Online Appendix:
"Moonshot: Public R&D and Growth"

Shawn Kantor & Alexander Whalley

May 2025

Contents

1 Variable Definitions and Sources **3**

1.1 Local Manufacturing 3

1.2 National Manufacturing 4

1.3 Space Capability 4

1.4 Patent Counts 6

1.5 Population 8

1.6 NASA and Military Spending 8

1.7 Patent Inventor Migration 9

1.8 Transportation Costs and Other Variables 10

2 Data Construction **10**

2.1 Manufacturing Census Panel 10

2.2 Space Capable Places: Patent and Soviet Space Intelligence Text 11

2.3 Patent Inventor Panel 14

2.3.1 Feature set and similarity profiles 15

2.3.2 Disambiguation blocks 18

2.3.3 Algorithm performance 18

2.3.4 Analysis Dataset 19

3 Modelling Nationwide Effects of Public R&D **19**

3.1 Set Up 20

3.2 Empirical Implementation 23

3.3 Proof ρ is Constant 25

4 References **26**

1 Variable Definitions and Sources

1.1 Local Manufacturing

Value Added: Total difference between value of shipments and expenditure on materials in the county-industry-year unit, 1958 $.

Employment: Total employment in the county-industry-year unit.

Labor Income: Average annual labor income in the county-industry-year unit, 1958 $.

Non-Production Worker Share: Ratio of non-production to production workers in the county-industry-year unit.

Capital Investment: Total investment by manufacturing firms in the county-industry-year unit, 1958 $.

Capital: Measured by the capital stock in the industry-county-year unit. We construct our capital stock measure from the reported investment series using the perpetual inventory method. We follow Bloom, Shankerman, and Van Reenen (2013) and choose the baseline value of the capital stock in 1958: $k_{ij1958} = \frac{i_{ij1958}}{0.08+g_{kij1958}}$, where $g_{kij1958}$ is the growth rate of investment between 1954 and 1958 in the county-industry cell. We follow Bloom, Bond and Van Reenen (2007) in assuming a 8% depreciation rate. Our capital stock measure in years other than 1958 is given by $k_{ijt} = i_{ijt} + (1 - 0.08)^{2.5} k_{ijt-1}$. We assume that investment occurs at the mid-point between the five year differences in manufacturing census years. If investment is missing we assume it is zero, and depreciate the prior capital. If the lagged capital stock is missing, we use the SIC2 capital-employment ratio and observed employment in the industry-county cell to impute lagged capital stock. Measured in nominal dollars.

TFPR, Solow: Estimated total factor revenue productivity using a Cobb-Douglas value added production function with capital and labor as inputs, nominal $. We estimate the production function $Y_{ijt} = A_{ijt} K_{ijt}^{\beta_1} L_{ijt}^{\beta_2}$ to recover manufacturing revenue productivity at the industry-county-year level, A_{ijt}. We estimated this using OLS.

Unit of measurement for all variables: County × SIC 2 digit × year.

Sources for all manufacturing variables – U.S. Bureau of the Census (1947, 1954, 1958, 1963,

1967, 1972, 1977) and U.S. Bureau of the Census (1982, 1987, 1992)

Modality for data access – U.S. Bureau of the Census (1947, 1954, 1958, 1963, 1967, 1972, 1977): Accessed from each year's published volume related to "Area Statistics" (i.e., Volume 3 of each Census) and transcribed by the authors; U.S. Bureau of the Census (1982, 1987, 1992): Accessed electronic records on CD-ROM provided to the authors by the U.S. Bureau of the Census on May 24, 2012.

1.2 National Manufacturing

Output: Real Value of Shipments

Employment: Total Employment

Capital: Total Real Capital Stock

TFP: 4-Factor TFP Index

Unit of measurement for all variables: SIC 2 digit × year.

Sources for all manufacturing variables – Becker, Gray, and Marvakov (2021).

Modality for data access – Data are available electronically from the NBER website. We use the 1987 SIC version.

1.3 Space Capability

Space Score: The county median cosign similarity between full-text pre-1958 patent texts and post-1958 CIA National Intelligence Estimates of Soviet Space Capabilities texts. The *space score* ($\tilde{\rho^C}_i$) is the median value of $\tilde{\rho}_{p<1958}$ across all pre-1958 patents in county i, as describe in appendix section 2.2. This variable is defined at the county-industry level.

High Space Capability: Space capable places are county-industry cells with above median values of the space score variable, as describe in appendix section 2.2. This indicator variable is defined at the county-industry level.

High Space Market: High Space Market are counties above median in terms of our space

market measure $SpaceMarket_o \approx \sum_d \tau_{od}^{-\theta_d} SpaceSpending_d^{\theta}$, and zero otherwise. $\theta = 6.74$. See appendix section 3 for the derivation and discussion of this measure. This indicator variable is defined at the county-industry level level.

County-High Space Capability: Space capable places are county cells with above median values of the space score variable computed at the county level, as describe in appendix section 2.2. This indicator variable is defined at the county level.

US High Space Capability: Space capable places are county-industry cells with above median values of the US space score variable. The US space score measures similarity to US space technology. It is computed based on similarity of technologies in the pre-1958 patents to technologies in the text of post-1958 US NASA patents instead of CIA NIE of Soviet space documents as in our main measure. This indicator variable is defined at the county-industry level.

Mean Space Score: We compute the *space capability score* as the *mean* of $\tilde{\rho}_{p<1958,ij}$ for each county i-industry j, as describe in appendix section 2.2. This variable is defined at the county-industry level.

High Space Capability, Unstemmed: Space capable places are county-industries with above median values of the space score variable, as describe in appendix section 2.2. We compute the *space score* as the median of $\tilde{\rho}^{nostem}_{p<1958,ij}$ for each county i-industry j, as describe in appendix section 2.2. This version calculates the cosign similarity using unstemmed versions of the science direct terms ($\tilde{\rho}^{nostem}_{p<1958,i}$). This indicator variable is defined at the county-industry level.

High Space Capability, All: Space capable places are county-industries with above median values of the space score variable, as describe in appendix section 2.2. We compute the *space score* as the median of $\tilde{\rho}^{nostem}_{p<1958,ij}$ for each county i-industry j, as describe in appendix section 2.2. This version calculates the cosign similarity using stemmed versions of the science direct terms ($\tilde{\rho}^{nostem}_{p<1958,i}$). This version uses the full set of NIE documents that cover space technology to compute the median as indicated in appendix Table A1 column 'All'. This indicator variable is defined at the county-industry level.

High Space Capability, Exclusive: Space capable places are county-industries with above median values of the space score variable, as describe in appendix section 2.2. We compute the *space score* as the median of $\tilde{\rho}^{nostem}_{p<1958,ij}$ for each county i-industry j, as describe in appendix

section 2.2. This version calculates the cosign similarity using stemmed versions of the science direct terms ($\tilde{\rho}^{nostem}_{p<1958,i}$). This version only uses the subset of NIE documents that cover space technology exclusively to compute the median as indicated in appendix Table A1 column 'Exclusive'. This indicator variable is defined at the county-industry level.

High Space Capability, County Level: Space capable places are county cells with above median values of the space score variable, as describe in appendix section 2.2. This indicator variable is defined at the county level.

High Space Capability 75^{th} Percentile, County Level: Space capable places are county cells with above 75^{th} percentile values of the space score variable, as describe in appendix section 2.2. This indicator variable is defined at the county level.

Sources for all patent variables – All Space-capable variables are derived from the authors' calculations using the full text of patent documents, the full text of CIA national intelligence estimates of Soviet space capabilities, and the Science Direct Technology Corpus of terms. More details on the creation of the variables are provided in appendix section 2.2 below. The patent full text data are available for download at: Patent Grant Full Text Data/APS of the USPTO, available at https://bulkdata.uspto.gov/ (post-1976 data) and Google Cloud Console to access the Google Patents Public Datasets with their BigQuery language. The link is: https://cloud.google.com/blog/topics/public-datasets/google-patents-public-datasets-connecting-public-paid-and-private-patent-data and clicking on the hyperlink Google Patents Public Datasets brings you to the cloud console. See online appendix table A1 for sources to the CIA intelligence documents. The Science Direct Technology Corpus list is available here: https://www.sciencedirect.com/topics/index. The authors accessed this list on 10/20/2021.

1.4 Patent Counts

Total Patents: Total patents measured at the county-industry-year unit level. Patents with authors from more than one county are assigned fractionally to multiple counties based on county authorship share.

Pre-1958 Patents: Total patents between 1947 and 1958 at the county-industry-year level.

NASA Patents: Total NASA patents in the county-industry-year unit. NASA patents are

those in which NASA is either the assignee or listed as a funder on the patent. Patents with authors from more than one county are assigned fractionally to multiple counties based on county authorship share.

Army Patents: Total Army patents in the county-industry-year unit. Army patents are those in which the Army is either the assignee or listed as a funder on the patent. Patents with authors from more than one county are assigned fractionally to multiple counties based on county authorship share.

Navy Patents: Total Navy patents in the county-industry-year unit. Navy patents are those in which the Navy is either the assignee or listed as a funder on the patent. Patents with authors from more than one county are assigned fractionally to multiple counties based on county authorship share.

Non-NASA Patents: Total Non-NASA patents in the county-industry-year unit. These patents include those for which NASA is neither an assignee nor listed as a funder on the patent. Patents with authors from more than one county are assigned fractionally to multiple counties based on county authorship share.

Military Text Patents: Total patents in the county-industry-year unit with above median similarity to military technologies based on an inverse document term frequency measure. Military technologies are obtained from the intersection of the U.S. Department of Defense (2021) glossary of military terms with the Science Direct set of technology terms. Patents with authors from more than one county are assigned fractionally to multiple counties based on county authorship share.

Unit of measurement for all variables (unless stated otherwise): County $\times$ SIC 2 digit $\times$ year.

Sources for all patent variables – U.S. Patent and Trademark Office (2014, 2020); Lybbert and Zolas (2014); Marco, et al. (2015); Petralia, et al. (2016a,b); Fleming, et al. (2019); and authors' calculations. Information on assignment or funding of patents is drawn from Fleming, et al. (2019) for patents before 1976 and from USPTO (2020) PatentsView for those 1976 and after.

Modality for data access – U.S. Patent and Trademark Office (2014): Accessed electronic records from internet; U.S. Patent and Trademark Office (2020): Accessed electronic records

from internet (data used from the following files: "assignee," "g-inventor-location," "g-inventor-location-disambiguated," "g-location-not-disambiguated," "government-organization," "Patent-assignee," and "patent-gointorg"); Lybbert and Zolas (2014): Accessed electronic records from internet; Marco, et al. (2015): Accessed electronic records from internet; Petralia, et al. (2016b): Accessed electronic records from internet; Fleming, et al. (2019): Accessed electronic records from internet; U.S. Department of Defense (2021): Converted PDF to text file for textual analysis. A terms of use of the Petralia, et al. (2016b) data is a citation to their 2016a publication.

1.5 Population

Population: Total population in the county-year unit.

High School Graduate Percentage: Percentage of adults age 25+ who are high school graduates in the county-year unit.

Unit of measurement for all variables: County × year.

Source for all variables – Haines (2010).

1.6 NASA and Military Spending

NASA Spending (1958$): Total cumulative NASA contractor spending in a county-industry in the manufacturing census year, plus spending in years since the last the census. Locations of contractor activity, not corporate headquarters, are utilized. The top 100 contractors account for the vast majority of NASA prime contractor spending. The maximum share of NASA contractor spending occurring at the top 100 is 92.86% and the minimum share is 87.17%. NASA contractor data drawn from Van Nimmen and Bruno (1976, 191-208), Gawdiak (1994, 184-224), Rumerman (2000, 541-70), and Rumerman (2012, T-453-64). NASA total spending data are from Van Nimmen and Bruno (1976, 115), Gawdiak (1994, 148), Rumerman (2000, 507) and Rumerman (2012, T-414).

Military Spending: Total military contract spending in the county-year unit. The 1947-1966 data are based on state-level data, but allocated using 1967 SIC2-county weights from that year's military spending applied to each location and industry. The 1967 to 1992 data

are based on totals on individual contracts over $10,000 at the SIC2-county level. The earliest year for the state data is 1951. The data we use for 1947 is based on reported values in 1951. The post-1967 contract data only have industry in a few years. We use the federal supply codes for equipment cross-walked to SIC2 industries to get the industry level data. Unit of measurement: County × SIC 2 digit × year. Sources: U. S. Department of Defense (1975, 1981, and 2006); and authors' calculations.

Modality for data access – Van Nimmen and Bruno (1976), Gawdiak (1994), and Rumerman (2000, 2012): All data were accessed from published Volumes and transcribed by the authors. U.S. Department of Defense (1981) contains state-level contract data that are available in a published volume and were transcribed by the authors. U. S. Department of Defense (1975 and 2006): Accessed electronic files from internet.

1.7 Patent Inventor Migration

County In-Migration of Patent Inventors: We denote origin locations o and destination locations d. The number of inventors who move from o to d is P_{odt} and the number of inventors who begin in o and do not move, P_{oot}, so that $\log\left(\frac{P_{odjt}}{P_{oojt}}\right)$ is the log odds ratio for inventor migration. This variable is measured at the origin-destination-industry county pair level for the 1945 to 1992 patent application years. Source: Authors' calculation using patent inventor panel data as described in section 2.3 of the appendix.

Corporate Income Tax Rate: The additional tax burden accruing to a firm in the top tax bracket in state s for an additional one dollar of revenue if all of its operations were in s. Measured at the state-year level. Source: Akcigit, Grigsby, Nicholas, and Stantcheva (2021).

Personal Average Income Tax Rate, 90^{th} Percentile: The total tax burden for an individual at the 90^{th} percentile of the national income distribution divided by that individual's total income. Measured at the state-year level. Source: Akcigit, Grigsby, Nicholas, and Stantcheva (2021).

R&D Credit: Statutory credit rate adjusted for recapture and type of credit for a given state-year. Source: Akcigit, Grigsby, Nicholas, and Stantcheva (2021).

Modality for data access – Akcigit, Grigsby, Nicholas, and Stantcheva (2021): Accessed electronic files from internet.

1.8 Transportation Costs and Other Variables

Transportation Costs (τ_{ij}): County-to-county transportation costs in 1960. This measure is based on the 1959 Rand McNally Road Atlas highway network to compute the travel costs between all county pairs in the contiguous United States in each year. Transportation costs are computed by measuring the road surface, using historical sources for travel speed by road surface type and legislated speeds. Monetary travel costs are obtained by using the per mile wage of a truck driver multiplied by the travel time plus the per mile fuel cost times the distance. See Jaworski and Kitchens (2019) for more details. Unit of measurement: Origin-Destination County pair. Sources: Jaworski and Kitchens (2019); authors' calculations.

Defense Scientist (1962): Number of research scientists who have received funding from a defense agency before 1962 in the county. Source: National Register of Scientific and Technical Personnel (NSF 1962).

Research Scientist (1962): Number of research scientists in 1962 in the county. Source: National Register of Scientific and Technical Personnel (NSF 1962).

IBM Mainframe Computers (1961): Number of IBM mainframes installed before 1961. Unit of measurement: County. Source: IBM (1962).

Modality for data access – Jaworski and Kitchens (2019): Digital records obtained from authors; National Register of Scientific and Technical Personnel (NSF 1962): Digital records provided to the authors on CD-ROM from the National Archives (variables employed in this analysis: City, State, and Government Sponsorship); IBM (1962): Confidential records obtained from IBM archives and transcribed by authors.

2 Data Construction

2.1 Manufacturing Census Panel

We obtain data from the Censuses of Manufactures of 1947, 1954, 1958, 1963, 1967, 1972, 1977, 1982, 1987, and 1992. These census data provide reporting at the county and MSA levels. We standardize the reporting to measure consistent quantities and monetary values across data years for SIC 2-digit industries. Some large metro counties do not report separately from the MSA in 1963, 1967, 1972, and 1977. We apportion these MSA values to counties using the share of employment in an industry-county cell in a MSA-industry

cell. We take the average of this apportionment factor in 1958 and 1982 to apportion the 1963-1977 MSA-industry data to county-industry cells.

We drop observations that report: (1) missing value added or employment in a year; (2) less than eight times during the sample; (3) no space sector in the county; and (4) no pre-1958 patents. These restrictions lead to a loss of 2,597 observations. We also drop the observations that appear in ND, SD, or WY because only a single county in each state reported manufacturing data. Our analysis sample contains 6,759 county-industry observations from 86 counties and 19 two-digit SIC industries from 1947 to 1992.

2.2 Space Capable Places: Patent and Soviet Space Intelligence Text

Corpus of Technology Concepts: To compute the similarity between full-text USPTO patent documents and CIA National Intelligence Estimates of Soviet Space Capabilities documents we employ three sources. First, we obtain the Science Direct (SD) corpus of Technology Terms.[1] The SD corpus of technology concepts consists of 193,715 expressions comprising Science Direct (SD) Topics. Similar to the well-known Medical Subject Headings (MeSH) terms, the vocabulary in the SD Topics indexes articles in SD in order to improve information retrieval for researchers. Unlike the medical focus of MeSH terms, the SD Topics cover all scientific disciplines represented in SD.

Full Patent Text: Two sources are used for extracting full patent texts for the time period from January 1940 to December 1991, inclusive. Full patent text since 1976 is available in the US Patent and Trademark Office's Patent Full-Text and Image Database (PatFT) USPTO (2021a). For the time period of 1940-1975, description and claims text was extracted using the Google Patents Public Datasets on BigQuery (USPTO (2021b)). Access modality: accessed electronic files from internet.

National Intelligence Estimates (NIEs): The CIA's now-declassified NIEs are authoritative intelligence assessments related to the Soviet Union's capabilities with regard to space flight, among a number of other geopolitically sensitive areas.[2] During the period of 1946-

[1]The list is available here: https://www.sciencedirect.com/topics/index. The authors accessed this list on 10/20/2021.

