The Erie Canal and the Economic Transformation of Nineteenth-Century New York State

Youwei Xing*

October 12, 2023

Abstract

This paper presents the first causal estimates of the impact of canals on sectoral transition and long-term growth. My context is the Erie Canal in nineteenth-century New York State. Using a difference-in-differences approach, I compare the outcomes in towns located close to the canal relative to those that were not close to the canal, before and after its opening. In an alternative identification strategy, I construct a least-cost path connecting targeted towns as an instrument for the actual canal route. I find that towns located close the canal experienced faster population growth, lower employment shares in agriculture, higher employment shares in commerce, and faster employment growth in manufacturing. I argue that the mechanism underlying sectoral transition and long-term growth was the development of small-scale manufacturing that processed agricultural products and the expansion of commercial activities brought about by lower transportation costs and easier access to distant markets. My findings also indicate the extent of the canal's impact: the effects disappear in areas that were located more than 15 to 20 km from the canal.

*John E. Walker Department of Economics, Clemson University, Clemson, 29634, South Carolina, USA; email: youweix@g.clemson.edu. I am grateful for comments and guidance from my advisor, Howard Bodenhorn. I am grateful for comments from Jonathan Leganza, William Dougan, Ilya Polovnikov, Michael Makowsky, and Robert Fleck. I thank participants at Clemson Public Economics Workshop, Clemson Labor Economics Workshop, Clemson Job Market Workshop, and the 2023 World Cliometrics Conference for their valuable comments.

1 Introduction

The canal-building boom predated the railroad boom by a generation or more. However, relative to railroads, canals have not received much academic attention. Fogel (1964) shows that railroads provided a marginal cost savings relative to canals. Yet, the reduction in transportation costs provided by canals, relative to wagon haulage was more substantial than the subsequent reduction associated with the railroads (Goodrich et al., 1961).

Previous studies look at the effects of canals on market integration (Albion, 1939), government promotion of canals (Goodrich, 1960), public financing of canals (Wallis, 2003), and the association of canals with inventive activity (Sokoloff, 1988). However, the causal effects of canal operation on sectoral transition in employment and long-term growth have not been studied in depth, and identifying the causal effects of canals on economic change is challenging. In particular, there is a potential selection problem because canals may have been constructed in populated areas with active commercial or manufacturing activities.

This paper addresses the potential selection problem to provide a clearer picture of how canals impacted sectoral transition in employment and long-term growth. My context is the Erie Canal in nineteenth-century New York State. For identification, I use the rollout of the canal as a quasi-experiment. This setting offers a way to identify the impacts of the canal. The design and construction of the Erie Canal was led by the state. Unlike private transport entities, the state had a heightened capacity to navigate regional disputes, reconcile political rivalries, and pursue larger-scale construction. The goal was to connect several targeted towns from the West to the East. Consequently, many non-targeted towns incidentally gained access to the canal, which alleviates some endogeneity concerns.

I begin by using a standard difference-in-differences (DiD) design. I estimate changes in employment in three broad sectors, as well as in the sectoral transition in employment from 1820 to 1840 among towns proximate to the Erie Canal relative to those that were not. The DiD estimates reveal lower employment shares in agriculture, higher employment shares in commerce, and faster manufacturing employment growth in areas neighboring the canal.

Quantitatively, towns that were close to the canal had 6.9% lower employment shares in agriculture, 2.3% greater employment shares in commerce, and 41.1% faster manufacturing employment growth than observed in towns that were not close to the canal. Notably, people did not stop being or becoming farmers. The lower employment share in agriculture was driven not by decreasing employment in agriculture but rather by faster employment growth in manufacturing and commerce. Subsequently, I estimate the impact of the Erie Canal on long-term growth using a longer time horizon population data between 1790 and 1870. The DiD estimates show areas located close to the canal witnessed 38.6% faster population growth. My findings also indicate the extent of canal's impact. The long-term growth and sectoral transition effects disappear in areas that were located more than 15 km and 20 km from the canal respectively.

The estimated sectoral transition and long-term effects are robust to multiple alternative robustness checks and an alternative identification strategy. First, I exclude towns immediately adjoining the treated areas from the control group to address potential spillover effects, where the DiD estimates are upward biased if people migrated from noncanal towns to canal towns after canal's opening; the DiD estimates are downward biased if people migrated from canal towns to noncanal towns after its opening. Second, I exclude targeted towns from the treatment group to further mitigate the potential selection problem. Third, I exclude towns in the middle section of the canal that was completed in 1820 from the treatment group. Fourth, I construct an alternative control group that includes towns that shared similar pre-existing characteristics across the New York State by using a propensity score matching approach.

Finally, I follow the approach of Faber (2014), Berger (2019) and Bogart et al. (2022) and use an instrumental variable (IV) strategy that exploits the variation induced by geographic locations, where towns near the Least Cost Paths (LCPs) were more likely to be connected to the canal. I construct the LCPs that correspond to the canal that state planners would have built if the objective had been cost minimization subject to connecting a few targeted towns along the route. Those untargeted towns are traversed by LCPs only because of their geography. While the targeted areas were arguably not randomly chosen, the validity of the instrument relies on the argument that areas that were located along LCPs were more likely to be traversed by the canal solely due to their location along these LCPs. I find broadly similar results using those specification checks and the IV strategy.

Further data on shipments and manufacturing establishments elucidate the mechanism underlying sectoral transition in employment and long-term growth. This shift stemmed from the development of small-scale manufacturing that processed agricultural products and the expansion of commercial activities brought about by lower transportation costs and easier access to distant markets. People moved into canal towns, established manufacturing mills, worked as manufacturing workers, and shipped processed agricultural staples along the canal as merchants.

This paper contributes to several strands of the literature on the economic impact of canals. One strand of this literature documents the impact of canals on western settlement and market integration (Whitford, 1906; Albion, 1939; Goodrich et al., 1961). The other studies the impact of canals on regional specialization (Ransom, 1967; Niemi, 1970; Ransom, 1971; Niemi, 1972; Ransom, 1972). The key contribution of this paper is to identify the causal effect of canals on sectoral transition in employment and long-term economic growth. It thus aligns with Ransom and Niemi's narrative debate about whether canals encourage regional specialization, using employment to measure the degree of regional specialization. However, these studies do not pay adequate attention to the endogeneity problem. Furthermore, these papers study the effects of canals on regional specialization using county-level data. But counties are too broad an analysis when investigating the canal effects in early nineteenth-century because the complementary land transportation is mainly wagon haulage. For example, the Erie Canal only traversed through the southern part of Herkimer County, leaving the northern three-quarters—a significant portion—with less direct and more costly access. By using town level data, this paper offers more precise estimates of effects of canals.

Second, as this study assesses long-term economic growth via population metrics, it contributes to the body of research on urbanization in nineteenth-century America. Atack et al. (2022) provide a literature overview emphasizing the influence of railroads on urbanization. This paper adds canals to the discussion, presenting a causal relationship between canals and urbanization.

Third, this paper contributes to the expansive literature examining the economic impact of transportation infrastructure. A significant portion of existing research is centered on railroads (Fogel, 1964; Fishlow, 1965; Atack et al., 2008; Donaldson and Hornbeck, 2016; Berger, 2019; Bogart et al., 2022). Recognizing that canals served as the main transportation medium before the advent of railroads, this paper introduces insights from the pre-railroad era to the discourse on transportation infrastructure.

Finally, this paper complements the literature on structural transformation by underscoring the commercial revolution preceding America's second industrial revolution and delineating the role canals played in it. While discussions pertaining to the commercial revolution in early nineteenth-century America are rare, there exists a body of work on market revolution (Sellers, 1991), where the dynamics of commerce are extensively explored.

In summary, identifying the impact of the Erie Canal on nineteenth-century New York State suggests both sectoral transition and long-term growth effects attributed to the canal. A notable point of comparison is the impact of the railroad. As for the sectoral transition effect, while railroads typically spurred a shift from agriculture to manufacturing (Atack et al., 2008; Berger, 2019; Bogart et al., 2022; Américo, 2022), my study indicates a transition primarily from agriculture to commerce (with early-stage manufacturing development) due to the canal. This difference highlights the commercial revolution preceding America's second industrial revolution. As for the long-term effects, my study shows that between 1790 and 1870, towns located close to the canal witnessed an average 38.6% faster population growth. In the context of railroads of Sweden, the rate stands at 25% between 1850 and 1900 (Berger, 2019). In the context of railroads of England and Wales, Bogart et al. (2022) reports a 0.875%

in annual population growth rate increase from 1851 to 1891. In the U.S., Donaldson and Hornbeck (2016) suggests that without railroads, population would have declined by 58.4% in 1890, and Atack et al. (2010) shows a 3.7% higher urbanization rate in counties with railroad access. Of course, the reader should proceed with caution when comparing those figures due to differences in context, methodology, measurement, study period, and data sources.

2 Background

Built between 1817 and 1825, the Erie Canal stretched 586 kilometers, connecting the Hudson River with Lake Erie. Prior to 1860, the ton-kilometers of shipping of the Erie Canal consistently surpassed that of any other singular transportation route (Poor, 1860). Unlike the railroads that emerged later, the Erie Canal epitomized state-led intervention. One outstanding characteristic of New York State's economic landscape during this period was the government's entrepreneurial role, underscored by state subsidies, such as state-backed loans to manufacturing corporates, tax exemption granted to certain manufactories, and financial support to agricultural societies (Miller, 1962). Canal construction and operation would entail a change in form and scale of prevailing government interference.

There were two main reasons favoring state-led construction of the canal over private enterprises. First, private entities grappled with financial limitations. Prior to the state's takeover of the Erie Canal's construction, two private companies tried to construct and operate specific sections of the canal, but both failed due to financial shortage. The Western Inland Lock Navigation Company and the Northern Inland Lock Navigation Company received authorization in 1792 to sell up to 1,000 shares at an introductory price of $25 each, alongside a state grant of $12,500 to facilitate canal work. Subsequent financial strain prompted the state to acquire 200 shares from each company and extend a $37,500 loan to the Western Company. Nonetheless, the canal still could not be completed because of shortages of funds despite the state's assistance.

When state involvement in canal construction grew, New York State briefly harbored aspirations of acquiring federal support, particularly when President Jefferson advocated for reallocating surplus Treasury revenue to internal improvements. However, hopes of federal intervention waned due to inter-state rivalry and the War of 1812, propelling the state to assume full responsibility. Legislation enacted in 1816 established the Canal Fund, designated to finance the canal's construction through revenue streams such as canal tolls, a tax on steamboat passengers, and a levy on salt production in the western district, among others. Further legislation in 1817 permitted loans on the credit of the state that allowed the Board of Commissioners of the Canals Fund to issue state bonds. Consequently, the Erie Canal reached completion in eight years at an expenditure of $7.143 million, approximately $185 million when adjusted to 2022 values. Furthermore, the Erie Canal was enlarged between 1836 and 1862 at a cost of nearly $32 million, or four and a half times the amount spent to build the original canal.

