The Commercial Path: The Erie Canal and Transformation of Early Nineteenth-Century New York

Youwei Xing*

January 31, 2023

Abstract

This article studies the impact of the Erie Canal on growth and structural transformation in nineteenth-century New York State. A network of least cost paths is created to exploit the feature that the state-owned canal traversed towns that were not targeted by planners. The IV-DiD estimates show areas plausibly randomly traversed by the canal experienced more rapid population growth, lower employment share in agriculture, and higher employment share in commerce over the next 20 years. I show that the mechanism of structural transformation was increased market orientation brought by lower transportation costs and easier access to distant markets.

*John E. Walker Department of Economics, Clemson University, Clemson, 29634, South Carolina, USA; email: youweix@g.clemson.edu. I would like to thank Howard Bodenhorn for his suggestions and guidance. I am grateful for comments and suggestions by participants in Clemson Public Workshop.

1 Introduction

The canal-building boom predated the railroad boom by a generation or more and canals were similarly as transformative as railroads. Relative to railroads, canals have not received their fair share of academic attention. The increased interregional trade and the reduction of transportation costs caused by the railroad have been well documented by economic historians (Fishlow, 1965; Fogel, 1964). The reduction in transportation costs provided by canals, as compared to wagon haulage, was more substantial than the subsequent differential between railroads and canals (Goodrich et al., 1961).

Previous studies look at the effects of canals on market integration (Albion, 1939), government promotion of canals (Goodrich, 1960), public finance of canals (Wallis, 2003), and the association of canals with inventive activity (Sokoloff, 1988). However, the causal effects of canal operation on structural transformation of local economies have not been studied in depth.

Identifying the causal effects of canal on economic changes is challenging. First, there is a potential selection problem, as canals may have been constructed in populated areas with active commercial or manufacturing activities. Second, previous research has discussed the effects of canals on regional specialization using county-level data (Ransom, 1967; Niemi, 1970), but counties are too broad an analysis when investigating the canal effects in the early nineteenth-century given that the complementary land transportation is mainly wagon haulage.[1]

This paper solves these potential problems and provides a clearer picture of how the Erie Canal impacted the transformation of early nineteenth-century New York State.[2] For identification, I use the rollout of the Erie Canal as a quasi-natural experiment.

[1] The Erie Canal passed southern part of Herkimer County which is narrow in direction of north and south. 3/4 of the county gets access to the Erie Canal with high transportation costs. In the early nineteenth-century a wagon traveled 2 to 3 miles per hour depending upon weather, roadway conditions, and the health of the travelers.

[2] "The ton-mileage upon the Erie was always greater before 1860 than upon any other single transportation route."–Henry V. Poor

This setting offers a useful feature to identify the impacts of the canal. Unlike private transportation companies whose priority is to maximize profits by constructing canals in populated places with active commercial activities, the design and construction of the Erie Canal were led by the state. The goal was to connect several consequential cities from the west to the east. Consequently, many non-targeted cities and towns gained access to the canal, which alleviates some endogeneity concerns. To solve the data issue, I use a dataset on population, manufacturing activity, employment in three broad sectors of the economy, and other socioeconomic characteristics for towns and cities before and after canal operation took place (Bodenhorn and Cuberes, 2018).

I begin by using a standard difference-in-differences (DiD) approach and I estimate changes in sectoral employment, population, and structural transformation in employment between 1820 and 1840 among towns and cities located close to the Erie Canal relative to those that were not proximate to the canal.[3] The DiD estimates reveal a lower employment share in agriculture and higher employment shares in commerce in areas neighboring the canal. These effects are highly localized. Effects of structural transformation in employment attenuate in towns once they are located more than 9 km away from the canal. Thus, an important question is whether these localized impacts are offset by a reallocation of manufacturing and commerce workers from nearby areas. Yet, there is limited evidence of negative spillovers onto nearby areas in flexible specifications. Estimates also remain similar when comparing treated areas to the same regional units but excludes areas immediately adjoining the treated areas.

Further estimates plus data in manufacturing establishments and shipment show the mechanism of structural transformation was increased market orientation brought by lower transportation costs and easier access to distant markets where people moved into canal-towns, established manufacturing mills, worked as manufacturing worker, and

[3] The alternative transportation, wagon transportation, rate is 6 cents to 12 cents per ton-kilometer in early nineteenth-century (McClelland, 1968).

shipped processed agricultural staples along the canal as businessmen. People did not stop being or becoming farmers. The lower employment share in agriculture is not driven by the decreasing total employment in agriculture but because of the increasing population and more rapid growth in other sectors.

Overall, these DiD estimates indicate more rapid population growth and structural transformation of employment over a 20-year horizon in areas that were penetrated by the canal. However, still a concern could be that the estimates are potentially biased if the canal traversed towns with better or worse growth prospects. To address this issue, I follow the approach of Berger (2019) and Bogart et al. (2022) and use an instrumental variable (IV) strategy that exploits the fact that the canal was explicitly to be built along the shortest route between the endpoints of the canal network. I construct the Least Cost Paths (LCPs) that correspond to the canal that state planners would have built if the sole objective had been cost minimization subject to connecting a few targeted cities along the route. Those untargeted towns and cities are traversed by LCPs only because of their geography, which places them on a cost-minimizing route. 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 not because of their other economic characteristics. Nested with the DiD design in the beginning, the IV strategy improves early DiD design by using the exogenous variation of the untargeted areas. The IV-DiD estimates reveal large increases in population, employment in commerce, and structural transformation of employment in areas traversed by the canal. IV-DiD estimates are larger in magnitude than the corresponding DiD estimates, showing the DiD estimates are downward biased.

A key contribution of this paper is to identify the causal effect of canals on the transformation of employment and long-term economic growth. Because I study the effects of canals on sectoral employment and population to understand the mechanism behind the structural transformation process better, and those variables are used by scholars to measure other variables of interest, this paper is also related to the literature on canals' other economic impacts: accelerating western settlement, market integration, and regional specialization (Whitford, 1906; Albion, 1939; Goodrich et al., 1961; Ransom, 1967; Niemi, 1970)

2 Background

Built between 1817 and 1825 and operated by the state, the Erie Canal, 364 miles in length, linked the Hudson River to Lake Erie. Unlike the railroads that would come later, the Erie Canal was a product of state intervention. One outstanding characteristic of the economy of New York State during this period is subsidies which emphasizes the state's entrepreneurial role in economic activity in terms of state loans to manufacturing corporates, tax exemption of certain manufactories, and contributing funds to agricultural societies (Miller, 1962). Canal construction and operation would entail only a change in form and extent of prevailing government interference.

There are two advantages for the state to construct the canal over private companies. The first is the financial issue faced by private companies. Before the construction of the Erie Canal was taken over by the state, two private companies tried to construct and operate limited sections of the canal but both failed due to financial constraints. The Western Inland Lock Navigation Company and the Northern Inland Lock Navigation Company were granted to sell a maximum of 1,000 shares for an initial payment of $25 a share and a state gift of $12,500 to complete the canal work in 1792. Later, the state subscribed to 200 of the shares of each company and loaned the Western Company $37,500 when the companies faced financial difficulties. But those tries did not save the companies. Different from private companies, there was a chance for New York State to get federal aid to construct the canal when president Jefferson suggested the Treasury surplus be devoted to internal improvement. But this hope for federal aid broke because of the competition among rival states and the war of 1812.

Second, compared with private companies, the state is easier to reconcile political competition and regional differences. Although the western state wanted the canal, the Long Island counties and mid-Hudson valley farmers were opposed to the canal because of the threat of extensive Midwestern and upstate produce entering New York City. Southern counties along the Pennsylvania border complained that the canal would be of no use to them because the benefits of the canal would go to residents of only a limited area. Also, there were different political groups that tried to take credit for the canal. One group was led by DeWitt Clinton, who took office as Governor of New York State in 1817 and served until 1822, and consisted of wealthy Federalist landholders in the west, great merchants in the cities, western farmers, and Irish immigrants who expected the canal would increase their land value and boost in-state commerce (Rubin, 1961). The opposite group was led by Martin Van Buren and consisted of the majority of Democrats who were in control of New York City and cities in the Hudson River Valley. To reconcile the political and regional differences, DeWitt Clinton built a political coalition that wanted the canal to transcend pure political allegiance and overcame regional differences in priorities.

When it came to building the canal, given the minimum expectation of connecting several major cities like Buffalo, Syracuse, Utica, and Albany, engineering feasibility rather than local commerce employment was the main priority before breaking ground. Surveys were conducted by canal commissioners[4] and engineers[5]. The canal commissioners went by land or by boat up the Mohawk River to Rome and thence down to

[4]They were Gouverneur Morris, De Witt Clinton, Stephen Van Rensselaer, Simeon De Witt, William North, Thomas Eddy, and Peter B. Porter in 1811. In 1816, the new commissioners are De Witt Clinton, Stephen Van Rensselaer, Samuel Young, Joseph Ellicott, and Myron Holley.

[5]James Geddes and Benjamin Wright, father of American Civil Engineering, for examples.

