The View from Above: Applications of Satellite Data in Economics

Dave Donaldson and Adam Storeygard

Photographic images of the Earth from above have been recorded since the late 19th century, initially by cameras attached to balloons, kites, or pigeons. This remotely sensed information has been used for economic analysis since at least the 1930s; for example, Monmonier (2002) discusses how several dozen photographic crews started flying over US farmland in the 1930s to study production and conservation, and by 1941 had recorded images of more than 90 percent of total acreage. However, the past decade or so has seen a dramatic change in the way that economists can learn by watching our planet from above. A revolution has taken place in remote sensing and allied fields such as computer science, engineering, and geography. Petabytes of satellite imagery have become publicly accessible at increasing resolution, many algorithms for extracting meaningful social science information from these images are now routine, and modern cloud-based processing power allows these algorithms to be run at global scale. This paper seeks to introduce economists to the science of remotely sensed data, and to give a flavor of how this new source of data has been used by economists so far and what might be done in the future.

We group the main advantages of such remote sensing data to economists into three categories: 1) access to information difficult to obtain by other means; 2) unusually high spatial resolution; and 3) wide geographic coverage.

■ *Dave Donaldson is Associate Professor of Economics, Stanford University, Stanford, California. Adam Storeygard is Assistant Professor of Economics, Tufts University, Medford, Massachusetts. Their email addresses are ddonald@stanford.edu and Adam.Storeygard@ tufts.edu.*

The first advantage is simply that remote sensing technologies can collect panel data at low marginal cost, repeatedly, and at large scale on proxies for a wide range of hard-to-measure characteristics. We discuss below economic analysis that already uses remotely sensed data on night lights, precipitation, wind speed, flooding, topography, forest cover, crop choice, agricultural productivity, urban development, building type, roads, pollution, beach quality, and fish abundance. Many more characteristics of potential interest to economists have already been measured remotely and used in other fields. Most of these variables would be prohibitively expensive to measure accurately without remote sensing, and there are settings in which the official government counterparts of some remotely sensed statistics (such as pollution or forestry) may be subject to manipulation.

As an example of this first advantage, access to previously unavailable data, consider Burgess et al. (2012), who study deforestation in Indonesia. Forestry is highly regulated in Indonesia, but local officials are sometimes bribed to overlook illegal logging. Thus, administrative statistics, produced in collaboration with such officials, are subject to concerns about misreporting. Satellite data allowed Burgess et al. (2012) to measure deforestation—an example of which can be seen in Figure 1—in an objective manner, and thereby to shed new light on the political economy of natural resource appropriation. Consistent with Cournot competition across jurisdictions, these authors find that each additional district created in a province induces an 8.2 percent increase in the province's deforestation rate. In principle, with higher spatial and temporal resolution, enough satellites and improved algorithms, future work could track forest cover or agricultural output (even at the level of individual plots) globally.

Figure 2 illustrates a second example where novel satellite data made an important study possible. Jayachandran (2009) estimates the impact of air pollution (particulate matter) resulting from Indonesia's devastating late-1997 forest fire on infant and fetal mortality. Using daily satellite sensor readings about airborne smoke and dust, Jayachandran measures the spread of pollution and estimated that this two-month event caused 16,400 infant and fetal deaths. New technologies are already being developed in order to measure other pollutants such as methane, a major greenhouse gas, perhaps even with the ability to resolve individual plumes.
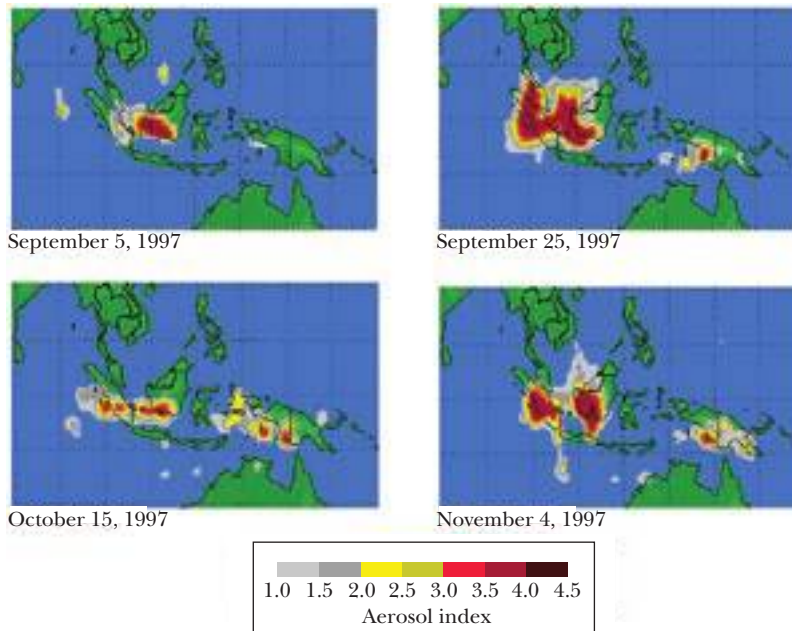
The second advantage of remote sensing data sources is that they are typically available at a substantially higher degree of spatial resolution than are traditional data. Much of the publicly available satellite imagery used by economists provides readings for each of the hundreds of billions of 30-meter-by-30-meter grid cells of land surface on Earth. Many economic decisions (particularly land use decisions such as zoning, building types, or crop choice) are made at approximately this same level of spatial resolution. But since 1999, private companies have offered submeter imagery and, following a 2014 US government ruling, American companies are able to sell imagery at resolutions below 0.5 meters to nongovernment customers for the first time. This is important because even when a coarser unit of *analysis* is appropriate, 900 1-meter pixels provide far more information available for signal extraction than a single 30-meter pixel covering the same area. In addition, some

Figure 1
Forest Cover over Time in Riau Province, Indonesia
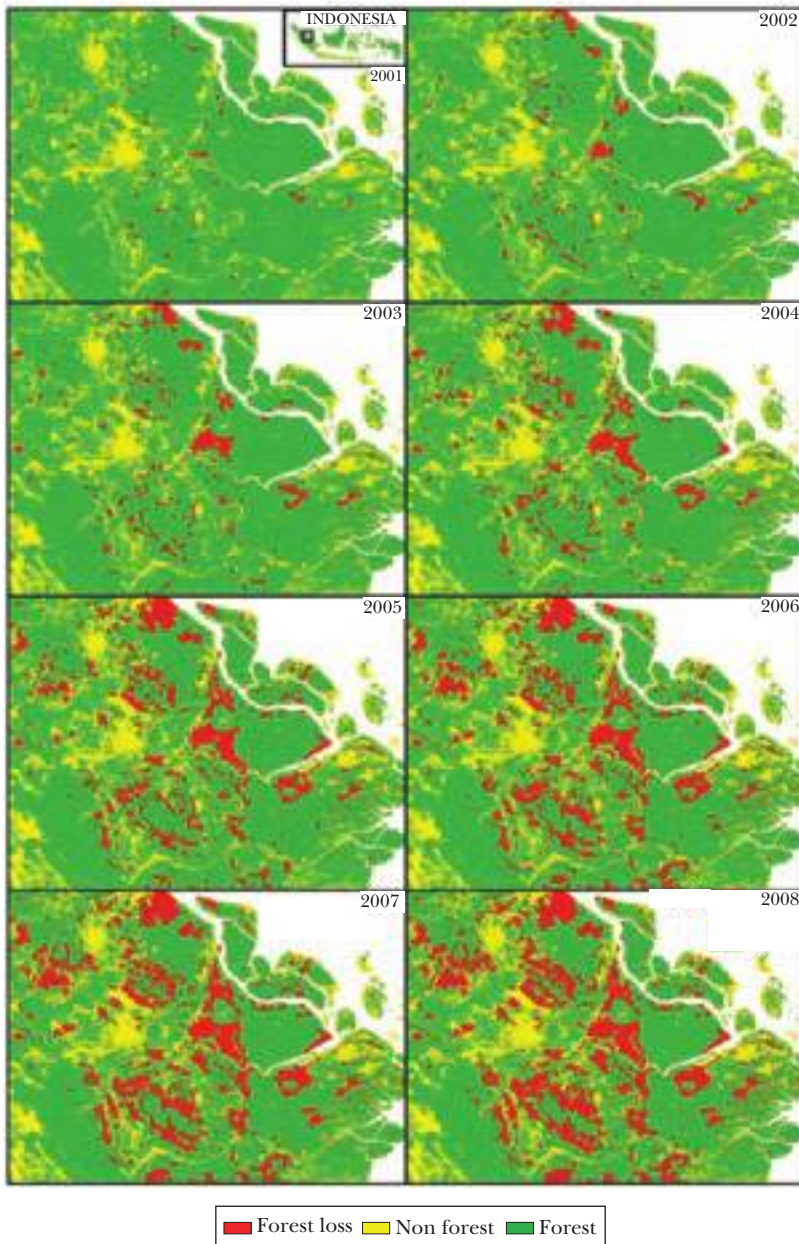
Note: Readers of the printed journal can check out the color versions of this figure in the article at the journal website.

Figure 2

Aerosol Index of Particulate Air Pollution in Indonesia during Massive Wildfires in 1997

September 5, 1997

September 25, 1997

October 15, 1997

November 4, 1997

1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	

Aerosol index

Source: Jayachandran (2006).

innovative identification strategies used by economists exploit stark policy changes that occur at geographic boundaries; these high-spatial-resolution research designs rely intimately on high-spatial-resolution outcome data (for example, Turner, Haughwout, and van der Klaauw 2014).

