Did the US Interstate Contribute to Agglomeration?

Thor Dodson

October 27, 2020

Abstract

The US Interstate Highway System had a significant impact on market accessibility and transportation costs between regions. Whether this should lead to increased clustering of economic activity due to increased 'economic centripetal forces' or a dispersal from 'centrifugal forces' depends on factors that differ by industry. This paper suggests the impact depends on truck transportation utilization and input-output linkages. Utilizing travel time data constructed using GIS techniques along with BEA data on the spatial distribution of economic activity, a simple panel estimation is conducted to test the theory.

Keywords: transportation, new economic geography, agglomeration

JEL Codes: would go here

1 Introduction

Roads foster the movement of goods and people between regions, facilitating the clustering of economic activity (agglomeration) as well as dispersal. The spatial distribution of economic activity has important implications for regional policy—especially programs concerned with regional inequality and targeting growth. I consider roads as part of the economic environment, shaping firms' location decisions and the patterns of where economic activity happens.

My paper examines how the US Interstate Highway System impacted the spatial distribution of different industries, as well as properties of industries that can explain the varying responses. Based on location theory and the benefits of agglomeration and dispersal I suggest that industries with a higher truck transportation share of inputs are more likely to disperse. This question is related to a line of research examining the interaction of roads and spatial distribution including Rothenberg (2011) who finds that road surface quality improvements in Indonesia lead to a dispersal of durable goods manufacturers using the Ellison and Glaeser index, Redding and Turner (2014) who survey the existing literature finding that highways tend to decentralize urban populations and manufacturing activity while different sectors appear to respond differently, and Frye (2016) who examines the effect of the US Interstate Highway System on employment concentration finding substantial growth in highway counties relative to non-highway counties.

I construct a novel data set of travel times between metropolitan regions in the US for each year between 1950 and 1993 as the Interstate Highway System was constructed. The travel time is an important component of the transportation cost between regions, affecting trade and decisions of where to locate economic activity. My data adds to the literature explicitly representing road systems as a network of transportation costs, such as Rothenberg (2011) who utilizes a mapping between road quality and speeds to estimate the travel time changes in Indonesia, Faber (2014) who constructs least cost path spanning tree networks examining China's National Trunk Highway System, Donaldson and Hornbeck (2016) who calculate lowest-cost county-to-county freight routes in the US, Alder (2016) who constructs a grid of cells with different speeds to use a shortest path algorithm examining bilateral travel times in India, and Jaworski et al (2018) who utilize decennial maps with surface information, mileage, and travel time estimates to construct internal trade costs for the US.

Using data on county level earnings by industry in the US I construct a spatial GINI measuring how inequal the distribution of economic activity is across all counties for each year. This is similar to Rey and Smith (2012) who introduce a spatial decomposition of the GINI coefficient that exploits the contiguity matrix, Sutton (2012) who constructs spatial GINI from nighttime satellite imagery and population density, and Panzera and Postiglione (2019) who propose an index based on the GINI that introduces regional importance weighting.

I utilize the regional variation in road completion timing to identify the causal effect of the changes to the road system on the spatial distribution of the different industries. This strategy is a similar to timing of treatment as in Brand and Xie (2007), Steiner et al (2016), and Imai (2020), but in my case the treatment is not an event but a process, so I construct indices representing the timing of road construction and change in spatial distribution.

I use a panel data set with interaction affects to detect the industry varying affect the change in the travel time index has on the spatial GINI index. I construct 'meaningful' marginal affects and standard errors as in Brambor, Clark, and Golder (2004). Furthermore, I perform simulations verifying the appropriateness of the specification.

I find that industries with a higher trucking share of inputs disperse more when the road system was improved. The average highway travel time between metropolitan regions decreased by about 18%, with varying declines across regions. The spatial GINI for total personal income declined slightly between 1969 and 1985, but rose to it's previous level by 2000 and does not change much after that, while the spatial GINI for population declined slightly until 1980 and has been slightly increasing ever since. This combined with the significant movements in industry specific spatial GINI suggest there is not a large change in the overall spatial distribution of economic activity, but there is significant relocation of where specific types of industry occur.

The paper proceeds as follows: Section 2 discusses the theory surrounding why different industries will respond differently to an improvement in the road system, Section 3 describes the data, Section 4 reports the estimation results, and Section 5 presents the model.

2 Theory of Roads and Spatial Distribution

Roads alter the time it takes to traverse an area, effectively warping the space and bringing regions closer together by facilitating the movement of cars and trucks. This reduction in travel time lowers the cost of moving goods by lowering the wage paid to the drivers, and reduces the need for large inventories as stocks can more quickly be replenished. This second effect is particularly important as observed in the rise of "just-in-time" manufacturing/inventory during the 1970s and 80s, as well as the importance industry places on overnight shipping. Although rail and water can typically transport materials at a lower cost, the speed of roads is crucial for supply coordination, and the access provided by roads to regions not adjacent to rail or water lead to their involvement in the 'first and last mile' for intermodal shipping. These benefits from roads influence the desirability of locations, as elaborated in Weber's (1909) conception of contours of total transport costs from multiple sources of input and output in his theory of the location of industries, and expanded upon by the central place theory of Christaller (1933) and Loschian (1941) discussing how producers position themselves in the market to realize maximum utility. By providing access and lowering the cost of transportation between regions, roads play a crucial role in shaping the location decisions of firms.

Agglomeration is the clustering of economic activity in space. This applies to multiple scales, including countries, cities, and districts. The theories for why this occurs are aptly summarized by Marshall (1890) who points to three sources of benefits from agglomeration: 1) knowledge spillovers–the idea that information is "in the air" and propagated by interactions, 2) pooled labor–the benefits for employers and employees from being able to access a pool of , 3) forward and backward linkages (proximity to markets and sources of inputs, as transport is costly). The third type is the most explored by the new economic geography and 'market access' literature, spearheaded by Fujita and Krugman (1999). We can think of these as 'centripetal forces' that pull activities towards each other, resulting in clustering. However, being near other firms has a trade-off: wages and the price of land (either pur-

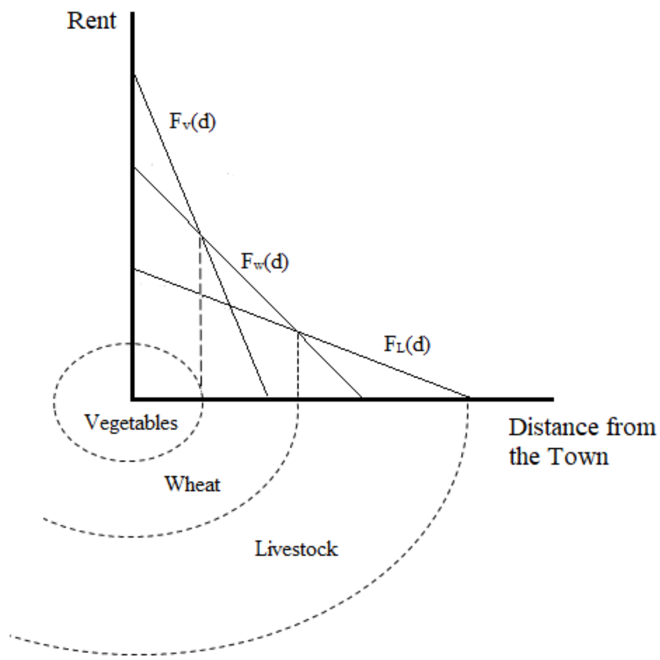

Figure 1: A Bid-Rent Curve

chasing or renting) are pushed up due to competition, acting as 'centrifugal forces' pushing firms to locate away from clusters. Furthermore, proximity to multiple sources of demand and inputs may be a relevant consideration pushing a plant away from any particular market center.