Oswego, and up from Three River Point to Geneva, where the boats were sold, the party proceeding by carriage to the Niagara (Kirkland, 1934). The surveys determined whether a sufficient supply of water for all seasons could be obtained without diminishing the supply of water to hydraulic works, the feasibility of constructing a stone lock, a culvert, and an aqueduct, and any other plans.[6] The majority of contracts of excavation, embankment, and stone locks were assigned to local farmers, mechanics, merchants, and professional men, residing in the vicinity of the line. Three-fourths of all the laborers involved were born locally. Machinery that had been used in the heavy business of grubbing and clearing could be used to build the canal. For example, by means of an endless crew, connected with a roller, a cable, a wheel, and a crank, one man was able to bring down a tree of the largest size. Also, the narrow plow was invented for cutting small roots and fibers which overspread the surface.[7]

2.1 New York State before the Canal

The western expansion and trade developed in-state and internationally are the two demand factors for internal improvement. Between 1790 and 1814, the population of New York City more than doubled from 54,182 to 123,980 and it emerged as the nation's chief port as measured by the number of exports and imports loaded and unloaded at its wharves. In the year of 1821, New York port accounted for approximately 20% of U.S. total exports and 37% of total imports. Tonnage entered from foreign countries and cleared for foreign counties accounted for 20% and 19% of total U.S. respectively (Albion, 1939). This commercial expansion was not limited to the city. Farther north on the Hudson River and Mohawk River were emerging towns with local hinterlands that served as markets or as transshipment points in the river trade.

The rapid western expansion, with the population of the state tripling from 339,582

[6] Laws of the State of New-York, 1828.

[7] State of New York. *Annual Report of the Canal Commissioners.* Columbia University.

to 1,038,058 during this time, heightened the market consciousness of rural New York, and signified that the state as a whole was moving toward a higher level of economic complexity and integration. Population concentrated on Long Island and Hudson River Valley in 1790 with several towns, like Palatine, Canajoharie, German Flattes, Herkimer, and Whitestown, growing along the Mohawk river. Only ten years later, in 1800, there were more towns emerging in the Mohawk River Valley. Rome with a population of 1479, one of the important stops of the Erie Canal later, was emerging during that 10 years. People kept moving to the western towns around modern Syracuse and towns around the Finger Lakes Region. By 1810, people reached Buffalo, and by 1820, the population sprawled the state (see Figure 1). Populated towns were concentrated in three areas with water navigation: Hudson River Valley, Mohawk River Valley, and Finger Lakes Region. For merchants and farmers of these towns and surrounding areas, the need for improved transportation to maintain contact between the west-expanding line of commercial agriculture and the river towns increased steadily as time passed.

2.2 New York State after the Canal

The period after 1825 was referred to as the agricultural revolution where the old life of self-sufficiency was vanishing and all but the most isolated farmers were raising foodstuffs for sale by 1850 (Ellis, 1946). In the navigable season from April to December, thousands of boats passed on the Erie Canal. For example, The total number of boats and other floats passed at lock No. 26, 3 miles west of Schenectady, increased from 6,166 in 1824 to 35,981 in 1854, with almost even distribution among each month.

Flood of western produce poured into the Hudson-Mohawk valley by the way of the Erie Canal. Raw materials and basic resources, such as 311,256 feet of lumber, 3,640 bushels of wheat, 4,335 barrels of flour, 4,754 barrels of provisions (pork and beef), and 1,705 tons of ashes, which listed top 5 among properties cleared at Buffalo were

(a) (b)

(c) (d)

Figure 1: Population changes from 1790 to 1820

passing east on Erie Canal in 1829. By 1835, double amounts of provision, 4 times more ashes, 7 times more lumber, 23 times more flour, and 46 times more wheat were passing east on the Erie Canal from Buffalo (see Figure 2). Similar composition of articles passed Utica was observed from 1830 to 1832. On the other hand, 7,151 tons of Merchandize, 935 tons of furniture and mechanics' tools, and 65,431 barrels of salt were shipped to Buffalo from the east on the Erie Canal in 1830. By 1835, 1.2 times more salt, 4 times more merchandise, and 9 times more furniture and mechanics' tools arrived at Buffalo from the east. Properties that arrived at Buffalo were not only consumed

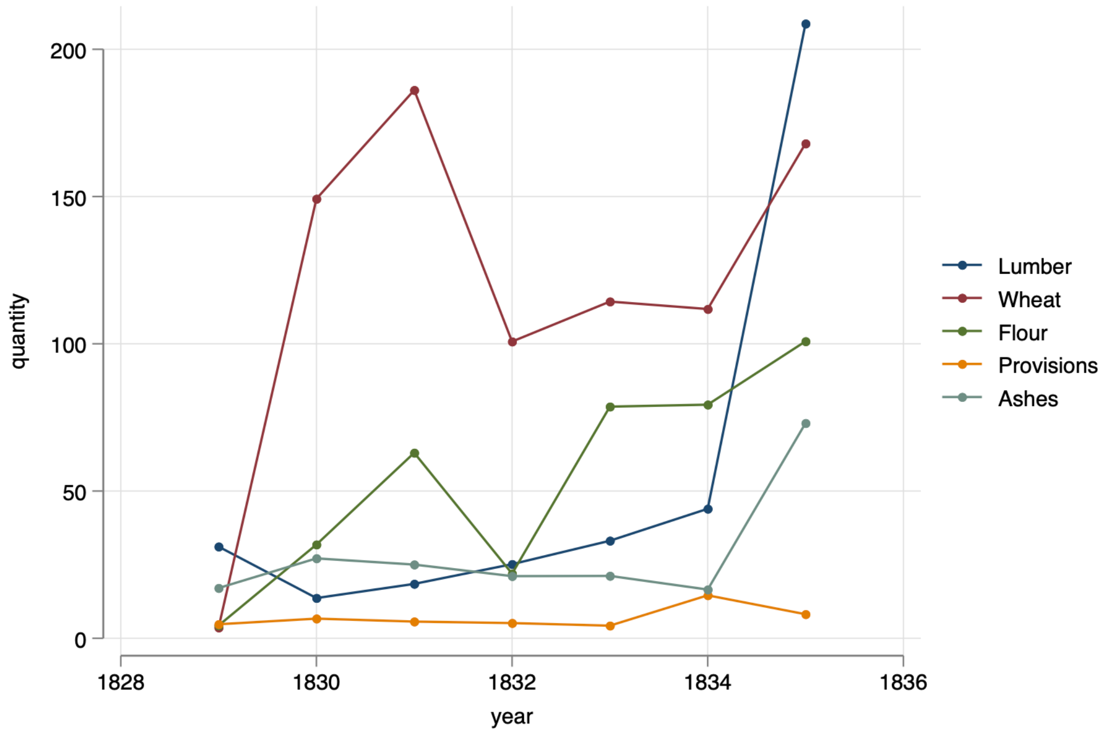


Figure 2: Top 5 property cleared at Buffalo and passing east on the Erie Canal from 1829 to 1835

by those in New York State but also destined out from New York State to places like Upper Canada, Pennsylvania, Ohio, Michigan, Indiana, Illinois, Virginia, Kentucky, Tennessee, Missouri, and Alabama.[8] The wide variety of articles passing on the Erie Canal, including 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. portrays a lifelike early 19th century New York state.

3 Data

To analyze the impact of the Erie Canal on population and structural transformation in employment, I use a dataset on population, manufacturing activity, employment in

[8]State of New York. *Annual Report of the Canal Commissioners.* Columbia University.

three broad sectors of the economy, and other socioeconomic characteristics for towns and cities before and after canal operation took place from Bodenhorn and Cuberes (2018). Then combine the dataset with Geographic Information System (GIS) data to capture differences in canal access across towns and cities. The employment data in the dataset are from the federal census. The federal government only canvasses employment in the years 1820 and 1840. The 1820 federal census records employment in agriculture, commerce, and manufacturing. The 1840 federal census also includes employment in mining, ocean navigation, inland navigation, and professions. Population data in the dataset are from the federal census and state census. The federal government canvasses the population at decade intervals in years ending in zero. Beginning in 1825, New York State conducted statewide canvasses of its population at decade intervals in years ending in five; it also conducted a census in 1814.

Town population and employment data are supplemented with information on towns' topographical, geographical, social, and economic data. The topographical and geographical data include its elevation above sea level, its access to Iroquois Trail, and whether it serves as a county seat (i.e., a regional governmental center, which typically included the county courthouse, the county jail, and other administrative offices). Social data include the number of electors, the number of churches, the number of schools, and the number of youth attending school in 1824. Other economic data except population and employment include the amount of public money spent, the value of taxable properties, and the number of a variety of manufacturing establishments like distilleries, grist mills, fulling mills, tanneries, etc. These social and economic data that only consist of the year of 1824 are from the Gazetteer of the State of New York 1824 (Spafford, 1981). Data of articles traded through the Erie Canal are digitalized from the Annual Report of the Canal Commissioners.

The New York State city and town boundaries data are from the NYS Office of

Information Technology Services GIS Program Office (GPO). Since there is no historical New York State civil boundaries data available, I use the year 2021' boundaries as the base map. City and town boundaries do change from the study period to now. There were 985 towns in nineteenth-century New York State. There are 62 cities and 932 towns in 2021. One of the trends of township change is that new towns are established from existing towns from the 19th century. To make the map consistent over time, I add data from new towns back to existing towns. For example, Middletown in Orange County was part of Newburgh in Orange County before 1888; then, I add the population of Middletown to Newburgh. Another trend is that there are places incorporated as a city within original towns. For example, Amsterdam city in Montgomery County was incorporated as a city in 1885. To deal with these cases, I put the historical population data into the modern town. City and town boundaries are adjusted according to the Gazetteer of the State of New York 1824 and Wikipedia.

Using GIS software, I calculate the Euclidean distance between towns and the Erie Canal. The City hall or town hall is chosen as a point to represent a city or a town. Distance from cities or towns to the Erie Canal is either the perpendicular or the closest vertex. The historical canal path, Mohawk river, Hudson River, and New York State Thruway shapefiles are from New York State GIS Clearinghouse or from ArcGIS online portal. To construct the instrument variable, I use the Digital Elevation Models with 10 meters pixel resolution from the U.S. Geological Survey.

