Appendices (For Online Publication Only)

Note: This appendix has been provided here for reference only. The main manuscript is meant to be self-sufficient and the contents of this appendix only serve to add additional detail to the material presented.

Table of Contents

Appendix A: Data description

This appendix describes in additional detail the data sets used in the analysis.

A1. Landsat Coverage Data

It is useful to review some technical details of the Landsat satellite program before understanding how the data on Landsat coverage are generated.

Landsat Program: The Landsat program is a forty year-old program to collect imagery of the earth's surface. There have in total been seven successful Landsat satellite launches including Landsat 1, 2 and 3 which form the first generation of satellites launched in the 1972, 1975 and 1978 respectively. These first generation satellites had a similar technical design and are the focus of this paper. Each operated at an orbit of about 900 km above the earth's surface, took images of "moderate resolution" covering an area of approximately 185km X 185km in each image. Each satellite orbited the earth once every 18 days, and consequently was designed to collect repeat images of the earth's surface over this interval. Each satellite carried the "Multispectral Scanner System" (MSS sensor) that captured information in the spectral resolution of $0.5 – 1.1$ μm (Landsat Data Users Handbook). Because images were taken using the MSS rather than a standard optical camera, different bands of information were captured including the visible IR and reflected near-IR portions of the spectrum, and all of this data was available for analysis. Note that Landsat-1 also contained another sensor, the "Return Beam Vidicon" (RBV) sensor, which proved to be a subsidiary sensor and provided very little data.[5] I exclude the RBV sensor from the analysis and focus on the MSS only. The Landsat system operated under the "Worldwide Reference System" (WRS) that is a referencing system to identify different locations around the earth and their corresponding image in the Landsat system. By my calculations, about 15,000 of these Landsat image locations intersect land features on the earth, and these 15,000 185kmx185km "blocks" form the sample for my analysis.

Calculating Landsat Intensity by block: In 1973, The US Geological Survey (USGS) constructed a facility near Sioux Falls, South Dakota known as the Earth Resources Observation and Science (EROS) data center to archive and distribute Landsat imagery. This data center is the main repository of Landsat information for follow-on use. In order to quantify variation in the availability of imagery it is necessary to study these archives and arrive at estimates of Landsat holdings at a given point in time for a given block, a non-trivial exercise, given both the size of the data holdings and the difficulties in accessing the data. I rely on Goward et al. (2006), a study conducted at the behest of the Advisory Committee to the USGS National Satelite Land Remote Sensing Data Archive to produce estimates of Landsat historical holdings at different points in time. This study analyzed all available data at the EROS center and produced maps that visualized geographical variation in Landsat holdings on a yearly basis and also provided estimates of cloud cover for each block-year. These data are publicly available and I downloaded them from `http://edcftp.cr.usgs.gov/pub//data/richness/` on Nov 13, 2014. Specifically I focus on the `mss1` files available (from `mss1972.tar.gz` upto `mss1983.tar.gz`) as the raw maps used to calculate my Landsat intensity scores. Each of these files provides a shape file for the corresponding year, that show the number of images collected for each block as well as measures of cloud cover in 10% increments. Using these block-year level observations of intensity and cloud cover, I calculate overall measures of intensity and cloud cover.

Note that while the EROS center was the major repository of Landsat imagery charged with providing data without discrimination, globally and at a reasonable cost – it was not the only source of Landsat data. There were a few countries that collected local Landsat data through "International Cooperator" (IC) stations, and some US departments maintained their personal repository of Landsat data. For the purposes of this paper, I am unable to survey these data and rely on the estimates from Goward et al. (2006) for the analysis in this paper.

---

[5] see `https://lta.cr.usgs.gov/rbv.html.`

A2. MinEx Consulting Discovery Database

There exists no canonical database that tracks global mineral discoveries. In this paper I use a proprietary database developed by MinEx Consulting, Australia to track gold discoveries. These data have been compiled manually over many years by Mr. Richard Schodde of MinEx consulting. These data are based on information sourced from company annual reports, press releases, NR 43-101 disclosure documents under Canadian law, technical and trade journals (like Economic Geology, Northern Miner and Mining Journal), government files from various national geological surveys and personal communications with key people in the industry. These data were made available for the current research project under a non-disclosure agreement with MinEx consulting and are not available for redistribution given their commercial value.

Coverage: The data was compiled from MinEx's master database which contains information on over 55,000 mineral deposits across a wide range of metals. A large number of these deposits are smaller than "Minor" – and as such are of limited commercial interest. It is unlikely that the database has 100% coverage, because a large number of deposits (especially of smaller sizes) are not reported systematically and many companies and countries do not provide full breakdowns of their current inventory. Notwithstanding these issues, according to MinEx estimates that "its database (including information on discovery date) for Gold and Base Metals captures at least 99% of all giant-sized deposits and 93% of the major deposits and 65% of the moderate deposits" across all minerals. Coverage for larger deposits for gold is estimated to be significantly better than this baseline estimate.

Additional Notes: It is important to note a few important aspects of these data. First, while I analyze "gold" deposits, deposits are often made up of more than one mineral. Gold for example, is often found along with Copper. I only include deposits where the "primary metal" has been identified as Gold according to the MinEx data. MinEx makes this evaluation based on the economic value of the different mineral deposits reported at a given location.

Second, the discovery year refers to when the deposit was recognized as having significant value. This is usually set as the date of the first economic drill intersection. It should be noted that a review of the discovery history of the deposit may show that there were small-scale workings on the site. For purposes of these data, if there is a order-of-magnitude step change in the known endowment of the deposit (i.e. from 100 koz to >1 Moz of Gold) the date of the upgrade is viewed as the discovery date for the main deposit.

Appendix B: Theoretical Framework

Theoretical Appendix

This appendix provides additional detail on the theoretical framework. Section B.1 provides a simple numerical example that illustrate the basic mechanics of the model. Section B.2 provides additional details on the model without competition across projects, while section B.3 derives some results for model with competition across projects.

B.1 Simple Example

To fix ideas consider the following simple example. Let $V = 100$ and $p_0 = 0.1$, so that expected value of exploration is 10 units. Now, let $C_L = 5$ and $C_H = 20$. This is the intermediate case; before public data, all low-cost firms explore and all high cost firms do not. The high-cost and low-cost case would be when both C_H and C_L are greater and less that 10 respectively. Let us also assume that there are 20 low-cost firms and 80 high cost ones so that $x = 0.2$. Before public data, all 20 low cost firms explore and the overall number of discoveries is 2.

Now, public data increases firms priors to p^+ after a positive signal and decreases it to p^- after a negative signal. In one case, the negative update is relatively small (i.e. $p^- > 0.05$) so that the low-cost firms do not exit, even when they receive a negative signal and the positive update is sufficiently strong to get high-cost firms to invest after a positive signal (i.e. $p^+ > 0.2$). Imagine for example that $p^- = 0.08$ and $p^+ = 0.28$, such that the probability of a positive signal, $q = 0.1$. In this case, all 20 low-cost firms invest and 2 of them report discoveries – but the 10% of the 80 high-cost firms, i.e. 8 firms receiving a positive signal also invest, leading to 2.24 (8×0.28) more discoveries.