Different industries have different sensitivities to each of these forces based on what they do and how they do it. Von Thünen (1826) captured this idea with his model of agricultural land use and this was extended by Alonso's (1960) bid-rent theory; an example is shown in Figure 1. The key idea is how much 'land rent'[1] a firm is able to generate at a particular location, based on the difference between the value of their product at the market and the input and transportation costs incurred from operating at that position[2]. In its simplest form we conceive a single market existing at a point in a uniform plane where economic activity can take place, but it can be extended to incorporate multiple market centers and

[1] Von Thünen defines land rent as value generated in excess of all input costs, although there are some competing definitions of this concept

[2] for agricultural models the yield of a location is also a factor

surfaces with varying transportation costs such as a river or road system. For the single market framework, the intercept represents the rent an industry can offer for being at the center of the market—the point where the benefits of agglomeration are the highest, and the slope represents how the rent an industry can offer changes with distance from the market—a combination of the transportation cost for that industries' product and how the total cost of inputs changes with distance. Industries that benefit from agglomeration tend towards the market, and industries with goods that can be moved cheaply tend to be pushed away from the market. In a multiple market framework this is more complicated as firms within industries may choose to deal with just one market or multiple markets, but still we would observe that industries benefitting more from agglomeration would tend towards market centers and industries with costs that decline more rapidly with distance would locate away from market centers. In reality markets do not operate at single points in space, but the same logic applies for distributed markets as long as there is some varying concentration of market activity. Additionally, from a large enough scale even a city appears as a point.

From this lens, an improvement in the road system does two things. By lowering the cost of transporting materials, the slope of the bid-rent curve is flattened as it is less costly to be located away from the market center. This effect pushes industries outward from market centers and makes more distant locations viable points of operation. However, an improvement in the road system also facilitates increased access to a market center—meaning employees from a wider radius can commute in. This increases the agglomeration benefits of an area by creating a larger labor pool firms can pull from, as well as potentially increasing the suite of interactions that lead to knowledge spillovers. Effectively, the market center becomes larger and has increased capacity for agglomeration. Improved roads, then, through lowering transportation costs push industries out and pull industries in—but as we will see how much of each varies by industry.

Industries that have a larger truck transportation share of inputs benefit more from the decline in transportation costs due to improvements to the road system. While the reduction in transportation costs reduces the slope of the bid-rent curve for all industries, the slope becomes flatterer for industries that utilize trucking more. This makes it comparatively less costly for these industries to be farther away and hence pushes them outward, away from the market centers/central business districts. Based on this, we suggest the main hypothesis of the paper—-that the coefficient on the interaction term between travel time and truck transport share of inputs will be positive.

As industries grow in scale the efficiency gains and limits to the scope of operations differ depending on what they do. Here I broadly classify two types of economic processes: information services and material transformations. Under information services I identify operations such as accounting, legal, finance, insurance, consulting, logistics, software development, planning, and design. Material transformation includes manufacturing, construction, agriculture, extraction, refining, chemical processing, anything that involves physical goods. 'Information services' tend to scale much more than material transformations due to the technologies involved, the increased scope of a specialist, and ability for a single firm to influence many operations. Because of this, information services benefit more from the knowledge spillovers and pooled labor of agglomeration. They receive a larger return from high skilled labor and innovation because their product can potentially have enormous reach, such as an accounting software or knowledge of how to structure an operation. In the bid-rent context, the intercept is higher. 'Material transformations' scale very much, but they are in a sense bound by the physical processes and input requirements–they cannot 'copy and paste' for near free. Therefore, they do not benefit as much from agglomeration when scaled. This means the intercept will be relatively lower for these industries, representing a lower benefit from being at the market center where agglomeration is highest. Considering this, when the road improves access to the cities, information services benefit more from the increased agglomeration capacity and want to cluster more, while material transformation processes are pushed out. In the bid-rent context the intercept rises more for the information services when access improves as being near the center is more valuable for them. Furthermore, material transformation operations tend to utilize more low-skilled labor, require large spaces, and ship to multiple sources of demand, implying they are disproportionately affected by the dispersal forces of wage, cost of land, and proximity to multiple sources of demand.

The stage in the product life-cycle will influence the sensitivity to the benefits of agglomeration and dispersal. The product life-cycle is a concept from management and marketing that distinguishes four stages in a product's life: introduction, growth, maturity, and saturation. The first two and last two can be grouped together as early and late respectively. Early stage products involve design, the supply chain is not well formed, demand must be created, and there is low competition; thus they benefit more from the knowledge spillovers and access to pooled high skilled labor of agglomeration. Late stage products face high competition and low prices, deal with complex supply chains and mass production, and profitability/survival is more based on production/distribution efficiency– thus they benefit more from the lower wages, cost of land, and central distribution offered by dispersal. When the road is improved both agglomeration and dispersal are further facilitated, exasperating the location preferences for both early and late stage products.

In summary, because of the differing effects of centripetal and centrifugal economic forces on industries, when the road is improved we suspect that industries that utilize trucking more will disperse, industries categorized as information services and material transformations will agglomerate and disperse respectively, and industries dealing in early and late stage products will agglomerate and disperse respectively.

3 Data on Travel Time and Spatial Distribution[3]

The Interstate Highway System began construction in 1956, in part due to President Eisenhower's dismal experience crossing the country on a military expedition in 1919, but the call for an updated national highway system had been around since the 1930's. While there was already a sizeable road system in place and most places could be accessed, the road conditions were often poor[4], many of them unpaved and not up to the high engineering standards the Interstate offered. The Interstate Highway System can be viewed as accomplishing two things: 1) connecting and providing or improving access to regions, 2) lowering the cost of moving goods and people through reductions in travel time and facilitating larger trucks. The key statistic I utilize is the average transportation time between metropolitan regions for each year during the time it was built.

I construct a representation of the US road system for each year between 1950 and 1993 as the Interstate Highway System was developed and use this to estimate the travel times between metropolitan statistical areas with a shortest path algorithm. I do this by combining two geographic information system (GIS) road files and converting them to an edge weighted network that the Dijkstra algorithm can be performed on.