[2]These declassified documents, which are readily available in the public domain, were sourced primarily

1991, these documents provide estimates of Soviet scientific and technical capabilities in space. Some of the documents also focus on military technology. We exclude the documents that primarily focus on military technology from our baseline analysis. We compute an additional patent text similarity measure using only documents with an exclusive space focus as a robustness exercise. Access modality: accessed PDF files from internet, converted to text files for textual analysis. The source locations of the specific documents we use are listed in appendix Table A1.

Data Pre-Processing: English stop words were removed, and Porter stemming was applied to the SD Topics corpus, as well as to the full patent and NIE texts.[3] To reduce the dimension of the SD Topics, we then dropped stemmed topics appearing less than 1,000 times over the full set of patent texts, and then dropped the top 1% most frequently occurring terms. The most frequently occurring terms that were dropped include "Copper," "Dye," "Gridding," and "Duct," for example. SD Topics containing more than four words are also dropped. The combination of stop word removal, stemming, dropping infrequent and too frequent terms, and dropping Topics comprising more than four words result in a dictionary of 25,767 technology concepts.

Text Similarity: We first compute the text similarity between each patent and each NIE document in appendix table A1 using a cosine similarity measure. This process is implemented by following the steps:

1. Construct a document term matrix containing frequency counts for each SD Topic in this representative NIE document (after data cleaning);

2. Construct document term matrices for each U.S. patent (after data cleaning);

3. For each patent document term matrix, compute the cosine similarity against the representative NIE:

$$\rho_{p<1958,n>1958} = \frac{\sum_{i=1}^{n} TF_{i,n>1958} TF_{i,p<1958}}{\sqrt{\sum_{i=1}^{n} TF_{i,n>1958}^2} \sqrt{\sum_{i=1}^{n} TF_{i,p<1958}^2}} \tag{A1}$$

where $\rho_{p<1958,n>1958}$ is cosine similarity between a patent document issued before 1958 ($p < 1958$) and an NIE document issued after 1958 ($n > 1958$). $TF_{i,n>1958}$ is the term frequency

from the CIA's Freedom of Information Act Electronic Reading Room. Links to specific documents are recorded in appendix Table A1. More information regarding these records and obtaining physical access is available at https://www.cia.gov/readingroom/collection/declassified-national-intelligence-estimates-soviet-union-and-international-communism.

[3]Both the English stop words removal and the Porter2 stemming were achieved using the SnowballStemmer function from the Natural Language Toolkit (NLTK) Python module.

for SD term i in NIE document $n > 1958$, and $TF_{i,p<1958}$ is the term frequency for SD term i in patent document $p < 1958$. Exhibit A1 in the online appendix provides a visual example of highly similar pages in an NIE document and patent document captured by this approach with the SD technology terms highlighted.

Space Capable Places: We aggregate patent level NIE similarities to the county-industry level to determine space capable places. We first compute median of $\rho_{p<1958,n>1958}$ at the patent level across all NIE documents after 1958 to obtain $\tilde{\rho}_{p<1958}$. The *space score* ($\tilde{\rho^C}_i$) is the median value of $\tilde{\rho}_{p<1958}$ across all pre-1958 patents in county i - indusry j. County-industries with high values of $\tilde{\rho_i^C}$ (high space place capability score) have pre-Sputnik technologies represented in the county-industry, as evidenced through their patent records, similar to the space technologies that the Sovets possessed after 1958. County-industries with low values of $\tilde{\rho_i^C}$ (low space place capability score) had pre-Sputnik technologies within the county-industry that were not similar to later Soviet space technologies. Our *space capability place* variable takes a value of 1 for county-industries with above median values of $\tilde{\rho_i^C}$, and a value of zero otherwise.

Validation: Our textual similarity measure captures spaceflight technological similarity regardless of how patents were classified by the Patent Office. Examples of SD technology terms most frequent in patents owned or funded by NASA, shown in appendix table A2, include "Aircraft," "Antennae," and "Propellant." Examples of SD technology terms most frequent in NIE space technology intelligence reports, shown in appendix table A3, include "Missiles," "Satellites," and "Orbitals." Appendix table A4 reports the SD terms occurring frequently in *both* NIE and patent documents. Such terms as "Aircraft," "Spacecraft," and "Satellites" are frequently found in both types of documents. While the majority of frequent technology terms in our examples are closely connected to spaceflight, they do suggest one potential limitation of our measure. NIE intelligence and U.S. patent documents may be textually similar because of matching non-spaceflight technologies. For instance, military technology terms such as "Warhead" or "Missile" – even though they were commonly used to describe rocket technology at the time – could lead to a high similarity score even if their space-relevance might be low. In appendix table A5 we show that our patent-level space score based on textual similarity between the patent and NIE technologies strongly predicts NASA ownership or funding of a patent, conditional on military funding, technological area,

and county fixed effects.[4] We further address this concern by showing that our main results below are robust to a broad range of controls for military research and spending (Table 4) and point out here that the spatial correlation between military and NASA R&D activity is small (appendix table A6).

2.3 Patent Inventor Panel

We build a panel of inventors by disambiguating the inventors listed on all USPTO patents from 1947 to 1992. We follow the disambiguation procedure of Li, et al. (2014) to determine if a pair of patent-inventor records belongs to a single individual. This task is a clustering problem which is addressed using an Authority machine learning approach (see Torvik and Smalheiser 2009). Given a training data set of disambiguated inventors, we cluster inventors in our historical patent data based on a similarity profile. Following Akcigit, Grigsby, Nicholas and Stantcheva (2022), we measure similarity across a pair of inventors using combinations of inventor-level features – inventor name and location – and patent-level features – patent assignee, technology class, and coauthor network. Intuitively, the algorithm assigns a higher probability of two patent-inventor records belonging to the same individual when the two patents are technologically similar, or share the same assignees, trace back to geographically close locations, etc.

The ideal implementation of the disambiguation algorithm considers the similarity across all pairs of inventors in the historical patent records available through Google Patents (GP). With over 3.8 million patent-inventor records during our period of analysis (1920-1980), this translates to over 14.4 trillion inventor pairs. To reduce the computational burden of the ideal implementation, we adopt the iterative blocking approach from Akcigit, Grigsby, Nicholas and Stantcheva (2022). The starting point is to compare only record pairs within a block of inventors sharing an exact first and last name. Later iterations allow for increasingly

[4]The regression results reported in online appendix table A5 are patent level and are specified as:

$$NASA_l = \omega_1 + \omega_2 \text{Space Score}_l + \gamma_t + \nu_l, \tag{A2.2.1}$$

where $NASA_l$ takes a value of 1 if the patent is a NASA patent and zero otherwise. $SpaceScore_l$ measures the cosign similarity between patent l and all NIE documents, as discussed above. We expect ω_2 to be positive if our measure captures technological similarity to NASA demands. We report versions of this model that also control for NBER technology subcategory fixed effects, other government involvement in the patent, and county fixed effects.

larger blocks by comparing, for example, inventors with a same first initial and exact last name. The purpose of the iterative blocking approach is to (1) reduce the computational cost of the algorithm, and (2) allow for different feature sets when constructing the similarity of a pair of patent-inventors. The exact implementation of our disambiguation algorithm is described below.

2.3.1 Feature set and similarity profiles

Feature set. We compare two records by constructing pairwise similarity profiles $\mathbf{x}$ using a set of features $x_1, ..., x_k$. Each available feature x_i is encoded as follows.

- Middle name: middle names (and first and last names as well) are constructed from the inventor full name field. Once constructed, we compare the middle name feature for a pair of records and assign one of of the following alternatives.
  (a) The middle names match exactly (e.g. "JAMES" and "JAMES");
  (b) One record has a full middle name (with length greater than a single letter). The other record has only a middle initial which matches the initial of the other record (e.g. "JAMES" and "J");
  (c) Both records have missing middle name;
  (d) One of the two records have a missing middle name (e.g. "JAMES" and " ");
  (e) Otherwise (e.g. "JAMES" and "EDWARD").
- Location: we geolocate patent-inventors by linking patent numbers and inventor names with latitude and longitude information available at the Comprehensive Universe of U.S. Patents data.[5] Once each patent-inventor record is geolocated, we measure geodesic distance for a pair of records and assign one of the following alternatives:
  (a) The two inventors are located less than 1 mile apart;
  (b) The two inventors are located from 1 to under 10 miles apart;
  (c) The two inventors are located from 10 to under 25 miles apart;
  (d) The two inventors are located from 25 to under 50 miles apart;

[5]The data set was compiled by Enrico Berkes. Data inquiries should be directed to `enrico.berkes@gmail.com`.

(e) Either the two inventors are located 50 or more miles apart, or at least one record has a missing location).

- Patent technology classes: this feature uses the first reported U.S. Patent Class (USPC) code of a patent record. Comparison across a pair of records yields the following assignment:

  (a) The USPC codes are identical;

  (b) The USPC codes are not identical.

- Assignee: the assignee feature is constructed from the harmonized_assignee field in the Google Patents Big Query data base. We retain only the first assignee listed in the patent record. Comparison of a pair of records yields the following assignment:

  (a) The two patent records yield a Jaro-Winkler (JW)-based similarity metric of at least 0.9;

  (b) $0.8 \leq \text{JW} < 0.9$;

  (c) $0.7 \leq \text{JW} < 0.8$;

  (d) One of the two patent records have a missing assignee;

  (e) Otherwise.

- Coauthor network: we construct the coauthor network feature by assigning, to each patent, a list of all the patent's inventors. The list is alphabetically sorted, and uses each inventor's first and last name. We compare coauthor networks between two patent-inventor records and assign:

  (a) The patent coauthors list is identical between the pair of records;

  (b) Otherwise.

Similarity profile. We construct pairwise similarity profiles using the features above. For example, in an iteration where similarity is defined using the middle name and location features, the similarity profile for two records with middle names "JAMES" and "J" geolocated within a mile from one another is the vector $\mathbf{x} =< b, a >$.

Treating the disambiguated inventor data from Li, at al. (2014) as a "ground truth" training set, we compute the probability that each profile $\mathbf{x}$ belongs to the same inventor.

We construct this probability from the count of records with profile X that belong to the same inventor versus the count that belong to different inventors. Let M denote the event that a patent-inventor pair is a match (i.e., belong to the same individual) and N the event that it is a non-match. Using Bayes rule, the probability of a match M given an observed similarity profile X is:

$$P(M \mid \mathbf{x}) = \frac{P(\mathbf{x} \mid M)P(M)}{P(\mathbf{x} \mid M)P(M) + P(\mathbf{x} \mid N)(1 - P(M))}. \tag{A2}$$

The posterior probability $P(M \mid X)$ has a one-to-one relationship with the posterior odds of a match, defined as:

$$\frac{P(M \mid \mathbf{x})}{1 - P(M \mid \mathbf{x})} = \frac{P(M)}{1 - P(M)} \frac{P(\mathbf{x} \mid M)}{P(\mathbf{x} \mid N)}. \tag{A3}$$

Eq A3 can be converted back to *Eq A*2, and defining the likelihood ratio $r(\mathbf{x}) = \frac{P(\mathbf{x}|M)}{P(\mathbf{x}|N)}$, we get:

$$P(M \mid \mathbf{x}) = \frac{1}{1 + (\frac{1-P(M)}{P(M)})(\frac{1}{r(\mathbf{x})})}. \tag{A4}$$

From Eq A4, two components are needed to determine the posterior probability of a match given an observed similarity profile: the matching prior and the likelihood ratio. The prior match probability $P(M)$ at each iteration of the algorithm is the ratio of within-cluster matched pairs in a block over the total number of pairs in the block. The likelihood ratio $r(\mathbf{x})$ is determined directly from the training set by taking the ratio of times the similarity profile $\mathbf{x}$ lead to matched events M versus non-matched events N. The training set consists of disambiguated data from Li, at al. (2014)[6], which we use to compute posterior match probabilities for each similarity profile at each blocking algorithm. This data set contains over 9 million patent-inventor instances from over 4 million patents issued between 1975-1999. The underlying assumptions from using these data for our disambiguation task are that the inventor identifiers are accurately assigned in the training data, and also that there would be no systematic differences in the posterior match probabilities for patents in our historical records of 1920-1980 and patents in the training set years of 1975-1999.

[6]Available at the Harvard Dataverse Network at `https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/5F1RRI`. Use the compressed file "invpat_final.zip"

2.3.2 Disambiguation blocks

After each iteration, we say that two records originate from the same inventor if the computed posterior match probability exceeds 0.85. The blocking iterations are described below:

Iteration 1. Block inventors based on exact first and last name. Construct similarity profiles based on middle name and location.

Iteration 2. Same as Iteration 1.

Iteration 3. Same as Iteration 1.

Iteration 4. Block inventors based on exact first and last name. Construct similarity profiles based on assignee, patent technology class, coauthor network, and middle name.

Iteration 5. Block inventors based on first five letters of first name and exact last name. Construct similarity profiles based on assignee, patent technology class, coauthor network, and middle name.

Iteration 6. Block inventors based on first three letters of first name and exact last name. Construct similarity profiles based on assignee, patent technology class, coauthor network, and middle name.

Iteration 7. Block inventors based on initial of first name and exact last name. Construct similarity profiles based on assignee, patent technology class, coauthor network, and middle name.

Iteration 8. Block inventors based on initial of first name and exact last name. Construct similarity profiles based on middle name and location.

2.3.3 Algorithm performance

For the purposes of measuring algorithm performance and optimizing the cutoff parameter c of the posterior match probability, we subset a random sample 73,562 patent-inventor instances from 67,443 patent records from the Li, et al. (2014) data. We refer to this subset of the data as the held-out test set. Upon computing the posterior match probabilities in the training set (while holding out the test set), we ran the disambiguation iterations described above on this hold-out test data. We varied the cutoff parameter c in the set

$\{0.8, 0.85, 0.9, 0.95, 0.99\}$ and computed splitting (S) and lumping (L) performance statistics, defined as:

$$S = \frac{\sum_i \{\mathbf{x} \mid \mathbf{x} \in U_i, \mathbf{x} \notin V_i\}}{\sum_i |U_i|}, \tag{A5}$$

$$L = \frac{\sum_i \{\mathbf{x} \mid \mathbf{x} \in V_i, \mathbf{x} \notin U_i\}}{\sum_i |V_i|}. \tag{A6}$$

In Eqs A5 and A6, U_i denotes the set of patents for inventor i on the ground truth disambiguation of the Li, et al. (2014) data, while V_i denotes the largest set of patents for inventor i based on our disambiguation algorithm. The splitting and lumping statistics using the held-out test set in the searched range are shown below. For our disambiguation of historical patent records, we chose the cutoff $c = 0.85$ which minimizes the sum of splitting and lumping in the held-out test set.

2.3.4 Analysis Dataset

Our disambiguation identifies 882,072 U.S. inventors who were jointly granted 2.4 million patents for our sample period 1945 to 1992. Our main analysis in Table 5 uses a subsample of inventors in the top 50% of lifetime inventors, who have at least 5 lifetime patents. This subsample has 105,458 U.S. inventors who were jointly granted 1,198,146 patents.

Splitting and lumping statistics in held-out test set using different cutoffs of the posterior match probability

Cutoff	.80	.85	.90	.95	.99
Splitting	.01504	.02069	.03636	.15756	.43979
Lumping	.08360	.07626	.06326	.03212	.01447
Splitting + Lumping	.09864	.09696	.09961	.18969	.45426

3 Modelling Nationwide Effects of Public R&D

In this section we describe our approach to estimate the effect of public R&D on nationwide economic outcomes. We use a simple county to county trade model based on Donaldson and Hornbeck (2016) and for ease of exposition we follow their notation and presentation closely. We differ from their model, however, in that we focus on the impact of public

R&D spending, holding transportation infrastructure fixed and introducing market-level consumption externalities.

3.1 Set Up

We index counties by o if they are origin of trade and d if they are destinations. Consumers have CES preferences over a continuum of differentiated product varieties, where the elasticity of substitution across varieties is given by σ. Producers in each county combine a fixed factor land (L_o), and mobile factors labor (N_o) and capital (K_o) using a Cobb-Douglas technology to produce varieties. Public R&D reduces unit costs for firms in location o. The marginal cost of each variety is given by:

$$MC_o(j) = \frac{s_o^{-1} q_o^{\alpha} w_o^{\gamma} r^{1-\alpha-\gamma}}{z_o(j)} \tag{A8}$$

where s_o captures the unit cost effect of public R&D, q_o is the land rental rate, w_o is the wage, r is the interest rate, and $z_o(j)$ is the local productivity shifter drawn from a Frechet distribution with a CDF $F_o(z) = e^{-A_o z^{-\theta}}$. A_o captures the local knowledge stock, and θ captures the standard deviation of the knowledge stock.

Trade costs between o and d are iceberg: for each unit shipped from o to d, $\tau_{od} \geq 1$ is the cost to ship. That is, if a variety is produced and sold in the same county the price is $p_{oo}(j)$, while the same variety sold in a different county has price $p_{od}(j) = \tau_{od} p_{oo}(j)$.

Production and Prices. By assuming perfect competition, unit costs (including marginal and trade costs, as well as public R&D effects) are equal, letting consumers buy from the cheapest origin county. Using the assumption that $r_c = r$, Donaldson and Hornbeck (2016) note that the price index in destination d is defined by

$$(P_d)^{-\theta} = \kappa_1 \sum_o A_o (s_o^{-1} q_o^{\alpha} w_o^{\gamma})^{-\theta} \tau_{od}^{-\theta} = CMA_d \tag{A9}$$

where $\kappa_1 = \left[\Gamma\left(\frac{\theta+1-\sigma}{\theta}\right)\right]^{\frac{1}{1-\sigma-\theta}} r^{(1-\alpha-\gamma)\theta}$.