One example of this second advantage—high spatial resolution—offered by remote sensing comes from Marx, Stoker, and Suri (2015), who use 0.5-meter resolution imagery to measure a proxy for dwelling investments in a Nairobi slum. Figure 3 contains an example. In a context where most houses have corrugated iron roofs, newly replaced or improved roofs are more reflective than older, more rusted ones. These authors refined algorithms that can detect such differential reflectivity and used the resulting data to better understand the role of ethnic favoritism in residential markets in this setting. Market research firms similarly use imagery at this spatial resolution to count cars in parking lots to estimate retail demand, and in principle related algorithms could be used to measure automobile traffic or to count crowds at political rallies. With more high-resolution sensors and improved algorithms, these tasks could be done continuously and at low cost. Relatedly,

Figure 3
Roofs in Kibera, Nairobi

A: July 2009   B: August 2012
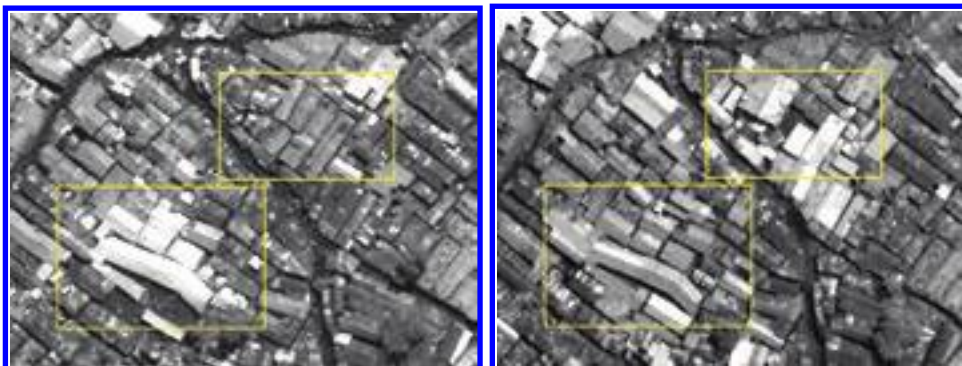
Source: Marx, Stoker, and Suri (2015).
Notes: This is a 50-cm resolution panchromatic image from Kibera, Nairobi. Roofs in the lower left rectangle dulled over time, while those in the upper right rectangle were replaced.

one could imagine trying to improve worldwide measures of building stocks and construction activity from the observed materials and height of built structures.

Of course, high spatial resolution is a relative concept. Even the 1-kilometer resolution NASA data on nighttime lights (Elvidge et al. 1997; Henderson, Storeygard, and Weil, 2012) offers a proxy for settlement patterns and wealth at substantially higher spatial resolution than is typically available. This has enabled creative studies of the spatial distribution of economic activity such as Bleakley and Lin (2012), who document the enduring persistence of US economic activity at sites of early commerce. Prior to the widespread use of canals or railroads, travel inland from the US southeast coast was impeded at the point where rivers cross a geomorphological feature known as the "fall line." Further navigation necessitated portage—reloading of cargo, overland, from one boat to another—and the sites at which such portage took place are almost without exception thriving cities to this day, well over a century after portage became obsolete. Figure 4 shows how this remarkable example of high-spatial-resolution path dependence in urban location is clearly visible in the pattern of night lights, while this relationship would be obscured by conventional US data sources at, for example, the county level.

The third key advantage of remotely sensed data lies in their wide geographic coverage. Only rarely do social scientists enjoy the opportunities, afforded by satellites, to study data that have been collected in a consistent manner—without regard for local events like political strife or natural disasters—across borders and with uniform spatial sampling on every inhabited continent. Equally important, many research satellites (or integrated series of satellites) offer substantial temporal coverage, capturing data from the same location at weekly or even daily frequency for several decades and counting.

Figure 4
**Night Lights in the United States and the Fall Line**

Source: Bleakley and Lin 2012.
Notes: The figure shows night lights across part of the United States in 2003, illustrated with rivers (dotted lines) and the fall line (solid line). Substantial cities on the fall line, starting from the northeast, include Trenton, Philadelphia, Washington DC, Richmond, Augusta, Columbia, Little Rock, Fort Worth, Austin, and San Antonio.

An example of this third feature—global scope—can be seen in work on the economic impacts of climate change in agriculture by Costinot, Donaldson, and Smith (2016). These authors draw on an agronomic model that is partly based on remotely sensed data. The agronomic model, when evaluated under both contemporary and expected (2070–99) climates, predicts a change in agricultural productivity for any crop in any location on Earth. For example, the relative impact for two of the world's most important crops, rice and wheat, is shown in Figure 5. Costinot, Donaldson, and Smith feed these pixel-by-pixel changes into a general equilibrium model of world agricultural trade and then use the model to estimate that climate change can be expected to reduce global agricultural output by about one-sixth (and that international trade is unlikely to mitigate this damage, despite the inherently transnational nature of the shock seen in Figure 5). Given the rate at which algorithms for crop classification and yield measurement have improved in recent years, future applications of satellite data are likely to be particularly rich in the agricultural arena.

Figure 5

Predicted Yield Changes due to Climate Change for (A) Wheat and (B) Rice, Based on Global Agro-Ecological Zones (GAEZ) Models
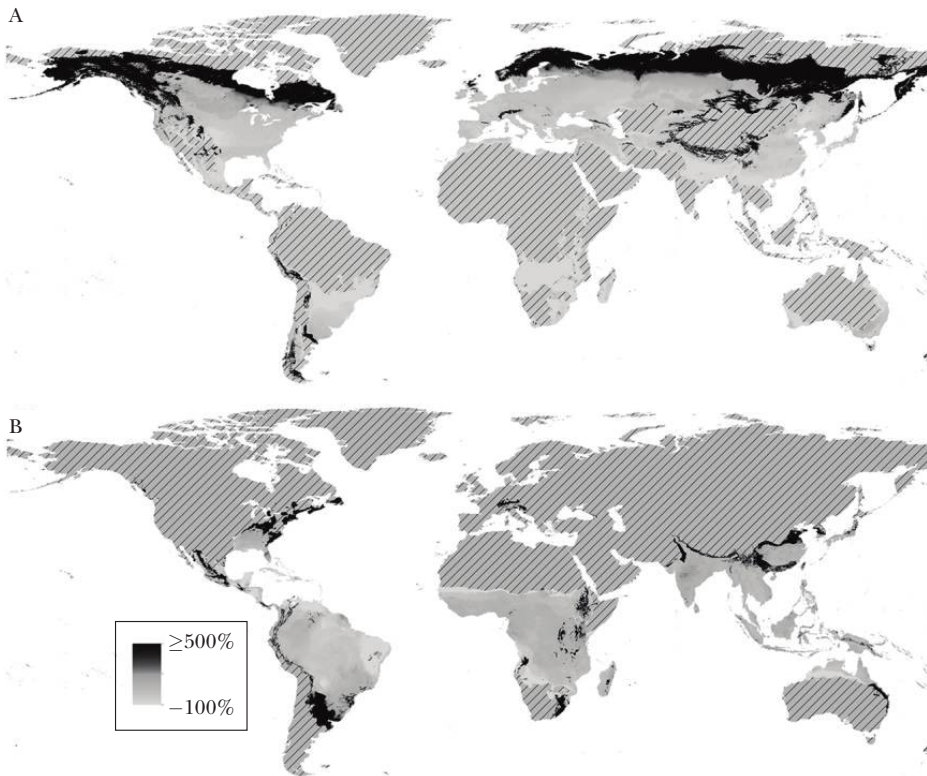
Source: Costinot, Donaldson, and Smith (2016).
Note: Percentage change in predicted yield (between current baseline and the 2070–2099 period) in the GAEZ model. Areas with diagonal stripes indicate regions for which predicted changes are zero.

As these five examples illustrate, economists have learned a great deal from satellites in recent years. In addition to the three advantages described above, this is in no small part due to the fact that imagery from prominent satellites is publicly available, that platforms such as Google's Earth Engine dramatically ease the process of accessing and analyzing such data, and that there is a robust community of remote sensing users from which one can rapidly learn. In the next two sections, we describe the basics of remote sensing technologies and techniques, and then offer a summary of the many areas where satellite data have advanced the study of economics to date.

It is important to keep in mind that, while the science of remote sensing has advanced enormously since the pigeon era, serious challenges remain. A human who looks at a picture can tell whether the image is of an empty field, a pile of rocks, or the roof of an apartment building. It is no surprise that small-scale applications

of remote sensing—for example, to monitor China's construction of military installations on islands in the South China Sea or Boko Haram's destruction of Nigerian villages—have recently enjoyed such prominence. But the more ambitious scale required of typical economics applications requires teaching a machine to make these distinctions, and that is no simple feat. We describe these and other challenges and potential pitfalls below. Ultimately, we believe that we are only at the beginning of a remote sensing data and software revolution, and that as these tools develop they will be of increasing importance to economists. We conclude with some speculative examples that we hope will help the reader to think of the possibilities that lie ahead.

A Primer on Remote Sensing For Economists

What do economists need to learn about remote sensing concepts?[1] In the broadest sense, remote sensing data are collected from above the Earth's surface using aircraft or spacecraft. Here we focus predominantly on satellite data, though a wealth of information is also collected from other spacecraft (such as the International Space Station and the Space Shuttle), as well as from planes and drones. We describe the essentials here, moving from the paths that satellites traverse, to the data that they can collect, to the ways in which these data can be processed and interpreted.

Orbits

Most remote sensing satellites operate in one of two orbits: geostationary or sun-synchronous. *Geostationary* satellites orbit the Earth in a way that keeps them directly above a fixed point on the Equator. Thus, they have the advantage of continuously observing the same area of the Earth's surface at all times—a key feature for national weather services, for example—and the corresponding disadvantage of not observing the rest of the world. Also, the physics behind geostationary orbits means that such satellites must locate approximately 36,000 kilometers above the Earth's surface, which means that their images are relatively low resolution.

Sun-synchronous satellites orbit typically within 6,000 kilometers and often much more closely. As their name implies, they observe the whole Earth at approximately the same time each day, ensuring constant solar illumination (conditional on latitude) at a given time of the year. Sun synchronous orbits are a type of (nearly) polar orbit, so named because they operate in a plane that (nearly) contains the North and South Poles. As a sun-synchronous satellite passes the equator on successive overpasses, the Earth has rotated to the east, and thus the satellite passes above an area to the west of its previous overpass. Sun-synchronous satellites can orbit Earth between 7 and 16 times per day, depending on their altitude. The lowest altitude

---

[1]A more comprehensive treatment can be found in remote sensing textbooks, including Jensen (2006); Campbell and Wynne (2011); Lillesand, Kiefer, and Chipman (2015); and Tempfli, Kerle, Huurneman, and Janssen (2009).