A second case is when low-cost firms with a negative signal dont explore (i.e. $p^- < 0.05$) and high-cost firms receiving a positive signal do (i.e. $p^+ > 0.2$). Specifically, assume that $p^- = 0.04$ and $p^+ = 0.28$, then I have the probability of a positive signal $q = 0.25$. In this case, 25 firms will receive positive signals and total discoveries therefore increases from 2 to 7. However, now, assume that three-quarters of the firms are low-cost. In this case, without the map I would have 7.5 discoveries, while with public data I would still have only 7 discoveries. In other words, discoveries would reduce. In fact, this result holds any time the share of low-cost firms is greater than 70% ($\frac{qp^+}{p_0}$). A similar result is obtained by manipulating the informativeness of the public signal, i.e. values of q and p^+. The less informative the signal, the less likely that discoveries increase. Overall, as long as a relatively smaller proportion of firms are low cost, and the signal is informative enough, I expect the total number of discoveries to go up. In other cases, public information could decrease investment and the number of discoveries.

Now, consider heterogeneous effects by firm type. Assume that 20 out of 100 S firms are low cost, while only 5 of the 100 J firms are low cost ($x_S = 0.2, x_J = 0.05$). Before public information, S firms make 2 discoveries and J firms make 0.5 discoveries giving them a 20% market share. After the public signal, consider the case where low-cost firms don't exit but high-cost firms with positive signals enter. As described before, S firms will make 4.24 discoveries. For J firms, in addition to the 5 firms who continue investing and make 0.5 discoveries, 10 percent of the remaining 95 firms, i.e. 9.5 firms also invest making 2.66 discoveries, for a total of 3.16 discoveries. The J market share therefore goes up from 20% to 42%. More generally, since there are a greater share of high-cost J firms (by assumption), the increase in J discoveries will be greater than the increase in S discoveries, thereby increasing their market share.

Now consider the case where only firms with positive signals invest and low-cost firms with negative signals do not. In this case, assuming the same structure of signals as before ($p^- = 0.04$ and $p^+ = 0.28$), S firms move from 2 discoveries to 7 discoveries and J firms move from 0.5 discoveries to 7 discoveries as well. Therefore J firms move to a market share of 50% from a market share of 20%. In fact, as long as total discoveries increase, J share increases as well. The only case where the J share decreases is when C_H is so low that all J firms invest before the map.

B.2 Model without competition.

B.2.1 Derivations for baseline model.

As clarified by the model and the example, in all cases, except the one described below, the effect on the total discoveries is straightforward: 1) the increase in total number of discoveries happens when group characterized by some cost type starts exploring, while no group stops exploring, 2) decrease happens when some group stops exploring, while none starts exploring, 3) no effect means that none of the groups change their behavior. In the case below both high and low cost groups change their behavior.

We have $C_L \leq p_0 V < C_H$. Suppose that then $C_L > p^- V$, $C_H \leq p^+ V$. Before the public information, the total number of discoveries is given by $x \times p_0$ — share of low-cost types by average probability of discovery equal to common prior, and after the public information signal it is given by $qp^+ = \left(\frac{p_0 - p^-}{p^+ - p^-}\right) \times p^+$ — probability of getting positive signal by probability of discovering gold conditional on the signal. Comparison of two values implies that number of discoveries increases with public information if $x < q\frac{p^+}{p_0}$, decreases if $x > q\frac{p^+}{p_0}$, and stays the same if $x = q\frac{p^+}{p_0}$.

B.2.2 General cost distributions.

Overall Discoveries Suppose that costs can be drawn from some distribution with cdf $F(\cdot)$ (which can represent a discrete or continuous distribution). Then condition on the strict increase in total number of discoveries can be rewritten as:

$$\frac{p^+}{p_0} > \frac{(F(p_0V) - F(p^-V))/(p_0V - p^-V)}{(F(p^+V) - F(p^-V))/(p^+V - p^-V)}$$

The left-hand side here contains term describing increased precision due to receiving positive signal, while the right-hand side gives ratio of average density across firms that can be induced by negative signal to abstain from exploration and average density across both firms that can be induced by positive signal to initiate the exploration and those that can be induced to abstain. It is intuitive as the effect on the total number of discoveries depends on both the increase in success rate of exploration attempts (that depends on precision of signal) and relative densities of those firms affected in terms of their decision to explore.

Several examples help illustrating this expression.

1) If distribution F is uniform and values p^-V and p^+V lie in its interior, then the right-hand side , i.e.

$$\frac{(F(p_0V) - F(p^-V))/(p_0V - p^-V)}{(F(p^+V) - F(p^-V))/(p^+V - p^-V)}$$

simplifies to 1 and we unambiguously get the increase in total number of discoveries: $p^+/p_0 > 1$.

2) If $p^- \to 0$, so that negative signal almost perfectly predicts that discovery is impossible and as long as $F(0) = 0$, the condition simplifies to $F(p_0V) < F(p^+V)$, so as long as there any additional firms that are induced to start exploring due to positive signal, the effect on total discoveries is positive.

3) In general, the condition above can fail. We can rewrite it in the following way:

$$\frac{p^+}{p_0} > \frac{p^+ - p^-}{p_0 - p^-} - \frac{(F(p^+V) - F(p_0V))/(p_0V - p^-V)}{(F(p^+V) - F(p^-V))/(p^+V - p^-V)}$$

Suppose $F(p^+V) - F(p_0V) \approx 0$ and $\frac{p^+}{p_0} < \frac{p^+ - p^-}{p_0 - p^-}$ which is implied by $p^- > 0$, and then the condition above must fail. For continuous distributions such failures would happen if the density of cost function near p_0V is decreasing sufficiently fast — for example, if we are in the right tail of normal or log-normal distribution (vast majority of firms are already exploring their projects), then increase from p_0V to p^+V in expected value would attract a small mass of firms, while decrease from p_0V to p^-V would have a strong negative effect on exploration. However, if we are initially in the part of distribution where density is increasing or stable, the failure of the condition above is unlikely to happen.

Junior share: If we assume that costs for junior firms are drawn from distribution $F_J(\cdot)$ and costs for senior firms are drawn from distribution $F_S(\cdot)$, then the general condition for share of juniors to increase can be written down as:

$$\left(\frac{F_J(p^+V)}{F_J(p_0V)} - \frac{F_S(p^+V)}{F_S(p_0V)}\right) + \left(\frac{p^+ - p_0}{p_0 - p^-}\frac{p^-}{p^+}\right)\left(\frac{F_J(p^-V)}{F_J(p_0V)} - \frac{F_S(p^-V)}{F_S(p_0V)}\right) > 0$$

As one can expect, results depends on the relative cumulative distribution function for juniors and seniors. In particular if the negative signal is sufficiently precise, the second expression becomes negligible and juniors share increases as long as there is higher relative growth in exploration by juniors than by seniors due to the positive information shock. One example of a condition under which this would happen is when cost distribution of juniors is a parallel shift from continuous cost distribution of seniors towards higher values of costs and f/F is decreasing for the distributions (that would hold for uniform or normal distribution among others).

B.3 Model with competition.

B.3.1 Setup.

1) There are N firms and K projects. Each firm i decides whether to: 1) not engage in any activity and get payoff 0, 2) explore on one of the projects to evluate whether it is profitable (project k), pay cost C_i and get value V with probability p_k/N_k if success is achieved given that p_k is probability of success for project k and N_k is total number of entrants on the project k. Here we assume that all firms are equally likely to make a discovery while exploring the project, and gains go to the firm that makes the discovery. Firms are risk-neutral and they explore if they are indifferent between exploration and inaction. This assumption also implies that number of discoveries is equal to the number of projects on which discoveries are made.

