

Did High Speed Rails Promote  
Local Economic Activities?:  
Application of Night Time Satellite Data

Kanghyock Koh <sup>1</sup> Hyunjoo Yang <sup>2</sup>

<sup>1</sup>Korea University

<sup>2</sup>Incheon National University

11/15/2019

# Motivation

- ▶ Cost of construction of Korean High Speed Rail (KTX) about \$ 27 billion USD  
(Yonhap News 2016.9.21.)
- ▶ Few studies evaluating the overall impacts of KTX on local economies
- ▶ Difficulty of doing empirical analysis at the sub-national level
  - ▶ Ideal data are regional level balanced panel data
  - ▶ Lack of county- or township-level GDP data
  - ▶ Lack of high frequency data
  - ▶ Frequent administrative unit changes

# Contribution of this paper

## **Contribution of this paper**

- ▶ Novel data set of 20-year balanced panel with a local economic activity measure
- ▶ Empirical analysis of the effects of KTX on local economic activity

# Mixed evidence on the effects of High Speed Rail (HSR)

## Positive impacts of HSR

- ▶ Better transportation infrastructure improves economic activities
- ▶ Housing boom and new industrial parks
- ▶ (Sands, 1993; Banister and Berechman, 2003; Kwon 2014)
- ▶ Increase in firm productivity (Baek and Park 2015)
- ▶ New entry of firms, employment, productivity (Kim and Ahn 2009, 2013; Ahn and Kim 2009)

# Effects of High Speed Rail on Local Economic Activity

## Negative impacts of HSR

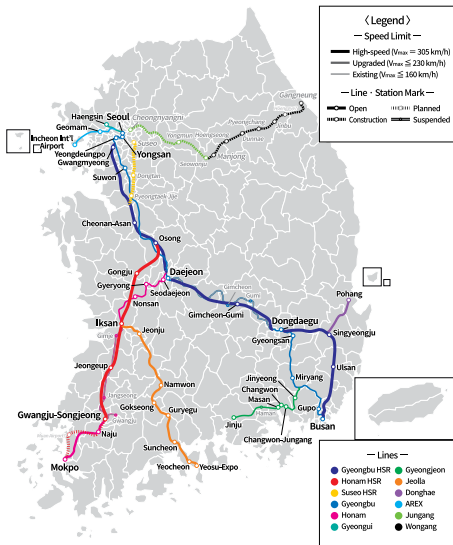
- ▶ Economic activities being drained away from larger cities (Whitelegg and Holzapfel, 1993; Thompson, 1995; Vickerman, 1997)
- ▶ Cancer patients switch to hospitals in larger cities (Kim et al., 2010)

→ Empirical investigation is needed for assessing overall effects

## Preview of Findings

- ▶ Introduction of the HSR system increases local economic activities of rural townships near HSR stations by 10% in 10 years

# The KTX Network



# Data

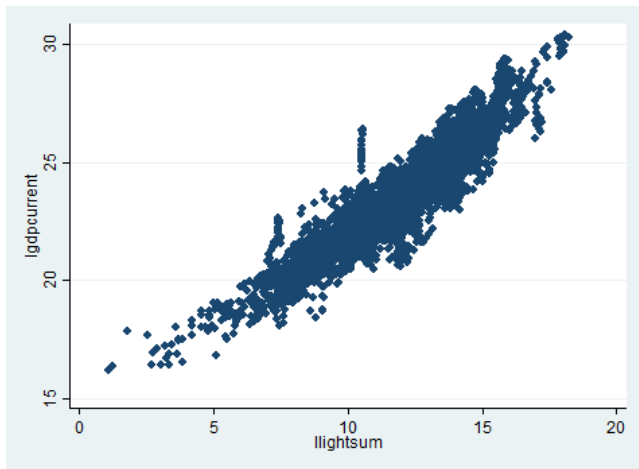
- ▶ Township-level measure of economic activity
  - ▶ National Oceanic and Atmospheric Administration (NOAA) Night Time Light Data from Defense Meteorological Satellite Program (DMSP), 1994-2013
  - ▶ Reliable measure of local economic activities (Henderson et a. 2012; Donaldson and Storeygard 2016)
  - ▶ lights-GDP elasticity of 0.3 (Henderson et a. 2012)
  - ▶ About 600m × 600m pixel with light intensity 0-63
- ▶ Construction of a 20-year township level balanced panel data set
  - ▶ 1,265 townships × 20 years = 25,300 observations



## Data (cont.)

- ▶ Construction of a 20-year township level balanced panel data set
  - ▶ 1,265 townships  $\times$  20 years = 25,300 observations
- ▶ Distance between KTX stations and townships
  - ▶ Location of KTX stations are geo-coded using KTX line maps
  - ▶ Euclidean distance from the nearest KTX station
- ▶ Excluding Gangwon province, islands and urban districts

## World GDP and Light Intensity 1992-2013

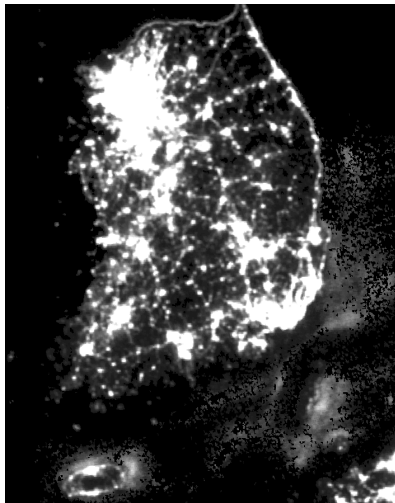


Source: Kyoochul Kim. (2017) North Korean Economy Seen from the Sky,  
*KDI Policy Research Series Working Paper*