(and highest frequency) sun-synchronous satellites fly at an altitude of 282 kilometers, with an orbital period of 90 minutes.

The area covered in a given overpass of a satellite or other sensor is known as its "swath." This is essentially the width of the sensor's field of vision as it looks down at the Earth. With a wide enough swath, a sensor can record the whole Earth surface every day. Perhaps the easiest way to think of orbital swaths is through an analogy. If one were to cover a ball with a thin ribbon, one could do so as follows. Starting at the North Pole, wrap to the South Pole, and back around to the North pole, so that the ribbon has completed one orbit (circumference). Now rotate the ball about its axis and wrap again. Repeating this process in a way that minimized overlap at the equator, one would fully cover the ball after $C/2w$ wraps, where C is the circumference of the ball, and w is the width of the ribbon. With a wider ribbon (or swath), fewer orbits would be required to cover the ball.

A tradeoff arises here: holding constant other aspects of satellite and sensor design, the higher the frequency of repeat coverage, the lower the resolution. Thus, given current technology, satellites providing imagery at 1-meter resolution take weeks to obtain full Earth coverage, while kilometer-resolution sensors can provide daily global coverage. More regular coverage can be important. For example, Burgess et al. (2012) used 250-meter Moderate Resolution Imaging Spectroradiometer (MODIS) data with near-daily overpasses—unlike many other studies of deforestation based on 30-meter Landsat data with 16-day repeat overpasses—because they focused on rain forests that are covered by clouds most of the time. A given plot of land in such an environment might not have a cloud-free Landsat image for a year or more.

Sensors and Bands

A remote sensing satellite (or other platform) hosts one or more sensors, each of which may generate one or more independent data streams. A given sensor can observe energy in one or more bands, or ranges of the electromagnetic spectrum, corresponding to segments of the spectrum known as microwaves, infrared, ultraviolet, or visible light. Sensors emphasizing spectral resolution may record information about dozens (multispectral) or hundreds (hyperspectral) of bands. By comparison, the human eye can sense three bands, centered in the red, blue, and green regions of the spectrum.

Different bands, and combinations thereof, have different useful properties, as we will discuss in the applications below. As one example, plants reflect at different sets of frequencies at different stages of their life cycle. For this reason, functions of reflectance in specific portions of the visible and infrared spectra can provide information about vegetation growth. This insight is used to produce the commonly used Normalized Difference Vegetation Index (NDVI). Infrared data can also be used to measure temperature and, indirectly, precipitation when applied to clouds (for example, Novella and Thiaw 2012).

Most sensors used by social scientists are "passive," meaning that they observe energy emitted directly from the Earth or, more typically, reflected off the Earth from the Sun. In contrast, "active" sensors, such as radar and LiDAR, emit radiation

and record some property of that radiation when it is reflected back, typically its transit time. Radar is the technology behind the Shuttle Radar Topography Mission (SRTM), a common source of global elevation data recorded by the Space Shuttle in 2000. LiDAR is a technology similar to radar that uses reflectance of laser emissions rather than sound. Especially on aircraft, LiDAR is used to precisely measure elevation and building heights.

Intermediate Processing

Social scientists almost never use raw satellite data. Extensive processing is typically required to make data comparable and interpretable, and there is an active remote sensing literature developing and refining these methods and applying them to new and existing datasets. In remote sensing parlance, raw data are called Level 0 data, while data providers often release Level 1 or Level 2 data, with each higher level reflecting further processing. We briefly describe some key steps here.

Within a given swath, a satellite views the locations it is passing directly over (at nadir) from a right angle, but it is simultaneously viewing other locations at different angles. "Orthorectification" transforms the resulting images to account for these differences in angles, as well as an analogous distortion due to elevation differences. Similarly, raw data are typically in a reference frame corresponding to a swath, but they can be projected to the user's geographic projection of choice. The most common projection for distributing global remote sensing data, as well as many other datasets, is Plate Carrée—sometimes referred to as "latitude-latitude," or, confusingly, "unprojected" or "geographic"—in which a spherical degree of longitude is assigned the same Cartesian length as a degree of latitude. Of course, no projection of a spherical surface to a plane can preserve distances, areas, and directions. Plate Carrée distorts East-West distances (and therefore areas) in proportion to the cosine of latitude. In general, any such re-projection necessarily involves combining information from multiple input pixels.

Data for a given region of interest may be drawn from multiple overlapping swathes. Such data can be combined in a process sometimes known as "mosaicking," which can help in reducing random error (noise) or other unwanted features. For example, when investigating land cover, areas covered in clouds do not contain useful information (in bands for which clouds are opaque). Multiple images, captured on several different days, may be combined (that is, averaged or mosaicked) in order to obtain a cloud-free composite image.

Interpretation

While some applications discussed below involve relatively direct conversion of sensed quantities for individual pixels to physical quantities of interest (such as night lights, greenness, elevation, particulate concentration, or temperature), many others instead involve classifying pixels, each of which is a vector of quantities in different bands, into a discrete set of land cover categories. More complex forms of classification use contextual information (that is, information on neighboring pixels) to help

classify a given pixel. This potentially allows algorithms to identify shapes consisting of multiple pixels, typically known as objects, and find edges between them.[2] Objects might include cars, buildings, or agricultural plots. For any given object size, higher spatial resolution, more bands in relevant spectra, and a more frequent repeat cycle all provide more information to help with identifying the object.

Classification is generally subjective, in the sense that the same underlying data can be summarized in substantially different classifications. At the same time, classification is guided by decades of prior art in the remote sensing and computer science literature. While much of this work has been done on individual scenes on the order of hundreds or thousands of square kilometers, global datasets are increasingly common.[3] One broad distinction is between *unsupervised* classification, in which the data are partitioned into clusters without additional information, and *supervised* classification, which relies on training datasets to assign data to pre-specified classes. As an example of supervised classification that is particularly intensive in the use of external information, the Vogelmann et al. (2001) land cover data (used by Burchfield et al. 2006, as discussed below) supplements satellite data with aerial photography and census data to classify land cover and land use.

Sophisticated supervised classification systems increasingly rely on machine learning techniques. A critical input to these and other methods is the availability of training data on the variable of interest that assigns ground truth values to sample sites. For example, delineating imaged urban neighborhoods as residential, or even more specifically as slums, requires first providing a set of areas pre-defined as slums by other means. Doing so well requires a training dataset that reflects the full diversity of distinct neighborhoods within the category of slums. This is especially challenging when the object of interest is heterogeneous or imprecisely defined. But even for well-defined land cover, such as an individual crop species, differences in density, growth stage, solar angle, cloud cover, and neighboring land cover will generate a variety of signals, so that ground truth data reflecting the full heterogeneity will improve classification accuracy. Relatedly, training data should ideally be as close to the desired final outcome as possible. One could imagine economists using remotely sensed information on buildings to estimate a region's capital stock; in such a case, the ideal training data would concern building values instead of building types.

Because these training datasets are used to define the classes underlying a classification algorithm, they must be produced outside the algorithm. Thus, they are typically a labor-intensive analog constraint on a technology that otherwise can operate with all the scale benefits of computer processing.

---

[2] These methods are related to those used in image analysis more generally. See Glaeser, Kominers, Luca, and Naik (2015) and Naik et al. (2015) for two recent examples of economic analysis drawing on Google Streetview images.

[3] For a review of recent work in human-related land cover classification at the global scale, see Small and Sousa (2016).

Prominent Existing Data Sources

Economists have used data derived from many different satellites and sensors. Table 1 lists a few of the most prominent. A comprehensive database of essentially all declassified and unclassified remote sensing satellites ever launched except for the very smallest, complete with information on onboard sensors and application areas, can be found in the World Meteorological Organization's OSCAR database: http://www.wmo-sat.info/oscar/. A series of eight satellites in the Landsat program have been collecting data since 1972. The MODIS sensor has flown on two satellites since 1999. It provides coarser resolution data than Landsat, but has been useful in producing high-frequency imagery, and data on important phenomena including particulate pollution and fires. Night lights data from the Defense Meteorological Satellite Program–Operational Linescan System (DMSP-OLS), and Visible Infrared Imaging Radiometer Suite (VIIRS) sensors are even coarser, spatially and spectrally, but their simplicity has made them easy for economists to use. The 2000 Shuttle Radar Topography Mission (SRTM) provides the most comprehensive elevation data to date. Finally, a rapidly increasing trove of high-resolution images is being captured by private sector firms, including Spot Image and Digital Globe and, more recently, Planet and Terra Bella.

Auxiliary Gridded Datasets

Remote sensing datasets are typically recorded and stored in the form of large grids of pixels, with each pixel corresponding to a geographic location. Grids are more efficiently stored than points because once one corner, the cell size, and the overall dimensions are defined, no pixel-specific location information is required. Several other datasets, not collected by satellite, have been organized in the same gridded fashion to facilitate analysis in combination with the satellite data. For example, Gridded Population of the World is a lightly modeled population surface, while LandScan and WorldPop are more heavily modeled population datasets. G-Econ offers a set of modeled GDP estimates (Nordhaus 2006), while the FAO-IIASA's Global Agro-Ecological Zones (GAEZ) project from the Food and Agriculture Organization and the International Institute for Applied Systems Analysis has seen widespread use among agricultural economists.[4]

Applications of Remote Sensing Data in Economics

We now discuss a series of applications of remote sensing data that have been used in the economics literature to date—though our remit is to provide examples rather than to be comprehensive. For ease of reference, we organize our discussion

---

[4]See the Gridded Population of the World (GPW) at http://sedac.ciesin.columbia.edu/gpw/; LandScan at http://web.ornl.gov/sci/landscan/; WorldPop at http://www.worldpop.org.uk/; and Geographically based Economic data (G-Econ) at http://gecon.yale.edu/; and Global Agro-Ecological Zones (GAEZ) http://gaez.fao.org/.