4 Empirical Strategy

4.1 DiD Strategy

To examine the impact of the Erie Canal, the main empirical analysis is based on a difference-in-differences approach that compares variables of interest in areas that are physically close to the canal (treated towns) and towns that share similar characteristics but physically far away from the canal (control towns) in the period before and after the year 1825:

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

where Y_{it} either corresponds to population or ln employment, or the share of employment in agriculture, manufacturing, and commerce of town i in year t. $Canal_i$ is a distance measure to the canal, which corresponds to an indicator taking the value 1 for towns located within 9 km of the canal and 0 for other towns. $Post_t$ is the indicator variable equal to 1 after year 1825 and 0 otherwise. θ_i is the town fixed effect, capturing any county invariant characteristics like elevation, terrain ruggedness, whether the town is county seat, and the access to the Iroquois Trail. ϕ_t is a year fixed effect to model possible nonlinearities common for all towns in the evolution of dependent variable not captured by linear time trends. If it is not emphasized, the regression is always clustered at the town level to avoid understating the standard errors and drawing unwarranted inferences. α is a constant term common for all towns. The success of the construction of the Erie Canal was determined in large part by politics (Rubin, 1961) and county seats served as governmental and political centers in which population may concentrate. Bleakley and Lin (2012) show that North American cities tended to locate at native portage points and the Iroquois Trail was an ancient, Native-American trade route used by colonial military forces in troop deployments that crossed several rivers. Elevation posed a veritable engineering challenge for canal construction, and also may have a direct impact on economic development (Nunn and Puga, 2012).

Under the DiD approach, the identification assumptions are 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 the interference

Table 1: Comparison of town characteristics before 1825 between canal-towns (9km) and noncanal-towns (20km)

	Treated towns	Obs.	Control towns	Obs.	Diff.	$p - value$
population(N)	2827.25	64	2787.71	68	-39.54	0.90
ln empl_agri	6.06	66	6.08	66	0.02	0.89
ln empl_comm	2.12	48	1.85	54	-0.27	0.26
ln empl_manu	4.50	66	4.38	65	-0.12	0.48
share_empl_agri	0.78	66	0.82	66	0.04	0.15
share_empl_comm	0.02	66	0.02	66	-0.01	0.48
share_empl_manu	0.20	66	0.16	66	-0.04	0.12
public money(Dollar)	534.78	51	359.35	51	-175.43	0.35
taxable property(k)	387.77	62	347.81	66	-39.96	0.58
iroquois trail(Dummy)	0.96	24	1.00	16	0.04	0.42
county seat(Dummy)	0.11	120	0.05	115	-0.06	0.12
elevation	531.44	115	771.11	114	239.66***	0.00
schools(N)	12.48	58	13.20	60	0.72	0.56
taught(N)	712.65	52	759.90	51	47.25	0.58
custom	0.02	108	0.01	110	-0.01	0.55
children age 5-15(N)	761.64	50	706.24	50	-55.40	0.51
electors(N)	537.79	62	489.16	63	-48.63	0.39
distillery(N)	1.50	64	1.96	68	0.46	0.19
grist mill(N)	3.91	64	3.66	68	-0.24	0.67
saw mill(N)	7.50	64	8.21	68	0.71	0.50
oil mill(N)	0.56	64	0.18	68	-0.39**	0.01
fulling mill(N)	1.98	63	2.35	68	0.37	0.36
carding machine(N)	1.95	63	2.57	68	0.62	0.30
ashery(N)	2.09	64	2.56	68	0.47	0.36
cotton woollen factory(N)	0.40	62	0.40	68	-0.01	0.97
tannery(N)	0.05	62	0.09	68	0.04	0.70
trip hammer(N)	0.44	62	0.26	68	-0.17	0.33
cloth made(yard_k)	17.35	61	31.88	66	14.53	0.28
churches(N)	3.33	15	3.95	20	0.62	0.52
iron works(N)	0.23	61	0.34	67	0.11	0.36

across units and spillover effects. This assumption likely held in the early nineteenth-century when a wagon traveled 2 to 3 miles per hour depending upon weather, roadway conditions, and the health of the travelers. I will also compare treated areas to a larger control area but excludes areas immediately adjoining the treated areas to eliminate the spillover effects.

The non-anticipation assumption is highly likely to hold since the year of operation of the whole Erie Canal is 1825 and the pre-treatment year I utilize is 1820. Even though people in the agriculture sector or commerce sector were incentivized by the construction of the Erie Canal to change jobs, the assumption here is people seldom made decisions 5 years before. I will also create subsets of the data, and remove towns located beside the middle section of the Erie Canal which operated before 1820, for robustness. The "parallel trends" assumption is tested by comparing the pre-treatment town characteristics using 1820 federal census data and the 1824 Gazetteer of the State of New York. The results from Table 1 demonstrate that there are not highly significant baseline differences in employment measures of interest; transportation, Iroquois trail; political influence, the county seat and the number of electors; economy, the value of taxable property; the number of a variety of manufacturing establishment; education, schools and the number of people were taught.

While areas immediately adjoining the treated towns are excluded to mitigate some concerns about spillover effects, it is difficult to make sure the control towns followed the same trend as canal-towns after the operation of the canal. Take population, one of the outcome variables of interest, for illustration. If the population migrated from noncanal-towns to canal-towns after the operation of the canal, the DiD estimates are overestimated. If the population in noncanal-towns saw a higher growth rate than in canal-towns after the operation of the canal because of other factors, the DiD estimates are underestimated. An alternative interpretation of biased DiD estimates could be they are potentially underestimated or overestimated if the canal traversed towns with better or worse growth prospects. In order to study this issue deeper and provide another estimate to compare, I proceed to develop an IV strategy that fit the construction background of the canal and isolates plausibly exogenous variation in access to the canal.

4.2 IV-DiD Strategy

As the basis of the IV strategy, I construct a hypothetical network that connects the targeted consequential places along approximate LCPs. The key idea, applied to my setting, is that some spatial units were selected by canal builders to pass through because of their existing traffic volumes and/or their potential to generate freight traffic. They are known as consequential places. There were also inconsequential places receiving canal only because of their geography, which places them on a cost-minimizing canal route connecting the consequential places. Areas that were not directly targeted by planners "accidentally" traversed by the canal.

The IV-DiD strategy involves two steps. In the first step, I estimate the first-stage regression of the form:

$$Canal_i = \mu + \tau_j + \delta DistanceToLCP_i + \nu_i \tag{2}$$

where $DistanceToLCP_i$ is the instrument variable, which is the Euclidean distance between town hall and the LCP. τ_j is the county fixed 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 second-stage IV regression using DiD equation given in equation 1:

$$Y_{it} = \alpha + \theta_i + \phi_t + \beta Post_t \hat{Canal_i} + \epsilon_{it} \tag{3}$$

where $\hat{Canal_i}$ is the predicted values of the first-stage regression.

To construct the instrument, I proceed in three steps. The first step is choose the consequential places that LCPs need to pass through. More consequential places chosen means more restrictions are put on LCPs, and less accidental canal-towns. I start from only choosing the start point, Baffalo, and the end point, Waterford town bordering on Albany county, Saratoga county, Rensselaer county, and Hudson River. Second, to identify the LCPs that connect Baffalo and Waterford, I follow Faber (2014) and Berger (2019) using GIS data on elevation. Digital Elevation Models having a pixel resolution of approximately 10 meters are used (Survey, 1995). The LCPs follow the topographical features to minimize construction costs. Third, I modify Bogart et al. (2022)' method as a robustness check of reasonableness of LCPs from market demand perspective. A simple gravity equation[9] and Flater (2022)'s algorithm are used to approximate the relative value of connecting town pairs. A planar distance is used to measure the distance of town pairs. I use the 1820's population in the gravity equation because the Erie Canal was constructed from 1817 to 1825, and the year 1820 lies between them. I get a 562 $\times$ 562 table of gravity index for all the cities and towns that have population data in 1820 except New York City. New York City has a 142,260 population which is 11 times more than the second-largest town, Albany. I exclude New York City, which is not along the canal, from the dataset. I then rank the gravitational value of the other 561 towns that pair with waterford and find the townHalfmoonthat has the largest gravitational value. Then find the town that has the largest gravitational value with Halfmoon. If the town that has the largest gravitational value with Halfmoon is Waterford, then select the town that has the second-largest gravitational value with Halfmoon. Continue this process until the endpoint, Buffalo, is included. Considering the nature of the canal, which needs to be supplied with water, a constraint is being considered in this process. If two candidate towns have similar gravitational values, I choose the town closer to the

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

water resources. LCPs are not necessarily passing through the center of selected towns. If LCPs approximate to those town, then it is believed that LCPs are constructed by considering economic features like population from the benefit side and topographical features like elevation from the construction costs side.