# Raw Satellite Images 1994 & 2013

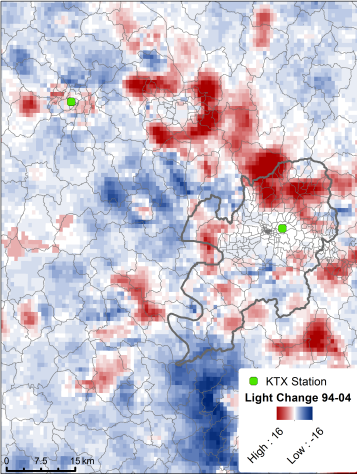


(a) 1994

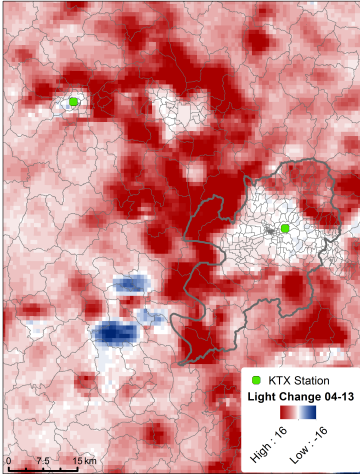


(b) 2013

# Growth of Lights near Daegu & Gimcheon

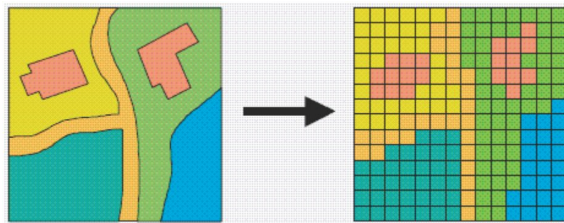


(c) 1994-2004



(d) 2004-2013

## GIS data: vector and raster



**Vector:** points, lines, polygons

**Raster:** satellite images, altitudes, land cover

**Create outcome variable:**

**Proxy for economic activity using night time satellite images**

# Data preparation - download data set from NOAA



ngdc.noaa.gov/eog/dmsp/downloadV4composites.html



NOAA > HESDIS > HCEI (formerly HGOI) > EOG



EOG Home

DMSPP Archive Description

Description of DMSPP Sensors

Data Availability

Data Services and Pricing

Data Download

Online Maps and Web Services

Nighttime Lights Posters

Presentations

Publications

News & Media

Items of Interest

Documents About DMSPP

Nighttime Lights Temporal Loops

Nightsat

MicMurdo Ground Project Data Resources

FAQ

## Version 4 DMSPP-OLS Nighttime Lights Time Series

The files are cloud-free composites made using all the available archived DMSPP-OLS smooth resolution data for calendar years. In cases where two satellites were collecting data - two 30 arc second grids, spanning -180 to 180 degrees longitude and -65 to 75 degrees latitude.

[More Information](#)

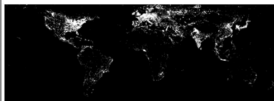
[Download Average Visible, Stable Lights, and Cloud Free Coverages](#) || [Download Average Lights X Pct](#)

**NOTE:** Disk space required for compressed data is ~300 MB and uncompressed data is ~3 GB.

**NOTE:** These files are gzipped and added to tar files under linux. Winzip and some other windows utilities may convert newlines in the data into a windows linefeed. To avoid this in Win Miscellaneous and uncheck the 'TAR file smart CR/LF conversion'.

Average Visible, Stable Lights, & Cloud Free Coverages						
Year\Sat.	F10	F12	F14	F15	F16	F18
1992	<a href="#">F101992</a>	-----	-----	-----	-----	-----
1993	<a href="#">F101993</a>	-----	-----	-----	-----	-----
1994	<a href="#">F101994</a>	<a href="#">F121994</a>	-----	-----	-----	-----
1995	-----	<a href="#">F121995</a>	-----	-----	-----	-----
1996	-----	<a href="#">F121996</a>	-----	-----	-----	-----
1997	-----	<a href="#">F121997</a>	<a href="#">F141997</a>	-----	-----	-----
1998	-----	<a href="#">F121998</a>	<a href="#">F141998</a>	-----	-----	-----
1999	-----	<a href="#">F121999</a>	<a href="#">F141999</a>	-----	-----	-----
2000	-----	-----	<a href="#">F142000</a>	<a href="#">F152000</a>	-----	-----
2001	-----	-----	<a href="#">F142001</a>	<a href="#">F152001</a>	-----	-----
2002	-----	-----	<a href="#">F142002</a>	<a href="#">F152002</a>	-----	-----
2003	-----	-----	<a href="#">F142003</a>	<a href="#">F152003</a>	-----	-----
2004	-----	-----	-----	<a href="#">F152004</a>	<a href="#">F162004</a>	-----
2005	-----	-----	-----	<a href="#">F152005</a>	<a href="#">F162005</a>	-----
2006	-----	-----	-----	<a href="#">F152006</a>	<a href="#">F162006</a>	-----
2007	-----	-----	-----	<a href="#">F152007</a>	<a href="#">F162007</a>	-----
2008	-----	-----	-----	-----	<a href="#">F162008</a>	-----
2009	-----	-----	-----	-----	<a href="#">F162009</a>	-----
2010	-----	-----	-----	-----	-----	<a href="#">F182010</a>
2011	-----	-----	-----	-----	-----	<a href="#">F182011</a>
2012	-----	-----	-----	-----	-----	<a href="#">F182012</a>
2013	-----	-----	-----	-----	-----	<a href="#">F182013</a>

F15 2003 Nighttime Lights Composite



# Data preparation - download QGIS software

3.10.0  
3.4.13 LTR

DISCOVER QGIS FOR USERS GET INVOLVED DOCUMENTATION

Search English

Time until freeze 2020-01-17 12:00:00 UTC 66d 23h 59m  
Time until packaging 2020-02-21 12:00:00 UTC 101d 23h 59m  
Time until next pointrelease 2019-11-22 12:00:00 UTC 10d 23h 59m

## QGIS

### A Free and Open Source Geographic Information System

# QGIS 3.10 A Coruña

has been released!

**New release: 3.10!**  
Get the [installer](#) or [packages](#) for your Operating System!

Create, edit, visualise, analyse and publish geospatial information on Windows, Mac, Linux, BSD (Android coming soon)

For your desktop, server, in your web browser and as developer libraries

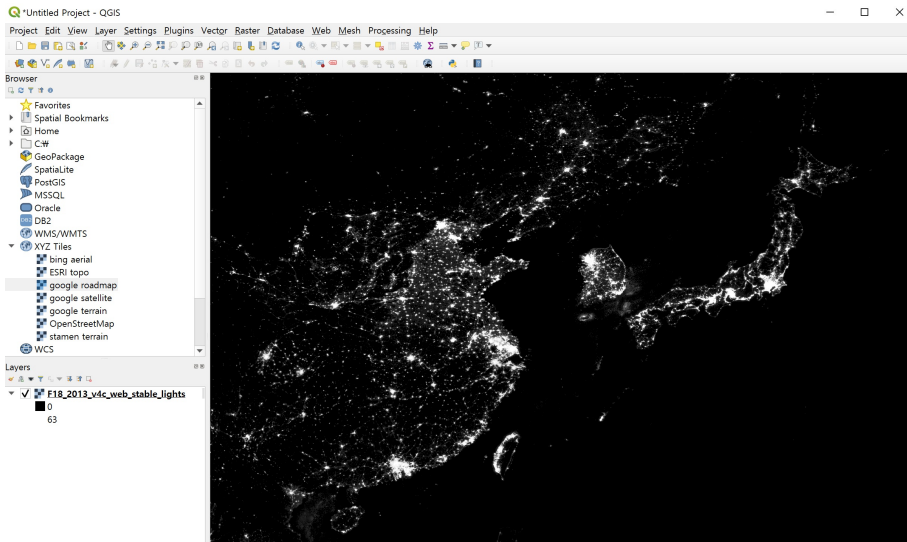
[Download Now](#) [Support QGIS](#)

Version 3.10.0  
Version 3.4.13 LTR

Donate now!



