HKUST IEMS
INSTITUTE FOR EMERGING MARKET STUDIES

Eye in the sky: private satellites and government macro data

Abhiroop MUKHERJEE, George PANAYOTOV, and Janghoon SHON

HKUST IEMS Working Paper No. 2019-68

October 2019

HKUST Institute for Emering Market Studies
The Hong Kong University of Science and Technology
Clear Water Bay, Kowloon, Hong Kong
T: +852 3469 2215 E: iems@ust.hk W: iems.ust.hk

Eye in the sky: private satellites and government macro data

Abhiroop MUKHERJEE, George PANAYOTOV, and Janghoon SHON
HKUST IEMS Working Paper No. 2019-68

Abstract

We develop an approach to identify whether recent technological advancements – such as the rise of commercial satellite-based macroeconomic estimates – can provide an effective alternative to government data. We measure the extent to which satellite estimates are affecting the value of government macro news using the asset price impact of scheduled announcements. Our identification relies on cloud cover, which prevents satellites from observing economic activity at a few key hubs. Applying our approach, we find that some satellite estimates are now so effective that markets are no longer surprised by government announcements. Our results point to a future in which the resolution of macro uncertainty is smoother, and governments have less control over macro information.

Keywords: Alternative data, Satellite Imagery, Asset price impact, Macroeconomic Estimates

JEL classification: G14, E44

Author's contact information

Abhiroop Mukherjee
Department of Finance,
The Hong Kong University of Science and Technology.
E: amukherjee@ust.hk

George Panayotov
Department of Finance,
The Hong Kong University of Science and Technology.
E: panayotov@ust.hk

Janghoon Shon
Department of Finance,
The Hong Kong University of Science and Technology.
E: jshon@connect.ust.hk

We are grateful to Sumit Agarwal, Steffen Andersen, Utpal Bhattacharya, Darwin Choi, James Choi, Zhi Da, Sudipto Dasgupta, Paul Gao, John Griffin, Harrison Hong, Andrew Karolyi, Praveen Kumar, Stefan Lewellen, Xuewen Liu, Dong Lou, Peter Mackay, Ian Martin, Ron Masulis, John Nash, Kasper Nielsen, Christine Parlour, Rik Sen, Kelly Shue, Sheridan Titman, Jialin Yu, Alminas Zaldokas, Chu Zhang and seminar participants at HKUST for many valuable insights on this paper.

1. Introduction

Macroeconomic data plays a central role in economics. Macro variables are "state" variables, key to decision-making in most models of individuals, businesses, and governments. Not surprisingly, then, the collection, processing, and publication of such data has been of great importance, at least since the 17th century, when Sir William Petty[1] proclaimed, *"just accounts might be kept of the People, with the respective Increases and Decreases of them, their Wealth and Foreign Trade"*.

Since the very beginning of economic measurement, markets have relied on government announcements for macro information. This reliance can historically be traced back to the prohibitive cost that any private entity would have incurred to provide aggregate data. However, it raises two main issues. First, since such macro information is also used to measure the government's economic performance, this reliance creates a potential conflict of interest. Second, government announcements are intermittent and often come with delays. This leads to uncertainty about macroeconomic conditions in periods without announcements, and implies that decision-makers typically do not have access to the latest information before they act.

We develop an approach to identify whether recent technological advancements, such as improvements in satellite imagery, are changing the way markets get macro information. We apply our approach in two different settings – primarily in the U.S. crude oil market, and also in Chinese manufacturing – where satellite-based macro estimates have drawn significant recent interest. Our evidence points to such estimates substantially changing the market's reliance on government macro data in both settings.

This change is important because it has the potential to alleviate both of the above issues with government announcements. First, alternative sources of macro data can help resolve conflict-ofinterest issues due to governments providing data on their own performance. Concerns surrounding such conflicts prevail for large parts of the world – for example, data from countries in Africa, South America, or from China, India, etc., is often treated with suspicion (see, e.g., *New York Times* (2018),

[1] Sir Petty was an adviser to Oliver Cromwell, and later to King Charles II of England. See Richard Stone's Nobel Prize Lecture in 1984 for attribution.

New York Times (2019)). Even in Europe, countries like Hungary, Ukraine, Greece, and Italy, among others, have produced suspect macro numbers at some point in recent history.[2] Moreover, trust in government data is far from absolute even in advanced economies like the U.S., where it varies significantly along partisan lines (typically higher among supporters of the party in power at the time, see, e.g., *Edison Research* (2018)). If alternative data and techniques can provide a way to independently verify government numbers – satellites, for example, do not require government approval to monitor economic activity in a country, as they can orbit above restricted national airspace – such problems could be mitigated.

Second, alternative data can also help reduce the uncertainty associated with infrequent arrivals of government macro news. Under the current paradigm, macro uncertainty typically builds up through periods without any government announcements, before its sharp resolution on announcement day. Such lumpiness is perhaps most conspicuously reflected in asset prices, particularly in the fact that over 60% of the cumulative annual equity risk premium is earned on macro announcement days (e.g., Savor and Wilson (2014)), often associated with large price jumps. Given that macroeconomic data from some of these new sources is available more frequently, such estimates may allow firms and investors to react more quickly to changing fundamentals, and reduce the severity of asset price jumps.

The question of alternative macro estimates is particularly relevant today, due to recent advances in large-scale estimation strategies born out of alternative data (e.g., Da, Engelberg, and Gao (2011)). Data collected from internet searches or social media, aggregate credit card transaction information, and satellite images are a few such examples. Satellite-based estimates, in particular, have seen a sharp increase in popularity. Several companies now track planes, ships, trucks, roads, buildings, military activity, crops, oil storage, etc.; many of them with an explicit goal to estimate macroeconomic quantities.

Still, it is far from obvious that such macro estimates based on alternative data can be as accurate as government data. Even large private entities might still find it prohibitively expensive to collect information at the scale of an entire economy. This might force them to focus on limited-sample surveys,

[2] The Hungarian government lied about the state of the economy in order to win the elections in 2006, as surfaced later from a leaked statement by its Prime Minister (*Financial Times* (2007)). Ukraine manipulated its level of reserves, as reported to the International Monetary Fund between 1996 and 1998 (IMF (2000)). Further examples from Argentina, Greece, and Italy, among many others, are given in Michalski and Stoltz (2013).

making their estimates noisy. Moreover, respondents are not legally required to truthfully disclose information to private firms, creating uncertainty about data quality. To what extent can these new estimation methods be effective in overcoming such data concerns is, therefore, an empirical question. If these estimates are truly effective, then markets will be able to incorporate the information contained in government macro data before it is announced. The government's role in resolving macro uncertainty will then diminish, eroding the value of government macro news.

Two key issues, however, need to be resolved by any approach that attempts to study the effectiveness of alternative macro estimates. First, how does one measure the value of macroeconomic information that the government brings through its announcements? We make progress by focusing on asset price moves in a short time window around government announcements. The magnitude of these price moves (i.e., the "price impact" of the announcement) constitutes a summary measure of the value of government macro news. For example, if all relevant information contained in a particular government number has already been impounded in asset prices through individual trades – reflecting the proverbial "wisdom of the crowd" – before its announcement, then it would not move the price.

Second, even armed with a measure of the value of government information, how does one causally identify the extent to which an alternative data source affects this value? A simple beforeafter study, which compares an earlier period when such a data source is not available, to a later period when this source becomes more widely used, can suffer from endogeneity issues. For example, a change in the price impact of a government announcement after such data becomes available could be due to other concurrent developments, e.g., the government starting to provide other related information to the market prior to this particular announcement. Even if one adopts a difference-in-differences strategy – based on alternative data vendors starting to provide certain macro numbers and not others (cross-sectional variation) – endogeneity concerns will remain. This is because the key assumption of random assignment into treatment and control groups is likely to be violated, due to the fact that commercial estimate providers *choose* which macro number to cover.

The ideal empirical experiment to assess causality would involve an exogenous change in the quality of an estimate based on alternative data, which can be achieved, for example, by turning on and shutting off *at random* (Fuchs-Schu¨ndeln and Hassan (2016)) the alternative data source. Then, one could test whether the price response is different on announcement days when the source is switched on, as compared to announcement days when it is shut off.

To design an approach that gets close to this ideal, we rely on two main insights. First, one does not necessarily have to randomize the alternative data source over *entire* national geography for proper identification. This is due to the fact that measuring economic activity at a few select locations, such as production hubs or bottlenecks in the supply chain, is often critically important for a macro estimate provider. Therefore, it may be sufficient to randomize the availability of alternative data *only* over such specific locations to obtain large exogenous differences in the quality of the overall estimate.

We illustrate this idea using the U.S. crude oil market. Crude is typically transported via pipelines, and there are a handful of places – often small towns like Cushing, Oklahoma (population 7,826) or Patoka, Illinois (population 584) – where multiple pipelines intersect, creating central hubs in the supply chain. Such hubs host a substantial proportion of oil storage facilities (for example, ten storage locations used in our analysis account for up to a third of the entire U.S. inventory storage capacity). We exploit this concentration in our test design, and focus on a series of natural experiments that randomize the availability of satellite data over these few hubs.

The second insight underlying our identification strategy is that satellites cannot "see" if clouds obscure their view. Satellite-based estimates of oil inventories are likely to be noisier when clouds cover key supply hubs. More specifically, oil is often stored in tanks with floating roofs, allowing satellites to observe differences in the shadows cast inside each tank, which are used to estimate the level of oil therein. However, shadows cannot be observed if clouds cover the storage location. Figure 1 illustrates the difference between what a satellite sees on a clear day (left panel), and a cloudy day (right panel).

Our identification strategy, then, is simply to test whether the crude oil price responds more to government announcements of oil inventories in cloudy weeks (when satellites are unlikely to provide a good estimate before the announcement, so that the market relies on the announcement to resolve uncertainty), compared to clear weeks (when such estimates are likely to be more accurate).

We find that in weeks when a few key oil storage locations have predominantly cloudy skies, government announcements move oil prices significantly. However, in weeks with clear skies – when satellites can accurately monitor changes in storage, and hence some traders know this information beforehand – prices do not respond to the same announcement. Illustrating the main result of our paper, Figure 2 contrasts the average oil price response to a one-standard-deviation increase in oil inventories around announcements in clear vs. cloudy weeks.

We conduct a number of robustness and placebo checks to test the validity of this evidence. Perhaps the most interesting one of these checks finds no significant difference between the price impacts of the same announcement in cloudy and clear weeks in an earlier period ("pre-period"), when few commercial satellite-based predictions were available. This is exactly what one would expect if cloudiness affected the oil price impact of government announcements *only* through its effect on satellite-based estimates.

Next, we address the mechanisms underlying our findings in more detail. First, we find that cloud cover over key hubs indeed affects the accuracy of satellite-based inventory estimates. Such estimates have substantially higher errors in cloudy weeks, both in terms of economic magnitudes as well as statistical significance. Next, we investigate whether cloudiness affects uncertainty about oil inventories, using the implied variance of USO (the largest U.S. Oil ETF). We find that cloudy periods have an 11% higher variance relative to clear periods, consistent with our hypothesis. No such relationship, however, is observed in the pre-period, when satellite-based estimates were not prevalent.

We end our discussion on oil markets by examining extreme price movements, i.e., jumps, often associated with infrequent information arrivals. If satellites are indeed effective in providing more frequent oil market information, there would be smaller oil price jumps in clear periods, when satellites are able to provide accurate estimates. On the other hand, cloudy periods could still see large jumps, reflecting the lumpiness of information flows in the absence of such estimates. We use the non-parametric price jump detection method of Lee and Mykland (2008), and find that the average size of an oil price jump is 25% higher in cloudy weeks relative to clear weeks. This evidence is consistent with a *causal* link between price jumps and the availability of more frequent sources of information in financial markets. Overall, our evidence points to cloudiness affecting the resolution of uncertainty in the oil market through its impact on the accuracy of satellite-based estimates.

Next, we show that our approach can also be applied in other settings, such as to assess the effectiveness of satellite-based estimates of manufacturing activity in China. There is, of course, significant interest worldwide in understanding the pace of macroeconomic activity in China, imparting relevance to such an assessment. The manufacturing PMI (Purchasing Manager Index) is considered to be a major monthly barometer, and, given the extent of investor interest in independently verifying official numbers, commercial satellite-based estimates of the Chinese PMI are even available on the

Bloomberg platform. Moreover, Chinese manufacturing also happens to be geographically concentrated, which facilitates the application of our approach.