The price index in county d will fall with an increase in public R&D in any origin market, with a weight that is declining in trading costs. Bilateral trade between counties then implies that location specific public R&D may affect prices in other markets, where the size of the cross-market effect depends on the cost of trading. The inverse transformation of the price

index reflecting customer access to cheap products is commonly termed Consumer Market Access and denoted CMA_d.

Workers and Amenities. Turning to workers, we assume that they are perfectly mobile across space as our goal is to understand spatial equilibrium implications. We depart from Donaldson and Hornbeck (2016), however, by including both consumer market access as a positive amenity of a location and an exogenous fixed utility level ($\bar{u}$) that is common across locations. As a result of workers' endogenous location choice, workers' utility levels are equalized across counties in equilibrium and, hence, real wages satisfy:

$$\frac{w_o}{P_0} = \bar{u} CMA_o^{\epsilon} \tag{A10}$$

We include a consumer city amenity where access to a larger set of varieties (Glaeser, Kolko and Saiz, 2001) increases utility where the spatial scope for agglomeration amenities is beyond just a county of residence. For example, a worker in Princeton, NJ obtains amenities not just from the variety of products available in Princeton, NJ, but also those accessible in New York City. The strength of the market access amenity is captured by the parameter ϵ.

Output. We obtain output in a county by summing up exports to all other locations. Eaton and Kortum (2002) give the following gravity equation for exports from o to d.

$$X_{od} = A_o (s_o^{-1} q_o^{\alpha} w_o^{\gamma})^{-\theta} \tau_{od}^{-\theta} \kappa_1 CMA_d^{-1} Y_d \tag{A11}$$

Total output in county o is the summation of exports to all other counties, so that

$$Y_o = \sum_d X_{od} = \kappa_1 A_o (s_o^{-1} q_o^{\alpha} w_o^{\gamma})^{-\theta} \sum_d \tau_{od}^{-\theta} CMA_d^{-1} Y_d \tag{A12}$$

Multilateral market access for the origin county, termed Firm Market Access, is defined as:

$$FMA_o = \sum_d \tau_{od}^{-\theta} CMA_d^{-1} Y_d \tag{A13}$$

Thus, output in county o is given by

$$Y_o = \kappa_1 A_o (s_o^{-1} q_o^{\alpha} w_o^{\gamma})^{-\theta} FMA_o \tag{A14}$$

This expression suggests intuitively that output is increasing in location specific productivity, A_o, cost reductions from public R&D, s_o, and firm market access, FMA_o. Output is falling in local factor prices for labor, w_o, and land, q_o.

Solving the Model. We solve the model for equilibrium output to obtain our estimation equation. First, we solve for the county labor supply relationship in (A10) in terms of nominal wages to substitute into equation (A14) and obtain

$$Y_o = \kappa_2 A_o s_o^{\theta} q_o^{-\theta\alpha} CMA_o^{\frac{\gamma-\epsilon\gamma\theta}{\theta}} FMA_o \tag{A15}$$

where $\kappa_2 = \kappa_1 \bar{u}$.

Next, we express CMA_d in terms of FMA_o. We can use equation (A14) to solve for $(\kappa_1 A_0 s_o^{-1} q_o^{\alpha} w_o^{\gamma})$ and substitute into (A9) to get:

$$CMA_d = \sum_o \tau_{od}^{-\theta} FMA_o^{-1} Y_o \tag{A16}$$

Under symmetric trade costs (i.e., $\tau_{od} = \tau_{od}$), equation (A16) and the definition of FMA_o implies there exists a constant ρ such that $FMA_o = \rho CMA_o$.[7] We further define $MA_0 \equiv FMA_o = \rho CMA_o$. Substituting for $FMA_o = MA_o$ and $CMA_0 = \frac{MA_o}{\rho}$ into equation (A15) and rearranging we obtain,

$$Y_o = \kappa_3 A_o s_o^{\theta} q_o^{-\theta\alpha} MA_o^{1+\frac{\gamma-\epsilon\gamma\theta}{\theta}} \tag{A17}$$

where $\kappa_3 = \kappa_2 \rho^{(\epsilon\gamma\theta-\gamma)}$.

Market Effects of Public R&D. Finally, we express MA_0 as a function of s_d so that market effects can incorporate public R&D unit cost shocks. Solving for market access $MA_o \equiv FMA_o = \sum_d \tau_{od}^{-\theta} CMA_d^{-1} Y_d$, $CMA_d = \left(\frac{MA_d}{\rho}\right)$, and $Y_d = \kappa_3 A_d s_d^{\theta} q_d^{-\theta\alpha} MA_d^{1+\frac{\gamma-\epsilon\gamma\theta}{\theta}}$ from equation (A14), we obtain

$$MA_o = \kappa_3 \rho \sum_d \tau_{od}^{-\theta} A_d s_d^{\theta} q_d MA_d^{1+\frac{\gamma-\epsilon\gamma\theta}{\theta}} \tag{A18}$$

Market access in county o increases in response to a distance-weighted sum of county d's productivity (A_d), public R&D (s_d), land values (q_d), and multilateral market access (MA_d). These supply side fundamentals affect the income level in county d and thus take the place of county d's output in Donaldson and Hornbeck (2016)'s market access formulation.

Estimation Equation. We can obtain an estimation equation by taking logs of equation (A15) which leads to:

$$\log(Y_o) = \psi_Y + \theta \log(s_0) + \left(1 + \frac{\gamma - \epsilon\gamma\theta}{\theta}\right) \log(MA_o) - \theta\alpha \log(q_o) + \log(A_o) \tag{A19}$$

[7] See Appendix section 3 for the proof of this proposition. Donaldson and Hornbeck (2016) and Allen and Arkolakis (2014) make use of a similar relationship.

where $\psi_Y = \log(\kappa_3)$ is a constant. The local effect of public R&D cost shocks in the same county is captured by θ, the trade elasticity. The local public R&D effect is positive.

The market effect of public R&D cost shocks – that is, $\left(1 + \frac{\gamma - \epsilon\gamma\theta}{\theta}\right)$ – can be either positive or negative, depending on two key factors. A first force is that public R&D in location d increases location d's income, which, in turn, increases exports of goods from location o, which results in an increase in output by firms in the origin county. A positive market access effect is standard in this class of models. A second mitigating force is migration. Public R&D in locations near d increase amenities that may induce workers in county o to move. With a sufficiently strong amenity effect the negative effects of migration away from o can outweigh the positive effects of increasing income in location d so that market-level public R&D results in a reduction in output. Thus, the sign of the market effect of public R&D is an empirical question.[8]

3.2 Empirical Implementation

Our goal is to use an econometric model to estimate the national manufacturing effects of NASA activity. To do so we add market effects from the spatial model into our baseline econometric model using the specification,

$$\begin{aligned}\log(Y_{ijt}) =& \beta_1 + \beta_2 \text{Space Capable Place}_{i,<1958} \times \text{ Space Race}_t + \\ & \beta_3 \text{Space Capable Place}_{i,<1958} \times \text{post-Space Race}_t + \\ & \beta_4 \text{High Space Market}_{i,<1958} \times \text{Space Race}_t + \\ & \beta_5 \text{High Space Market}_{i,<1958} \times \text{Post-Space Race}_t \\ & + \delta_i + \gamma_t + \text{Total Pre-1958 Patents}_i \times \gamma_t + S_i \times \gamma_t + \nu_{ijt}.\end{aligned} \tag{A22}$$

Here the outcome variables are the log of a manufacturing variable in county i, industry j and year t, such as value added or employment. Again, we expect β_2 and β_3 to be positive

[8] We can also derive an expression for employment in county o using equation (10) and $W_o N_0 = Y_o \gamma$ to obtain

$$N_o = \kappa_4 A_o s_o^{\theta} q_o^{-\theta\alpha} M A_o^{\frac{1+\theta-\gamma-\epsilon\gamma\theta-\epsilon\gamma}{\theta}} \tag{A20}$$

where $\kappa_4 = \kappa_3 \frac{\gamma}{u}$. Taking logs we obtain an estimation equation for employment as,

$$\log(N_o) = \psi + \theta \log(s_o) + \left(\frac{1+\theta-\gamma-\epsilon\gamma\theta-\epsilon\gamma}{\theta}\right) \log(\sum_d \tau_{od}^{-\theta} s_d^{\theta}) - \theta\gamma \log(q_o) + \log(A_o) \tag{A21}$$

where $\psi_Y = \log(\kappa_4)$ is a constant. Again, local public R&D has a positive effect on employment. Market level public R&D has an ambiguous effect on employment depending on the parameter values.

as places that were specialized in Space Race-relevant technologies before it began in 1958 were likely to experience more NASA activity after 1958.

Market effects of public R&D are captured by β_4 and β_5. Market effects may be positive implying national effects would be larger than local effects due to strong cross-county demand or productivity effects. Alternative, market effects may be negative implying national effects are smaller than local effects due to strong cross-county migration effects. The sign and magnitude of market effects are an empirical question.

Our empirical implementation in (A22) differs from our model-derived estimation equations in (A19) for a number of reasons. First, our market access measures build on our research design using the same source of variation as our main analysis. We follow Donaldson and Hornbeck (2016) in applying further assumptions to make our market access term empirically tractable.

First, we assume that our space score measure indexes the unit cost effect of space R&D in our model (s_d) and that we can approximate space-driven market access in equation (A19) in county o with

$$SpaceMarket_o \approx \sum_{d \neq o} \tau_{od}^{-\theta} \text{S}paceCapabilityScore_d^{\theta} \tag{A23}$$

Our approximation for space market access for county o in era e focuses on a distance-weighted average of space scores across destination locations. It allows us to obtain a market access measure when land value (q_d) and productivity (A_d) are not reported in manufacturing census years. It does not include differences across regions in terms of multilateral effects (MA_d since we treat the market access from county d as part of the constant term). It does not include income in county d unlike Donaldson and Hornbeck (2016) because our supply side fundamentals take the place of income in our formulation.

We do not include origin county space scores in the market activity measure to separately identify only market-level effects. All variation in market activity comes from shocks in space scores elsewhere that are likely exogenous to county-level outcomes. For this reason we do not seek instruments for space market activity.

Second, construction of our market-level activity measure, $\text{S}paceMarket_i$, requires θ values. We use $\theta = 8.28$, the preferred estimate from the meta analysis in Head and Meyer (2014).

Lastly, we retain the median contrast in our main specification for estimation of market effects. The variable $HighSpaceMarket_i$ takes a value of 1 for counties with above median $SpaceMarket_i$ and zero otherwise.

3.3 Proof ρ is Constant

Define key equations

$$FMA_o = \sum_d \tau_{od}^{-\theta} CMA_d^{-1} Y_d \tag{A24}$$

$$CMA_d = \sum_o \tau_{od}^{-\theta} FMA_o^{-1} Y_o \tag{A25}$$

and

$$FMA_o = \sum_d \tau_{od}^{-\theta} \left(FMA_d\right)^{-1} Y_d \tag{A26}$$

and

$$FMA_o = \rho_o CMA_o \tag{A27}$$

Step 1: Rearrange (A27) for CMA_0, change index from o to be d to substitute into (A24) for CMA_d so that:

$$FMA_o = \sum_d \tau_{od}^{-\theta} \left(\frac{FMA_d}{\rho_d}\right)^{-1} Y_d \tag{A25}$$

and

$$FMA_o = \sum_d \rho_d \tau_{od}^{-\theta} \left(FMA_d\right)^{-1} Y_d \tag{A26}$$

Step 2: Substitute in for CMA_d into (A25) so that:

$$\frac{FMA_d}{\rho_d} = \sum_o \tau_{od}^{-\theta} FMA_o^{-1} Y_o \tag{A27}$$

then

$$FMA_d = \rho_d \sum_o \tau_{od}^{-\theta} FMA_o^{-1} Y_o \tag{A28}$$

Changing indexes we get

$$FMA_o = \rho_o \sum_d \tau_{od}^{-\theta} FMA_d^{-1} Y_d \tag{A29}$$

Noting that (A26) and (A29) are both expressions for FMA_o, we can see that only $\rho_o = \rho_d = \rho$ can be a solution to this system of equations.

4 References

Akcigit, Ufuk, John Grigsby, Tom Nicholas, and Stefanie Stantcheva. 2021, *Replication Data for: "Taxation and Innovation in the 20th Century."* Cambridge, MA: Oxford University Press; distributed by Harvard Dataverse, Cambridge, MA. https://doi.org/10.7910/DVN/SR410I.

Akcigit, Ufuk, John Grigsby, Tom Nicholas, and Stefanie Stantcheva. 2022. "Taxation and Innovation in the 20th Century." *Quarterly Journal of Economics* 137(1): 329–85.

Becker, Randy A., Wayne B. Gray, and Jordan Marvakov. 2021. "NBER-CES Manufacturing Industry Database (1958-2018, version 2021a)." Cambridge, MA: National Bureau of Economic Research. https://www.nber.org/research/data/nber-ces-manufacturing-industry-database (accessed October 20, 2021).

Bloom, Nick, Stephen Bond, and John Van Reenen. 2007. "Uncertainty and Investment Dynamics." *Review of Economic Studies* 74(2): 391-415.

Bloom, Nick, Mark Schankerman, and John Van Reenen. 2013. "Technology Spillovers and Product Market Rivalry." *Econometrica* 81(4): 1347–93.

Central Intelligence Agency. 1946-1991. "National Intelligence Estimates." Washington, DC: Central Intelligence Agency Freedom of Information Act Electronic Reading Room. https://www.cia.gov/readingroom/home (accessed August 14, 2024).

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

Eaton, Jonathan, and Samuel Kortum. 2002. "Technology, Geography, and Trade." *Econometrica* 70(5): 1741-79.

Fleming, Lee, Hillary Green, Guan-Cheng Li, Matt Marx, and Dennis Yao. 2019. "Replication Data for: Government-Funded Research Increasingly Fuels Innovation." Washington, DC: American Association for the Advancement of Science; distributed by Harvard Dataverse, Cambridge, MA. https://doi.org/10.7910/DVN/DKESRC.

Gawdiak, Ihor, with Helen Fedor. 1994. *NASA Historical Data Book, Vol. IV: NASA Resources, 1969-1978.* Washington, DC: National Aeronautics and Space Administration.

Glaeser, Edward L., Jed Kolko, and Albert Saiz. 2001. "Consumer City." *Journal of Economic Geography* 1(1): 27–50.

Haines, Michael R. 2010. "Historical, Demographic, Economic, and Social Data: The United States, 1790-2002." Distributed by Inter-university Consortium for Political and Social Research, Ann Arbor, MI. https://doi.org/10.3886/ICPSR02896.v3.

Head, Keith, and Thierry Mayer. 2014. "Gravity Equations: Workhorse, Toolkit, Cookbook." In *Handbook of International Economics Vol. 4*, edited by Gita Gopinath, Elhanan Helpman and Kenneth Rogoff. New York: Elsevier.

IBM. 1962. "IBM1401 Customer List." International Business Machines Archives, Armonk, NY.

Jaworksi, Taylor, and Carl T. Kitchens. 2019. "National Policy for Regional Development: Historical Evidence from Appalachian Highways." *Review of Economics and Statistics* 101(5): 777–90.

Li, Guan-Cheng, Ronald Lai, Alexander D'Amour, David M. Doolin, Ye Sun, Vetle I. Torvik, Amy Z. Yu, and Lee Fleming. 2014. "Disambiguation and Co-authorship Networks of the U.S. Patent Inventor Database (1975–2010)." *Research Policy* 43(6): 941-55

Lybbert, Travis J., and Nikolas J. Zolas. 2014. "Getting Patents and Economic Data to Speak to Each Other: An 'Algorithmic Links with Probabilities' Approach for Joint Analyses of Patenting and Economic Activity." *Research Policy* 43(3): 530-42.

Marco, Alan C., Michael Carley, Steven Jackson, and Amanda F. Myers. 2015. "The USPTO Historical Patent Data Files: Two Centuries of Innovation." USPTO Economic Working Paper 2015-1, Available at https://ssrn.com/abstract=2849514. Data distributed by USPTO, Washington, DC. https://www.uspto.gov/ip-policy/economic-research/research-datasets/historical-patent-data-files.

National Science Foundation. 1962. "National Register of Scientific and Technical Personnel, 1962." Distributed by the National Archives, College Park, MD (Record Group 307: electronic public-use data files available by request from the National Archives).

Petralia, Sergio, Pierre-Alexandre Balland, and David L. Rigby. 2016a. "Unveiling the Geography of Historical Patents in the United States from 1836 to 1975." *Scientific Data*, 3

(160074).

Petralia, Sergio, Pierre-Alexandre Balland, and David L. Rigby. 2016b. Replication data for "Unveiling the Geography of Historical Patents in the United States from 1836 to 1975." London, United Kingdom: Springer-Nature; distributed by Harvard Dataverse, Cambridge, MA. http://dx.doi.org/10.7910/DVN/BPC15W.

Rummerman, Judy A. 2000. *NASA Historical Data Book, Vol. VI: NASA Space Applications, Aeronautics and Space Research and Technology, Tracking and Data Acquisition/Support Operations, Commercial Programs, and Resources, 1979-1988.* Washington, DC: National Aeronautics and Space Administration.

Rummerman, Judy A. 2012. *NASA Historical Data Book, Vol. VIII: NASA Earth Science and Space Applications, Aeronautics, Technology, and Exploration, Tracking and Data Acquisition/Space Operations, Facilities and Resources, 1989–1998 — Tables.* Washington, DC: National Aeronautics and Space Administration.

Torvik, Vetle I, and Neil R. Smalheiser. 2009. "Author Name Disambiguation in MEDLINE." *ACM Transactions on Knowledge Discovery from Data (TKDD)* 3(3): 1-29.

U.S. Bureau of the Census. 1947, 1954, 1958, 1963, 1967, 1972, 1977. *Census of Manufactures, Volume 3.* Washington, DC: Government Printing Office.