# Data preparation - load satellite images (drag and drop)



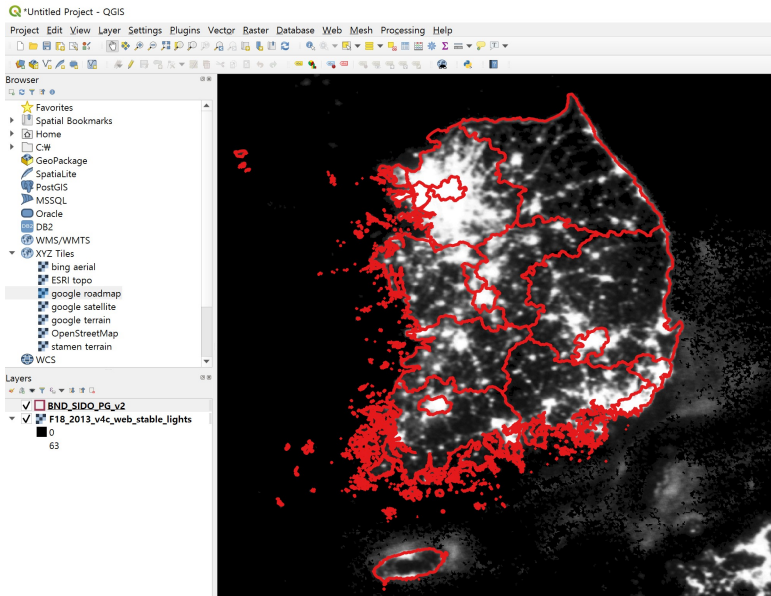
# Data preparation - download admin boundaries

sgis.kostat.go.kr/contents/shortcut/shortcut\_05.jsp

## 통계지역경계

대상자료명	기준년도	자료형식	공개여부	대상지역	가격
센서스용 행정구역경계(전체)	2019, 2018, 2017, 2016, 2015, 2014, 2013, 2012, 2011, 2010, 2009, 2008, 2007, 2006, 2005, 2004, 2003, 2002, 2001, 2000, 1995, 1990, 1985, 1980, 1975	SHP	공개	전국	무료
센서스용 행정구역경계(시도)	2019, 2018, 2017, 2016, 2015, 2014, 2013, 2012, 2011, 2010, 2009, 2008, 2007, 2006, 2005, 2004, 2003, 2002, 2001, 2000, 1995, 1990, 1985, 1980, 1975	SHP	공개	전국	무료
센서스용 행정구역경계(시군구)	2019, 2018, 2017, 2016, 2015, 2014, 2013, 2012, 2011, 2010, 2009, 2008, 2007, 2006, 2005, 2004, 2003, 2002, 2001, 2000, 1995, 1990, 1985, 1980, 1975	SHP	공개	전국	무료
센서스용 행정구역경계(읍면동)	2019, 2018, 2017, 2016, 2015, 2014, 2013, 2012, 2011, 2010, 2009, 2008, 2007, 2006, 2005, 2004, 2003, 2002, 2001, 2000, 1995, 1990, 1985, 1980, 1975	SHP	공개	전국	무료

# Data preparation - load admin. boundaries (drag and drop)



# Data preparation - zonal statistics

The screenshot shows the QGIS 'Zonal Statistics' dialog box. The 'Parameters' tab is active. The 'Raster layer' is set to 'F18\_2013\_v4c\_web\_stable\_lights [EPSG:4326]'. The 'Raster band' is 'Band 1 (Gray)'. The 'Vector layer containing zones' is 'BND\_SIDO\_PG\_v2 [EPSG:5179]'. The 'Output column prefix' is empty. Under 'Statistics to calculate', '1 options selected' is shown. A 'Multiple selection' dialog box is open in the foreground, listing various statistical options with checkboxes. The 'Mean', 'Count', 'Sum', 'Median', and 'Std. dev.' options are checked. The 'Min', 'Max', 'Range', 'Minority', 'Majority (mode)', 'Variety', 'Variance', and 'All' options are unchecked. The 'Multiple selection' dialog has buttons for 'Select All', 'Clear Selection', 'Toggle Selection', 'OK', and 'Cancel'. The 'Zonal statistics' help text on the right states: 'This algorithm calculates statistics of a raster layer for each feature of an overlapping polygon vector layer.' At the bottom of the main dialog, there is a progress bar at 0%, a 'Run as Batch Process...' button, and 'Run', 'Close', and 'Help' buttons.

Parameters Log

Raster layer  
F18\_2013\_v4c\_web\_stable\_lights [EPSG:4326]

Raster band  
Band 1 (Gray)

Vector layer containing zones  
BND\_SIDO\_PG\_v2 [EPSG:5179]

Output column prefix  
-

Statistics to calculate  
1 options selected

**Multiple selection**

- Mean
- Count
- Sum
- Median
- Std. dev.
- Min
- Max
- Range
- Minority
- Majority (mode)
- Variety
- Variance
- All

Zonal statistics  
This algorithm calculates statistics of a raster layer for each feature of an overlapping polygon vector layer.