Following our insight on the importance of hubs in estimating macro quantities, we focus on four key provinces (Guangdong, Jiangsu, Shandong, and Zhejiang), which together account for 35-40% of the manufacturing activity in China. We use local cloud cover over eight hub cities, two in each of these provinces, to randomize the accuracy of satellite-based PMI estimates. Similar to our U.S. results, we find that in predominantly cloudy months, (i) satellite-based PMI estimates are significantly less accurate, and (ii) government announcements of the PMI move a broad Chinese stock market index (CSI300) significantly. On the other hand, the impact of such announcements is substantially smaller in clear months. We also find that there are larger price jumps in cloudy months, as detected using the Lee-Mykland approach. Our results using Chinese data, however, are weaker than those that we find in the U.S. – while the differences between cloudy and clear months are economically large, they are not statistically significant in most cases. One possible reason is that the manufacturing PMI may be more difficult to estimate using data that satellites can provide, relative to our U.S. example where satellites can directly observe the quantity of oil stored. Another reason could be that the monthly (as opposed to weekly) frequency of observations in the Chinese case reduces our number of observations and limits our power to demonstrate differences between clear and cloudy periods with sufficient precision, and Section 5.2 provides further discussion.

Overall, the evidence from our tests points towards several implications of alternative data for financial markets, and beyond. Our evidence on the reduction in implied volatility and in price jumps in the U.S. crude oil market as a result of satellite-based inventory estimates suggests a shift in the way macroeconomic uncertainty is resolved in markets. Our evidence on Chinese manufacturing suggests a shift in the way markets get macro information in a country where investors do not always trust official announcements. Besides these, however, there are many other possible consequences of effective alternative macro estimates that are beyond the scope of our paper. For example, effective estimates based on alternative data can help reduce noise in macroeconomic measurement, which has always been an important issue faced by governments (Morgenstern (1963)). On the other hand, if the estimation technology remains prohibitively expensive for most market participants, advance macro information can become concentrated in the hands of a select few, raising a different range of issues on fairness in markets and in society. We leave these issues for future research.

Finally, it is important to clarify that our goal in this paper is to suggest an approach towards understanding the effectiveness of satellite-based estimates, rather than to claim that our evidence in some particular case is conclusive. We recognize that it may take a long time before alternative data sources can provide a viable alternative to a wide range of government-produced macro data, if ever. Yet, given the growing interest in such new sources, it might be timely to develop a rigorous framework to study these developments. Our approach is only a first step in this direction.

2. Related literature

The advantages of satellite imagery as a data source for economic studies are well-established in a recent economics literature. These advantages include wide geographic coverage, access to otherwise unavailable information, and high spatial resolution (e.g., Donaldson and Storeygard (2016)).[3]

Satellite data has been used to measure aggregate economic activity or land usage (e.g., Henderson, Storeygard, and Weil (2012), Saiz (2010), Michalopoulos and Papaioannou (2013)), monitor pollution, agricultural land and crops, electricity use or deforestation (e.g., Foster and Rosenzweig (2003), Burchfield, Overman, Puga, and Turner (2006), Burgess, Hansen, Olken, Potapov, and Sieber (2012), Holmes and Lee (2012)), and study the economic impact of climate and weather, among others. We focus on assessing the impact of satellite-based estimates on the resolution of macro uncertainty, rather than on using such imagery to measure certain activity.

This paper is also related to nascent literature that documents the effects of satellite data on individual firm stocks. Zhu (2019) examines Orbital Insight's sales forecasts that use satellite-based estimates of parking lot traffic for a sample of retailers, and the role of such forecasts as a disciplining mechanism for these firms. Her study shows that the stock price response to earnings surprises has weakened for firms for which Orbital Insight predictions are available. The implications of satellite data usage for equity trading and information asymmetries in the stock market are also investigated in Katona, Painter, Patatoukas, and Zeng (2019). They find that satellite-based information on parking lot

[3] The wide geographic coverage is a clear advantage of satellites, in contrast to aircraft like helicopters and drones, which fly at limited height. Satellites are also not subject to national airspace restrictions.

fill rates is not fully impounded into retailer stock prices prior to the public disclosure of company performance.

We differ from these papers, first, in our focus on government macroeconomic announcements. Macro numbers are key inputs for most economic decisions made by individuals, firms, and other organizations, and the resolution of macroeconomic uncertainty has been shown to be important for welfare by a large literature in economics. Second, our identification strategy, based on a series of natural experiments involving cloud cover, differs from existing approaches. Our strategy does not rely on a single satellite-based estimate provider's numbers, or on differences between assets that such a provider chooses to cover and those it does not, thereby avoiding endogeneity concerns about coverage choice.

Our paper is also broadly related to the literature on the diffusion of information in financial markets (e.g., Hong and Stein (1999), Hong, Torous, and Valkanov (2007)), and in particular to Hong and Yogo (2012), who study this issue in the context of commodity markets. We contribute to this literature by developing an approach to identify the causal impact of satellite-based estimates on the diffusion of macro information in asset markets.

Finally, we are also not the first to use weather-related data in a finance context. Previous studies have documented a relation between asset trading and weather-induced mood fluctuations of traders (e.g., Hirshleifer and Shumway (2003), Goetzmann, Kim, Kumar, and Wang (2015)). Note that this particular channel is not a concern in our study, because the hubs over which we measure local cloud cover are not where trading takes place; therefore, weather-related changes in traders' mood are unlikely to affect our results.

3. Conceptual Framework

This section presents the three steps that our approach takes to assess the relevance of alternative data sources for generating macroeconomic information. While the discussion in Sections 3.1 and 3.2 applies more generally to data sources of various nature, the instrument introduced in Section

3.3 applies specifically to satellite-based data.

3.1. Measuring the value of government macro information

We are interested in measuring how estimates based on alternative data, such as those derived from satellite images, may be changing the value created by the government as an aggregator of macroeconomic information. The first step, then, is to measure this value. Such a measure needs to account for the fact that the government is not the only source of macro information. Various market participants have access to different pieces of such information and may trade on it in financial markets. Their collective trading activity can then impound macro information into asset prices, without a need for government aggregation. For example, individual agents might sell assets in response to a decrease in their own income. If a sufficient number of agents do so, asset prices can reflect an aggregate drop in national income, even before the government announces this drop. In sum, asset prices also serve as an information aggregator. Moreover, analysts, traders, institutional investors, etc., try explicitly to estimate macro numbers, using a variety of different sources. Speculative trades based on these estimates are also likely to impound information from such sources into the prices of related assets.

In this sense, the government creates value as an aggregator of information only to the extent that the macro numbers it announces are not already reflected in asset prices. Therefore, following a large literature in finance that examines asset price changes to gauge the arrival of new information, we measure the value of a particular government announcement focusing on its price impact. If an announcement contains valuable new information, it should have a price impact.

A simple way to measure how, e.g., satellite-based estimates may be changing the value of government information, then, is to examine how much its asset price impact causally depends on the availability of high-quality satellite imagery. The following two sections elucidate our identification strategy that permits us to provide such causal evidence.

3.2. The role of hubs and bottlenecks in estimating macro quantities

While the motivation for market participants to estimate macroeconomic numbers is clear, it is far from obvious whether such estimates can be effective. The main issue stems from the scale of the information-gathering activity in this case, which is typically beyond the reach even of large private entities. For example, governments collect information throughout the *entire* economy, via censuses, which would be prohibitively expensive for private firms. Furthermore, respondents (citizens or firms) are legally required to truthfully disclose information to governments, but not to private firms. Governments also have access to direct information sources (e.g., tax records), unlike private firms.

However, it may not be necessary to monitor entire economies to estimate macro quantities; it may instead be enough just to focus on measuring economic activity at key production hubs or supply-chain bottlenecks. To provide an analogy, if we need to know the number of people attending an event at a large theatre, we do not necessarily have to count every audience member ("census"), or take a *random* sample of a small geo-space in the theatre ("noisy sampling"). Instead, it may be more efficient to focus on counting arrivals through the doors in the few minutes before the event starts. Similarly, to measure Chinese oil imports, it may be sufficient to measure oil tanker traffic in the straits around Singapore, where the shipping routes from the Middle East to China run into a bottleneck. Or, to understand developments in certain industries, it might be enough to monitor the mining of a geographically concentrated natural resource, e.g., cobalt in the Democratic Republic of Congo (essential for rechargeable batteries).

We highlight this insight, using two distinct examples – the concentration of key storage hubs for WTI (West Texas Intermediate) crude oil in places like Cushing, Oklahoma, in the U.S., and the concentration of manufacturing activity in four provinces in China (e.g., Guangdong, Jiangsu, Zhejiang, and Shandong).

This insight, which allows our approach to focus on a limited number of hubs, is important because it makes it easier for the econometrician to identify the causal effect of a particular macro estimate on some variable of interest. Suppose we want to understand how the availability of such an estimate causes a change in the value of a government announcement. A typical approach would require random variation in the accuracy of this estimate, for example through randomly shutting off and switching on the availability of the data used in the estimation process. Finding such random variation for a large country would be much more difficult than for a handful of hubs. Yet, if these hubs are crucial for the activity being estimated, random variation in their observability might generate differences in the quality of the estimate that are sufficient for identification purposes.

3.3. Clouds as an instrument for identification

Satellite imagery is gaining popularity as a source of economic estimates. Several companies have been established in recent years, providing satellite-based forecasts for a variety of economic variables. Among such companies, Planet Labs, Spire Global, and SpaceKnow track planes, ships, roads, buildings, and containers worldwide, Tellus Labs tracks global crops, and Orbital Insight and Ursa Space Systems focus on the real-time estimation of the amount of oil stored in various facilities around the world. Many of these providers estimate macroeconomic variables, which is of particular interest to traders in financial markets (e.g., https://blog.quandl.com/alternative-data-satellitecompanies). However, doubts remain around the accuracy of such estimates. For example, Risk.net mentions the director of stock selection research at Acadian Asset Management saying that as many as eight out of nine efforts to build strategies based on alternative data fail (https://www.risk.net/ asset-management/5305971/quants-look-to-image-recognition-to-process-alternative-data). The same article also mentions quantitative traders at JP Morgan doubting the ability of such strategies to generate signals with sufficient accuracy. This lack of consensus makes it important to assess the effectiveness of commercially available satellite-based estimates. Can satellite-based estimates indeed be so effective that asset markets are able to incorporate the information content of the government's macro numbers before they are announced?

One major hurdle to answering this question lies in establishing causality. For example, it is not sufficient simply to show that an asset's price moves less around certain government announcements after some satellite-based estimate becomes available. First, such an outcome could be due to other contemporaneous market developments, e.g., the government starting to provide other related information to the market prior to this particular announcement. Second, the emergence of the satellite-based estimate is endogenous. It could be due to a change in the demand for information on a particular asset, which prompts alternative data coverage, but at the same time induces analysts to provide better quality forecasts *not using* satellite data. A change in the price impact can then be driven by such better analyst forecasts, and not the availability of alternative data. Furthermore, the advent of satellite-based estimates or forecasts on a particular macro variable also reflects a *choice* made by a satellite data provider to cover this variable, and not another. This is a major concern for any difference-in-differences study, as we mentioned in the introduction.

We design a strategy to identify the effectiveness of satellite-based estimates that sidesteps these issues and is based on random variations in cloud cover. In particular, we relate the price impact of a particular macro announcement to cloud cover over a few hubs, key to the economic activity reflected in the announcement. If an estimate is effective, this price impact should differ across clear and cloudy periods. With clear skies, when the satellite can observe economic activity at these key hubs, the market is informed about macro conditions before the government announcement, leading to little price impact. With cloudy skies – when satellite-based estimates are unlikely to provide precise information beforehand – the market has to wait for the announcement to resolve any uncertainty, leading to a larger price impact.

If the satellite-based estimate is not effective (i.e., if the asset price, in its role as an information aggregator, already incorporates the estimate's content; maybe through other estimates), the price impact of the announcement should be the same, whether the estimate is available (clear periods) or not (cloudy periods).

This intuition leads to a simple test of our main hypothesis: satellite-based estimates are effective if the price impact of the target macro announcement differs across cloudy and clear periods. We describe the advantages and limitations of this approach in the following section.

3.4. Advantages and limitations of our approach

We recognize that another way to assess a satellite-based estimate might be to compare its error to the errors of other estimates/forecasts. This has two disadvantages with regard to our goal of studying the effectiveness of such estimates. First, the asset price can incorporate macro information even without relying on any estimates, if a sizeable proportion of individuals trade their own information into the price, as mentioned in Section 3.1. If it does, then no estimate – however accurate – is effective. This is because asset prices already incorporate the information content of the government's macro announcement in advance, regardless of any estimate. Second, while it may be possible to benchmark one particular forecast error against a specific list of alternatives, it is practically impossible for the econometrician to ensure that this list is *exhaustive* – a key condition for assessing the effect of the estimate under study on the *value* of government information. Some of these alternative estimates, for example, can be proprietary, and hence inaccessible.

