ONLINE APPENDIX

Natural Disasters, Industrial Policy, and Innovation

Davide M. Coluccia & Mara P. Squicciarini

September, 2025

TABLE OF CONTENTS

A DATA APPENDIX

This section provides further information on the primary data sources underlying the datasets used in the main analysis and the methods adopted to construct them.

A.I Patent Data

We collect a novel dataset of US patents spanning 1853–1900. Compared to existing datasets, we develop a new methodology to extract information from patent texts. This section describes this new repository.

A.I.1 Motivation

Our analysis requires detailed information on each inventor's location to assign the patents to the closest metropolitan area. We also need text data to identify construction, wood- and non-wood-related innovations, and technological classes. Additionally, the United States Patent Office number allows to match patent records to the novelty measure produced by Kelly et al. (2021). We now briefly discuss why publicly available datasets do not satisfy these data requirements.

There are three datasets of US patents covering the period 1850–1900 (Andrews, 2021). Among those, the data produced by Sarada et al. (2019) and Petralia et al. (2016) are publicly available. Sarada et al. (2019) digitize the Annual Reports of the Commissioner of Patents and the Annual Indices of Patents. This dataset starts in 1870 and does not contain geographical information on inventors' addresses beyond their state. In addition, it does not include text data, as the primary source is not the text of the patent but an index.

Petralia et al. (2016) extracted patent information from digitized images of US patents, as we do. The data contain information on the county of the inventors. This data is, therefore, not suited for our analysis, which is performed at a more granular level. Importantly, this dataset does not contain the text of the patents, which we use to identify construction and wood- and non-wood-related construction innovations.

A.I.2 Methodology to Assemble the Dataset

We rely on the corpus of digitized US patents produced by Google Patents. We collect the universe of patent images by adapting the algorithm developed by Moser and San (2020). On top of the patent images, we collect from Google Patents the CPC technological classes of each patent and the backward and forward citations, which we do not use in this paper. Then, we employ a commercial optical character recognition (OCR) software—Amazon's `textract`—to convert the images into machine-readable text.

From the patent texts, we follow a procedure similar to Coluccia and Dossi (2025), who apply it to historical British patents. We use a large language model—`GPT 4o-mini`—to extract the name and surname of the inventors, their address of residence, the filing and issue date of the patent, and information on whether the patent has a firm assignee. Compared to standard extraction methods relying on regular expressions, such as those employed by Berkes (2018), the flexibility of large language models allows us to parse the data even in the presence of minor OCR errors or syntactical inconsistencies.

Lastly, we use a commercial georeference software—`Google Maps API`—to assign latitude and longitude coordinates to the address of each inventor. The address listed on the patent is the residence of the inventor. In small towns, the address indicates the town. In larger cities, the proper address appears on the patent. For consistency, we georeference the town given that the analysis is performed at that level of spatial aggregation.

A.I.3 Comparison with Existing Datasets

The dataset contains information on all patents listed on Google Patents issued between 1853 and 1900. We have the full digitized text of the patent, the name and surname of the inventor(s), their address(es), their latitude and longitude, the filing and issue date of the patent, and an assignee flag. Since the dataset contains the unique patent number, we can match our records with the novelty measure assembled by Kelly et al. (2021). From Google Patents, we also have the CPC technological classification associated with each patent.

Compared to Andrews (2021), whose coverage spans 1870 to 1942, our data has more detailed geographical information, more systematic inventors' names and surnames, full-text access, CPC technology classification information, filing, as well as issue date, and more precise dates. Compared to Petralia et al. (2016), our dataset contains more granular geographical information—down to the georeferenced town of each inventor, as opposed to the county—, CPC classes, filing and issue dates, and full-text access. Our data essentially mimics the state-of-the-art dataset produced by (e.g. Berkes, 2018), which is unfortunately not publicly accessible, over a shorter time window.

A.I.4 External Validation

Even though the existing datasets are not suited for our empirical application, we use them to validate the coverage of our new sample. Figure C.1a compares the number of patents in our dataset with Petralia et al. (2016)—dashed gray line—and Sarada et al. (2019)—dotted gray line. Figure C.1b reports coverage rate, defined as the number of patents in our dataset relative to Petralia et al. (2016)—red solid line—and Sarada et al. (2019)—gray dashed line, expressed in percentage.

Our series mimics Petralia et al. (2016) throughout the period. The coverage rates vary between 90% before the 1870s and 110% during the later part of the period, indicating that our dataset contains more patents. The data by Sarada et al. (2019) is less comprehensive—our dataset contains one-third more patents between 1870 and 1880 and one-tenth more towards the end of the century—but displays a similar co-movement with our series. Importantly, even though we cannot perform validation with the dataset developed by Berkes (2018), their data and that of Petralia et al. (2016) display substantial overlap before 1900, which indicates that our coverage rate of his dataset will mimic that with Petralia et al. (2016).

A.I.5 Methodology to Identify Construction Patents

We adopt a simple dictionary-based procedure to identify patents related to construction, non-wood construction, and wood construction.

First, we generate a dictionary of 30 words related to construction using `GPT-o3`. The words are displayed in Table D.1. Then, we search for the number of instances each word appears in the text of each patent. A patent is identified as "construction-related" if at least one word appears at least five times in its text. The results are not sensitive to alternative, more demanding thresholds, but increasing the required threshold increases the rate of false negatives without significantly impacting the rate of false positives. We provide a more detailed discussion below.

Second, *among construction patents*, we search for wood- and non-wood-related innovations. Using `GPT o1`, we generate a list of 12 wood-related and 11 non-wood-related words. We then count the number of instances each word appears in the patents' texts. A patent is identified as "wood-related construction" if it mentions at least one word from column (1) of Table D.1 at least five times and at least one word from column (2). Analogously, a patent is identified as "non-wood-related construction" if it mentions at least one word from column (1) of Table D.1 at least five times and at least one word from column (3).

We manually check the plausibility of the results obtained using this dictionary-based approach on a random sample of 200 patents. Within this sample, 41 patents were flagged as construction-related, five as wood-construction, and six as non-wood-construction. We did not find false positives, and the algorithm missed four patents that a human would have coded as construction, thus yielding a 10% false negatives rate among construction patents. Increasing the threshold to ten patents over the same sample decreased the rate of false negatives to 5% but decreased the rate of true positives by almost one-third. While suggestive, this exercise indicates that the five-word threshold is reasonable, if ad-hoc, heuristic to identify construction-related innovations.

A.II Population Census

We use individual-level data mapped to CPP locations to compute demographic characteristics of the metropolitan areas and counties. The synthetic control and synthetic difference-in-differences estimates use, as balancing variables, the share of natives, blacks, imputed income per capita (OCCSCORE), and the employment shares by occupation and industry.

To compute the employment share by occupation, we use the OCC1950 standardized codes to construct a coarser occupational taxonomy that follows from the categories provided by IPUMS: professionals (OCC1950 between 0 and 99), farmers (OCC1950 between 100 and 123 and 810 and 840), managers (OCC1950 between 200 and 290), clerical (OCC1950 between 300 and 390), sales (OCC1950 between 300 and 390), craftsmen (OCC1950 between 500 and 595), operatives (OCC1950 between 600 and 690), services (OCC1950 between 700 and 790), and laborers (OCC1950 between 910 and 970).

Similarly, to construct the employment share by industry, we use the IND1950 standardizes codes to construct a coarser industry taxonomy following IPUMS: agriculture (IND1950 between 105 and 126), chemicals (IND1950 between 466 and 478), construction (IND1950 between 246 and 246), engineering (IND1950 between 367 and 388 or 898), liberal professions (IND1950 between 716 and 897), metallurgy (IND1950 between 336 and 348), miscellaneous manufacturing (IND1950 between 306 and 358, 406 and 429, and 456 and 459), public administration (IND1950 between 906 and 946), textiles (IND1950 between 436 and 449), trade (IND1950 between 606 and 699), transportation (IND1950 between 506 and 579), and utilities (IND1950 between 596 and 598 and 826 and 859).

All variables from the population census are measured in 1870, i.e., the year before the Great Chicago Fire.

A.III Manufacturing Census

We compile manufacturing data—number of establishments, value of production, labor cost, material cost, and capital—from county-by-industry data from the Census of Manufactures transcribed by Hornbeck and Rotemberg (2024). The sample comprises all counties with at least one metropolitan area.

We use the industry concordance table provided by the authors to construct a county-by-industry panel at the decade level between 1860 and 1880. The final sample is a decadal county-level panel that records the various variables for different industries. In particular, we group the "construction," "construction materials," and "furniture" as one "construction" industry, the "brick, stone, and tile," "marble and stone work," and "lime and cement" titles as one "non-wood manufacturing" industry, and the "lumber, sawed," "wood products, other," "wood, turned and carved," "wooden ware," and “saws” titles as one “wood manufacturing” industry. All other industry classifications remain unchanged.

A.IV Historical Landmarks

We collect all historical landmarks from the “National Register of Historic Places” (Stutts, 2024). We georeference each entry using the provided address and `Google Maps API`. We assign latitude and longitude to 99% of the landmarks in the sample. Each landmark is then allotted to the closest CPP location within 20 Km, following the same procedure we apply to the patent records. The National Register does not contain information on the construction year. We thus manually search each entry on Wikipedia and assign a construction year to 77% of the entire dataset. Among the landmarks in the metropolitan analysis sample, 85% have a recorded construction year. We include in the sample entries that do not refer to “buildings” (64% of the sample), thus excluding “districts” (20%), “objects” (0.25%), and “sites” (13%).

B SUMMARY OF THE ROBUSTNESS ANALYSES

This section provides complementary information on the robustness exercises mentioned in passing in the main text.

B.I Patent Novelty Data

We use the text-based measure by Kelly et al. (2021) to explore the effect of the 1871 Fire on economically relevant innovation. Intuitively, a patent that is more similar to future patents than to previous patents is labeled as "more innovative." More formally, let the backward inverse-document frequency associated with word w be defined as

$$\text{BIDF}_{w,t} \equiv \log\left(\frac{\text{Number of Patents Issued Before } t}{1 + \text{Number of Patents Issued Before } t \text{ that contain word } w}\right). \tag{B.1}$$

To each patent-word pair, it associates a variable equal to the number of instances word w appears in patent i, normalized by the patent length. Let i denote the patent and the set of words it contains. The term-frequency weight is equal to

$$\text{TF}_{wi} \equiv \frac{\sum_{c \in i} 1(c = w)}{\sum_{c \in i} 1(c)}, \tag{B.2}$$

where the numerator returns how many times word w appears in patent i, and the denumerator is the number of words in patent i. The TF-BIDF associated with word w, patent i at time t is the product between the TF and the BIDF:

$$\text{TF-BIDF}_{wi,t} \equiv \text{TF}_{wi} \times \text{BIDF}_{w,t}. \tag{B.3}$$

The vector TF-BIDF$_{i,t}$ thus collects the TF-BIDF$_{wi,t}$ for all words w in i, normalized by its norm to have unit length.

The approach allows the representation of each patent as a TF-BIDF vector. One can thus compute a measure of similarity—in their case, the cosine similarity—between each patent pair. In particular, the backward similarity is the average similarity between i and all previous patents within τ_1 years:

$$\text{Backward Similarity}_i^{\tau_1} \equiv \frac{1}{|\mathcal{F}_i^{-\tau_1}|} \sum_{j \in \mathcal{F}_i^{-\tau_1}} \rho_{i,j}, \tag{B.4}$$

where the set $\mathcal{F}_i^{-\tau_1}$ denotes the set of US patents issued within τ_1 years from the issue year of patent i and $\rho_{i,j}$ is the cosine similarity between the vectors TF-BIDF$_{i,t}$ and TF-BIDF$_{j,t}$. Analogously, the forward similarity is the average similarity with all patents in the later τ years:

$$\text{Forward Similarity}_i^{\tau_2} \equiv \frac{1}{|\mathcal{F}_i^{+\tau_2}|} \sum_{j \in \mathcal{F}_i^{+\tau_2}} \rho_{i,j}, \tag{B.5}$$

where the set $\mathcal{F}_i^{+\tau_2}$ denotes the set of US patents issued τ_2 years after the issue year of patent i. Given these measures, one can compute the similarity of p with future relative to previous patents:

$$\text{Excess Forward Similarity}(\tau_1, \tau_2) \equiv \frac{\sum_{j \in \mathcal{F}_j^{+\tau_2}} \rho_{i,j}}{\sum_{j \in \mathcal{F}_i^{-\tau_1}} \rho_{i,j}}. \tag{B.6}$$

In our application, we take $\tau_1 = \tau_2 = 5$ to have a symmetric 5-year window around each patent. The results remain qualitatively unchanged when using 1- and 10-year symmetric windows. Following Kelly et al. (2021), we partial out year fixed effects from the raw Excess Forward Similarity(τ_1, τ_2) measure to ensure that aggregate trends in language and patent redaction do not influence the results. A patent is then defined as "novel" if it is in the top 20% of the excess forward similarity distribution. The results remain qualitatively unchanged using different thresholds at the 5% and 10%.