0%

Run as Batch Process... Run Close Help

# Data preparation - zonal statistics (results)

BND\_SIDO\_PG\_v2 :: Features Total: 17, Filtered: 17, Selected: 1

OBJECTID	BASE_YEAR	SIDO_CD	SIDO_NM	SIDO_NM_EN	_count	_sum	_mean	_median	_stdev
1	1 2015	11	서울특별시	Seoul	894.0000000000...	56146.000000000...	62.80313199105...	63.00000000000...	0.922834383977...
2	2 2015	21	부산광역시	Busan	1135.0000000000...	67250.000000000...	59.25110132158...	63.00000000000...	7.227909448695...
3	3 2015	22	대구광역시	Daegu	1268.0000000000...	60506.000000000...	47.71766561514...	54.00000000000...	16.49799061204...
4	4 2015	23	인천광역시		1701.0000000000...	72724.000000000...	42.75367430922...	48.00000000000...	21.22525241581...
5	5 2015	24	광주광역시	Gwangju	714.0000000000...	34614.000000000...	48.47899159663...	56.00000000000...	16.07363068708...
6	6 2015	25	대전광역시	Daejeon	781.0000000000...	37611.000000000...	48.15749039692...	55.00000000000...	15.72016934716...
7	7 2015	26	울산광역시	Ulsan	1526.0000000000...	61828.000000000...	40.51638269986...	41.00000000000...	18.75221420622...
8	8 2015	29	세종특별자치시	Sejong	678.0000000000...	25698.000000000...	37.90265486725...	37.00000000000...	13.69751044393...
9	9 2015	33	충청북도	Chungcheong N.	10764.000000000...	218648.000000000...	20.31289483463...	14.00000000000...	16.03393769907...
10	10 2015	34	충청남도	Chungcheong S.	12068.000000000...	287709.000000000...	23.84065296652...	19.00000000000...	14.95513834931...
11	11 2015	35	전라북도	Jeolla N.	11518.000000000...	233580.000000000...	20.27956242403...	15.00000000000...	14.76021755640...
12	12 2015	36	전라남도	Jeolla S.	18019.000000000...	281761.000000000...	15.63688328986...	10.00000000000...	13.94421573721...
13	13 2015	37	경상북도	Gyeongsang N.	27540.000000000...	462284.000000000...	16.78591140159...	11.00000000000...	14.68771728295...
14	14 2015	38	경상남도	Gyeongsang S.	15147.000000000...	334136.000000000...	22.05954974582...	14.00000000000...	17.92753745139...
15	15 2015	39	제주특별자치도	Jeju	2635.0000000000...	51117.000000000...	19.39924098671...	14.00000000000...	14.49779147588...
16	16 2015	31	경기도	Gyeonggi	15284.000000000...	627223.000000000...	41.03788275320...	45.00000000000...	19.67272309940...
17	17 2015	32	강원도	Gangwon	24930.000000000...	296691.000000000...	11.90096269554...	8.00000000000...	12.86967984535...

**Create explanatory variable:**

**Distance from the nearest KTX station**

# Data preparation - load ktx stations (digitization)

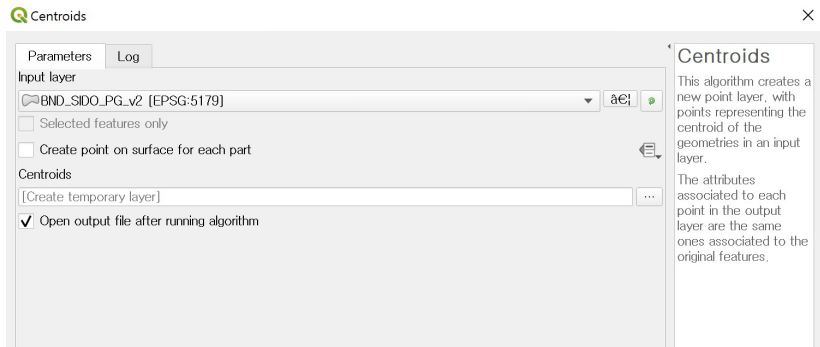
The screenshot displays the QGIS interface for an 'Untitled Project'. The top menu bar includes Project, Edit, View, Layer, Settings, Plugins, Vector, Raster, Database, Web, Mesh, Processing, and Help. Below the menu is a toolbar with various GIS tools. The left sidebar contains a 'Browser' panel with a tree view of data sources, including Favorites, Spatial Bookmarks, Home, C:\, GeoPackage, SpatialLite, PostGIS, MSSQL, Oracle, DB2, WMS/WMTS, and XYZ Tiles. Under XYZ Tiles, several providers are listed: bing aerial, ESRI topo, google roadmap, google satellite, google terrain, OpenStreetMap, stamen terrain, and WCS. The 'Layers' panel at the bottom left shows a list of loaded layers: 'new\_ktx\_stn\_wgs' (checked), 'BND\_SIDO\_PG\_v2' (checked), and 'F18\_2013\_v4c\_web\_stable\_lights' (unchecked). The 'F18\_2013\_v4c\_web\_stable\_lights' layer is expanded to show two sub-layers: '0' and '63'. The main map area shows a geographical region with a red outline representing the boundary. Numerous blue dots are scattered across the map, representing the locations of ktx stations. The background of the map is white, and the red outline is composed of many small segments, suggesting a digitized boundary.

# Data preparation - create centroids for admin units

## Calculate linear distance (Point A to Point B)

KTX stations are points but admin. units are not

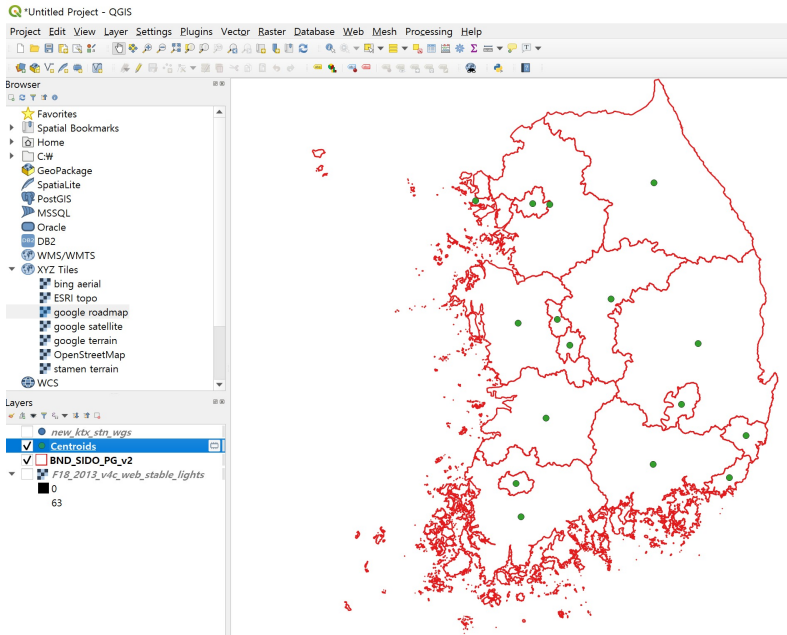
→ create points for admin units (centroids)



The screenshot shows the 'Centroids' algorithm dialog box in QGIS. It has two tabs: 'Parameters' and 'Log'. The 'Parameters' tab is active. Under 'Input layer', there is a dropdown menu showing 'BND\_SIDO\_PG\_v2 [EPSG:5179]' and a 'Selected features only' checkbox which is unchecked. Below that is a 'Create point on surface for each part' checkbox, also unchecked. The 'Centroids' section has a text field containing '[Create temporary layer]' and a 'Log' button. At the bottom, there is a checked checkbox for 'Open output file after running algorithm'. To the right of the dialog is a help panel titled 'Centroids' with the following text: 'This algorithm creates a new point layer, with points representing the centroid of the geometries in an input layer. The attributes associated to each point in the output layer are the same ones associated to the original features.'



# Data preparation - centroids for admin. units



# Data preparation - calculate linear distance (distance matrix)

Distance Matrix

Parameters Log

Input point layer  
admin\_centroid [EPSG:5179] Selected features only

Input unique ID field  
abc SIDO\_CD

Target point layer  
new\_ktx\_stn\_wgs [EPSG:32652] Selected features only

Target unique ID field  
123 Id

Output matrix type  
Linear (N\*k x 3) distance matrix

Use only the nearest (k) target points  
1

Distance matrix  
[Create temporary layer]

Open output file after running algorithm

Distance matrix  
This algorithm creates a table containing a distance matrix, with distances between all the points in a points layer.