Second, the state, in contrast to companies, was better positioned to reconcile regional disputes and political rivalries. Although residents in western New York State wanted the canal, Long Island residents and Mid-Hudson Valley farmers resisted it, competitive from influx of Midwestern and Upstate produce into New York City. People in southern along the Pennsylvania boundary complained that the canal would be of no use to them because its benefits would be limited to areas proximate to the canal. Political rivalries further complicated matters. There were different political groups that tried to take credit for the canal. One group, led by DeWitt Clinton, who took office as Governor in 1817 and served until 1822, consisted of wealthy Federalist landholders in the West, great merchants in the towns, and western farmers who expected the canal to increase their land value and boost in-state commerce (Rubin, 1961). Counteracting this was a group led by Martin Van Buren that comprises the majority of Democrats who were in control of New York City and towns in the Hudson River Valley. To reconcile these political and regional differences, Governor

Clinton built a political coalition that transcended pure political allegiances and overcame regional disputes about financing and routing the canal.

In planning the route of the canal, the overriding objective was its engineering feasibility, especially given the imperative of connecting certain targeted towns: Buffalo, Syracuse, Utica, and Albany. Commercial employment spillovers to local regions was a secondary consideration at the early stage. Surveys were conducted by appointed canal commissioners and engineers[1]. Their expeditions, often conducted by land or along water routes, took them from the Mohawk River to Rome, from there to Oswego, and further from Three River Point to Geneva where the boats were sold and the party proceeding by carriage to the Niagara (Kirkland, 1934). These surveys ascertained vital details such as the consistent availability of water throughout the year without depleting hydraulic works and the practicality of constructing essential infrastructure including stone locks, culverts, and aqueducts (Laws of the State of New-York, 1828). Many of the tools and machinery used for land clearing and initial groundwork were repurposed for canal construction. For instance, with the right apparatus—a combination of rollers, cables, wheels, and cranks—a single individual could fell a large tree. Additionally, specialized equipment like the narrow plow was devised to cut through the dense mesh of roots and fibers on the surface (Canal Commissioners, 1817). Given the commissioners's objective of identifying a feasible and low-cost route, the canal's path was exogenous to most areas that were proximate to the canal. This exogeneity affords the opportunity to determine its effects on growth and development.

2.1 New York State before the Canal

The drive for internal improvement was spurred by two primary factors: western expansion and the evolution of intra-state and international trade. Between 1790 and 1814 New York

[1] Initial commissioners in 1811 included Gouverneur Morris, De Witt Clinton, Stephen Van Rensselaer, Simeon De Witt, William North, Thomas Eddy, and Peter B. Porter. By 1816, the roster had changed to feature De Witt Clinton, Stephen Van Rensselaer, Samuel Young, Joseph Ellicott, and Myron Holley. Notable engineers involved were James Geddes and Benjamin Wright, often dubbed the "father of American Civil Engineering".

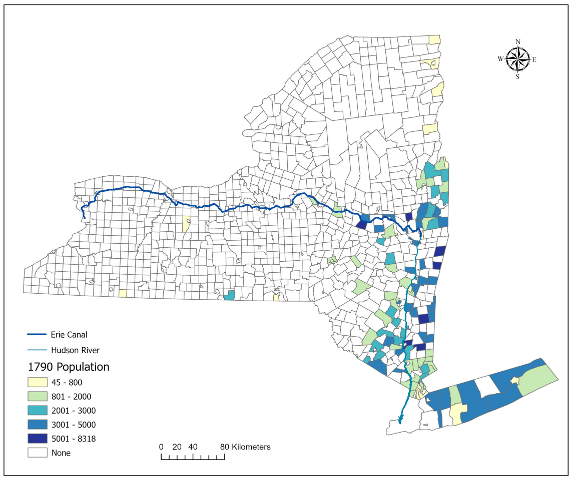
City's population more than doubled, from 54,182 to 123,980. Concurrently, it became one of the nation's primary port, a status indicated by the number of exports and imports loaded and unloaded at its wharves. By 1821, the port of New York was responsible for approximately 20% of total U.S. exports and 37% of its total imports value. The tonnage entered from foreign countries and cleared for foreign counties accounted for 20% and 19% of the U.S. totals, respectively (Albion, 1939). Yet, this commercial expansion was not confined to the city. Towns further north along the Hudson and Mohawk Rivers emerged as vital hubs, serving either as markets or as pivotal points in river trade, bolstered by their growing hinterlands.

The rapid western expansion, with New York State's population tripling from 339,582 in 1790 to 1,038,058 in 1814, heightened the demand for a market in rural New York State, and signified a state-wide shift towards a higher level of economic integration. In 1790, the majority of the population was clustered on Long Island and in the Hudson River Valley, with burgeoning towns such as Palatine, Canajoharie, German Flattes, Herkimer, and Whitestown emerging along the Mohawk River. By 1800 even more towns surfaced in the Mohawk River Valley, notably Rome, which later became a significant stop on the Erie Canal. The westward migration continued, with settlers establishing themselves around present-day Syracuse and in towns dotting the Finger Lakes Region. By 1810 settlers had made their way to Buffalo, and by 1820, the population centers emerged across the state (see Figure 1).

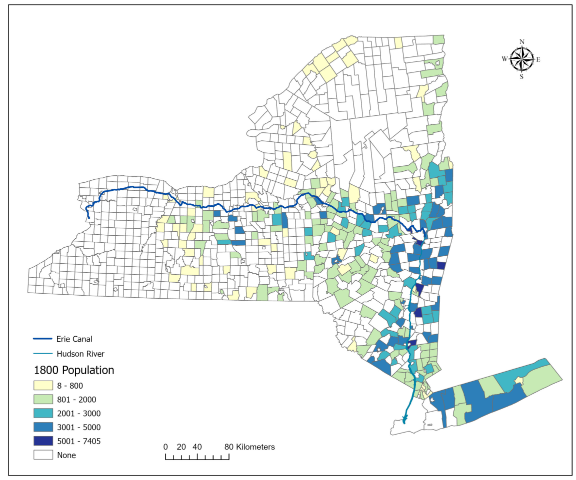
Populated towns were concentrated in three areas with navigable waters: the Hudson River Valley, the Mohawk River Valley, and the Finger Lakes Region. For merchants and farmers within these towns and their surrounding regions, the growing demand for enhanced transportation became increasingly evident.

2.2 New York State after the Canal

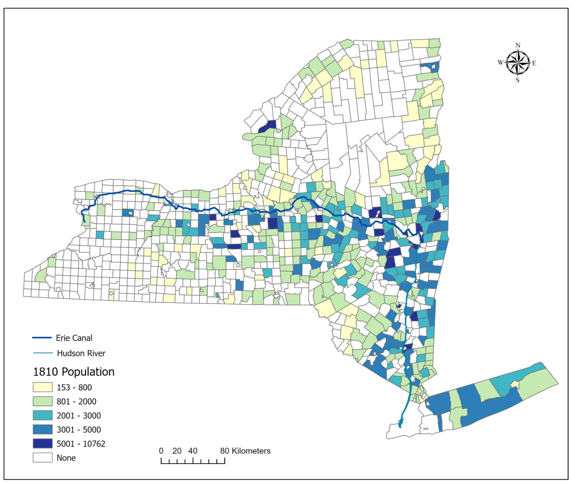
The Erie Canal helped the New York City established its dominance as the nation's primary port. By 1840, the port of New York was responsible for approximately 26% of total U.S.



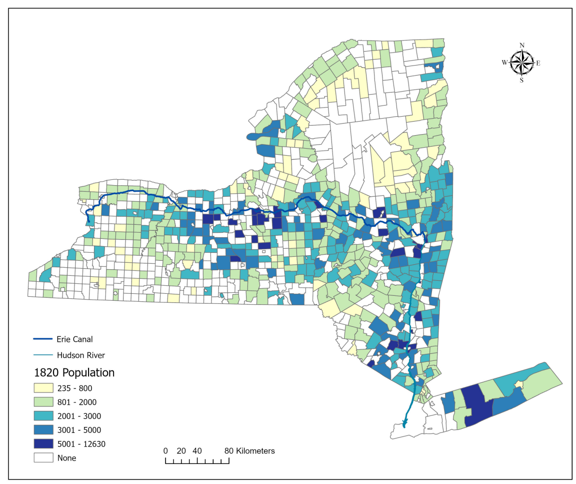
(a) 1790 Population



(b) 1800 Population



(c) 1810 Population



(d) 1820 Population

Figure 1: Town level population in New York State from 1790 to 1820

Notes: The town boundaries maps shown here are from the year 2021. Town boundaries do change from the study period to now. Detail of boundary adjustment is in section 3. New York City is excluded from the sample because of its extreme large population. The population in New York City was 54,182, 83,849, 125,805, and 123,980 respectively in the above 4 years. Population in the second largest town do shown in the maps as indicated by the changing number of the upper bar of the legend. Towns in white did not have population data in corresponding years.

exports and 56% of its total imports value. The tonnage entered from foreign countries and cleared for foreign counties accounted for 24% and 17% of the U.S. totals, respectively (Albion, 1939).

The post-1825 period is referred to as the era of market revolution (Sellers, 1991). During this time, nascent manufactures developed around the port towns, the adjacent countryside became specialized in producing perishable items such as vegetables, fruits, and dairy products for urban populations. Simultaneously, grain and livestock production extended farther out, with interior towns taking on the role of processing lumber, hides, and grain. The reduction in transportation cost gave areas a comparative advantage in more specialized production. Echoing Sellers (1991), Ellis (1946) describes this era as an agricultural revolution where by 1850 the old life of self-sufficiency was vanishing and all but the most isolated farmers were raising foodstuffs for sale. During the navigable season from late April to early December, the Erie Canal witnessed a bustling flow of boats. For instance, the total number of boats passing lock No. 26, located 5 kilometers west of Schenectady, surged from 6,166 in 1824 to 35,981 in 1854.

The Erie Canal facilitated a surge of western produce into the Hudson and Mohawk Valleys. In 1829, raw materials and basic resources cleared at Buffalo, such as 311,256 feet of lumber, 3,640 bushels of wheat, 4,335 barrels of flour, 4,754 barrels of provisions (comprising pork and beef), and 1,705 tons of agricultural ashes, were transported eastward on the Erie Canal. By 1835, the canal saw double the provisions, fourfold agricultural ashes, sevenfold lumber, twenty-threefold flour, and forty-sixfold wheat moving from Buffalo (refer to Table 7). Observations from 1830 to 1832 highlighted a similar product composition flowing through Utica. On the other hand, 7,151 tons of merchandise, 935 tons of furniture and mechanics' tools, and 65,431 barrels of salt were shipped to Buffalo from the East via the Erie Canal in 1830. These numbers increased by 1835, with salt inflow growing by 20%, merchandise by 300%, and furniture and tools by 800%. The goods arriving at Buffalo were consumed not just within New York State, but also in Upper Canada, Pennsylvania, Ohio, Michigan, Indiana, Illinois, Virginia, Kentucky, Tennessee, Missouri, and Alabama. The wide variety of articles shipped on the Erie Canal—lard, cheese, flax seed, hops, oysters and clams, whiskey, quinces, horn tips, wool, brick, iron ore, soap, plows, window sash, wooden ware, shingles, posts, and rails, etc.—paints a vivid picture of early nineteenth-century New

York State.

3 Data

To analyze the impact of the Erie Canal on long-term growth and sectoral transition in employment, I employ a town level dataset detailing population, employment in three broad sectors of the economy, manufacturing activity, and other socioeconomic characteristics, sourced from Bodenhorn and Cuberes (2018). The employment data originates from the federal census. The federal government only canvassed employment in the years 1820 and 1840. The 1820 federal census recorded employment in agriculture, commerce, and manufacturing. The 1840 federal census added employment in mining, ocean navigation, inland navigation, and professions. As for population, the dataset draws from both federal and state censuses. The federal government conducted decennial population in years ending with a zero. Beginning in 1825, New York State conducted statewide canvasses of its population every ten years in years ending in five and also held a census in 1814.