Table 1

Main Satellite Data Sources Used in Economics

Source	Economics applications	Highest resolution	Pricing	Availability by year	Examples	For more information
Landsat	Urban land cover, beaches, forest cover, mineral deposits	30 m	Free	1972– (8 satellites)	Foster and Rosenzweig (2003); Faber and Gaubert (2015)	http://landsat.usgs.gov/
MODIS	Airborne pollution, fish abundance	250 m	Free	1999– (Terra); 2002– (Aqua)	Foster, Gutierrez, and Kumar (2009); Burgess et al. (2012)	http://modis.gsfc.nasa.gov/
Night lights (DMSP-OLS, VIIRS)	Income, electricity use	~1 km	Free digital annual (DMSP- OLS) and monthly (VIIRS) composites	Digital archive 1992–2013+ (VIIRS 2012–; film archive 1972–1991)	Henderson et al. (2012); Chen and Nordhaus (2011)	http://ngdc.noaa.gov/eog/
SRTM	Elevation, terrain roughness	30 m	Free	2000 (static)	Costinot, Donaldson, and Smith (2016) via Global Agro-Ecological Zones (GAEZ) data	https://lta.cr.usgs.gov/SRTM
DigitalGlobe (including Quickbird, Ikonos)	Urban land cover, forests	< 1 m	Not free	1999– (6 satellites)	Marx, Stoker, and Suri (2015); Jayachandran, de Laat, Lambin, and Stanton (2016)	https://www.digitalglobe.com/resources/satellite-information

around the type of physical phenomenon being considered in each application. In each case, the work of economists is based on decades of work on sensor and algorithm design and interpretation by researchers and practitioners in the remote sensing community. For brevity, we generally limit discussion to a few papers on each topic in the economics literature, and encourage interested readers to seek out the relevant remote sensing literature cited in those papers.

Night Lights

Total visible light emitted from Earth's surface at night has become a commonly used proxy for local economic activity. These data have been obtained primarily from the DMSP-OLS sensor, which offers nominal 1-kilometer resolution, and are now easy to access and process. They can plausibly be used as a proxy for economic activity under the assumption that lighting is a normal good. In this sense, the use of lights data follows the same logic behind a great deal of earlier work that uses data on consumption decisions to proxy for income under the assumption that Engel curves are stable (for example, Bils and Klenow 2001; Costa 2001; Young 2012).

While the cross-sectional correlation between lights and human activity has been noted since shortly after declassification of these data in 1972 (Croft 1973; Doll, Muller, and Morley 2006), the use of lights data expanded dramatically after the development of new processing methods and the distribution of a digital archive by the National Oceanic and Atmospheric Administration (NOAA) in the 1990s and 2000s (Elvidge et al. 1997). The accuracy of lights as a proxy for economic growth has been the subject of formal econometric scrutiny only recently, starting with Henderson, Storeygard, and Weil (2012). In an annual panel of countries from 1992 to 2008, as well as the corresponding long difference, these authors estimate a lights-GDP elasticity of 0.28 to 0.32, with no evidence of nonlinearity or asymmetry between increases and decreases in lights. In the long difference, the lights-GDP relationship has a correlation coefficient of 0.53. Under a range of assumptions about measurement error of GDP in countries with good data, they estimate a structural elasticity of lights growth with respect to GDP growth of between 1.0 and 1.7.[5]

Henderson, Storeygard, and Weil (2012) focused on aggregate country-level lights in order to make comparisons with traditional GDP data, harnessing the independence of errors between the two data sources. By contrast, a subsequent and innovative literature has used lights as a proxy for economic activity within fine geographic units for which no alternative data source is available. Such units include cities (Harari 2016; Storeygard 2016), ethnic homelands (Michalopoulos and Papaioannou 2013, 2014), subnational administrative units (Hodler and Raschky 2014), larger uniform grid squares (Henderson, Squires, Storeygard, and Weil 2016), and around natural features such as rivers (Bleakley and Lin 2012, as discussed earlier). Lee (2016) has also used the night lights data to estimate economic activity in North Korea, whose government produces no credible economic statistics. The many additional novel uses of night lights data in recent work are too voluminous to fully capture here.

Thus, even one of the most rudimentary sources of satellite data (luminosity at night, aggregated to annual frequency, with 1-kilometer resolution, at best, and just one useable band) has enabled economists to answer important questions in new and convincing ways. This only serves to underscore the great potential that lies ahead for the use of remote sensing data in economics.

Climate and Weather

A vast literature in economics has studied the impact of short-run weather fluctuations and longer-term climate trends on human activity (for a survey, see Dell, Jones, and Olken 2014). This weather–economy relationship is both of interest in its own right (for example, as an input into an improved understanding of the impacts of climate change) and as a source of plausibly exogenous income or cost variation in agricultural or resource-dependent settings.

---

[5] Chen and Nordhaus (2011) develop an alternative method and arrive at more pessimistic view of the lights data. For a bibliography of early lights work at NOAA, see http://ngdc.noaa.gov/eog/pubs_new.html.

But in many parts of the world, data-collecting weather stations are extremely sparse. Thus, researchers estimate measures of interest by combining three ingredients: weather data from meteorological stations (when available); remotely sourced readings that are available for a wider range of locations, such as cloud-cover and cloud-top temperature; and a climate model. In effect, the researcher uses a climate model and remotely sensed readings about certain meteorological phenomena that are not of direct interest (cloud-top temperature, for example) to interpolate intelligently between station-level readings of direct interest (on rainfall, for example).[6]

Two examples of recent work on Africa draw particularly strongly on the fine spatial resolution of this weather data. First, Harari and La Ferrara (2015) combine 1-degree gridded weather data with data on conflict events in Africa (the PRIO/Uppsala Armed Conflict Location and Event Dataset). The within-country spatial resolution of such datasets allows these authors to estimate spatial spillovers of conflict—how drought in one cell can cause armed conflict both in that cell and in other cells nearby—that are surprisingly large relative to the existing literature that had used country-level data. For example, these authors find that every conflict event in one subnational grid cell causes the risk of conflict in each of the eight adjacent grid cells to rise by 3.6 percentage points. Naturally, an additional advantage of satellite-sourced weather data in this setting is that ground-based weather reporting may be jeopardized during periods of conflict. Second, Kudamatsu, Persson, and Stromberg (2016) match weather data to the Demographic and Health Surveys in all 28 African countries in which GPS coordinates were recorded for individual survey clusters. This allows them to conduct a detailed investigation into the extent to which infant mortality in Africa is related to weather variation; in fact, the spatial precision involved allows them to estimate the extent to which weather affects infant mortality through particular channels such as malaria or malnutrition.

A number of other innovative weather attributes derived from satellite imagery have been used by economists. These have been incorporated into studies of the impact of flooding (Guiteras, Jina, and Mobarak 2015) and cyclones (Yang 2008; Hsiang and Jina 2014), as well as the use of wind speed and direction as a creative source of variation in predicting colonial settlement patterns (Feyrer and Sacerdote 2009) and the extent to which oceanic trade routes were disrupted by the adoption of steam shipping in the late nineteenth century (Pascali 2015). Guiteras, Jina, and Mobarak (2015) make the surprising observation that households do not recall inundation events very accurately, especially in areas where floods are routine, thereby highlighting the importance of the objectivity that satellites can provide.

Topography

High-resolution maps of elevation—and hence a reasonable approximation to topography—commonly known as "digital elevation models" (DEMs)—have

---

[6]Auffhammer, Hsiang, Schlenker, and Sobel (2013) provide details about these climate models and derived weather products, particularly as used by economists.

seen many recent applications to economic questions. Easily accessible DEMs have been produced globally at 30-meter resolution using remote sensing instruments, including the Shuttle Radar Topography Mission (SRTM) and the satellite-based Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER).[7] Such DEMs superseded an earlier generation of DEMs that were built from decades of aerial photography (via stereophotogrammetry) and centuries of surveying, providing globally uniform high resolution.

One prominent application of a digital elevation model comes from the urban literature. Following Burchfield et al. (2006), Saiz (2010) measures the geographic constraints on expansion that US cities face.[8] Drawing on data from the US Geological Survey, Saiz calculates the fraction of area within a 50-kilometer radius of a city that neither has a slope greater than 15 percent nor is covered by a body of water. This measure is important in its own right, but its widest use to date has been to provide a source of exogenous variation in urban land supply for the large literature that aims to estimate determinants of urban housing demand. Recent examples include Mian and Sufi (2014) on the extent to which housing decline caused unemployment and Hsieh and Moretti (2015) on the aggregate misallocative consequences of housing supply regulations.

Topographical variation can also be applied to study the economic impact of large infrastructure investments. For example, Duflo and Pande (2007) seek to estimate the effects of large dams in India, which are used for both irrigation and hydropower. In order to isolate plausibly exogenous variation in the spatial placement of dams, these authors argue that dams are most likely to be viable when they are built in locations in which the course of a river is neither too shallow nor too steep—locations that that can be identified using a digital elevation model. This geographical measure, interacted with further time-series determinants of the rate of dam-building, is indeed highly predictive of where dams were built in India over the past 50 years. The authors' resulting instrumental variable estimates of the impact of dams on poverty and agricultural productivity reveal that these controversial projects do not appear to generate enough economic benefit to justify their (substantial) construction costs. Dinkelman (2011) and Lipscomb, Mobarak, and Barham (2013) developed refinements of this approach in their studies of the impacts of electrification in South Africa and Brazil, respectively.

Remotely sourced topographic data also serve as an indirect input into a large body of work on the economics of agriculture. A number of studies model a production possibilities frontier, which requires that the researcher be able to predict the suitability of a given crop at a given location, regardless of whether that crop is actually grown in that location. Conventional output data cannot be used for such

---

[7]SRTM data are not available at high latitudes, beyond 60 degrees north and 54 degrees south. Prior to 2014, the US government only released SRTM data outside the United States at 90-meter resolution, but they are now available at full 30-meter resolution globally. ASTER elevation data are newer and have different limits and advantages, including availability at high latitudes.