# Data preparation - distance matrix calculation results

distance\_matrix :: Features Total: 17, Filtered: 17, Selected: 0

	InputID	TargetID	Distance
1	21		40 5481.527122396...
2	11		49 1883.815571582...
3	29		22 9129.183636982...
4	32		49 118742.0085239...
5	26		29 8930.817285475...
6	25		43 2000.324040809...
7	24		38 4547.781472031...
8	36		24 22035.69575824...
9	35		36 23285.04105391...
10	34		42 37470.36517924...
11	33		22 46800.49058272...
12	31		49 18006.80122811...
13	39		39 156508.5493409...

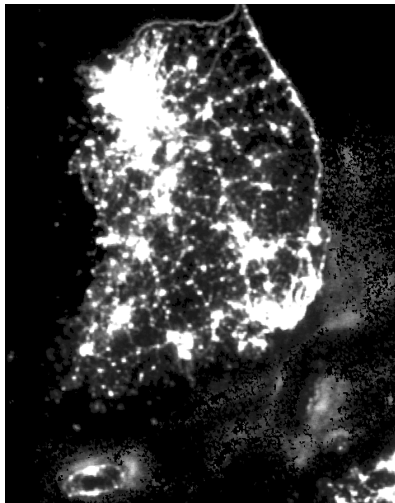
## Final step

- ▶ Export created tables to CSV / Excel file
- ▶ Merge using ids with Stata
- ▶ Repeat for other years

# Raw Satellite Images 1994 & 2013

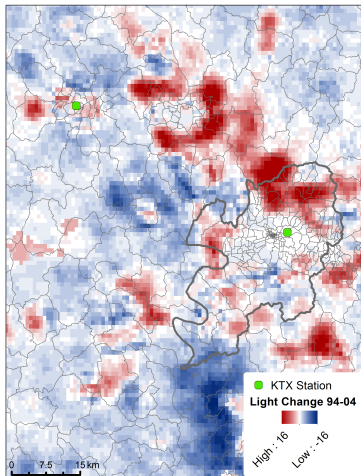


(e) 1994

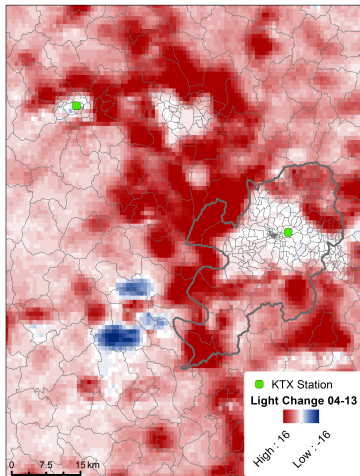


(f) 2013

# Growth of Lights near Daegu & Gimcheon



(g) 1994-2004



(h) 2004-2013

# Empirical Strategy

## We use difference in differences approach

- ▶ Two groups
  - ▶ Treatment: townships “closer” to KTX station
  - ▶ Control: townships “further away” from KTX station
  - ▶ → definition of “closer” will be discussed
- ▶ Causal effect of treatment (introduction of KTX)
  - ▶ How economic activities of treatment group change compared to control group?
    - Before and after the introduction of KTX
- ▶ Important assumption
  - ▶ The pre-trend of outcome should be parallel for both groups

# Empirical Strategy - Model 1

## Difference in differences approach (treatment = 0 or 1)

$$\begin{aligned} LIGHT_{k,m,t} = & \alpha_0 + \alpha_1 TREAT_{k,m} + \alpha_2 POST_t \\ & + \alpha_3 TREAT_{k,m} \cdot POST_t \\ & + X'_{k,m,t} \Psi + \tau_k + \epsilon_{k,m,t} \end{aligned}$$

- ▶  $LIGHT_{k,m,t}$  = average light intensity of township  $m$  in county  $k$ , year  $t$
- ▶  $TREAT_{k,m} = 1$  if KTX distance  $\leq 34$  kilometers (median value)
- ▶  $POST_t =$  post 2004 dummy
- ▶  $\tau_k =$  county fixed effects (No township effects yet)
- ▶ Distance measured from KTX stations in 2016
- ▶ Treatment was assumed to begin in 2004 (underestimated)



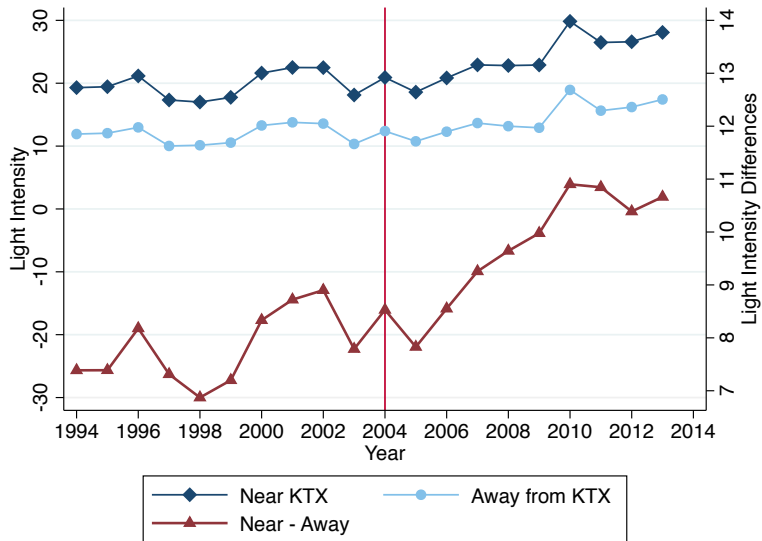
# Identification

- ▶ Key identifying assumption: common trends
  - ▶ Trends of economic activity would be the same in both treatment and control groups in the absence of treatment (within the same county)
- ▶ Unobserved characteristics of rural units are inconsequential to the choice of routes
  - ▶ (Chandra and Thompson, 2000; Michaels, 2008; Redding and Turner, 2015)
- ▶ Endogenous selection of KTX stations
  - ▶ KTX stations were built where the increased economic activity is predicted
  - ▶ Policy makers to choose KTX stations in 1990 where the increase in economic activity in nearby townships would occur exactly 14 years later