The first GIS file is formed by isolating the interstate highways from the PA_NHS 2012 shapefile[5] detailing all US roads at that time. The second GIS shape file I form by manually tracing a 1954 map image[6] produced by the US government detailing the principle highways and arterials in existence at that time, what we refer to now as the US numbered highways. I approach the road system in this way because in addition to entirely new roads the Interstate Highway System replaced many segments of the previous highway system, so many portions of Interstate Highways were still beneficial although the entire road was not yet completed. This method does leave out additional non-interstate highways that were constructed during

[3] For an interactive data visualization see `https://spatial-gini-dash.herokuapp.com/`

[4] see Figure 6 in the appendix

[5] Accessed from the FHWA website, `https://www.fhwa.dot.gov/policyinformation/hpms/shapefiles_2017.cfm`

[6] `https://www.raremaps.com/gallery/detail/38608/a-pictorial-map-of-the-united-states-of-america-show`

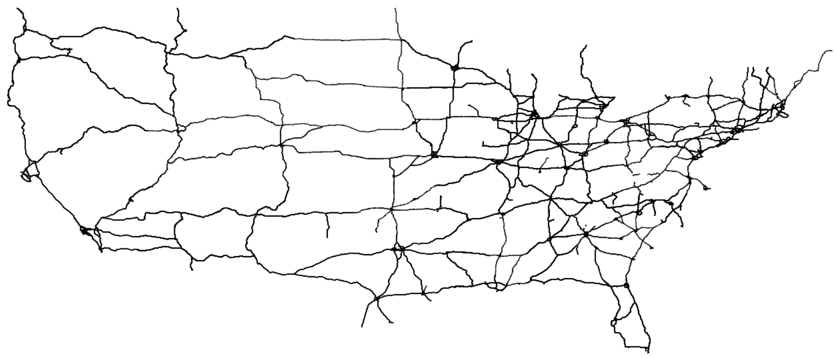

Figure 2: The US Interstate Highway System

this period, which biases the travel time reduction estimates downward.

Next, using the "PR-511" dataset, a construction log[7] detailing the completion date of each Interstate segment, the active segments of Interstate Highway are overlaid with the pre-existing highway system to construct a representation of the total highway system for each year between 1953-1994.[8]

Finally, with the highway system in place and converted to a network, the Dijkstra algorithm[9] finds the shortest weighted path between any two points in the network to estimate the travel time for each year. The weights on each road segment are the travel time based on the distance and speed. 65 mph is assumed for Interstate Highways; 50 mph is assumed for the non-interstate highways, differing slightly from the assumptions made in Jaworski et al (2018)[10]. This is done for every metropolitan-statistical-area (MSA) pair to generate a travel time matrix for each year. Figure 3 shows the average of this travel time index matrix for each year. On average the Interstate Highway System reduced travel times between MSA's

[7]This dataset was digitized and made available by Baum-Snow (2007), available here `https://www.dropbox.com/s/wq5cp6gm4ocxjo4/CD-ROM.rar?dl=0`

[8]The PR-511 has a range of statuses 1-6. Status 1 is fully complete and up to standards. Status 2 is mostly complete and open to traffic, and this is the measure of completion used.

[9]I use a combination of 'networkx' to shape the network, and 'igraph' in python to implement

[10]These speed assumptions are a simplification based on travel time estimates given by AAA maps from 1955, 1996, and 2018, to isolate the speed changes from the road and vehicle improvements. Routes without an interstate segment experienced a rise in speed of about 5mph, likely from improvements in car technology, while routes receiving interstate segments experienced rises in speed between 10-20mph. Thanks to John King for providing his personal copy of the 1955 AAA map.

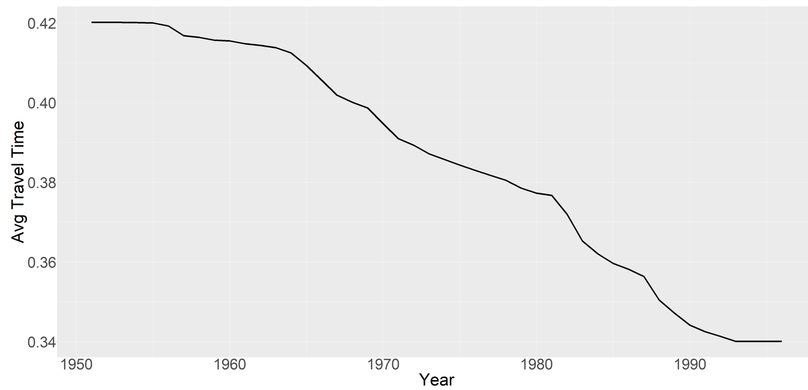

Figure 3: Average Travel Time between MSA's

by about 18%, although the actual reduction in travel time (unobserved) is partially due to vehicle improvements[11]. By assuming the same speed for all years, this number more accurately reflect the improvement coming purely from the new roads.

Using detailed BEA data on county earnings by industry I construct a measure of spatial inequality over time using the same principle as the GINI coefficient of income inequality. For each industry, I calculate the share of the nation's income from each of the 3081 US counties. Arranging these in order of lowest to highest forms a cumulative distribution, from which the GINI coefficient can be calculated. Figure 4 shows such a distribution for farming and non-farming income in 1969. The non-farming income is more bowed in, revealing the income is concentrated in fewer counties, indicating a higher degree of spatial inequality. The spatial GINI is calculated for each industry for each year between 1969-2000[12], as shown in Figure 7 in Appendix B. One issue in using this data is how the industries are categorized and where the earnings are attributed versus where the economic activity takes place. The earnings by industry are based on census data and taxes compiled by the Bureau of Labor Statistics and further processed by the BEA before being published for the public; from the users' perspective how the earnings are compiled into different industries is unknown

[11]There were some policy changes during this period that may interfere with these estimates–the National Speed Limit established in 1974 and the Motor Carrier Act of 1980. I discuss these in the appendix.

[12]for several observations negative earnings appear, heavily distorting the GINI. These observations are set to zero. This does not change the GINI significantly other than cases where the GINI was greater than one.

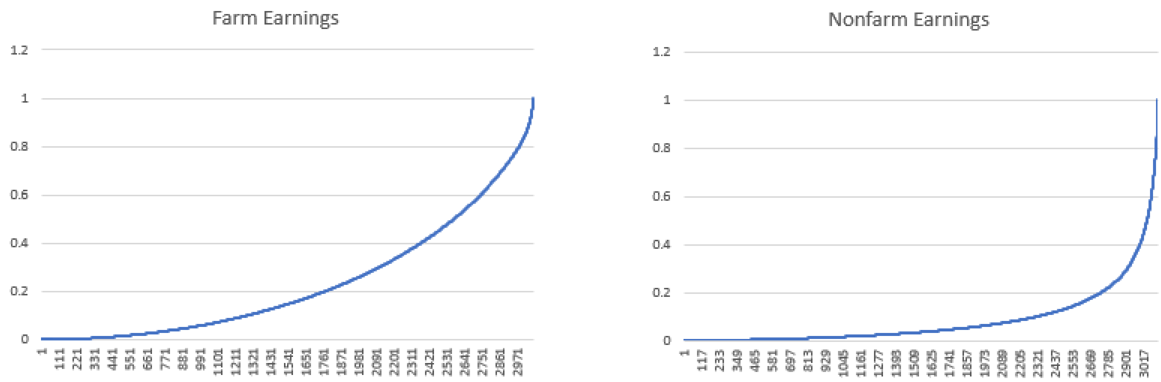

Figure 4: Cumulative Share of County Farm and Nonfarm Earnings in 1969

and could represent multiple possibilities given large firms often do many different things, for instance accounting activity done by a trucking firm internally versus an outside agency could seemingly be attributed to either truck transportation or accounting.