Figure C.3 displays the synthetic control results of the effect of the Great Chicago Fire on novel patenting in construction (Figure C.3a), non-wood construction (Figure C.3b) and wood construction (Figure C.3c). The divergent trajectories between Chicago and the synthetic control unit indicate that the effect of the Fire on innovation is not disproportionately driven by economically irrelevant or unoriginal innovation.

B.II Synthetic Difference-in-Differences

The synthetic difference-in-differences estimator developed by Arkhangelsky et al. (2021) nests the insights of standard synthetic control and difference-in-differences estimators. As highlighted by Arkhangelsky et al. (2021), the synthetic control method is typically applied in settings with one or a few treated units, where the parallel trends assumption required by the difference-in-differences estimator is unlikely to hold. The synthetic difference-in-differences estimator weights units in the control group to match the treated unit in terms of a set of specified pre-treatment observable characteristics and the outcome to maximize the empirical plausibility of the parallel trends assumption.

Formally, let i and t denote units and time periods. Exposure to the treatment is $W_{it} = \{0, 1\}$, and Y_{it} denotes the outcome. the synthetic control estimator selects weights ω_i^{sc} to minimize the distance between treated and control units and estimates the treatment effect as

$$(\hat{\tau}^{sc}, \hat{\mu}, \hat{\beta}) = \arg\min_{\mu,\beta,\tau} \left\{ \sum_i \sum_t (Y_{it} - \mu - \beta_t - W_{it}\tau)^2 \hat{\omega}_i^{sc} \right\}, \tag{B.7}$$

where $\hat{\tau}^{sc}$ is the estimated treatment effect. The difference in differences estimator, by contrast, weights all units in the same way, but it includes unit fixed effects to leverage within-unit variation:

$$\left(\hat{\tau}^{did}, \hat{\mu}, \hat{\alpha}, \hat{\beta}\right) = \arg\min_{\mu,\alpha,\beta,\tau} \left\{ \sum_i \sum_t (Y_{it} - \mu - \alpha_i - \beta_t - W_{it}\tau)^2 \right\}. \tag{B.8}$$

The synthetic difference-in-differences estimator nests these two approaches. First, it selects weights ω_i^{sdid} to minimize the distance between treated and control units in terms of pre-treatment outcome values and characteristics. Moreover, it selects λ_t^{did} that balance pre-exposure time periods with post-exposure ones. Then, it solves for the average treatment effect as in the DiD estimator, applying the so-computed weights:

$$\left(\hat{\tau}^{sdid}, \hat{\mu}, \hat{\alpha}, \hat{\beta}\right) = \arg\min_{\mu,\alpha,\beta,\tau} \left\{ \sum_i \sum_t (Y_{it} - \mu - \alpha_i - \beta_t - W_{it}\tau)^2 \hat{\omega}_i^{sdid} \hat{\lambda}_t^{sdid} \right\}. \tag{B.9}$$

By weighting observations to minimize the distance between treated and control units, the SDiD estimator emphasizes units that are, on average, similar to the treated unit and periods that are, on average, similar to the target periods.

In our application, we compute the SDiD weights by including the population, share of whites and foreign-born, employment shares by occupation, and employment shares by industry. As in the synthetic control baseline case, these variables were measured in the 1870 census, before the Chicago and Boston fires.

We display the SDiD results in terms of standard panel event-study estimates to visualize the evolution of treatment effects over time. In addition, by looking at the pre-treatment differences between the treated (Chicago and Boston) and control units, we can assess the empirical plausibility of the parallel trends assumption. We follow the logic outlined in Clarke, Pailañir, Athey and Imbens (2023) to construct these figures. We wish to estimate, for each period t,

$$(\bar{Y}_t^T - \bar{Y}_t^C) - (\bar{Y}_0^T - \bar{Y}_0^C), \tag{B.10}$$

where $\bar{Y}_t^T$ and $\bar{Y}_t^C$ denote the average outcome for treated and control units at time t, and $\bar{Y}_0^T$ and $\bar{Y}_0^C$ denote the average pre-treatment outcome values for treated and control units. These are computed as

$$\bar{Y}_0^T = \sum_{t=1}^{\tau-1} \hat{\lambda}_t^{sdid} \bar{Y}_t^T, \tag{B.11}$$

for the treatment group and, similarly, as

$$\bar{Y}_0^C = \sum_{t=1}^{\tau-1} \hat{\lambda}_t^{sdid} \bar{Y}_t^C, \tag{B.12}$$

for the control group. The term τ denotes the treatment period. To compute the confidence intervals, we apply block-bootstrap resamples with 100 replications.

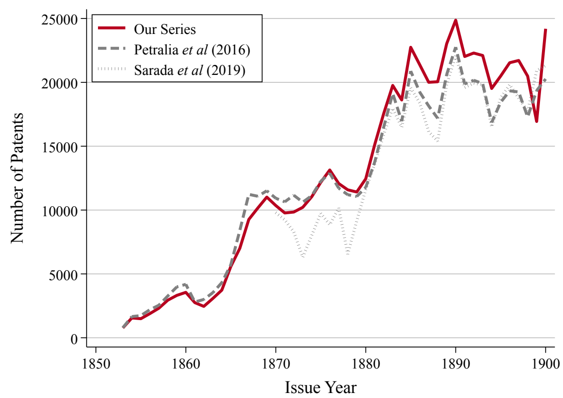
We display SDiD estimates for all the results shown in the main text. Figure C.4 displays the effect of the Great Chicago (Figure C.4a) and Boston (Figure C.4b) Fires on construction innovation. Figure C.10 focuses on the effect of the Chicago Fire on non-wood-related (Figure C.10a) and wood-related (Figure C.10b) patenting. Figure C.9 displays the results on historical landmarks. Analogously, we report the SDiD estimates on construction manufacturing (Figure C.8), as well as non-wood construction manufacturing (Figure C.12).

In all cases, the synthetic DiD estimates confirm the baseline results obtained through the more traditional synthetic control approach. Additionally, in most cases, the pre-treatment differences between treated and control units are statistically insignificant and are always quantitatively very small. These patterns confirm the empirical plausibility of the parallel trends assumption that requires that, in the absence of the Fires, the outcomes in Chicago and Boston and in the control units would have evolved similarly.

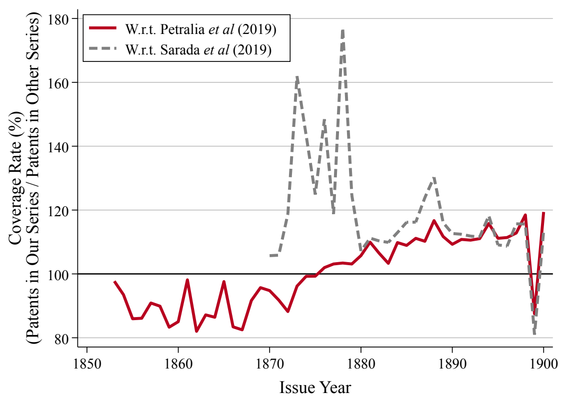
C ADDITIONAL FIGURES

Figure C.1. Comparison of Own and External Patent Repositories

(a) Total Number of Patents

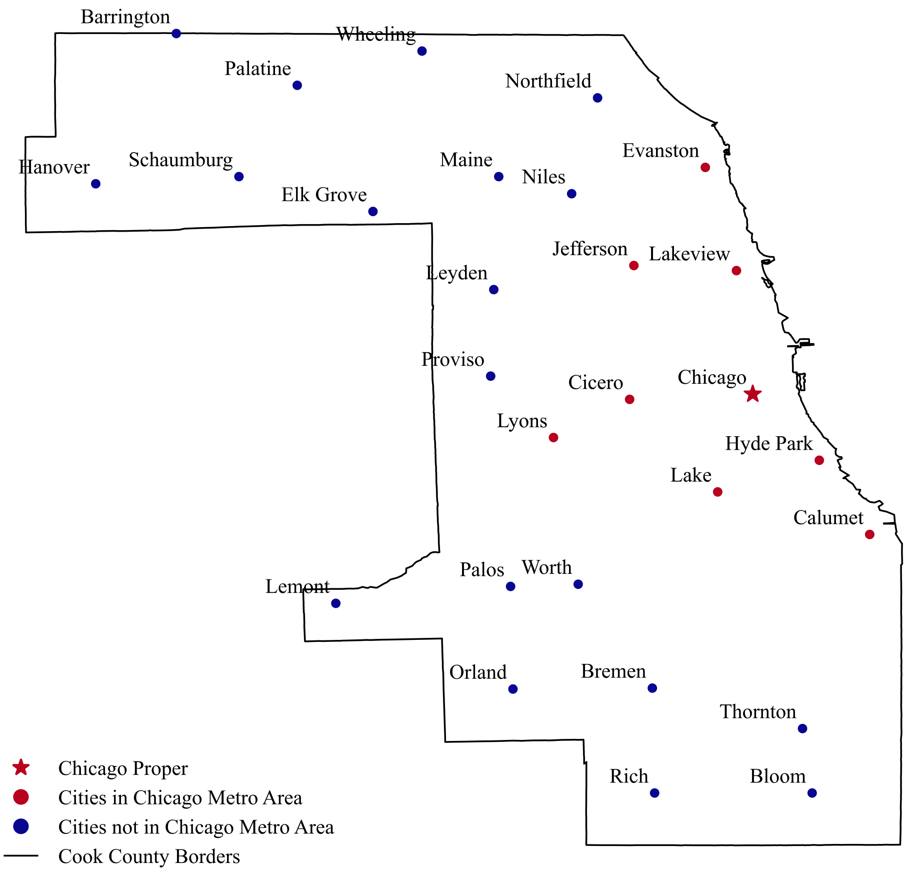


(b) Coverage Rate (%)



Notes. This Figure compares the number of patents in the dataset produced for this paper and data from Petralia et al. (2016) and Sarada et al. (2019). Panel C.1a reports the total number of patents in our data (red line), Petralia et al. (2016) (gray dashed line), and Sarada et al. (2019) (gray dotted line). Panel C.1b reports the coverage rate, computed as the ratio between the number of patents in our dataset and the number of patents in Petralia et al. (2016) (red line) and Sarada et al. (2019) (gray dashed line). The dataset by Sarada et al. (2019) starts in 1870; for comparability, we restrict it to "Utility" patents only. Referenced on page: A3.

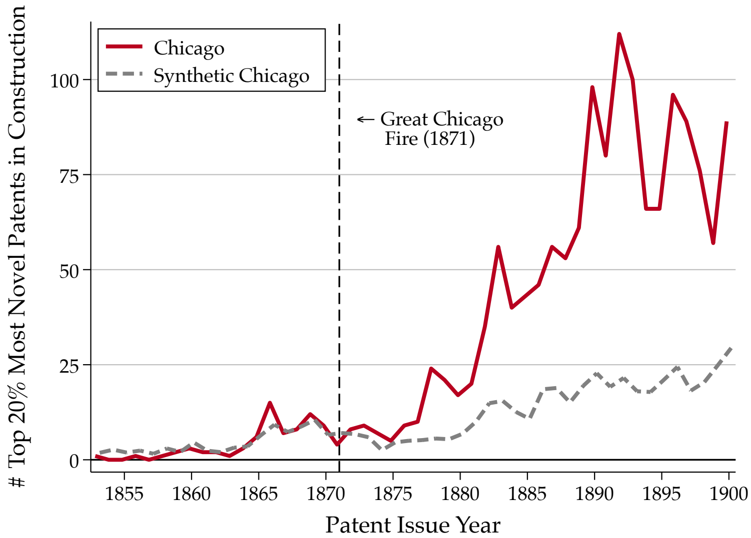
Figure C.2. Example of Metropolitan Area: Chicago



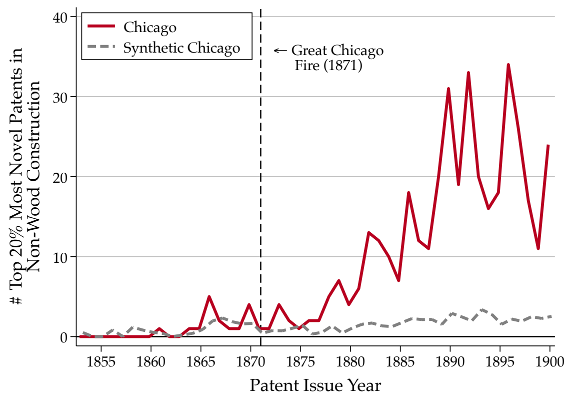
Notes. This Figure displays the construction of the Chicago Metropolitan Area according to the procedure described in Section II.E. Each dot reports the coordinates of a single CPP location in Cook County, whose 1870 borders are displayed in the black solid line. The red star displays the location of Chicago, which is the only CPP location with more than 20,000 inhabitants in Cook County. The red dots are the minor towns—i.e., the CPP locations with less than 20,000 inhabitants—that are closer than 20 kilometers from the center of Chicago and thus are considered part of the Chicago Metropolitan Area. In this case, the Chicago Metropolitan Area includes Evanston, Lakeview, Jefferson, Cicero, Lyons, Lake, Hyde Park, and Calumet. The blue dots are towns below 20,000 inhabitants that are further than 20 kilometers from the center of Chicago and are thus excluded from its metropolitan area. Referenced on page: 12.