[8]Pioneering work by Qian (2008) and Nunn and Puga (2012) also aims to understand how a region's terrain ruggedness impacts its economic characteristics.

a purpose but agronomic models, such as GAEZ, enable such predictions by using characteristics of the location (including its topography, from a remotely sourced digital elevation model) and crop as modeling inputs. The result is a prediction about the yield that could potentially be achieved for any crop at any location on Earth. As discussed above, Costinot, Donaldson, and Smith (2016) draw on such predictions about the global agricultural production possibility frontier (both in the current climate and in the future climate predicted by climatologists) to study economic consequences of climate change in an integrated world economy. Earlier versions of the GAEZ data were used in Nunn and Qian's (2011) study of the impact of the discovery of the potato in the New World on Old World living standards and Alesina, Giuliano, and Nunn's (2013) test of the hypothesis that modern-day gender roles have roots in traditional agricultural practices such as the use of the plough. In both cases, the GAEZ model provides a unique input into the agricultural possibilities of any given location, regardless of whether those possibilities are being exploited at a given point in time.

Agricultural Land Use and Crop Choice

Satellite data have been used with rapidly improving accuracy to measure aspects of agricultural land use. Two recent studies draw on the USDA's Cropland Data Layer (CDL), which classifies 30-meter pixels throughout the United States into crop-specific categories, to evaluate some of the incentives that American farmers face. First, Scott (2014) investigates how US agricultural policy alters dynamic incentives for the preparation of land for crop agriculture. Given the lumpy nature of dynamic cultivation decisions, this question would be far more challenging without such high spatial precision. Holmes and Lee (2012) harness the CDL's crop-specific nature to estimate the extent to which crop choice patterns within land ownership parcels are more correlated than those across such parcels, as would be expected if economies of density at the farm level are substantial.

Another area of focus aims to measure the volume of agricultural output at a location. The simplest such efforts—used, for example, by the Kudamatsu, Persson, and Stromberg (2016) study referred to above—aim to capture a broad measure of agricultural productivity, which is typically done by measuring a normalized difference vegetation index (NDVI). The challenge here is that remote sensors capture increase in greenness, which may include growth in nonagricultural plants that are not the user's focus. For this reason, the use of NDVI data in agricultural applications can be enhanced with land-use classification that identifies cropland. Improved measures of remotely sensed yield measurement—such as those used in the production function estimates from Nebraska by Farmaha et al. (2016)—suggest that future applications in this area could be particularly rich.

Urban Land Use

Perhaps the starkest distinction in land use is between urban and nonurban areas. Small, Pozzi, and Elvidge (2015) and Small, Elvidge, Balk, and Montgomery (2011) carried out early work on urban extent using the night lights data. In a

path-breaking study of the determinants of urban sprawl, Burchfield et al. (2006) draw on two datasets, Vogelmann et al. (2001) and US Environmental Protection Agency (1994), that classify US land cover in 1992 and approximately 1976, respectively. After careful integration of the two datasets, these authors can track both urban (residential, commercial, and industrial development and transportation networks) and nonurban (water, bare rock and sand, forest, range and grassland, agricultural land, and wetlands) land cover categories over time. The resulting 30-meter data allow Burchfield et al. (2006) to document the striking extent to which recent US urban development has taken place in grid cells in close proximity to previously developed cells. Urban growth appears to have taken place with very little spatial "leapfrogging" across undeveloped cells.

Another classic topic in urban economics concerns the incidence of land use regulation policy. Turner, Haughwout, and van der Klaauw (2014) use the satellite-derived (30-meter resolution) National Land Use and Land Cover dataset (NLCD) in order to measure the impact of municipal land use regulations in the United States (as measured by the Wharton Land Use Regulation data). From a spatial regression discontinuity analysis across straight-line municipal boundaries, these authors conclude that a one standard deviation increase in regulations reduces land values by about one-third. In addition, the richness of the data allows Turner, Haughwout, and van der Klaauw (2014) to estimate spatially decaying spillover effects of land regulations (which appear to be small).

Building Types

An active research frontier is using remote sensing data to identify individual buildings, and to classify these buildings by type. For example, Casaburi and Troiano (2016) highlight an innovative use of aerial photography by the Italian government to find buildings not registered for property tax collection. Though these authors do not use the photographs directly, their analysis draws on the resulting counts of unregistered buildings obtained from aerial photographs to study the political ramifications of the intensity with which this program was used to prevent tax evasion.

Two recent studies from Nairobi have made important progress with extremely high (0.5-meter) resolution imagery. First, as described above, Marx, Stoker, and Suri (2015) use satellite data to detect the new, or better maintained, corrugated iron roofs that are indicative of higher-quality housing. Interestingly, even though these authors had also conducted a household survey with rich attribute information, they find that the satellite-derived measure of housing quality is a useful additional measure because it is not subject to recall bias. Second, Henderson, Regan, and Venables (2016) rely on imagery spanning multiple years to consider a dynamic model of city land use with formal and informal sectors. These authors combine data on building heights and imagery from aerial photography and satellites on building and road footprints in order to consider changes in building volumes in the data-poor environment of this low-income city. This study is unique because the novel data employed allow the authors to test the implications that urban models

typically have for building density with direct data about buildings, rather than with indirect proxies such as population.

Natural Resources

Large-scale deforestation has been quantified via satellite since at least Skole and Tucker (1993). But the first application of such data in an economic model came from work by Foster and Rosenzweig (2003), who harnessed three decades of Landsat data in seeking to understand the rise in forest cover that India (in common with many other countries) has seen since the late 1950s. These authors combine satellite data with village-level survey data to document that forest growth appears to be caused by growth in the demand for forest products (such as wood fuel) due to income growth, and not by supply-side factors (such as land and labor prices) in the forestry sector.

Other work has focused on the ability for satellite data to monitor illegal logging activity, as discussed in our introduction. Two recent studies have used innovative research designs in this context. First, Burgess, Costa, and Olken (2016) demonstrate, via a spatial regression discontinuity design using 30-meter resolution satellite data, just how suddenly Brazil's 2006 anti-deforestation policy affected land use choices there. Second, Jayachandran, de Laat, Lambin, and Stanton (2016) combine a randomized controlled trial with bespoke satellite images (commissioned from the QuickBird satellite by the authors) and algorithms that are sufficiently rich as to detect individual trees owned by farmers in Uganda. In doing so, they document the powerful impact of a payment-for-ecosystem services program in reducing deforestation.

The possibilities for tourism are another resource that can be inferred from remotely sensed data. Faber and Gaubert (2016) pursue such a strategy in order to estimate the effect of tourism on economic development. They isolate plausibly exogenous variation in tourism supply across municipalities in Mexico by focusing on a metric of beach quality, which can be constructed by isolating the distinctive features (in six Landsat bands) of four of Mexico's top-ranked beaches in an industry journal, and then searching for other areas of coastline with similar spectral characteristics. This measure of beach quality is a strong predictor of the relative extent of tourism along Mexico's coastline (excluding the top four beaches) today. A surprising result that emerges when the authors use this variation is that tourism appears to have positive effects on the size of even tradable, nontourism sectors (such as manufacturing) in the same municipality.

Finally, Axbard (2016) documents a connection between income shocks due to reduced fishing opportunities and increased piracy in Indonesia using remotely-detected data on chlorophyll and building on the established correlation (in the marine biology of this study region) between chlorophyll and fish abundance. This sea-based data proves useful in this study because it varies in a way that is orthogonal to land-based incomes and policies. Furthermore, chlorophyll does not have a direct effect on other actors; critically in the context of studying piracy, it does not directly constrain sea travel for cargo or enforcement.

Pollution Monitoring

Environmental monitoring is another promising research area, particularly involving air conditions. Ground-based air pollution monitoring stations are not widespread in developing countries, and they are potentially subject to government manipulation. For this reason, as discussed above, Jayachandran (2009) draws on remote sensing work establishing the relationship between aerosol optical depth from the satellite-based Total Ozone Mapping Spectrometer and standard measures of airborne particulate matter in order to measure the air pollution caused by forest fires in Indonesia. Foster, Gutierrez, and Kumar (2009) document the effect of air pollution, measured similarly using MODIS data, on infant mortality, using an instrumental variable based on a voluntary certification policy applied to manufacturing firms. Chen, Jin, Kumar, and Shi (2013) and Bombardini and Li (2016) evaluate potential causes of air pollution (government policy and international trade patterns, respectively), and are able to compare the satellite data with ground-based monitoring systems in a context where pollution is a politically contentious matter.

Combinations of Sources

An exciting research avenue concerns the use of broad sets of satellite data, combined with techniques from machine learning, to obtain predictors of a phenomenon of interest. This type of work has not yet seen much application in economics. But in one important recent example, Xie, Jean, Burke, Lobell, and Ermon (2016) develop high-resolution predictors of household income, wealth, and poverty rates in five African countries. Their algorithm combines information from a standard image classification database (drawing on both night lights data and 3-meter resolution daytime imagery) with real measures of consumption and assets, available for larger spatial units, from national-level Demographic and Health Surveys and the Living Standards Measurement Survey. This method has the potential to provide estimates of economic well-being at spatial and temporal frequencies that are an order of magnitude higher than those that are commonly available. These estimates could then be used as dependent or independent variables in future research, at least subject to the caveat that these estimates would be problematic for use in a study involving any variable that was itself used as an input into these well-being estimates.

Potential Pitfalls in the Use of Remote Sensing Data

Economists who use these and other remote sensing data will face some challenges that are unique to these data, and some that arise in other settings as well. Here, we focus on issues concerning dataset size, spatial dependence, measurement error, and privacy concerns. For an overview of the econometric issues associated with spatial data and large datasets, useful starting points are Gibbons, Overman, and Pattacchini (2015) and Varian (2014), respectively.

One set of challenges arises from the complexity of many remote sensing data-sets. The land area on Earth can be subdivided into hundreds of billions of 30-meter grid cells, and hundreds of trillions of 1-meter cells. Such high-dimensional data can be difficult to model, even with the simplest linear functions. Yet remotely sensed dependent variables are often discrete or truncated in some way, so linear analysis is unlikely to be appropriate. Relatedly, part of the excitement of the new remote sensing data lies in its promise for informing and disciplining economic theory, but solving even simple models with the sort of spatial heterogeneity afforded by enormous datasets will require advances in the tools that economists use to solve for equilibria in high-dimensional spatial models. Some recent progress has been made in this area, including Allen and Arkolakis (2014), Krusell and Smith (2015), and Desmet, Nagy, and Rossi-Hansberg (2015).