In contrast, our approach avoids both these issues. The first issue is already accounted for by the fact that we use changes in price impacts to understand the effectiveness of estimates. If the price already contains all value-relevant information, the price impact would be small, regardless of whether satellites can observe economic activity (in clear periods) or cannot (in cloudy periods). We avoid the second issue relying on the insight that any value-relevant information (even that contained in proprietary estimates that we do not know about explicitly) will be reflected in the asset price, through the trading process. To the extent that the random variation in the accuracy of a satellitebased estimate is uncorrelated with the quality of any other estimate that does not depend on local cloud cover, our approach will be able to distil the effectiveness of satellite estimates with respect to every other source of information embedded in the asset price.

This is not to suggest that our framework is always applicable. We depend on the existence of a traded asset to calculate the price impact of an announcement. In markets or countries where no asset related to the macro announcement is available, or the econometrician does not have access to high-frequency price data on such an asset, error comparisons might still be the only way of assessing the value of estimates.

When it is possible to apply our approach, though, our identification strategy provides another distinct advantage. Our strategy is constructed to examine the effectiveness of satellite-based estimation in general, i.e., of all estimates of the relevant macro variable that use images from *some* satellite, rather than relying on one particular satellite-based estimate. This is because no satellite using optical imagery (which is easier to process, and hence most common) can observe activity on the ground, if its view is obscured by clouds – cloud cover allows us to practically "switch off" estimates made using imagery from *every* satellite over the cloudy area.

It is important to note that such an identification strategy can only be valid if it satisfies the "exclusion restriction", that is, if local cloudiness over the selected hubs is unrelated to factors determining the demand and supply of the variable of interest. This, of course, can never be directly tested. We discuss the plausibility of this assumption in Section 4.4 for the U.S. oil market, and in Section 5.2 in the Chinese manufacturing context.

Finally, we also recognize that our approach does not predetermine the choice of hubs to be used, if multiple of these are available in some settings. It also does not specify how exactly to measure cloudiness. These methodological choices not only depend on the particular question at hand, but are also likely to be determined by specific technological considerations.[4] However, these choices can – and should – be validated in each case. This can be done by directly examining whether the particular measure of cloudiness constructed for the chosen hubs does indeed relate to the noisiness of a satellite-based estimate in an economically and statistically significant manner. Such tests of validity are similar in spirit to the first stage in a two-stage instrumental variables design. In our context, we provide related evidence in Sections 4.6 and 5.2.

4. Satellites and crude oil inventories in the U.S.

Oil is the source of more than a third of the world's energy – more than coal, and more than twice as much as nuclear, hydroelectric and renewable energy sources combined. A recent BBC article states, "No wonder the price of oil is arguably the most important single price in the world."[5] The oil market is also by far the largest commodity market (the value of oil consumed in 2018 is about $1.7 trillion at current prices). It has attracted multiple providers of satellite-based estimates in recent years (see Section 3.3), which highlights the practical relevance of using this example in our paper. Moreover, the oil market provides a unique advantage for our study, as crude oil is the only traded asset, for which there is an official U.S. government announcement at the *weekly* level. The weekly frequency yields enough data points for us to obtain statistical power in our tests, in spite of the popularity of satellite-based estimates being a relatively recent phenomenon.

4.1. Background

4.1.1. Locations where oil inventories are concentrated

Essential for our study is the choice of oil storage hubs over which to measure cloud cover. On one hand, the locations should sufficiently represent overall oil inventories, so that a measure based on them can be meaningfully related to aggregate oil market quantities, and hence to the price of oil. On the other hand, these locations should be limited in number, to avoid the possibility that local cloud

[4] Computer vision algorithms, for example, may be differently impacted by various types of cloud cover, depending on what they are trying to detect.

[5] https://www.bbc.com/news/business-49499443.

cover at these locations is somehow directly related to the overall demand or supply for oil in the U.S. (exclusion restriction).

Figure 3 plots, in different colors, the five U.S. PADDs (i.e., Petroleum Administration for Defense Districts, which date back to World War II, and are nowadays relevant mostly for data collection purposes). The figure also shows the amounts stored in each PADD at the end of 2016 (excluding those in the Strategic Petroleum Reserve). As we can see, PADDs 2 and 3 account for over 80% of the total storage, so we focus only on them.

Within PADDs 2 and 3, there are a few key points where multiple pipelines intersect. These are the natural locations of oil storage hubs, as they maximize the flexibility in sourcing and directing crude oil flows over the pipeline network in response to changes in demand or supply conditions. The most conspicuous of these junctions is at Cushing, Oklahoma, which is the site of vast oil storage facilities, accounting for 14% of the total U.S. oil inventories, as of the end of 2016. Cushing is also the delivery and price settlement point for the world's most actively traded crude oil futures contract - the NYMEX WTI Light Sweet Crude Oil contract (denoted further as "WTI").

Following these observations, we focus attention on ten specific locations, shown in red circles in Figure 3, that we use to construct our weekly cloudiness measure. Besides Cushing, these locations include the Louisiana Offshore Oil Port (part of the Houma area and important for waterborne crude oil), Houston, and Midland (a storage hub for the Permian Basin, currently one of the world's top oil producing regions). They also include Patoka, Illinois, where several pipelines that supply Midwest refineries intersect, as well as several key locations on the Gulf coast.[6] Storage facilities in these ten locations account for about a third of the total for the U.S. in our sample period.

4.1.2. How satellites "see" oil inventories

Although any technical analysis of satellite imagery is beyond the scope of this paper, a basic idea of the methods is still important for understanding our approach, and we provide some clarification.

[6] The full list is: Cushing, OK, Patoka, IL, Clovelly, LA, Saint James, LA, Houston, TX, Beaumont-Nederland, TX, Corpus Christi, TX, Midland, TX, Wink, TX, and Wichita Falls, TX.

Oil is often stored in tanks with floating roofs, to avoid losses from evaporation in the space between the oil surface and the tank ceiling. The floating roof is the main feature that makes it possible to use satellite data to measure shadows, key to estimating quantities stored in such tanks. In particular, a full tank has a very small inner shadow, (i.e., shadow thrown by the tank wall on the floating roof), whereas for a completely empty tank this inner shadow is as wide as the outer shadow, thrown by the tank wall on the ground. Figure IA-1 in the Internet Appendix illustrates the geometric approach to these shadows, and shows how they help to derive the oil content of a tank.

While the geometry is simple, the actual image-processing technology is not. It exploits advances in computer vision and data science, but also involves qualified personnel at certain steps, for example for initial identification of the tanks and their heights and diameters. Cloud cover presents an important additional hurdle. As we can see from Figure 1, for example, even scattered clouds can significantly affect the measurement of shadows, that are critical for an accurate reading. Advances are being made in this direction, by using infrared sensors or radar techniques that can allow peeking through clouds, yet these new methods are still in their development stages. In Section 4.6 we provide evidence that even in the most recent years, cloudiness continues to impact strongly the accuracy of satellite-based estimates of oil inventories.

4.2. Data and measurement

We collect weekly data for crude oil inventories from the EIA (the U.S. government's Energy Information Administration). The data is obtained by aggregating survey responses from a large number of oil market participants (Form EIA-803). The EIA surveys require strict disclosure from the surveyed companies, and failure to file accurate and timely data makes them liable to penalties. The surveyed companies collectively account for at least 90% of the total oil inventories in the U.S. They report inventories as of Friday at 7:00 a.m. After aggregating the responses, the EIA announces the results, typically just after 10:30 a.m. (Eastern Standard Time) on the following Wednesday, i.e., five days later, in its Petroleum Status Report. If there is a holiday on Monday, Tuesday or Wednesday in a certain week, the announcement for that week is delayed to Thursday or Friday at 11:00 a.m., and we adjust the respective oil price and cloud data accordingly.

The price of oil should respond only to *un*expected inventory changes (or inventory surprise). Therefore, we need to calculate what the expected inventory change was before each EIA announcement. Expectations, of course, are not directly observable under most circumstances, and hence are subject to concerns about mismeasurement. One strategy to alleviate such concerns would be to use a variety of expectations measures to understand the robustness of a set of results. More importantly, however, our instrument for identification (cloud cover) offers an additional advantage in this respect. While we might get the precise model for market expectations incorrect, our errors in measurement are likely to be uncorrelated with cloud cover over a handful of places. As a result, if we find significant differences between the price impact on clear and on cloudy weeks, for example, these differences cannot be driven by measurement errors.

Our expectations measure should, ideally, exclude any satellite-based estimates, the effect of which we wish to separately tease out using our instrument. We calculate this measure as a moving average of the percentage changes in inventories over the preceding four weeks, and subtract it from the actual to obtain the unexpected inventory change used in our tests. Our results are robust to different estimation windows. We also show robustness results where the expected inventory change is the change implied by the inventory number for the respective week reported by the American Petroleum Institute (API) in their Weekly Statistical Bulletin.[7] We do not use the API estimate as our main measure of expectations because we do not have API data in the early period when satellite-based estimates were not as prevalent, and hence cannot conduct a placebo test, key to proper identification.

To construct our measure of cloudiness, we collect cloud cover data from the ISD (Integrated Surface Database) via Climate Data Online, provided by NOAA (National Oceanic and Atmospheric Administration), and available at https://www.ncdc.noaa.gov/isd. Similar data has been used in Hirshleifer and Shumway (2003). We collect hourly cloud cover data from the airport nearest to each of our 10 storage locations, as airports typically have higher quality weather data. We average the cloud cover measure over all daylight hours, as we do not know the precise time at which a satellite might observe the location, and there might be multiple data providers that use different satellites.

[7] While the API estimate is also based on a survey, the response to it is voluntary, and non-compliance is not punished.

The EIA announces oil inventories, measured as of 7:00 a.m. on every Friday, just after 10:30 a.m. on the following Wednesday. So, satellite images on any day in-between can, in principle, be informative about the *latest* level of inventories, which determines prices. This timing requires special attention in the construction of our measure of cloudiness relevant to satellite observability of oil inventories, which we describe below.

As discussed earlier, the price impact of the EIA announcement depends on whether the EIA brings in any new information about the *latest* inventory situation, and the Friday inventory number announced by the EIA on Wednesday may not necessarily be the latest information available to the market at that time. This is especially likely to be true in clear weeks, when satellites can provide a better inventory estimate before the announcement. We elucidate what constitutes a clear week in this context using two examples. First, consider the case when at least one of the days between the measurement day (Friday) and the announcement on the following Wednesday is clear. For example, say only the Monday is clear. Then a satellite can get a clear view of oil inventories on that Monday, making that day's estimate the market's latest information on inventory. In such a week, when the government announces the Friday number on Wednesday, an efficient market will view this number as stale, and will not respond to it – it already has inventory information from an even later period. On the other hand, if every day between Friday and the announcement Wednesday is cloudy, the market will not have received any good estimate of inventories before the announcement. In that case, the number announced by the EIA on Wednesday would indeed constitute the *latest* piece of news on the inventory situation, and hence the price should respond to it.

A cloudy week, then, should be defined as one when none of the days between the measurement day and the announcement day is clear. Accordingly, we regard a week as cloudy if its *least* cloudy day has cloudiness above a certain threshold, say the 75th percentile of the sample.[8] We define all other weeks as "clear". By that definition, in a clear week, the least cloudy day must have a sufficiently low

[8] Our week covers the Thursday preceding the measurement until the following Tuesday. The choice of days here follows from the following logic: Given that the EIA's measure refers to inventory at 7:00 a.m. on Friday, the Thursday number should not differ much from the EIA's, so we start measuring clouds on Thursday. On the other hand, satellitebased data providers like OI typically provide the Tuesday estimate on Wednesday morning, making Tuesday the last day with an inventory estimate before the EIA announcement. We show robustness to this choice of window in the Internet Appendix; for example, dropping the Thursday does not change our conclusions.

level of cloudiness. More details on variable construction, such as on how we aggregate hourly cloudiness into lower frequency units, are reported in the Internet Appendix.

Our WTI crude oil futures price data comes from the Thomson Reuters Tick History database. We also use proprietary data on satellite-based estimates of weekly oil inventories from Orbital Insight, and the OVX index of crude oil implied volatility from the CBOE.

Table 1 shows summary statistics for the U.S. cloudiness measure, oil inventory, oil returns and their option-implied variance, together with the main Orbital Insight variables used. As we can see from this table, oil returns are centered around zero, but are slightly negative on average, reflecting a general trend of a slight expansion in inventories in our sample period. Orbital Insight's estimated inventory numbers have a mean value of 435.6 million barrels, close to the true value of 434.7 million, with a standard deviation of 46.6 million barrels, which is somewhat smaller than the volatility of the EIA's numbers at 57.6 million. The average (absolute) error in OI's estimate is 26 million barrels, which is 0.45 of a standard deviation of EIA's inventory numbers.