Figure C.3. Synthetic Control Estimates of the Effect of the Great Chicago Fire on Construction Innovation: Novel Patents

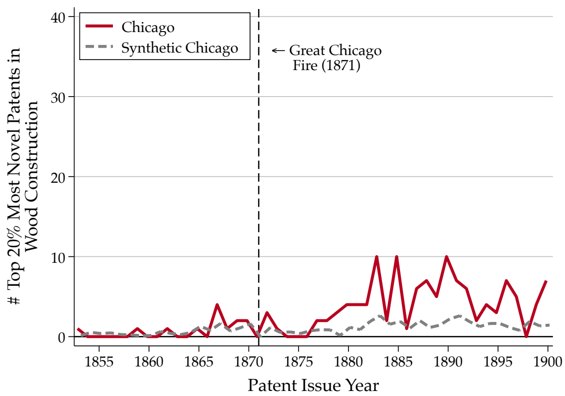
(a) All Construction Patents



(b) Non-Wood Construction Patents



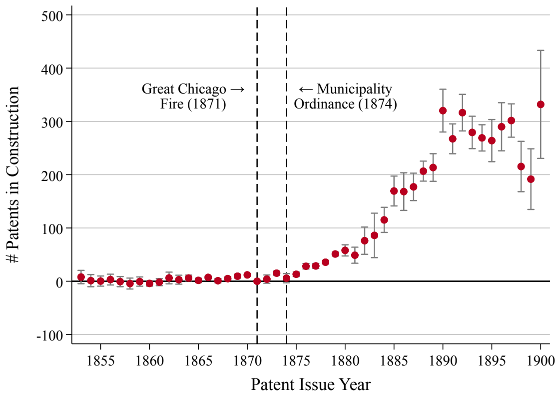
(c) Wood Construction Patents



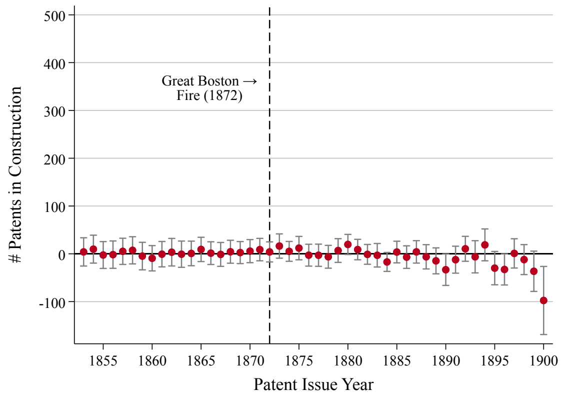
Notes. This Figure reports the effect of the Great Chicago Fire (1871) on construction innovation in Chicago. The dependent variable is the number of patents in the top 20% of the novelty distribution of the text-based novelty measured developed by Kelly et al. (2021). The unit of observation is a metropolitan area at a yearly frequency between 1853 and 1900. In Panel C.3a, we look at construction patents; Panel C.3b reports the effect on non-wood construction patents. In Panel C.3c, we report the effect on wood construction patents. The black dashed line marks the year of the Great Chicago Fire (1871). The red line refers to Chicago; the dashed gray line refers to the synthetic control. Referenced on pages: 17, 23, B8.

Figure C.4. Synthetic Difference-in-Differences Estimates of the Effect of the Great Chicago and Boston Fires on Construction Innovation

(a) Chicago

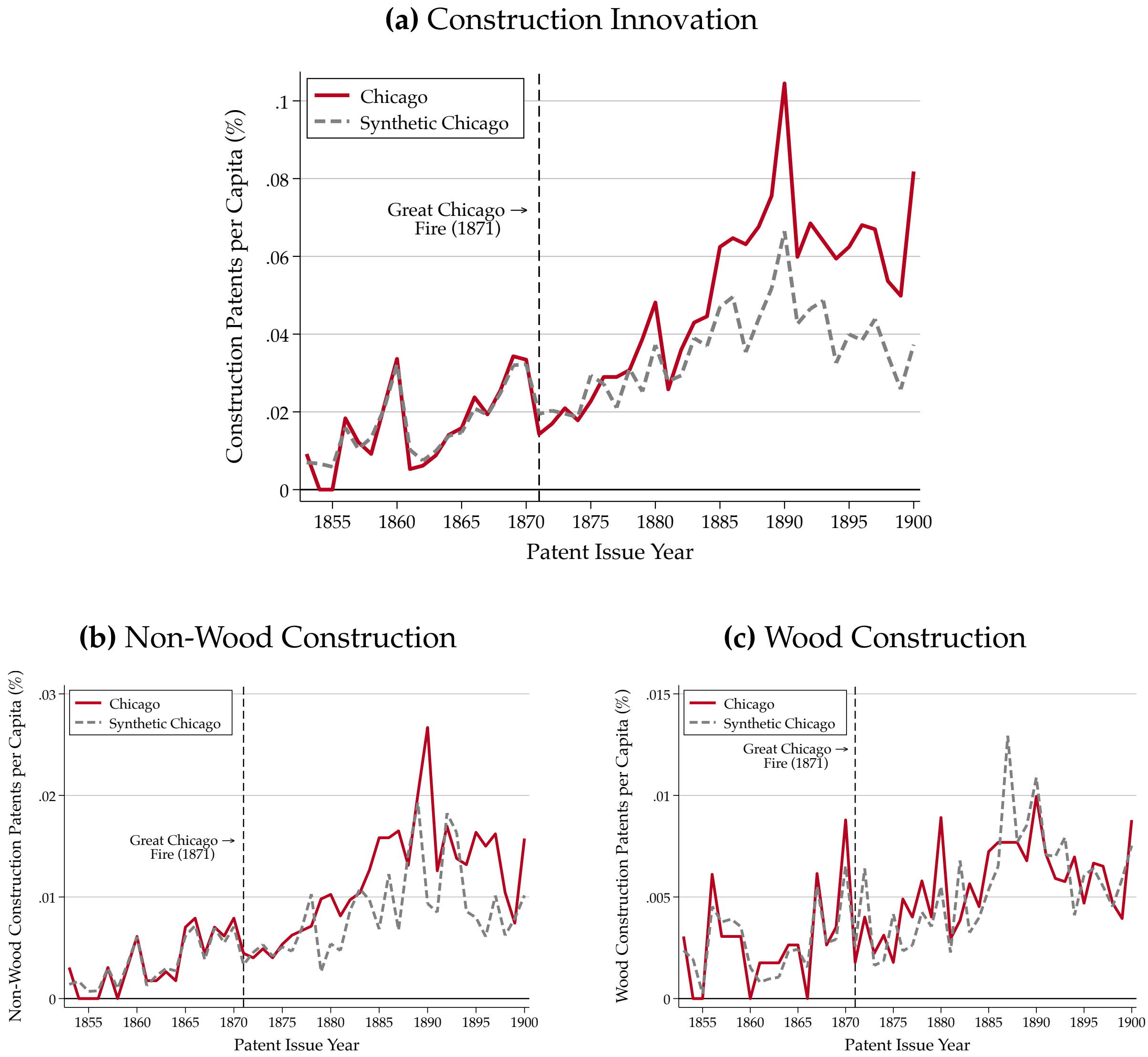


(b) Boston



Notes. This Figure reports the effect of the Great Chicago (1871) and Boston (1872) Fires on construction innovation in Chicago (Panel C.4a) and Boston (Panel C.4b). The dependent variable is the number of patents in construction. The unit of observation is a metropolitan area at a yearly frequency between 1853 and 1900. The Figures report event-study coefficients obtained using the synthetic difference-in-differences methods developed by Arkhangelsky et al. (2021) and obtained following the procedure described in Clarke et al. (2023). Standard errors are estimated using bootstrap sampling with 100 replications; gray bars report 95% confidence intervals. The covariates included in the synthetic difference-in-differences estimation are the same as in the synthetic control estimation, except for the pre-treatment outcome values. The black dashed line marks the year of the Great Chicago Fire (1871), the Chicago Municipality Ordinance prohibiting wood construction (1874), and the Great Boston Fire (1872). Referenced on pages: 18, 27, B10.

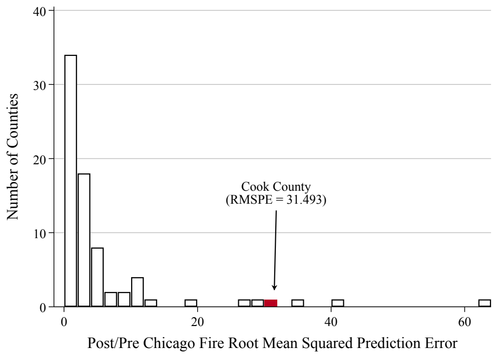
Figure C.5. Synthetic Control Estimates of the Effect of the Great Chicago Fire on Patenting per Capita



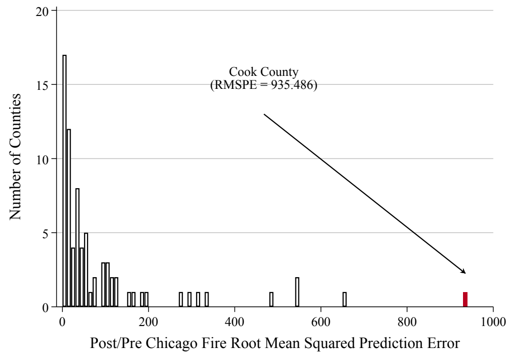
Notes. This Figure reports the effect of the Great Chicago Fire (1871) on construction innovation in Chicago. The dependent variable is the number of patents in construction (Panel C.5a), non-wood construction (Panel C.5b), and wood construction (Panel C.5c). Patenting activity is normalized by each metropolitan area's total employment in the decade, as measured in the population census, and is expressed in percentage terms. The unit of observation is a metropolitan area at a yearly frequency between 1853 and 1900. The black dashed line marks the year of the Great Chicago Fire (1871). The red line refers to Chicago; the dashed gray line refers to the synthetic control. Referenced on pages: 19, 23.

Figure C.6. Construction Manufacturing: Pre-Post Great Chicago Fire Synthetic Control Root Mean Squared Prediction Error Comparison