Satellite data often display substantial spatial dependence: local correlation among nearby units. Here an important distinction is whether the data are being used as a dependent or independent variable. When spatial dependence appears in the dependent variable, then the error term in a multivariate regression that seeks to explain that variable will no longer be distributed independently. As a result, statistical inference should be conducted with attention to the dependence across units; for example, Pinkovskiy (2013) discusses how this can be done in the context of regression discontinuity designs when using remote sensing data.

When the remote sensing data is used as an independent variable, then estimates of all the regression coefficients are potentially biased (if one hopes to use them to make counterfactual predictions), unless adjustments are made to allow for the possibility that treatment in one spatial unit could have a causal effect on outcomes not just in that unit, but also on other units. For example, as Bleakley and Lin (2012) show, the causal effect of historical portage sites on the location of economic activity today is clearly evident from Figure 4. But, as these authors make clear, it would be wrong to conclude from this evidence that aggregate US economic activity would be higher today if only there had been more sites for portage in 1800—it is far more likely that when activity concentrated at portage sites, it displaced activity elsewhere. In more technical terms, the problem of spatial dependence here is a violation of the so-called "stable unit treatment value assumption," as in Rubin (1990). Recent examples of strategies for addressing this bias include Crepon et al. (2013), Acemoglu, Garcia-Jimeno, and Robinson (2015), and Donaldson and Hornbeck (2016).

These concerns about spatial dependence arise in many kinds of geographical data, but they are likely to be more acute at high spatial resolution. For example, for most questions that arise in spatial economics, the spatial dependency between the cities of Boston and New York is likely to be weaker than that between two neighboring 30-meter grid cells within Boston. In an agricultural context, if a region faces a completely inelastic demand curve for a given crop, then the total amount of that crop produced in the region will be fixed by the demand side, implying that any increase in production by one grid cell must displace, one-for-one, production by some other grid cell in the region.

Related to the issue of spatial dependence is the importance of understanding the nature of spatial relationships in any gridded dataset. There can be a tendency to equate dataset resolution with information density, but this intuition may not hold true in certain cases, especially for derived products that combine input data sources with multiple resolutions.[9] Relatively raw data can also exhibit mechanical spatial correlation. For example, the commonly used night lights data are distributed as a grid of cells that are one kilometer wide at the Equator and decreasing with the cosine of latitude, but each of these cells records light for a footprint roughly 3–5 kilometers in diameter, so in practice measurements are autocorrelated at that distance.[10]

Derived products combining multiple inputs also require care in overall interpretation. For example, across various literatures in natural and social science journals, nighttime lights have been used as inputs to gridded maps of income, poverty, electricity, impervious surface, in-use metal stocks, carbon emissions, and rice consumption. While each of these quantities is surely positively correlated with night lights, at least at the national level, the assumptions underlying each use are different and not always explicitly spelled out.

Measurement error raises a different set of concerns. Classification methods are generally subjective data reduction exercises, albeit ones refined by decades of algorithm development. False positives and false negatives are to be expected from any such exercise. While assessing and minimizing such errors is the stock-in-trade of remote sensing science, it is important to recognize that the types of errors that would seriously jeopardize a typical economic application of remotely sensed data are not always a main concern of the remote sensing community. Such issues can arise even when the relevant data product reports a relatively well-defined physical quantity. For example, the elevation recorded in the SRTM dataset generally corresponds to a point in the tree canopy, rather than the ground, in areas where tree cover is particularly dense (Hofton, Blair, and Rabine 2006). Thus, in applications for which ground elevation differences on the order of tree heights matter, there is a potentially problematic correlation between the measurement error in elevation and the type of local tree cover.

---

[9] An example from agriculture is the M3 Crops Database (Monfreda, Ramankutty, and Foley 2008). This product combines the best publicly available data on agricultural output (by crop) for relatively large subnational land units (for example, counties in the US case) with satellite data on the location of crop agriculture (relative to noncrop agriculture). Assuming that production is allocated proportionally (within each crop) to each grid cell within each subnational unit, this dataset provides interpolated estimates of output by crop at the resolution of the satellite data. The MapSPAM database (You et al. 2014) goes further and uses the GAEZ model described earlier, along with FAO price data at the national level, to assign crop totals to grid cells by revenue-maximization rather than pure proportionality. The result in both cases is a high-resolution proxy for crop-level output around the world, but neither dataset presents raw data by crop at the grid cell-level.

[10] A newer generation of sensor, the Visible Infrared Imaging Radiometer Suite, launched on the Suomi National Polar Partnership satellite in 2011, has a true sub-kilometer footprint and has the potential to overcome some other limitations of the DMSP-OLS sensor, though current versions have more noise, and the overpass time is much later, at roughly 1:30 AM (Elvidge, Baugh, Zhizhin, and Hsu 2013; Chen and Nordhaus 2015).

Even within the same data product or class of products, successive overpasses of the same location will in general be at (at least) slightly different times of day, angle of view, solar angle, and atmospheric conditions. The orbit of an individual satellite generally decays over its lifetime, sensors degrade, and replacement sensors, even within the same series, operate with subtle differences. Whether these differences affect interpretation will depend on the nature of the study. On the plus side, there are many contexts where error in remotely sensed measurements is plausibly orthogonal to error in human-collected data, so that the two can be used to instrument for each other to reduce measurement error. And while the continuous data underlying discrete datasets are often not made available, in principle they could be, enhancing economists' ability to consider classification error in an econometric framework.

Overall, it is essential for economists using remote sensing data to understand, and be skeptical of, the data and assumptions underlying them. Increased communication, and in some cases active collaboration, between economists and remote sensing scientists will be important in ensuring that the next wave of this kind of economics research is fruitful.

Finally, we note that at very high (submeter) resolution, satellites increasingly provide information that, if disclosed, might violate reasonable expectations of privacy. If such information comes from an image that a researcher orders from a satellite data provider (as opposed to purchasing from an inventory of already public images), thereby making it (potentially) public, this raises ethical and regulatory concerns related to human subjects research.

Concluding Remarks: The Promise of Satellite Data for Research in Economics

The cheap availability of high-resolution satellite imagery has already become a powerful tool in economics, just as it has in other research fields and in the private sector.[11] And costs are likely to continue to fall as entry in the satellite sensing industry has accelerated of late. Cheaper microsatellites (up to 70 or 100 kilograms), nanosatellites (up to 10 kilograms), and even picosatellites (up to 1 kilogram) are becoming more prevalent and more useful due to increasingly powerful lightweight components. Even among the mostly larger satellites covered in the OSCAR database, launches have been steadily increasing from the 5–7 launches per year that were common in the 1970s and 1980s to 15–18 launches per year from 2010–2015 (World Meteorological Organization OSCAR database at http://www.wmo-sat.info/oscar/). Indeed, such launches are becoming commonplace—within 48 hours of the date in September 2016 on which we happen to be writing this

---

[11] For example, Nagaraj (2015) shows—remarkably—that regions of the world with high idiosyncratic cloud cover at the instants the heavily used Landsat satellite was overhead have seen less local gold discovery in recent decades.

sentence, 11 satellites, each with best resolution near or below one meter, operated by two firms, are scheduled to be launched on three different rockets.

As the data collected by these orbiting machines increasingly offer greater spatial, temporal, and spectral resolution, and the algorithms and processing power used to interpret these data become more and more sophisticated, new potential applications are increasingly becoming feasible. We conclude with a few such examples.

In studying urban areas, aerial photography and some satellites can be used to count crowds of people and cars at events such as protests, political rallies, or peak shopping periods (for example, Kearns 2015; Swanson 2015). Similarly, traffic volumes could be measured from snapshots of car densities. This approach might be especially useful in areas including border controls and ports. Facebook is in the process of producing a map of buildings for several countries as a proxy for population density using billions of 0.5-meter resolution images as inputs (Simonite 2016). Measurements of disturbed soil or impervious materials could potentially track construction on an annual or even monthly basis.

In agricultural and resource economics, higher spectral resolution and continued algorithm development might allow for more automated distinctions between crops and noncrop vegetation, and increased temporal frequency may allow tracking of crop growth within seasons (Xie, Sha, and Yu 2008; Lobell 2013). NASA's GRACE satellite provides spatially coarse (1-degree) measures of soil moisture and groundwater depletion, via small changes in the local gravitational field. Economically important minerals located at the Earth's surface also have distinctive spectral and topographic properties that are routinely exploited in geology (for early and more recent reviews, see Goetz and Rowan 1981 and van der Meer et al. 2012, respectively).

New opportunities in environmental economics will also become available. Pollutants other than particulates, such as nitrogen dioxide, can be detected remotely (Geddes, Martin, Boys, and van Donkelaar 2015). Invasive plant species might be detected much like crops (Huang and Asner 2009). Precise trends in global forest cover change can be used in climate modelling (Hansen et al. 2013). Changes in sea level and its effects on coastal communities and ecosystems can be monitored from space.

As one example of many other possibilities, satellite imagery has been used for real-time conflict monitoring, for example in Darfur and Sri Lanka (UNITAR 2011), but not yet in long-term analysis of conflict, as far as we are aware. This research agenda for economists and other social scientists based on remote sensing data has made great strides in the last decade, but it seems safe to say that the real excitement lies ahead.

References

Acemoglu, Daron, Camilo Garcia-Jimeno, and James Robinson. 2015. "State Capacity and Economic Development: A Network Approach." *American Economic Review* 105(8): 2364–2409.

Alesina Alberto, Paola Giuliano, and Nunn Nathan. 2013. "On the Origins of Gender Roles: Women and the Plough." *Quarterly Journal of Economics* 128(2): 469–530.

Allen, Treb, and Costas Arkolakis. 2014. "Trade and the Topography of the Spatial Economy." *Quarterly Journal of Economics* 129(3): 1085–1140.

Auffhammer, Maximilian, Solomon M. Hsiang, Wolfram Schlenker, and Adam Sobel. 2013. "Using Weather Data and Climate Model Output in Economic Analyses of Climate Change." *Review of Environmental Economics and Policy* 7(2): 181–98.