2) Information structure is the same as before. Without public signal the probability of discovery on each project is equal to p_0. With public signal each project has probability of discover equal to $p^+ > p_0$ with probability $q \in (0,1)$, or p^- otherwise. Here $q = K^+/K$, where K^+ is the total number of projects receiving positive signal.

3) Firms are indexed in such way that $C_1 \leq C_2 \leq \ldots \leq C_N$. Costs can take two levels: there are N_L firms with lower value of costs C_L and N_H firms with higher value of costs C_H. Let $x = N_L/N \in (0,1)$.

4) We assume that decisions are made sequentially, with firms that have lower indices making their decisions earlier. This ensures uniqueness of equilibrium in terms of number of entrants. This is commonly made assumption in industrial organization literature on firm's entry (Bresnahan and Reiss, 1990; Berry and Reiss, 2007).

B.3.2 Equilibrium.

In this section we describe behavior along equilibrium paths

Without public signal: Entry takes place until the following two conditions hold for some I:

$$\frac{p_0 V}{N_{min}(I-1)+1} \geq C_I, \frac{p_0 V}{N_{min}(I)+1} < C_{I+1}$$

Here $N_{min}(i)$ is the minimal number of entrants on a project across all possible projects, when i firms have entered. In this case $N_{min}(i) = \lfloor i/K \rfloor$. If the latter condition never holds, then all firms enter exploration. If $K \geq N$, these conditions simplify to:

$$p_0 V \geq C_I, p_0 V < C_{I+1}$$

If $K \geq N$ and two cost types we have one of three cases depending on the parameters — no firms entering, all low cost firms entering and all high cost firms not entering, and all firms entering. This corresponds to three cases described in the model without competition.

With public signal: Entry takes place until the following two conditions hold for some I:

$$\max\left(\frac{p^+}{N^+_{min}(I-1)+1}, \frac{p^-}{N^-_{min}(I-1)+1}\right)V \geq C_I,$$
$$\max\left(\frac{p^+}{N^+_{min}(I)+1}, \frac{p^-}{N^-_{min}(I)+1}\right)V < C_{I+1}$$

Here $N^+_{min}(i)$ is the minimal number of entrants on a project with positive signal and $N^-_{min}(i)$ is the minimal number of entrants on a projects with negative signal, when i firms have entered. If $K \geq N$, these conditions simplify to:

$$\max\left(\frac{p^+}{N^+_{min}(I-1)+1}, p^-\right)V \geq C_I,$$
$$\max\left(\frac{p^+}{N^+_{min}(I)+1}, p^-\right)V < C_{I+1}$$

B.3.3 Intermediate cost case results.

Here we consider in detail the intermediate cost case, which is the central part of the analysis without competition.

Assume that $C_L \leq p_0V < C_H$. Without the public signal the expected number of discoveries in equilibrium is given by p_0K if the low cost firms explore all projects. This is possible if $N_L \geq K$. If $K > N_L$, then all low cost firms explore different projects and total expected number of discoveries is equal to p_0N_L. Therefore, total number of discoveries is equal to $p_0\text{min}(N_L, K) = p_0\text{min}(xN, K)$.

Firms receive public signal, so that qK projects get positive signals and $(1-q)K$ projects get a negative one. Now consider two cases. First, suppose that $xN \leq qK$ — number of projects receiving high signal is higher than number of low cost firms. Then as at least N_L firms enter projects receiving high signal in equilibrium — if it does not happen at least one of low cost firms would want to deviate to enter one of high signal projects. Therefore, the expected number of discoveries is no smaller than $p^+N_L > p_0\text{min}(N_L, K)$. Second, suppose that $xN > qK$. Then there is at least one entrant on each of the qK projects receiving positive signal — again if it does not happen at least one of low cost firms would want to deviate to enter one of high signal projects. Therefore, the expected number of discoveries is no smaller than $p^+qK > p_0\text{min}(N_L, K)$ as long $p^+qK > p_0xN \Leftrightarrow x < q(p^+/p_0)(K/N)$. That means that a sufficient condition for number of discoveries to increase is:

$$x < (q) \times \left(\frac{p^+}{p_0}\right) \times \left(\frac{K}{N}\right)$$

Note that this condition is very similar to the one in the case with no competition, except being adjusted by term regulating toughness of competition — total number of projects divided by total number of firms. Public signal gives firms more information to the firms. If projects are numerous relative to number of firms, allows for better sorting of firms across projects and makes it more likely that total number of discoveries goes up in response to arrival of public signal. If total number of firms is larger than number of projects, concentration of firms within projects getting high signals makes it less likely that total number of discoveries goes up. This happens as from the perspective of maximizing total number of discoveries, too many firms might enter into high signal projects maximizing their individual benefit. Intuition here is similar to the business-stealing effect discussed as reason for inefficiently high entry levels in Mankiw and Whinston (1986).

While the condition above is sufficient, it is not necessary. Suppose $x \geq q(p^+/p_0)(K/N)$. In addition, after public signal $p^-V > C_L$, $qK < xN$ and $xN < K$, and $p^- > p^+/2$ so that firms prefer to be the first entrant on a low signal project rather than second entrant on a high signal project. Then the expected number of discoveries is no smaller than $p^+K^+ + p^-(xN - K^+)$, as all low cost firms enter, first on high signal and then on low signal projects. Note that $(p^+(K^+/xN) + p^-(xN - K^+)/xN)xN = (p^- + (p^+ - p^-)(K^+/xN))xN > (p^- + (p^+ - p^-)(K^+/K))xN = p_0xN = p_0\text{min}(N_L, K)$, so the total number of discoveries goes up even if $x \geq q(p^+/p_0)(K/N)$. On the other hand, one can construct examples of decrease in total number of discoveries. Suppose that $p^+/p^- > (N/qK) + 1$ — a sufficient condition for all entry to happen on high signal projects, and $p^- > 0$. If at the same time we have $xN \geq K$, or $qK < xN < K$ and $p^+/p^- < ((1-q)/q)(xN/(K - xN))$, the total number of discoveries goes down as: $p^+qK < p_0\text{min}(N_L, K)$.

Suppose that we were to introduce heterogeneity across types of firms similar to the model without competition: there are senior and junior firms, where juniors are more likely to have high cost. Does the share of discoveries made by young firms increase in response to public signal? If we consider intermediate cost case, then initially only low cost firms explore, and so share of young firm among exploring firms is minimized. Therefore, in response to signal share of young firms entering projects has to weakly increase.[6] It increases strictly as long as at least one of high-cost firms enters in a new equilibrium — for example, $xN + 1 \leq qK$ and $p^{+}V \geq C_H$ would be a sufficient condition for that.

Appendix C: Instrumental Variables Specification

This section presents additional information on the IV analysis described in section 4.3.