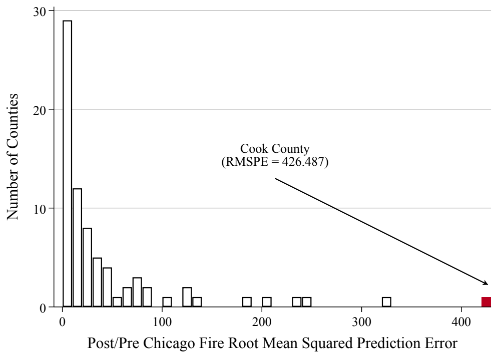
(a) Number of Establishments



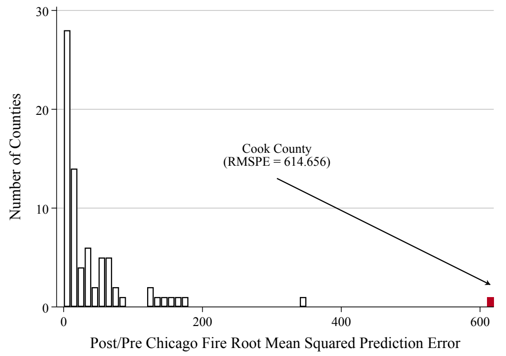
(b) Value of Production



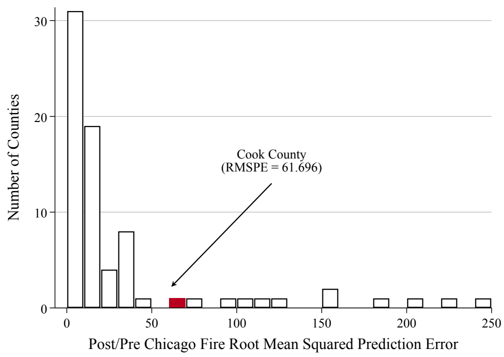
(c) Fixed Capital



(d) Cost of Materials

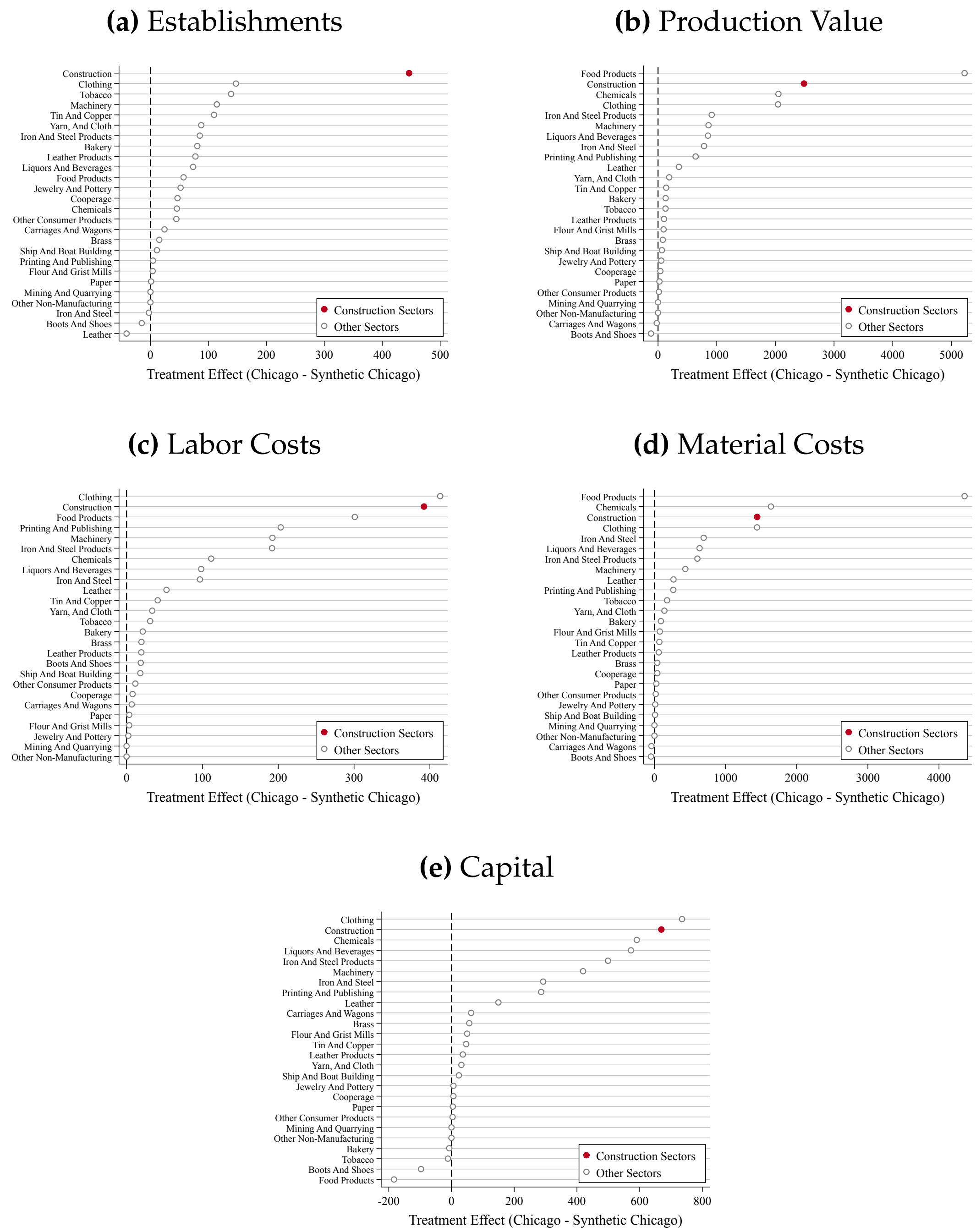


(e) Cost of Labor



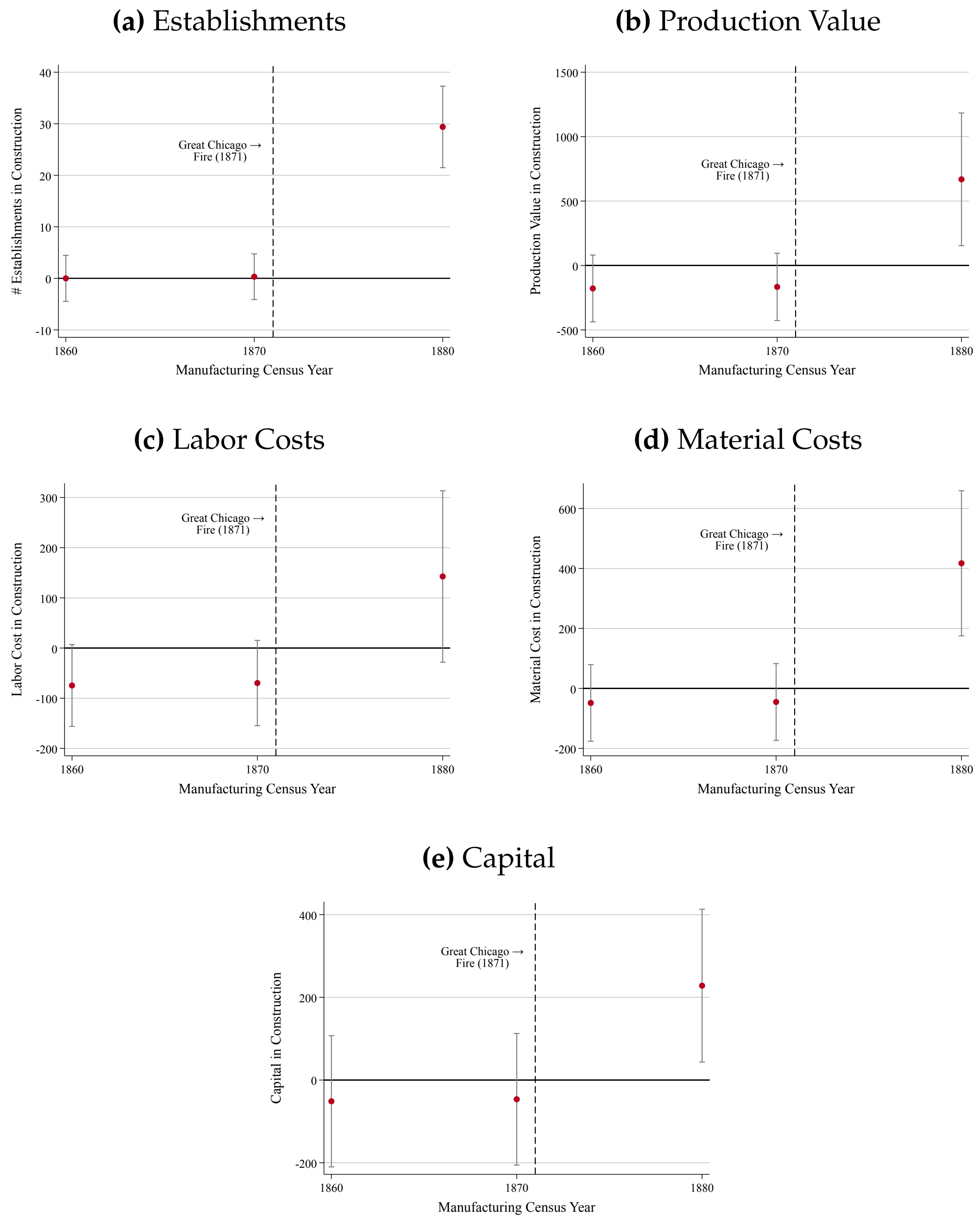
Notes. This Figure reports the impact of the Great Chicago (1871) on construction manufacturing activity as measured in the Census of Manufacturing. The unit of observation is a county at a decade frequency between 1860 and 1880. The dependent variable is: in Panel C.6a, the number of establishments; in Panel C.6b, production value; in Panel C.6c, fixed capital; in Panel C.6d, the cost of materials; in Panel C.6e, the cost of labor. Each figure reports the ratio between the post-Fire and pre-Fire mean squared prediction error across counties and highlights Cook County (IL) in red. The sample includes all counties with at least one metropolitan area. Referenced on page: 20.

Figure C.7. Synthetic Control Estimates of the Effect of the Great Chicago Fire on Manufacturing: Industry-by-Industry Estimates



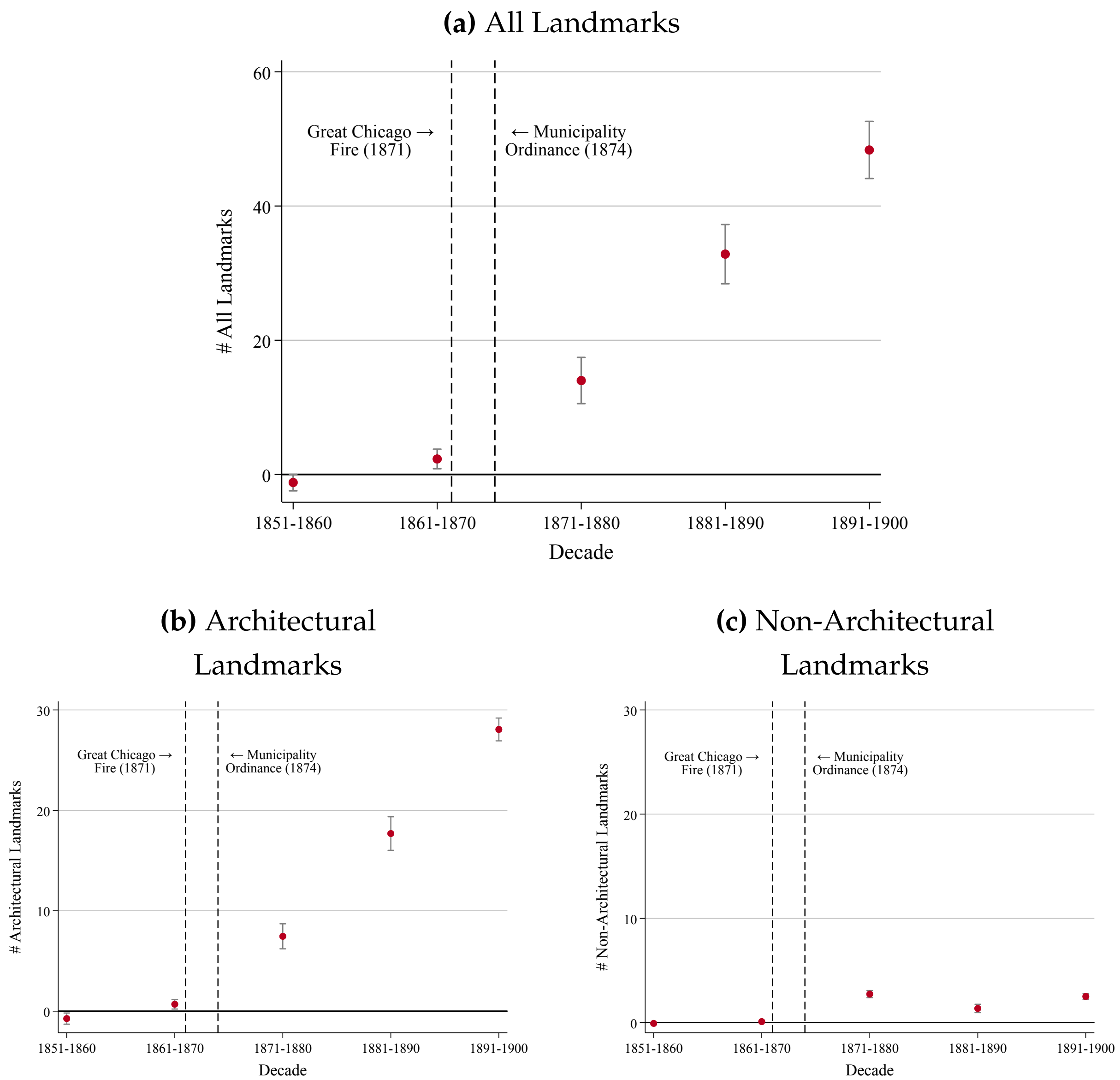
Notes. This Figure reports the impact of the Great Chicago Fire (1871) on manufacturing activity as measured in the Census of Manufacturing. The unit of observation is a county at a decade frequency between 1860 and 1880. The dependent variable is: in Panel C.7a, the number of establishments; in Panel C.7b, production value; in Panel C.7c, the cost of labor; in Panel C.7d, the cost of materials; in Panel C.7e, fixed capital. Each dot plots the treatment effect—i.e., the difference between Chicago and the synthetic control in 1880—by industry. Construction is displayed in red; all other sectors are displayed in gray. The sample includes all counties with at least one metropolitan area. Referenced on page: 20.

Figure C.8. Synthetic Difference-in-Differences Estimates of the Effect of the Great Chicago Fire on Construction Manufacturing



Notes. This Figure reports the impact of the Great Chicago Fire (1871) on construction manufacturing activity as measured in the Census of Manufacturing. The unit of observation is a county at a decade frequency between 1860 and 1880. The dependent variable is: in Panel C.8a, the number of establishments; in Panel C.8b, production value; in Panel C.8c, the cost of labor; in Panel C.8d, the cost of materials; in Panel C.8e, fixed capital. The Figures report event-study coefficients obtained using the synthetic difference-in-differences methods developed by Arkhangelsky et al. (2021) and obtained following the procedure described in Clarke et al. (2023). Standard errors are estimated using bootstrap sampling with 100 replications; gray bars report 95% confidence intervals. The sample includes all counties with at least one metropolitan area. Referenced on pages: 21, B10.