The dataset, comprising population and employment figures for towns, includes additional geographical, social, and economic information. The geographical data include elevation relative to sea level and proximity to the Iroquois Trail and turnpike. Social and political data encompass the count of residents eligible to vote, churches, schools, children aged 5 to 15, students attending school in 1824, and whether a location functioned as a county seat. Beyond population and employment, the economic data detail public expenditure, taxable property values, and the number of a variety of manufacturing establishments like distilleries, grist mills, fulling mills, tanneries, etc. These social and economic data, specific to 1824, are from the Gazetteer of the State of New York 1824 (Spafford, 1981).

I merge this dataset with Geographic Information System (GIS) data. The boundary data for cities and towns within New York State are from the NYS Office of Information Technology Services GIS Program Office (GPO). Due to the absence of historical boundary data, the 2021 boundaries serve as the base map. Over time, the delineations town boundaries have experienced modifications. The dataset shows usable information on cities and towns, which increased from 131 in 1790 to 573 in 1820 to 977 in 1870. In 2021, this number stands at 932 towns and 62 cities. One notable transformation is the establishment of new towns carved out from existing ones. To make the map consistent over time, data from newer towns are incorporated into their originating towns. For example, Middletown in Orange County was part of Newburgh in Orange County before 1888. In such cases, Middletown's data is merged with that of Newburgh. Another trend is that some regions have been incorporated as cities within their original town boundaries, as observed with Amsterdam city in Montgomery County, incorporated in 1885. In these scenarios, data are aligned with the contemporary township. I made town boundaries adjustments in reference to the Gazetteer of the State of New York 1824 and auxiliary data from Wikipedia.

Using Geographic Information System (GIS) tools, I measure the Euclidean distance between towns and the Erie Canal. The reference point for each town is its respective town hall. Distances are computed either perpendicularly from these halls to the canal or directly to the nearest canal vertex. The data for the historical canal path, Mohawk River, Hudson River, and New York State Thruway are from New York State GIS Clearinghouse or the ArcGIS online portal. For the instrumental variable construction, I use the Digital Elevation Models with ten meters pixel resolution from the U.S. Geological Survey. I digitalize data on articles transported through the Erie Canal from 1829 to 1835 from the Annual Report of the Canal Commissioners.

4 Difference-in-Differences Strategy

To examine the impact of the Erie Canal on sectoral transition in employment and long-term growth, the main empirical analysis is based on a difference-in-differences design. This design compares variables of interest in towns located close to the canal (termed "canal towns") with towns that shared similar characteristics but were not proximate to the canal (termed "noncanal towns") in periods before and after 1825, the year that the Erie Canal commenced operations. The main model is specified as

$$Y_{it} = \alpha + \theta_i + \phi_t + \beta Post_t Canal_i + \epsilon_{it}, \tag{1}$$

where Y_{it} either corresponds to log(population), log(employment), or the share of employment in agriculture, manufacturing, and commerce for town i in year t. $Canal_i$ is an indicator taking the value 1 for canal towns, defined as those located within 5 km of the canal, and 0 for noncanal towns, defined as those located between 5 and 15 km from the canal. Alternative distance measure to the canal was used to assess the sensitivity of my estimates. $Post_t$ is an indicator variable equal to 1 after year 1825 and 0 otherwise. θ_i is the town fixed effect, capturing any town invariant characteristics like elevation and terrain ruggedness. ϕ_t is the year fixed effect. Standard errors are clustered at the town level. α is a constant term common for all towns.

The DiD design rests on three core identification assumptions: the stable unit treatment value assumption (SUTVA), no-anticipation assumption, and parallel trends assumption. The SUTVA assumption implicitly implies that potential outcomes for unit i are not affected by the treatment of unit j. In other words, SUTVA rules out interference across units and spillover effects. This assumption likely held in the early nineteenth-century when a wagon traveled only 3 to 5 kilometers per hour depending upon weather, roadway conditions, and the health of the travelers. To further mitigate concerns over spillover effects, I exclude areas immediately adjoining the treated towns from the control group in the robustness checks.

The non-anticipation Assumption is likely to hold given that the Erie Canal commenced operations in 1825 and the pre-treatment year considered is 1820. While the canal might have influenced the occupational choices of the population by adjusting the occupational rewards, it's posited that such a transformation was less probable five years prior in early nineteenth-

Table 1: Balance test of dependent variables in 1820 between canal towns (0-5km) and noncanal towns (5-15km)

	Canal towns	Obs.	Noncanal towns	Obs.	Diff.	p-value
Population	2750.09	47	3115.11	53	365.03	0.32
Total Employment	647.92	48	720.15	54	72.23	0.34
Employment in agriculture	483.58	48	565.57	54	81.99	0.16
Employment in manufacturing	144.98	48	137.56	54	-7.42	0.81
Employment in commerce	19.35	48	17.02	54	-2.34	0.85
Agriculture employment share	0.75	48	0.81	54	0.06	0.12
Manufacturing employment share	0.22	48	0.17	54	-0.05	0.11
Commerce employment share	0.03	48	0.02	54	-0.01	0.29

Notes: Data are from the federal and state census in 1820 and the Gazetteer of the State of New York 1824. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

century New York State. As a robustness check, I will segment the data, excluding towns located in the middle section of the Erie Canal, which began operations before 1820. The Parallel Trends Assumption is tested by estimating the effect of the canal on population using population data from 1790 to 1820. I also provide a balance test between the treated control groups by comparing the pre-treatment town characteristics using data from the 1820 federal census and the 1824 Gazetteer of the State of New York.

Table 1 demonstrates that there are not significant baseline differences in 1820 between canal towns and noncanal towns among eight dependent variables: population, total employment, employment shares in three sectors, and employment count in three sectors. Table 2 shows the balance test in 1824 between canal towns and noncanal towns concerning other town characteristics: economic indicators (such as taxable property value, the public expenditure, and number of cloth made), old transportation access (the Iroquois trail and turnpike access), political influence (the county seat status and the number of residents eligible to vote), education (number of schools and people were taught), religious institutions (the number of churches), elevation, the breadth of manufacturing establishment and their products and equipments, and others. North American towns tended to locate at native portage points, as demonstrated by (Bleakley and Lin, 2012). The Iroquois Trail was an an

Table 2: Balance test of other town charactersitics in 1824 between canal towns (0-5km) and noncanal towns (5-15km)

	Canal towns	Obs.	Noncanal towns	Obs.	Diff.	p-value
Economic indicator						
Public expenditure	301.03	35	648.88	43	347.86	0.16
Taxable property(thousand dollar)	410.34	43	387.73	54	-22.61	0.81
No. of cloth made(thousand yards)	15.19	43	36.87	53	21.68	0.22
Old transportation						
Iroquois trail(0/1)	0.95	21	1.00	11	0.05	0.48
Turnpike(0/1)	0.50	46	0.60	55	0.10	0.41
Political influence						
County seat(0/1)	0.17	71	0.05	96	-0.12**	0.01
No. of Electors	516.30	44	561.24	51	44.94	0.51
Education						
No. of schools	11.83	41	14.35	49	2.52*	0.09
No. of students	660.28	36	847.09	43	186.82*	0.08
Religion						
No. of churches	2.83	12	4.53	15	1.70	0.12
Manufacturing establishment						
No. of distilleries	1.42	45	2.02	55	0.60	0.16
No. of grist mills	3.60	45	4.02	55	0.42	0.53
No. of saw mills	7.02	45	8.65	55	1.63	0.18
No. of oil mills	0.47	45	0.42	55	-0.05	0.80
No. of fulling mills	1.80	44	2.51	55	0.71	0.13
No. of asheries	2.29	45	2.56	55	0.27	0.65
No. of cotton woollen factories	0.41	44	0.50	54	0.09	0.70
No. of tanneries	0.07	44	0.11	54	0.04	0.76
Others						
Elevation(feet)	472.62	71	676.19	96	203.57***	0.00
Custom(0/1)	0.03	71	0.01	96	-0.02	0.40
No. of children age 5-15	733.54	35	778.32	41	44.77	0.66
No. of carding machine	1.73	44	2.78	55	1.05	0.16
No. of trip hammers	0.25	44	0.56	54	0.31	0.18

Notes: Data are from the federal and state census in 1820 and the Gazetteer of the State of New York 1824. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

ancient Native-American trade route used by colonial military forces in troop deployments that crossed several rivers. As Table 2 shows, canal towns and noncanal towns do not differ in terms of the access to the Iroquois Trail. Elevation not only presented engineering challenges for canal construction but also potentially influenced economic development (Nunn and Puga, 2012). The Erie Canal's construction was influenced by political dynamics, as noted by (Rubin, 1961). I then compare canal towns and noncanal towns in terms of elevation and whether a town functioned as a county seat (i.e., a regional governmental center, which typically included the county courthouse, the county jail, and other administrative offices). While differences in elevation and county seat exist between groups, these are accounted for in the town fixed effects, as they remained constant from 1820 to 1840. The data provides weak evidence of differences in the number of schools and students, with respective p-value of 0.09 and 0.08. Due to the availability of only one year's educational data, it's not feasible to control for educational characteristics in the primary regression. However, when excluding targeted towns in the subsequent robustness checks, no significant disparities in educational attributes emerge. While differences in elevation and county seat exist between groups, these are accounted for in the town fixed effects, as they remained constant from 1820 to 1840. Byound these factors, Table 2 shows there are not significant differences in economics indicator, old transportation access, religion, the breadth of manufacturing establishment and their products and equipments, and others.

In 1820, there were 48 canal towns and 54 non-canal towns. By 1840, these numbers grew to 66 canal towns and 89 non-canal towns. By 1870, they further expanded to 71 canal towns and 96 non-canal towns. Correia (2015) shows maintaining singleton groups—groups with only one observation—in linear regressions where fixed effects are nested within clusters can overstate statistical significance and lead to incorrect inference. Given that employment data is available for only two years and the federal government only canvassed employment in one year for some towns, the main regression, focusing on the canal's impact on sectoral transition, includes 48 canal towns and 54 non-canal towns with data from both 1820 and 1840.

A closer look at the towns dropped is useful (see Table 3). The 53 towns excluded—

comprising 18 canal towns and 35 noncanal towns—emerged between 1820 and 1840. A balance test for 1840 shows that, on average, the newly formed canal towns exhibited higher total employment and employment in manufacturing and commerce. They also displayed lower agricultural employment shares and higher manufacturing employment shares. These findings suggest the DiD estimates for sectoral transition might be downward biased, implying that this study presents a lower bound for the canal's effects on sectoral transition.

Regarding population, given that I use 15 years' observation between 1790 and 1870 and no towns has only one year of observation, the long-term effects estimated will not be biased due to the same reason. I include all 71 canal towns and 96 non-canal towns up to 1870 in the analysis of long-term effects.