Introduction: While a number of specification checks and qualitative fieldwork suggest that the timing of when certain blocks were imaged was unrelated to the evolving prospectivity of different blocks (a key assumption in the differences-in-differences estimation), I also investigate a set of specifications that uses another exogenous source of variation – cloud-cover. Conceptually, the basic idea is to use the fact that some regions (for example, Los Angeles) are rarely cloudy, while others (for example, Seattle) have a significant amount of cloud-cover, to generate variation in the *timing* of the mapping effort. Specifically, I use a variable measuring cross-sectional variation in the average cloud cover in different regions interacted with an indicator variable that equals one if the Landsat program is operational in the continent in which the block is located as the instrument in my analysis.[7] This interacted variable is necessary because the purely cross-sectional cloud-cover variable is absorbed by the block-fixed effects in the analysis. Once the Landsat program has began collecting some images in a given region, low-cloud blocks are more likely to receive low-cloud imagery earlier as compared to blocks with extensive cloud cover. This is partly because these blocks are harder to image because of cloud-cover, but also because NASA administrators anticipated low-cloud cover and imaged such blocks less frequently. The IV analysis then evaluates whether the instrument predicts the timing of the mapping effort and consequently the timing of gold discovery. While the baseline panel specifications help to establish the main effect, these IV estimates provide a separate way to estimate the impact of Landsat mapping on regional gold discovery (while preserving the block and year fixed effects) and helps provide additional confidence in the difference-in-difference results.

Additional Data: I collect additional data from a number of different data-sets at the block and block-year level, to implement the IV strategy. First, I develop a measure of scientific interest in the area of gold geology at the block-year level. This is useful because past work has shown that the scientific literature has an important role in guiding technological search (Fleming and Sorenson, 2004). Accordingly, as a first step, I compile a list of about all 3500 publications related to gold exploration from Scopus that matched my search criteria, which provides a relatively complete index of all major scientific publications. Specifically, I search for terms related to gold mining in journals that belong to the category of "Earth and Planetary Sciences" and "Environmental Science." For each publication, using a "geo-parsing" algorithm, I identify all the geographical entities referenced in the title and abstract of the article, typically the region of the field site of the study. For example, for the article "Glacial fans in till from the Kirkland Lake fault: a method of gold exploration" (Lee, 1963), the geo-parsing algorithm would identify "Kirkland Lake" as a geological feature and a separate geocoding algorithm would provide the exact latitude and longitude of the feature, which can be used to match to a given Landsat block. Using these data and the date of the publication, I link the observation to a block-year observation in my dataset. This procedure helps me calculate the total number of gold-related publications linked to each block-year as the main covariate of interest, allowing me to create a $Pubs_{it}$ measure which captures the time-varying level of scientific research about a given block in the study period.

Second, I use the "Global Earthquake Hazard Frequency and Distribution" database (Dilley et al., 2005; CHRR and CIESIN - Columbia University, 2005), which provides a census of seismic activity for constructing a block-year level measure of earthquake frequency. Geological research has shown that gold mineral-

---

[6] Share among firms making discoveries weakly increases as well, as long as we do not make additional assumptions on differential probability of success for young/senior firms.

[7] As defined by whether at least 2% of all blocks in a continent receive any images.

ization is often associated with earthquakes and related structural activity in the Earth's crust (Weatherley and Henley, 2013). These data capture time-varying measures of seismic activity at the block-year level. I use these data combined with data on scientific publication data to create a gold "prospectivity" score (the potential of a block to contain gold) at the block-year level. In order to construct this score, I regress total gold discoveries before 1972 in a given block on total gold-mining related publications in this period as well as average number of earthquakes and then use the fitted values to predict the likelihood of gold-discovery in the post-Landsat era at the block level. As show in the scatterplots in Appendix Figure D.3, both the level of publications as well as earthquake activity prove to be reasonable predictors of gold discovery at the block level, providing confidence in the validity of the prospectivity score measure.

As a final step, I use data on average cloud cover at the block-level as the basis for an instrument for the timing of Landsat mapping. These data are derived from the MODIS satellites by NASA and measure the average level of cloud cover in the year 2005 at a resolution of 5km X 5km (MODIS Atmosphere Science Team, 2005). These data provide a reasonable proxy for the average cloud cover that any block experiences in a given year, and thereby a good measure for the probability that Landsat map might have been obscured by clouds. I match these data and create a measure of average cloud cover percent corresponding to each Landsat block. This measure is employed in the instrumental variables specifications and a map of this measure at the block-level is presented in Figure D.4.

Validity of IV Specification: To test the validity of the IV specification, Table D.14 Panel B, compares regions that typically have a low level of cloud cover, with regions that are usually cloudy. Contrary to Panel A, these data show that these two types of blocks are comparable in the cross-section in terms of the number of discoveries before 1972, and the prospectivity score.[8] This analysis provides some preliminary evidence to suggest the validity of the IV strategy. The results section will investigate the exclusion restriction more formally.

Consider Panel A in Figure D.3. This figure plots the average cloud cover at the block level collected from weather databases and the average year in which a block was first mapped with a low-cloud image using binned scatterplots. Blocks are binned according to average cloud cover measured up to two decimal digits (i.e. 0.01 to 0.99) on the x-axis, and average first low-cloud year for each of these approximately hundred groups are on the y-axis. Panel A shows a strong positive correlation, indicating that regions with higher levels of cloud cover are more likely to receive mapping information later rather than sooner. Similarly, Panel A, Figure 2 shows the relationship between cloud cover and the average value of the Post Low-Cloud indicator variable. This scatterplot also confirms the intuition that regions with more cloud cover on average have poorer image quality as compared to less cloudy areas.

These data suggest that one could rely on cloud cover to construct a potentially good instrument to understand the role of the Landsat mapping effort on gold discovery. However, for cloud cover to form the basis for a valid instrument, it needs to satisfy the exclusion restriction. In other words, cloud cover must predict gold discovery only through its role in influencing the quality and timing of Landsat mapping, rather than through other channels. For example, if increasingly earthquake-prone regions are also more likely to be cloud-free, then we might doubt the validity of the exclusion restriction because geological research suggests that earthquake-prone regions are also useful targets for gold exploration. Figure D.3 Panel B tests whether the exclusion restriction seems plausible, although it is hard to test it formally. Panel B, Figure 1 analyzes the relationship of the prospectivity score of a block calculated based on the number of publications and the earthquake hazard index with the cloud cover of a region. Panel B, Figure 1 shows a relatively flat relationship between predicted prospectivity and cloud-cover. Similarly, cloud cover does not predict the number of discoveries of gold pre-1972, as indicated by the scatter plot in Panel B, Figure 2, a more direct test of the exclusion restriction. The IV specification relies on the cloud-cover variable and these plots provide confidence that the IV specification is likely to satisfy the exclusion restriction.