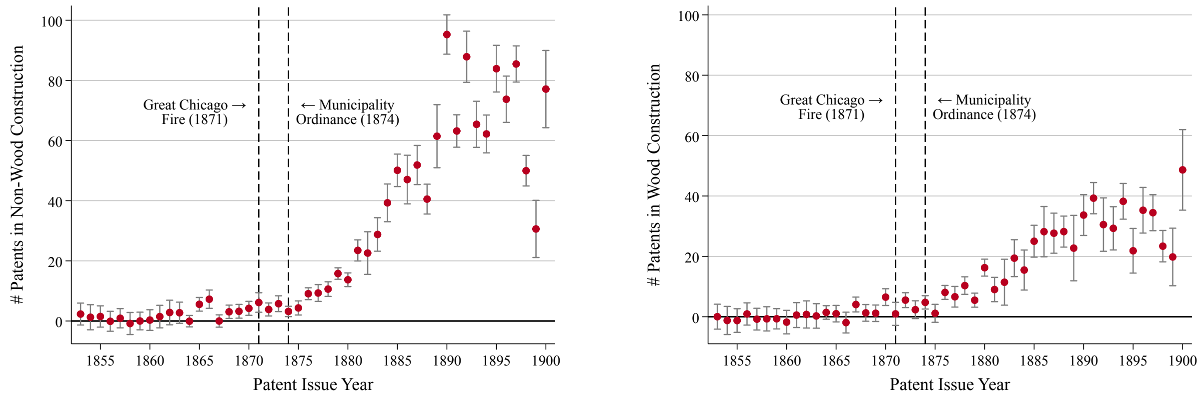
Figure C.9. Synthetic Difference-in-Differences Estimates of the Effect of the Great Chicago Fire on Historical Landmarks



Notes. This Figure reports the effect of the Great Chicago Fire (1871) on historical landmarks in Chicago. The dependent variable is the number of all historical landmark buildings (Panel C.9a), those listed due to architectural significance (Panel C.9b), and all other significant buildings (Panel C.9c). The unit of observation is a metropolitan area at a decade frequency between 1850 and 1900. The Figures report event-study coefficients obtained using the synthetic difference-in-differences methods developed by Arkhangelsky et al. (2021) and obtained following the procedure described in Clarke et al. (2023). Standard errors are estimated using bootstrap sampling with 100 replications; gray bars report 95% confidence intervals. Referenced on pages: 22, B10.

Figure C.10. Synthetic Difference-in-Differences Estimates of the Effect of the Great Chicago Fire on Wood and Non-Wood Construction Innovation

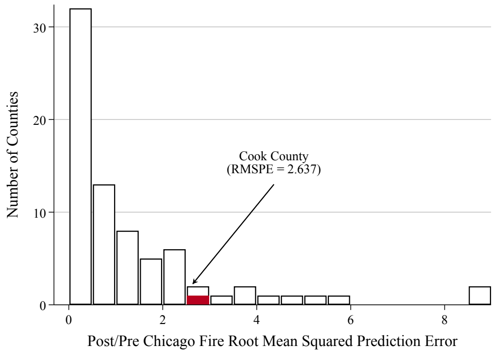
(a) Non-Wood Construction Patents **(b)** Wood Construction Patents



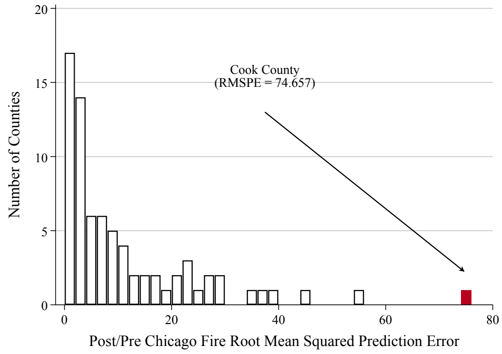
Notes. This Figure reports the effect of the Great Chicago Fire (1871) on innovation in non-wood (Panel C.10a) and wood (Panel C.10b) technologies. The dependent variable is the number of patents in either class. The unit of observation is a metropolitan area at a yearly frequency between 1853 and 1900. The Figures report event-study coefficients obtained using the synthetic difference-in-differences methods developed by Arkhangelsky et al. (2021) and obtained following the procedure described in Clarke et al. (2023). Standard errors are estimated using bootstrap sampling with 100 replications; gray bars report 95% confidence intervals. The covariates included in the synthetic difference-in-differences estimation are the same as in the synthetic control estimation, except for the pre-treatment outcome values. The black dashed line marks the year of the Great Chicago Fire (1871) and the Chicago Municipality Ordinance prohibiting wood construction (1874). Referenced on pages: 23, B10.

Figure C.11. Non-Wood Construction Manufacturing: Pre-Post Great Chicago Fire Synthetic Control Root Mean Squared Prediction Error Comparison

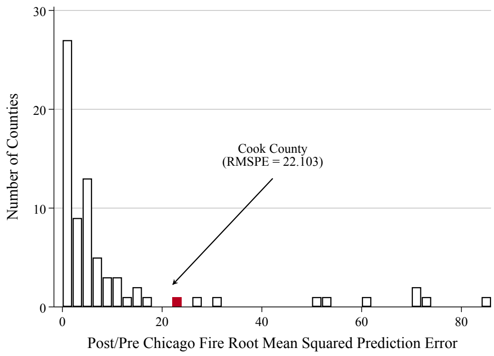
(a) Number of Establishments



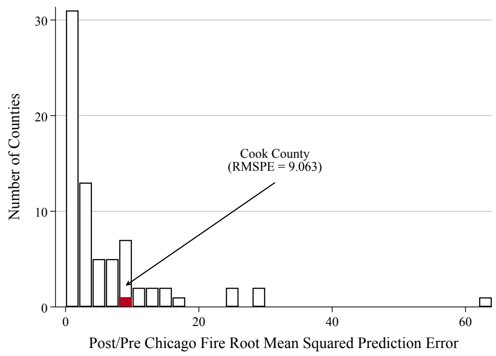
(b) Value of Production



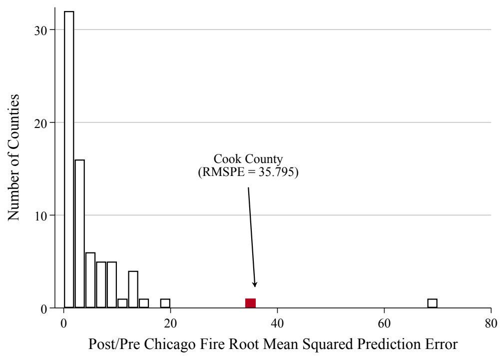
(c) Fixed Capital



(d) Cost of Materials



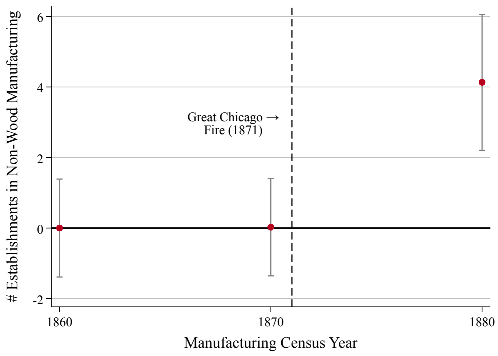
(e) Cost of Labor



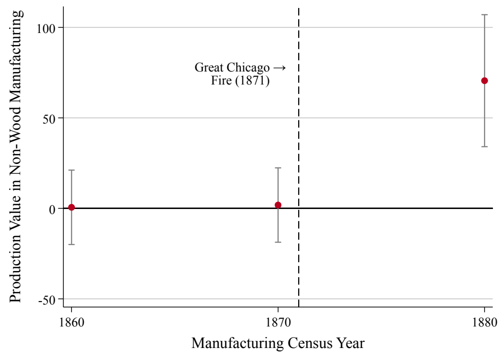
Notes. This Figure reports the impact of the Great Chicago (1871) on non-wood construction manufacturing activity as measured in the Census of Manufacturing. The unit of observation is a county at a decade frequency between 1860 and 1880. The dependent variable is: in Panel C.11a, the number of establishments; in Panel C.11b, production value; in Panel C.11c, fixed capital; in Panel C.11d, the cost of materials; in Panel C.11e, the cost of labor. Each figure reports the ratio between the post-Fire and pre-Fire mean squared prediction error across counties and highlights Cook County (IL) in red. The sample includes all counties with at least one metropolitan area. Referenced on page: 24.

Figure C.12. Synthetic Difference-in-Differences Estimates of the Effect of the Great Chicago Fire on Non-Wood Manufacturing

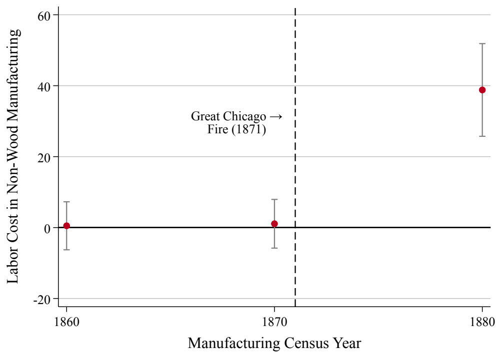
(a) Establishments



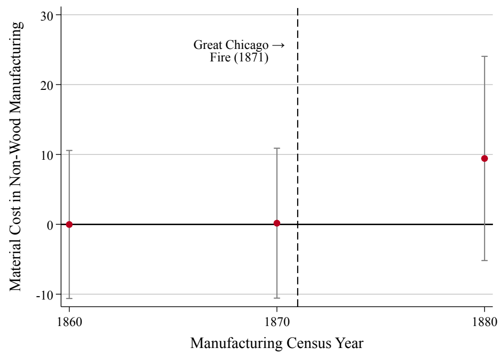
(b) Production Value



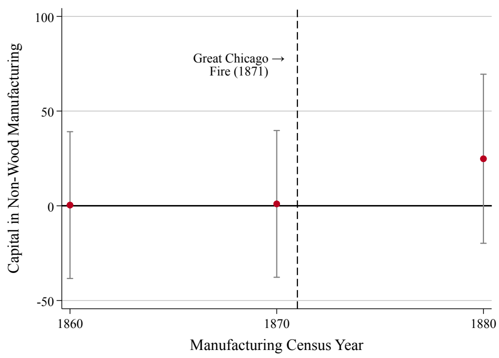
(c) Labor Costs



(d) Material Costs



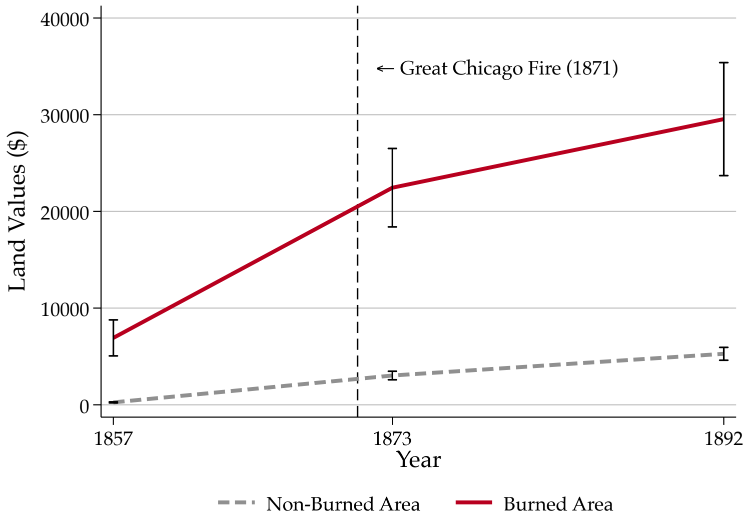
(e) Capital



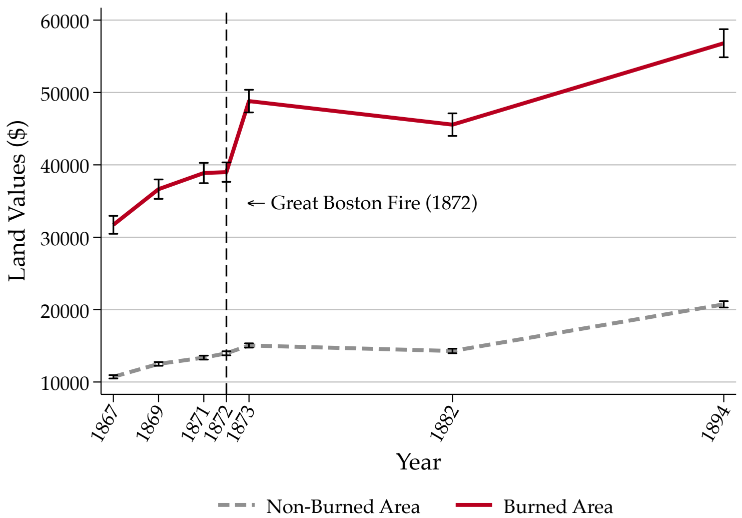
Notes. This Figure reports the impact of the Great Chicago Fire (1871) on non-wood manufacturing activity as measured in the Census of Manufacturing. The unit of observation is a county at a decade frequency between 1860 and 1880. The dependent variable is: in Panel C.12a, the number of establishments; in Panel C.12b, production value; in Panel C.12c, the cost of labor; in Panel C.12d, the cost of materials; in Panel C.12e, fixed capital. The Figures report event-study coefficients obtained using the synthetic difference-in-differences methods developed by Arkhangelsky et al. (2021) and obtained following the procedure described in Clarke et al. (2023). Standard errors are estimated using bootstrap sampling with 100 replications; gray bars report 95% confidence intervals. The sample includes all counties with at least one metropolitan area. Referenced on pages: 24, B10.

Figure C.13. Land Values in Chicago and Boston Before and After the Fires

(a) Chicago



(b) Boston



Notes. This Figure reports the average land values per square mile in Chicago (Panel C.13a and Boston (Panel C.13b) before and after the 1871 and 1872 fires, respectively. Data for Chicago are digitized from Hoyt (1933), while data for Boston are from Hornbeck and Keniston (2017). In each panel, the red line reports land values in areas exposed to the fires, and the gray lines report average land values for the remaining parts of the city. Bands report one-standard-deviation intervals around the mean. Referenced on page: 26.