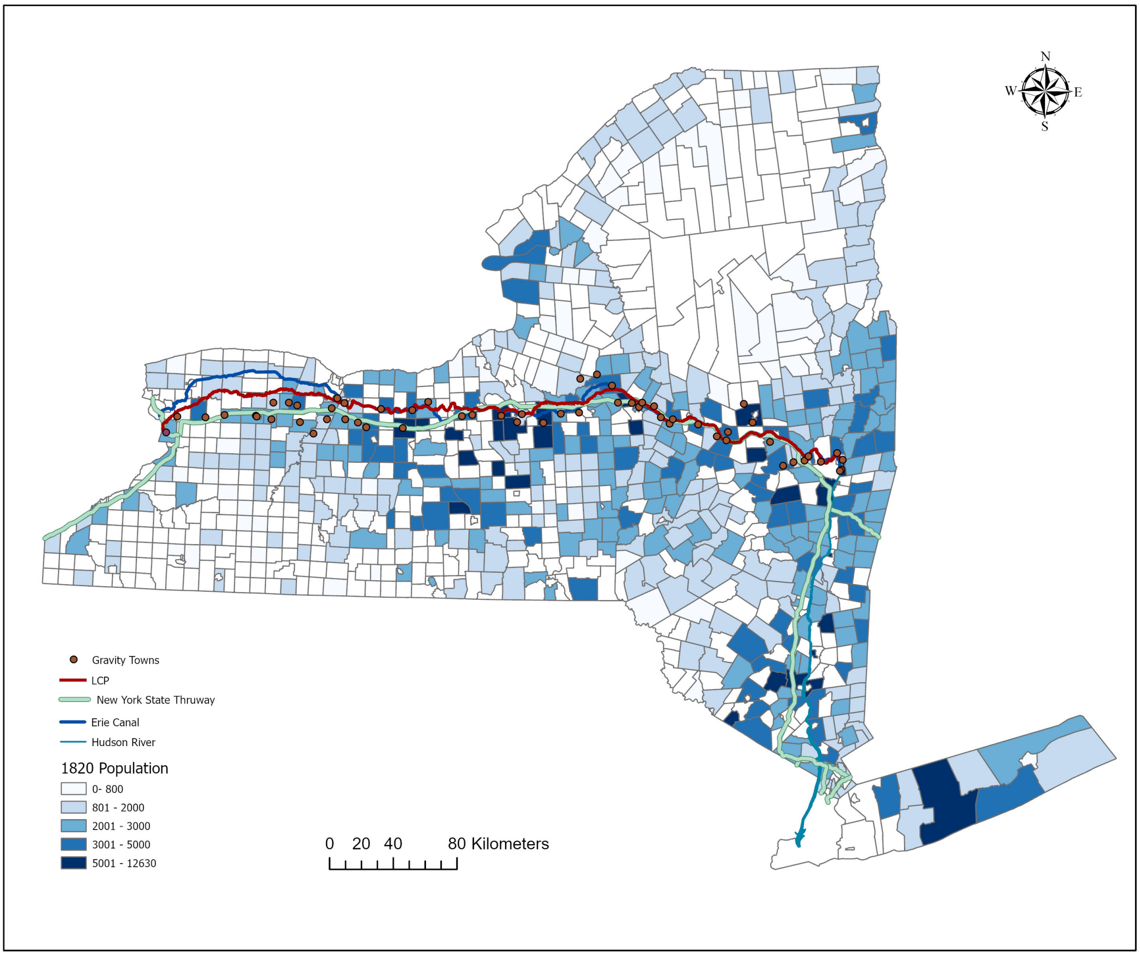


Figure 3: Least cost path, gravity towns, and population in 1820

Shown in Figure 3, the LCPs pass through consequential towns like Utica, Syracuse, and Rochester besides the Waterford and Baffalo. It also fits selected towns, selected by gravity index, very well. The New York State Thruway operated by the New York State Thruway Authority and constructed in 1954 when the cost-minimized technology was improved relative to the 1820, can be a good reference to the reasonableness of the

LCP. The thruway mostly overlaps with the LCP.

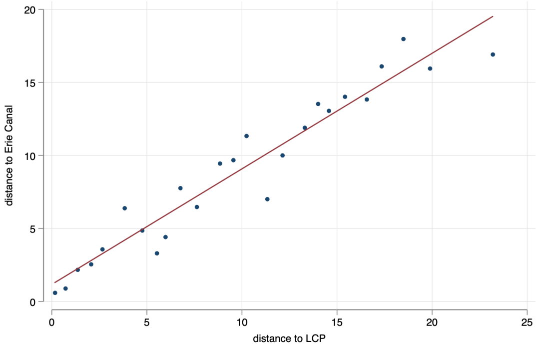
To interpret β as the LATE of being treated by LCP, the IV approach requires identifying assumptions: first stage, monotonicity, conditional independence, and exclusion. I summarize empirical evidence that supports the validity of the identifying assumptions.

4.2.1 First Stage

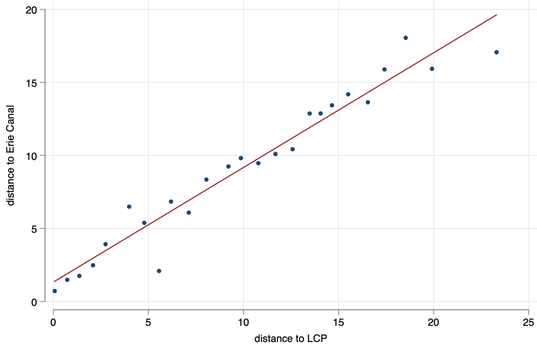
To document that towns along the LCPs also are more likely to be traversed by the Erie Canal, Figure 4 visualizes the first stage using binned scatterplots. I use the exact same towns used in DiD estimates. All towns are grouped into 25 equal-sized bins based on the distance to the nearest LCP. Each dot denotes to the mean distance to the nearest Erie Canal (a and b) or the share of towns winthin 9 km of the Erie Canal (c and d), and the mean distance to the nearest LCP respectively within each bin. Also shown are the best-fit lines estimated. Areas in proximity to the LCPs are clearly more likely to be traversed by the Erie Canal, including county fixed effects.

In the presence of heterogeneous treatment effects, we need to assume monotonicity 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, monotonicity requires that the probability of being treated by Erie Canal is negatively correlated with the distance with LCPs. I examine monotonicity assumption by subfigure c and d in Figure 4. The share of towns treated by Erie Canal declines as distance to LCPs increases.

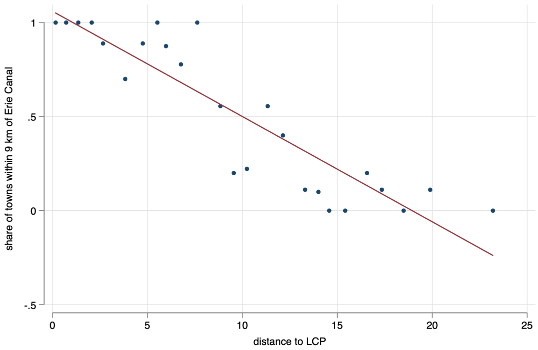
Table 2 present the first-stage regression. Columns (1) and (2) show that town probability of being treated by the canal decreases as the distance to LCP increasing. Distance to the nearest LCP is a statistically significant predictor of the actual canal constructed, also when conditioning on the county fixed effects. The $F-$statistics reported at the bottom are always sufficiently large to satisfy the (Stock and Yogo, 2005) weak instrument criterion, namely a first-stage $F-$statistics in excess of 10.00.



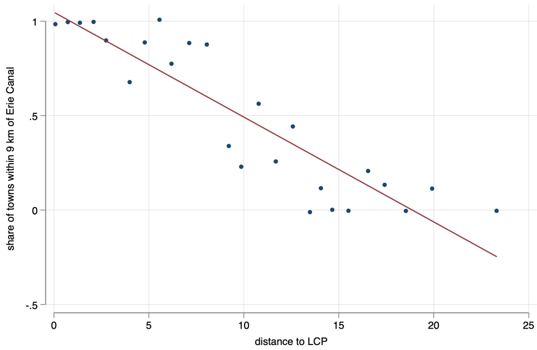
(a) Baseline



(b) With county fixed effect



(c) Baseline



(d) With county fixed effect

Figure 4: First stage

4.2.2 Balance Tests

For the instrument to be valid, it requires that there is no correlation (conditional on county fixed effects) between factors that affect the potential outcomes of a town and proximity to these LCPs. Interpreting the IV estimates as identifying the causal effect of being treated by the Erie Canal requires an exclusion restriction. That is, the LCP only impacts town potential outcomes through the Erie Canal, not through any other channels. As an indirect test of these two assumptions, I show that town observed characteristics are not significantly correlated with the instrument in 1820.

Table 3 presents OLS estimates from regressing outcomes variables and a variety of town characteristics on the distance to the nearest LCP, while conditioning on county fixed effects. Town observed characteristics are well balanced across the instrument, the

Table 2: First stage: Erie canal and distance to LCPs

	(1) Canal	(2) Canal	(3) distance to Erie canal	(4) distance to Erie canal
distance to LCP	-0.055***	-0.061***	0.784***	0.878***
	(0.005)	(0.004)	(0.071)	(0.045)
Constant	1.046***	1.100***	1.335**	0.376
	(0.048)	(0.035)	(0.640)	(0.451)
Observations	218	217	218	217
County FE	No	Yes	No	Yes
R-squared	0.502	0.676	0.652	0.828
F-statistics	109.844	299.754	121.629	373.518

Standard errors in parentheses

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

distance to a LCP, with two exceptions. Towns closer to the LCPs had lower employment share in agriculture sector (column 4) and higher employment share in manufacturing sector (column 5). Yet the coefficients, 0.005 and -0.005 are relatively small given the initial level of employment shares in agriculture and manufacturing sector are about 75% and 23% respectively at baseline in 1820. Thus, the instrument is presumably valid conditional on county fixed effects.