IV Results: Table D.11 provides estimates from a differences-in-differences specification similar to the baseline, where the $Post_{it}$ variable is instrumented by the IV. Note that because the cloud-cover variable

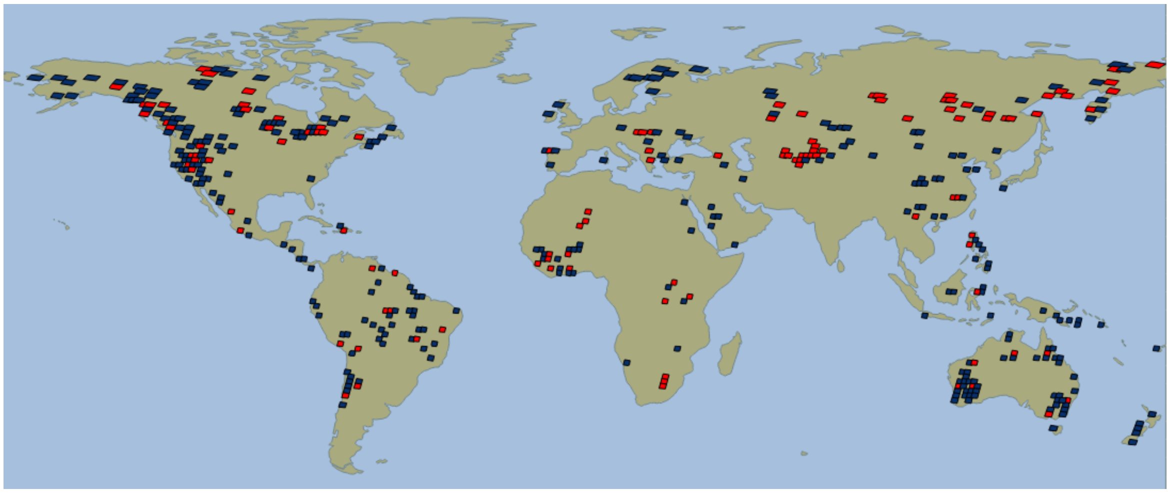
---

[8] There seems to be some difference in the number of publications. However, the difference is in the opposite direction to what would be a concern for the IV specification, i.e. cloudier regions have a higher level of publications as compared to less cloudy regions.

is time-invariant and would be absorbed by the block-level fixed effects, I interact this variable with an indicator variable that equals one if the Landsat program is operational in the continent in which the block is located, and use this interacted variable as $Cloud-cover\ Instrument_{it}$ as the IV in my estimation. Column (1) suggests a strong first-stage between the two variables, i.e. a higher value of cloud cover indicates that the block is likely to receive a low-cloud image later rather than sooner. The IV estimates are presented in Column (2). This estimate is about 1.126— much greater than the baseline estimate. This estimate implies that compared to the average rate of discoveries in a block-year, mapped blocks are about 6 times more likely to report a new discovery, a large and economically significant effect. This large difference between OLS and IV estimates could be attributed to differences in the local average treatment effect of the IV specification. More specifically, it is is possible that the variation arising from cloud-cover is more localized than the satellite mapping variation. This localized variation combined with large and positive inter-block spillovers could generate the patterns I observe, although I'm unable to test this hypothesis directly. In sum, I interpret these results as providing a validity check for the baseline specifications, but continue to emphasize the baseline estimates because they offer more conservative estimates of the impact of the Landsat mapping on gold discovery.

Appendix D: Additional Figures and Tables

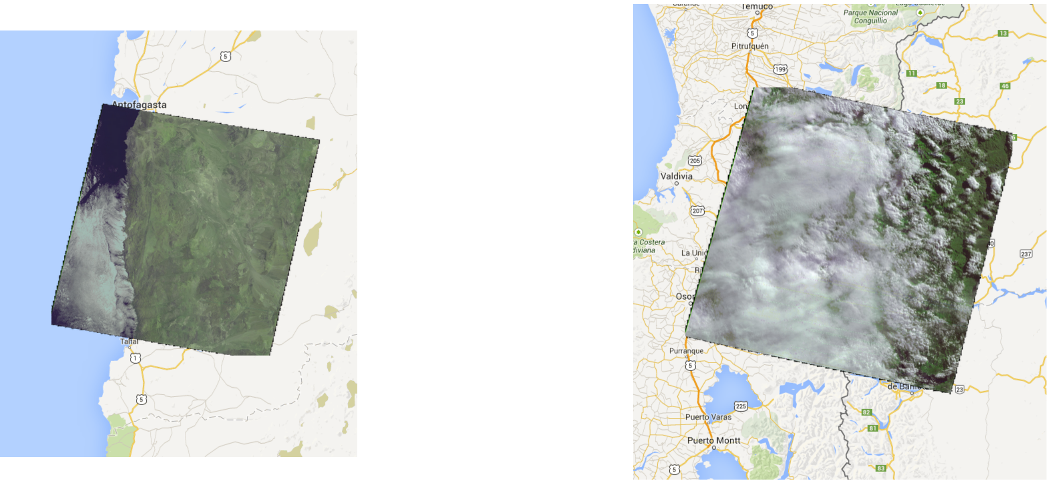
Figure D.1. **Blocks with Gold Discoveries**



Key– 1958-1973: Red and 1973-1988: Blue

Note: This map plots blocks reported gold discoveries of significant size. The blocks are color coded by the first year that discovery was reported since 1958; red if this year was before 1973 and blue if it was after. Note that blocks can (and sometimes do) report discoveries in multiple years, in which case, they are color coded by the first year in which the discovery was reported.

Figure D.2. **Example of Variation in Landsat Coverage**



(1) Block 25177, Chile
Year mapped (*Post Mapped*) = 1973
Year low-cloud map (*Post Low-Cloud*) = 1973

Outcome: Amax Gold Discovery reported in 1980

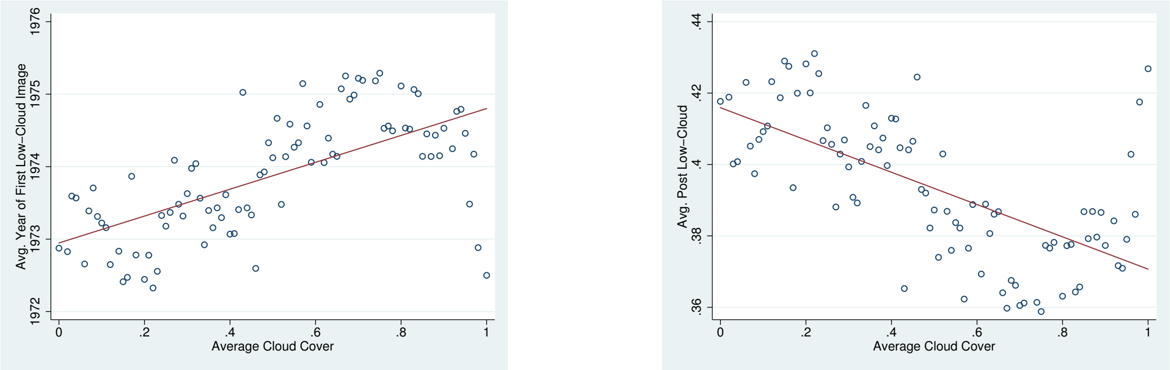
(2) Block 24988, Chile
Year mapped (*Post Mapped*) = 1975
Year low-cloud map (*Post Low-Cloud*) = 1976

Outcome: No discovery reported to date

Note: This figure provides an example that illustrates the research design and data used in the baseline specification using two Landsat blocks in Chile. Figure (1) on the left, shows the best available image for Block 25177 which arrived in 1973 and Figure (2) on the right shows the best available image (with 39% cloud cover) for Block 24988 which arrived in 1976. The block on the left reported a discovery by Amax gold in 1980, while no discovery had been made in the block on the right by 1990.