D ADDITIONAL TABLES

Table D.1. List of Words Related to Construction, Wood-Related, and Non-Wood-Related Patents

Words related to...		
Construction	Wood-Related Construction	Non-wood Related Construction
(1)	(2)	(3)
Construction	Wood	Iron
Building	Timber	Brick
Edifice	Beam	Ston
Frame	Joist	Mortar
Roof	Mortise	Concrete
Wall	Tenon	Steel
Floor	Plank	Cast
Ceiling	Bracing	Lime
Joist	Joinery	Terracotta
Pillar	Lumber	Cement
Foundation	Plywood	Glass
Footing	Veneer	
Slab		
Stair		
Staircase		
Railing		
Balustrade		
Baluster		
Fence		
Gate		
Door		
Window		
Sill		
Lintel		
Arch		
Vault		
Dome		

Notes. This Table reports the keywords that we use to implement the dictionary-based approach to identify patents in construction (column 1), wood-related construction (column 2), and non-wood-related construction (column 3). We flag a patent as construction-related if the total number of mentions of words in column (1) is five or more. A wood-related (resp. non-wood-related) patent is one that further mentions at least one word in column (2) (resp 3). Referenced on pages: 9, A4.

Table D.2. List of Metropolitan Areas in the Sample

Metro Area (1)	City (2)	Metro Area (3)	City (4)	Metro Area (5)	City (6)
Albany	Albany	Albany	Bethlehem	Albany	Coeymans
Albany	Greenbush	Albany	Guilderland	Albany	New Baltimore
Albany	New Scotland	Albany	Niskayuna	Albany	Schodack
Allegheny	Allegheny	Allegheny	Bellevue	Allegheny	Chartiers
Allegheny	Dixmont	Allegheny	Green Tree	Allegheny	Kilbuck
Allegheny	Marshall	Allegheny	Mccandless	Allegheny	Mount Washington
Allegheny	Neville	Allegheny	Ohio	Allegheny	Peters
Allegheny	Pine	Allegheny	Richland	Allegheny	Ross
Allegheny	Scott	Allegheny	South Fayette	Allegheny	Stowe
Allegheny	Upper Saint Clair	Allegheny	Wexford		
Atlanta	Atlanta	Atlanta	Clayton	Atlanta	Decatur
Atlanta	Dekalb	Atlanta	Panthersville		
Augusta	Augusta	Augusta	Hamburg	Augusta	Richmond Factory Pond
Baltimore	Baltimore	Baltimore	Baltimore Zoo	Baltimore	Brooklandville
Baltimore	Brooklyn	Baltimore	Catonsville	Baltimore	Cockeysville
Baltimore	Lutherville	Baltimore	Saint Denis	Baltimore	Texas
Baltimore	Warren				
Boston	Boston	Boston	Braintree	Boston	Milton
Boston	Quincy	Boston	Randolph	Boston	Stoughton
Boston	Weymouth				
Brooklyn	Brooklyn	Brooklyn	College Point	Brooklyn	Columbusville
Brooklyn	Flatbush	Brooklyn	Flatlands	Brooklyn	Flushing
Brooklyn	Gravesend	Brooklyn	Jamaica	Brooklyn	New Lots
Brooklyn	New Utrecht	Brooklyn	Westchester	Brooklyn	Whitestone
Buffalo	Amherst	Buffalo	Buffalo	Buffalo	Cheektowaga
Buffalo	Eden	Buffalo	Grand Island	Buffalo	Hamburg
Buffalo	Tonawanda	Buffalo	West Seneca	Buffalo	Wheatfield
Buffalo	Williamsville				
Camden	Blackwood	Camden	Camden	Camden	Cinnaminson
Camden	Deptford	Camden	Gloucester	Camden	Gloucester City
Camden	Haddon	Camden	Haddonfield	Camden	Harrison
Camden	Mantua	Camden	Merchantville	Camden	Washington
Camden	Woodbury				
Charleston	Charleston	Charleston	James Island	Charleston	Johns Island
Charleston	Saint Andrews				
Charlestown	Charlestown	Charlestown	Chelsea	Charlestown	East Boston
Charlestown	Everett	Charlestown	Malden	Charlestown	Melrose

Charlestown	Wakefield				
Chatham	Chatham	Chatham	Whitmell		
Chicago	Calumet	Chicago	Chicago	Chicago	Cicero
Chicago	Evanston	Chicago	Hyde Park	Chicago	Jefferson
Chicago	Lake	Chicago	Lakeview	Chicago	Lyons
Cincinnati	Bromley	Cincinnati	Cincinnati	Cincinnati	Colerain
Cincinnati	Cumminsville	Cincinnati	Delhi	Cincinnati	Glendale
Cincinnati	Green	Cincinnati	Ludlow	Cincinnati	Millcreek
Cincinnati	Springdale	Cincinnati	Springfield	Cincinnati	Taylorsport
Cincinnati	West Covington				
Cleveland	Bedford	Cleveland	Brecksville	Cleveland	Brooklyn
Cleveland	Cleveland	Cleveland	East Cleveland	Cleveland	Independence
Cleveland	Newburg	Cleveland	Parma	Cleveland	Rockport
Cleveland	Royalton				
Columbus	Blendon	Columbus	Clinton	Columbus	Columbus
Columbus	Franklin	Columbus	Groveport	Columbus	Hamilton
Columbus	Harrison	Columbus	Jackson	Columbus	Madison
Columbus	Mifflin	Columbus	Norwich	Columbus	Orange
Columbus	Perry	Columbus	Scioto	Columbus	Sharon
Columbus	Truro	Columbus	Westerville		
Covington	Alexandria	Covington	Anderson	Covington	Bank Lick
Covington	Cold Spring	Covington	Columbia	Covington	Covington
Covington	Dayton	Covington	Florence	Covington	Independence
Covington	Johns Hill	Covington	Kenton	Covington	Newport
Covington	Pendleton	Covington	Scott	Covington	Visalia
Davenport	Andalusia	Davenport	Blackhawk	Davenport	Bowling
Davenport	Buffalo	Davenport	Butler	Davenport	Coal Valley
Davenport	Davenport	Davenport	Hampton	Davenport	Lincoln
Davenport	Moline	Davenport	Pleasant Valley	Davenport	Preemption
Davenport	Rock Island	Davenport	Rockingham	Davenport	Sheridan
Davenport	Winfield				
Dayton	Bath	Dayton	Beavercreek	Dayton	Bellbrook
Dayton	Bethel	Dayton	Brandt	Dayton	Butler
Dayton	Clear Creek	Dayton	Dayton	Dayton	Harrison
Dayton	Jefferson	Dayton	Mad River	Dayton	Madison
Dayton	Miami	Dayton	Miamisburg	Dayton	Monroe
Dayton	Randolph	Dayton	Sugarcreek	Dayton	Vandalia
Dayton	Washington	Dayton	Wayne	Dayton	West Charleston
Detroit	Dearborn	Detroit	Detroit	Detroit	Ecorse
Detroit	Fort Wayne	Detroit	Grosse Pointe	Detroit	Hamtramck
Detroit	Roseville	Detroit	Royal Oak	Detroit	Springwells
Detroit	Warren	Detroit	Wyandotte		
Edgefield	Edgefield	Edgefield	Fruit Hill	Edgefield	Johnston
Edgefield	Meeting Street				

Elizabeth	Clark	Elizabeth	Elizabeth	Elizabeth	Keyport
Elizabeth	Linden	Elizabeth	Livingston	Elizabeth	Millburn
Elizabeth	New Dorp	Elizabeth	New Springville	Elizabeth	Perth Amboy
Elizabeth	Port Richmond	Elizabeth	Rahway	Elizabeth	Richmond
Elizabeth	South Amboy	Elizabeth	South Orange	Elizabeth	Springfield
Elizabeth	Summit	Elizabeth	Tottenville	Elizabeth	Westfield
Elizabeth	Woodbridge				
Eufaula	Eufaula	Eufaula	Georgetown		
Evansville	Armstrong	Evansville	Campbell	Evansville	Center
Evansville	Evansville	Evansville	German	Evansville	Henderson
Evansville	Knight	Evansville	Ohio	Evansville	Perry
Evansville	Pigeon	Evansville	Scott	Evansville	Spottsville
Evansville	Union				
Fall River	Berkley	Fall River	Bristol	Fall River	Dighton
Fall River	Fall River	Fall River	Little Compton	Fall River	Middletown
Fall River	Portsmouth	Fall River	Rehoboth	Fall River	Somerset
Fall River	Swansea	Fall River	Taunton	Fall River	Tiverton
Fall River	Warren	Fall River	Westport		
Harrisburg	Dauphin	Harrisburg	Duncannon	Harrisburg	East Pennsboro
Harrisburg	Fairview	Harrisburg	Goldsboro	Harrisburg	Halifax
Harrisburg	Hampden	Harrisburg	Harrisburg	Harrisburg	Highspire
Harrisburg	Hummelstown	Harrisburg	Lewisberry	Harrisburg	Lower Allen
Harrisburg	Lower Paxton	Harrisburg	Lower Swatara	Harrisburg	Marysville
Harrisburg	Mechanicsburg	Harrisburg	Middle Paxton	Harrisburg	Middletown
Harrisburg	Monaghan	Harrisburg	New Buffalo	Harrisburg	New Cumberland
Harrisburg	Newberry	Harrisburg	Penn	Harrisburg	Reed
Harrisburg	Rockville	Harrisburg	Rye	Harrisburg	Silver Spring
Harrisburg	South Hanover	Harrisburg	Susquehanna	Harrisburg	Upper Allen
Harrisburg	Warrington	Harrisburg	Watts	Harrisburg	West Hanover
Hartford	Avon	Hartford	Berlin	Hartford	Bloomfield
Hartford	Cromwell	Hartford	East Granby	Hartford	East Hartford
Hartford	East Windsor Hill	Hartford	Farmington	Hartford	Glastonbury
Hartford	Granby	Hartford	Hartford	Hartford	Manchester
Hartford	Middletown	Hartford	New Britain	Hartford	Portland
Hartford	Rocky Hill	Hartford	Simsbury	Hartford	South Windsor
Hartford	West Hartford	Hartford	Wethersfield	Hartford	Windsor
Hoboken	Hackensack	Hoboken	Hoboken	Hoboken	North Bergen
Hoboken	Weehawken				
Indianapolis	Allisonville	Indianapolis	Carmel	Indianapolis	Center
Indianapolis	Clay	Indianapolis	Decatur	Indianapolis	Eagle
Indianapolis	Franklin	Indianapolis	Indianapolis	Indianapolis	Lawrence
Indianapolis	Millersville	Indianapolis	Perry	Indianapolis	Pike

Indianapolis	Pleasant	Indianapolis	Warren	Indianapolis	Washington
Indianapolis	Wayne	Indianapolis	White River		
Jersey City	Bayonne	Jersey City	Greenville	Jersey City	Jersey City
Jersey City	New Brighton	Jersey City	Rutherford Park	Jersey City	Tompkinsville
Kansas City	Gallatin	Kansas City	Independence	Kansas City	Kansas City
Kansas City	Pettis	Kansas City	Quindaro	Kansas City	Shawnee
Kansas City	Westport				
Knoxville	Beaver Ridge	Knoxville	Knoxville	Knoxville	Louisville
Knoxville	Maryville	Knoxville	Rockford		
Lancaster	Clay	Lancaster	Conestoga	Lancaster	East Hempfield
Lancaster	East Lampeter	Lancaster	Eden	Lancaster	Elizabeth
Lancaster	Ephrata	Lancaster	Lancaster	Lancaster	Leacock
Lancaster	Manheim	Lancaster	Manor	Lancaster	Martic
Lancaster	Millersville	Lancaster	New Providence	Lancaster	Paradise
Lancaster	Penn	Lancaster	Pequea	Lancaster	Rapho
Lancaster	Strasburg	Lancaster	Upper Leacock	Lancaster	Warwick
Lancaster	West Earl	Lancaster	West Hempfield	Lancaster	West Lampeter
Lawrence	Andover	Lawrence	Atkinson	Lawrence	Bradford
Lawrence	Danville	Lawrence	Georgetown	Lawrence	Groveland
Lawrence	Hampstead	Lawrence	Haverhill	Lawrence	Lawrence
Lawrence	Methuen	Lawrence	Newton	Lawrence	North Andover
Lawrence	North Reading	Lawrence	Plaistow	Lawrence	Salem
Lawrence	Sandown	Lawrence	Wilmington		
Louisville	Carr	Louisville	Charlestown	Louisville	Franklin
Louisville	Harrods Creek	Louisville	Jeffersonville	Louisville	Lafayette
Louisville	Louisville	Louisville	New Albany	Louisville	Newburg
Louisville	Saint Matthews	Louisville	Shively	Louisville	Silver Creek
Louisville	Springdale	Louisville	Union	Louisville	Utica
Louisville	nan				
Lowell	Acton	Lowell	Bedford	Lowell	Billerica
Lowell	Carlisle	Lowell	Chelmsford	Lowell	Concord
Lowell	Dracut	Lowell	Dunstable	Lowell	Hudson
Lowell	Lincoln	Lowell	Lowell	Lowell	Pelham
Lowell	Tewksbury	Lowell	Tyngsborough	Lowell	Westford
Lowell	Windham				
Lynchburg	Amherst	Lynchburg	Brookville	Lynchburg	Coolwell
Lynchburg	Elon	Lynchburg	Forest	Lynchburg	Lynchburg
Lynchburg	New London				
Lynn	Hingham	Lynn	Hull	Lynn	Lynn
Lynn	Lynnfield	Lynn	Middleton	Lynn	Nahant
Lynn	Saugus	Lynn	Swampscott	Lynn	Winthrop
Macon	Clinton	Macon	Macon		
Manchester	Allenstown	Manchester	Auburn	Manchester	Bedford
Manchester	Bow	Manchester	Concord	Manchester	Derry