## Identification (cont.)

1. Did policy makers in 1990 predict the increased economic activity in 2004?
  - ▶ Planned operation year delayed multiple times due to financial constraints (Moon and Kim 2012)
  - ▶ Actual year of operation was 2004
2. Was the major determinant of the location of original KTX stations increased future economic activity? (Korea Rail Network Authority 2011; Moon and Kim 2012)
  - ▶ Not likely (e.g., Daejeon, Daegoo, Kyungjoo)
3. What about newly established KTX stations?
  - ▶ Began operation in 2010 (construction began in 2007, 2008)
  - ▶ New stations announced in 2003
  - ▶ Possible announcement effect (use Kangwon province to check)

# Results 1



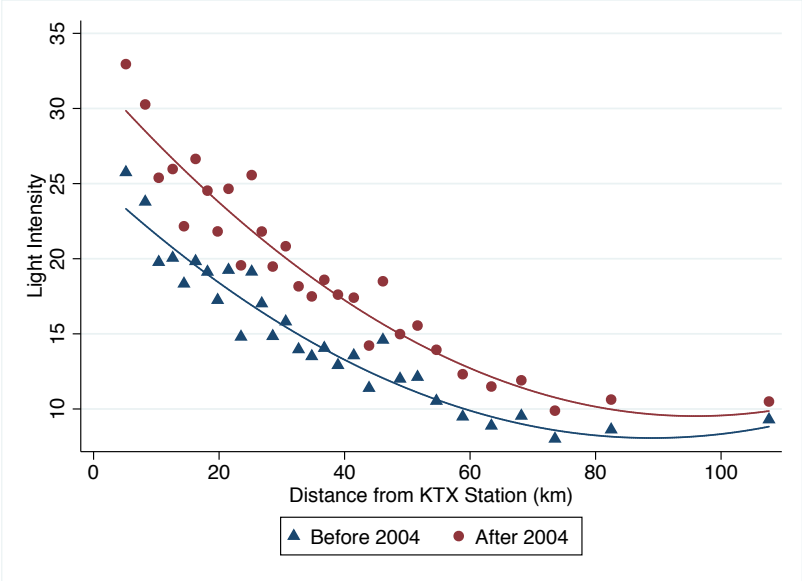
**Table 1:** Diff-in-Diffs Estimates of the Impact of the KTX on Light

	(1)	(2)	(3)	(4)
<i>treatment</i> × <i>post</i>	2.3*** (0.28)	2.3*** (0.27)	2.3*** (0.55)	2.3*** (0.55)
Observations	25,300	25,300	25,300	25,300
R-squared	0.15	0.21	0.59	0.63
County FE			X	X
Covariate		X		X

Source: U.S. National Oceanic and Atmospheric Administration night time light data, 1994-2013. Standard errors are clustered at the county level.

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

# KTX Distance and Light Intensity



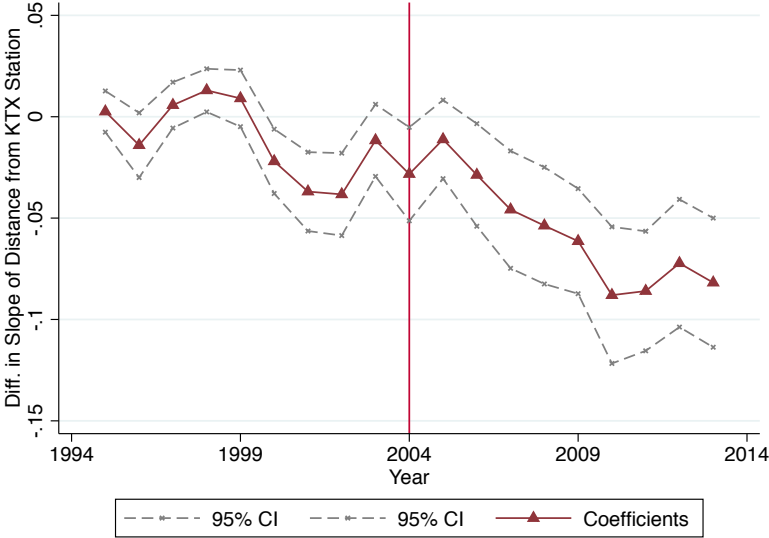
## Empirical Strategy - Model 2

### Difference in differences approach (continuous treatment)

$$\begin{aligned} LIGHT_{k,m,t} = & \delta_0 + \delta_1 DIST_{k,m} + \gamma_t \\ & + \sum_t \lambda_t DIST_{k,m} \cdot \gamma_t \\ & + X'_{k,m,t} \Theta + \sigma_k + \xi_{k,m,t} \end{aligned}$$

- ▶  $LIGHT_{k,m,t}$  = average light intensity in county  $k$ , township  $m$ , year  $t$
- ▶  $DIST_{k,m}$  = KTX distance (kilometer)
- ▶  $\gamma_t$  = year fixed effects
- ▶  $\sigma_k$  = county fixed effects

# Results 2



	(1) Coeff.	(1) S.E.	(2) Coeff.	(2) S.E.
1995	0.00	0.02	0.00	0.01
1996	-0.01	0.02	-0.01*	0.01
1997	0.01	0.01	0.01	0.01
1998	0.01	0.01	0.01**	0.01
1999	0.01	0.02	0.01	0.01
2000	-0.02	0.02	-0.02***	0.01
2001	-0.04**	0.02	-0.04***	0.01
2002	-0.04**	0.02	-0.04***	0.01
2003	-0.01	0.01	-0.01	0.01
2004	-0.03*	0.02	-0.03**	0.01
2005	-0.01	0.02	-0.01	0.01
2006	-0.03*	0.02	-0.03**	0.01
2007	-0.05***	0.02	-0.05***	0.01
2008	-0.05***	0.02	-0.05***	0.01
2009	-0.06***	0.02	-0.06***	0.01
2010	-0.09***	0.02	-0.09***	0.02
2011	-0.09***	0.02	-0.09***	0.01
2012	-0.07***	0.02	-0.07***	0.02
2013	-0.08***	0.02	-0.08***	0.02
Observations	25,300		25,300	
R-squared	0.21		0.70	
County FE			X	
Covariate			X	

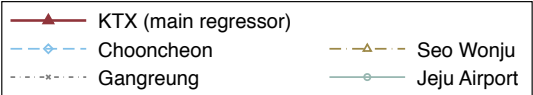
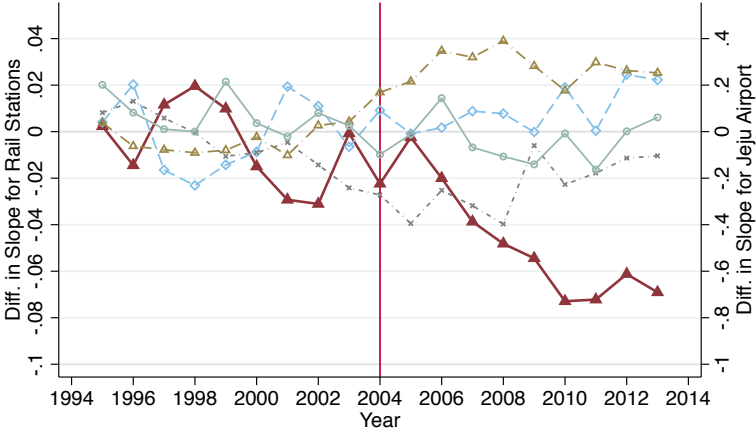
Source: U.S. National Oceanic and Atmospheric Administration night time light data, 1994-2013. Standard errors are clustered at the county level. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .



# Falsification Checks

- ▶ Increased economic activities near rail stations after 2004
- ▶ Use Gangwon Province without any KTX connections
  - ▶ Distance to major rail stations and future KTX stations
- ▶ Use Jeju Island
  - ▶ Proximity to Jeju International Airport

# Results 2



# Conclusions

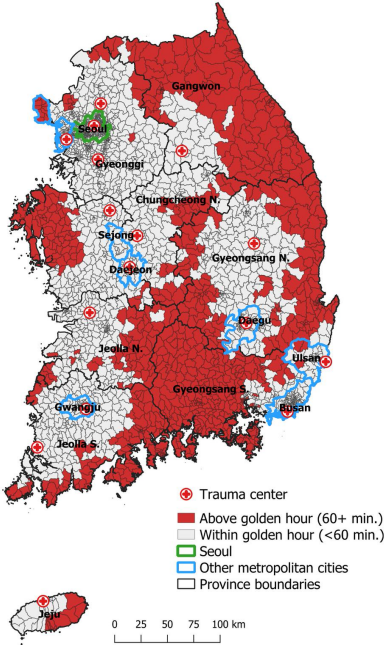
## Summary

- ▶ Evaluation the impact of HSR system on local economic activity
- ▶ Introduction of novel panel data on local economic activity

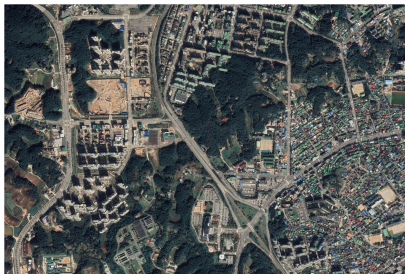
## Future work

- ▶ Economic growth or reorganization?
  - ▶ Currently collecting data on population, business, etc.
- ▶ Mechanisms?
  - ▶ Housing booms, new jobs, region-based policies
  - ▶ Entry of firms, employment, productivity
  - ▶ Plan to use population and land price data
- ▶ From light intensity to \$ value

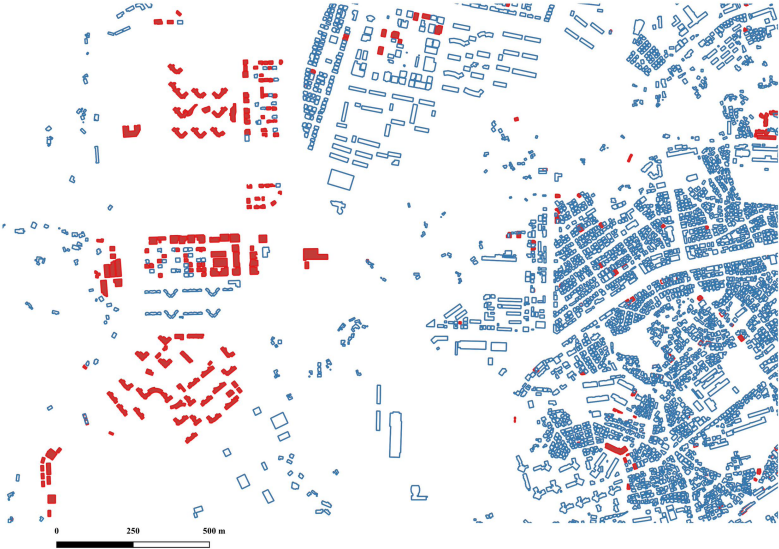
# Medical accessibility: trauma centers



# 강릉시 유천동 (평창올림픽 이전과 이후)



# 행안부 건물 데이터를 통해 올림픽 이후 새로 생긴 건물 identify

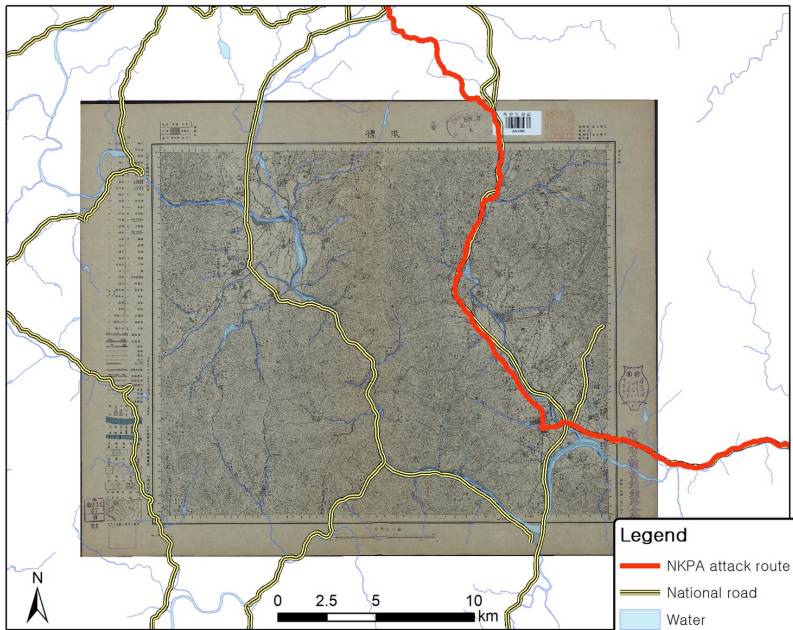


# 행안부 건물 데이터를 통해 올림픽 이후 새로 생긴 건물 identify



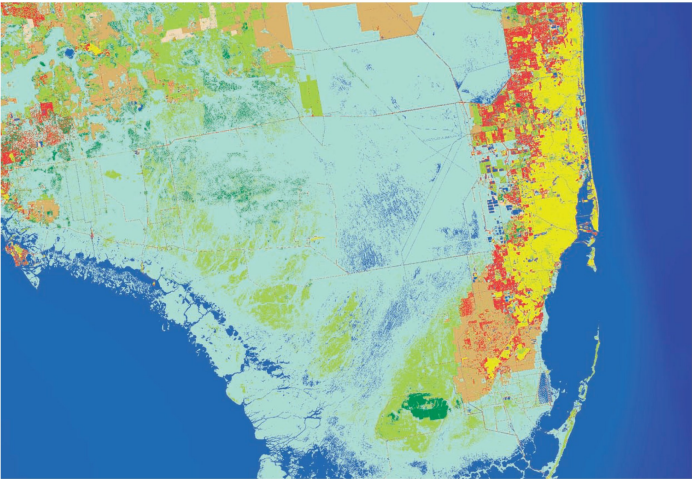


# Digitizing 1918 road networks & NKPA attack routes



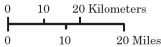
# Burchfield et al. 2006. QJE

Urban expansion in Miami (1976-1992)