4.3. Regression specification

In order to assess the differential price impact of EIA announcements in cloudy vs. clear weeks, we run regressions of the following type:

$$ret_t = \alpha + \beta_{clear} * \Delta Oil\_Inv_t * Clear_t + \beta_{cloudy} * \Delta Oil\_Inv_t * Cloudy_t + \epsilon_t, \tag{1}$$

where ret_t denotes crude oil futures returns at time t, calculated from the front-month WTI futures contract, and expressed in percent. $\Delta Oil\ Inv_t$ is the (unexpected) change in oil inventories, as announced on day t, scaled to unit standard deviation to facilitate the interpretation of the coefficients β_{clear} and β_{cloudy}. $Cloudy_t$ is a dummy variable that takes the value of one on a cloudy week, defined as a week when our cloudiness variable exceeds its 75th percentile in the sample. $Clear_t$ equals one minus $Cloudy_t$. To avoid distributional assumptions, we report bootstrap p-values.

Our results remain virtually identical if we also include $Clear_t$ as a separate regressor, to reflect any possible differences in average oil returns in clear and cloudy weeks, as we show in Table IA-2. This invariance in the results is consistent with our evidence on clouds not affecting oil returns directly, as we show in Table IA-1 and discuss in Section 4.4. We also recognize that many other macroeconomic variables can potentially affect price impact, but we cannot incorporate such other controls because we do not have high-frequency data on them. Note, however, that this is unlikely to affect our conclusions on the difference in price impacts in cloudy and clear weeks, as long as cloudiness is not correlated with these variables (as we show for several macro variables, measured at a lower frequency, also in Table IA-1).

The baseline period for these regressions is 01/2014 to 12/2018. This choice is guided by a 2014 U.S. government ruling, that allowed U.S. satellite companies to sell high-resolution imagery (below 0.5 meters) to non-government customers for the first time.[9] As we show in the Internet Appendix, our results are robust to this choice of baseline period.

4.4. Main result: Oil price reactions around EIA announcements

Our main result is presented in Table 2. The first row in the table shows four return horizons *i*, each straddling the time of announcement (typically just after 10:30 a.m. on each Wednesday), constructed to reflect the immediate impact of an oil inventory announcement on the oil price. The following four rows show the regression coefficients β_{clear} and β_{cloudy}, together with their bootstrap p-values in parentheses. We also show the differences between these coefficients, with bootstrap p-values.

Since ΔOil_Inv_t in equation (1) is scaled, β_{clear} and β_{cloudy} can be interpreted as average returns (in percent) *per* one standard deviation unexpected increase in oil inventories, in clear and cloudy weeks respectively. All beta estimates are negative, reflecting the negative relation between excess supply (larger inventories) and oil prices.

The β_{clear} estimates, associated with clear weeks, are all small in magnitude (five to ten basis points) and statistically insignificant in all cases. This shows that the EIA's announcements have practically no impact on the oil price in clear weeks, implying that the information contained in the announcements has already been reflected in the oil price. In contrast, the β_{cloudy} estimates, which are associated with cloudy weeks, are five to ten times bigger in magnitude (51 to 55 basis points), and

[9] Perhaps as a reflection of this ruling, as well as a general decrease in the size and cost of satellites, the annual growth in active satellites has been, on average, 199 per year for 2014-2018, whereas it was on average 32, 31, 24 and 48

significant at the one percent confidence level in all cases. The differences between the respective β_{clear} and β_{cloudy} estimates are all statistically significant at the 1% confidence level.

We present two further sets of results in the Internet Appendix. First, in Table IA-1, we show results from regressing several macro and financial market variables on cloudiness. These include indexes for the stocks, energy, industrials, as well as for international trade (the Baltic Dry Index). We find no relationship between these variables and our cloudiness measure. This finding is consistent with cloudiness affecting our results only through its effect on the quality of satellite-based estimates.

Second, we establish the robustness of our findings. Table IA-2 shows that neither the exact choice of cutoff used to define the $Cloudy_t$ dummy, nor the model we use to calculate expectations of inventory changes is critical for our findings. Our results are also not affected if we use data from 2013-2018 instead of our baseline period, or calculate cloudiness with the daylight period defined as 9:00 to 15:00, or 10:00 to 14:00 (in contrast to our baseline regression, where it is 7:00 to 18:00). Table IA-3 shows that our results are also robust to using a model-free expectation measure, such as the API estimate mentioned in Section 4.2.

Overall, our evidence suggests that unlike cloudy weeks, clear weeks are associated with EIA announcements that have little impact on the oil price in recent years.

4.5. Placebo tests

Table 3 presents results from tests analogous to those in Table 2, but using placebo samples. First, we repeat our analysis using data from an earlier period (01/2007 to 12/2011), when the use of satellite data by oil market players was less prevalent (the "pre-period"). This choice of pre-period accounts for the fact that high-frequency oil futures price data for intervals before 10:00 a.m., as used in our tests, is missing before 2007 in the Thomson Reuters Tick History database. The pre-period ends in 2011, and is thus of the same length as the baseline period (five years). While the intermediate two years (2012 and 2013) are left out in order to sharpen the distinction between the baseline and pre-period, Table IA-2 shows that our results are robust to including these years.

in the five-year periods ending in 1998, 2003, 2008, and 2013, respectively, as per https://www.planet4589.org/space/.

Figure 4 illustrates the contrasting price patterns around EIA announcements in the two periods. The top two panels plot, separately for the baseline and pre-period, the slope coefficients β_{clear} and β_{cloudy} for several return horizons. The bottom two panels plot the difference between such coefficients in clear and cloudy weeks, and a bootstrap confidence interval around it, for the baseline and pre-period.[10]

The plot for the baseline period is identical to what was shown in Figure 2; however, in the pre-period, there is no noticeable difference between the price patterns in clear and cloudy weeks. Note that the price moves in the pre-period are of the same magnitude as in the cloudy weeks in the baseline period, which indicates that the market conditions prevailing in the two periods are similar, except for the role that observability of oil inventories in clear weeks plays in the later sample. The top panel of Table 3 presents statistical support, in the format of Table 2.

The second panel in Table 3 shows that if we shift by two hours the return horizons from Table 2, so that they no longer straddle the announcement, we obtain slope coefficient estimates for cloudy weeks that are much smaller on average and statistically insignificant. We observe similarly small magnitudes and lack of significance on non-announcement days around 10:30 a.m., as seen in the bottom panel of the table. These placebo tests, therefore, provide support for our earlier conclusions.

4.6. Evidence from satellite-based oil inventory estimates

The results so far indicate that the oil price responds to the EIA announcement in cloudy, but not in clear weeks. This finding is consistent with the hypothesis that oil inventory estimates are less accurate when satellites do not have a clear view of key oil storage hubs. Here we provide evidence on this mechanism.

We examine daily data from Orbital Insight (OI), a major provider of oil market information based on satellite images. This data allows us to evaluate directly the impact of clouds on the accuracy of satellite-based estimates of oil inventories. OI started providing data on oil inventories to individual clients since 02/2017. This data contains the sampling error of their inventory estimate, which reflects

[10] We use regression slope coefficients on standardized inventory surprise, rather than a simple event study chart, to account for the time variation in the sign and magnitude of inventory surprise. Such coefficients facilitate comparison across clear and cloudy weeks, in addition to being easily interpretable as the price response to a one-standard-deviation unexpected inventory change.

the staleness of oil storage tank observability. Their data guideline clarifies that this sampling error increases when a tank has not been observed for a few days. Therefore, this error is likely to decrease on clear days when tanks are observable, which we verify in the first column of Table 4.

This column shows a slope estimate of 0.31, which indicates that on a day when our 10 hubs have completely cloudy skies (as described in Section 4.2), the OI sampling error increases by 72.9% relative to a completely clear day at these locations (a coefficient of 0.31 vs. an average of 0.43 for the dependent variable). This indicates an economically non-trivial relation between our cloudiness measure and the observability of oil storage tanks.

In the second column of Table 4 we measure the error in OI's estimates directly. We calculate this error as the absolute percentage difference between OI's estimates and the true value (i.e., the EIA's announcement). The EIA only announces the inventory number as of Friday, so we can calculate the error in OI's estimate only for that day, unlike the sampling error which is available daily. We expect OI's estimates to have higher errors when it is cloudy over our 10 key oil storage hubs on a given Friday.

Our results show that going from completely clear to completely cloudy skies increases the error in OI's estimate by 172.0% (the slope estimate is 9.82, and the average estimation error is 5.71, both measured in percent). The difference between the economic magnitudes in the two columns shows that when clouds obscure the view of major oil storage hubs, the error in OI's estimate can increase disproportionately more than the corresponding drop in the tank sampling error. Such a difference is consistent with the notion that the visibility of key storage hubs is particularly important for estimating aggregate oil inventories.

In sum, cloudiness seems to have a large impact on the accuracy of satellite-based estimates of oil inventories, even in the most recent period. Going beyond the scope of our paper, the importance of this issue is highlighted by ongoing technological efforts to find ways allowing satellites to "see under the clouds" (e.g., OI is now adding Synthetic Aperture Radar (SAR) tools to its methodology). Advances in this direction may attenuate the impact of cloudiness on satellite-based estimates, weakening the validity of our instrument, if and when this happens; but the technology is not there yet.

4.7. Evidence from oil implied volatility

In the following two subsections, we examine the relation between clouds over our few hubs and uncertainty in the U.S. oil market.[11] First, we consider the implied oil return variance, obtained from the OVX volatility index. This index is constructed by applying the VIX methodology to options on the United States Oil Fund (ticker USO). USO is an exchange-traded security that holds near-term oil futures contracts and cash, aiming to replicate the(WTI)oil price.

If satellite-based estimates are more informative in clear periods – as we found using data from Orbital Insight – they should help resolve uncertainty about aggregate oil inventories in such periods. This would result in higher implied oil variance following cloudy periods relative to clear periods.

To test this hypothesis, we regress oil variance on a constant and a daily cloudiness dummy variable, similar to that defined in Table 2. Since we observe implied variance at the daily frequency, we can relate it here to daily cloudiness. We also note that satellite-based data providers typically deliver the inventory estimate for each day on the following day (for example, at 10:00 a.m. in the case Orbital Insight). Therefore, we regress implied variance on a particular day on cloudiness measured as of the previous day: yesterday's cloud cover affects the accuracy of that day's estimate, which reaches the market – and hence affects uncertainty resolution – today.

Table 5 shows that oil variance is on average 14.3% (annualized) following cloudy days, and 12.8% following clear days in the baseline period (2014-2018). The difference amounts to 10.5% of the average variance, with a p-value of 0.02, consistent with our hypothesis. We also present results for a pre-period when satellite-based estimates were not as prevalent. In that period we expect to find no relation between clouds and variance, if clouds indeed affect uncertainty only through their impact on what satellites can see. Our evidence supports this view: the difference between the implied variances in cloudy and clear days over 2007-2011 is 0.2% (i.e., 1% of the average variance in this period), with a p-value of 0.82.

Overall, our results in this Section 4 point to the following mechanism: Clouds over a few key oil storage hubs affect the accuracy of satellite-based inventory estimates (Table 4). Less accurate estimates in cloudy periods lead to higher uncertainty about oil inventories (Table 5). In such periods the market has to wait for the EIA's announcement to resolve this uncertainty, unlike in clear periods.

[11] We thank Ian Martin and Kelly Shue for their suggestions in this context.

This distinction is reflected in the higher price impact of EIA's announcements in cloudy weeks that we document as our main result (Table 2).

4.8. Evidence from oil price jumps

Here we present further evidence on the role played by satellite-based estimates in the resolution of oil market uncertainty, by contrasting oil price jumps in cloudy vs. clear weeks. Jumps are typically associated with infrequent arrivals of important pieces of information, and jumps in equity market indexes or major commodities are often triggered by macroeconomic news announcements. The role of jumps as a distinct aspect of uncertainty in the financial markets is well recognized in the literature (e.g., Merton (1976)).

If satellites indeed provide information on oil market conditions in a more continuous manner, this could potentially be reflected in larger price jumps in cloudy weeks (when high-quality satellite-based estimates are lacking) than in clear weeks (when satellites are able to provide accurate estimates).