5 Empirical Results

5.1 DiD Estimates

To identify the areas affected by the canal, Figure 5 displays estimated effects of the canal on employment share in commerce sector between 1850 and 1900 among towns located at different distances to the canal. The underlying DiD regression is based on equation (1) where I change the canal indicator, $Canal_i$, from 6 to 12. Each individual point estimate and its confidence interval is plotted in Figure 5. As evident from these estimates, the impact of the canal attenuate beyond 9 km with no statistically significant

Table 3: Balancedness of instrument

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A					
	ln empl_agri	ln empl_manu	ln empl_comm	share_empl_agri	share_empl_manu	share_empl_comm
distance to LCP	0.000	-0.023	-0.021	0.005***	-0.005***	-0.001
	(0.011)	(0.015)	(0.014)	(0.002)	(0.002)	(0.000)
Constant	6.071***	4.661***	2.181***	0.748***	0.227***	0.025***
	(0.123)	(0.176)	(0.185)	(0.031)	(0.027)	(0.006)
Observations	132	131	102	132	132	132
County FE	Yes	Yes	Yes	Yes	Yes	Yes
	Panel B					
	ln pop	public money(Dollar)	taxable property(k)	schools(N)	children age 5-15(N)	churches(N)
distance to LCP	-0.011	5.731	-3.351	0.099	2.023	0.002
	(0.010)	(5.953)	(3.748)	(0.079)	(5.240)	(0.033)
Constant	7.884***	393.562***	399.148***	11.910***	715.125***	3.661***
	(0.101)	(122.914)	(37.283)	(0.927)	(74.975)	(0.584)
Observations	132	102	128	118	100	35
County FE	Yes	Yes	Yes	Yes	Yes	Yes
	Panel C					
	iroquois trail(Dummy)	distillery(N)	grist mill(N)	saw mill(N)	fulling mill(N)	cotton woollen factory(N)
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)
Constant	0.965***	1.751***	3.641***	7.331***	1.828***	0.430**
	(0.039)	(0.312)	(0.583)	(1.098)	(0.402)	(0.194)
Observations	40	132	132	132	131	130
County FE	Yes	Yes	Yes	Yes	Yes	Yes

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

effects beyond this distance. In the subsequent analysis, I use this 9 km cutoff to define a single access indicator that assumes no impact on areas further away. Figure 6 visualizes the 9 km buffer in the map.

Table 4: DiD estimates of effects of canal operation on employment and population

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	share_empl_agri	share_empl_manu	share_empl_comm	ln empl_agri	ln empl_manu	ln empl_comm	ln pop
PostCanal	-0.072***	0.020	0.022***	-0.026	0.345**	0.840***	0.259***
	(0.020)	(0.019)	(0.006)	(0.108)	(0.139)	(0.236)	(0.088)
Constant	0.782***	0.190***	0.017***	6.241***	4.665***	2.247***	7.810***
	(0.005)	(0.005)	(0.002)	(0.027)	(0.035)	(0.059)	(0.021)
Observations	264	264	264	264	262	184	264
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Town FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

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

The baseline estimates of the impact of the Erie Canal on local employment and population, from estimating equation (1), are presented in Table 4. Column (1) (2) (3) test whether the Erie Canal operation had a significant impact on the employment

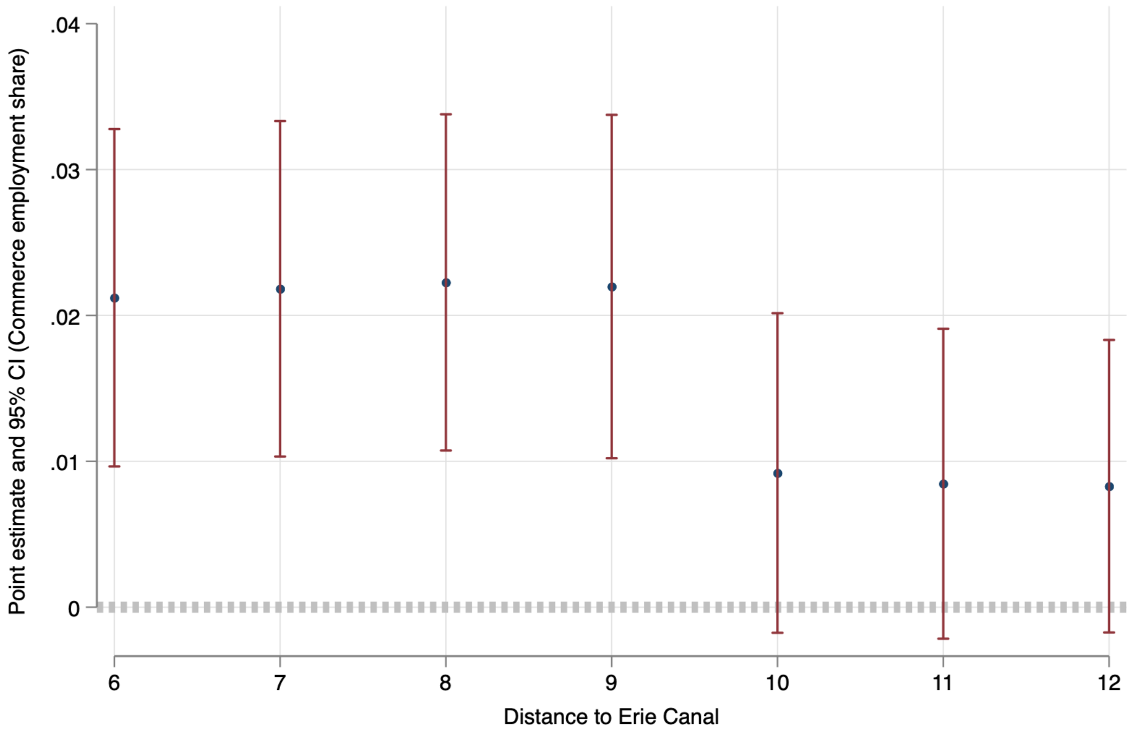


Figure 5: Distance to Erie Canal and commerce employment share

share of the agriculture sector, manufacturing sector, and commerce sector respectively. The results suggest that towns that located within 9 km of the Erie Canal had a lower employment share in the agriculture sector and a higher employment share in the commerce sector. In 1840, the average town located within 9km from the canal had 7.2% less employment share in agriculture and 2% more employment share in commerce compared with towns located between 9km and 20km from the canal, given the employment shares in agriculture, manufacturing, and commerce sector are 78.2%, 19%, and 1.7% respectively for average noncanal-towns in 1820. The operation of the Erie Canal did not have a significant effect on employment share in the manufacturing sector. I change the dependent variable to (log) sectoral employment to study the employment change within each sector. Columns (4) (5) (6) (7) show the higher employment share in the commerce sector in canal towns was attributed to the increasing employment in the

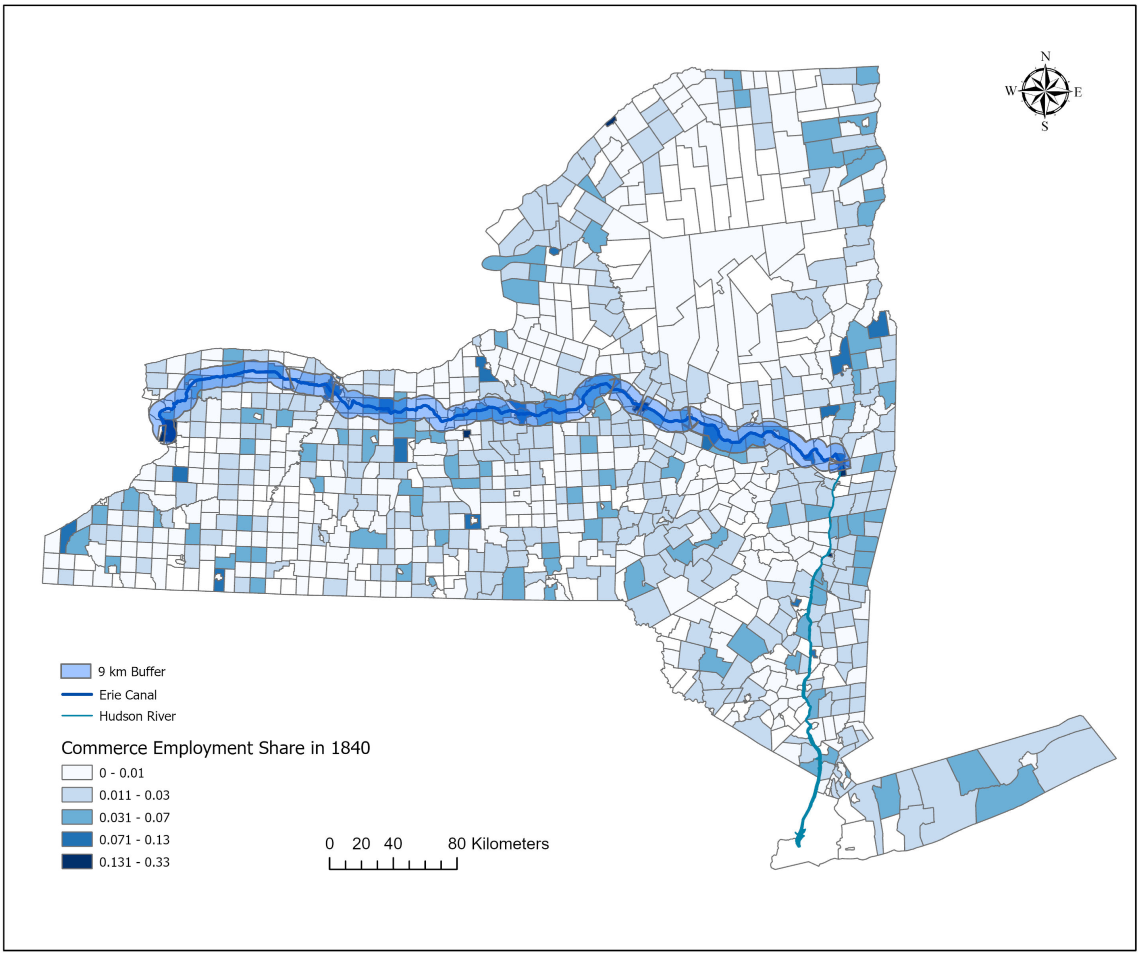


Figure 6: 9 km buffer and the employment share in commerce sector in 1840

commerce sector. The lower employment share in the agriculture sector in canal towns was not driven by the decreasing employment in the agriculture sector but because of the increasing population base. Towns located within 9km from the canal had 132.6% more employment in commerce and 41.2% more employment in the manufacturing sector in 1840, given the town-average employment in agriculture, manufacturing, and commerce sector are 494, 106, and 9 respectively in 1820.