Manchester	Dunbarton	Manchester	Goffstown	Manchester	Hooksett
Manchester	Litchfield	Manchester	Londonderry	Manchester	Manchester
Manchester	Merrimack	Manchester	Nashua	Manchester	Pembroke
Marion	Marion	Marion	Perry		
Memphis	Cuba	Memphis	Desoto	Memphis	Hopefield
Memphis	Horn Lake	Memphis	Memphis		
Milwaukee	Caledonia	Milwaukee	Franklin	Milwaukee	Granville
Milwaukee	Greenfield	Milwaukee	Lake	Milwaukee	Mequon
Milwaukee	Milwaukee	Milwaukee	Oak Creek	Milwaukee	Wauwatosa
Montgomery	Autauga	Montgomery	Elmore	Montgomery	Montgomery
Montgomery	Prattville	Montgomery	Wetumpka		
Nashville	Brentwood	Nashville	Goodlettsville	Nashville	Madison
Nashville	Nashville				
New Bedford	Acushnet	New Bedford	Dartmouth	New Bedford	Fairhaven
New Bedford	Freetown	New Bedford	Gosnold	New Bedford	Lakeville
New Bedford	Marion	New Bedford	Mattapoisett	New Bedford	New Bedford
New Bedford	Rochester				
New Haven	Ansonia	New Haven	Bethany	New Haven	Branford
New Haven	Cheshire	New Haven	Derby	New Haven	East Haven
New Haven	Hamden	New Haven	Milford	New Haven	New Haven
New Haven	North Branford	New Haven	North Haven	New Haven	Orange
New Haven	Prospect	New Haven	Seymour	New Haven	Wallingford
New Haven	West Haven	New Haven	Woodbridge		
New Orleans	Carrollton	New Orleans	Kenner	New Orleans	Metairie
New Orleans	New Orleans	New Orleans	Shrewsbury		
New York	Astoria	New York	Belmont	New York	Fordham
New York	Long Island City	New York	Morrisania	New York	New York
New York	Tremont	New York	West Farms		
Newark	Belleville	Newark	Bloomfield	Newark	East Orange
Newark	Harrison	Newark	Kearny	Newark	Montclair
Newark	Newark	Newark	Orange	Newark	West Orange
Newburgh	Cold Spring	Newburgh	Cornwall	Newburgh	Fishkill
Newburgh	Fishkill Landing	Newburgh	Glenham	Newburgh	New Windsor
Newburgh	Newburgh	Newburgh	Philipstown	Newburgh	Plattekill
Newburgh	West Point				
Norfolk	Deep Creek	Norfolk	Fort Monroe	Norfolk	Great Bridge
Norfolk	Hampton	Norfolk	Kempsville	Norfolk	Norfolk
Norfolk	Old Point Comfort Marina	Norfolk	Portsmouth	Norfolk	Tanner Creek
Norfolk	Western Branch Park	Norfolk	Wythe		
North Providence	Bellingham	North Providence	Blackstone	North Providence	Cumberland

North Providence	Johnston	North Providence	North Providence	North Providence	Scituate
North Providence	Smithfield	North Providence	Woonsocket		
Old Cambridge	Arlington	Old Cambridge	Belmont	Old Cambridge	Brighton
Old Cambridge	Brookline	Old Cambridge	Burlington	Old Cambridge	Canton
Old Cambridge	Dedham	Old Cambridge	Hyde Park	Old Cambridge	Lexington
Old Cambridge	Medford	Old Cambridge	Needham	Old Cambridge	Newton
Old Cambridge	Old Cambridge	Old Cambridge	Reading	Old Cambridge	Somerville
Old Cambridge	Stoneham	Old Cambridge	Waltham	Old Cambridge	Watertown
Old Cambridge	West Roxbury	Old Cambridge	Winchester	Old Cambridge	Woburn
Oswego	Fulton	Oswego	Granby	Oswego	Hannibal
Oswego	Ira	Oswego	Martville	Oswego	Oswego
Oswego	Scriba	Oswego	Sterling	Oswego	Sterling Valley
Oswego	Volney				
Paterson	Caldwell	Paterson	Clinton	Paterson	Hohokus
Paterson	Little Falls	Paterson	Lodi	Paterson	Passaic
Paterson	Paterson	Paterson	Pequannock	Paterson	Pompton
Paterson	Ramapo	Paterson	Ramsey	Paterson	Saddle River
Paterson	Washington	Paterson	Wayne		
Peoria	Akron	Peoria	Cincinnati	Peoria	Dillon
Peoria	Elm Grove	Peoria	Fondulac	Peoria	Groveland
Peoria	Hallock	Peoria	Hollis	Peoria	Kickapoo
Peoria	Limestone	Peoria	Medina	Peoria	Morton
Peoria	Pekin	Peoria	Peoria	Peoria	Radnor
Peoria	Richwoods	Peoria	Spring Bay	Peoria	Tremont
Peoria	Worth				
Philadelphia	Abington	Philadelphia	Cheltenham	Philadelphia	Conshohocken
Philadelphia	Darby	Philadelphia	Greenwich	Philadelphia	Haverford
Philadelphia	Horsham	Philadelphia	Lower Merion	Philadelphia	Oreland
Philadelphia	Philadelphia	Philadelphia	Ridley	Philadelphia	Springfield
Philadelphia	Springfield	Philadelphia	Tinicum	Philadelphia	Upper Darby
Philadelphia	Upper Dublin	Philadelphia	Whitemarsh		
Pittsburgh	Allentown	Pittsburgh	Baldwin	Pittsburgh	Braddock
Pittsburgh	East Pittsburgh	Pittsburgh	Elizabeth	Pittsburgh	Etna
Pittsburgh	Hampton	Pittsburgh	Indiana	Pittsburgh	Lincoln
Pittsburgh	Mckeesport	Pittsburgh	Mifflin	Pittsburgh	Millvale
Pittsburgh	Pittsburgh	Pittsburgh	Reserve	Pittsburgh	Shaler
Pittsburgh	Sharpsburg	Pittsburgh	Snowden	Pittsburgh	Surgeon Hall
Pittsburgh	Union	Pittsburgh	West Elizabeth	Pittsburgh	Wilkins
Portland	Cape Elizabeth	Portland	Cumberland	Portland	Falmouth
Portland	Gray	Portland	North Yarmouth	Portland	Portland
Portland	Scarborough	Portland	Westbrook	Portland	Yarmouth
Poughkeepsie	Clinton	Poughkeepsie	East Fishkill	Poughkeepsie	Esopus

Poughkeepsie	Fishkill Plains	Poughkeepsie	Hyde Park	Poughkeepsie	La Grange
Poughkeepsie	Lloyd	Poughkeepsie	Marlborough	Poughkeepsie	New Paltz
Poughkeepsie	Pleasant Valley	Poughkeepsie	Port Ewen	Poughkeepsie	Poughkeepsie
Poughkeepsie	Rhinebeck	Poughkeepsie	Rondout	Poughkeepsie	Sleightsburg
Poughkeepsie	Wappingers Falls				
Providence	Attleboro	Providence	Barrington	Providence	Cranston
Providence	East Greenwich	Providence	East Providence	Providence	Pawtucket
Providence	Providence	Providence	Seekonk	Providence	Warwick
Quincy	Burton	Quincy	Ellington	Quincy	Fabius
Quincy	Fall Creek	Quincy	La Grange	Quincy	Liberty
Quincy	Melrose	Quincy	Mendon	Quincy	Miller
Quincy	Palmyra	Quincy	Quincy	Quincy	South River
Quincy	Ursa				
Reading	Adamstown	Reading	Alsace	Reading	Bern
Reading	Brecknock	Reading	Brecknock	Reading	Caernarvon
Reading	Caernarvon	Reading	Centre	Reading	Cumru
Reading	Exeter	Reading	Hamburg	Reading	Maiden Creek
Reading	Muhlenberg	Reading	Oley	Reading	Ontelaunee
Reading	Penn	Reading	Perry	Reading	Reading
Reading	Richmond	Reading	Robeson	Reading	Ruscombmanor
Reading	South Heidelberg	Reading	Spring	Reading	Union
Richmond	Bermuda	Richmond	Brookland	Richmond	Chester
Richmond	Chesterfield	Richmond	Clover Hill	Richmond	Dale
Richmond	Fairfield	Richmond	Manchester	Richmond	Richmond
Richmond	Tuckahoe	Richmond	Varina		
Rochester	Brighton	Rochester	Chili	Rochester	Gates
Rochester	Greece	Rochester	Henrietta	Rochester	Honeoye Falls
Rochester	Irondequoit	Rochester	Mendon	Rochester	Penfield
Rochester	Pittsford	Rochester	Rochester	Rochester	Rush
Rochester	Scottsville				
Sacramento	Brighton	Sacramento	Franklin	Sacramento	Fremont
Sacramento	Sacramento				
Saint Louis	Brooklyn	Saint Louis	Carondelet	Saint Louis	Caseyville
Saint Louis	Columbia	Saint Louis	East Saint Louis	Saint Louis	French
Saint Louis	Gartside	Saint Louis	Madison	Saint Louis	Millstadt
Saint Louis	Saint Ferdinand	Saint Louis	Saint Louis	Saint Louis	Venice
Saint Paul	Centerville	Saint Paul	Cottage Grove	Saint Paul	Eagan
Saint Paul	Inver Grove	Saint Paul	Mendota	Saint Paul	Minneapolis
Saint Paul	Mounds View	Saint Paul	Newport	Saint Paul	Oakdale
Saint Paul	Rosemount	Saint Paul	Roseville	Saint Paul	Saint Anthony
Saint Paul	Saint Paul	Saint Paul	West Saint Paul	Saint Paul	White Bear
Saint Paul	Woodbury				

Salem	Beverly	Salem	Boxford	Salem	Danvers
Salem	Essex	Salem	Hamilton	Salem	Ipswich
Salem	Manchester	Salem	Marblehead	Salem	Peabody
Salem	Rowley	Salem	Salem	Salem	Topsfield
Salem	Wenham				
San Francisco	Alameda	San Francisco	Oakland	San Francisco	San Bruno
San Francisco	San Francisco	San Francisco	San Pablo	San Francisco	Sausalito
Savannah	Hardeeville	Savannah	Savannah	Savannah	Thunderbolt
Savannah	White Bluff				
Scranton	Archbald	Scranton	Benton	Scranton	Blakely
Scranton	Dunmore	Scranton	Exeter	Scranton	Greenfield
Scranton	Hyde Park	Scranton	Jefferson	Scranton	Jenkins
Scranton	North Abington	Scranton	Old Forge	Scranton	Olyphant
Scranton	Pittston	Scranton	Ransom	Scranton	Scott
Scranton	Scranton	Scranton	South Abington	Scranton	Spring Brook
Scranton	Waverly	Scranton	West Pittston		
Shreveport	Bossier City	Shreveport	Shreveport		
Springfield	Agawam	Springfield	Chicopee	Springfield	Easthampton
Springfield	Enfield	Springfield	Granby	Springfield	Hadley
Springfield	Holyoke	Springfield	Longmeadow	Springfield	Ludlow
Springfield	Northampton	Springfield	Somers	Springfield	South Hadley
Springfield	Southampton	Springfield	Springfield	Springfield	Suffield
Springfield	West Springfield	Springfield	Westfield	Springfield	Wilbraham
Springfield	Windsor Locks				
Syracuse	Amboy	Syracuse	Belgium	Syracuse	Belleisle
Syracuse	Brewerton	Syracuse	Camillus	Syracuse	Cardiff
Syracuse	Caughdenoy	Syracuse	Central Square	Syracuse	Cicero
Syracuse	Clay	Syracuse	Dewitt	Syracuse	Euclid
Syracuse	Fayetteville	Syracuse	Geddes	Syracuse	Jamesville
Syracuse	Lafayette	Syracuse	Liverpool	Syracuse	Manlius
Syracuse	Navarino	Syracuse	Onondaga	Syracuse	Otisco
Syracuse	Pompey	Syracuse	Salina	Syracuse	South Onondaga
Syracuse	Syracuse	Syracuse	Threerivers	Syracuse	Van Buren
Toledo	Bedford	Toledo	Erie	Toledo	Lake
Toledo	Maumee	Toledo	Oregon	Toledo	Perrysburg
Toledo	Sylvania	Toledo	Toledo	Toledo	Troy
Toledo	Washington	Toledo	Webster	Toledo	Whiteford
Trenton	Bordentown	Trenton	Bristol	Trenton	Burlington
Trenton	Chesterfield	Trenton	Ewing	Trenton	Falls
Trenton	Hamilton	Trenton	Hamilton Square	Trenton	Hopewell
Trenton	Lawrence	Trenton	Lower Makefield	Trenton	Mansfield
Trenton	Middletown	Trenton	Morrisville	Trenton	Mount Holly
Trenton	Princeton	Trenton	Springfield	Trenton	Trenton