Table 3: Balance test of employment in 1840 between dropped singleton canal towns (0-5km) and noncanal towns (5-15km)

	Canal towns	Obs.	Noncanal towns	Obs.	Diff.	p-value
Total Employment	1190.61	18	660.09	35	-530.53**	0.01
Employment in agriculture	644.28	18	502.89	35	-141.39	0.13
Employment in manufacturing	421.56	18	123.23	35	-298.33*	0.01
Employment in commerce	64.17	18	16.23	35	-47.94*	0.04
Agriculture employment share	0.66	18	0.79	35	0.13*	0.02
Manufacturing employment share	0.26	18	0.17	35	-0.10*	0.02
Commerce employment share	0.04	18	0.02	35	-0.02	0.10

Notes: Data are from the federal and state census in 1820 and the Gazetteer of the State of New York 1824. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5 Results

5.1 Main Results

Table 4 presents the estimated effects of the Erie Canal on sectoral transition in employment and population, based on Equation 1. Columns (1), (2), and (3) assess the canal's influence on employment shares in agriculture, manufacturing, and commerce respectively. Columns

(1) and (3) reveal that canal towns had a lower employment share in agriculture and a higher employment share in commerce. Column (3) shows canal towns had higher manufacturing employment shares but not statistically significant. Quantitatively, canal towns recorded 6.9% less employment share in agriculture, 1.5% more employment share in manufacturing, and 2.3% more employment share in commerce when compared with noncanal towns. In 1820, five years before Erie Canal's operation, canal towns had 75%, 22%, and 3% across agriculture, manufacturing, and commerce, respectively.

I shift the dependent variable to log(employment) in three sectors and log(total employment) to analyze the composition change of total employment in each sector. Column (4) indicates that the Erie Canal's operation did not significantly decrease employment in agriculture in canal towns. As the agricultural employment share is quantified as the ratio of agricultural employment to total employment, the lower agriculture employment share in canal towns was not due to declining employment in agriculture, but rather the increasing employment in total employment, as elucidated in columns (5), (6), and (7). Specifically, column (6) attributes the augmented commerce employment share in canal towns to the sector's growing employment. Column (5) reflects a growth in manufacturing employment in canal towns. However, this growth, though significant, was relatively small to the growth of total employment, which explains the insignificant change in the sector's employment share. Quantitatively, compared to noncanal towns, canal towns exhibited 41.1% faster employment growth in manufacturing, 135.3% faster employment growth in commerce, and 26% faster total employment growth. The 1820 employment counts for agriculture, manufacturing, commerce, total employment stood at 484, 145, 19, and 648, respectively.

Building on the preceding section's discussion, the regression for Equation 1 excludes towns with employment data from only the year of 1840. The 53 towns excluded—comprising 18 canal towns and 35 noncanal towns—emerged between 1820 and 1840. Given that the newly incorporated canal towns exhibited higher total employment, higher employment in manufacturing and commerce, and lower agricultural employment shares (see Table 3), the

DiD estimates presented in columns(1), (5), (6), and (7) present a lower bound for the canal's effects on agricultural employment share, manufacturing employment, commerce employment, and total employment. While these new canal towns also had higher manufacturing employment shares, it remains uncertain whether canal towns experienced significantly higher manufacturing employment shares. This is contingent upon the relative growth rates of manufacturing employment to total employment. Nonetheless, column (5) shows that canal towns had faster employment growth in manufacturing compared with noncanal towns.

Table 4: DiD estimates of effects of canal operation on sectoral transition in employment

	(1) Agriculture employment share	(2) Manufacturing employment share	(3) Commerce employment share	(4) log(employment) in agriculture	(5) log(employment) in manufacturing)	(6) log(employment) in commerce	(7) log(total employment)
PostCanal	-0.069***	0.015	0.023***	0.027	0.345**	0.856***	0.231*
	(0.024)	(0.023)	(0.008)	(0.126)	(0.166)	(0.273)	(0.121)
Observations	204	204	204	204	204	204	204
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Town FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The row of Post Canal shows the effect of the canal, β, on different outcome of interest from the DiD estimates. The constant term is estimated but not shown in the table. Standard errors in parentheses are clustered on towns.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

While the employment data spans only two years, population records from 1790 to 1870 offer an opportunity to assess the long-term growth effect of the Erie Canal. I proceed by deploying an event study to estimate the canal's effects on log(population), employing the subsequent approach

$$log(Population)_{it} = \alpha + \theta_i + \phi_t + \sum_{k=1790}^{1820} \beta_k lead_k Canal_i + \sum_{k=1825}^{1870} \beta_k lag_k Canal_i + \epsilon_{it}, \qquad (2)$$

where $lead_k$ is the time indicator for years before 1825. The $lead_k$ variable includes years of 1790, 1800, 1810, and 1814. lag_k is the time indicator for years after the operation of the canal. The lag_k variable includes years from 1825 to 1870 with 5 years interval. Other parameters and variables are defined as in Equation 1.

Figure 2 shows the results from an event study of population between 1790 and 1870. There is no significant difference in population between canal towns and noncanal towns from

1790 to 1820, which provides further evidence to the validity of the parallel trend assumption. Following the Erie Canal's completion, canal towns registered a pronounced acceleration in population growth from 1825 to 1870, in contrast to their noncanal counterparts. The canal's enduring growth ramifications spanned at least 45 years, within the confines of this study's timeframe. Transforming natural *log* to exponential form, canal towns witnessed a faster population growth at rate of 21.2%, 27.1%, 39.9%, 44.8%, 44.1%, 49.8%, 58.7.0%, 52.6%, 56.1%, and 67.3% from 1825 to 1870 respectively, given 1820 as the benchmark year. By substituting $log(population)_{it}$ for Y_{it} in Equation 1, the DiD estimate of β is 0.326. By transforming natural *log* to exponential form, canal towns witnessed 38.6% faster population growth.

Overall, Table 4 and Figure 2 highlight that the Erie Canal's operation had significant effects on sectoral transition in employment and the long-term growth in population. There was a decline in the agricultural employment share, a rise in the commercial employment share, and faster employment growth in manufacturing for towns in proximity to the Erie Canal. These canal towns witnessed a pronounced long-term growth effect, registering a more substantial population increase compared to towns situated further from the canal.

5.2 Sensitivity Checks and the Extent of Canal's Effects

First, I assess the sensitivity of my estimates to the choice of canal towns definition. In Figure 3, each subfigure displays 25 separate DiD regressions, derived from Equation 1, adjusting the definition of canal towns in increments from 1 km to 25 km. Noncanal towns—serving as the control group—are those located within 35 km of the canal but not categorized as canal towns. The time frame for this analysis spans 1820 to 1840. Subfigures (a), (b), and (c) display DiD estimates of the canal's effects on sectoral transition, with agriculture employment share, manufacturing employment share, and commerce employment share as the dependent variable respectively. Subfigure (d) shows the DiD estimates of canal's effects on long-term growth with $log(Population)$ as the dependent variable. The collective findings

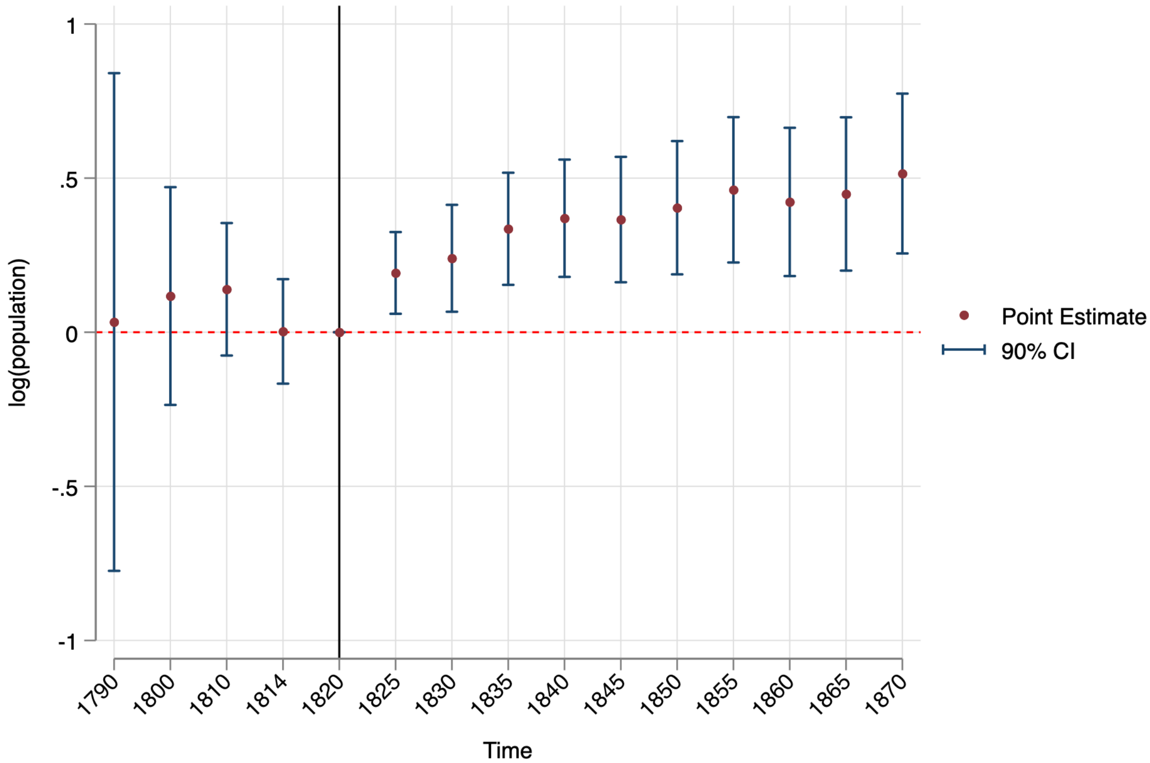
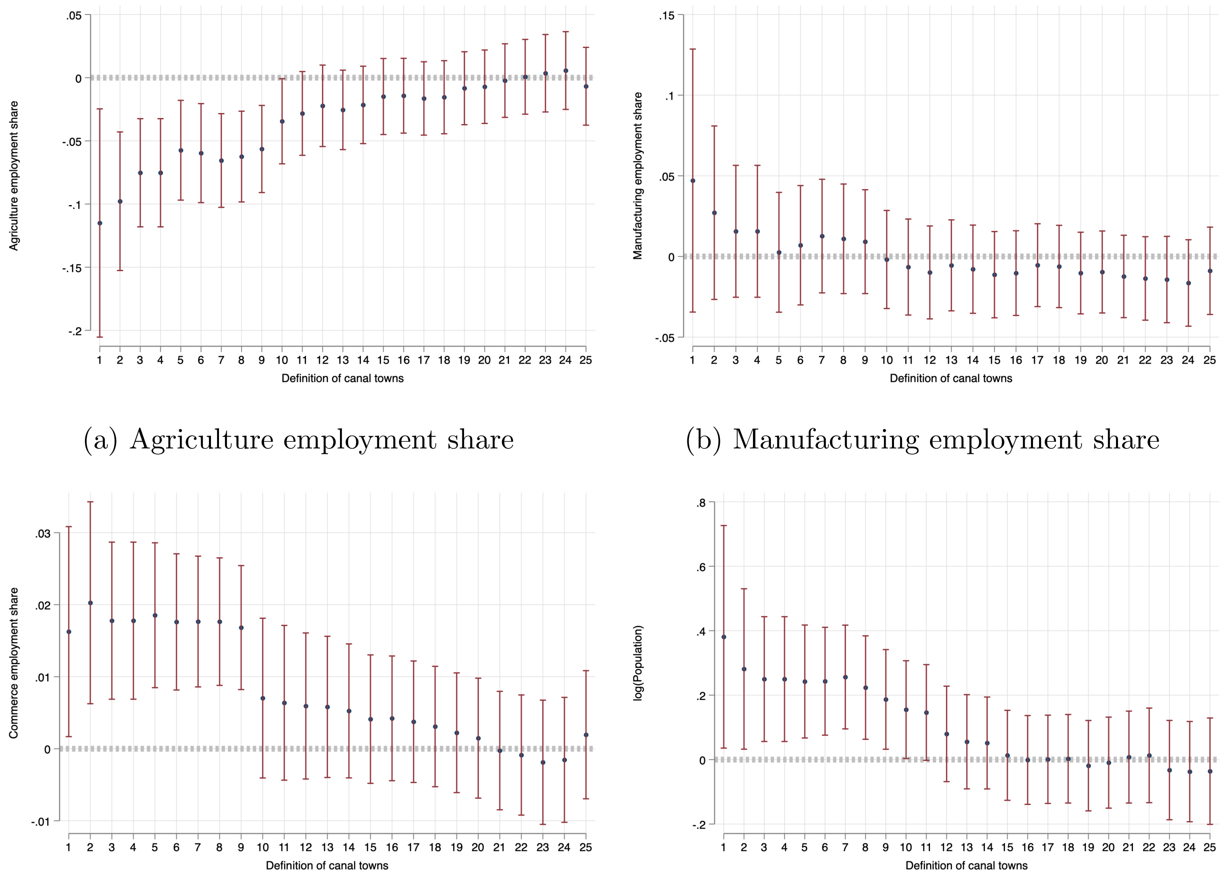