Axbard, Sebastian. 2016. "Income Opportunities and Sea Piracy in Indonesia: Evidence from Satellite Data." *American Economic Journal: Applied Economics* 8(2): 154–94.

Bils, Mark, and Peter J. Klenow. 2001. "Quantifying Quality Growth." *American Economic Review* 91(4): 1006–30.

Bleakley, Hoyt, and Jeffrey Lin. 2012. "Portage and Path Dependence." *Quarterly Journal of Economics* 127(2): 587–644.

Bombardini, Matilde, and Bingjing Li. 2016. "Trade, Pollution and Mortality in China." February 26. http://www.sfu.ca/content/dam/sfu/economics/Documents/SeminarPapers/Matilde_Bombardini_sem_pap_March_1_2016.pdf.

Burchfield, Marcy, Henry G. Overman, Diego Puga, and Matthew A. Turner. 2006. "Causes of Sprawl: A Portrait from Space." *Quarterly Journal of Economics* 121(2): 587–633.

Burgess, Robin, Francisco Costa, and Benjamin A. Olken. 2016. "The Power of the State: National Borders and the Deforestation of the Amazon." Unpublished paper, London School of Economics.

Burgess, Robin, Matthew Hansen, Benjamin A. Olken, Peter Potapov, and Stefanie Sieber. 2012. "The Political Economy of Deforestation in the Tropics." *Quarterly Journal of Economics* 127(4): 1707–54.

Campbell, James B., and Randolph H. Wynne. 2011. *Introduction to Remote Sensing*, 5th edition. New York: The Guildford Press.

Casaburi, Lorenzo, and Ugo Troiano. 2016. "Ghost-House Busters: The Electoral Response to a Large Anti-Tax Evasion Program." *Quarterly Journal of Economics* 131(1): 273–314.

Chen, Xi, and William D. Nordhaus. 2011. "Using Luminosity Data as a Proxy for Economic

Statistics." *PNAS* 108(21): 8589–94.

Chen, Xi, and William D. Nordhaus. 2015. "A Test of the New VIIRS Lights Data Set: Population and Economic Output in Africa." *Remote Sensing* 7(4): 4937–47.

Chen, Yuyu, Ginger Zhe Jin, Naresh Kumar, and Guang Shi. 2013. "The Promise of Beijing: Evaluating the Impact of the 2008 Olympic Games on Air Quality." *Journal of Environmental Economics and Management* 66(3): 424–43.

Costa, Dora. 2001. "Estimating Real Income in the United States from 1888 to 1994: Correcting CPI Bias Using Engel Curves." *Journal of Political Economy* 109(6): 1288–1310.

Costinot, Arnaud, Dave Donaldson, and Cory Smith. 2016. "Evolving Comparative Advantage and the Impact of Climate Change in Agricultural Markets: Evidence from 1.7 Million Fields Around the World." *Journal of Political Economy* 124(1): 205–248.

Crepon, Bruno, Esther Duflo, Marc Gurgand, Roland Rathelot, and Philippe Zamora. 2013. "Do Labor Market Policies Have Displacement Effects? Evidence from a Clustered Randomized Experiment." *Quarterly Journal of Economics* 128(2): 531–80.

Croft, Thomas A. 1973. "Burning Waste Gas in Oil Fields." *Nature* 245: 375–76

Dell, Melissa, Benjamin F. Jones, and Benjamin A. Olken. 2014. "What Do We Learn from the Weather? The New Climate–Economy Literature." *Journal of Economic Literature* 52(3): 740–98.

Desmet, Klaus, Dávid Kriaztián Nagy, and Esteban Rossi-Hansberg. 2015. "The Geography of Development: Evaluating Migration Restrictions and Coastal Flooding." NBER Working Paper 21087.

Dinkelman, Taryn. 2011. "The Effects of Rural Electrification on Employment: New Evidence from South Africa." *American Economic Review* 101(7): 3078–3108.

Doll, Christopher N. H., J.-P. Muller, and Jeremy G. Morley. 2006. "Mapping Regional Economic Activity from Night-time Light Satellite Imagery." *Ecological Economics* 57(1): 75–92.

Donaldson, Dave, and Richard Hornbeck. 2016. "Railroads and American Economic Growth: A 'Market Access' Approach." *Quarterly Journal of Economics* 131(2): 799–858.

Duflo, Esther, and Rohini Pande. 2007. "Dams." *Quarterly Journal of Economics* 122(2): 601–46.

Elvidge, Christopher D., Kimberly E. Baugh, Eric A. Kihn, H. W. Kroehl, E. R. Davis. 1997. "Mapping of City Lights Using DMSP Operational

Linescan System Data." *Photogrammetric Engineering and Remote Sensing* 63: 727–34.

Elvidge, Christopher D., Kimberly E. Baugh, Mikhail Zhizhin, and Feng-Chi Hsu. 2013. "Why VIIRS Data Are Superior to DMSP for Mapping Nighttime Lights." *Proceedings of the Asia-Pacific Advanced Network* 35: 62–69.

Faber, Benjamin, and Cecile Gaubert. 2016. "Tourism and Economic Development: Evidence from Mexico's Coastline." September. https://www.dropbox.com/s/w08bdts16g19d3k/Faber_Gaubert_Tourism.pdf.

Farmaha, Bhupinder S., David B. Lobell, Katherine E. Boone, Kenneth G. Cassman, Haishun S. Yang, and Patricio Grassini. 2016. "Contribution of Persistent Factors to Yield Gaps in High-Yield Irrigated Maize." *Field Crops Research* 186(1): 124–32.

Feyrer, James, and Bruce Sacerdote. 2009. "Colonialism and Modern Income: Islands as Natural Experiments." *Review of Economics and Statistics* 91(2): 245–62.

Foster, Andrew, Emilio Gutierrez, and Naresh Kumar. 2009. "Voluntary Compliance, Pollution Levels, and Infant Mortality in Mexico." *American Economic Review* 99(2): 191–97.

Foster, Andrew D., and Mark R. Rosenzweig. 2003. "Economic Growth and the Rise of Forests." *Quarterly Journal of Economics* 118(2): 601–637.

Geddes, Jeffrey A., Randall V. Martin, Brian L. Boys, and Aaron van Donkelaar. 2015. "Long-Term Trends Worldwide in Ambient NO_2 Concentrations Inferred from Satellite Observations for Exposure Assessment." *Environmental Health Perspectives* 124(3): DOI:10.1289/ehp.1409567.

Gibbons, Steve, Henry G. Overman, and Eleonora Patacchini. 2015. "Spatial Methods." In *Handbook of Regional and Urban Economics*, vol. 5, edited by Gilles Duranton, J. Vernon Henderson, and William C. Strange, 115–68. North-Holland, Amsterdam.

Glaeser, Edward L., Scott Duke Kominers, Michael Luca, and Nikhil Naik. 2015. "Big Data and Big Cities: The Promises and Limitations of Improved Measures of Urban Life." NBER Working Paper 21778.

Goetz, Alexander F. H., and Lawrence C. Rowan. 1981. "Geologic Remote-Sensing." *Science* 211(4484): 781–91.

Guiteras, Raymond, Amir Jina, and A. Mushfiq Mobarak. 2015. "Satellites, Self-Reports, and Submersion: Exposure to Floods in Bangladesh." *American Economic Review* 105(5): 232–36.

Hansen M.C., P. V. Potapov, R. Moore, M. Hancher, S. A. Turubanova, A. Tyukavina, D. Thau, S. V. Stehman, S. J. Goetz, T. R. Loveland, A. Kommareddy, A. Egorov, L. Chini, C. O. Justice, and J. R. G. Townshend. 2013. "High-Resolution Global Maps of 21st-Century Forest Cover Change." *Science* 342(6160): 850–53.

Harari, Mariaflavia. 2016. "Cities in Bad Shape; Urban Geometry in India." February. http://real.wharton.upenn.edu/~harari/Harari_Papers/CityShapeHarariMarch2016_updated.pdf.

Harari, Mariaflavia, and Eliana La Ferrara. 2015. "Conflict, Climate and Cells: A Disaggregated Analysis." http://real.wharton.upenn.edu/~harari/Harari_Papers/Harari_Climate.pdf

Henderson, J. Vernon, Tanner Regan, and Anthony J. Venables. 2016. "Building the City: Sunk Capital, Sequencing, and Institutional Frictions." CEPR Discussion Paper 11211.

Henderson, J. Vernon, Tim L. Squires, Adam Storeygard, and David N. Weil. 2016 "The Global Spatial Distribution of Economic Activity: Nature, History, and the Role of Trade." April. NBER Working Paper 22145.

Henderson, J. Vernon, Adam Storeygard, and David N. Weil. 2012. "Measuring Economic Growth from Outer Space." *American Economic Review* 102(2): 994–1028.

Hodler, Roland, and Paul A. Raschky. 2014. "Regional Favoritism." *Quarterly Journal of Economics* 129(2): 995–1033.

Hofton, Michelle A., Ralph O. Dubayah, J. B. Blair, and D. Rabine. 2006. "Validation of SRTM Elevations over Vegetated and Non-vegetated Terrain Using Medium Footprint Lidar. *Photogrammetric Engineering and Remote Sensing* 72(3): 279–85.

Holmes, Thomas J., and Sanghoon Lee. 2012. "Economies of Density versus Natural Advantage: Crop Choice on the Back Forty." *Review of Economics and Statistics* 94(1): 1–19.

Hsiang, Solomon M., and Amir S. Jina. 2014. "The Causal Effect of Environmental Catastrophe on Long-Run Economic Growth: Evidence From 6,700 Cyclones." NBER Working Paper 20352.

Hsieh, Chang-Tai, and Enrico Moretti. 2015. "Why Do Cities Matter? Local Growth and Aggregate Growth." http://eml.berkeley.edu/~moretti/growth.pdf.

Huang, Cho-ying, and Gregory P. Asner. 2009. "Applications of Remote Sensing to Alien Invasive Plant Studies." *Sensors* (Basel, Switzerland) 9(6): 4869–89.

Jayachandran, Seema. 2006. "Air Quality and Early-Life Mortality: Evidence from Indonesia's Wildfires." Unpublished manuscript, January.