U.S. Bureau of the Census. 1982, 1987, 1992. *Census of Manufactures.* Washington, DC: Government Printing Office. Digital files provided to authors on nine CD-ROMs from the Bureau of the Census on 5/24/2012. CD-ROM description identifiers: C1-E82-CMFG-26-US1; C1-E87-EC1E-09-US1; C1-E87-EC2B-09-US1; C1-E92-EC1J-07-US1; C1-E92-EC2B-07-US1; C1-E97-NAFI-17-US1; C1-E97-NA3Z-17-US1; V1-E02-NAIA-17-US1; and V1-E02-NA2Z-17-US1.

U.S. Department of Defense. 1975. "Records of Prime Contracts Awarded by the Military Services and Agencies, created, 7/1/1965 - 6/30/1975, documenting the period 7/1/1965 - 6/30/1975." Distributed by the National Archives, College Park, MD. https://aad.archives.gov/aad/series-description.jsp?s=492 (accessed October 20, 2021).

U.S. Department of Defense, Washington Headquarters Services, Directorate for Information, Operations and Reports. 1981. "Military Prime Contract Awards by State: Fiscal Year 1951 to Date." Distributed by Hathi Trust, Ann Arbor, MI. https://hdl.handle.net/2027/

uiug.30112001844346.

U.S. Department of Defense. 2006. "Records of Prime Contracts Awarded by the Military Services and Agencies, created, 7/1/1975 - 9/30/2006, documenting the period 7/1/1975 - 9/30/2006." Distributed by the National Archives, College Park, MD. https://aad.archives.gov/aad/series-description.jsp?s=503 (accessed October 20, 2021).

U.S. Department of Defense. 2021. "DOD Dictionary of Military and Associated Terms, As of November 2021," Joint Publication 1-02. Washington, DC: Government Printing Office.

U.S. Patent and Trademark Office. 2014. "Historical Master File." Washington, DC: United States Patent and Trademark Office. https://bulkdata.uspto.gov/data/patent/historical/2014/annual.dta (accessed January 15, 2016).

U.S. Patent and Trademark Office. 2020. "PatentsView." Washington, DC: United States Patent and Trademark Office. https://www.patentsview.org/download/ (accessed January 15, 2017). Data have been archived at: https://zenodo.org/records/13835830.

U.S. Patent and Trademark Office. 2021a. "Patent Grant Full Text Data (No Images) (JAN 1976 - PRESENT)." Washington, DC: United States Patent and Trademark Office. https://developer.uspto.gov/product/patent-grant-full-text-dataxml (accessed January 15, 2019). Data have been archived at: https://doi.org/10.5281/zenodo.13137561.

U.S. Patent and Trademark Office. 2021b. "Google Patents Public Datasets on BigQuery." Washington, DC: United States Patent and Trademark Office. https://cloud.google.com/blog/topics/public-datasets/google-patents-public-datasets-connecting-public-paid-and-private-patent-data (follow directions to Google Patents Public Data; data set ID: "patents-public-data:patents"; accessed January 15, 2019). Data have been archived at: https://doi.org/10.5281/zenodo.13137561.

Van Nimmen, Jane, and Leonard C. Bruno, with Robert L. Rosholt. 1976. *NASA Historical Data Book, Vol. I: NASA Resources, 1958-1968.* Washington, DC: National Aeronautics and Space Administration.

Table A1: Documents Detailing National Intelligence Assessments of Soviet Space Capabilities

Document	Date	Title	Baseline; Post-1958	All	Space Exclusive
ORE 3/1	1946-10-31	Soviet Capabilities For The Development And Production Of Certain Types Of Weapons & Equipment (https://www.cia.gov/readingroom/docs/CIA-RDP67-00059A000200130011-3.pdf)		Included	
NIE 3	1950-11-15	Soviet Capabilities And Intentions (https://www.cia.gov/resources/csi/static/Watching-the-Bear-5-Chap5-Mil-Intentions-Capabilities.pdf)		Included	
SE 10	1951-09-15	Soviet Capabilities For A Surprise Attack On The Continental United States Before July 1952 (https://www.cia.gov/readingroom/docs/DOC_0000269249.pdf).		Included	
SE 36	1953-03-05	Soviet Capabilities For Attack On The US Through Mid-1955 (https://www.cia.gov/readingroom/docs/CIA-RDP79S01011A000800050011-5.pdf)		Included	
NIE 65	1953-06-16	Soviet Bloc Capabilities Through 1957 (https://www.cia.gov/readingroom/docs/nie_65.pdf)		Included	
NIE 11-6-54	1954-10-05	Soviet Capabilities And Probable Programs In The Guided Missile Field (https://www.cia.gov/readingroom/docs/nie_11_6_54.pdf)		Included	
NIE 11-12-55	1955-12-20	Soviet Guided Missile Capabilities And Probable Programs (https://www.cia.gov/readingroom/docs/nie_11_12_55.pdf)		Included	
NIE 11-5-57	1957-03-12	Soviet Capabilities And Probable Programs In The Guided Missile Field (https://www.cia.gov/readingroom/docs/DOC_0006707795.pdf)		Included	
NIE 11-5-58	1958-08-19	Soviet Capabilities In Guided Missiles And Space Vehicles (https://www.cia.gov/readingroom/docs/CIA-RDP79R01012A011100040006-0.pdf)	Included	Included	
NIE 11-5-59 (Adv)	1959-09-08	Soviet Capabilities In Guided Missiles And Space Vehicles-Advance Portion (https://www.cia.gov/readingroom/docs/DOC_0000267655.pdf)	Included	Included	
NIE 11-5-59	1959-11-03	Soviet Capabilities In Guided Missiles And Space Vehicles (https://www.cia.gov/readingroom/docs/nie_11_5_59.pdf)	Included	Included	
NIE 11-5-61	1961-04-25	Soviet Technical Capabilities In Guided Missiles And Space Vehicles (https://www.cia.gov/readingroom/docs/DOC_0000278407.pdf)	Included	Included	
NIE 11-1-62	1962-12-05	The Soviet Space Program (https://www.cia.gov/readingroom/docs/DOC_0000283833.pdf)	Included	Included	Included
NIE 11-1-65	1965-01-27	The Soviet Space Program (https://www.cia.gov/readingroom/docs/DOC_0000283836.pdf)	Included	Included	Included
NIE 11-1-67 M/H	1968-04-04	The Soviet Space Program (https://www.cia.gov/readingroom/docs/DOC_0000283812.pdf)	Included	Included	Included
NIE 11-1-69	1969-06-19	The Soviet Space Program (https://www.cia.gov/readingroom/docs/DOC_0000283814.pdf)	Included	Included	Included
NIE 11-8-68 M/H	1969-06-23	Soviet Strategic Attack Forces (https://www.cia.gov/readingroom/docs/DOC_0000278487.pdf)	Included	Included	

NIE 11-1-69 M/H	1970-03-26	The Soviet Space Program (https://www.cia.gov/readingroom/docs/DOC_0000283816.pdf)	Included	Included	Included
NIE 11-1-71	1971-07-01	The Soviet Space Program (https://www.cia.gov/readingroom/docs/DOC_0000283819.pdf)	Included	Included	Included
NIE 11-1-73	1973-12-20	Soviet Space Programs (https://www.cia.gov/readingroom/docs/DOC_0000283822.pdf)	Included	Included	Included
IIR 74-018	1974-10-01	A Soviet Land-Mobile ICBM: Evidence Of Development And Considerations Affecting A Decision On Deployment (https://www.cia.gov/readingroom/docs/CIA-RDP79R01099A001600080003-0.pdf)	Included	Included	
IIM (1975)	1975-11-01	Soviet Dependence on Space Systems (https://www.cia.gov/readingroom/docs/DOC_0000380221.pdf)	Included	Included	Included
NIE 11-1-80	1980-08-06	Soviet Military Capabilities And Intentions In Space (https://www.cia.gov/readingroom/docs/DOC_0000284010.pdf)	Included	Included	
NI IIM 82-10006	1982-06-01	The Technology Acquisition Efforts Of The Soviet Intelligence Services (https://www.cia.gov/readingroom/docs/CIA-RDP84B00049R001403390009-8.pdf)	Included	Included	
NIE 11-1-83	1983-07-15	The Soviet Space Program (https://www.archives.gov/declassification/iscap/pdf/2011-030)	Included	Included	Included
NIE 11-1/7-84	1984-11-26	Potential For The Transfer Of Space Technology To The Soviet Union (https://www.archives.gov/declassification/iscap/pdf/2011-030)	Included	Included	
NI IIA 84-10006	1984-12-01	Soviet Approach To Nuclear Winter (https://www.cia.gov/readingroom/docs/CIA-RDP87T00413R000100170003-9.pdf)	Included	Included	
NIE 11-1-85J	1985-12-01	Soviet Space Programs (https://www.archives.gov/files/declassification/iscap/pdf/2011-030-doc-6.pdf)	Included	Included	Included
NI IIM 86-10002	1986-03-21	Soviet Military Production, 1974-85 (https://www.cia.gov/readingroom/docs/DOC_0000802748.pdf)	Included	Included	
NI IIM 87-10006JX	1987-05-22	Soviet Military Production, 1975-86 (https://www.cia.gov/readingroom/docs/CIA-RDP89M00699R002201800005-2.pdf)	Included	Included	
SOV 88-10063	1988-09-01	Soviet Reusable Space Systems Program: Implications for Space Operations in the 1990s (https://www.cia.gov/readingroom/docs/DOC_0000500641.pdf)	Included	Included	Included
NIE 11-3/8-91W	1991-08-08	Soviet Capabilities For Strategic Nuclear Conflict Through the Year 2000 (https://www.cia.gov/readingroom/docs/DOC_0000265648.pdf)	Included	Included	

Notes: This table lists the various intelligence estimates of Soviet space capabilities that are used in this paper. Our baseline measure of High Space Capability uses the documents noted as "Included" in the column "Baseline; Post-1958" (i.e., those from 1958 and beyond). Our All-NIE High Space Capability measure uses all available relevant intelligence documents listed in the table. Our space-exclusive-NIE High Space Capability measure uses the intelligence documents indicated as "Included" in the last column.

Table A2: 25 Most Frequent Science Direct Technology Topics in NASA Patent Documents

Unstemmed Term	Stemmed Term	Topic Rank in NASA Patents	Topic Rank in NIE Documents
Aircraft	aircraft	1	17
Antennae	antenna	2	99
Nationalism	nation	3	14
Transducer	transduc	4	13652.5
Amplitudes	amplitud	5	13652.5
Spacecraft	spacecraft	6	13
Specimen	specimen	7	649.5
Governance	govern	8	70.5
Modelers	model	9	78.5
Wavelength	wavelength	10	832.5
United States of America	unit state of america	11	1246.5
Instrumentalism	instrument	12	53.5
Propellant	propel	13	32.5
Reflectors	reflector	14	1246.5
Waveform	waveform	15	1246.5
Equator	equat	16	399.5
Provisioning	provis	17	155.5
Satellites	satellit	18	3
Emittance	emitt	19	649.5
Multiplication	multipl	20	73.5
Acceleration	acceler	21	399.5
Ceramer	ceram	22	13652.5
Factorization	factor	23	23
Minimality	minim	24	214
Actualization	actual	25	51

Notes: The first column reports the Science Direct Technology Topic in unstemmed form and the second column reports the stemmed form of the Science Direct topic. The third and fourth columns report the ranking of each term with respect to its appearance in NASA patent documents and in the CIA national intelligence estimates of Soviet space capabilities documents, respectively.

Table A3: 25 Most Frequent Science Direct Technology Topics in National Intelligence Estimates of Soviet Space Capabilities Documents

Unstemmed Term	Stemmed Term	Topic Rank in NASA Patents	Topic Rank in NIE Documents
Missiles	missil	762	1
USSR	ussr	16565	2
Satellites	satellit	18	3
Estimability	estim	138	4
Orbitals	orbit	34	5
Secretions	secret	8846	6
Intelligibility	intellig	1782.5	7
Defensiveness	defens	8115.5	8
Scientification	scientif	947.5	9
Warhead	warhead	18245.5	10
Directorate	director	2392	11
Payload	payload	165.5	12
Spacecraft	spacecraft	6	13
Nationalism	nation	3	14
Westernization	western	6151.5	15
Germanate	german	4295	16
Aircraft	aircraft	1	17
Mobilization	mobil	476	18
Altitude	altitud	198	19
Space Stations	space station	183	20
Reconnaissance	reconnaiss	13452.5	21
Lates	late	1745	22
Factorization	factor	23	23
Basicity	basic	81	24
Economics	econom	493	25

Notes: The first column reports the Science Direct Technology Topic in unstemmed form and the second column reports the stemmed form of the Science Direct topic. The third and fourth columns report the ranking of each term with respect to its appearance in NASA patent documents and in the CIA national intelligence estimates of Soviet space capabilities documents, respectively.

Table A4: 25 Most Frequent Science Direct Technology Topics Occurring in Top 0.5% of both NASA patents and NIE documents

Unstemmed Term	Stemmed Term	Topic Rank in NASA Patents	Topic Rank in NIE Documents
Nationalism	nation	3	14
Aircraft	aircraft	1	17
Spacecraft	spacecraft	6	13
Satellites	satellit	18	3
Orbitals	orbit	34	5
Propellant	propel	13	32.5
Factorization	factor	23	23
Instrumentalism	instrument	12	53.5
Actualization	actual	25	51
Governance	govern	8	70.5
Observability	observ	40	41.5
Modelers	model	9	78.5
Multiplication	multipl	20	73.5
Antennae	antenna	2	99
Basicity	basic	81	24
Publicity	public	73	46
Calibrator	calibr	52	91
Identifiability	identifi	95	50
Physicalism	physic	61	88.5
Criticality	critic	108	45
Pastes	past	130.5	27
Simulators	simul	54	106.5
Affectivity	affect	127	40
Interference	interfer	128	55.5
Commercialization	commerci	106	97.5

Notes: The first column reports the Science Direct Technology Topic in unstemmed form and the second column reports the stemmed form of the Science Direct topic. The third and fourth columns report the ranking of each term with respect to its appearance in NASA patent documents and in the CIA national intelligence estimates of Soviet space capabilities documents, respectively. The set of terms represented in this table are the top 25 terms in the intersection of the top 132 terms from the NASA patents and the top 132 terms in the NIE documents.

Table A5: Soviet Space Intelligence Similarity and NASA patents

Dependent Variable=	NASA Patent			
	(1)	(2)	(3)	(4)
Space Capability Score	0.181	0.175	0.091	0.062
	(0.004)	(0.004)	(0.004)	(0.004)
Army Patent			-0.080	-0.091
			(0.001)	(0.001)
Navy Patent			-0.080	-0.077
			(0.001)	(0.001)
Government Patent			0.078	0.077
			(0.000)	(0.000)
Year Fixed Effect	Y	Y	Y	Y
NBER Technology Subcategory Fixed Effects		Y	Y	Y
County Fixed Effects				Y
R^2	0.003	0.005	0.079	0.125
Observations	897,732	897,732	897,732	897,732

Notes: Each column in the table reports the results from estimating one version of equation (A2.2.1) in the appendix. The unit of observation is patent level. The space score variable measures the cosign similarity between the CIA national intelligence estimates of Soviet space capabilities texts between 1958 and 1992 and the text of the reference patent using the Science Direct technology terms corpus, as described in the text and data appendix. The model in column (1) includes year fixed effects, the model in columns (2) also includes NBER technology subcategory fixed effects, the model in column (3) further includes indicator variables for whether the Army, Navy, or other government agency was the owner or funder of the patent. Finally, column (4) further adds county fixed effects.

TABLE A6: Correlation Between NASA and Military Measures

	NASA Spending	Military Spending	NASA Patents	Army Patents	Navy Patents	Gov. Patents
NASA Spending	1					
Military Spending	0.0230	1				
NASA Patents	0.5109	0.0375	1			
Army Patents	0.0511	-0.0039	0.1630	1		
Navy Patents	0.1947	0.0136	0.4556	0.3365	1	
Gov. Patents	0.2990	-0.0071	0.6252	0.3765	0.6394	1

Notes: Definitions and sources for the various spending and patent measures are in the data appendix. The data are reported at the industry-county level, except for Military Spending which is a county-level measure that is applied to all observations within a specific county.

TABLE A7: Space Capability and Manufacturing – Dynamic Effects

Dependent Variable =	Log(Value Add$_{ijt}$)	Log(Employ$_{ijt}$)	Log(Capital$_{ijt}$)	Log(TFP$_{ijt}$)
	(1)	(2)	(3)	(4)
High Space Capability$_{ij<1958}$ × Year=1947$_t$ × Space Ind$_j$	-0.02	0.15	-0.22	-0.09
	(0.11)	(0.11)	(0.30)	(0.08)
High Space Capability$_{ij<1958}$ × Year=1954$_t$ × Space Ind$_j$	0.04	0.08	-0.29	0.04
	(0.16)	(0.16)	(0.27)	(0.06)
High Space Capability$_{ij<1958}$ × Year=1958$_t$ × Space Ind$_j$	0	0	0	0
High Space Capability$_{ij<1958}$ × Year=1963$_t$ × Space Ind$_j$	0.46	0.51	0.10	0.04
	(0.09)	(0.10)	(0.18)	(0.04)
High Space Capability$_{ij<1958}$ × Year=1967$_t$ × Space Ind$_j$	0.45	0.57	0.22	-0.04
	(0.07)	(0.08)	(0.17)	(0.05)
High Space Capability$_{ij<1958}$ × Year=1972$_t$ × Space Ind$_j$	0.21	0.35	0.21	-0.10
	(0.10)	(0.09)	(0.15)	(0.03)
High Space Capability$_{ij<1958}$ × Year=1977$_t$ × Space Ind$_j$	0.45	0.45	0.30	0.04
	(0.13)	(0.10)	(0.16)	(0.05)
High Space Capability$_{ij<1958}$ × Year=1982$_t$ × Space Ind$_j$	0.63	0.65	0.66	-0.02
	(0.17)	(0.18)	(0.17)	(0.03)
High Space Capability$_{ij<1958}$ × Year=1987$_t$ × Space Ind$_j$	0.20	0.24	0.39	-0.07
	(0.09)	(0.11)	(0.16)	(0.04)
High Space Capability$_{ij<1958}$ × Year=1992$_t$ × Space Ind$_j$	0.16	0.17	0.37	-0.06
	(0.15)	(0.13)	(0.23)	(0.05)
County Fixed Effects	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y
Pre-1958 Patents × Year Fixed Effects	Y	Y	Y	Y
Industry Fixed Effects	Y	Y	Y	Y
Industry × Year Fixed Effects	Y	Y	Y	Y
R^2	0.68	0.56	0.46	0.86
Observations	6,759	6,759	6,759	6,759

Notes: Data are drawn from CIA national intelligence estimate, Census of Manufactures, and United States Patent and Trademark data from 1947 to 1992, as described in the data appendix. The unit of observation is 2-digit SIC industry × county × year, indexed by *j*, *i*, *t*, respectively. Each column in the table reports the results from estimating one version of equation (2) in the text. Log(TFP) is defined as log(A_{ijt}) from estimating the production function $Y_{ijt} = A_{ijt}K_{ijt}^{\alpha}L_{ijt}^{\beta}$ by OLS. High Space Capability$_{ij<1958}$ is an indicator variable reflecting a county-industry's being above median in terms of the similarity between the technologies present in its pre-1958 patents and the national intelligence estimates of Soviet space capabilities between 1958 and 1992, as described in the text and appendix.