Figure D.3. **Binned Scatterplots To Evaluate the Validity of the IV Specification**

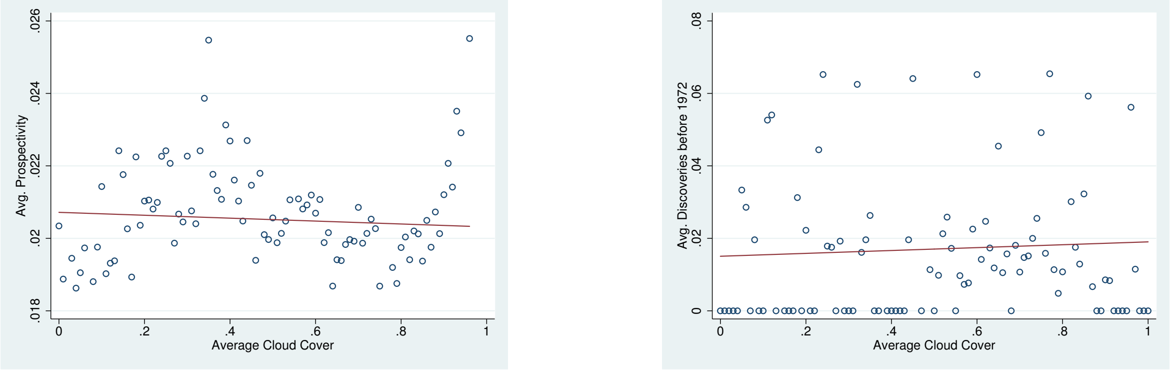
Panel A: First Stage: Cloud Cover and Image Timing



(1) Avg. Year of First Low-Cloud Image

(2) Avg. Post Low-Cloud Indicator

Panel B: Exclusion Restriction: Cloud Cover and Correlates of Gold Discovery

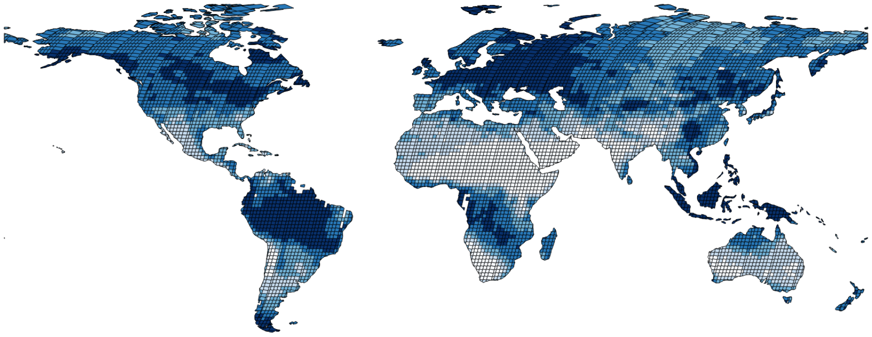


(1) Avg. Predicted Prospectivity

(2) Avg. Discoveries before 1972

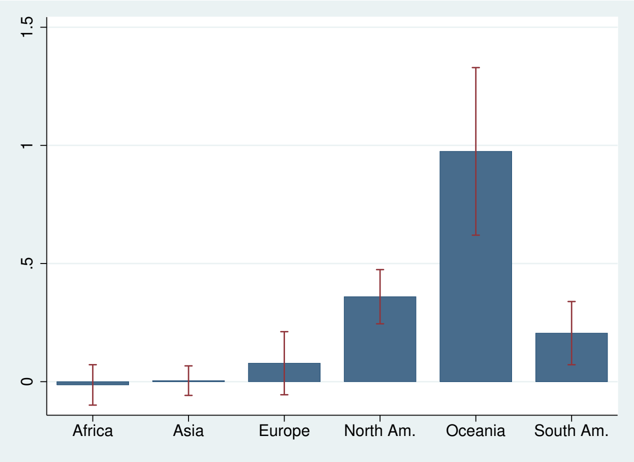
Note: This figure plots the relationship between average annual cloud cover and the timing of Landsat images (Panel A) and between the average annual cloud cover and correlates of gold discovery at the block-level (Panel B). For all four charts, blocks are binned by the level of average annual cloud cover rounded to two decimal digits, and mean value of the variable on the y-axis is calculated. Panel A records the first-stage relationship between cloud-cover and the endogenous variable. Outcome variable in Panel A, Figure 1 is the year in which the block recieved a low cloud cover image, while the variable in Panel A, Figure 2 is the average of the indicator variable for whether a low-cloud image is available. Panel B, tests the correlates of cloud cover with other variables that could affect gold discovery. I predict a “prospectivity” score for a given block as a function of gold-mining publications (before 1972) and earthquake-risk index based on the geology of the region. A mean value of this score is graphed on the y-axis. Similarly, Panel B, Figure 2 plots the average number of discoveries before 1972 on the y-axis.

Figure D.4. **A Heatmap of Average Cloud Cover**



Note: This map plots a heatmap of the distribution of the average cloud cover variable over Landsat blocks. Average cloud cover in the year 2005 is calculated based on MODIS satellite data from NASA (MODIS Atmosphere Science Team, 2005). Blocks in white indicate an average cloud cover of 0-20%, the darkest blocks indicate a cloud cover of 80-100%, and the other three colors indicate the other three quintiles in between.

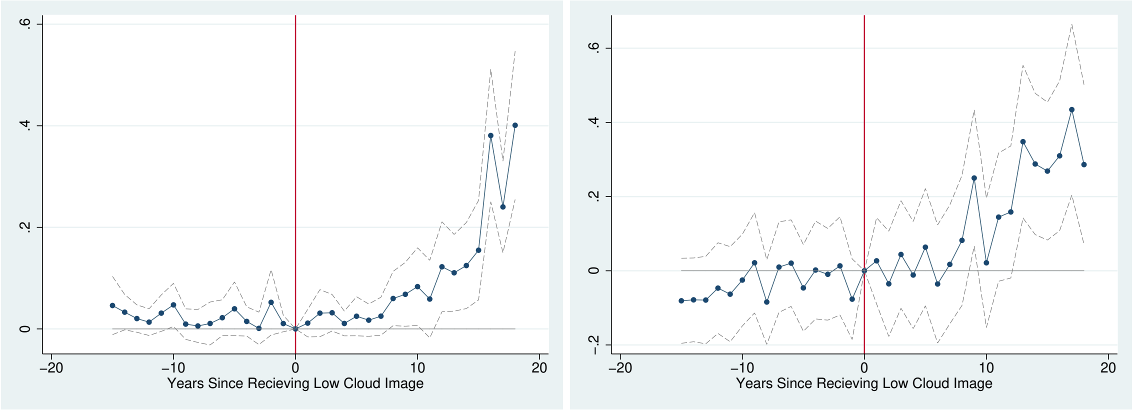
Figure D.5. **Baseline Effects by Region**



Note: This chart shows the baseline estimates separately for each continent on the surface of the earth. Specifically, rather than estimate one coefficient for $Post\ Low - Cloud$, six co-efficients are estimated for $Post\ Low - Cloud\ X\ 1(Continent)$ for each of the six continents depicted above.

Figure D.6. **Juniors vs. Seniors: Impact of Landsat on Gold Discoveries Over Time**

Panel A: Juniors **Panel B: Seniors**



Note: This figure plots estimates (and 95 percent confidence intervals) of β_t from the event study specification specified in Section 4.1B separately for juniors and seniors in Panels A and B. This figure describes the estimated difference between treatment and control blocks for years relative to a "zero year", which marks the year when a block first received low-cloud imagery.