Jayachandran, Seema. 2009. "Air Quality and Early-Life Mortality: Evidence from Indonesia's Wildfires." *Journal of Human Resources* 44(4): 916–54.

Jayachandran, Seema, Joost de Laat, Eric F. Lambin, and Charlotte Y. Stanton. 2016. "Cash for Carbon: A Randomized Controlled Trial of

Payments for Ecosystem Services to Reduce Deforestation." Unpublished manuscript, Northwestern University, July.

Jensen, John R. 2006. *Remote Sensing of the Environment: An Earth Resource Perspective,* 2nd edition. Pearson.

Kearns, Jeff. 2015. "Satellite Images Show Economies Growing and Shrinking in Real Time." *Bloomberg Business*, July 8. http://www.bloomberg.com/news/features/2015-07-08/satellite-images-show-economies-growing-and-shrinking-in-real-time.

Krusell, Per, and Anthony Smith. 2015. "A Global Climate-Economy Model with High Regional Resolution." Unpublished paper, IIES Stockholm.

Kudamatsu, Masayuki, Torsten Persson, and David Stromberg. 2016. "Weather and Infant Mortality in Africa." Unpublished manuscript, IIES, August.

Lee, Yong Suk. 2016. "International Isolation and Regional Inequality: Evidence from Sanctions on North Korea." https://web.stanford.edu/~yonglee/NKsanctions_100516.pdf.

Lillesand, Thomas, Ralph W. Kiefer, and Jonathan Chipman. 2015. *Remote Sensing and Image Interpretation,* 7th Edition. Wiley

Lipscomb, Molly, A. Mushfiq Mobarak, and Tania Barham. 2013. "Development Effects of Electrification: Evidence from the Topographic Placement of Hydropower Plants in Brazil." *American Economic Journal: Applied Economics* 5(2): 200–231.

Lobell, David B. 2013. "The Use of Satellite Data for Crop Yield Gap Analysis." *Field Crops Research* 143(1): 56–64.

Marx, Benjamin, Thomas M. Stoker, and Tavneet Suri. 2015 "There is No Free House: Ethnic Patronage in a Kenyan Slum." http://www.mit.edu/~tavneet/Marx_Stoker_Suri.pdf.

Mian, Atif, and Amir Sufi. 2014. "What Explains the 2007–2009 Drop in Employment?" *Econometrica* 82(6): 2197–2223.

Michalopoulos, Stelios, and Elias Papaioannou. 2013. "Pre-colonial Ethnic Institutions and Contemporary African Development." *Econometrica* 81(1): 113–52.

Michalopoulos, Stelios, and Elias Papaioannou. 2014. "National Institutions and Subnational Development in Africa." *Quarterly Journal of Economics* 129(1): 151–213.

Monfreda, C., N. Ramankutty, and J. A. Foley. 2008. "Farming the Planet: 2. Geographic Distribution of Crop Areas, Yields, Physiological Types, and Net Primary Production in the Year 2000." *Global Biochemical Cycles* 22(1). DOI: 10.1029/2007GB002947.

Monmonier, Mark. 2002. "Aerial Photography at the Agricultural Adjustment Administration: Acreage Controls, Conservation Benefits, and Overhead Surveillance in the 1930s." *Photogrammetric Engineering and Remote Sensing* 76(11): 1257–61.

Nagaraj, Abhishek. 2015. "The Private Impact of Public Maps—Landsat Satellite Imagery and Gold Exploration." http://web.mit.edu/nagaraj/files/nagaraj_jmp_nov6.pdf.

Naik, Nikhil, Scott Duke Kominers, Ramesh Raskar, Edward L. Glaeser, and César A. Hidalgo. 2015. "Do People Shape Cities, or Do Cities Shape People? The Co-evolution of Physical, Social, and Economic Change in Five Major U.S. Cities." October. NBER Working Paper 21620.

Nicholas Novella, and Wassila Thiaw. 2012. "Africa Rainfall Climatology Version 2." National Oceanic and Atmospheric Administration, Washington.

Nordhaus, William D. 2006. "Geography and Macroeconomics: New Data and New Findings." *PNAS* 103(10): 3510–17.

Nunn, Nathan, and Diego Puga. 2012. "Ruggedness: The Blessing of Bad Geography in Africa." *Review of Economics and Statistics* 94(1): 20–36.

Nunn, Nathan, and Nancy Qian. 2011. "The Potato's Contribution to Population and Urbanization: Evidence from a Historical Experiment." *Quarterly Journal of Economics* 126(2): 593–650.

Pascali, Luigi. 2015. "The Wind of Change: Maritime Technology, Trade and Economic Development." November, http://www.econ.upf.edu/~pascali/LUIGI_PASCALI/Research_files/sail%2009-11-2015.pdf.

Pinkovskiy, Maxim L. 2013. "Economic Discontinuities at Borders: Evidence from Satellite Data on Lights at Night." March. https://www.newyorkfed.org/medialibrary/media/research/economists/pinkovskiy/Economic_Discontinuities_at_Borders.pdf.

Qian, Nancy. 2008. "Missing Women and the Price of Tea in China: The Effect of Sex-Specific Earnings on Sex Imbalance." *Quarterly Journal of Economics* 123(3): 1251–85.

Rubin, Donald B. 1990. "Formal Models of Statistical Inference for Causal Effects." *Journal of Statistical Planning and Inference* 25(3): 279–92.

Saiz, Albert. 2010. "The Geographic Determinants of Housing Supply." *Quarterly Journal of Economics* 125(3): 1253–96.

Scott, Paul T. 2014. "Dynamic Discrete Choice Estimation of Agricultural Land Use." May 22. http://www.tse-fr.eu/sites/default/files/medias/doc/wp/env/wp_tse_526.pdf.

Simonite, Tom. 2016. "Facebook's New Map of World Population Could Help Get Billions

Online." *MIT Technology Review*, February 22.

Skole, David, and Compton Tucker. 1993. "Tropical Deforestation and Habitat Fragmentation in the Amazon: Satellite Data from 1978 to 1988." *Science* 260(5116): 1905–10.

Small, Christopher, Christopher D. Elvidge, Deborah Balk, and Mark Montgomery. 2011. "Spatial Scaling of Stable Night Lights." *Remote Sensing of Environment* 115(2): 269–80.

Small, Christopher, Francesca Pozzi, and C. D. Elvidge. 2005. "Spatial Analysis of Global Urban Extent from DMSP-OLS Nighttime Lights." *Remote Sensing of Environment* 96(3–4): 277–91.

Small, Christopher, and Daniel Sousa. 2016. "Humans on Earth: Global Extents of Anthropogenic Land Cover from Remote Sensing." *Anthropocene* 14(June): 1–33.

Storeygard, Adam. 2016. "Farther on Down the Road: Transport Costs, Trade and Urban Growth in Sub-Saharan Africa." *Review of Economic Studies* 83(3): 1263–95.

Swanson, Ana. 2015. "A Stunning New Look from Space at Nature, North Korea and Chipotle." *Washington Post Wonkblog*, July 23. https://www.washingtonpost.com/news/wonk/wp/2015/07/23/stunning-satellite-images-make-us-look-at-nature-north-korea-and-chipotle-as-never-before/.

Tempfli, Klaus, Norman Kerle, Gerrit C. Huurneman, and Lucas L. F. Janssen, eds. 2009. *Principles of Remote Sensing.* Enschede: The International Institute for Geo-Information Science and Earth Observation (ITC). (Warning large file: http://www.itc.nl/library/papers_2009/general/PrinciplesRemoteSensing.pdf.)

Turner, Matthew A., Andrew Haughwout, and Wilbert van der Klaauw. 2014. "Land Use Regulation and Welfare." *Econometrica* 82(4): 1341–1403.

United Nations Institute for Training and Research (UNITAR). 2011. *UNOSAT Brief: Satellite Applications for Human Security.* Geneva: UNITAR.

US Environmental Protection Agency. 1994. 1:250,000 Scale Quadrangles of Landuse/ Landcover GIRAS Spatial Data in the Conterminous United States. Washington, DC: United States Environmental Protection Agency, Office of Information Resources Management.

van der Meer, Freek D., Harald M. A. van der Werff, Frank J. A. van Ruitenbeek, Chris A. Hecker, Wim H. Bakker, Marleen F. Noomen, Mark van der Meijde, E. John M. Carranza, J. Boudewijn de Smeth, and Tsehaie Woldai. 2012. "Multi- and Hyperspectral Geologic Remote Sensing: A Review." *International Journal of Applied Earth Observation and Geoinformation,* 14(1): 112–28.

Varian, Hal R. 2014. "Big Data: New Tricks for Econometrics." *Journal of Economic Perspectives* 28(2): 3–28.

Vogelmann, James E., Stephen M. Howard, Limin Yang, Charles R. Larson, Bruce K. Wylie, and Nick Van Driel. 2001. "Completion of the 1990s National Land Cover Data Set for the Conterminous United States from Landsat Thematic Mapper Data and Ancillary Data Sources." *Photogrammetric Engineering & Remote Sensing* 67(6): 650–84.

Xie, Yichun, Zongyao Sha, and Mei Yu. 2008. "Remote Sensing Imagery in Vegetation Mapping: A Review." *Journal of Plant Ecology* 1(1): 9–23.

Xie, Michael, Neal Jean, Marshall Burke, David Lobell, and Stefano Ermon. 2016. "Transfer Learning from Deep Features for Remote Sensing and Poverty Mapping." February 27, arXiv:1510.00098v2. http://arxiv.org/pdf/1510.00098v2.pdf.

Yang, Dean. 2008. "Coping with Disaster: The Impact of Hurricanes on International Financial Flows, 1970–2002." *B. E. Journal of Economic Analysis & Policy* 8(1): 13.

You, L., U. Wood-Sichra, S. Fritz, Z. Guo, L. See, and J. Koo. 2014. Spatial Production and Allocation Model (SPAM) 2005 v2.0. Available from http://mapspam.info.

Young, Alwyn. 2012. "The African Growth Miracle." *Journal of Political Economy* 120(4): 696–739.