Space Race years are 1963, 1967 and 1972. Post-Space Race years are 1977, 1982, 1987, and 1992. Space Ind_j is an indicator variable for the industry *j* being a space industry. The models in all columns include county, industry, and year fixed effects, the count of pre-1958 patents in a county × year fixed effects, and industry × year fixed effects. Main entries report coefficient estimates. Standard errors are two-way clustered at the MSA × Industry level and are reported in parentheses.

TABLE A8: Space Capability and Manufacturing – Saturated Models

Dependent Variable =	Log(Value Add$_{ijt}$)		Log(Employ$_{ijt}$)		Log(Capital$_{ijt}$)		Log(TFP$_{ijt}$)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
High Space Capability$_{ij<1958}$ × Space Race$_t$ × Space Ind$_j$	0.42	0.00	0.49	-0.02	0.20	0.80	-0.01	-0.16
	(0.18)	(0.25)	(0.18)	(0.27)	(0.20)	(0.26)	(0.08)	(0.05)
High Space Capability$_{ij<1958}$ × Post- Space Race$_t$ × Space Ind$_j$	0.24	-0.13	0.32	-0.14	0.45	1.03	-0.10	-0.24
	(0.13)	(0.07)	(0.11)	(0.10)	(0.15)	(0.25)	(0.05)	(0.07)
County Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Industry Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Pre-1958 Patents$_{ij}$ × Year Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Industry × Year Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
County × Year Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Industry × County Fixed Effects		Y		Y		Y		Y
R^2	0.72	0.95	0.60	0.93	0.48	0.80	0.88	0.92
Observations	6,715	6,715	6,715	6,715	6,715	6,715	6,715	6,715

Notes: Data are drawn from CIA national intelligence estimate, Census of Manufactures, and United States Patent and Trademark data from 1947 to 1992, as described in the data appendix. The unit of observation is 2-digit SIC industry × county × year, indexed by *j*, *i*, and *t*, respectively. Each column in the table reports the results from estimating one version of equation (1) in the text. Log(TFP) is defined as log(A_{ijt}) from estimating the production function $Y_{ijt} = A_{ijt}K_{ijt}^{\alpha}L_{ijt}^{\beta}$ by OLS. High Space Capability$_{ij<1958}$ is an indicator variable reflecting a county-industry's being above median in terms of the similarity between the technologies present in its pre-1958 patents and the national intelligence estimates of Soviet space capabilities between 1958 and 1992, as described in the text and appendix. Space Race years are 1963, 1967 and 1972. Post-Space Race years are 1977, 1982, 1987, and 1992. Space Ind$_j$ is an indicator variable for the industry *j* being a space industry. The models in all columns includes county, industry, and year fixed effects, and the count of pre-1958 patents in a county × year fixed effects. The models in column (2), (4), (6) and (8) also include MSA × year fixed effects. Main entries report coefficient estimates. Standard errors are two-way clustered at the MSA × industry level and are reported in parentheses.

TABLE A9: Space Capability and Manufacturing– Broad Space Industry Categorization

Dependent Variable =	Log(Value Add$_{ijt}$)		Log(Employ$_{ijt}$)		Log(Capital$_{ijt}$)		Log(TFP$_{ijt}$)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
High Space Capability$_{ij<1958}$ × Space Race$_t$ × Space Ind$_j$	0.23	0.24	0.27	0.28	0.14	0.12	0.00	0.00
	(0.09)	(0.10)	(0.10)	(0.11)	(0.13)	(0.14)	(0.04)	(0.05)
High Space Capability$_{ij<1958}$ × Post- Space Race$_t$ × Space Ind$_j$	0.49	0.47	0.46	0.43	0.48	0.46	0.03	0.03
	(0.07)	(0.07)	(0.06)	(0.07)	(0.12)	(0.13)	(0.06)	(0.06)
County Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Industry Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Pre-1958 Patents$_{ij}$ × Year Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
MSA$_i$ × Year Fixed Effects		Y		Y		Y		Y
R^2	0.67	0.68	0.53	0.56	0.40	0.46	0.85	0.86
Observations	6,759	6,759	6,759	6,759	6,759	6,759	6,759	6,759

Notes: Data are drawn from CIA national intelligence estimate, Census of Manufactures, and United States Patent and Trademark data from 1947 to 1992, as described in the data appendix. The unit of observation is 2-digit SIC industry × county × year, indexed by *j*, *i*, and *t*, respectively. Each column in the table reports the results from estimating one version of equation (1) in the text. Log(TFP) is defined as log(A_{ijt}) from estimating the production function $Y_{ijt} = A_{ijt}K_{ijt}^{\alpha}L_{ijt}^{\beta}$ by OLS. High Space Capability$_{ij<1958}$ is an indicator variable reflecting a county-industry's being above median in terms of the similarity between the technologies present in its pre-1958 patents and the national intelligence estimates of Soviet space capabilities between 1958 and 1992, as described in the text and appendix. Space Race years are 1963, 1967 and 1972. Post-Space Race years are 1977, 1982, 1987, and 1992. Space Ind$_j$ is an indicator variable for the industry *j* being a space industry. The models in all columns includes county, industry, and year fixed effects, and the count of pre-1958 patents in a county × year fixed effects. The models in column (2), (4), (6) and (8) also include MSA × year fixed effects. Main entries report coefficient estimates. Standard errors are two-way clustered at the MSA × industry level and are reported in parentheses.

TABLE A10: Space Capability and Manufacturing – Alternative Inference Procedures

Dependent Variable =	Log(Value Add$_{ijt}$)		Log(Employ$_{ijt}$)		Log(Capital$_{ijt}$)		Log(TFP$_{ijt}$)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
High Space Capability$_{ij<1958}$ × Space Race$_t$ × Space Ind$_j$	0.35	0.38	0.42	0.45	0.31	0.25	-0.04	-0.03
Cluster = County	(0.12)	(0.13)	(0.11)	(0.12)	(0.26)	(0.21)	(0.06)	(0.05)
Cluster = County × Industry	(0.08)	(0.09)	(0.09)	(0.10)	(0.13)	(0.15)	(0.03)	(0.03)
Cluster = State	(0.09)	(0.10)	(0.08)	(0.09)	(0.24)	(0.19)	(0.06)	(0.06)
Cluster = State × Industry	(0.06)	(0.06)	(0.08)	(0.08)	(0.08)	(0.15)	(0.03)	(0.04)
Spatial = HAC-100	(0.11)	(0.11)	(0.10)	(0.10)	(0.21)	(0.21)	(0.06)	(0.06)
High Space Capability$_{ij<1958}$ × Post-Space Race$_t$ × Space Ind$_j$	0.38	0.36	0.36	0.34	0.56	0.50	-0.02	-0.02
Cluster = County	(0.13)	(0.13)	(0.12)	(0.12)	(0.29)	(0.23)	(0.06)	(0.05)
Cluster = County × Industry	(0.08)	(0.11)	(0.10)	(0.13)	(0.17)	(0.15)	(0.03)	(0.03)
Cluster = State	(0.12)	(0.13)	(0.11)	(0.11)	(0.27)	(0.20)	(0.07)	(0.06)
Cluster = State × Industry	(0.07)	(0.11)	(0.09)	(0.13)	(0.11)	(0.13)	(0.03)	(0.03)
Spatial = HAC-100	(0.13)	(0.13)	(0.12)	(0.12)	(0.22)	(0.22)	(0.05)	(0.05)
County Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Pre-1958 Patents × Year Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Industry Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Industry × Year Fixed Effects		Y		Y		Y		Y
R^2	0.66	0.68	0.53	0.55	0.40	0.46	0.85	0.86
Observations	6,759	6,759	6,759	6,759	6,759	6,759	6,759	6,759

Notes: This table is analogous to Table 3 in the main paper with alternative inference procedures applied. See Table 3 in the main paper for additional information on notes and sources. The differences here are that the first parentheses row clusters standard error by county, the second parentheses row clusters standard errors two-way by county × industry, the third parentheses row clusters the standard errors by state, the fourth parentheses row clusters standard errors two-way by state × industry, and the last parentheses row computes standard errors using a spatial HAC procedure with a 100km cutoff.

TABLE A11: Space Capability and Manufacturing – Heterogeneous Effects and Correlated Share Robustness

Dependent Variable =	Log(Value Add$_{ijt}$)		Log(Employ$_{ijt}$)		Log(Capital$_{ijt}$)		Log(TFP$_{ijt}$)	
Method:	HR	SS	HR	SS	HR	SS	HR	SS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
High Space Capability$_{ij<1958}$ × After Sputnik$_t$ × Space Ind$_j$	0.809	0.441	0.822	0.461	0.098	0.305	0.155	0.019
	(0.373)	(0.056)	(0.417)	(0.053)	(0.972)	(0.061)	(0.292)	(0.015)
County Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Observations	3,158	6,759	3,158	6,759	3,158	6,759	3,158	6,759

Notes: This table is analogous to Table 3 in the main paper, where the reader can find notes and sources. The differences here are that After Sputnik takes a value of one in years 1963, 1967, 1972, 1977, 1982, 1987, and 1992, and zero before. The estimates in columns (1), (3), (5), (7) are computed using the heterogeneous effect robustness methods in de Chaisemartin and D'Haultfœuille (2020). The estimates in columns (2), (4), (5), (6) compute standard errors allowing for correlation across units depending on the sector shares utilizing methods in Adão, Kolesár, and Morales (2019). The unit of observation is 2-digit SIC industry × county × year, indexed by *j*, *i*, *t*, respectively. The models in all columns include county fixed effects and year fixed effects. Standard errors are clustered at the MSA level.

TABLE A12: Space Capability and Manufacturing Productivity Growth

Dependent Variable =	$\frac{\text{Log}(\text{TFP}_{ijt})-\text{Log}(\text{TFP}_{ijt-1})}{\text{Log}(\text{TFP}_{ijt-1})}$	
	(1)	(2)
High Space Capability$_{ij<1958}$ × Space Race$_t$ × Space Ind$_j$	-7.54	-8.68
	(10.11)	(10.52)
High Space Capability$_{ij<1958}$ × Post-Space Race$_t$ × Space Ind$_j$	-7.65	-4.94
	(6.84)	(4.91)
County Fixed Effects	Y	Y
Year Fixed Effects	Y	Y
Pre-1958 Patents × Year Fixed Effects	Y	Y
Industry Fixed Effects	Y	Y
Industry × Year Fixed Effects		Y
R^2	0.03	0.24
Observations	6,024	6,024

Notes: Data are drawn from CIA national intelligence estimate, Census of Manufactures, and United States Patent and Trademark data from 1947 to 1992, as described in the data appendix. The unit of observation is 2-digit SIC industry × county × year, indexed by *j*, *i*, *t*, respectively. Each column in the table reports the results from estimating one version of equation (2) in the text. Log(TFP) is defined as log(A_{ijt}) from estimating the production function $Y_{ijt} = A_{ijt}K_{ijt}^{\alpha}L_{ijt}^{\beta}$ by OLS. High Space Capability$_{ij<1958}$ is an indicator variable reflecting a county-industry's being above median in terms of the similarity between the technologies present in its pre-1958 patents and the national intelligence estimates of Soviet space capabilities between 1958 and 1992, as described in the text and appendix. Space Race years are 1963, 1967 and 1972. Post-Space Race years are 1977, 1982, 1987, and 1992. Space Ind$_j$ is an indicator variable for the industry *j* being a space industry. The models in all columns include county, industry, and year fixed effects, and the count of pre-1958 patents in a county × year fixed effects. The model in column (2) also includes industry × year fixed effects. Main entries report coefficient estimates. Standard errors are two-way clustered at the MSA × Industry level and are reported in parentheses.

TABLE A13: NASA Spending and Manufacturing – IV Models

	(1)	(2)	(3)	(4)
	First Stage: Dependent Variable = Any NASA Spending$_{ijt}$			
High Space Capability$_{ij<1958}$ × After Sputnik$_t$ × Space Ind$_j$	0.07	0.07	0.07	0.07
	(0.02)	(0.02)	(0.02)	(0.02)
F-Statistic:	22.63	22.63	22.63	22.63
	[0.0002]	[0.0002]	[0.0002]	[0.0002]
Second Stage: Dependent Variable=	Log(Value Add$_{ijt}$)	Log(Employ$_{ijt}$)	Log(Capital$_{ijt}$)	Log(TFP$_{ijt}$)
Any NASA Spending$_{ijt}$	4.96	5.24	6.11	-0.43
	(1.84)	(2.01)	(2.52)	(0.39)
County Fixed Effects	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y
Pre-1958 Patents × Year Fixed Effects	Y	Y	Y	Y
Industry Fixed Effects	Y	Y	Y	Y
Observations	6,759	6,759	6,759	6,759

Notes: Data are drawn from CIA national intelligence estimate, Census of Manufactures, and United States Patent and Trademark data from 1947 to 1992, as described in the data appendix. The unit of observation is 2-digit SIC industry × county × year, indexed by *j*, *i*, *t*, respectively. Each column in the table reports the results from estimating a modified version of equation (1) in the text where Space Race and Post Space Race interactions are replaced with a single interaction After Sputnik. Log(TFP$_{ijt}$) is defined as log(A$_{ijt}$) from estimating the production function $Y_{ijt} = A_{ijt}K_{ijt}^{\alpha}L_{ijt}^{\beta}$ by OLS. High Space Capability$_{ij<1958}$ is an indicator variable reflecting a county-industry's being above median in terms of the similarity between the technologies present in its pre-1958 patents and the mational intelligence estimates of Soviet space capabilities between 1958 and 1992, as described in the text and appendix. After Sputnik year are 1963, 1967,1972, 1977, 1982, 1987, and 1992. Space Ind$_j$ is an indicator variable for the industry *j* being a space industry. The models in all columns include county, industry, and year fixed effects, and the count of pre-1958 patents in a county × year fixed effects. Panel A reports the first-stage estimate where Any NASA Spending$_{ijt}$ is the left hand side variable, which is an indicator variable taking the value of 1 if the industry-county combination received any NASA funding in year *t*. Panel B reports the instrumental variables estimate where High Space Capability$_{ij<1958}$ × After Sputnik$_t$ × Space Ind$_j$ is the excluded instrument. Main entries report coefficient estimates. Standard errors are two-way clustered at the MSA × Industry level and are reported in parentheses.

TABLE A14: Space Industries and Manufacturing – Industry Level

Dependent Variable =	Log(Value Add$_{jt}$)	Log(Employ$_{jt}$)	Log(Capital$_{jt}$)	Log(TFP$_{jt}$)
	(3)	(4)	(5)	(6)
Space Ind$_j$ × Space Race$_t$	0.33	0.25	0.13	0.04
	(0.09)	(0.14)	(0.03)	(0.08)
Space Ind$_j$ × Post-Space Race$_t$	0.41	0.31	0.42	0.05
	(0.19)	(0.19)	(0.14)	(0.10)
Year Fixed Effects	Y	Y	Y	Y
Industry Fixed Effects	Y	Y	Y	Y
R^2	0.93	0.93	0.97	0.97
Observations	190	190	190	190

Notes: Data are drawn from Manufacturing Census data from 1947 to 1992, as described in the data appendix. The unit of observation is 2-digit SIC industry × year, indexed by *j* and *t*, respectively. Log(TFP$_{jt}$) is defined as log(A$_{ijt}$) from estimating the production function $Y_{ijt} = A_{ijt}K_{ijt}^{\alpha}L_{ijt}^{\beta}$ by OLS, then taking the mean of this measure by industry to compute industry-level productivity. Each column in the table reports the results from estimating a modified version of equation (1) in the text without county fixed effects or 1958 patent-count controls included. Space Race years are 1963, 1967 and 1972. Post-Space Race years are 1977, 1982, 1987, and 1992. All models include industry and year fixed effects. Main entries report coefficient estimates. Standard errors are clustered at the industry level and are reported in parentheses.