The results from an event study of population between 1790 and 1870 are shown in

Figure 7.[10] There is no significant difference in population between treated towns and control towns from 1790 to 1820. Compared with noncanal-towns, canal-towns saw a more rapid population growth after the completion of the Erie Canal.

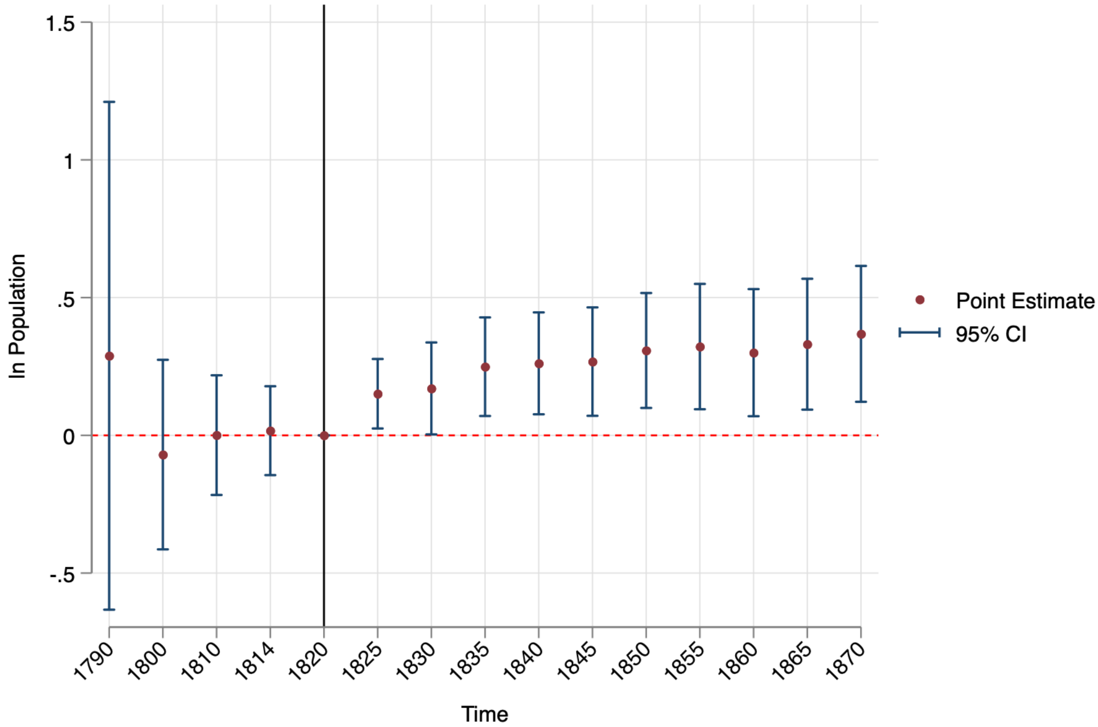


Figure 7: Population difference between canal-towns and noncanal-towns

5.2 IV-DiD Estimates

Table 5 displays the IV-DiD estimates of equation (3). When it comes to the structural transformation effects in employment, IV-DiD estimates are consistently larger in magnitude than the corresponding OLS estimates in Table 4, showing the OLS estimates are underestimated. IV-DiD estimate in columns (1) and (3) shows towns that located

---

10

$$lnPopulation_{it} = \alpha + \theta_i + \phi_t + \sum_{k=1790}^{1814} \beta_k Post_t Canal_i + \sum_{k=1825}^{1870} \beta_k Post_t Canal_i + \epsilon_{it}$$

was used in figure 3. Population data is available for years shown in the figure.

within 9km from the canal had 8.1% less employment share in agriculture and 2.4% more employment share in commerce compared with other towns located between 9km and 20km from the canal. Same as the DiD estimates, IV-DiD also shows the operation of the Erie Canal did not have a significant effect on employment share in manufacturing sector. The higher employment share in commerce sector in canal towns was attributed to the increasing employment in commerce sector (column 6). The lower employment share in the agriculture sector in canal towns was not driven by the decreasing employment in the agriculture sector but because of the increasing population base (column 4 and 7).

Table 5: IV-DiD estimates of effects of canal operation on employment and population

	(1) share_empl_agri	(2) share_empl_manu	(3) share_empl_comm	(4) ln empl_agri	(5) ln empl_manu	(6) ln empl_comm	(7) ln pop
$Post\hat{Canal}$	-0.081***	0.010	0.024***	-0.158	0.266	0.758**	0.250**
	(0.028)	(0.026)	(0.006)	(0.147)	(0.170)	(0.296)	(0.117)
Constant	0.785***	0.193***	0.016***	6.284***	4.697***	2.270***	7.816***
	(0.007)	(0.007)	(0.002)	(0.037)	(0.044)	(0.080)	(0.029)
Observations	270	270	270	270	268	184	270
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Town FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

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

6 Robustness Analysis

In this section I provide alternative estimates of DiD and IV-DiD regression to document that the main results are robust by excluding the immediately adjoining areas, using another instrument variable, using a subset of the sample by excluding towns in the middle section of the canal, and using matching methods.

6.1 Excluding Adjoining Towns

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 the interference across units and spillover effects. Thus, an important question in my context is whether these localized impacts are offset by a reallocation of manufacturing and commerce workers from nearby areas. Therefore, I exclude areas immediately adjoining the treated areas from both DiD and IV-DiD regression.

In the DiD and IV-DiD regression, the control towns are those located more than 9 km but less than 20 km from the canal. I split the control towns into two parts, and exclude towns between 9 km and 14 km. Now the control towns are only those located between 14 km and 20 km from the canal. The assumption is that even though there is a spillover effect, towns located between 14 km and 20 km from the canal are not affected by the treatment.

6.1.1 DiD estimates

A balance test is presented between treated towns and new control towns. Table 6 demonstrate that there are not highly significant baseline differences in population and employment measures of interest; transportation, Iroquois trail; political influence, the county seat and the number of electors; public economy; the number of a variety of manufacturing establishment; education, schools and the number of people were taught. The elevation difference is controlled by town fixed effects. The structural transformation effects in agriculture and commerce employment are still significant with the same signs.(Table 7)

Table 6: Comparison of town characteristics before 1825 between canal-towns (9km) and noncanal-towns (14km - 20km)

	Treated towns	Obs.	Control towns	Obs.	Diff.	$p - value$
population(N)	2827.25	64	2451.05	39	-376.20	0.21
ln empl_agri	6.06	66	6.09	37	0.03	0.83
ln empl_comm	2.12	48	1.78	29	-0.34	0.21
ln empl_manu	4.50	66	4.25	36	-0.25	0.20
share_empl_agri	0.78	66	0.84	37	0.07	0.06
share_empl_comm	0.02	66	0.01	37	-0.01	0.15
share_empl_manu	0.20	66	0.15	37	-0.05	0.06
public money(Dollar)	534.78	51	325.50	30	-209.28	0.39
taxable property(k)	387.77	62	279.81	38	-107.96*	0.05
iroquois trail(Dummy)	0.96	24	1.00	9	0.04	0.55
county seat(Dummy)	0.11	120	0.03	64	-0.08	0.07
elevation	531.44	115	834.14	63	302.70***	0.00
schools(N)	12.48	58	12.51	35	0.03	0.98
taught(N)	712.65	52	704.13	30	-8.52	0.93
custom	0.02	108	0.00	61	-0.02	0.29
children age 5-15(N)	761.64	50	633.47	30	-128.17	0.18
electors(N)	537.79	62	445.43	37	-92.36	0.11
distillery(N)	1.50	64	1.74	39	0.24	0.52
grist mill(N)	3.91	64	3.62	39	-0.29	0.67
saw mill(N)	7.50	64	7.92	39	0.42	0.74
oil mill(N)	0.56	64	0.13	39	-0.43*	0.02
fulling mill(N)	1.98	63	2.15	39	0.17	0.71
carding machine(N)	1.95	63	2.44	39	0.48	0.46
ashery(N)	2.09	64	2.08	39	-0.02	0.97
cotton woollen factory(N)	0.40	62	0.26	39	-0.15	0.48
tannery(N)	0.05	62	0.00	39	-0.05	0.43
trip hammer(N)	0.44	62	0.23	39	-0.20	0.34
cloth made(yard_k)	17.35	61	18.64	38	1.29	0.63
churches(N)	3.33	15	3.09	11	-0.24	0.80
iron works(N)	0.23	61	0.37	38	0.14	0.36

Table 7: DiD estimates of effects of canal operation on employment and population excluding adjoining towns