- Urban land circa 1976
- Urban land built 1976-92
- Water
- Bare rock and sand

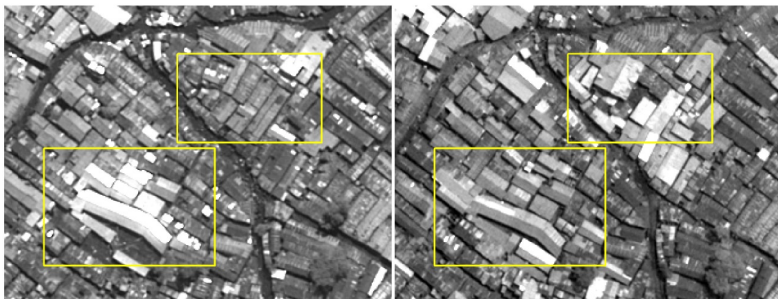
- Forest
- Range and grassland
- Agricultural land
- Wetlands



# Marx et al. 2019. **AEJ: Applied Economics**

## Roof types as a proxy for poverty

Appendix Figure A2: Old and New Roofs in Kibera



Note: Both pictures are taken over the same area of the slum with the same resolution (0.5 meters panchromatic).

The picture in the left panel was taken in July 2009 and that in the right panel in August 2012.

The yellow rectangles highlight clusters of roofs that markedly evolved over the period.

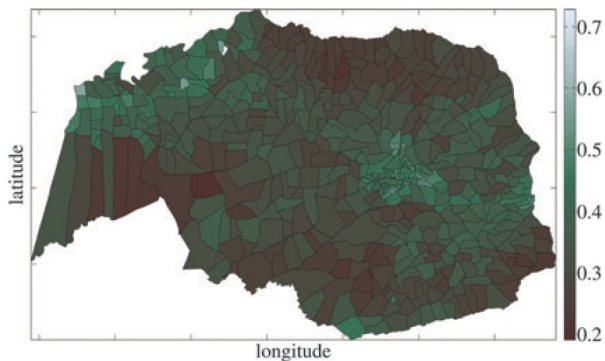
Roofs highlighted in the bottom rectangle degraded while roofs within the top rectangle were upgraded in the same timeframe.

The picture area is approximately 175 meters long and 140 meters wide.

Shiny new roofs v.s. old roofs in villages in Kenya



**FIG. 1.** Sample satellite image patches of the region with **(a)** thatched roofs in the center and **(b)** a metal roof in the center.



**FIG. 3.** Estimated proportion of roofs that are metal per village.

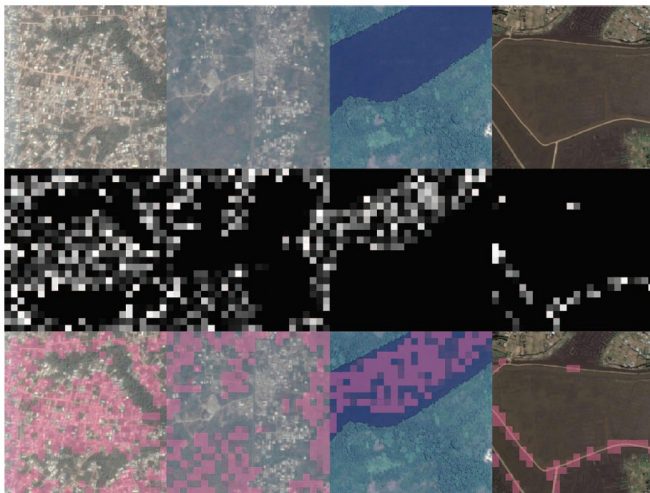
## RESEARCH ARTICLES

### ECONOMICS

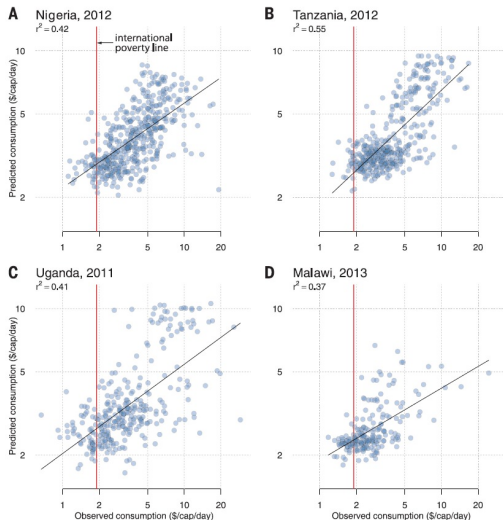
# Combining satellite imagery and machine learning to predict poverty

Neal Jean,<sup>1,2\*</sup> Marshall Burke,<sup>3,4,5\*†</sup> Michael Xie,<sup>1</sup> W. Matthew Davis,<sup>4</sup>  
David B. Lobell,<sup>3,4</sup> Stefano Ermon<sup>1</sup>

Reliable data on economic livelihoods remain scarce in the developing world, hampering efforts to study these outcomes and to design policies that improve them. Here we demonstrate an accurate, inexpensive, and scalable method for estimating consumption expenditure and asset wealth from high-resolution satellite imagery. Using survey and satellite data from five African countries—Nigeria, Tanzania, Uganda, Malawi, and Rwanda—we show how a convolutional neural network can be trained to identify image features that can explain up to 75% of the variation in local-level economic outcomes. Our method, which requires only publicly available data, could transform efforts to track and target poverty in developing countries. It also demonstrates how powerful machine learning techniques can be applied in a setting with limited training data, suggesting broad potential application across many scientific domains.



**Fig. 2. Visualization of features.** By column: Four different convolutional filters (which identify, from left to right, features corresponding to urban areas, nonurban areas, water, and roads) in the convolutional neural network model used for extracting features. Each filter “highlights” the parts of the image that activate it, shown in pink. By row: Original daytime satellite images from Google Static Maps, filter activation maps, and overlay of activation maps onto original images



**Fig. 3. Predicted cluster-level consumption from transfer learning approach (y axis) compared to survey-measured consumption (x axis).** Results are shown for Nigeria (A), Tanzania (B), Uganda (C), and Malawi (D). Predictions and reported  $r^2$  values in each panel are from fivefold cross-validation. Black line is the best fit line, and red line is international poverty line of \$1.90 per person per day. Both axes are shown in logarithmic scale. Countries are ordered by population size.