TABLE A15: Space Capability and Manufacturing – Non-Space Industries

Dependent Variable =	Log(Value Add$_{ijt}$)		Log(Employ$_{ijt}$)		Log(Capital$_{ijt}$)		Log(TFP$_{ijt}$)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
High Space Capability$_{ij<1958}$ × Space Race$_t$	-0.06	-0.06	-0.07	-0.07	-0.05	-0.05	0.00	0.00
	(0.07)	(0.07)	(0.06)	(0.06)	(0.09)	(0.09)	(0.02)	(0.02)
High Space Capability$_{ij<1958}$ × Post-Space Race$_t$	0.04	0.04	0.06	0.06	-0.01	-0.01	-0.01	-0.01
	(0.07)	(0.08)	(0.06)	(0.07)	(0.08)	(0.08)	(0.02)	(0.02)
County - High Space Capability$_{ij<1958}$ × Space Race$_t$		-0.14		-0.09		-1.09		0.18
		(0.43)		(0.33)		(1.38)		(0.28)
County - High Space Capability$_{ij<1958}$ × Post-Space Race$_t$		0.86		1.02		-0.81		0.26
		(0.62)		(0.57)		(1.45)		(0.28)
County Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Pre-1958 Patents × Year Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Industry Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Industries Included in Sample	Non-Space	Non-Space	Non-Space	Non-Space	Non-Space	Non-Space	Non-Space	Non-Space
R^2	0.66	0.66	0.52	0.52	0.40	0.40	0.85	0.85
Observations	6,077	6,077	6,077	6,077	6,077	6,077	6,077	6,077

Notes: This table is analogous to Table 3 in the main paper. See Table 3 for sources and definition of Log(TFP$_{ijt}$). The version here differs from Table 3 in that only data from the non-space industries are used. The unit of observation is 2-digit SIC industry × county × year, indexed by *j*, *i*, *t*, respectively. High Space Capability$_{ij<1958}$ is an indicator variable reflecting a county-industry's being above median in terms of the similarity between the technologies present in its pre-1958 patents and the national intelligence estimates of Soviet space capabilities between 1958 and 1992, as described in the text and appendix. Space Race years are 1963, 1967 and 1972. Post-Space Race years are 1977, 1982, 1987, and 1992. County - High Space Capability$_{ij<1958}$ is the average of the space score in county *i* amongst the space sector firms. The models in all columns include county, industry, and year fixed effects, and the count of pre-1958 patents in a county × year fixed effects. Main entries report coefficient estimates. Standard errors are two-way clustered at the MSA × industry level and are reported in parentheses.

TABLE A16: Space Capability and Manufacturing – US Space Technology Similarity Measures

Dependent Variable =	Log(Value Add_{ijt})		Log($Employ_{ijt}$)		Log($Capital_{ijt}$)		Log(TFP_{ijt})	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
US High Space $Capability_{ij<1958}$ × $Space\ Race_t$ × $Space\ Ind_j$	0.34	0.36	0.40	0.43	0.47	0.39	-0.08	-0.05
	(0.09)	(0.10)	(0.09)	(0.10)	(0.09)	(0.13)	(0.03)	(0.04)
US High Space $Capability_{ij,<1958}$ × Post-$Space\ Race_t$ × $Space\ Ind_j$	0.40	0.39	0.36	0.36	0.71	0.60	-0.04	-0.02
	(0.11)	(0.13)	(0.11)	(0.13)	(0.08)	(0.15)	(0.02)	(0.04)
County Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Pre-1958 Patents × Year Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Industry Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
Industry × Year Fixed Effects	N	Y	N	Y	N	Y	N	Y
R^2	0.66	0.68	0.53	0.55	0.40	0.46	0.85	0.86
Observations	6,747	6,747	6,747	6,747	6,747	6,747	6,747	6,747

Notes: This table is analogous to Table 3 in the main paper. This version differs in that the High Space Capability measure uses similarity to US space technology to create the High Space Capability measure. US High $Space\ Capability_{ij<1958}$ is an indicator variable reflecting a county-industry's being above median in terms of the similarity between the technologies present in its pre-1958 patents and NASA patents between 1958 and 1992, as described in the text and appendix. Otherwise, notes and sources in Table 3 remain the same.

TABLE A17: Space Capability and Manufacturing – Alternative Text Processing Methods

Dependent Variable =	Log(Value Add$_{ijt}$)	Log(Employ$_{ijt}$)	Log(Capital$_{ijt}$)	Log(TFP$_{ijt}$)
	(1)	(2)	(3)	(4)
Panel A: Mean-Based Space Score				
Mean Space Score$_{ij<1958}$ × Space Race$_t$ × Space Ind$_j$	43.28	46.84	41.52	-1.98
	(5.02)	(8.72)	(16.64)	(5.13)
Mean Space Score$_{ij<1958}$ × Post-Space Race$_t$ × Space Ind$_j$	51.38	44.95	70.03	1.24
	(9.45)	(12.11)	(22.35)	(5.38)
Panel B: Unstemmed Text Processing				
High Space Capability$_{ij<1958}$ × Space Race$_t$ × Space Ind$_j$	0.36	0.38	0.22	0.01
	(0.10)	(0.10)	(0.23)	(0.03)
High Space Capability$_{ij<1958}$ × Post-Space Race$_t$ × Space Ind$_j$	0.42	0.41	0.46	0.00
	(0.16)	(0.14)	(0.27)	(0.02)
Panel C: All Documents				
High Space Capability$_{ij<1958}$ × Space Race$_t$ × Space Ind$_j$	0.30	0.36	0.27	-0.04
	(0.05)	(0.07)	(0.13)	(0.03)
High Space Capability$_{ij<1958}$ × Post-Space Race$_t$ × Space Ind$_j$	0.33	0.30	0.50	-0.02
	(0.08)	(0.11)	(0.17)	(0.03)
Panel D: Space Exclusive Documents Only				
High Space Capability$_{ij<1958}$ × Space Race$_t$ × Space Ind$_j$	0.31	0.36	0.36	-0.04
	(0.03)	(0.06)	(0.09)	(0.03)
High Space Capability$_{ij<1958}$ × Post-Space Race$_t$ × Space Ind$_j$	0.40	0.35	0.62	0.00
	(0.05)	(0.08)	(0.09)	(0.03)
Panel E: Above Median – County Level				
High Space Capability$_{j<1958}$ × Space Race$_t$ × Space Ind$_j$	0.23	0.28	0.28	-0.04
	(0.09)	(0.08)	(0.13)	(0.03)
High Space Capability$_{j<1958}$ × Post-Space Race$_t$ × Space Ind$_j$	0.33	0.35	0.57	-0.06
	(0.08)	(0.06)	(0.18)	(0.03)
Panel F: Above 75th Percentile – County Level				
High Space Capability$_{j<1958}$ × Space Race$_t$ × Space Ind$_j$	0.18	0.24	0.22	-0.05
	(0.12)	(0.10)	(0.10)	(0.04)
High Space Capability$_{j<1958}$ × Post-Space Race$_t$ × Space Ind$_j$	0.29	0.31	0.51	-0.06
	(0.09)	(0.09)	(0.16)	(0.04)
County Fixed Effects	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y
Pre-1958 Patents × Year Fixed Effects	Y	Y	Y	Y
Industry Fixed Effects	Y	Y	Y	Y
Observations	6,759	6,759	6,759	6,759

Notes: This table is analogous to Table 3 in the main paper but with alternative text processing procedures applied to create the High Space Capability measures. Panel A replaces the above-median space score with the mean of the space score in the county-industry cell. Panel B computes the above-median High Space Capability score using scientific terms without applying stemming. Panel C computes the above-median High Space Capability score using all national intelligence estimates of Soviet space capabilities documents. Panel D computes the above-median High Space Capability score using only national intelligence estimates of Soviet space capabilities documents that are space exclusive. Panel E computes the above-median High Space Capability score at the county instead of the county-industry level. Panel F computes the above-high High Space Capability score as being above the 75th percentile in the space score distribution instead of the median, at the county level. All other details follow those in Table 3 of the main paper.

TABLE A18: High Space Capability and Manufacturing: Text Based Military Patent Controls

Dependent Variable =	Log(Value Add$_{ijt}$)		Log(Employ$_{ijt}$)		Log(Capital$_{ijt}$)		Log(TFP$_{ijt}$)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
High Space Capability$_{ij<1958}$ × Space Race$_t$ × Space Ind$_j$	0.34	0.34	0.41	0.41	0.31	0.33	-0.04	-0.04
	(0.07)	(0.06)	(0.09)	(0.08)	(0.13)	(0.17)	(0.03)	(0.04)
High Space Capability$_{ij<1958}$ × Space Race$_t$ × Space Ind$_j$	0.37	0.36	0.35	0.34	0.56	0.55	-0.02	-0.02
	(0.09)	(0.09)	(0.10)	(0.10)	(0.17)	(0.20)	(0.03)	(0.04)
Additional Military Patent Controls:								
Military Text Patents	Y	Y	Y	Y	Y	Y	Y	Y
Pre-1958 Patent Military Similarity × Year Fixed Effects		Y		Y		Y		Y
R^2	0.66	0.66	0.53	0.53	0.40	0.40	0.85	0.85
Observations	6,759	6,741	6,759	6,741	6,759	6,741	6,759	6,741

Notes: Data are drawn from CIA national intelligence estimate, Census of Manufactures, and United States Patent and Trademark data from 1947 to 1992, as described in the data appendix. The unit of observation is 2-digit SIC industry × county × year, indexed by *j*, *i*, *t*, respectively. Each column in the table reports the results from estimating one version of equation (1) in the text. Log(TFP) is defined as log(A_{ijt}) from estimating the production function $Y_{ijt} = A_{ijt}K_{ijt}^{\alpha}L_{ijt}^{\beta}$ by OLS. High Space Capability$_{ij<1958}$ is an indicator variable reflecting a county-industry's being above median in terms of the similarity between the technologies present in its pre-1958 patents and the national intelligence estimates of Soviet space capabilities between 1958 and 1992, as described in the text and appendix. Space Race years are 1963, 1967 and 1972. Post-Space Race years are 1977, 1982, 1987, and 1992. Space Ind$_j$ is an indicator variable for the industry *j* being a space industry. The models in all columns include county, industry, and year fixed effects, and the count of pre-1958 patents in a county × year fixed effects. The models in column (1), (3), (5) and (7) also include counts of text-based military patents as controls. The models in column (2), (4), (6) and (8) also include pre-1958 similarity of patents to military technology × year fixed effects. Main entries report coefficient estimates. Standard errors are two-way clustered at the MSA × industry level and are reported in parentheses.

Exhibit A1: Science Direct Technology Terms and Intelligence-Patents Similarity Examples

A. Intelligence Document (NIE 11-1-62, p27)

TABLE 3

OTHER POSSIBLE SOVIET SPACE MISSIONS

We have estimated that the chances are better than even that the Soviets will attempt a manned lunar landing ahead of or in close competition with the US. In addition, the Soviets will probably undertake other programs including scientific satellites, military support satellites, and interplanetary probes. This table lists space missions estimated to be within Soviet capabilities, but we do not believe that all these missions could be accomplished within the time periods indicated. If the Soviets are not committed to a lunar race with the US, other programs will probably receive greater emphasis.

PROGRAMS	POSSIBLE DATE
UNMANNED SATELLITES	
Space Science Operations *	
(1) Magnetic measurement	1962 on
(2) Radiation measurement	
(3) Study of electromagnetic propagation	
(4) Study of upper atmosphere	
(5) Study of meteorites	
(6) Orbital astronomical observatory	
Military Systems	
(1) Early warning satellite	1962 on
(2) Reconnaissance satellite	1962 on
(3) Defensive space weapons systems *	
(i) Inspection of single nonmaneuvering satellite	1964
(ii) More sophisticated satellite with inspection, neutralization, and damage assessment capability	Later in decade
(4) Offensive space weapons systems	
(i) Demonstration of orbital bombardment satellite *	1962 on
(ii) Developmental system of limited effectiveness	1965
Commercial or Military Application	
(1) Meteorological satellites	1962–1963
(2) Communications satellites	
(3) Geodetic satellites	
(4) Navigation satellites	
MAN-IN-SPACE	
Manned Earth Orbital Flights	
(1) Orbit of multimanned spacecraft	1962–1963
(2) Rendezvous and docking	1962–1963
(3) Demonstration of 10-day life support system	1962–1963
(4) Transfer of man from one space vehicle to another	1963–1964
Large Manned Space Station	1965–1966 *
Manned Lunar Flights	
(1) Circumlunar	1965–1966 *
(2) Lunar satellite	1966–1967
UNMANNED LUNAR AND PLANETARY EXPLORATION	
Circumlunar, Lunar Satellite, Lunar Soft Landing	1962 on
Probes to Mars and Venus	1962 on
Probes to More Distant Planets	1963 *
Solar Probe	1963
Ejection of Vehicle from Solar System	1963 *

B. Similar Patent (3907225)

SPACECRAFT FOR DEPLOYING OBJECTS INTO SELECTED FLIGHT PATHS

BACKGROUND OF THE INVENTION

1. Field of the Invention

This invention relates generally to the space field and more particularly to a spacecraft for deploying objects into selected multiple space flight paths.

2. Prior Art

Spacecraft of the class to which this invention pertains are launched into space carrying a number of objects which are deployed at intervals from the spacecraft in a manner such that each object is inserted into a selected space flight path. The flight paths into which the objects are inserted may be either orbital paths, outer space paths, or ballistic trajectories. The invention will be described in the context of launching objects into ballistic trajectories.

According to a typical ballistic trajectory deployment sequence, the spacecraft is launched into an initial flight path and any of the objects which are to follow this path are then deployed from the spacecraft. Insertion of objects into other flight paths is accomplished by propelling the spacecraft along the local range insensitive axis, i.e., an axis passing through the spacecraft and the center of the earth, and across the desired flight paths and deploying the objects along these paths. Such deployment may be accomplished either passively or actively. Passive deployment of an object is accomplished by releasing the object for separation from the spacecraft and backing the spacecraft away from the object. Active deployment involves forcible ejection of the object from the spacecraft by spring action or the like.

Actual separation of the deployable objects from the spacecraft along the seleted flight paths may be only one of the requirements for proper insertion of the objects into these paths. Another deployment requirement may involve proper orientation of each object relative to its path at the time of separation from the spacecraft. Thus, objects which are deployed in the manner described herein generally have an axis, referred to as a reference axis, which must be oriented at a predetermined angle relative to the respective flight paths at the time of deployment.

My prior U.S. Pat. No. **3,652,042** discloses a spacecraft of the general class described. Another patent of interest with regard to such spacecraft is U.S. Pat. No. **3,547,375.**

SUMMARY OF THE INVENTION

This invention provides a novel spacecraft for deploying objects into selected multiple flight paths. The spacecraft comprises two separable, individually powered units, referred to herein as deployment vehicles, with unique interfitting configurations which permit assembly or mating of the vehicles into a compact unitary spacecraft structure. Each deployment vehicle has propulsion, braking and attitude control thrusters, an object mounting platform, fuel tanks, control systems, and the like. A number of objects to be deployed are mounted on the platform of each vehicle by either active or passive deployment means.

The two deployment vehicles are initially assembled in mating or interfitting relation to form a unitary spacecraft which is stowed within the nose section of a launch vehicle for launch of the spacecraft into a pre-

C. Intelligence Document (NIE 11-1-69, p15)

43. There are several likely near term applications of the automatic rendezvous and docking system. When Soyuz 4 and 5 docked, the Soviets claimed that they had fulfilled the rendezvous techniques that would be used to assemble space stations in earth orbit; they would also be used for resupply of such stations and crew rotation. Moreover, the Soviets have indicated that rendezvous and docking would be used to assemble lunar and planetary spaceships in earth orbit.

44. The need for a cooperating target limits the use of the automatic rendezvous and docking technique in many applications such as rescue and in-orbit repair and maintenance. However, the Soyuz automatic rendezvous and docking system equipped with suitable sensors could be used by a maneuverable satellite for passive targets.

45. In the Soyuz and Zond programs the Soviets employed aerodynamic lift re-entry techniques similar to those used in the Gemini and Apollo missions. The Soyuz vehicle is designed to follow a preprogrammed deceleration profile, which is calculated to reduce re-entry "g" loadings during re-entry and to provide some control over the point of landing. The Zond re-entry is more complicated. The spacecraft normally re-enters over the Indian Ocean, and its lifting capability is used to move the landing point some five thousand miles north into the Soviet Union. The lifting re-entry techniques indicate a desire to recover manned lunar spacecraft in the Soviet Union rather than to rely regularly on a water recovery system. The ocean recovery capability is probably provided mainly as a backup.

46. The prime electric power source for the Soyuz and the Zond is solar energy. Solar power, however, is inadequate for some interplanetary missions and orbital missions having high continuous power requirements in excess of 50 kilowatts; nuclear powered systems will be required for such flights. The Soviets are actively engaged in R&D on various nuclear electric power systems. By the mid-1970's they could have a nuclear auxiliary power system generating a few kilowatts.

47. The Soviets have made significant strides in improving the reliability of their planetary spacecraft. They are using an improved thermal control system and also are equipping spacecraft with redundant component subsystems for backup in the event of failure. In their discussions of the Venus 4, the Soviets for the first time mentioned the use of a space simulator to check out the completed spacecraft before launch. These modifications should improve spacecraft and insure better success in the future.

D. Similar Patent (3232560)

3,232,560
RECOVERABLE SPACE VEHICLE
John C. Moise, Carmichael, and John Eut Tileston, Sacramento, Calif., assignors to Aerojet-General Corporation, Azusa, Calif., a corporation of Ohio
Filed Feb. 25, 1963, Ser. No. 260,355
3 Claims. (Cl. 244—1)

This invention relates in general to space vehicles and more particularly to a space tanker or freighter which may be recovered and reused.

It is expected that within the next decade space stations will be in use either for military or for scientific purposes or for both. Such space stations will be set up for purposes of refueling space craft, effecting transfer of personnel from one mission to another and for other such uses. It is seen, then, that a vehicle capable of placing these stations in orbit and keeping them supplied, at a minimum total operating cost per pound of payload in orbit, is needed. It is to this latter requirement that the present invention is primarily directed.

Various means have been proposed for conveying large payloads to low orbit space stations, some including water recoverable vehicles and others winged vehicles such as an aerospace plane. However, in expected cost per pound of payload the present invention is greatly superior to any of the above.