Figure 2: The event study of canal operation on long-term population growth

Notes: Each point shows the DiD estimate by comparing $log(population)$ in canal towns relative to noncanal towns with 1820 as the benchmark year. Standard errors are clustered on towns.

from Figure 3 show that the main results, determined by a 5 km canal town definition, are robust and not sensitive to marginal distance change. Canal towns, in comparison to their noncanal counterparts, experienced faster population growth, lower agricultural employment share, and higher commerce employment share. Additional outcomes for other variables are detailed in Figure 6 in appendix.

Second, Figure 3 suggests the extent of canal's impact on sectoral transition and long-term growth. Specifically, subfigures (a) and (c) indicate the effect on sectoral transition in employment from agriculture to commerce diminishes in towns as they lie increasingly distant from the canal. Subfigure (b) in Figure 6 in appendix also indicates the effect on manufacturing employment diminishes in towns as they lie increasingly distant from the canal. The sectoral transition effect becomes statistically insignificant beyond 10 km



(a) Agriculture employment share

(b) Manufacturing employment share

(c) Commerce employment share

(d) Population

Figure 3: Sensitivity of DiD estimates to the choice of canal towns definition

Notes: The horizontal axis is the definition of $canal_i$ by using different distance from the towns to the canal. The distance unit is kilometer. Each point is estimated by Equation 1 with 95% confidence interval. Standard errors are clustered on towns.

and disappear beyond 20 km. Subfigure (d) indicates the long-term growth effect becomes statistically insignificant beyond 10 km and disappear beyond 15 km. Therefore, the extent of canal's effects on sectoral transition and long-term population growth was 15 to 20 km between 1820 to 1840.

Given the era's transportation limitations—absent cars and trains—a 15 to 20 km distance was consequential. In the early nineteenth-century, a wagon traveled 3 to 5 kilometers per hour depending upon weather, roadway conditions, and the health of the travelers. Evaluating transportation costs, wagon transport rates ranged from 6 to 12 cents per ton-kilometer during this period (McClelland, 1968). Consequently, the cost associated with this 15 to 20 km journey was between 90 to 240 cents per ton, exclusive of the time investment—approximately 3 to 6 hours for a horse or mule. To contextualize the transportation cost from a wage perspective, the daily wage for a common labor in the northeast U.S. was 69 cents in 1830 (Ba4254 in Historical Statistics of the U.S.).

To provide context regarding the value of goods transported via the canal, I reference the wholesale prices of wheat and wheat flour (Cc206 and Cc211 in Historical Statistics of the U.S.). In 1830, in the northern U.S., the wholesale price for one ton of wheat stood at \$39.31. Thus, the cost of transporting this commodity 15 to 20 km by wagon amounted to between 2.3% and 6.1% of its wholesale value. Meanwhile, in New York City, the wholesale price for one ton of wheat flour was \$50.92. This implies that transportation costs, for the same distance by wagon, equaled 1.8% to 4.7% of its wholesale value in 1830. Incorporating time costs would amplify these figures. In 2023, the transportation cost of shipping one ton of wheat from farms to rail in Kansas State is around 4% of its farm value (U.S. Department of Agriculture, 2023).

5.3 Robustness Checks

5.3.1. Excluding the Immediately Adjoining Towns to Tackle the Potential Spillover Effects. A potential issue to consider is the bias in the estimates due to spillover effects between canal towns and noncanal towns. Take population, one of the outcome variables of interest, for illustration. If migration occurred from noncanal towns to canal towns post-canal operation, the DiD estimates would be upward biased. Conversely, the DiD estimates would be downward biased if there have been a migration from canal towns to noncanal towns. In addressing this concern, I exclude areas immediately adjoining the treated towns from the control group to mitigate spillover effects. The assumption here is that while migration may have occurred between canal towns and their neighboring noncanal towns, such movement was less probable between canal towns and more distant noncanal towns, given the substantial wagon transportation costs in early nineteenth-century New York State.

In the baseline DiD estimates, control towns are those located between 5 km and 15 km from the canal. I split the control towns into two parts and exclude those from 5 km to 10 km. Now the control towns are solely those situated from 10 km to 15 km away from the canal. A balance test, available in Table 10 in appendix, compares treated towns to this revised control group. Notably, no significant baseline disparities emerge in population or employment metrics. Any time invariant variable is accounted for by town fixed effects. The effects of canal operation on sectoral transition in employment and population remain significant with same signs and slightly smaller magnitudes, as evidenced in Panel A of Table 5.

Table 5: Robustness checks for DiD estimates

	(1) Agriculture Employment share	(2) Manufacturing employment share	(3) Commerce employment share	(4) log(employment) in agriculture	(5) log(employment) in aanufacturing)	(6) log(employment) in commerce	(7) log(employment) total	(8) log(Population)
Panel A: Excluding adjoining towns								
PostCanal	-0.077***	0.021	0.019***	0.140	0.501**	1.117***	0.352**	0.436***
	(0.027)	(0.025)	(0.005)	(0.166)	(0.194)	(0.319)	(0.159)	(0.121)
Observations	150	150	150	150	150	150	150	148
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Town FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Panel B: Excluding targeted towns								
PostCanal	-0.064***	0.027	0.012***	0.036	0.306*	0.760***	0.163	0.259***
	(0.023)	(0.021)	(0.004)	(0.120)	(0.166)	(0.279)	(0.119)	(0.095)
Observations	194	194	194	194	194	194	194	190
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Town FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Panel C: Excluding the middle section								
PostCanal	-0.075**	0.029	0.018**	0.033	0.415**	0.855***	0.259**	0.342***
	(0.029)	(0.024)	(0.008)	(0.132)	(0.181)	(0.311)	(0.126)	(0.113)
Observations	178	178	178	178	178	178	178	176
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Town FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Panel D: Matching method								
PostCanal	-0.057***	-0.012	0.022**	0.081	0.367***	1.266***	0.332***	0.370**
	(0.022)	(0.029)	(0.011)	(0.137)	(0.116)	(0.346)	(0.106)	(0.145)
Observations	1082	1082	1082	1084	1048	1090	1082	1098
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Town FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The constant term is estimated but not shown in the table. Standard errors in parentheses are clustered on towns.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5.3.2. Excluding the Targeted Towns. In the process of constructing the canal, several towns were planned to be traversed by the canal. They were Buffalo, Syracuse, Utica, and Albany. Many other towns were either of lesser importance or did not exist when the canal was constructed like Lockport. A potential issue arises if employment in manufacturing and commerce in these major towns was already higher than in other control towns, which would create a selection bias. To address this concern, I exclude these four targeted towns from the treated group. Panel B in Table 5 provides the DiD estimates in the absence of these towns. The impacts of canal operation on sectoral transition in employment and population persist, exhibiting consistent significance and direction, albeit with marginally reduced magnitudes.

5.3.3. Excluding Towns in the Middle Section. The middle section of the Erie Canal, extending from Utica in Oneida County to Montezuma in Seneca County, began its construction in 1817 and was operational by 1820. This segment traversed various towns within the counties of Cayuga, Madison, Oneida, and Onondaga. Towns situated within 5 km of the canal in these counties might have experienced the canal's effects as early as 1820, contingent on the immediacy of employment reactions to canal operations. Consequently, I remove from the treated group those towns lying within this 5 km proximity in the aforementioned counties. The resultant impacts of canal operation on sectoral transition in employment and population remain significant, with consistent direction and comparable magnitudes, as illustrated in Panel C of Table 5.

5.3.4. Changing Control Towns Using Matching Approach. Towns shared similar pre-treatment employment structures and population may be located across New York State. That is, candidate control towns located more the 15 km from the Erie Canal are previously excluded from our analyses. An alternative approach to selecting a control group by including those towns is to employ the matching method. I first use propensity score matching (PSM) to choose suitable control towns. For the matching process, data from the pre-treatment year, 1820, including population, employment shares, and log(employment) in both agriculture and commerce, were utilized. Towns with extreme propensity scores, falling below 0.01 or exceeding 0.9, were excluded. Figure 7 in appendix compares propensity scores of treated towns with those of control towns. Subsequent to this, I apply the DiD method, as articulated in Equation 1, utilizing propensity score-weighted regressions. Panel D in Table 5 presents the PSM-DID estimates. The canal's operational effects on sectoral transition in employment and population persist in significance, maintaining consistent directions and magnitudes.

6 Alternative Indentification Strategy: Instrumental Variables

Consistent with facts of the Erie Canal construction and the methodologies of Faber (2014), Berger (2019), and Bogart et al. (2022), I leverage variation induced by geographic location as an instrument for canal accessibility. Specifically, towns situated near least-cost paths (LCPs) had a higher likelihood of canal connection. As the basis of the IV strategy, I constructed a hypothetical network that connects the targeted places on the canal. These targeted places were selected by the state to pass through because of their existing traffic volumes and/or their potential to generate freight traffic. Conversely, many other locations were incorporated into the canal's route solely due to geographic convenience. They bridged the targeted places in the cost-minimizing manner. Such areas were not directly targeted by planners and were "accidentally" traversed by the canal.

The IV strategy consists of two steps. In the first step, I undertake the first-stage regression in the following format

$$Canal_i = \mu + \tau_j + \delta DistanceToLCP_i + \nu_i, \tag{3}$$

where $DistanceToLCP_i$ is the instrumental variable, which is the Euclidean distance between town hall and the LCPs. τ_j is the county fixed effect for accuracy of the first stage regression. ν_i is clustered at county level. The first-stage prediction of the $Canal_i$ variable is then used to estimate the cross-section change equation

$$\Delta Y_i = \tau_j + \beta \hat{Canal}_i + \epsilon_i, \tag{4}$$

where $\hat{Canal}_i$ is the predicted values of the first-stage regression. ΔY_i either corresponds to changes of log(population) or log(employment), or the share of employment in agriculture, manufacturing, and commerce of town i between 1820 and 1840. These two steps are just for illustration. I use Stata to do the two steps since the OLS standard errors from the manual second stage will not be the correct (Angrist and Pischke, 2009).