Trenton	West Windsor	Trenton	Westampton		
Troy	Brunswick	Troy	Clifton Park	Troy	Cohoes
Troy	Green Island	Troy	Halfmoon	Troy	Lansingburgh
Troy	Mechanicville	Troy	Nassau	Troy	North Greenbush
Troy	Poestenkill	Troy	Sand Lake	Troy	Schaghticoke
Troy	Stillwater	Troy	Troy	Troy	Waterford
Troy	Watervliet	Troy	West Sand Lake		
Utica	Bridgewater	Utica	Cassville	Utica	Clark Mills
Utica	Clayville	Utica	Clinton	Utica	Deerfield
Utica	Floyd	Utica	Frankfort	Utica	Gravesville
Utica	Holland Patent	Utica	Kirkland	Utica	Litchfield
Utica	Marcy	Utica	Marshall	Utica	New Hartford
Utica	New York Mills	Utica	Oriskany	Utica	Paris
Utica	Prospect	Utica	Remsen	Utica	Sauquoit
Utica	Schuyler	Utica	South Trenton	Utica	Steuben

Notes. This Table reports the list of metropolitan areas (columns 1, 3, and 5) and all the cities below 20,000 inhabitants that are part of them. There is a total of 84 metropolitan areas comprising 1,048 smaller towns. Referenced on page: 12.

Table D.3. Synthetic Control Weights for Boston and Chicago

Metropolitan Area (1)	Synthetic Chicago (2)	Synthetic Boston (3)
New Orleans	0.0	0.024
San Francisco	0.093	0.413
Cincinnati	0.0	0.112
Kansas City	0.02	0.0
New York	0.254	0.166
Jersey City	0.304	0.0
Springfield	0.0	0.008
Charlestown	0.328	0.277

Notes. This Table presents the weights assigned to the metropolitan areas listed in column (1) to construct the synthetic Chicago (column 2) and Boston (column 3) control units. Weights are selected following a data-driven optimization algorithm that minimizes the distance between the treated unit and the synthetic control in terms of a set of balancing variables. The balancing variables are population, the share of men, the share of US-born, the share of literate, employment shares by occupation, and employment shares by industry. Shares are expressed in terms of population. All balancing variables are constructed from the 1870 population census. The weights are obtained by applying the synthetic control approach on construction patenting. Referenced on page: 14.

Table D.4. Comparison between Boston, the Other Metropolitan Areas, and Synthetic Boston

	Boston	All Other Cities			Synthetic Boston		
	Mean	Mean	Difference		Mean	Difference	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Demographics							
Literacy Rate (%)	70.645	60.780	9.865***	(1.685)	68.193	2.452	(1.739)
Imputed Income per Capita	893.628	707.792	185.837***	(14.929)	895.798	-2.170	(52.655)
Share of Whites (%)	98.886	85.715	13.171***	(2.443)	94.992	3.894	(2.684)
Share of Blacks (%)	1.110	14.023	-12.913***	(2.452)	1.697	-0.587	(1.079)
Share of Natives (%)	66.585	78.113	-11.528***	(1.475)	61.703	4.883	(7.343)
Panel B. Employment Share (%) by Occupation							
Liberal Profession	1.412	1.029	0.383***	(0.032)	1.632	-0.220	(0.254)
Farmer	0.620	8.410	-7.790***	(0.970)	1.008	-0.388	(0.372)
Manager	2.768	1.977	0.791***	(0.082)	3.382	-0.615	(0.404)
Clerical Worker	1.373	0.650	0.723***	(0.037)	1.202	0.171***	(0.057)
Sales	3.063	1.363	1.700***	(0.071)	2.862	0.201	(0.284)
Skilled Manufacture	8.168	5.640	2.528***	(0.213)	7.884	0.284	(0.293)
Low-Skill Manufacture	9.904	7.200	2.704***	(0.583)	8.651	1.253***	(0.242)
Service	6.692	4.389	2.303***	(0.218)	6.360	0.331	(1.134)
Panel C. Employment Share (%) by Industry							
Laborer	5.783	4.669	1.114***	(0.211)	6.055	-0.272	(0.712)
Agriculture	0.843	8.667	-7.824***	(0.965)	1.308	-0.465	(0.496)
Chemistry	0.092	0.078	0.014	(0.014)	0.187	-0.095	(0.124)
Construction	3.336	2.379	0.957***	(0.090)	3.201	0.136	(0.183)
Liberal Professions	10.275	6.597	3.678***	(0.245)	10.078	0.198	(1.578)
Metallurgy	0.730	0.766	-0.037	(0.066)	0.686	0.043	(0.051)
Public Administration	0.500	0.297	0.204***	(0.027)	0.673	-0.173	(0.144)
Textiles	0.918	1.949	-1.031**	(0.460)	0.641	0.277	(0.240)
Trade	7.158	3.606	3.552***	(0.171)	6.789	0.369	(0.280)
Transports	3.048	1.991	1.057***	(0.099)	3.266	-0.218	(0.407)
Utilities	8.025	4.962	3.063***	(0.230)	7.679	0.346	(1.362)
Residual Industries	3.946	2.986	0.959***	(0.149)	4.145	-0.200	(0.172)
Engineering	0.528	0.386	0.142***	(0.032)	0.606	-0.078	(0.088)

Notes. This Table compares the values of the balancing variables included in the synthetic control design in Boston and in the other metropolitan areas in the sample. Column (1) reports the average value of the various variables for Boston; columns (2) and (5) report the average across all control cities and in synthetic Boston, respectively. The weights used to compute the co-variates in the synthetic control are obtained by applying the synthetic control approach on construction patenting. In columns (3–4) (resp. 6–7), we report the difference between Boston and all other cities (resp. synthetic Boston). Robust standard errors are displayed in parentheses. All data are computed from the 1870 population census and expressed in population percentage. Referenced on page: 15.

*: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$

Table D.5. Synthetic Control Estimates of the County-Level Impact of the Great Chicago Fire on per Capita Manufacturing Activity

	Dependent Variable (Treated County - Synthetic Treated County)				
	(1) # Establishments	(2) Value of Production	(3) Fixed Capital	(4) Cost of Materials	(5) Cost of Labor
Panel A. Construction Manufacturing					
1860 (Pre-Fire)	0.000	3.247	-0.741	-1.726	-0.793
1870 (Pre-Fire)	-0.001	-0.070	-3.776	-2.314	-0.355
1880 (Post-Fire)	0.065	3320.363	749.794	1836.556	904.345
Mean Dep. Var. (Before 1870)	0.118	2596.728	1314.480	870.335	930.620
Panel B. Non-Wood Manufacturing					
1860 (Pre-Fire)	0.000	0.222	6.317	0.002	-0.118
1870 (Pre-Fire)	0.000	-0.727	6.214	-0.272	0.599
1880 (Post-Fire)	0.008	138.772	-23.062	11.504	120.543
Mean Dep. Var. (Before 1870)	0.015	425.407	255.699	74.280	219.966
Number of Counties	76	76	76	76	76
Number of Observations	228	228	228	228	228

Notes. This Table reports the impact of the Great Chicago (1871) Fire on construction and non-wood-related manufacturing activity as measured in the Census of Manufacturing. The unit of observation is a county at a decade frequency between 1860 and 1880. The dependent variable is: in column (1), the number of establishments; in column (2), production value; in column (3), fixed capital; in column (4), the cost of materials; in column (5), the cost of labor. Each dependent variable is normalized by the county's total employment, as measured in the population census, and expressed in percentage terms. Each column reports the difference between the observed outcome in Cook County and a synthetic control constructed using the baseline balancing variables and pre-treatment outcome values. The sample includes all counties with at least one metropolitan area. The Table reports separately the effects on overall construction manufacturing (Panel A) and non-wood construction manufacturing (Panel B). Referenced on pages: 19, 20, 24.

Table D.6. Synthetic Control Estimates of the County-Level Estimates of the Impact of the Great Chicago Fire on Wood Manufacturing

	Dependent Variable (Treated County - Synthetic Treated County)				
	(1) # Estab-lishments	(2) Value of Production	(3) Fixed Capital	(4) Cost of Materials	(5) Cost of Labor
1860 (Pre-Fire)	0.000	-38.942	-2.560	-31.377	-1.253
1870 (Pre-Fire)	1.300	313.823	96.100	239.464	25.101
1880 (Post-Fire)	0.400	360.009	9.851	317.960	18.798
Mean Dep. Var. (Before 1870)	2.000	358.670	104.075	269.963	40.599
Number of Counties	76	76	76	76	76
Number of Observations	228	228	228	228	228

Notes. This Table reports the impact of the Great Chicago (1871) Fire on wood-related manufacturing activity as measured in the Census of Manufacturing. The unit of observation is a county at a decade frequency between 1860 and 1880. The dependent variable is: in column (1), the number of establishments; in column (2), production value; in column (3), fixed capital; in column (4), the cost of materials; in column (5), the cost of labor. Each column reports the difference between the observed outcome in Cook County and a synthetic control constructed using the baseline balancing variables and pre-treatment outcome values. The sample includes all counties with at least one metropolitan area. Referenced on page: 24.

Table D.7. Synthetic Control Estimates of the County-Level Estimates of the Impact of the Great Boston Fire on Wood Manufacturing

	Dependent Variable (Treated County - Synthetic Treated County)				
	(1) # Establishments	(2) Value of Production	(3) Fixed Capital	(4) Cost of Materials	(5) Cost of Labor
Panel A. Construction Manufacturing					
1860 (Pre-Fire)	0.000	17.223	-0.012	12.848	-0.002
1870 (Pre-Fire)	-0.001	1.153	-0.041	0.799	-0.035
1880 (Post-Fire)	-0.310	-174.464	-5.451	-154.345	-11.866
Mean Dep. Var. (Before 1870)	1.350	221.230	57.305	165.920	28.675
Number of Counties	76	76	76	76	76
Number of Observations	228	228	228	228	228

Notes. This Table reports the impact of the Great Boston (1872) Fire on wood-related manufacturing activity as measured in the Census of Manufacturing. The unit of observation is a county at a decade frequency between 1860 and 1880. The dependent variable is: in column (1), the number of establishments; in column (2), production value; in column (3), fixed capital; in column (4), the cost of materials; in column (5), the cost of labor. Each column reports the difference between the observed outcome in Suffolk County and a synthetic control constructed using the baseline balancing variables and pre-treatment outcome values. The sample includes all counties with at least one metropolitan area. Referenced on page: 27.

APPENDIX REFERENCES

Andrews, Michael J, "Historical Patent Data: A Practitioner's Guide," *Journal of Economics & Management Strategy*, 2021, *30* (2), 368–397.

Arkhangelsky, Dmitry, Susan Athey, David A Hirshberg, Guido W Imbens, and Stefan Wager, "Synthetic Difference-in-Differences," *American Economic Review*, 2021, *111* (12), 4088–4118.

Berkes, Enrico, "Comprehensive Universe of US Patents (CUSP): Data and Facts," *Working Paper*, 2018.

Clarke, Damian, Daniel Pailañir, Susan Athey, and Guido W Imbens, "Synthetic Difference-in-Differences Estimation," *IZA Discussion Paper*, 2023, (15907).

Coluccia, Davide M and Gaia Dossi, "Return Innovation: The Knowledge Spillovers of the British Migration to the United States, 1870-1940," *Working Paper*, 2025.

Hornbeck, Richard and Martin Rotemberg, "Growth Off the Rails: Aggregate Productivity Growth in Distorted Economies," *Journal of Political Economy*, 2024, *132* (11), 3547–3602.

Hoyt, Homer, *One Hundred Years of Land Values in Chicago: The Relationship of the Growth of Chicago to the Rise in its Land Values, 1830–1933*, Chicago (IL): University of Chicago Press, 1933.

Kelly, Bryan, Dimitris Papanikolaou, Amit Seru, and Matt Taddy, "Measuring Technological Innovation Over the Long Run," *American Economic Review: Insights*, 2021, *3* (3), 303–320.

Moser, Petra and Shmuel San, "Immigration, Science, and Invention: Lessons from the Quota Acts," *Working Paper*, 2020.

Petralia, S, PA Balland, and DL Rigby, "Data Descriptor: Unveiling the Geography of Historical Patents in the United States from 1836 to 1975," *Nature: Scientific Data*, 2016, *3*, 1–14.

Sarada, Sarada, Michael J Andrews, and Nicolas L Ziebarth, "Changes in the Demographics of American Inventors, 1870–1940," *Explorations in Economic History*, 2019, *74*, 101275.