Table D.1. **Different Cutoffs for Low Cloud Cover**

	10 pct.	20 pct.	40 pct.	50 pct.
Post Low-Cloud	0.156***	0.148***	0.180***	0.165***
	(0.0287)	(0.0282)	(0.0262)	(0.0277)
Block FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
N	389213	389213	389213	389213

*+:p<0.15; *:p<0.10; **:p<0.05; ***:p<0.01, Standard errors clustered at block-level shown in parentheses.*

Note: This table presents estimates that use different cutoff points to calculate an image with low cloud cover. In the main specification, any image below 30 percent cloud cover is considered a low cloud image. This table evaluates the baseline regression with a block considered to be mapped with a low cloud image if an image was obtained with cloud cover values below 10%, 20%, 40% or 50% in Columns 1-4 above.

Table D.2. **Alternate Spatial Clustering**

	Any Disc.	Any Disc.	Any Disc.	Any Disc.	Any Disc.	Any Disc.
Post Mapped	0.152***		0.152***		0.152***	
	(0.0315)		(0.0349)		(0.0395)	
Post Low-Cloud		0.164***		0.164***		0.164***
		(0.0306)		(0.0350)		(0.0452)
Block FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Clust. Group	2x3	2x3	5x6	5x6	Country	Country
N	389213	389213	389213	389213	389213	389213

*+:p<0.15; *:p<0.10; **:p<0.05; ***:p<0.01, Standard errors clustered at block-level shown in parentheses.*

Note: Estimates similar to the baseline specification except with different assumptions for the clustering of standard errors. In all specifications, standard errors are clustered at "block groups", which are sets of blocks larger than any given block. In columns (1) and (2), the size of a block group is 2x3 blocks (i.e. 6 blocks in any given block group), while in columns (3) and (4), the size of a block group is 5x6. In columns (5) and (6), the standard errors are clustered at the country-level.

Table D.3. **Negative Binomial Specification**

	Any Discovery	Any Discovery	Any Discovery	Any Discovery
isfind				
Post Mapped	2.640***		1.626**	
	(0.581)		(0.661)	
Post Low-Cloud		1.526***		0.609*
		(0.329)		(0.358)
Block FE	No	No	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
N	389213	389213	18860	18860

*+:p<0.15; *:p<0.10; **:p<0.05; ***:p<0.01, Standard errors clustered at block-level shown in parentheses.*

Note: This table presents estimates for the Landsat program similar to the specification in the baseline specification, except the estimates are produced from negative-binomial regressions, rather than OLS models. Exponentiated versions of the above coefficients provide estimates of the treament elasticities.

Table D.4. **Large Finds Only**

	Any Discovery	Any Discovery	Any Discovery	Any Discovery	Any Discovery
Post Mapped	0.151***		0.0915***		0.0272
	(0.0193)		(0.0245)		(0.0353)
Post Low-Cloud		0.156***		0.0878***	0.0741**
		(0.0196)		(0.0222)	(0.0313)
Block FE	No	No	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
N	389213	389213	389213	389213	389213

*+:p<0.15; *:p<0.10; **:p<0.05; ***:p<0.01, Standard errors clustered at block-level shown in parentheses.*

Note: This table presents estimates for the Landsat program similar to the specification in the baseline specification except the main outcome variable includes only those discoveries categorized a being above 2.2 MOz. in average size, thereby excluding any that are classified as “moderate”, i.e between 0.32 and 2.2 M0z.

Table D.5. **Subsample analysis**

Panel A

	Excluding-USA	Excluding USA/CAN/AUS	Excluding USSR/China
Post Low-Cloud	0.141***	0.0929***	0.0977**
	(0.0270)	(0.0277)	(0.0440)
Block FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
N	363998	291551	289706

Panel B

	Excluding-Pre72	Excluding Unmapped	Only Tree-Covered
Post Low-Cloud	0.157***	0.109***	0.0492
	(0.0243)	(0.0323)	(0.103)
Block FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
N	383719	373592	80975

*+:p<0.15; *:p<0.10; **:p<0.05; ***:p<0.01, Standard errors clustered at block-level shown in parentheses.*

Note: This table presents estimates from baseline DD specification for different subsamples of the data. Panel A excludes key sets of countries, while Panel B excludes certain types of blocks. Panel B, Col 2, excludes all the blocks that were not mapped by the first generation of the Landsat program (i.e. by 1983) and Col 3, includes only those blocks that have substantial tree-cover, where tree cover is coded using codes 1-4 from the GLC2000 dataset indicating all “broadleaved” trees.

Table D.6. **Excluding Continents**

	Africa	Asia	Europe	North Am	Oceania	South Am
Post Low-Cloud	0.192***	0.139***	0.162***	0.128***	0.143***	0.180***
	(0.0306)	(0.0486)	(0.0289)	(0.0287)	(0.0265)	(0.0261)
Block FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	332100	231609	359693	299382	370107	353174

*+:p<0.15; *:p<0.10; **:p<0.05; ***:p<0.01, Standard errors clustered at block-level shown in parentheses.*

Note: This table presents estimates from baseline DD specification for different subsamples of the data. Every column excludes blocks from the focal continent one at a time and evaluates whether the results are driven by any one continent in particular.

Table D.7. **Excluding Blocks Treated Early**

	Excluding 1972	Excluding 1972	Excluding 1972-1973	Excluding 1972-1973
Post Mapped	0.133***		0.0825^{+}	
	(0.0398)		(0.0507)	
Post Low-Cloud		0.115***		0.0902*
		(0.0339)		(0.0473)
Block FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
N	137268	180851	50758	87494

*+:p<0.15; *:p<0.10; **:p<0.05; ***:p<0.01, Standard errors clustered at block-level shown in parentheses.*

Note: This table presents estimates for the Landsat program by excluding blocks that were treated in the first year, or the first two years of the program. A majority of the blocks were treated in this period, so this analysis restricts attention only within blocks that experienced a delay of at least two years or more. Columns (1) and (3) exclude 64.7% (6145) and 86.7% (8255) blocks that were mapped within the first or first two years of the program, respectively while columns (2) and (4) exclude 53.5% (5082) and 77.5% (7359) blocks which were mapped with low-cloud maps within the first, or first two years of the progam respectively.

Table D.8. **Differential Time Trends by Region Types**

	Any Discovery	Any Discovery	Any Discovery	Any Discovery
Post Low-Cloud	0.171***	0.0713**	0.0729**	0.173***
	(0.0260)	(0.0292)	(0.0307)	(0.0253)
Block FE	No	Yes	Yes	Yes
Time FE	Income Grp. X Year	Continent X Year	Subregion X Year	Age of Disc.
N	389213	389213	389213	389213

*+:p<0.15; *:p<0.10; **:p<0.05; ***:p<0.01, Standard errors clustered at block-level shown in parentheses.*

Note: This table presents estimates for the Landsat program similar to the specification in the baseline specification, except instead of a common year fixed effect across regions, time fixed effects are estimated using region-specific time trends. Column (1) includes separate time trends for five different income groups interacted with year dummies, Column (2) includes time trends for six continents (Africa, Asia, Europe, North America, South America, Oceania) interacted with year dummies while Column (3) includes separate time trends for 21 separate subregions (for example, “south-eastern asia” or “western europe”) interacted with year dummies. Column (4) provides these estimates while controlling for the number of years since a previous discovery was made in that block.