To construct the instrument, I adhere to a three-step process. The first step is to choose places that LCPs need to traverse. Incorporating more such places imposes stricter constraints on LCPs, reducing the likelihood of accidental canal towns. I commence by selecting only the canal's origin, Buffalo, and its terminal, Waterford, which is situated adjacent to Albany county, Saratoga county, Rensselaer county, and the Hudson River. Second, to discern the LCPs bridging Buffalo and Waterford, I employ Digital Elevation Models, which provide elevation data at a resolution of approximately 10 meters (U.S. Geological Survey, 1995). These LCPs, in their effort to minimize construction expenses, naturally align with ruggedness as calculated from elevation data. Third, I adapt the methodology of Bogart et al. (2022) to test the validity of the LCPs from an economic perspective. A simple gravity equation[2] and Flater (2022)'s algorithm are utilized to estimate the relative significance of connecting various town pairs. I measure town distances using the Euclidean distance and employ population data from the 1820s, reflecting the Erie Canal's construction era (1817-1825). Consequently, I produce a gravity index matrix for all towns from 1820, excluding New York City, given its disproportionate population size. Subsequently, I rank gravitational values relative to Waterford, pinpointing Halfmoon as the town with the highest value. I then determine the town that has the largest gravitational value with Halfmoon. If the town that has the largest gravitational value with Halfmoon is Waterford, then I choose the town

[2] $G_{ij} = (population_i * population_j)/distance_{ij}$

that has the second-largest gravitational value with Halfmoon. I persist with this iterative process until Buffalo is integrated. Inherent to the canal's design is its reliance on water; hence, I introduce a constraint. When faced with towns having comparable gravitational values, preference is given to the one in closer proximity to water resources. LCPs are not necessarily traverse the center parts of the selected towns. If LCPs approximate to those town, it's inferred that their design factored in economic considerations like population and topographical variables like ruggedness.

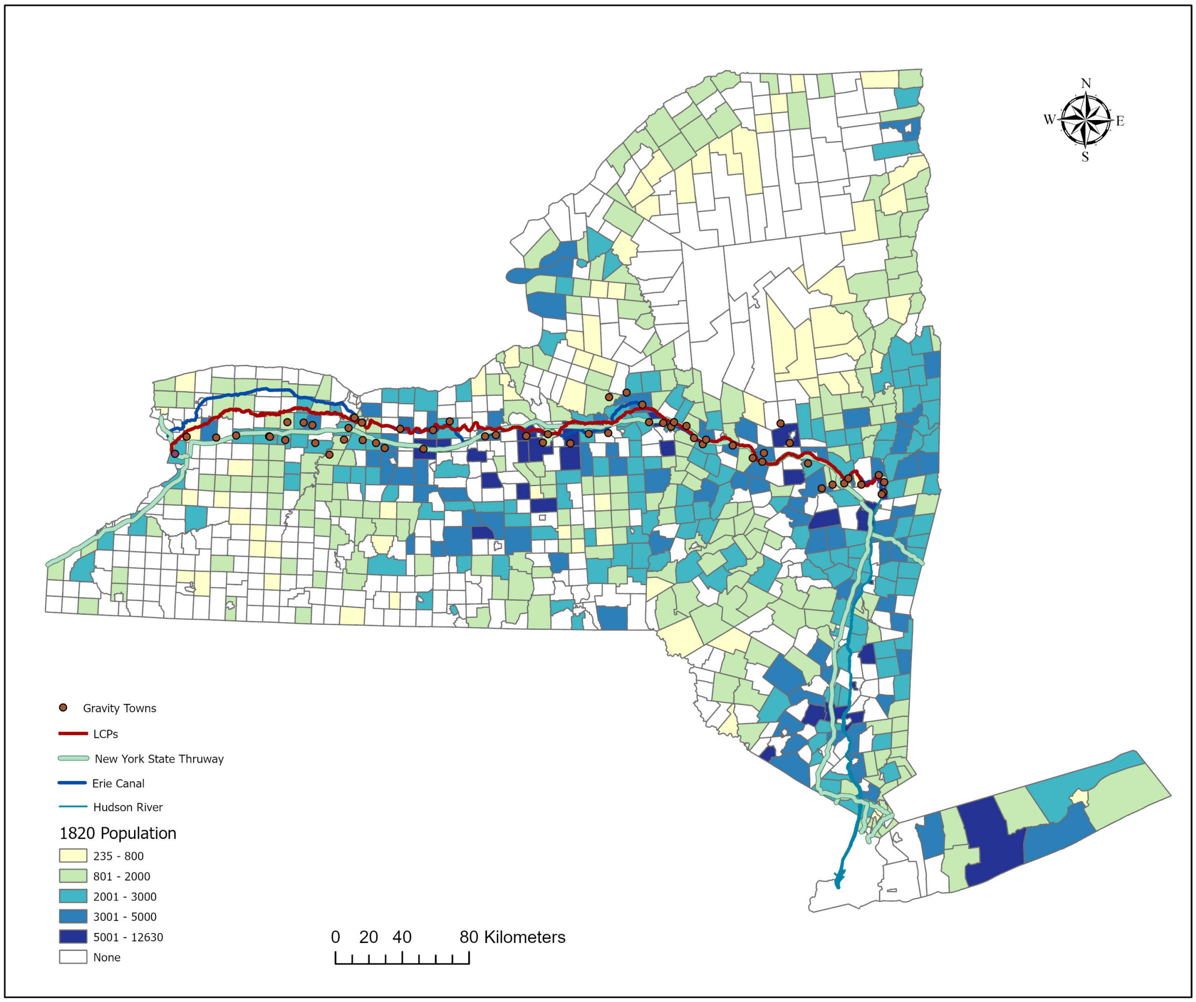


Figure 4: Least cost paths, gravity towns, and population in 1820

As depicted in Figure 4, the LCPs traverse all targeted towns, including Albany, Utica, Syracuse, and Buffalo. The LCPs almost overlap with the eastern section of the Erie Canal, exhibit slight deviations from its middle section, and largely diverge from the western section. The LCPs fit gravity towns, selected by gravity index, very well. An instructive point of reference for the LCP's plausibility is the New York State Thruway, which was constructed in 1954 by the New York State Thruway Authority. Given the technological advancements by 1954 relative to 1820, this thruway offers a more cost-efficient layout. Notably, the thruway's path largely coincides with the LCP.

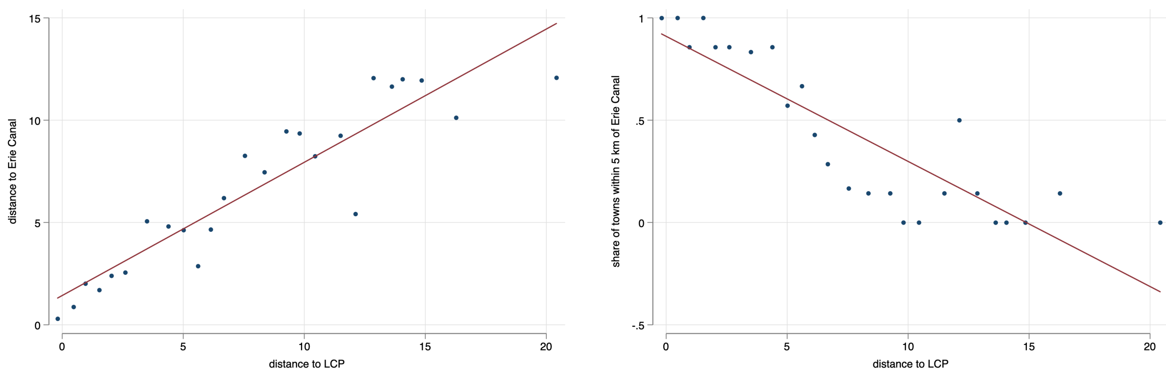
To interpret β as the local average treatment effect (LATE) of being treated by LCPs, the IV approach requires the following identifying assumptions: first stage, monotonicity, conditional independence, and exclusion. In the subsequent sections, I present empirical evidence that supports the validity of these assumptions.

First Stage. To demonstrate that towns proximate to the LCPs have a higher likelihood of being intersected by the Erie Canal, refer to Figure 5. This figure presents binned scatterplots for clarity. The towns used here align with those in the DiD estimates. Towns are grouped into 25 equal-sized bins based on the distance to the nearest LCP. In these scatterplots, each dot presents the mean distance to the closest point on the Erie Canal (panel a) or the share of towns situated within 5 km of the Erie Canal (panel b) against the mean distance to the nearest LCP for each bin. Also shown are the estimated best-fit lines. Evidently, towns closer to the LCPs are more likely to be traversed by the Erie Canal, accounting for county fixed effects.

Table 6: First stage regression

	(1) Canal	(2) Canal	(3) distance to Erie canal	(4) distance to Erie canal
distance to LCP	-0.061***	-0.073***	0.655***	0.782***
	(0.009)	(0.006)	(0.094)	(0.065)
Observations	167	167	167	167
County FE	No	Yes	No	Yes
R-squared	0.446	0.592	0.570	0.756
F-statistics	49.341	135.196	49.008	143.465

Notes: The constant term is estimated but not shown in the table. Standard errors in parentheses are clustered on towns.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.



(a) Distance to the Erie Canal

(b) Share of canal towns

Figure 5: First stage

Table 6 displays the results from the first-stage regression. As demonstrated in Columns (1) and (2), the likelihood of a town being traversed by the canal diminishes as its distance from the LCP increases. Notably, the distance to the closest LCP emerges as a significant predictor for the actual canal constructed, even after controlling for county fixed effects. The $F-$statistics provided are always sufficiently large to satisfy the weak instrument criterion (Stock and Yogo, 2005), namely a first-stage $F-$statistics in excess of 10.00. Furthermore, a strong correlation is evident between a town's proximity to the Erie Canal and its distance to the nearest LCP, as detailed in Columns (3) and (4) of Table 6.

In the presence of heterogeneous treatment effects, the monotonicity assumption is essential to interpret IV estimates as a LATE, i.e., the average causal effect among towns induced by the instrument into being treated by the canal. In my setting, this monotonicity implies a negative correlation between a town's likelihood of canal treatment and its distance from the LCPs. Subfigure (b) of Figure 5 lends support to this assumption, demonstrating a decrease in the proportion of towns traversed by the Erie Canal as their distance from LCPs grows.

Balance Tests. For the instrument to be valid, it requires that there is no correlation (conditional on county fixed effects) between factors that affect a town's potential outcomes and its proximity to the LCPs. To interpret the IV estimates as identifying the causal effect of being treated by the Erie Canal, the exclusion restriction must hold; specifically, the LCP should only influence town outcomes via the Erie Canal, without other intervening channels. As an indirect test of these two assumptions, I present evidence indicating no significant correlation between observed town characteristics in 1820 and the instrument.