We follow Lee and Mykland (2008) to detect price jumps. They develop a non-parametric approach that tests for jumps within each return interval, accounting for the instantaneous volatility by standardizing the returns. Consistent with the setup in Section 4.4, we consider the same weekly measures of cloudiness over our oil storage hubs. We examine returns from open on Friday till 11:00 a.m. on the day of EIA's next announcement. This timing reflects the fact that the inventories announced by the EIA are recorded on Friday morning, but the EIA announcement is scheduled at 10:30 a.m. on Wednesday.[12] We use 5% and 10% confidence levels to detect the jumps for each week, and then calculate the average size of these jumps (i.e., the absolute value of the corresponding price move). We assign a value of zero to weeks without any jumps. Lee and Mykland (2008), whom we follow closely in our implementation of the test, offer further methodological details on jump detection.

Table 6 presents our results. It shows, similar to Table 5, slope coefficient estimates from regressing the weekly sums of jump magnitudes on our dummy variables for clear and cloudy weeks.

[12] Note that while the EIA announcement may be one major source of price jumps, such jumps may not necessarily be linked only to these announcements. Any other source of major inventory news, e.g., company releases on major pipelines developing technical faults, etc., could also lead to price jumps if market participants do not have a sense of such disruptions beforehand through satellite estimates.

In the baseline period (first and third columns), the β_{cloudy} coefficient is 25% (29%) higher than β_{cloudy}, when using the 5% (10%) confidence levels in the jump-detection test. These differences are statistically significant, with p-values 0.09 and 0.03, respectively. There is no significant difference between the clear and cloudy weeks in the pre-period, as shown in the second and fourth columns of the table.

Overall, this evidence suggests that oil market information is more likely to come in larger jumps when oil storage hubs are not observable by satellites due to clouds. This evidence is consistent with a causal link between price jumps and the availability of more frequent sources of information in financial markets.

5. Satellites and manufacturing in China

Investors have long shown interest in verifying official macroeconomic numbers with independent estimates for various high-growth markets, including China and India. This has spurred the establishment of outfits that exploit alternative data to generate such estimates. In China, for example, satellite imagery is used to track agricultural output by TerraQuanta and GagoGroup, crude oil usage by Ursa Space Systems, and manufacturing output by SpaceKnow, among others.

In this section, we apply our approach to Chinese manufacturing, and this choice has three motivating reasons. First, as discussed in Section 3, we rely on measuring the price impact of macro announcements to assess the effectiveness of satellite-based estimates. In the Chinese context, the manufacturing PMI (Purchasing Managers Index) is the only macro variable announced by the government (on a regular basis) for which satellite-based estimates/forecasts are used by market participants. Second, Chinese manufacturing is highly concentrated in a few major industrial hubs. For example, four provinces – Guangdong, Jiangsu, Zhejiang, and Shandong – account for 35-40% of the nation's manufacturing GDP, but only 6% of the nation's area (see Figure 5). As described in Section 3.2, such concentration facilitates identification in our approach. Third, and going beyond the specifics of our approach, the PMI is a major index used to monitor the state of the Chinese economy, and is thus a key indicator of global economic growth. This makes it particularly important to understand the effectiveness of independent (satellite-based) estimates of the official PMI.[13]

5.1. Data and methodology

The analysis in this section follows the methodology implemented in Section 4. We evaluate the impact of the monthly government announcements of the PMI on a major Chinese stock price index.[14] We use a market index, because the PMI announcement is perceived to be generally informative about Chinese macroeconomic trends, and can hence move the aggregate stock market. Specifically, we examine the CSI300, which is a value-weighted index of the 300 largest stocks traded in the Shanghai and Shenzhen stock exchanges, for which we get high-frequency data from Tradeblazer.

We collect PMI data directly from the China Federation of Logistics and Purchasing (CFLP), which is the government agency that conducts monthly surveys among purchasing managers. We resort to this source because the PMI is first announced on the CFLP website (http://www.chinawuliu.com.cn/lhhkx/class_30.shtml). The PMI is announced around 9:00 a.m. (when the market is still not open for the day), so we measure CSI300 returns from the previous trading day's close to 10:00, 10:30, 11:00, and 11:30 a.m. on the announcement day, to ensure that our return horizons include the announcement. We also note that the announcement times are much more irregular prior to 2009, so we drop those years from our analysis, and use 2014-2018 as a baseline period and 2009-2013 as a pre-period, each of the same length (five years) as in our setup for the U.S. oil market. (In the Internet Appendix we also show robustness of our results in shorter sample periods.)

The PMI is scaled between 0 and 100, whereby the value of 50 is the cutoff for economic performance, with manufacturing viewed to be expanding (contracting) when the PMI is above (below) this value. Similar to our treatment of the oil inventory variable in Section 4, we calculate here the unexpected changes in the PMI, by subtracting its six-month moving average (robust to using the three-

---

[13] In China, manufacturing accounts for 42% of the total GDP. The manufacturing PMI includes construction and real estate, and is based on a large sample of about 3,000 firms. While a non-manufacturing (service) PMI also exists, it is considered to be less reliable (*CNBC* (2014)).

[14] Note that this methodology does not rely on governments telling the "truth"; government announcements should move prices as long as there are no other better sources.

or nine-month average, as shown in the Internet Appendix).[15] Table 7 shows summary statistics for the PMI, which indicate that the PMI has been associated with industrial expansion between 2014-2018 (average level of 50.69).

To construct our cloud cover variable, we use weather data from NOAA, as in Section 4.2. Focusing on the four manufacturing hubs, we select Nanjing and Liyang for Jiangsu province; Hangzhou and Jinhua for Zhejiang; Yantai and Qingdao for Shandong; and Zhaoqing and Guangzhou for Guangdong. The choice of these particular cities within a province is mostly driven by the quality of cloud data availability (further details are in the Internet Appendix). In view of the literature on the impact of weather-related mood on trading activity (e.g., Hirshleifer and Shumway (2003)), we deliberately exclude from the list Shanghai and Shenzhen, which host the two main (mainland) Chinese stock exchanges. In Table IA-5 we show that our results are robust to including these two cities in the calculation of the cloudiness measure.

In contrast to Section 4, where the government announces the *stock* of U.S. oil inventories, as observed on a particular day, the Chinese PMI reflects manufacturing activity throughout a month (a *flow*). Therefore, a satellite-based estimate of the PMI needs to average multiple observations of such activity over the respective month. Following this intuition, our cloudiness measure averages daily cloud cover for the selected eight cities over each month. We estimate the same regressions as specified in equation (1), using a clear (cloudy) dummy which equals one when our cloudiness measure is less (greater) than its 75th percentile, again as in Section 4.

We also use satellite-based estimates of the PMI, provided by SpaceKnow, to establish the validity of our cloud cover instrument. This Satellite Manufacturing Index (SMI) has been developed to track Chinese industrial output, based on satellite imagery of over 6,000 industrial sites across China. It synthesizes observations of a wide variety of markers, such as inventory changes, surface materials, and industrial transport. The SMI is expressed in the same units as the official PMI.

As we can see from Tables 1 and 7, the error in the satellite-based estimates of the PMI appears to be larger than that for oil inventory estimates, once we take into account differences in the volatility

[15] As discussed in Section 4.2, even if we do not measure expectations precisely, a lack of correlation between cloud cover and any potential measurement errors will ensure that any difference in price impact between cloudy and clear months is not an artifact of such errors.

of the variables being estimated. In particular, the average (absolute) error in OI's estimate, as mentioned in Section 4.2 is 0.45 of the standard deviation of EIA's inventory numbers. By comparison, the absolute error for the SMI estimate (relative to the actual PMI) is 0.82, while the PMI standard deviation in Table 7 is 0.79, hence the ratio is 1.04.[16] Satellite-based estimates, then, seem to be more error-prone for the Chinese PMI, as compared to U.S. oil inventories.

5.2. Results

Our main results are presented in Table 8. It shows results from regressions of CSI300 returns on the unexpected component of the PMI, interacted with dummy variables for clear and cloudy months, similar to Table 2. The results in all four columns, each referring to one of the four return horizons, show that unexpectedly high PMI readings lead to positive market reaction on average. This reaction is at least twice larger in cloudy months than in clear months, and is statistically significant for cloudy months, and never significant for clear months. However, the differences between the β_{clear} and β_{cloudy} coefficients are not statistically significant, with (bootstrapped) p-values between 0.16 and 0.57.

This lack of significance could be driven by insufficient power in our tests, due, for example, to the time series being relatively short. It could also be driven by satellite-based estimates still being less widely used in the Chinese stock market context, relative to the U.S. oil market. Or, estimating manufacturing activity in China using satellite data on, e.g., truck and ship traffic, could be more difficult than estimating oil inventories in the U.S. using images of storage tanks (as our evidence from Section 5.1 suggests).

Next, Table 9 presents results from tests analogous to those in Table 8, but using placebo samples, in the spirit of Table 3. First, we repeat our analysis using data from the pre-period (20092013), when the use of satellite data on Chinese manufacturing was less widespread. As expected, we find little difference in the price impact on clear vs. cloudy months in this period, with similar values of the β_{clear} and β_{cloudy} coefficient estimates. The p-values for these differences exceed 0.70 in all cases. We also note that while the coefficient estimates are not statistically significant in many specifications (possibly due

[16] This difference in relative error magnitudes is larger if we use the root mean squared error (RMSE) as a measure of estimate accuracy, instead of the absolute errors, which strengthens our conclusion here.

to power issues, as discussed above), at least for the full return interval (ending at 11:30 a.m.), both coefficient estimates are significant at the 10% confidence level.

The middle panel in Table 9 shows CSI300 returns measured on PMI announcement days, but from 13:00 p.m. to the day's close (i.e., in a time interval that does *not* contain the announcement). These returns are economically and statistically insignificant in both clear and cloudy months. Finally, the bottom panel of the table shows similarly small magnitudes of the estimates and lack of significance on non-announcement days, chosen here to be the working days in the week of the PMI announcement, excluding the actual announcement day. Combined with the results in Table 8, these placebo tests confirm the importance of the PMI announcements for the Chinese stock market, and are consistent with our earlier conclusions on the role of satellite-based estimates in the Chinese context.

Similar to our tests for the U.S. oil market in Table IA-2, we establish the robustness of our findings for the Chinese PMI in Table IA-5. It is seen again that the specific cutoff used to define the $Cloudy_t$ dummy, or the specific calculation of the expected PMI is not crucial for our findings, which are also robust to using shorter baseline and pre-periods, and to including Shanghai and Shenzhen in the calculation of the cloudiness measure.

Furthermore, Table IA-4 shows results analogous to those in Table IA-1, which demonstrate the lack of statistical relationship between a number of Chinese macro variables and our cloudiness measure, consistent with cloudiness affecting the stock market index only through its effect on the quality of satellite-based estimates. For example, these results speak to the concern that higher manufacturing activity itself might increase pollution, which our cloud measure may be picking up. Our evidence, which shows a lack of correlation between cloudiness and the level of manufacturing output (as measured by the PMI), helps alleviate this concern.

Next, in Table 10 we focus again on mechanism. In particular, we assess whether satellitebased estimates of the PMI are indeed noisier on cloudy months. For this purpose, we regress the error in the SMI estimate on a constant and our cloudiness measure. The error is calculated as the absolute percentage difference between the SMI and the actual PMI. The table shows that going from a completely clear to a completely cloudy month increases SMI's error by 262.4% (the slope coefficient is 4.25% and the average error in the SMI estimate is 1.62%). This is reassuring, given the concern that the monthly nature of the PMI forces us to average cloudiness over entire months in the Chinese context, which can smooth the fluctuations and lead to power issues in our tests. Our result shows that the averaged cloudiness measure retains enough power to detect large differences in SMI's errors across clear and cloudy months.

Finally, in Table 11, we follow the analysis in Section 4.8, but now in the Chinese context. We examine the interaction between cloudiness and price jumps in the CSI300 index. Since the PMI announcement comes at the start of each month and uncertainty builds through time, more unresolved uncertainty should accumulate towards the end of a month. This makes it easier to statistically detect larger jumps toward the month-end. Therefore, to improve the power of our tests we examine jumps only in the second half of each month.[17] If satellites indeed provide more frequent information, one should expect larger price jumps in cloudy months (when satellite-based estimates are less precise) than in clear months (when these estimates are more precise).

The first (last) two columns in Table 11 refer to jumps detected at the 5% (10%) confidence level.[18] In the baseline period, the average jump size in clear months is around 0.7%, which is substantially smaller than that in cloudy months at about 1.2%. This difference, however, is only statistically significant at the 10% level, in the third column. In the pre-period, the difference between clear and cloudy months is at least three times smaller, with p-values above 0.40. These results are consistent with our hypothesis.

6. Conclusion

This paper is motivated by the recent growth in the availability of various estimates based on alternative data, and their use by market participants. We focus on satellite-based estimates of macroeconomic variables. The main question that we seek to address is whether such satellite-based information can indeed be effective, in the sense that they help resolve macro uncertainty before any government announcement.