	(1) share_empl_agri	(2) share_empl_manu	(3) share_empl_comm	(4) ln empl_agri	(5) ln empl_manu	(6) ln empl_comm	(7) ln pop
PostCanal	-0.060**	0.008	0.017***	-0.022	0.263*	0.712***	0.210**
	(0.023)	(0.021)	(0.004)	(0.114)	(0.148)	(0.257)	(0.088)
Constant	0.779***	0.195***	0.018***	6.241***	4.674***	2.339***	7.807***
	(0.007)	(0.007)	(0.001)	(0.036)	(0.048)	(0.085)	(0.027)
Observations	206	206	206	206	204	140	206
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Town FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

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

6.1.2 IV-DiD estimates

An indirect test of the exclusion assumption is provided in Table 8. Although employment in three sectors is correlated with the instrument, distance to the canal, most of other observed town characteristics are not significantly correlated with the instrument in the year of 1820. It is reasonable to say the instrument affects the dependent variables only through the endogenous variable, not from other channels. Also, the comparison between canal-towns and noncanal-towns in Table 8 passes the balance test in the first place. Table 9 presents IV-DiD estimates, excluding immediately adjoining towns. Column (1) shows the result from first-stage regression. Whether a town is traversed by the LCP is a significant predictor of whether the town is penetrated by the Erie Canal, also when conditioning on the county fixed effects with R^2 0.78 and $F-$statistics 212.26. The structural transformation effects in agriculture and commerce employment are still significant with very similar magnitude.

6.2 Use an Alternative IV

Instead of using the distance to LCP as the IV, I use an indicator variable, LCP_i, as the IV. LCP_i is a distance measure to the LCP, which corresponds the value 1 for towns located within 9 km of the LCP and 0 for other towns.

Table 8: Balancedness of instrument

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A					
	ln empl_agri	ln empl_manu	ln empl_comm	share_empl_agri	share_empl_manu	share_empl_comm
distance to LCP	0.003	-0.027**	-0.026**	0.006***	-0.005***	-0.001**
	(0.009)	(0.011)	(0.012)	(0.002)	(0.001)	(0.000)
Constant	6.044***	4.649***	2.223***	0.745***	0.229***	0.026***
	(0.113)	(0.178)	(0.175)	(0.031)	(0.027)	(0.007)
Observations	103	102	77	103	103	103
County FE	Yes	Yes	Yes	Yes	Yes	Yes
	Panel B					
	ln pop	public money(Dollar)	taxable property(k)	schools(N)	children age 5-15(N)	churches(N)
distance to LCP	-0.012	3.741	-7.271***	0.033	-3.417	-0.033
	(0.007)	(6.808)	(2.483)	(0.078)	(5.272)	(0.048)
Constant	7.864***	424.450***	412.280***	12.202***	743.534***	3.545***
	(0.100)	(134.114)	(39.956)	(1.021)	(82.111)	(0.573)
Observations	103	81	100	93	80	26
County FE	Yes	Yes	Yes	Yes	Yes	Yes
	Panel C					
	iroquois trail(Dummy)	distillery(N)	grist mill(N)	saw mill(N)	fulling mill(N)	cotton woollen factory(N)
distance to LCP	0.001	-0.011	-0.008	0.028	0.020	-0.013
	(0.001)	(0.019)	(0.036)	(0.081)	(0.031)	(0.013)
Constant	0.965***	1.689***	3.866***	7.404***	1.869***	0.468**
	(0.039)	(0.321)	(0.611)	(1.214)	(0.436)	(0.200)
Observations	33	103	103	103	102	101
County FE	Yes	Yes	Yes	Yes	Yes	Yes

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

The alternative IV-DiD strategy also involves two steps. In the first step, I estimate the first-stage regression of the form:

$$Canal_i = \mu + \tau_j + \delta LCP_i + \nu_i \tag{4}$$

τ_j is the county fixed 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 second-stage IV regression using DiD equation given in equation 4:

$$Y_{it} = \alpha + \theta_i + \phi_t + \beta Post_t \hat{Canal_i} + \epsilon_{it} \tag{5}$$

where $\hat{Canal_i}$ is the predicted values of the first-stage regression.

The results from Table 10 demonstrate that there are not highly significant baseline differences in employment measures of interest; transportation, Iroquois trail; political

Table 9: IV-DiD estimates of effects of canal operation on employment and population excluding adjoining towns

	(1) Canal	(2) share_empl_agri	(3) share_empl_manu	(4) share_empl_comm	(5) ln empl_agri	(6) ln empl_manu	(7) ln empl_comm	(8) ln pop
distance to LCP	-0.058***							
	(0.004)							
$Post\hat{Canal}$_19th_5k		-0.079***	0.009	0.020***	-0.152	0.215	0.759***	0.175
		(0.028)	(0.027)	(0.005)	(0.135)	(0.184)	(0.268)	(0.125)
Constant	1.178***	0.786***	0.194***	0.017***	6.284***	4.687***	2.312***	7.814***
	(0.037)	(0.009)	(0.009)	(0.002)	(0.045)	(0.062)	(0.092)	(0.042)
Observations	168	206	206	206	206	204	140	206
Time FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Town FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes							

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

influence, the county seat and the number of electors; economy, the value of taxable property; the number of a variety of manufacturing establishment; education, schools and the number of people were taught.

IV-DiD estimates using a dummy variable IV are presented in Table 11. Column (1) shows the result from first-stage regression. Whether a town is traversed by the LCP is a significant predictor of whether the town is penetrated by the Erie Canal, also when conditioning on the county fixed effects with R^2 0.71 and $F-$statistics 97.77. The structural transformation effects in agriculture and commerce employment are still significant with very similar magnitude.

6.3 Drop the Middle Section of Erie Canal

The middle section of the Erie Canal from Utica in Oneida county to Montezuma in Sebeca county was first constructed in the year 1817 and was completed and operated immediately in 1820. The middle section also passed through many towns in Cayuga, Madison, Oneida, and Onondaga counties. Towns located within 9km from the canal in these counties may have already been treated in or before 1820 depending on how quickly employment responded to canal operation. I exclude towns located within 9 km of the canal in those counties from the IV-DiD analysis where using the distance to LCP as

Table 10: Comparison of town characteristics before 1825 between canal-towns (9km) and noncanal-towns (20km) using distance to LCP as a indicator

	Treated towns	Obs.	Control towns	Obs.	Diff.	$p - value$
population(N)	2823.23	60	2793.25	72	-29.98	0.92
ln empl_agri	6.06	61	6.08	71	0.02	0.85
ln empl_comm	2.16	45	1.83	57	-0.33	0.17
ln empl_manu	4.51	61	4.39	70	-0.12	0.49
share_empl_agri	0.77	61	0.82	71	0.05	0.09
share_empl_comm	0.02	61	0.02	71	-0.01	0.34
share_empl_manu	0.21	61	0.16	71	-0.04	0.08
public money(Dollar)	354.33	49	532.81	53	178.48	0.35
taxable property(k)	381.34	60	354.67	68	-26.67	0.71
iroquois trail(Dummy)	0.96	23	1.00	17	0.04	0.40
county seat(Dummy)	0.10	109	0.06	126	-0.04	0.30
elevation	539.90	105	744.62	124	204.73***	0.00
schools(N)	12.30	54	13.31	64	1.02	0.41
taught(N)	689.34	50	780.11	53	90.77	0.29
custom	0.02	97	0.01	121	-0.01	0.44
children age 5-15(N)	729.85	48	737.71	52	7.86	0.93
electors(N)	522.00	60	505.23	65	-16.77	0.77
distillery(N)	1.71	62	1.76	70	0.05	0.89
grist mill(N)	3.66	62	3.89	70	0.22	0.69
saw mill(N)	7.50	62	8.19	70	0.69	0.51
oil mill(N)	0.44	62	0.30	70	-0.14	0.36
fulling mill(N)	1.93	61	2.39	70	0.45	0.26
carding machine(N)	2.05	61	2.47	70	0.42	0.48
ashery(N)	2.16	62	2.49	70	0.32	0.52
cotton woollen factory(N)	0.40	60	0.40	70	0.00	1.00
tannery(N)	0.05	60	0.09	70	0.04	0.73
trip hammer(N)	0.27	60	0.41	70	0.15	0.40
cloth made(yard_k)	16.72	58	31.79	69	15.07	0.26
churches(N)	3.50	14	3.81	21	0.31	0.75
iron works(N)	0.17	59	0.39	69	0.22	0.07

Table 11: IV-DiD estimates of effects of canal operation on employment and population using an alternative IV

	(1) Canal	(2) share_empl_agri	(3) share_empl_manu	(4) share_empl_comm	(5) ln empl_agri	(6) ln empl_manu	(7) ln empl_comm	(8) ln pop
LCP	0.787***							
	(0.080)							
$Post\hat{Canal}$		-0.082***	0.021	0.023***	-0.120	0.365**	0.813***	0.228*
		(0.026)	(0.024)	(0.008)	(0.137)	(0.180)	(0.306)	(0.116)
Constant	0.146***	0.785***	0.190***	0.016***	6.265***	4.658***	2.248***	7.815***
	(0.036)	(0.007)	(0.006)	(0.002)	(0.035)	(0.046)	(0.079)	(0.029)
Observations	217	264	264	264	264	262	184	264
Time FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Town FE		Yes	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes							

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

the IV. The structural transformation effects in agriculture and commerce employment are still significant. Excluding the middle section, the average town located within 9km of the canal still had less employment share in agriculture and more employment share in commerce compared with towns that located between 9km and 20km from the canal (columns 1 and 3 in Table 12). Differing from estimates including all towns, the lower employment share in the agriculture sector in canal towns excluding towns in the middle section was attributed to the decreasing employment in agriculture sector, not because of the increasing population base. The coefficients are very similar to the IV-DiD estimates that do not exclude immediately adjoining towns.