Table D.9. **Different Panel Lengths**

	1950-1990	1951-1989	1952-1988	1953-1987	1954-1986	1955-1985
Post Low-Cloud	0.163^{***}	0.160^{***}	0.161^{***}	0.145^{***}	0.131^{***}	0.112^{***}
	(0.0274)	(0.0281)	(0.0278)	(0.0284)	(0.0288)	(0.0293)
Block FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	389213	370227	351241	332255	313269	294283

*+:p<0.15; *:p<0.10; **:p<0.05; ***:p<0.01, Standard errors clustered at block-level shown in parentheses.*

Note: This table presents estimates for the Landsat program by differing length of the panel. In each of the columns, data is only included for years 1950-90, 1951-89, 1952-88, 1953-87, 1954-86 and 1955-90 respectively.

Table D.10. **Placebo Treatment Year**

	Any Discovery	Any Discovery
Post Low-Cloud	-0.0521	-0.0362
	(0.0424)	(0.0395)
Block FE	No	Yes
Year FE	Yes	Yes
N	389213	389213

*+:p<0.15; *:p<0.10; **:p<0.05; ***:p<0.01, Standard errors clustered at block-level shown in parentheses.*

Note: This table presents estimates for the Landsat program similar to the specification in the baseline specification except the "Year Low-Cloud" is replaced by a random integer between 1972 and 1988 for each block, effectively randomizing the treatment year at the block-level. The other variables are calculated as usual.

Table D.11. **Instrumental-variables Estimates for the Impact of Landsat on Gold Discovery**

	Post Low-Cloud	Any Discovery
Cloud-cover Instrument	0.129^{***}	
	(0.00924)	
Post Low-Cloud		1.244^{***}
		(0.474)
Block FE	Yes	Yes
Year FE	Yes	Yes
N	389213	389213
F-Stat	196.09	

*+:p<0.15; *:p<0.10; **:p<0.05; ***:p<0.01, Standard errors clustered at block-level shown in parentheses.*

Note: This table presents instrumental variable estimates relating discovery and discovery-value to the indicator variable for whether a low-cloud image was obtained at the block-year level (Post Low-Cloud). The $Cloud - cover\ Instrument_{it}$ is defined as an interaction between a measure of average annual cloud cover at the block level ($Avg.\ Cloud\ Cover_i$) interacted with an indicator variable for whether the program is operational in the block's continent. Block-year level observations. All estimates are from OLS models and include block and year fixed effects. The sample includes all block-years from 1950 to 1990 (9493 blocks for 41 years implies 389,213 block-year observations). See text and appendix for data and variable descriptions.

Table D.12. **Cross-Sectional Specification: Are Delays Associated with Lower Rate of Discoveries?**

	Any Discovery (72-90)		Any SNL Discovery (82-90)		Years to Discovery Post-72	
	(1)	(2)	(3)	(4)	(5)	(6)
Delay (Years)	-0.148***	-0.105***	-0.0639	-0.0686***	0.0130***	0.00789***
	(0.0479)	(0.0330)	(0.0560)	(0.0153)	(0.00304)	(0.00197)
Fixed Effects	Subregion	Block-Group	Subregion	Block-Group	Subregion	Block-Group
adj. R^2	0.0168	0.0208	0.0463	0.104	0.0106	0.0178
N	9493	9493	9493	9493	9493	9493

*+:p<0.15; *:p<0.10; **:p<0.05; ***:p<0.01, Standard errors clustered at block-level shown in parentheses.*

Note: This table presents cross-sectional estimates from $Y_i = \alpha + \beta_1 \times Delay_i + \gamma_i + \epsilon_i$. $Delay_i$ is the difference between the mapped year and 1972 and γ_i represents spatial fixed effects. Columns (1), (3) and (5) include 21 separate subregion fixed effects (e.g. Central Asia or Northen Africa). Columns (2), (4) and (6) include separate "block-group" fixed effects where all blocks are divided into thirty-four large block-groups. Any SNL Discovery (82-90) refers to discoveries measured from an alternate database from the SNL Metals and Mining group for discoveries. Standard errors are clustered at the same level as the fixed effects.

Table D.13. **Junior and Senior Discovery Accounting for Joint Ventures**

	1(Junior)	1(Junior)	1(Senior)	1(Senior)
Post Mapped	0.0227***		0.125***	
	(0.00523)		(0.0283)	
Post Low-Cloud		0.0351***		0.119***
		(0.00566)		(0.0259)
Percent Gain	**294.64%**	**456.15%**	**182%**	**173.32%**
Block FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
N	389213	389213	389213	389213

*+:p<0.15; *:p<0.10; **:p<0.05; ***:p<0.01, Standard errors clustered at block-level shown in parentheses.*

Note: This table presents estimates for the Landsat program accounting for joint discoveries. All joint-ventures are assumed to be senior-led discoveries, while junior-led discoveries include only those projects where only one junior firm was involved in the discovery.

Table D.14. **Cross-sectional Comparison of Blocks by Landsat Coverage**

Panel A – Comparison by Landsat Timing

	(1) mapped-early	(2) mapped-late	(3) diff	(4) p-val
Discoveries (pre-72)	0.020	0.015	0.006	0.22
Prospectivity Score	1.925	1.796	0.129	0.00
Publications (pre-72)	0.032	0.001	0.031	0.05
Earthquake Hazard	0.683	0.440	0.242	0.00

Panel B – Comparison by Avg. Annual Cloud Cover

	(1) Low Cloud	(2) High Cloud	(3) diff	(4) p-val
Discoveries (pre-72)	0.018	0.020	-0.001	0.76
Prospectivity Score	1.893	1.899	-0.007	0.70
Publications (pre-72)	0.014	0.036	-0.022	0.10
Earthquake Hazard	0.628	0.629	-0.001	0.98

Note: This table compares cross-sectional differences between blocks in terms of four covariates for two different subsamples. Panel A compares blocks mapped early (before 1974) and late (in or after 1974) with low-cloud images by the Landsat program. Panel B compares blocks with low amount of cloudiness (below the median value of 67%) with blocks with high amount of cloudiness. Column (3) is the estimate for the difference in means, and column (4) is the p-value for the t-test that the difference in means is significantly different than zero.

Table D.15. **Robustness: Impact of Landsat on Juniors vs. Seniors**

	Excluding USA/Can/Aus		Excluding USSR/China		Large Disc. Only	
	Junior	Senior	Junior	Senior	Junior	Senior
Post Low-Cloud	0.0151***	0.0789***	0.0483***	0.0545	0.0248***	0.0668***
	(0.00406)	(0.0272)	(0.00827)	(0.0427)	(0.00455)	(0.0216)
Percent Gain	**376.6%**	**119.6%**	**438.9%**	**79%**	**412.8%**	**148.4%**
Block FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	291551	291551	289706	289706	389213	389213

*+:p<0.15; *:p<0.10; **:p<0.05; ***:p<0.01, Standard errors clustered at block-level shown in parentheses.*

Note: This table presents robustness of estimates from baseline DD specification for juniors vs. seniors. The first two columns exclude blocks in USA, Canada and Australia, the next two exclude blocks in the USSR and China and the last two columns include only large discoveries, i.e. those discoveries categorized a being above 2.2 MOz. in average size.