We suggest an approach towards understanding this issue, which has several key components. First, we measure the value of a government macro announcement by its price impact. Second, we focus on a handful of locations that are particularly important for estimating specific macro variables, as we

[17] We find qualitatively similar, but not statistically significant results if we examine jumps throughout the entire month.

[18] In China we have data on 30-minute returns, and we define jumps accordingly.

illustrate using storage hubs for crude oil in the U.S., and the concentration of manufacturing activity in China. Third, we use local cloud cover over these locations as an instrument that naturally provides random variation in the quality of satellite data, key to our identification strategy.

In both of the contexts considered, we find that when the hubs of interest have predominantly cloudy skies, (i) the satellite-based estimates are indeed less accurate and have significantly higher errors, (ii) the respective government announcement has a substantially larger, and statistically significant, price impact, and (iii) the resolution of macro uncertainty is lumpier, resulting in larger price jumps.

Can satellite-based estimates *replace* the government as a provider of macroeconomic information? This paper cannot answer that question. Even if such satellite-based information can become very accurate, governments may still have a role in validating these measures, or perhaps more importantly, in disseminating macro information more broadly and in a more equitable fashion (relative to commercial satellites). This paper's scope and contribution are limited to the approach we suggest to measure the effectiveness of estimates. Such measurement though is important; it is the first step to understanding some of these broader issues, which we leave for future research.

References

Burchfield, M., Overman, H. G., Puga, D., Turner, M. A., 2006. Causes of sprawl: A portrait from space. Quarterly Journal of Economics 121, 587–633.

Burgess, R., Hansen, M., Olken, B. A., Potapov, P., Sieber, S., 2012. The political economy of deforestation in the tropics. Quarterly Journal of Economics 127, 1707–1754.

Da, Z., Engelberg, J., Gao, P., 2011. In search of attention. Journal of Finance 66, 1461–1499.

Edison Research, 2018. Economic Anxiety Index poll. 18 October.

CNBC, 2014. Connie Ann. "Why manufacturing PMI still matters for China. 19 August.

Financial Times, 2007. Stefan Wagstyl. "Lies haunt a reformer's grip on power". 25 October.

Foster, A., Rosenzweig, M. R., 2003. Economic growth and the rise of forests. Quarterly Journal of Economics 118, 601–637.

Fuchs-Schu¨ndeln, N., Hassan, T. A., 2016. Natural experiments in macroeconomics. In: Handbook of macroeconomics Elsevier.

Goetzmann, W., Kim, D., Kumar, A., Wang, Q., 2015. Weather-induced mood, institutional investors, and stock returns. Review of Financial Studies 28, 73–111.

Henderson, J. V., Storeygard, A., Weil, D., 2012. Measuring economic growth from outer space. American Economic Review 102, 994–1028.

Hirshleifer, D., Shumway, T., 2003. Good day sunshine: Stock returns and the weather. Journal of Finance 58, 1009–1032.

Holmes, T. J., Lee, S., 2012. Economies of density versus natural advantage: Crop choice on the back forty. Review of Economics and Statistics 94, 1–19.

Hong, H., Stein, J., 1999. A unified theory of underreaction, momentum trading, and overreaction in asset markets. Journal of Finance 54, 2143–2184.

Hong, H., Torous, W., Valkanov, R., 2007. Do industries lead stock markets?. Journal of Financial Economics 83, 367–396.

Hong, H., Yogo, M., 2012. What does futures market interest tell us about the macroeconomy and asset prices?. Journal of Financial Economics 105, 473–490.

IMF, 2000. Release of PricewaterhouseCoopers report on the National Bank of Ukraine. News Brief 00/26.

Katona, Z., Painter, M., Patatoukas, P., Zeng, J., 2019. On the capital market consequences of alternative data: Evidence from outer space. Unpublished working paper. University of California, Berkeley.

Lee, S., Mykland, P., 2008. Jumps in financial markets: A new nonparametric test and jump dynamics. Review of Financial Studies 21, 2535–2563.

Merton, R., 1976. Option pricing when underlying stock returns are discontinuous. Journal of Financial Economics 3, 125–144.

Michalopoulos, S., Papaioannou, E., 2013. Pre-colonial ethnic institutions and contemporary African development. Econometrica 81, 113–152.

Michalski, T., Stoltz, G., 2013. Do countries falsify economic data strategically? Some evidence that they might. Review of Economics and Statistics pp. 591–616.

Morgenstern, O., 1963. On the accuracy of economic observations. Princeton University Press.

New York Times, 2018. Keith Bradsher. "China's economic growth looks strong. Maybe too strong.". 18 Jan.

New York Times, 2019. Keith Bradsher. "Yes, India's economy is growing, but can you trust the data?". 30 May.

Saiz, A., 2010. The geographic determinants of housing supply. Quarterly Journal of Economics 125, 1253–1296.

Savor, P., Wilson, M., 2014. Asset pricing: A tale of two days. Journal of Financial Economics 113, 171–201.

Zhu, C., 2019. Big data as a governance mechanism. Review of Financial Studies 32, 2021–2061.

Table 1

Summary statistics: U.S. oil market

The cloudiness measure is obtained from daily cloud cover data for few key locations, available from NOAA (https://www.ncei.noaa.gov/data/global-hourly/). The Sky Coverage observations from the Station-Hourly data are first aggregated for each hour and location, then these hourly measures are averaged over the daylight period (7:00 to 18:00), and then averaged across the ten locations to obtain the cloudiness for a given day. The oil inventory data (excluding the Strategic Petroleum Reserve) is from the EIA, announced weekly, typically on Wednesday, just after 10:30 a.m. Oil returns for different time intervals on an announcement day are calculated from the front-month WTI oil futures contract (traded on the NYMEX), and the respective statistics are shown in percent. Also shown are statistics for oil inventory estimates available from Orbital Insight, a major provider of satellite-based oil market information, and for the implied variance of crude oil returns, based on the OVX index. The sample period is 01/2014-12/2018.

	Obs.	Mean	Median	St.Dev.
Clouds				
Daily cloudiness	1,826	0.35	0.33	0.16
EIA				
Oil inventory (in million barrels)	261	434.37	438.45	57.56

Oil returns				
10:30-11:00 a.m.	261	-0.08	-0.03	1.13
10:00-11:00 a.m.	261	-0.14	-0.04	1.20
09:45-11:15 a.m.	261	-0.14	-0.08	1.30
09:30-11:30 a.m.	261	-0.13	-0.03	1.41
Orbital Insight (OI)				
Oil inventory estimate (in million barrels)	100	435.61	428.86	46.58
abs(OI-EIA)	100	26.04	20.75	23.73
OVX				
implied oil variance (daily)	1,258	0.13	0.10	0.10

Table 2

Oil price moves around oil inventory announcements in the baseline period

This table shows results from regressions of oil futures returns:

$$ret_t \;=\; \alpha + \beta_{clear} * \Delta Oil\_Inv_t * Clear_t + \beta_{cloudy} * \Delta Oil\_Inv_t * Cloudy_t + \epsilon_t.$$

Returns ret_t (in percent) are calculated over four return horizons, as shown in the first row of the table, t denotes a day when the EIA announces U.S. oil inventories, $\Delta Oil\ Inv_t$ is the (unexpected) change in oil inventories (excluding the SPR), as announced on day t, $Clear_t$ ($Cloudy_t$) is a dummy variable that is equal to one, when day t is associated with a clear (cloudy) week, and zero otherwise. $\Delta Oil\ Inv_t$ is calculated as the difference between (i) the percentage change in oil inventories from $t-1$ to t (as per the EIA's announcement), and (ii) the average of such changes over the preceding four weeks. This variable is scaled to unit standard deviation. A week is defined as clear or cloudy as per the description in Section 4.2. One, two and three stars denote statistical significance at the 10%, 5%, and 1% confidence level, and bootstrap p-values are shown in parentheses. Shown are also R^2's, and the differences between the respective slope coefficients β_{cloudy} and β_{clear}, with bootstrap p-values. The regression intercepts are not displayed, and the regressions employ 261 weekly observations in the baseline period 2014-2018.

	10:30-11:00	10:00-11:00	09:45-11:15	09:30-11:30
β_{clear}	-0.05	-0.10	-0.06	-0.07
	(0.50)	(0.18)	(0.46)	(0.38)
β_{cloudy}	-0.51***	-0.55***	-0.52***	-0.55***
	(0.00)	(0.00)	(0.01)	(0.00)
Difference	-0.46***	-0.45***	-0.46***	-0.48***
	(0.00)	(0.01)	(0.01)	(0.01)

R^2	0.05	0.06	0.04	0.04
Observations	261	261	261	261

Table 3

Placebo tests: Oil price moves at other times

In the format of Table 2, the top panel of this table shows the results from similar regressions, again for the EIA's announcement days, but now over an earlier five-year period (2007-2011, the pre-period). The second panel shows results for EIA's announcement days over the baseline period (2014-2018), but with return horizons shifted by two hours. The bottom panel shows results over the baseline period, but for all *non*-announcement days combined. β_{cloudy} and β_{clear} are slope coefficients on the $Clear_t$ and $Cloudy_t$ dummies interacted with oil inventories.

	Pre-period (2007-2011)			
	10:30-11:00	10:00-11:00	09:45-11:15	09:30-11:30
β_{clear}	-0.43***	-0.50***	-0.47***	-0.51***
	(0.00)	(0.00)	(0.00)	(0.00)
β_{cloudy}	-0.31**	-0.48***	-0.59***	-0.48***
	(0.02)	(0.00)	(0.00)	(0.01)
Difference	0.12	0.02	-0.12	0.03
	(0.43)	(0.86)	(0.50)	(0.89)
R^2	0.12	0.17	0.16	0.12
Observations	261	261 **ine**	261	261 **around**
		period (		**.2:30**
	Basel		**2014-2018),**	
	12:30-13:00	12:00-13:00	11:45-13:15	11:30-13:30
β_{clear}	0.03	0.06	0.02	0.05
	(0.35)	(0.15)	(0.53)	(0.42)
β_{cloudy}	0.01	-0.01	-0.04	-0.04
	(0.75)	(0.91)	(0.55)	(0.68)
Difference	-0.02	-0.07	-0.07	-0.09
	(0.71)	(0.36)	(0.43)	(0.48)
R^2	0.01	0.01	0.00	0.00
Observations	261	261	261	261 **nnouncement**
				lays
	Baseline period (2014-		**2018), non-a**	
β_{clear}	10:30-11:00	10:00-11:00	09:45-11:15	09:30-11:30

	-0.01	-0.02	0.02	0.00
	(0.69)	(0.53)	(0.56)	(0.99)
β_{cloudy}	0.03	0.03	-0.01	0.00
	(0.33)	(0.44)	(0.87)	(0.95)
Difference	0.03	0.05	-0.03	0.00
	(0.31)	(0.30)	(0.64)	(0.98)
R^2	0.00	0.00	0.00	0.00
Observations	992	992	992	992

Table 4

Clouds and Orbital Insight estimates

This table shows the results from regressions of the sampling errors and errors in the estimates of Orbital Insight (OI) on a constant and cloudiness, as discussed in Section 4.6. The sampling error reflects staleness of tank observability, and is provided daily by OI in million barrels; we take its natural logarithm. The error in OI's estimate is the absolute difference between OI's estimate and the true value (EIA's announcement) scaled by the true value; this error is available weekly, given the weekly frequency of the EIA's announcement. The row denoted "Cloudiness" shows the slope coefficient estimates. "Economic magnitude" is as described in Section 4.6. All regressions include a time trend, to reflect improvements in the technology used by OI over time. One, two and three stars denote statistical significance at the 10%, 5%, and 1% confidence level, and bootstrap p-values are shown in parentheses. The sample period is 02/2017-12/2018 (OI started providing its oil inventory estimates in 02/2017).

	Sampling error (log million barrels)	**OI error** (in %, absolute)
Cloudiness	0.31***	9.82**
	(0.00)	(0.03)
Time trend	Yes	Yes
R^2	0.60	0.12
Observations	699	100
Economic magnitude	72.9%	172.0%

Table 5

Clouds and oil market uncertainty

This table shows results from regressing oil return variance (obtained from the OVX crude oil implied volatility index, and available daily):

$$var_{t+1} = \beta_{clear} * Clear_t + \beta_{cloudy} * Cloudy_t + \epsilon_{t+1}.$$

$Clear_t$ ($Cloudy_t$) is a dummy variable that is equal to one, when day t is defined as clear (cloudy) as per the description in Section 4.2, and is zero otherwise. One, two and three stars denote statistical significance at the 10%, 5%, and 1% confidence level, and bootstrap p-values are shown in parentheses. Shown are also R^2's, and the differences between the respective slope coefficients β_{clear} and β_{cloudy}, with bootstrap p-values (in parentheses).