Table 7: Balancedness of instrument

	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A			
	log(Employment) in agriculture	log(Employment) in manufacturing	log(Employment) in commerce	Agriculture employment share	Manufacturing employment) share	Commerce employment share
distance to LCP	0.006	-0.020	-0.016	0.006*	-0.005*	-0.001
	(0.018)	(0.018)	(0.025)	(0.003)	(0.002)	(0.001)
Observations	102	102	102	102	102	102
County FE	Yes	Yes	Yes	Yes	Yes	Yes
			Panel B			
	log(Total employment)	Population	Public expenditure	Taxable property	Iroquois trail	No. of Schooles
distance to LCP	-0.003	-0.004	22.935	1.615	0.000	0.291**
	(0.016)	(0.014)	(16.209)	(8.979)	(0.002)	(0.137)
Observations	102	100	78	97	32	90
County FE	Yes	Yes	Yes	Yes	Yes	Yes
			Panel C			
	No. of children	No. of churches	No. of distilleries	No. of grist mills	No. of saw mills	No. of fulling mills
distance to LCP	0.001	-0.002	0.015	0.056	0.036	-0.003
	(0.002)	(0.021)	(0.041)	(0.079)	(0.029)	(0.009)
Observations	40	132	132	132	131	130
County FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The constant term is estimated but not shown in the table. Standard errors in parentheses are clustered on counties.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7 presents OLS estimates derived from regressing outcome variables and a variety of town characteristics on the distance to the nearest LCP, conditioning on county fixed effects. Town observed characteristics are well balanced across the instrument, the distance to a LCP, with two exceptions. Towns situated closer to the LCPs reported a lower employment share in agriculture (column 4) and a higher employment share in manufacturing (column 5). However, the coefficients, 0.006 and -0.005 are relatively small when considering that initial employment shares in agriculture and manufacturing stood at 75% and 22%, respectively, in 1820. Given these findings, the instrument is presumably valid conditional on county fixed effects.

IV Estimates. Table 8 displays the IV estimates. Both the sectoral transition effects in employment and the long-term population growth effects maintain their significance and direction. Quantitatively, canal towns recorded 7.6% less agriculture employment share, 2.6% more commerce employment share, 56% faster employment growth in manufacturing (albeit not statistically significant), 224.5% faster employment growth in commerce, 72.5% faster total employment growth, and 112.8% faster population growth when compared with noncanal towns. The magnitudes of canal's effect on all dependent variables are larger than the DiD estimates.

Table 8: IV estimates of effects of canal operation on sectoral transition in employment and population

	(1) Δ Agriculture employment share	(2) Δ Manufacturing employment share	(3) Δ Commerce employment share	(4) Δ log(employment) in agriculture	(5) Δ log(employment) in manufacturing	(6) Δ log(employment) in commerce	(7) Δ log(employment) total	(8) Δ log(Population)
$\hat{Canal}$	-0.076**	-0.001	0.026***	0.100	0.377	1.026**	0.284**	0.318***
	(0.029)	(0.030)	(0.008)	(0.130)	(0.223)	(0.470)	(0.131)	(0.104)
Observations	101	101	101	101	101	101	101	99
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Standard errors in parentheses are clustered on counties.* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

7 Mechanism

In this section, I elucidate the mechanism of sectoral transition in employment. The development of small-scale manufacturing that processed agricultural products and the expansion of commercial activities brought about by lower transportation costs and easier access to distant markets, attracted individuals to canal towns. There, they established manufacturing mills, worked as manufacturing workers, and shipped processed agricultural staples along the canal as merchants. As noted by Ransom (1971), manufacturing developed in response to eastern demands for western agriculture staples. This small-scale manufacturing expanded on the western region's agricultural strengths. The canal played a pivotal role in fostering regional economic specialization by integrating western and eastern New York State. Next, I present evidence considering the types of processing mills, the associated agricultural goods they process, and the types of good transported through the Erie Canal.

As evidenced by Table 2, the diversity of manufacturing establishments in canal towns offers insights. On average, canal towns hosted 1.5 distilleries, 4 grist mills, 7.5 sawmills, 0.5 oil mills, 2 fulling mills, 2 carding machines, 2 asheries, 0.5 cotton-woollen factories, 0.05 tannery, 0.5 trip hammers, 1,735,000 yards of cloth made in the year 1824. Processing agricultural staples such as wheat, corn, and meat into higher value-per-ton goods can mitigate their otherwise elevated transportation costs. For instance, wheat and corn were transformed into flour and whiskey, respectively. Livestock such as swine and cattle were processed into meat products for eastern transport via the Erie Canal. Similarly, deer skins and raw hides underwent tanning processes, and sheep-derived products like wool were manufactured in cotton-wool factories. Raw timber was processed into lumber, staves, shingles, and other wood products in sawmills. These manufacturing endeavors, driven by the abundant local agricultural produce, facilitated the shipment of processed goods to eastern New York State through the Erie Canal.

Table 9 underscores the nature of goods transported via the Erie Canal, giving strong evidence to the processing manufacturing. In 1829, raw materials and essential commodities, including 311,256 feet of lumber, 3,640 bushels of wheat, 4,335 barrels of flour, 4,754 barrels of provisions (comprising pork and beef), and 1,705 tons of agricultural ashes, ranked as the top five items cleared at Buffalo for eastward transit on the Erie Canal. By 1835, the volume of these goods saw a marked increase: provisions doubled, agricultural ashes quadrupled, lumber quantities rose sevenfold, flour volumes escalated by a factor of 23, and wheat shipments swelled by 46 times. A similar composition of goods was documented in Utica from 1830 to 1832.

8 Conclusion

During the early nineteenth-century New York State witnessed a western expansion where people moved into regions such as the Mohawk River Valley and the Figure Lake Region. Concurrently, the state's economy evolved from primarily agricultural to commercial, characterized by the rise of processing manufacturing. The construction of the state-led Erie Canal

Table 9: Property cleared at Buffalo and passing east on the Erie canal from 1829 to 1835

Articles	Unit	1829	1830	1831	1832	1833	1834	1835
Flour	barrels	4335	31810	62968	21932	78666	79324	100833
Provisions(pork & beef)	barrels	4754	6675	5668	5159	4273	14590	8160
Oil	barrels	214	802	1420	44	43	221	
Wheat	bushels	3640	149219	186148	100761	114337	111798	168012
Ashes	tons	1705	2713	2502	2110	2118	1655	7304
Tobacco	tons	32	62	222	386	535	1008	1765
Hemp	tons	22	20	70	29	17	5	0.5
Pig iron	tons	235	419	409	760	1167	1128	997
Castings	tons	241		422	468	757	689	768
Household furniture	tons	42	58	69	88	134	145	355
Furs	tons	86	82	96	107	101	154	136
Lumber	feet	311256	136499	184639	251504	331140	439643	2087024
Staves	meter	510	464	568	523	699	2400	2694
Fish	barrels		851	150	276	279	346	732
Whiskey	barrels	149	4182	3750	2208	2485	1347	614
Butter & lard	tons	70	174	205	394	449	119	503
Cheese	tons	68	122	127	74	95	138	34
Wool	tons			66	22	75	73	93
Deer skins & raw hides	tons					110	141	207
Grind-stones	tons	35	39	124	110	139	126	135
Lead	tons		41	9				
Sundries	tons							242
Shingles	meter							74062
Timber	feet							61430
Corn	bushels							12193

Notes: Data are from the Annual Report of the Canal Commissioners.

offers a quasi-experimental context, enabling this study to determine causal implications of canals on sectoral transition in employment and long-term growth population. Utilizing a difference-in-differences design, I compare towns located close to the canal relative to those were not close to the canal before and after the operation of the canal. I find that towns located close to the canal experienced more rapid population growth, lower employment shares in agriculture, higher employment shares in commerce, and faster employment growth in manufacturing. I argue that the mechanism underlying sectoral transition and long-term growth was the development of small-scale manufacturing that processed agricultural products and the expansion of commercial activities brought about by lower transportation costs and easier access to distant markets. A key contribution of this paper is to provide the first causal estimates of the economic impact of canals on the sectoral transition and long-term economic growth.

To thoroughly assess the impact of canals on early nineteenth-century New York State, a deeper examination of the political economy surrounding canal construction is warranted. The Erie Canal faced criticism from political adversaries, often derogatorily labeled as "Clinton's Folly" or "Clinton's Big Ditch," suggesting corruption and favoritism. Nevertheless, it is vital to acknowledge that the canal's realization was contingent upon a political coalition, orchestrated by DeWitt Clinton—who took office as Governor of New York State the same year the Erie Canal's construction commenced. This coalition, driven by a shared vision of the canal's potential, transcended pure political allegiances and overcame regional differences. Given the long-term economic growth highlighted in this study, an evaluation weighing the costs and benefits of the Erie Canal is essential, especially when accounting for potential inefficiencies tied to government intervention. This necessitates a more comprehensive exploration of the political economy associated with canal financing, as highlighted in sources such as (Goodrich et al., 1961; Wallis, 2003; Engerman and Sokoloff, 2004).

References

Albion, R. G. (1939). *The Rise of New York Port: 1815-1860*. New York and London: Charles Scribner's Sons.

Américo, P. (2022). The industrialization paths: Railroads and structural transformation in brazil 1872-1950. *Available at SSRN: https://ssrn.com/abstract=4057774 or http://dx.doi.org/10.2139/ssrn.4057774*.

Angrist, J. D. and J.-S. Pischke (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Atack, J., F. Bateman, M. Haines, and R. A. Margo (2010). Did railroads induce or follow economic growth? urbanization and population growth in the american midwest, 1850-1860. *Social Science History*.

Atack, J., M. R. Haines, and R. A. Margo (2008). Railroads and the rise of the factory: Evidence for the united states, 1850-70. *NBER Working Paper*.

Atack, J., R. A. Margo, and P. W. Rhode (2022). Industrialization and urbanization in nineteenth century america. *Regional Science and Urban Economics*.

Berger, T. (2019). Railroads and rural industrialization: evidence from a historical policy experiment. *Explorations in Economic History*.

Bleakley, H. and J. Lin (2012). Portage and path dependence. *The Quarterly Journal of Economics 127*(2).

Bodenhorn, H. and D. Cuberes (2018). Finance and urbanization in early nineteenth-century new york. *Journal of Urban Economics*.

Bogart, D., X. You, E. J. Alvarez-Palau, M. Satchell, and L. Shaw-Taylor (2022). Railways, divergence, and structural change in 19th century england and wales. *Journal of Urban Economics*.

Canal Commissioners (1817). *Annual Report of the Canal Commissioners*. Columbia University.

Correia, S. (2015). Singletons, cluster-robust standard erroes and fixed effects: A bad mix. *Technical Note, Duke University*.

Donaldson, D. and R. Hornbeck (2016). Railroads and american economic grwoth: A "market access" approach. *The Quarterly Journal of Economics*.

Ellis, D. M. (1946). *Landlords and Farmers in the Hudson-Mohawk Region: 1790-1850*. Ithca, New York: Cornell University Press.

Engerman, S. L. and K. L. Sokoloff (2004, December). Digging the dirt at public expense: Governance in the building of the erie canal and other public works. *NBER Working Paper*.

Faber, B. (2014). Trade integration, market size, and industrialization: Evidence from china's national trunk highway system. *The Review of Economic Studies 81*(3).

Fishlow, A. (1965). *Amercian Railroads and the Transformation of the Ante-Bellum Economy*. Cambridge, Massachusetts: Harvard University Press.

Flater, D. (2022). Gravity, spatial interaction, movement, and centrality modeling tools. `https://www.arcgis.com/home/item.html?id=0cdd43afa32b417681589e38ebd54097`.

Fogel, R. W. (1964). *Railroads and American Economic Growth: Essays in Econometric History*. Baltimore: Johns Hopkins Press.

Goodrich, C. (1960). *Government Promotion of American Canals and Railroads 1800-1890*. New York: Columbia University Press.

Goodrich, C., J. Rubin, H. J. Cranmer, and H. H. Segal (1961). *Canals and American Economic Development*. New York and London: Columbia University Press.

Kirkland, E. C. (1934). *A History of American Economic Life*. New York: F. S. Crofts & Co.