The truck transportation share of inputs is calculated from the BEA input-output 'use' table, detailing each industries use of other industries in dollars. Ideally, we would like a measure that reflects how much an industry relies on truck transportation for both inputs and outputs, and it is not clear how this is attributed in the input-output table.[13]

The theory suggests that scaling and product life-cycle stage will impact agglomeration and dispersal. The closest measure I have to capturing this is the 'Rasmussen backward linkages'–the column sum of the 'Leontief inverse' or 'total requirements matrix'. This measure reflects the total increase in production stemming from an increase in the final demand for a particular industry because of the additional inputs required to produce it, the additional inputs required to produce those, and so on. If an industry is in late stage production with a complex supply chain involving many industries as inputs, this will appear as a higher number in this measure.

[13]The input-output use table uses the NAICS industry codes, while the BEA county earnings uses the SIC industry code. Industries were matched as best possible based on the US BLS concordance guide and unmatched industries were dropped.
https://www.bls.gov/bls/exit_BLS.htm?a=true&url=https://www.census.gov/eos/www/naics/concordances/2002_NAICS_to_1987_SIC.xls

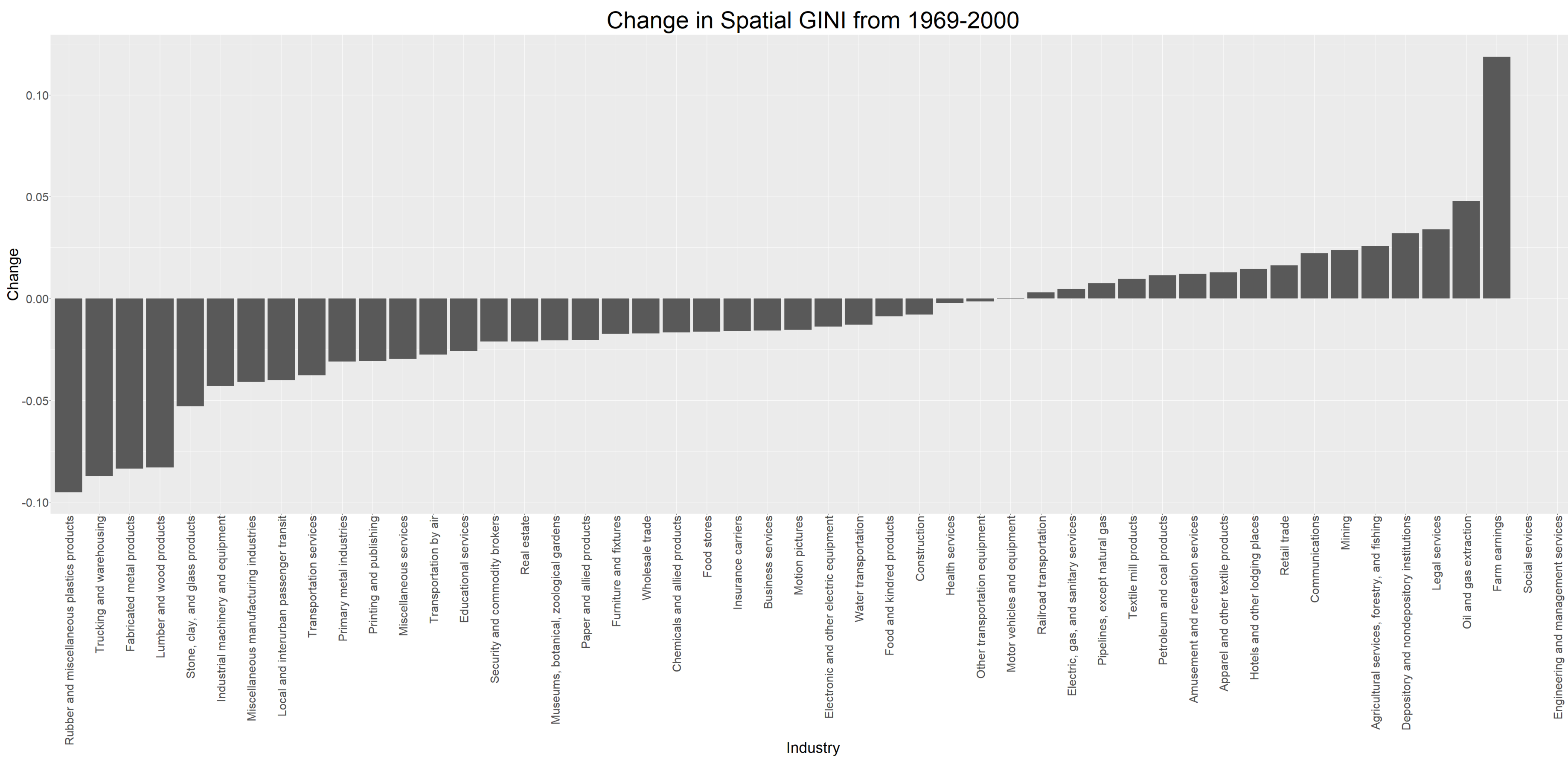

Figure 5: Change in Spatial GINI

4 Estimation

I utilize fixed effect regression with interaction terms to test if changes in travel time change the spatial GINI and if differences between industries explain the differences in the change of the spatial GINI across industries. Furthermore, I construct 'meaningful' marginal effects and standard errors as in Brambor, Clark and Golder (2014), I perform several robustness checks, and I argue quantitatively that the effect is causal as well as exploiting regional variation to identify the causal effect.

I estimate a model of the following form:

$$\text{spatialGINI}_{it} = \alpha + \alpha_i + \beta_0 tt_t + \beta_1 ts_{it} + \beta_2 bl_{it} + \beta_3 tt_t ts_{it} + \beta_4 tt_t bl_{it} + \epsilon_{it}$$

where tt_t is the index of average travel time between MSA's, ts_{it} is the truck transportation share of inputs, and bl_{it} is the Rasmussen measure of backward linkage for industry i at year t.

If tt and ts are not endogenous, the effect of reducing travel time on the spatialGINI is

$$\frac{\partial \text{spatialGINI}_{it}}{\partial tt_t} = \underset{(-)}{\beta_0} + \underset{(+)}{\beta_3 ts_{it}} + \underset{(+)}{\beta_4 bl_{it}}$$

where the signs we expect for the coefficients are noted. Conditional on a trucking input share of zero and a backwards linkage of zero, we expect the reduction in travel time to lead to an increase in the spatial GINI, that is, agglomeration. For industries with a high trucking input share and high backwards linkage, this effect will be mitigated to the point of being reversed so that a reduction in travel time leads to a decrease in the spatial GINI, that is, dispersion. Results and alternate specifications are shown below in Table 1.