	2014-2018	2007-2011
β_{clear}	0.128***	0.199***
	(0.00)	(0.00)
β_{cloudy}	0.143***	0.201***
	(0.00)	(0.00)
Difference	0.015**	0.002
	(0.02)	(0.82)
R^2	0.01	0.00
Observations	1,258	1,170

Table 6

Clouds and oil price jumps

This table shows results from applying the non-parametric jump detection test of Lee and Mykland (2008). We calculate the average jump size ($JumpSize_t$) each week, and then regress these on the $Clear_t$ and $Cloudy_t$ dummy variables, as defined in Table 2:

$$JumpSize_t = \beta_{clear} * Clear_t + \beta_{cloudy} * Cloudy_t + \epsilon_t.$$

The first (last) two columns show the slope coefficients obtained when the significance level of the jumps test is set at 5% (10%). These coefficients are shown in percent. In each case we show also the difference between the two slope coefficient estimates, and bootstrap p-values for all estimates (in parentheses). We use 15-minute oil futures returns from the Friday preceding an announcement to 11:00 a.m. on the respective announcement day. One, two and three stars denote statistical significance at the 10%, 5%, and 1% confidence level.

	Jumps at 5% significance		**Jumps at 10% significance**	
	2014-2018	2007-2011	2014-2018	2007-2011
β_{clear}	0.68***	0.79***	0.69***	0.86***
	(0.00)	(0.00)	(0.00)	(0.00)
β_{cloudy}	0.85***	0.70***	0.89***	0.78***
	(0.00)	(0.00)	(0.00)	(0.00)

Difference	0.16*	-0.09	0.20**	-0.08
	(0.09)	(0.45)	(0.03)	(0.45)
R^2	0.01	0.00	0.02	0.00
Observations	261	261	261	261

Table 7

Summary statistics: Chinese stock market index and PMI

The cloudiness measure for our Chinese stock market example is derived again from NOAA data. Daily cloudiness is calculated as the average sky coverage measure in eight cities (Nanjing, Liyang, Hangzhou, Jinhua, Qingdao, Yantai, Guangzhou, and Zhaoqing), aggregated over the daylight period (7:00 to 18:00). The monthly manufacturing PMI data is collected from the China Federation of Logistics & Purchasing (CFLP) website. Price data on the CSI300 index is obtained from Tradeblazer. CSI300 index returns are calculated over four return intervals that begin at the close of the last trading day before a PMI announcement, and end between 10:00 a.m. and 11:30 a.m. on the day of the announcement (which is made around 9:00 a.m.), the respective statistics are shown in percent. Also shown are statistics for the satellite-based estimates of the PMI, provided by SpaceKnow and denoted SMI (available on Bloomberg). The sample period is 01/2014-12/2018, as the baseline for the U.S. oil market.

	Obs.	Mean	Median	St.Dev.
Clouds				
Daily cloudiness	1,800	0.53	0.56	0.19
Chinese manufacturing PMI				
PMI (monthly)	60	50.69	50.45	0.79
CSI300 index returns				
prev. close-10:00 a.m.	60	-0.01	0.03	1.00
prev. close-10:30 a.m.	60	-0.03	0.02	1.19
prev. close-11:00 a.m.	60	0.04	0.15	1.22
prev. close-11:30 a.m.	60	0.07	0.11	1.29
SMI				
SMI estimate (monthly)	60	49.99	50.10	1.17
abs(SMI-PMI)	60	0.82	0.71	0.62

Table 8

Chinese stock market index returns and the PMI

Analogous to Table 2, the first four columns of this table show results from regressions of returns of the Chinese CSI300 stock market index, calculated on announcement days for the Chinese manufacturing PMI (Purchasing Managers Index). These returns are calculated as percentage changes over four return horizons, as shown in the first row, where “close” denotes the closing price of the index on the last trading day before the day of the PMI announcement. (Unlike in Table 2, we use here the closing price because the PMI is announced around 9:00 a.m., before trading starts.) The regressions are:

$$ret_t = \alpha + \beta_{clear} * PMI_t * Clear_t + \beta_{cloudy} * PMI_t * Cloudy_t + \epsilon_t.$$

ret_t denotes an index return on day t (in percent), t is a PMI announcement day (one each month), PMI_t is the *unexpected* component of the PMI, and $Clear_t$ ($Cloudy_t$) is a dummy variable that is equal to one, when the PMI announced on day t is associated with a clear (cloudy) month, and zero otherwise. PMI_t is calculated as the difference between the actual PMI announced at t and the average PMI over the preceding six months, and is scaled to unit standard deviation. A month is defined as clear or cloudy as per the description in Section 5.1. Shown are also R^2's and the differences between the respective slope coefficients β_{cloudy} and β_{clear}. The regression intercepts are not displayed. For all regressions, the sample period is 01/2014-12/2018, as the baseline in Table 2. One, two and three stars denote statistical significance at the 10%, 5%, and 1% confidence level, and bootstrap p-values for all estimates are shown in parentheses.

	close-10:00	close-10:30	close-11:00	close-11:30
β_{clear}	0.21	0.20	0.25	0.08
	(0.14)	(0.21)	(0.11)	(0.64)
β_{cloudy}	0.39***	0.40***	0.38**	0.44**
	(0.00)	(0.00)	(0.05)	(0.05)
Difference	0.18	0.21	0.13	0.36
	(0.32)	(0.31)	(0.57)	(0.16)
R2	0.07	0.05	0.06	0.03
Observations	60	60	60	60

Table 9

Placebo tests: Chinese stock index returns

Similar to Table 8, the top panel of this table shows results for PMI announcement days, but over an earlier five-year period (2009-2013, the pre-period). The middle panel shows results for PMI announcement days over the baseline period, but for CSI300 returns calculated over several afternoon intervals. The bottom panel shows results over the baseline period, but for the trading days in the week with a PMI announcement day (using their previous day closing price, and excluding the announcement day itself). β_{cloudy} and β_{clear} are as in Table 8.

	Pre-period (2009-2013)			
	close-10:00	close-10:30	close-11:00	close-11:30
β_{clear}	0.19	0.22	0.23	0.41*
	(0.35)	(0.37)	(0.33)	(0.10)
β_{cloudy}	0.25**	0.28*	0.22	0.36*

	(0.03)	(0.09)	(0.21)	(0.10)
Difference	0.06	0.05	0.00	-0.05
	(0.70)	(0.81)	(0.97)	(0.88)
R^2	0.07	0.06	0.04	0.08
Observations	60	60	60	60

	Baseline period (2014-2018), afternoon			
	13:00-13:30	13:00-14:00	13:00-14:30	13:00-15:00
β_{clear}	0.03	0.04	-0.05	-0.05
	(0.65)	(0.61)	(0.65)	(0.76)
β_{cloudy}	0.06	0.03	-0.03	-0.12
	(0.43)	(0.84)	(0.73)	(0.37)
Difference	0.03	-0.01	0.02	-0.07
	(0.76)	(0.82)	(0.97)	(0.58)
R^2	0.01	0.00	0.00	0.01
Observations	60	60	60	60

	Baseline period (2014-2018), non-announcement days			
	close-10:00	close-10:30	close-11:00	close-11:30
β_{clear}	-0.06	-0.08	-0.06	-0.11
	(0.52)	(0.38)	(0.57)	(0.33)
β_{cloudy}	0.01	-0.06	-0.09	0.06
	(0.99)	(0.71)	(0.55)	(0.83)
Difference	0.07	0.02	-0.03	0.17
	(0.72)	(0.92)	(0.83)	(0.53)
R^2	0.00	0.00	0.00	0.00
Observations	209	209	209	209

Table 10

Satellite-based estimates of the Chinese PMI

This table uses data for the satellite-based estimate of the manufacturing PMI, provided by SpaceKnow. This estimate is denoted SMI and available on Bloomberg. Similar to Table 4, we show the results from regressing the error in the SMI estimate on a constant and cloudiness. This error is the absolute difference between the SMI estimate and the PMI announced for the same month, scaled by the PMI. The row denoted "Cloudiness" shows the slope coefficient estimate, and a time trend is also included. "Economic magnitude" is the ratio between this slope estimate and the average error. The sample period is 01/2014-12/2018, as the baseline in Table 2. One, two and three stars denote statistical significance at the 10%, 5%, and 1% confidence level, and bootstrap p-values are shown in parentheses.

	SMI error (in %, absolute)
Cloudiness	
	4.25**
	(0.01)
Time trend	Yes
R^2	0.14
Observations	60
Economic magnitude	262.4%

Table 11
Clouds and Chinese equity index price jumps

Similar to Table 6, and in the same format, this table shows results from applying the non-parametric jump detection test of Lee and Mykland (2008) to the Chinese CSI300 equity index. We calculate the average jump size ($JumpSize_t$) over the second half of each month, and then regress these on the $Clear_t$ and $Cloudy_t$ dummy variables, as defined in Table 8:

$$JumpSize_t = \beta_{clear} * Clear_t + \beta_{cloudy} * Cloudy_t + \epsilon_t.$$

We show slope coefficients (times 100) for jump tests with significance level 5% and 10%, as well as the difference between the two slope coefficient estimates, with bootstrap p-values for all estimates (in parentheses). We use 30-minute CSI returns, starting two weeks before each PMI announcement and ending at 11:30 a.m. on the announcement day. One, two and three stars denote statistical significance at the 10%, 5%, and 1% confidence level.

	Jumps at 5% significance		**Jumps at 10% significance**	
	2014-2018	2009-2013	2014-2018	2009-2013
β_{clear}	0.67***	0.61***	0.72***	0.69***
	(0.00)	(0.00)	(0.00)	(0.00)
β_{cloudy}	1.22***	0.78***	1.28***	0.81***
	(0.00)	(0.00)	(0.00)	(0.00)
Difference	0.55	0.17	0.56*	0.12
	(0.13)	(0.48)	(0.08)	(0.60)
R^2	0.06	0.01	0.07	0.01
Observations	60	60	60	60

Figure 1: **Satellite images of oil inventory fields in Cushing, Oklahoma**

The two panels show photos taken above Cushing, Oklahoma, by Sentinel-2 on November 26-th and 22-nd, 2018. Sentinel-2 is a multi-spectral imaging mission that uses two satellites flying in the same orbit, designed to monitor the variability in the Earth's surface conditions. The two photos illustrate the difference between satellite images taken on a clear and cloudy day. In the left panel, one can see a large number of crude oil storage tanks, and in particular, the shadows thrown by their walls on their (floating) roofs, and on the ground. These shadows allow for the measurement of the amount of oil in a storage tank (see Section 4.1.2 for more detail). In the right panel, these shadows are not observed, even though many of the tanks can still be seen. The images are sourced from EOS, Land Viewer (https://eos.com/landviewer/).

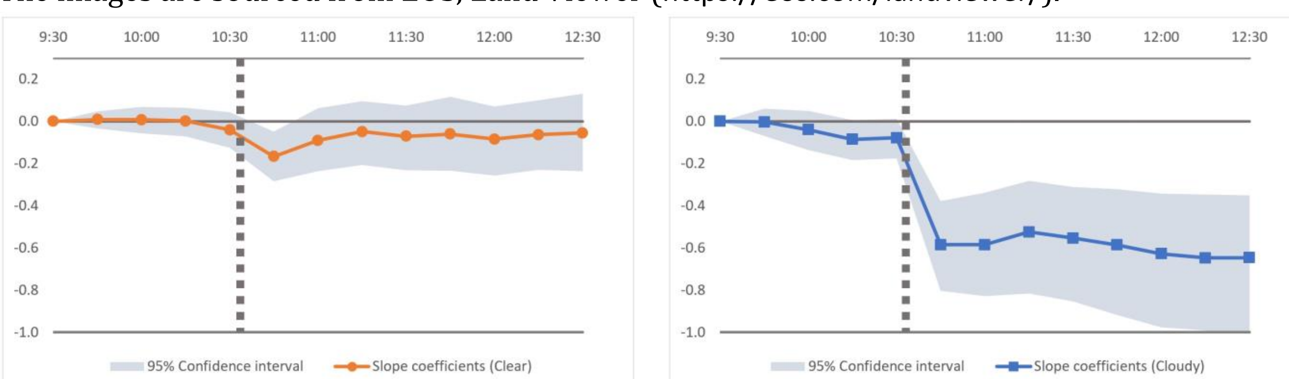

Figure 2: **Impact of oil inventory announcements on oil prices**

The figure shows slope coefficient estimates from regressions of crude oil returns around official announcements of oil inventories, in clear weeks (left panel) and cloudy weeks (right panel). The right-hand side variable in each regression is the *un*expected increase in the oil inventories announced by the Energy Information Administration (EIA), typically just after 10:30 a.m. each Wednesday. We scale this variable so that the displayed slope coefficients represent average oil returns per one standard deviation increase in inventories. The left-hand side variable in each regression is oil return, measured over return horizons that increase at 15 minutes intervals, all starting at 9:30 a.m. on an announcement day, as shown at the top of each panel. The grey areas in each panel show the 95% bootstrap confidence intervals. The dashed vertical lines are set just after 10:30 a.m., when EIA announcements are typically released. More details are in Section 4.2.

