study, including some background, the purpose of the study, or broad research questions.*

This study aims to investigate the importance of socio-economic events as drivers of land-use change in Latvia through the use of satellite imagery. Although the importance of socioeconomic events on land-use change is acknowledged, it remains unclear whether a recognisable, country-scale signature is left on the landscape. Using satellite imagery, pixel-scale analysis can be completed to determine specific land cover transitions over time, potentially unveiling a link between socio-economic events and land-use change.

Specific questions:

Q1: Is there a clear link between key socio-economic events and land-use change in Latvia?

Q2: Is the strength and direction of land-use change different among extensive, intensive and abandoned land-use types?

Q3: Is there a time lag between socio-economic events and the occurrence of land use change? Does this differ between land-use type?

## Hypotheses

*List specific, concise, and testable hypotheses. Please state if the hypotheses are directional or non-directional. If directional, state the direction. A predicted effect is also appropriate here. If a specific interaction or moderation is important to your research, you can list that as a separate hypothesis.*

H1: There is an observable, uniform link between both the Soviet Union collapse and EU accession and land-use change in Latvia at a country-scale.

H2: The strength and direction of land-use change are different for extensive, intensive and abandoned land types at pixel-scale, with a slight decrease in intensive land and a sharp increase in abandoned and extensive land following Soviet Union collapse. For the EU, I hypothesise a slight decrease in abandoned and extensive land-use types, but a sharp increase in intensive land.

H3: Intensive and abandoned land-use change is observed directly following the Soviet Union collapse at country-scale. However, there is a time lag when observing shifts towards intensive land-use change following the Soviet Union collapse. There is a time lag on when land-use change is observed at country-scale following Latvia joining the EU for abandoned land, but not for shifts towards intensive and extensive land-use types.

## Design Plan

### Study type

*Please check one of the following statements*

Other

## Blinding

*Blinding describes who is aware of the experimental manipulations within a study. Mark all that apply.*

No blinding is involved in this study.

Is there any additional blinding in this study?

*Blinding (Other)*

N/A

## Study design

*Describe your study design. Examples include two-group, factorial, randomized block, and repeated measures. Is it a between (unpaired), within-subject (paired), or mixed design? Describe any counterbalancing required. Typical study designs for observation studies include cohort, cross sectional, and case-control studies.*

I will be creating a classification of three agricultural land-use types: abandoned, intensive and extensive. All satellite data and land-use information are already freely available online on datasets such as LUCAS, MODIS, CORINE and LUH. By using data already collected and the computer program Google Earth Engine, I do not have to miss out on any points and will be "sampling" every pixel.

no file selected

## Randomization

*If you are doing a randomized study, how will you randomize, and at what level?*

N/A

## Sampling Plan

### Existing Data

*Preregistration is designed to make clear the distinction between confirmatory tests, specified prior to seeing the data, and exploratory analyses conducted after observing the data. Therefore, creating a research plan in which existing data will be used presents unique challenges. Please select the description that best describes your situation. Please see <https://cos.io/prereg> for more information.*

Registration prior to analysis of the data

### Explanation of existing data

*If you indicate that you will be using some data that already exist in this study, please describe the steps you have taken to assure that you are unaware of any patterns or summary statistics in the data. This may include an explanation of how access to the data has been limited, who has observed the data, or how you have avoided observing any analysis of the specific data you will use in your study.*

I have viewed satellite and land-use data on a country-level, but I have not looked at pixel scale information prior to conducting my analysis.

### Data collection procedures

*Please describe the process by which you will collect your data. If you are using human subjects, this should include the population from which you obtain subjects, recruitment efforts, payment for participation, how subjects will be selected for eligibility from the initial pool (e.g. inclusion and exclusion rules), and your study timeline. For studies that don't include human subjects, include information about how you will collect samples, duration of data gathering efforts, source or location of samples, or batch numbers you will use.*

To examine the relationship between land-use change and socio-economic events, I will use the Google Earth Engine (GEE), which is an online global spatial analysis platform. Specifically, I will be creating a classification of land-use types in Latvia to obtain land-use data, rather than solely land cover information. Using satellite data (e.g. Landsat), I will classify the area of interest (Latvia) into three categories: extensive, intensive and abandoned land. To classify the data, I will create a new layer for each land-use type with known points of each land-use type from historical records (training points). Layers then can be merged, converting each class into a value, or band. Next, I will extract the reflectance of each band and train the classifier on the known points of each land-use type (training data) and its reflectance. I will use a random forest supervised learning algorithm, which is an assemblage of decision trees. I will use this new, trained classifier to classify the rest of the imagery for each study year to uncover land-use types. I will be using data between 1978 to present time to allow for roughly the same time frame before and after each event studied. The accuracy of the classification will be assessed using a confusion matrix, as well as by using validation data. Validation data will be used as new testing data to assess the error. Following this analysis, I will obtain the land-use type for each pixel in Latvia, allowing me to compute, gain, loss and change of each land-use type at pixel-scale. Maps will also be produced to visually assess change. Exact datasets used will depend on the availability and time period of data.

no file selected

Sample size

*Describe the sample size of your study. How many units will be analyzed in the study? This could be the number of people, birds, classrooms, plots, interactions, or countries included. If the units are not individuals, then describe the size requirements for each unit. If you are using a clustered or multilevel design, how many units are you collecting at each level of the analysis?*

I will collect data on each pixel in Latvia over 40 years. With a pixel size of 30 x 30 metres (Landsat), roughly 70 million pixels will be analysed for each of 40 years.