We can see that a reduction in travel time is correlated with increased clustering, but for industries with a high truck transportation share of inputs and a high measure of backward linkage this is smaller and can even be negative, implying a correlation with dispersal rather than agglomeration. As detailed in Brambor, Clark and Golder (2014), when including in-

Table 1: Estimation Results

Coef.	RE	FE1	REWB	FE2	FE3	FD
α	1.20	-	1.20	-	-	.00
	(.05)		(.05)			(.00)
tt	-.81	-.82	-.82	-	-.74	.24***
	(.12)	(.12)	(.12)		(.13)	(.41)
ts	-1.31	-1.27	-1.27	-2.85	-1.23	-1.24**
	(.23)	(.23)	(.23)	(.90)	(.23)	(.73)
bl	-.15	-.16	-.15	-.18**	-.15	.14**
	(.02)	(.02)	(.02)	(.1)	(.02)	(.07)
*tt***ts*	3.58	3.50	3.50	6.85	3.45	3.69**
	(.56)	(.56)	(.56)	(2.44)	(.56)	(2.01)
*tt***bl*	.42	.42	.42	.52**	.43	-.30***
	(.06)	(.06)	(.06)	(.28)	(.06)	(.19)
$R^2_{adj,within}$	.09	.04	.09	.03	.06	.03
R^2_{adj}	.09	.96	.09	.03	.06	.03
DWH test	fail					
BP test	fail	fail	fail	fail	pass[?]	fail
BG test	fail	fail	fail	fail	fail	fail

Signif. codes: .01 ' ' .05 '*' .1 '**' 1 '***'

FE1 is individual 'within' fixed effects

FE2 is time 'within' fixed effects

FE3 is two ways 'within' fixed effects

FD is first difference

REWB is random effects with a control term for average industry truck share of inputs across time, this specification is included because of discussion in Bell and Jones (2015) about random effects vs fixed effects and controlling for heterogeneity bias

After running heteroscedastic and serially correlated robust SE the coef's are still significant, although that is not shown here teraction terms for testing conditional hypotheses, care must be taken in the implementation and interpretation of the results. Specifically, the constitutive effects must be included and must not be interpreted as unconditional marginal effects, and 'meaningful' marginal effects and standard errors should be reported. That is, for the specification above, the appropriate standard error formulation for the marginal effect of travel time is shown below. These standard errors and marginal effects are shown by industry in Figure 9 in Appendix C. The average z-score of the marginal effect of tt across time and across industries is 4.08 with a standard deviation of 1.99, indicating that the estimate is statistically significant for most industries most of the time.

$$\hat{\sigma}_{sg_{it}} = \sqrt{var(\hat{\beta}_0) + ts_{it}^2 var(\hat{\beta}_3) + bl_{it}^2 var(\hat{\beta}_4) + 2ts_{it}cov(\hat{\beta}_0\hat{\beta}_3) + 2bl_{it}cov(\hat{\beta}_0\hat{\beta}_4) + 2ts_{it}bl_{it}cov(\hat{\beta}_3\hat{\beta}_4)}$$

When regressing non-stationary trends spurious correlation is a major concern, however in this case I find it appropriate for a few reasons. First, because firms are forward looking, the road construction was generally known in advance, the plant lifetimes can potentially be very long, and there are potential benefits to being a first mover, it seems probable that some firms would relocate or expand operations in anticipation of the road completion. On the other hand, relocating is expensive, and firms may prefer to postpone relocation or expansion as the desirability of locations depends on the changing travel times as well as the locations of other firms. That is, the effect of the changing travel time index could lead or lag behind the effect on spatial GINI and the timing could vary by industry. This is supported by cross-correlation results between the industry specific spatial GINI's and the travel time index (see Figure 10 in the appendix). Because of this, transforming the series with first difference requires the regression to precisely specify the leads and lags structure, a problem noted in Vaisey and Miles (2014). By regressing the levels and not specifying leads or lags however, the long run effect is captured. I perform simulations with artificial data to verify the efficacy of this specification, finding that the levels regression with only contemporaneous variables accurately estimates the true long run effect regardless of the leads and lags distribution, while the first difference regression parameter estimates are extremely sensitive to the lead and lag specification. See the Appendix C for more information on this issue and the simulation results. Second, while a parallel trend could seemingly be driving the results, including a dummy variable for year does not substantially change the outcome (seen as FE2 in the regression table), nor does including a time trend. Additionally, leaving out *tt* entirely and replacing it with a simple time trend does not yield the same results, suggesting the movements in *tt* are meaningful beyond it's trend component. See Table 2 in the appendix for these regression results.

The spatial GINI discussed so far is based on county level earnings, but the land area of counties varies significantly, which could obscure the change in clustering when economic activity moves between counties of different sizes. To account for this I compute another set of spatial GINI's based on county earnings per land area, but these do not change the result significantly. Additionally, I construct alternate measures of spatial inequality: the Theil index and the 80:40 ratio, and find the same result. I also add controls for the boating, rail, and air transport shares of input but find these are not statistically significant and do not alter the main findings. The results of these robustness regressions can be seen in Table 3 in Appendix C.

However robust, these are still merely correlations, and two confounding sources present themselves: it's possible the regressors *ts* and *bl* are correlated with *tt* (if industries were restructuring their operations in response to new roads for instance) biasing the treatment estimates, and because the change in *tt* is monotonic, it's possible there is a parallel trend (such as technology change leading to industry restructuring) that is actually responsible for the change in spatial distribution of industries.

Considering the first, the change in spatial GINI from a change in travel time would be

$$\frac{\partial \text{spatialGINI}_{it}}{\partial tt_t} = \beta_0 + \beta_3 ts_{it} + \beta_4 bl_{it} + \beta_1 \frac{\partial ts_{it}}{\partial tt_t} + \beta_2 \frac{\partial bl_{it}}{\partial tt_t} + \beta_3 tt_{it} \frac{\partial ts_{it}}{\partial tt_t} + \beta_4 tt_{it} \frac{\partial bl_{it}}{\partial tt_t}$$

To refute this, we observe that the change in trucking input share and backward linkages is very low across time. The largest mean normalized variance (index of dispersion) across time among industries for ts is .017, while the mean is .0018, both of which are considered to be not very dispersed. For bl across industries the largest mean normalized variance across time is .022 and the mean is .0055, which again is not very dispersed. Because these two terms are changing very little across time we can consider the last four terms in the differential equation to be zero. See Figures 11 and 12 in the appendix.

The second claim is more difficult to refute; I consider two pieces of evidence. First, we note that the Interstate Highway was completed in 1993, and the spatial GINI is not changing much after the year 2000. The variance of the change in spatial GINI from 1969 to 2000 across industries is .0014, while it is .00063 from 2001 to 2017. Furthermore, the average of the absolute value of the change in spatial GINI is .028 and .013 for 1969-2000 and 2001-2017 respectively. The spatial GINI across the entire time frame is shown by industry in Reference to Figure 8 in the appendix. If there was a parallel trend driving the change in spatial GINI from 1969-2000 it would also have had to ended around the same time frame as the Interstate Highway completion.