Laws of the State of New-York (1828). *Laws of the State of New-York, Passed at the Fifty-First Session of the Legislature*.

McClelland, P. D. (1968). Railroads, american growth, and the new economic hisotry: A critique. *The Journal of Economic History 28*(1).

Miller, N. (1962). *The Enterprise of A Free People: Aspects of Economics Development in New York State during the Canal Period: 1792-1838*. Ithca, New York: Cornell University Press.

Niemi, A. W. (1970). A further look at interregional canals and economi specialization: 1820-1840. *Explorations in Economic History 7*(4).

Niemi, A. W. (1972). A closer look at canals and western manufacturing in the canal era: A reply. *Explorations in Economic History 9*(4).

Nunn, N. and D. Puga (2012). Ruggedness: The blessing of bad geography in africa. *The Review of Economics and Statistics 94*(1).

Poor, H. V. (1860). *History of the Railroads and Canals of the United States of America*. J.H. Schultz Company.

Ransom, R. L. (1967). Interregional canals and economic specialization in the antebellum united states. *Explorations in the Entrepreneurial History 5*(1).

Ransom, R. L. (1971). A closer look at canals and western manufacturing. *Explorations in Economic History 8*(4).

Ransom, R. L. (1972). A rebuttal. *Explorations in Economic History 9*(4).

Rubin, J. (1961). An innovating public improvement: The erie canal. In C. Goodrich (Ed.), *Canals and American Economic Development*. New York and London: Columbia University Press.

Sellers, C. (1991). *The Market Revolution: Jacksonian America, 1815-1846*. New York: Oxford University Press.

Sokoloff, K. L. (1988, December). Inventive activity in early industrial america: Evidence from patent records, 1790-1846. *The Journal of Economic History* (4).

Spafford, H. G. (1981). *A Gazetteer of the State of New York*. Heart of the Lakes Publishing.

Stock, J. and M. Yogo (2005). *Identification and Inference for Econometric Models*, Chapter Testing for Weak Instruments in Linear IV Regression, pp. 80–108. New York: Cambridge University Press.

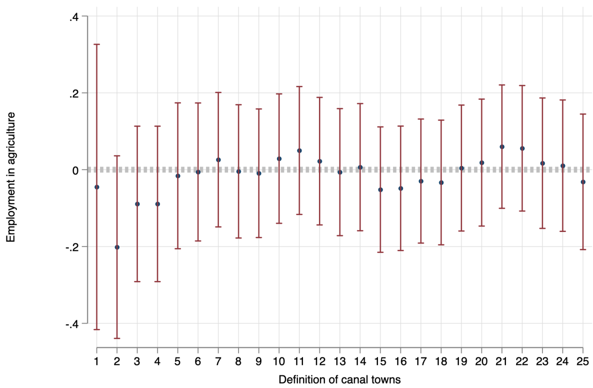
U.S. Department of Agriculture (2023). Grain transportation report. `http://dx.doi.org/10.9752/TS056.08-31-2023`.

U.S. Geological Survey (1995). Index of digital elevation models (dem), new york. `https://cugir.library.cornell.edu/catalog/cugir-008186?id=23`.

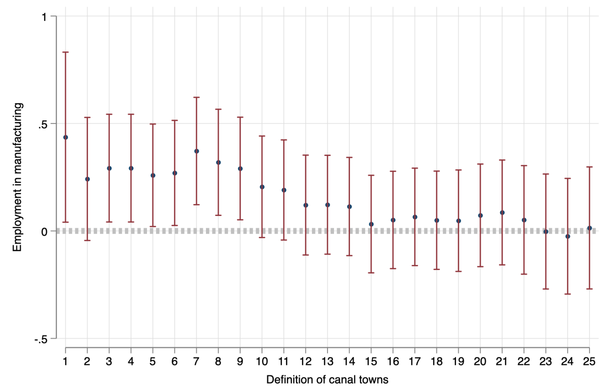
Wallis, J. J. (2003, July). The property tax as a coordinating device: Financing indiana's mammoth internal improment system, 1835-1842. *Explorations in Economic History 40*(3), 223–250.

Whitford, N. E. (1906). *History of the Canal System of the State of New York*, Volume 1. Albany: State Legislative Printer.

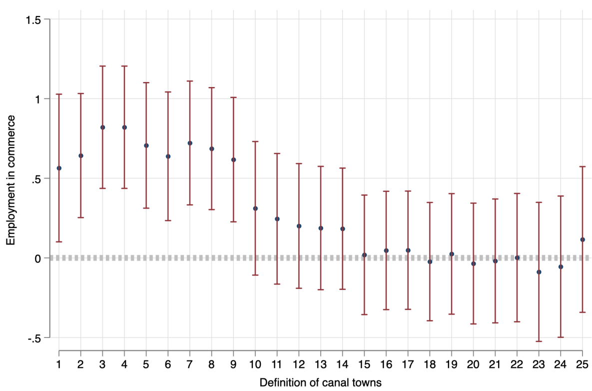
Appendices



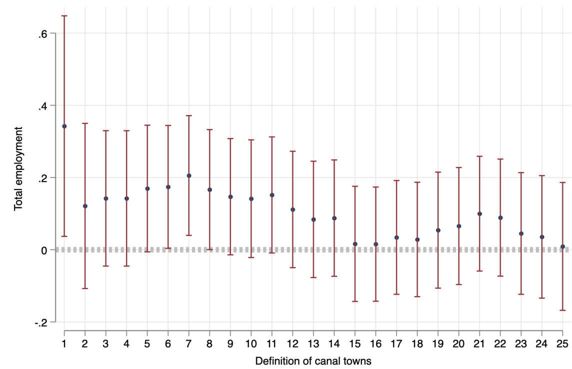
(a) Employment in agriculture



(b) Employment in manufacturing



(c) Employment in commerce



(d) Total employment

Figure 6

Notes: The vertical axis is the employment share in commerce. The horizontal axis is the measure of $canal_i$ by using different distance from the towns to the canal. The distance unit is kilometer. Each point is estimated by equation (1) with 95% confidence interval. Standard errors are clustered on towns.

Table 10: Balance test of town characteristics before 1825 between canal towns (0-5km) and noncanal towns (10-15km)

	Treated towns	Obs.	Control towns	Obs.	Diff.	p -value
Population	2750.09	47	2858.07	27	107.99	0.75
Total Employment	647.92	48	701.37	27	53.45	0.51
Employment in agriculture	483.58	48	578.56	27	94.97	0.16
Employment in manufacturing	144.98	48	116.00	27	-28.98	0.37
Employment in commerce	19.35	48	6.81	27	-12.54	0.26
Agriculture employment share	0.75	48	0.82	27	0.07	0.13
Manufacturing employment share	0.22	48	0.17	27	-0.05	0.16
Commerce employment share	0.03	48	0.01	27	-0.02	0.12
Public expenditure	301.03	35	379.64	22	78.61	0.23
Taxable property(thousand dollar)	410.34	43	325.71	27	-84.63	0.21
Iroquois trail(0/1)	0.95	21	1.00	4	0.05	0.67
Turnpike(0/1)	0.50	46	0.54	28	0.04	0.77
County seat(0/1)	0.17	71	0.07	45	-0.10	0.11
No. of Electors	516.30	44	530.08	26	13.78	0.82
No. of schools	11.83	41	15.84	25	4.01**	0.02
No. of students	660.28	36	897.86	22	237.59*	0.05
No. of churches	2.83	12	3.60	10	0.77	0.43
No. of distilleries	1.42	45	2.32	28	0.90*	0.07
No. of grist mills	3.60	45	3.82	28	0.22	0.74
No. of saw mills	7.02	45	9.25	28	2.23	0.12
No. of oil mills	0.47	45	0.25	28	-0.22	0.27
No. of fulling mills	1.80	44	2.71	28	0.92	0.10
No. of asheries	2.29	45	3.07	28	0.78	0.30
No. of cotton woollen factories	0.41	44	0.61	28	0.20	0.53
No. of tanneries	0.07	44	0.21	28	0.15	0.45
Elevation(feet)	472.62	71	713.33	45	240.71***	0.00
Custom(0/1)	0.03	71	0.02	45	-0.01	0.85
No. of children age 5-15	733.54	35	765.86	22	32.32	0.76
No. of carding machine	1.73	44	3.29	28	1.56*	0.06
No. of trip hammers	0.25	44	0.43	28	0.18	0.37
No. of cloth made(thousand yards)	15.19	43	21.14	27	5.95*	0.05

Notes: Data are from the federal and state census in 1820 and the Gazetteer of the State of New York 1824. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

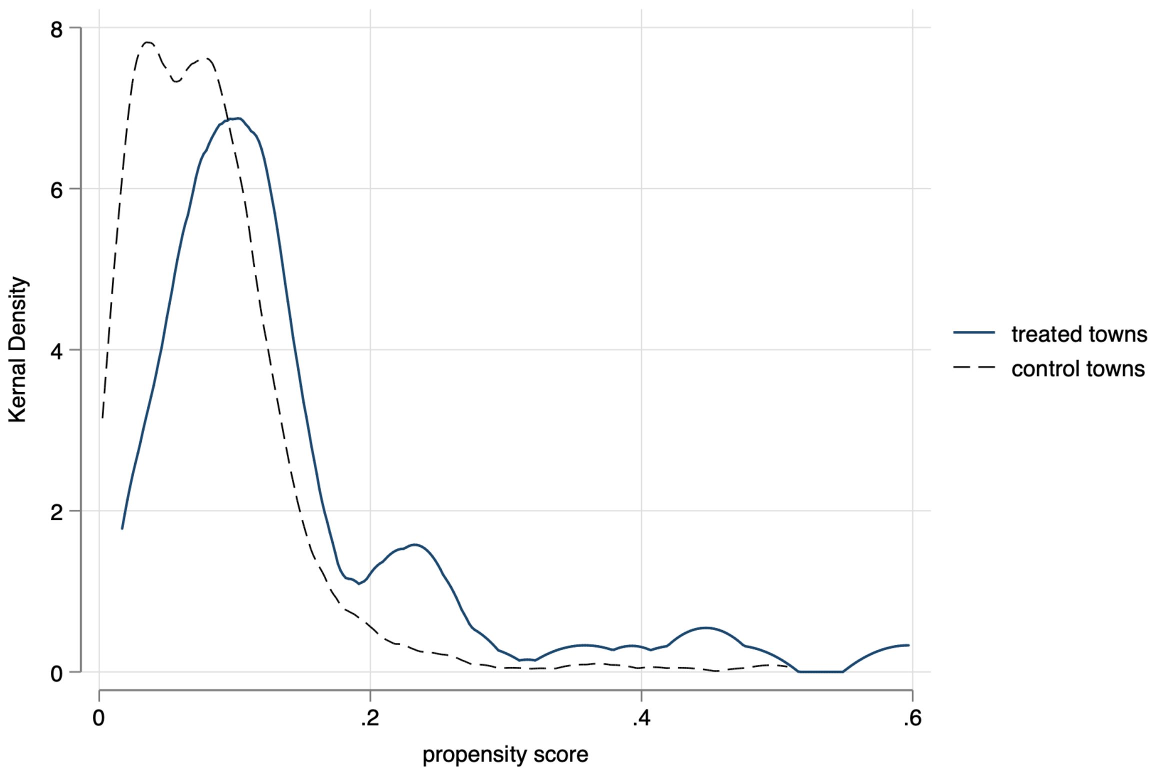


Figure 7: Density of propensity score of treated and control towns