## Sample size rationale

*This could include a power analysis or an arbitrary constraint such as time, money, or personnel.*

N/A - sampling on a computer allows infinite sampling.

## Stopping rule

*If your data collection procedures do not give you full control over your exact sample size, specify how you will decide when to terminate your data collection.*

N/A

## Variables

### Manipulated variables

*Describe all variables you plan to manipulate and the levels or treatment arms of each variable. This is not applicable to any observational study.*



N/A

no file selected

## Measured variables

*Describe each variable that you will measure. This will include outcome measures, as well as any predictors or covariates that you will measure. You do not need to include any variables that you plan on collecting if they are not going to be included in the confirmatory analyses of this study.*

Year: explanatory

Area of land-use type: dependent variable; variable of interest - sum of all pixels of each of the three land use types

no file selected

## Indices

*If any measurements are going to be combined into an index (or even a mean), what measures will you use and how will they be combined? Include either a formula or a precise description of your method. If you are using a more complicated statistical method to combine measures (e.g. a factor analysis), you can note that here but describe the exact method in the analysis plan section.*

Sum(area) of each land use type

no file selected

## Analysis Plan

### Statistical models

*What statistical model will you use to test each hypothesis? Please include the type of model (e.g. ANOVA, multiple regression, SEM, etc) and the specification of the model (this includes each variable that will be included as predictors, outcomes, or covariates). Please specify any interactions, subgroup analyses, pairwise or complex contrasts, or follow-up tests from omnibus tests. If you plan on using any positive controls, negative controls, or manipulation checks you may mention that here. Remember that any test not included here must be noted as an exploratory test in your final article.*

I am to do Bayesian statistics to analyse the impact that Year has on the area of each land-use type. However, if this becomes unattainable I will use mixed-effects models. Year will also be a random factor to account for the fact that pixels in one year are more likely to be similar to each other than other years.

no file selected

### Transformations

*If you plan on transforming, centering, recoding the data, or will require a coding scheme for categorical variables, please describe that process.*

N/A

### Inference criteria

*What criteria will you use to make inferences? Please describe the information you'll use (e.g. specify the p-values, Bayes factors, specific model fit indices), as well as cut-off criterion, where appropriate. Will you be using one or two tailed tests for each of your analyses? If you are comparing multiple conditions or testing multiple hypotheses, will you account for this?*

For priors, I will set the area of each land-use type to be between 0 and the total area of Latvia. I will also set a normal distribution prior.

## Data exclusion

*How will you determine which data points or samples if any to exclude from your analyses? How will outliers be handled? Will you use any awareness check?*

I will map my data to make sure that all points are on land and exclude any ones that are not.  
I will use known points of each land-use type to train my classifier and I will check the accuracy of my classifier.

## Missing data

*How will you deal with incomplete or missing data?*

This shouldn't be an issue, as I am creating my own classification. If satellite data consistently available across all years, I will use multiple sources.

## Exploratory analysis

*If you plan to explore your data set to look for unexpected differences or relationships, you may describe those tests here. An exploratory test is any test where a prediction is not made up front, or there are multiple possible tests that you are going to use. A statistically significant finding in an exploratory test is a great way to form a new confirmatory hypothesis, which could be registered at a later time.*

N/A

## Other

### Other

*If there is any additional information that you feel needs to be included in your preregistration, please enter it here. Literature cited, disclosures of any related work such as replications or work that uses the same data, or other context that will be helpful for future readers would be appropriate here.*

N/A

Copyright © 2011-2019 Center for Open Science (<https://cos.io>) | Terms of Use ([https://github.com/CenterForOpenScience/centerforopenscience.org/blob/master/TERMS\\_OF\\_USE.md](https://github.com/CenterForOpenScience/centerforopenscience.org/blob/master/TERMS_OF_USE.md)) | Privacy Policy ([https://github.com/CenterForOpenScience/centerforopenscience.org/blob/master/PRIVACY\\_POLICY.md](https://github.com/CenterForOpenScience/centerforopenscience.org/blob/master/PRIVACY_POLICY.md)) | Status (<https://status.cos.io/>) | API (<https://developer.osf.io/>) TOP Guidelines (<http://cos.io/top/>) | Reproducibility Project: Psychology (<https://osf.io/ezcuj/wiki/home/>) | Reproducibility Project: Cancer Biology (<https://osf.io/e81xl/wiki/home/>)



(<http://twitter.com/OSFramework>)



(<https://www.facebook.com/CenterForOpenScience/>)



(<https://groups.google.com/forum/#!forum/openscienceframework>)



(<https://www.github.com/centerforopenscience>)