The device of the present invention is fundamentally an aerodynamically configured grouping of liquid fuel and liquid oxidizer storage tanks plus propulsion means sufficient to permit vertical take-off, rendezvous and docking at a space station, payload transfer, retro-thrust and re-entry at a shallow angle followed by a controlled glide landing.

Accordingly, it is an object of the present invention to provide a vehicle capable of carrying a significant payload to orbit at a minimum cost.

It is another object of the present invention to provide a vehicle which may be recovered and reused a considerable number of times.

It is a further object of the present invention to provide a vehicle which permits full recovery of the vehicle through a single recovery technique.

Other objects and many of the attendant advantages of this invention will be readily appreciated as the same becomes better understood by reference to the following detailed description when considered in connection with the accompanying drawings, wherein:

FIG. 1 is a plan view of one embodiment of the invention.

FIG. 2 is a sectional view along line 2—2 of the embodiment of FIG. 1.

FIG. 3 is a side elevation of the embodiment of FIG. 1.

FIG. 4 is a sectional view along the line 4—4 of the embodiment of FIG. 1.

FIG. 5 is an alternate configuration for the section of FIG. 4.

FIG. 6 is an alternate configuration for the section of FIG. 4.

FIG. 7 is an alternate configuration for the section of FIG. 4.

FIG. 8 is an alternate configuration for the section of FIG. 4.

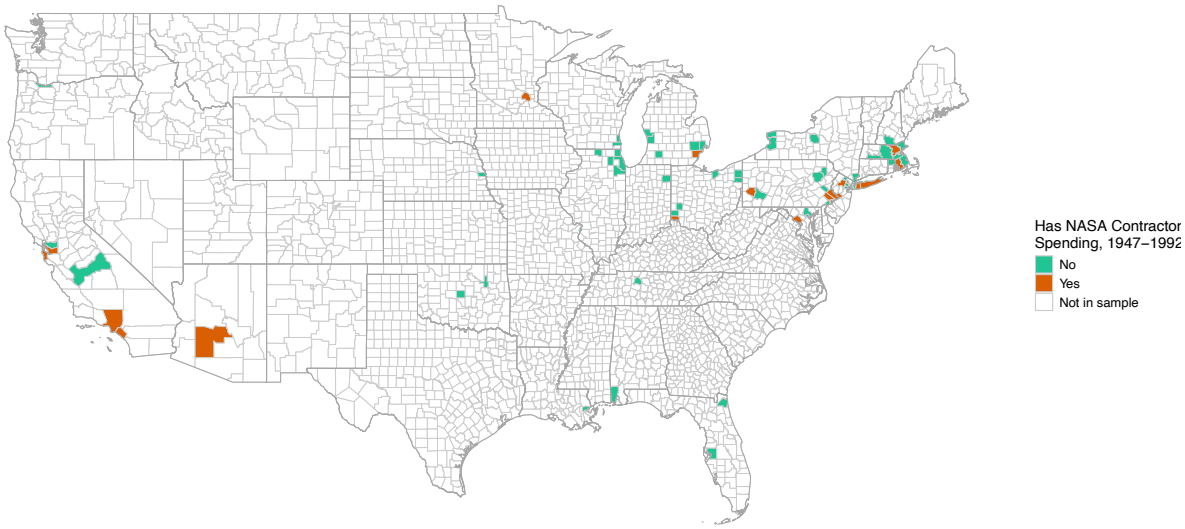
The space freighter or tanker of the present invention is primarily seen to be a series of aerodynamically configured tanks capable of generating sufficient lift during re-entry to enable the vehicle to return safely to the earth's surface, and, further, having means for propulsion and attitude control providing the capability of delivering an exceedingly high percentage of its total weight as payload to a space station.

Referring now to FIG. 1, space vehicle 11 is shown

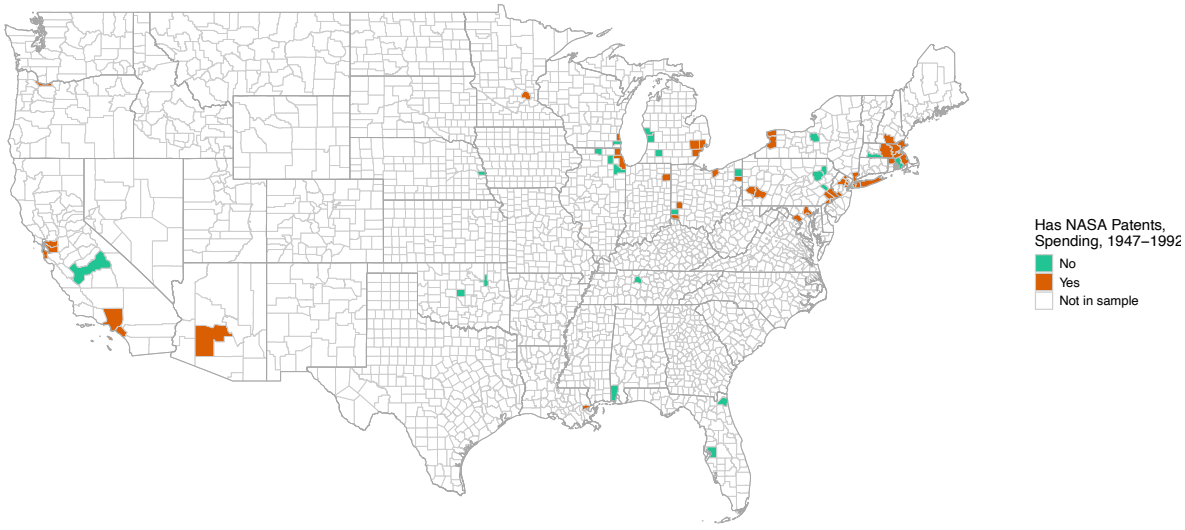
Notes: The highlighted words are Science Direct Technology Terms used to compute document similarity. The NIA document displayed in panel A would be considered similar to the patent in panel B. The NIE document in panel C is similar to the patent in panel D.

Map A1: NASA Locations, 1947-1992

A. NASA Spending



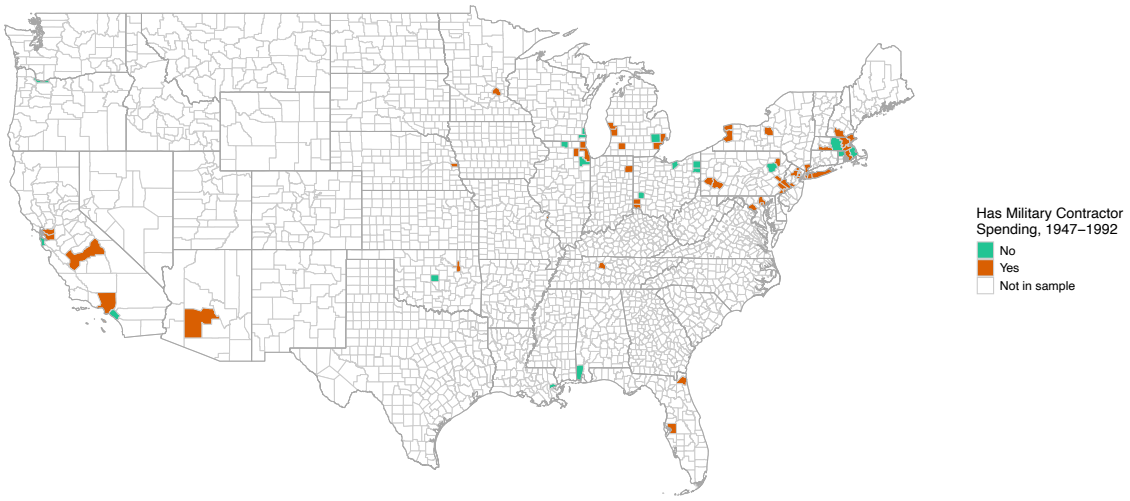
B. NASA Patents



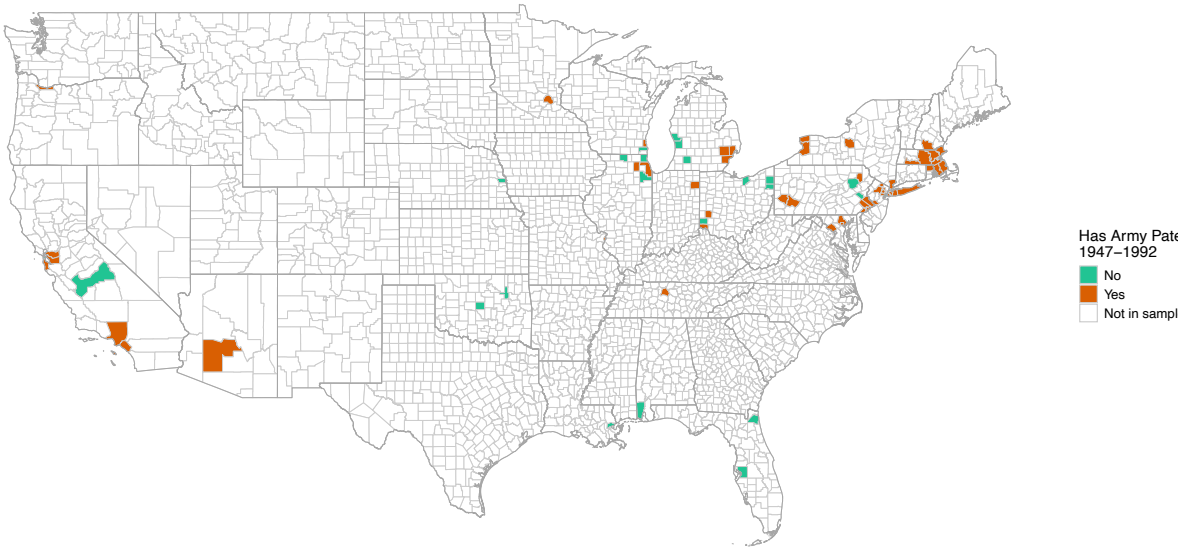
Sources: See the data appendix under NASA spending and NASA patents.

Map A2: Military Locations, 1947-1992

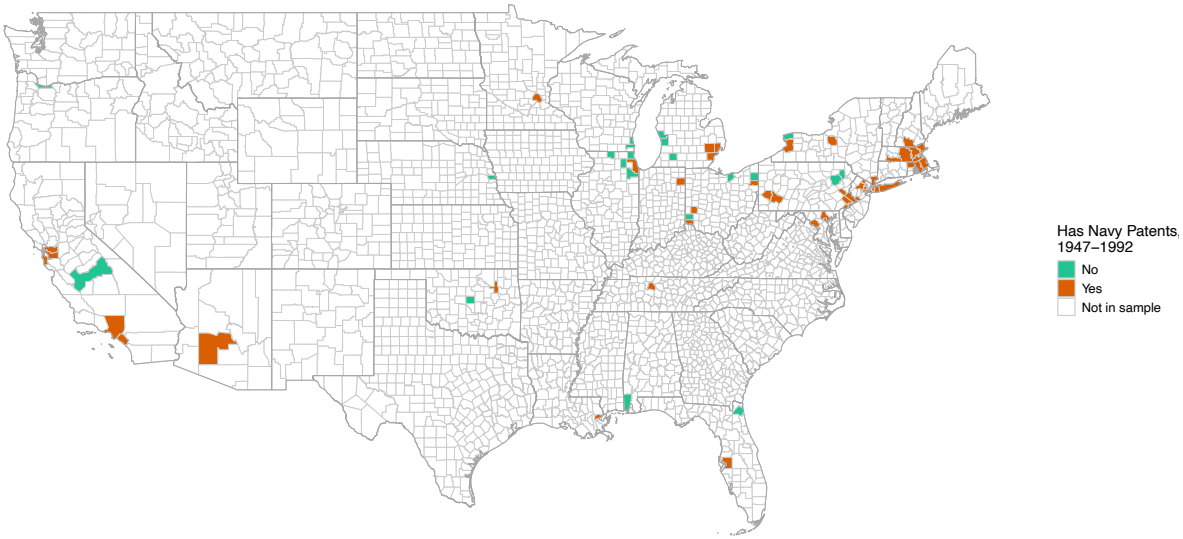
A. Military Contractor Spending



B. Army Patents

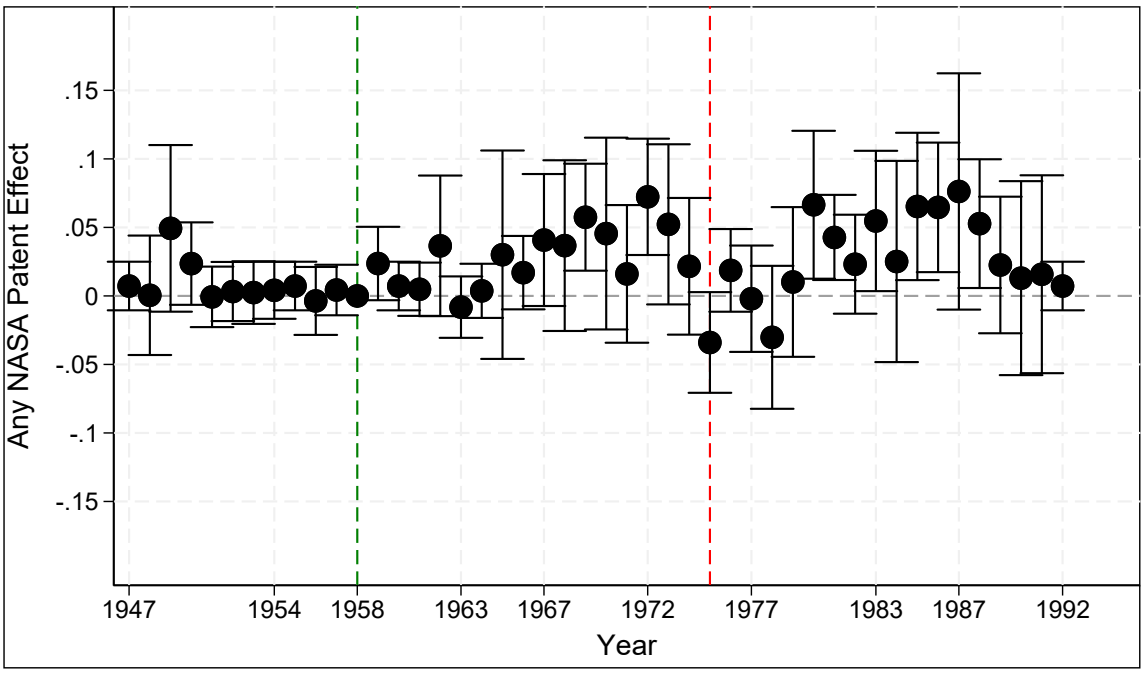


C. Navy Patents



Sources: See the data appendix under military spending and Army and Navy patents.

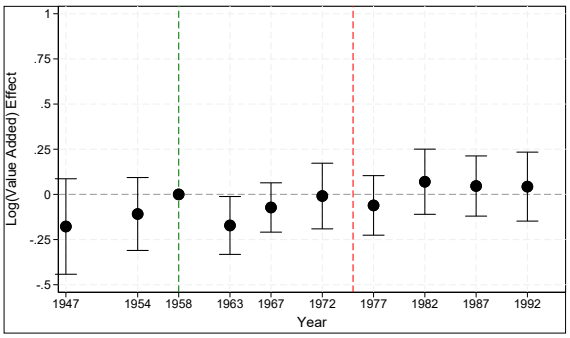
Figure A1: Dynamic Effects – NASA Patent Outcomes



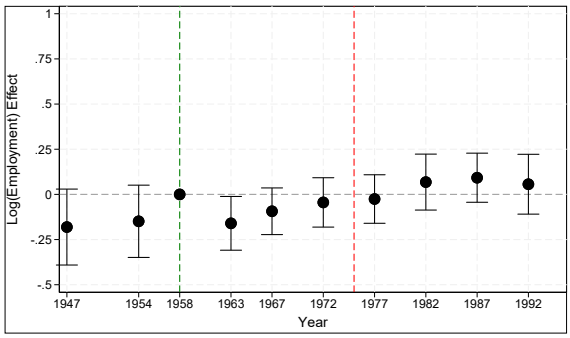
Notes and Sources: Data are drawn from CIA national intelligence estimate and United States Patent and Trademark data from 1947 to 1992, as described in the data appendix. The figure displays the results from estimating a modified version of equation (2) in the text. The outcome variable is a dummy variable taking the value of 1 if there were any patents in an industry-county in year *t*, and 0 otherwise, using annual data from 1958 to 1992 with the 1958 interaction with High Space Capability omitted. The points plot the year by year coefficient estimates on the High Space Capability$_{ij,<1958}$ variable with the 95% confidence intervals indicated by the range. The omitted year is 1958, as indicated in equation (2) in the text. Space Race years are 1963, 1967 and 1972. Post-Space Race years are 1977, 1982, 1987, and 1992. The unit of observation is 2 digit SIC industry × county × year. The model includes county, year fixed effects, the count of pre-1958 patents in a county × year fixed effects, state × year fixed effects, and include industry fixed effects. Standard errors are two-way clustered at the MSA × industry level.