Table 12: IV-DiD estimates of effects of canal operation on employment and population without the middle section

	(1) share_empl_agri	(2) share_empl_manu	(3) share_empl_comm	(4) ln empl_agri	(5) ln empl_manu	(6) ln empl_comm	(7) ln pop
$Post\hat{Canal}$	-0.098***	0.035	0.021***	-0.300*	0.260	0.778**	0.082
	(0.034)	(0.029)	(0.007)	(0.161)	(0.226)	(0.321)	(0.134)
Constant	0.803***	0.177***	0.015***	6.276***	4.576***	2.153***	7.784***
	(0.009)	(0.008)	(0.002)	(0.043)	(0.061)	(0.085)	(0.035)
Observations	196	196	196	196	194	126	198
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Town FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

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

6.4 Matching Approach

Towns shared similar preexisting employment structures in terms of sectoral employment and sectoral employment share may be scattered over New York State. That is, candidate control towns located more the 20 km from the Erie Canal were excluded from the analysis. Another way to choose control towns is using the matching method. I first use propensity score matching (PSM) to choose control towns. Population, employment share, and (log) employment in the agriculture sector, employment share, and (log) employment in the commerce sector in the pre-treatment year, 1820 year, were used in the matching process. Towns with extreme propensity scores that were lower than 0.01 and higher than 0.9 were excluded. **??** in the appendix compares propensity scores between treated towns and control towns. Then I deploy DiD (equation 1) using unweighted regression and weighted regression by propensity score respectively. Table 13 presents the unweighted and weighted (by propensity score) PSM-DID estimates. The effects of canal operation on employment share in agriculture and commerce are still significant and with the same signs. Towns located within 9 km of the canal still have lower employment share in agriculture and higher employment share in commerce compared with other towns with similar characteristics. The higher employment share in commerce sector is because of the increasing employment in commerce. The lower employment share in the agriculture sector in canal towns was not driven by the decreasing employment in the agriculture sector but because of the increasing population base.

7 Agricultural-Manufacturing as the Mechanism

In this section, I show the mechanism of structural transformation in employment. The decreasing transportation cost due to the canal attracted people pouring into canal-towns, established manufacturing mills, worked as manufacturing workers, and shipped

Table 13: PSM-DiD estimates of effects of canal operation on employment and population

	(1) share_empl_agri	(2) share_empl_manu	(3) share_empl_comm	(4) ln empl_agri	(5) ln empl_manu	(6) ln empl_comm	(7) ln pop
				Unweighted			
PostCanal	-0.063***	0.017	0.020***	-0.075	0.171	0.532***	0.144
	(0.022)	(0.020)	(0.005)	(0.089)	(0.125)	(0.178)	(0.089)
Constant	0.779***	0.183***	0.020***	6.169***	4.598***	2.213***	7.748***
	(0.001)	(0.001)	(0.000)	(0.005)	(0.007)	(0.012)	(0.005)
				Weighted			
PostCanal	-0.065***	0.004	0.024**	0.028	0.314**	0.513**	0.309**
	(0.023)	(0.028)	(0.010)	(0.130)	(0.134)	(0.214)	(0.145)
Constant	0.729***	0.226***	0.027***	6.230***	4.965***	2.489***	7.941***
	(0.002)	(0.002)	(0.001)	(0.011)	(0.011)	(0.020)	(0.012)
Observations	806	806	806	808	800	676	816
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Town FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

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

processed agricultural staples along the canal as businessmen. As Ransom (1971) pointed out manufacturing - like agriculture - developed in response to eastern demands for western agriculture staples. This processing manufacturing is an extension of the western comparative advantage in agriculture, and the canal encourages regional economic specialization by integrating west and east New York State.

The types of manufacturing establishments in canal towns shown by Table 1 gave some evidence. Averagely, canal towns had 1.5 distilleries, 4 grist mills, 7.5 sawmills, 0.5 oil mills, 2 fulling mills, 2 carding machines, 2 ashery, 0.5 cotton woollen factories, 0.05 tannery, 0.5 trip hammers, 1,735,000 yards of cloth made in the year 1824. The high transportation cost of agricultural staples like wheat, corn, and meat can be reduced by processing the staples into derivatives having higher value per ton. Wheat or corn was distilled into whiskey and ground into flour. Swine and cattle were slaughtered and sent east along the Erie Canal in the form of processed meat. Deer skins and raw hides were processed in the tannery. Sheep were processed into various derivatives like butter and wool in cotton woollen factories. Raw woods were cut into lumber, staves, shingles, and timber in sawmills. Such manufacturing activities result from the locational advantage created for processors by the production of local agricultural goods. And these processed goods were shipped to east New York State using the Erie Canal.

The types of goods shipped in the Erie Canal gave strong evidence (Table 14)[11]. Raw materials and basic resources, such as 311,256 feet of lumber, 3,640 bushels of wheat, 4,335 barrels of flour, 4,754 barrels of provisions (pork and beef), and 1,705 tons of ashes, which listed top 5 among properties cleared at Buffalo were passing east on Erie Canal in 1829. By 1835, double amounts of provision, 4 times more ashes, 7 times more lumber, 23 times more flour, and 46 times more wheat were passing east on the Erie Canal from Buffalo. Similar composition of articles passed Utica was observed from 1830 to 1832.

8 Conclusion

Moving forward, I will take two steps to improve the paper. First, I will gather and analyze more data that will permit new, more detailed analyses. The first step includes two parts: a) While Shares of employment in broad sectors are used to measure structural transformation in the literature, looking into greater detail of employment changes within sectors is helpful to better understand the transformation. Agriculture censuses starting from 1840 that have records of the value and number of produce, and employment of divisions within broad sectors will be studied, although they are county-level and available years are after the operation of Erie Canal. b) The town-level data on number of manufacturing establishments from state censuses before and after the operation of Erie Canal, and the data of commodities shipped along Erie Canal from the Annual Report of the Canal Commissioners will be further digitalized and used to provide a clearer picture of transformation of New York State.

Second, I will study the political economy of Erie Canal construction. DeWitt Clinton, who took office as Governor of New York State on the year of start of Erie Canal

[11] State of New York. *Annual Report of the Canal Commissioners.* Columbia University.

construction, built a political coalition that wanted the canal that transcended pure political allegiances and overcame regional differences in priorities. In order to better understand Erie Canal and the transformation of New York State, the political economy of canal finance (Goodrich et al., 1961; Wallis, 2003; Engerman and Sokoloff, 2004) has to be studied in more depth.

References

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

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*.

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.

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).

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

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).

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.

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.

Survey, U. G. (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.

Table 14: 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

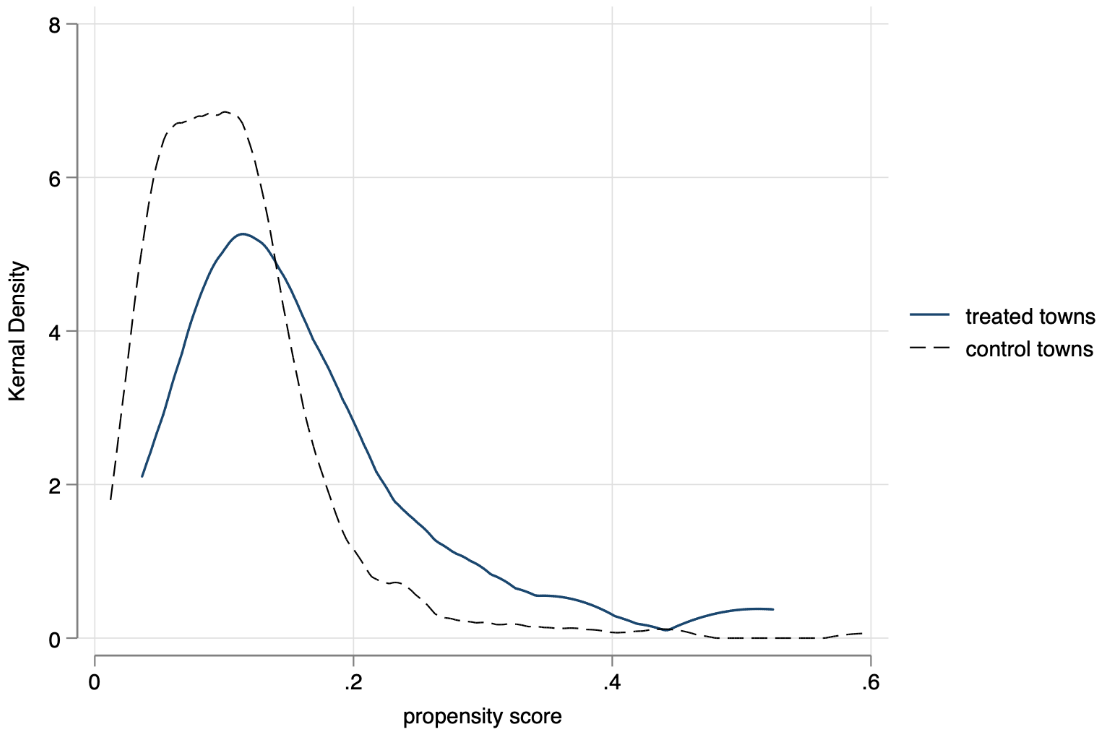


Figure 8: Density of propensity score of treated and control towns