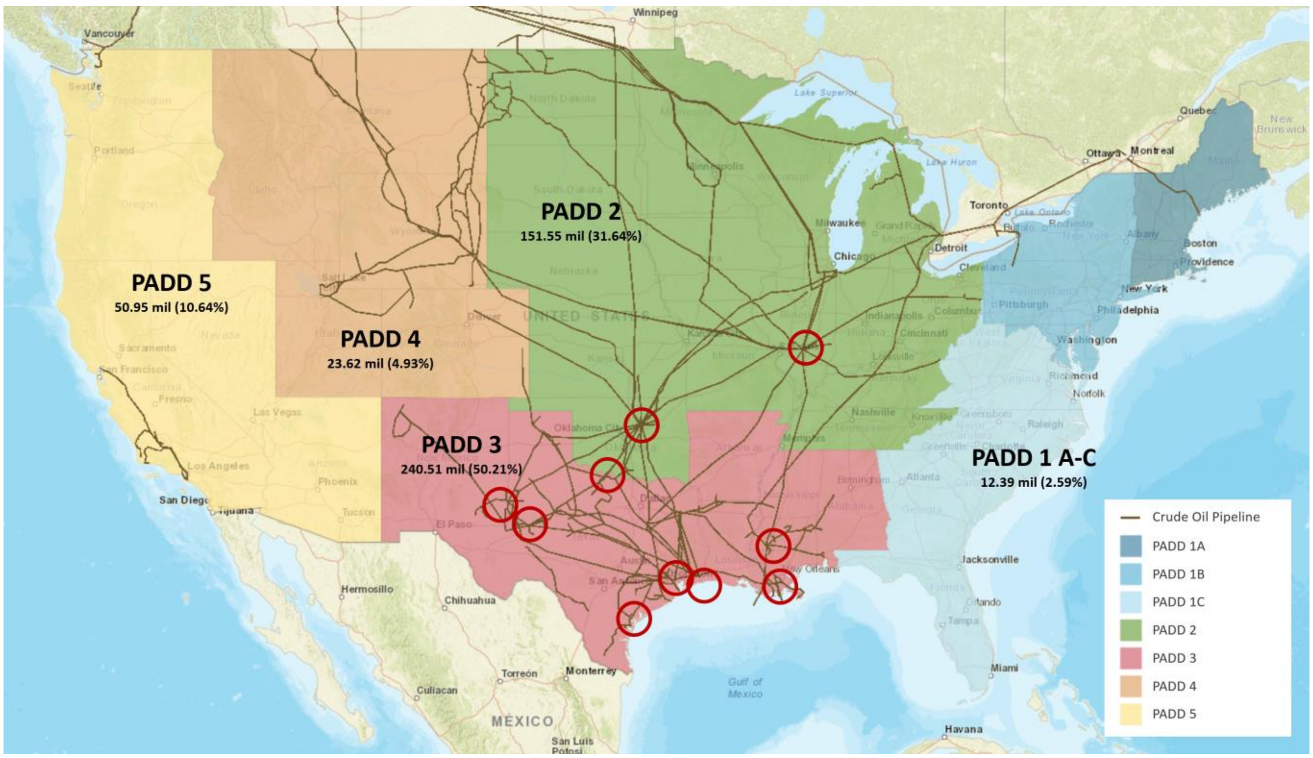

Figure 3: **U.S. PADDs and main oil storage locations**

Different colors on the graph denote the five different PADD's in the U.S. (i.e., Petroleum Administration for Defense Districts, dating back to World War II, nowadays used for data collection purposes). The graph also displays the oil inventories (in millions of barrels, as of the end of 2016, excluding the Strategic Petroleum Reserve), and the shares of each PADD in these inventories. Crude oil pipelines are shown in brown color. The red circles show the 10 locations over which we take cloud cover data to construct our cloudiness measure for the U.S.

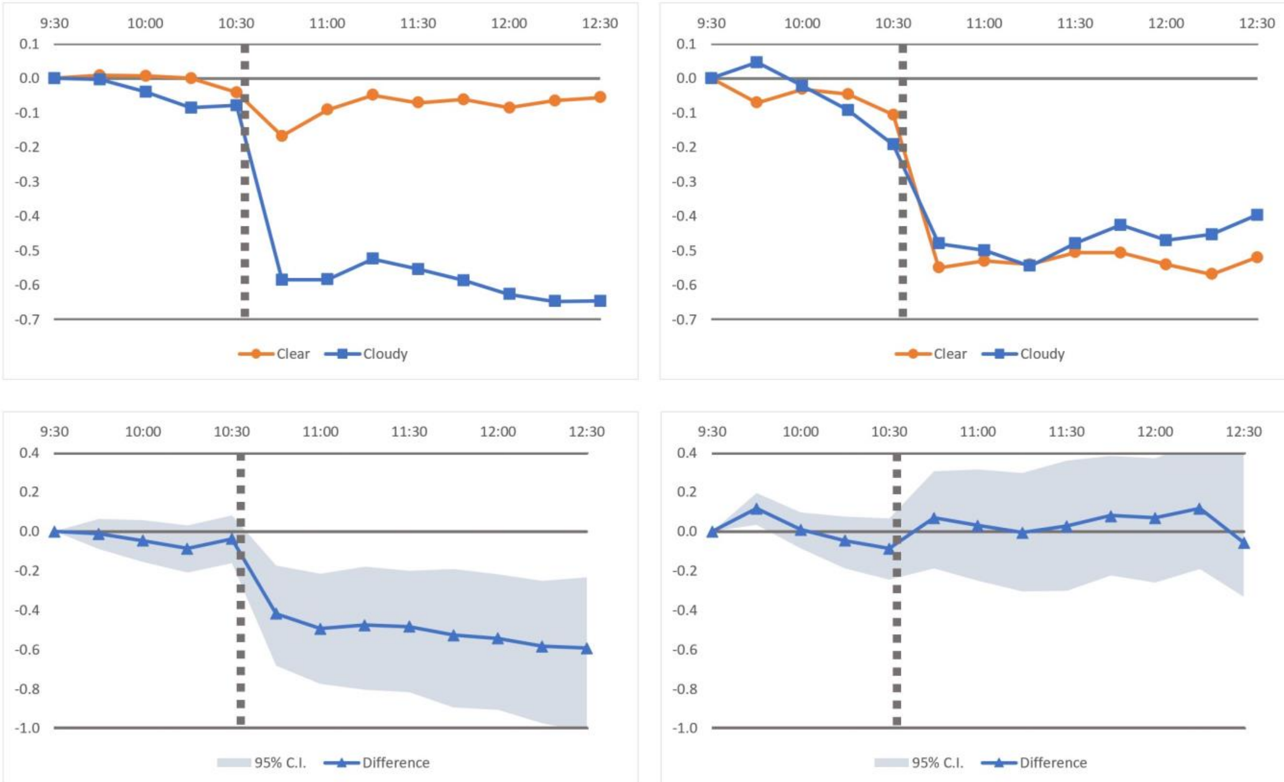

Figure 4: **Oil inventory announcements and oil prices: baseline vs. pre-period**

The top left panel in the figure displays the slope coefficient estimates that were already presented in Figure 2, obtained from regressions of oil futures returns calculated at different horizons on (unexpected) weekly changes in oil inventories, during the baseline period (20142018). These weekly changes are scaled so that the displayed slope coefficients represent average oil returns per one standard deviation increase in inventories. The return horizons increase at 15-minute intervals, all starting at 9:30 a.m., as shown at the top of the panel. The orange (blue) line refers to returns in clear (cloudy) weeks. The top right panel presents the same slope coefficient estimates, but now for the pre-period (2007-2011). The bottom left (right) panel shows the differences between the two sets of slope estimates for the baseline (pre-period), and the grey areas represent the 95% bootstrap confidence intervals for these estimates. The dashed vertical lines are set just after 10:30 a.m., when EIA announcements are typically released.

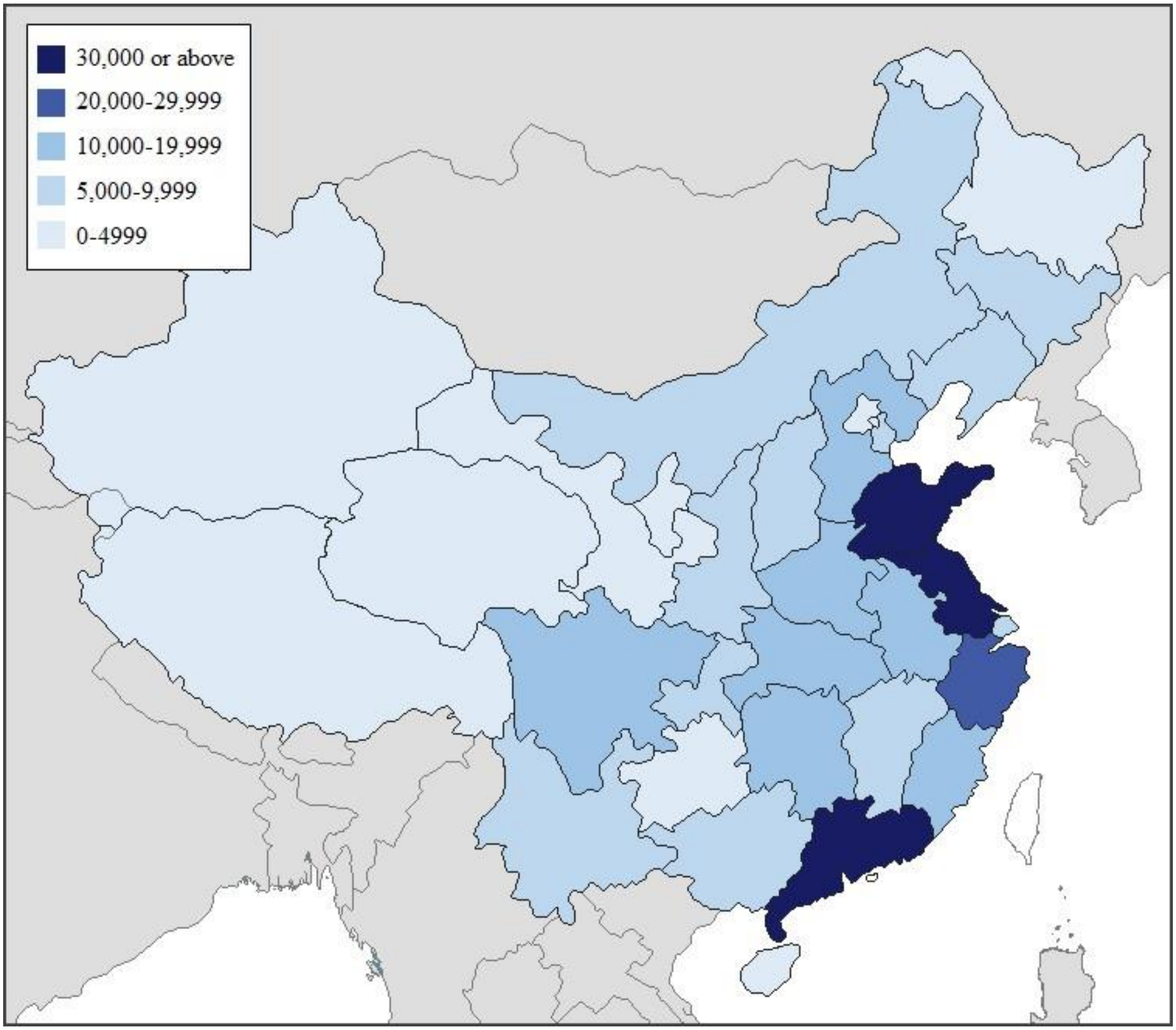

Figure 5: **Chinese industrial output by province**

The map illustrates the distribution of Chinese industrial output (value-added) in 2016, as provided by the National Bureau of Statistics of China, and shown in 100 million RMB (http://data.stats.gov.cn/english/easyquery.htm?cn=E0103). Provinces with larger industrial output are shown in darker shades of blue. The highest industrial concentration is in four provinces: Guangdong, Zhejiang, Jiangsu, and Shandong, which together accounted for 37% of aggregate manufacturing output in that year.