Figure A2: Space Capability and Manufacturing – Main Effect Dynamics

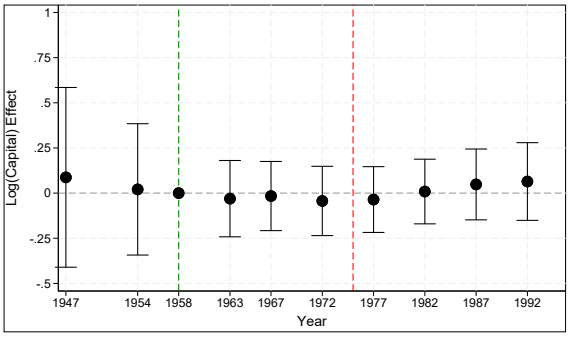
Panel A: Log(Value Added)



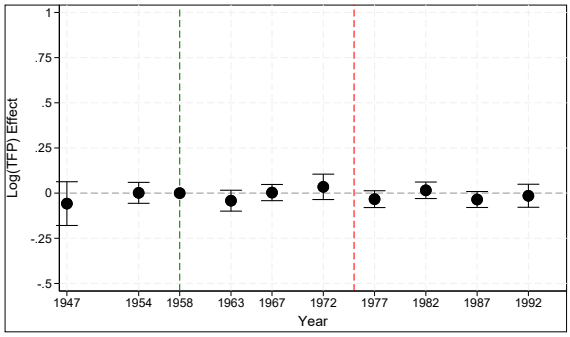
Panel B: Log(Employment)



Panel C: Log(Capital)



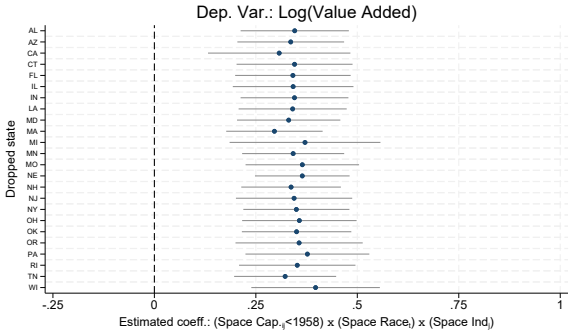
Panel D: Log (TFP)



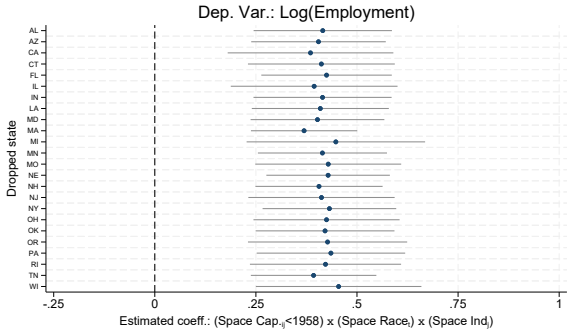
Notes: Data are drawn from CIA national intelligence estimate, Census of Manufactures, and United States Patent and Trademark data from 1947 to 1992, as described in the data appendix. The unit of observation is 2-digit SIC industry × county × year, indexed by *j*, *i*, and *t*, respectively. Each panel in the table displays the results from estimating one version of equation (2) in the text for the outcome indicated, with 1958 serving as the omitted base year. The unit of observation is 2 digit SIC industry × county × year. Log(TFP) is defined as $\log(A_{ijt})$ from estimating the production function $Y_{ijt} = A_{ijt}K_{ijt}^{\alpha}L_{ijt}^{\beta}$ by OLS. The points plot year by year coefficients of High Space Capability$_{ij<1958}$ and Space Race and Post-Space Race era interactions with the 95% confidence intervals indicated by the range. Space race years are 1963, 1967 and 1972. Post-Space Race years are 1977, 1982, 1987, and 1992. The models in all panels includes county, industry, and year fixed effects, and the count of pre-1958 patents in a county × year fixed effects. Standard errors are two-way clustered at the MSA × industry level.

Figure A3: Leave One State Out Estimates – Space Race Era Estimates

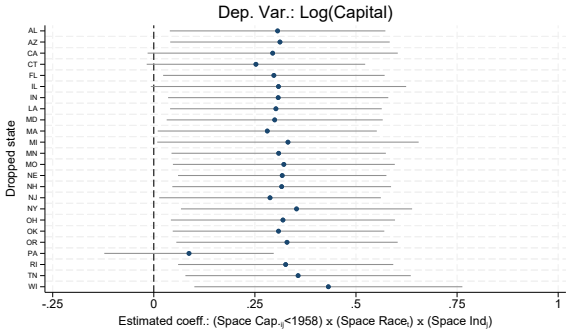
A. Log(Value Added)



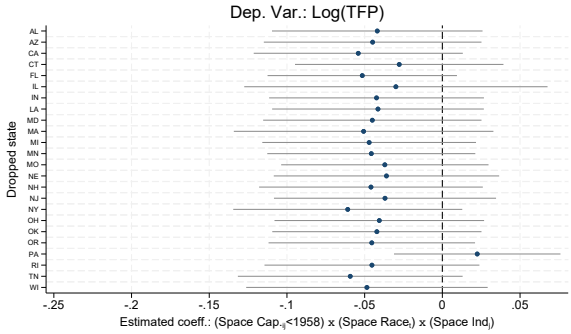
B. Log(Employment)



C. Log(Capital)



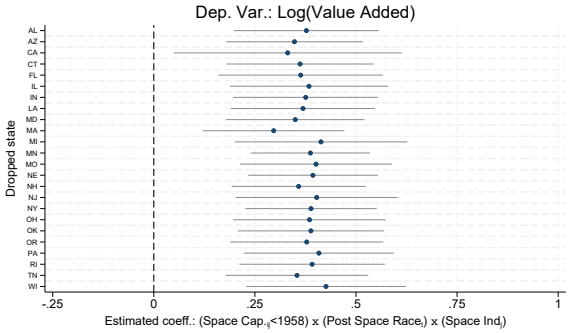
D. Log(TFP)



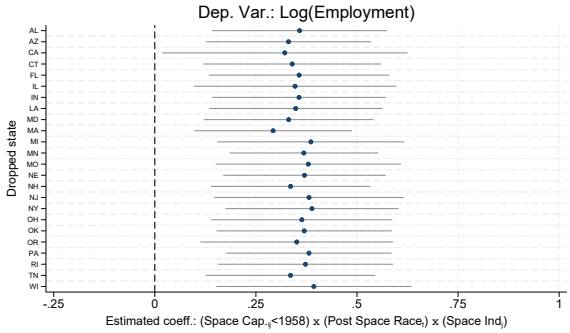
Notes: The estimates shown here graphically follow the regression estimation (equation (1)) presented in Table 3, column (1) in the main paper. Log(TFP) is defined as log(A_{ijt}) from estimating the production function $Y_{ijt} = A_{ijt}K_{ijt}^{\alpha}L_{ijt}^{\beta}$ by OLS. Each panel in the table displays the coefficient from estimating one version of equation (1) in the text, but omitting one state at a time. The Panels display coefficients and 95% confidence intervals for High Space Capability$_{i,<1958}$ × Space Race$_t$ × Space Ind$_t$ for each outcome.

Figure A4: Leave One State Out Estimates – Post- Space Race Era Estimates

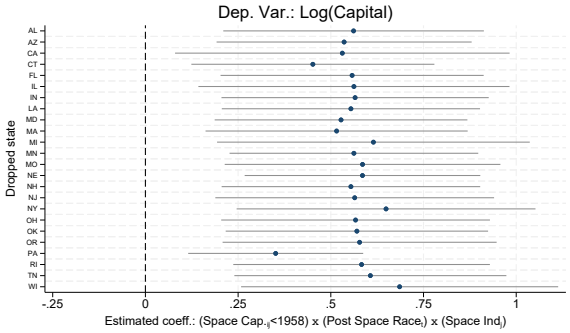
A. Log(Value Added)



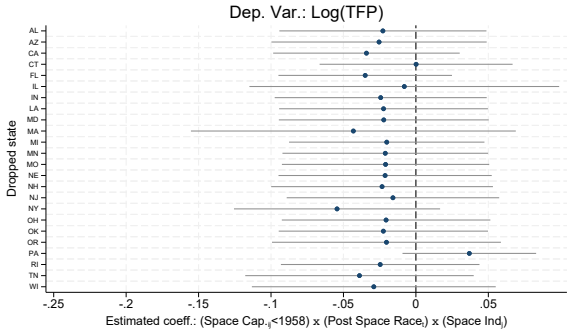
B. Log(Employment)



C. Log(Capital)



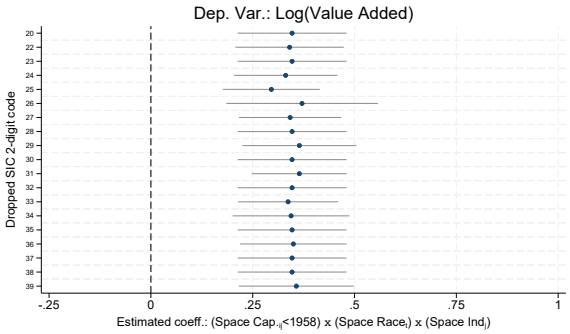
D. Log(TFP)



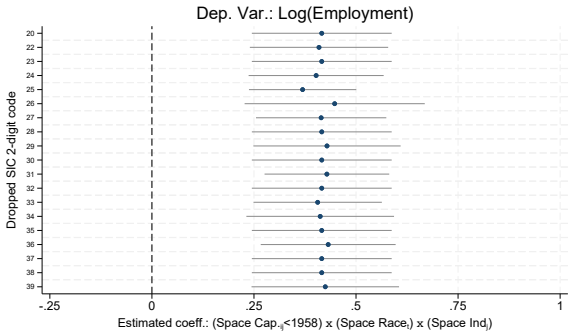
Notes: The estimates shown here graphically follow the regression estimation (equation (1)) presented in Table 3, column (1) in the main paper. Log(TFP) is defined as log(A_{ijt}) from estimating the production function $Y_{ijt} = A_{ijt}K_{ijt}^{\alpha}L_{ijt}^{\beta}$ by OLS. Each panel in the table displays the coefficient from estimating one version of equation (1) in the text, but omitting one state at a time. The Panels display coefficients and 95% confidence intervals for High Space Capability$_{i,<1958}$ × Post Space Race$_{t}$ × Space Ind$_{t}$ for each outcome.

Figure A5: Leave One Industry Out Estimates – Space Race Era Estimates

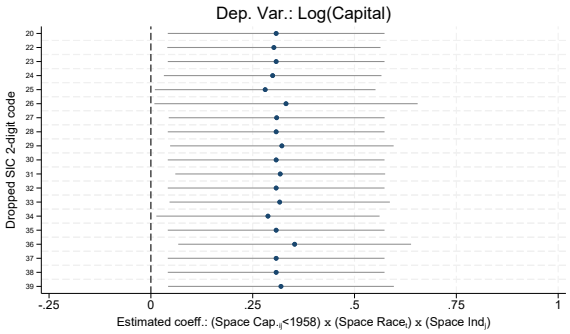
A. Log(Value Added)



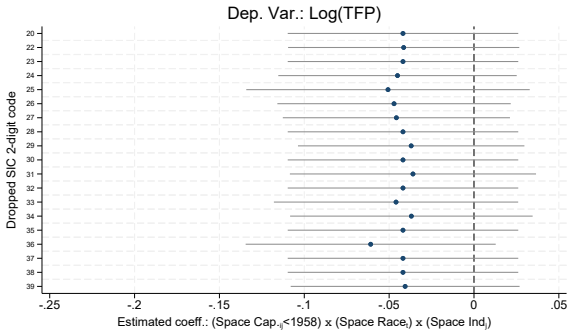
B. Log(Employment)



C. Log(Capital)



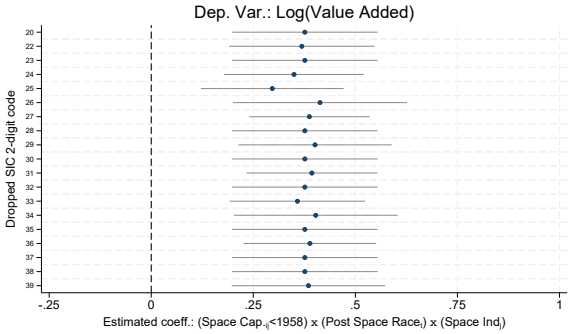
D. Log(TFP)



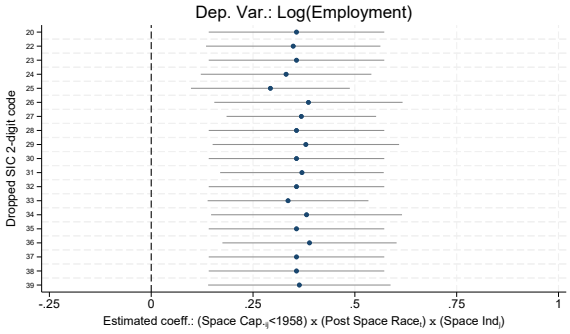
Notes: The estimates shown here graphically follow the regression estimation (equation (1)) presented in Table 3, column (1) in the main paper. Log(TFP) is defined as $\log(A_{ijt})$ from estimating the production function $Y_{ijt} = A_{ijt}K_{ijt}^{\alpha}L_{ijt}^{\beta}$ by OLS. Each panel in the table displays the coefficient from estimating one version of equation (1) in the text, but omitting one industry at a time. The Panels display coefficients and 95% confidence intervals for High Space Capability$_{i,<1958}$ × Space Race$_t$ × Space Ind$_t$ for each outcome.

Figure A6: Leave One Industry Out Estimates – Post- Space Race Era Estimates

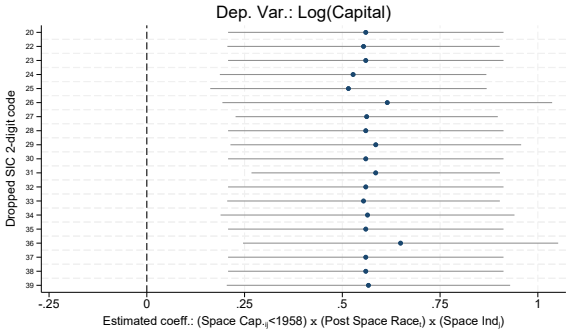
A. Log(Value Added)



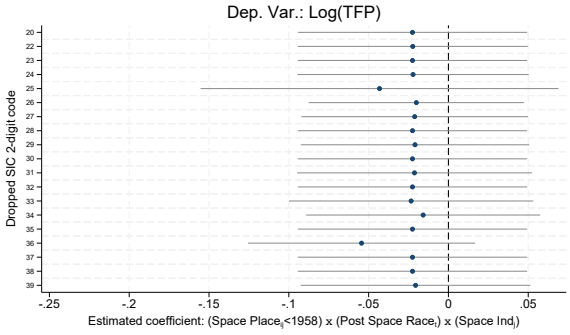
B. Log(Employment)



C. Log(Capital)



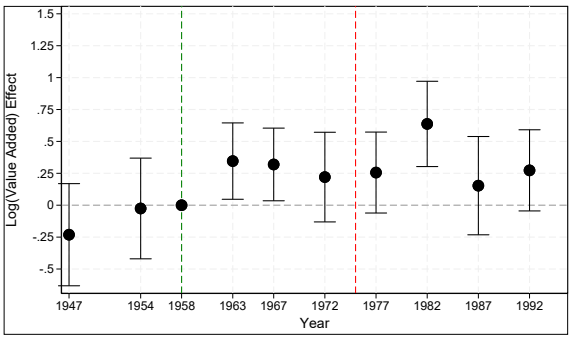
D. Log(TFP)



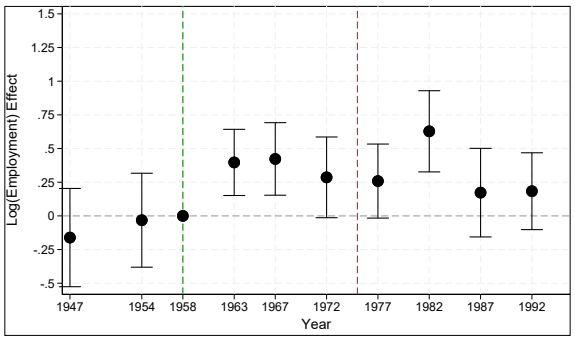
Notes: The estimates shown here graphically follow the regression estimation (equation (1)) presented in Table 3, column (1) in the main paper. Log(TFP) is defined as log(A_{ijt}) from estimating the production function $Y_{ijt} = A_{ijt}K_{ijt}^{\alpha}L_{ijt}^{\beta}$ by OLS. Each panel in the table displays the coefficient from estimating one version of equation (1) in the text, but omitting one industry at a time. The Panels display coefficients and 95% confidence intervals for High Space Capability$_{i,<1958}$ × Post Space Race$_t$ × Space Ind$_i$ for each outcome.

Figure A7: Dynamic Effects – US Space Technology Similarity Measure

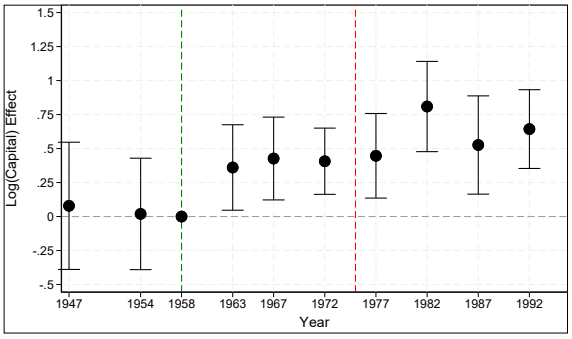
Panel A: Log(Value Added)



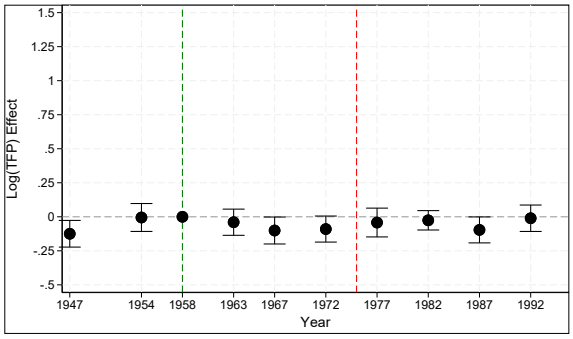
Panel B: Log(Employment)



Panel C: Log(Capital)



Panel D: Log (TFP)



Notes: This figure is akin to Figure 5 in the main paper. The only exception is that an industry-county's pre-1958 space capability was computed using post-Sputnik NASA patents instead of post-Sputnik Soviet technology.