Second, I will attempt to identify a causal effect by looking at *regional and state level variation in the timing of road completion and spatial GINI response*. If states that complete their highway portion earlier also agglomerate/disperse earlier, than this suggests that the change is due to the road completion. This identification strategy will be valid unless the unobserved parallel trend also varies at the state level in the same way as completion timing, or if there are unobserved state specific variables changing that happen to cause a change in spatial distribution at the same time the roads are being completed. This section is pending on further data from the BEA.[14]

5 Conclusion

Industries are subject to economic centripetal and centrifugal forces influencing the patterns of their relative positions in space. Differences between industries will result in differing sensitivities to these forces. As the road system is improved, both agglomeration and dispersal are facilitated, leading to some industries clustering more densely in fewer counties and some industries spreading throughout more counties. These differing responses can partially be explained by truck transportation utilization and backward linkages.

[14] in order to protect business confidentiality, many county earnings are suppressed for certain industries. The suppression rate in a given year varies from less than 5% to 50% depending on the industry. This should not interfere with the overall patterns of spatial distribution but when looking at the state or regional level these suppressions become a significant issue generating movements in the data that are more a product of suppression policy change than actual industry relocation.

Figures

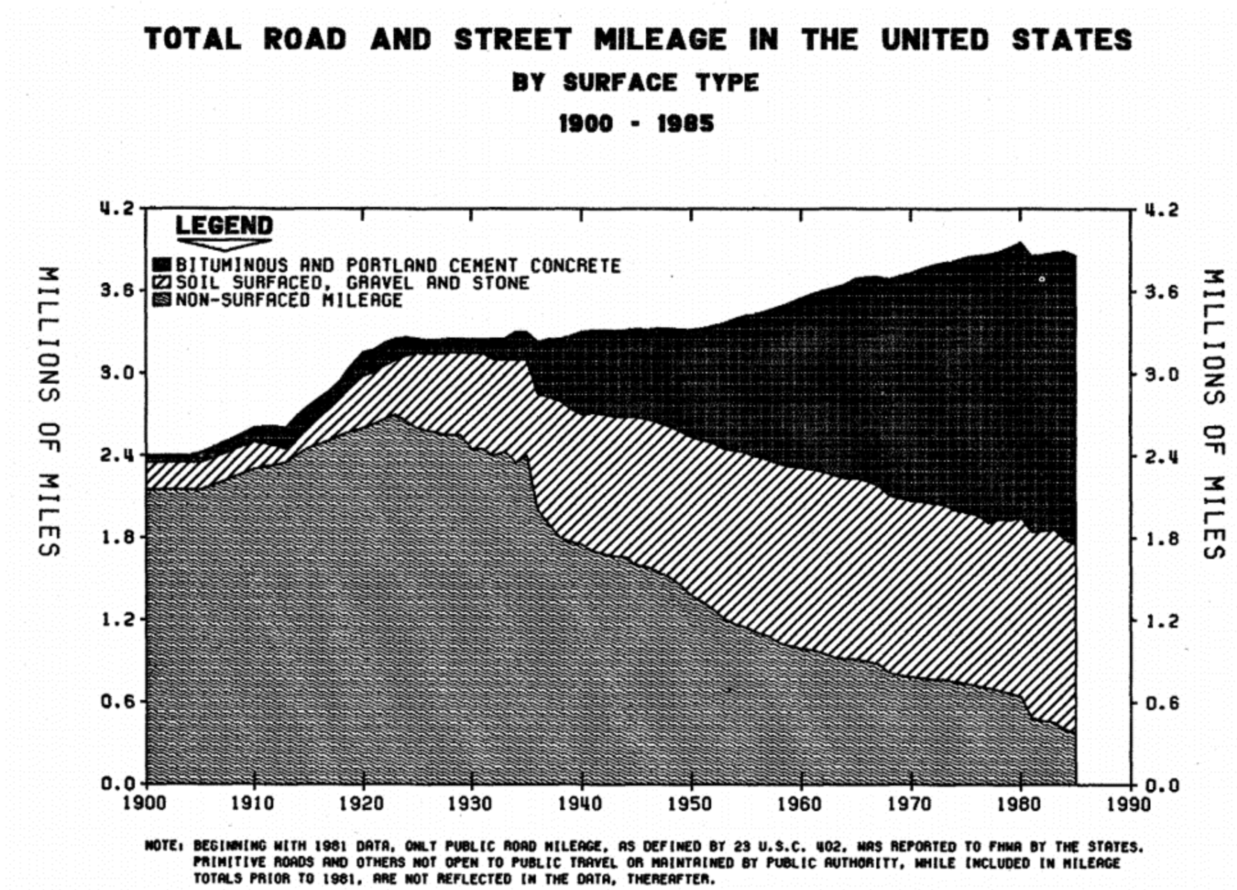

Figure 6: Surface Quality of US Roads, Source: FHWA

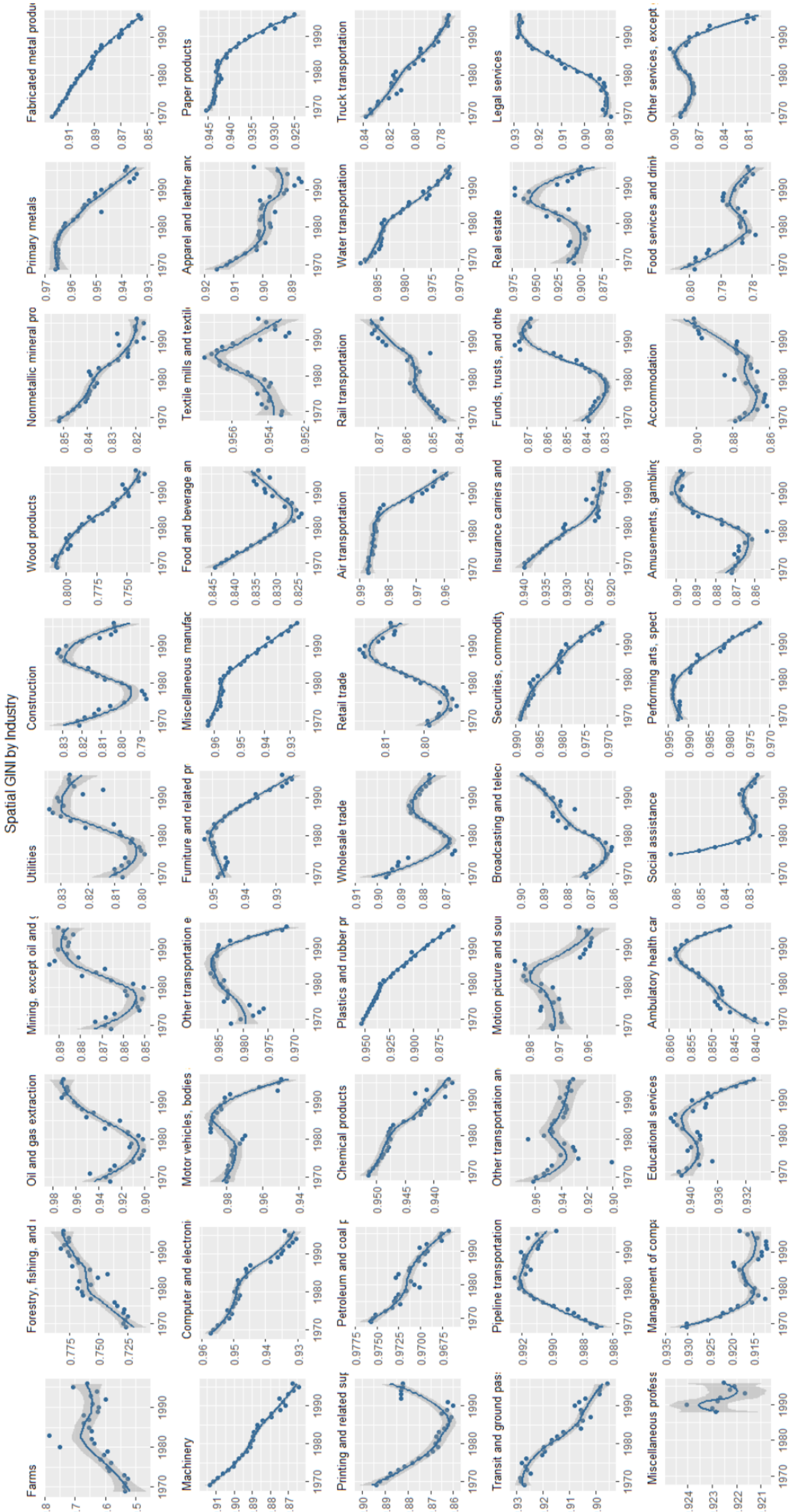

Figure 7: Spatial GINI by Industry

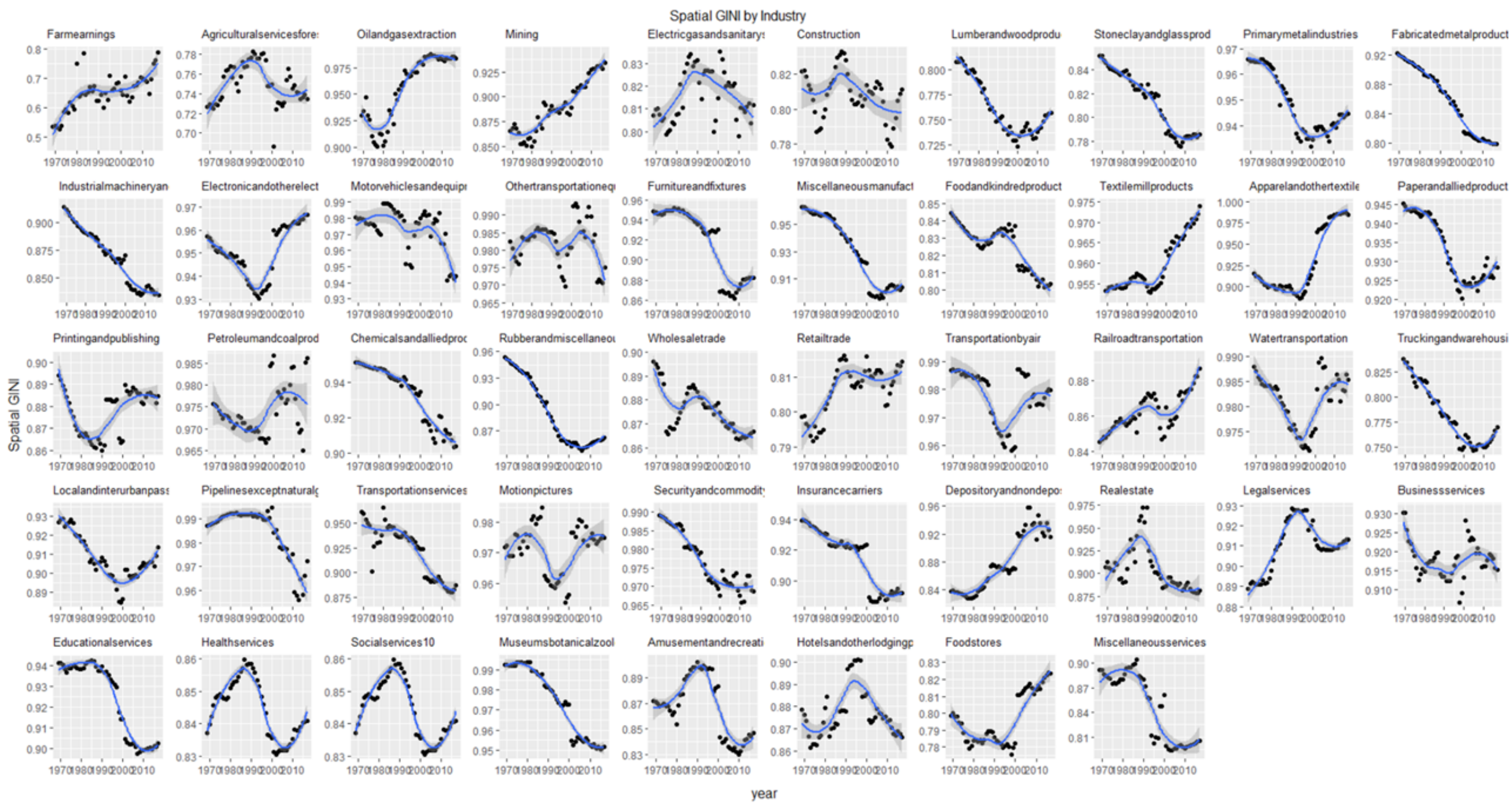

Figure 8: Spatial GINI by Industry before and after 2000

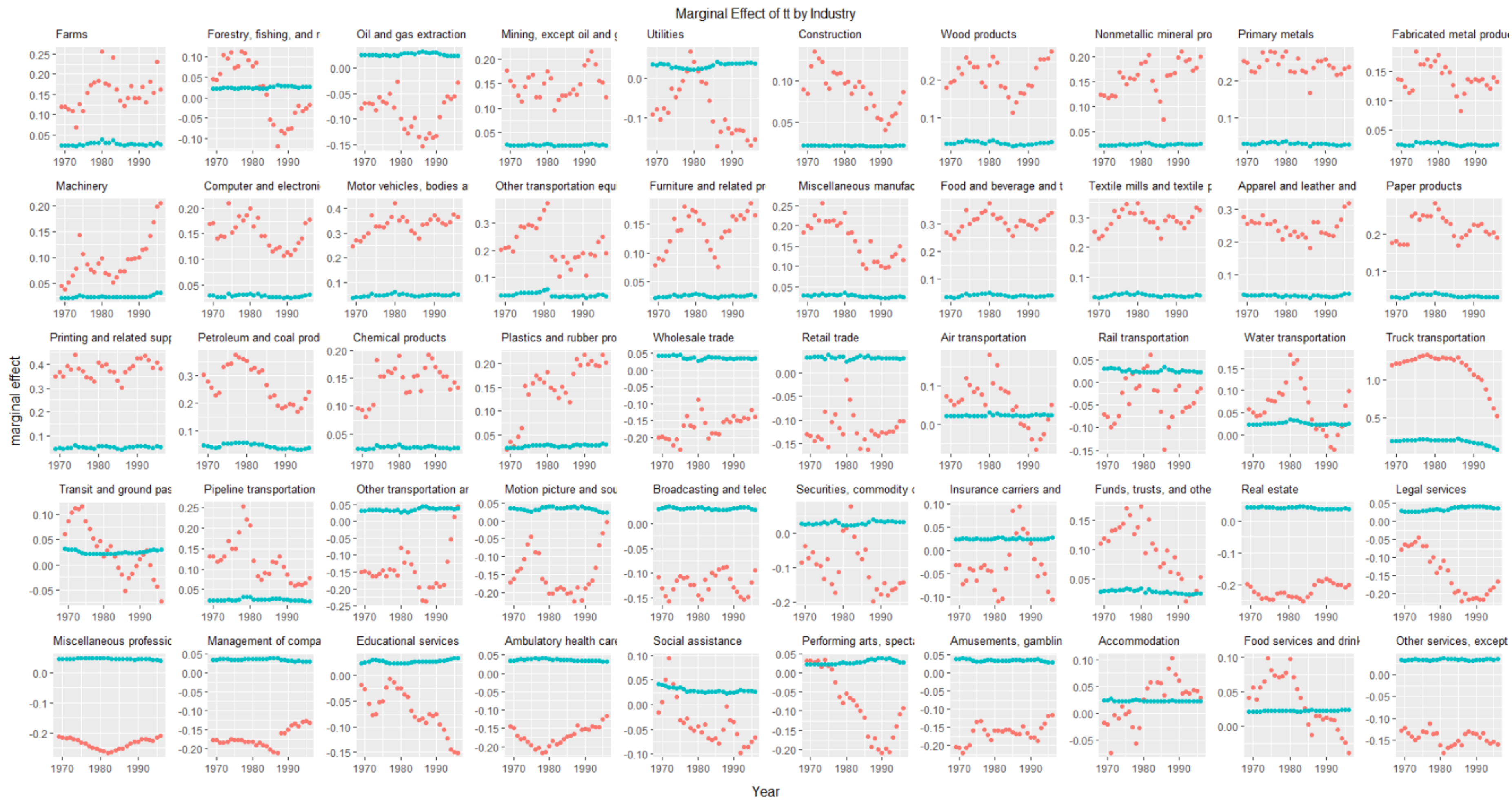

Figure 9: Marginal Effect of Travel Time on Spatial GINI and Standard Errors by Industry

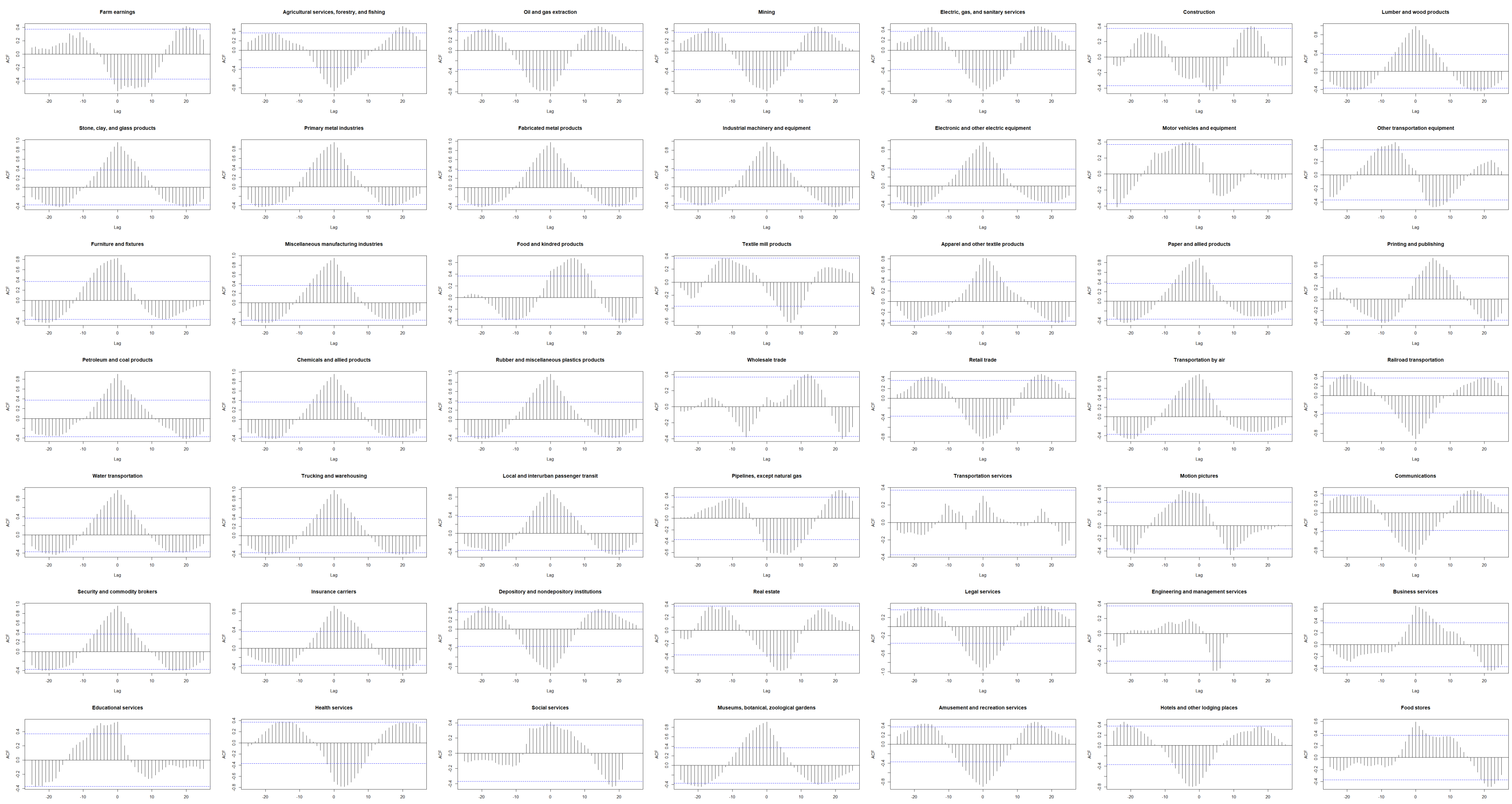

Figure 10: Cross Correlations for Lagged Values of Travel Time and Industry Spatial GINI

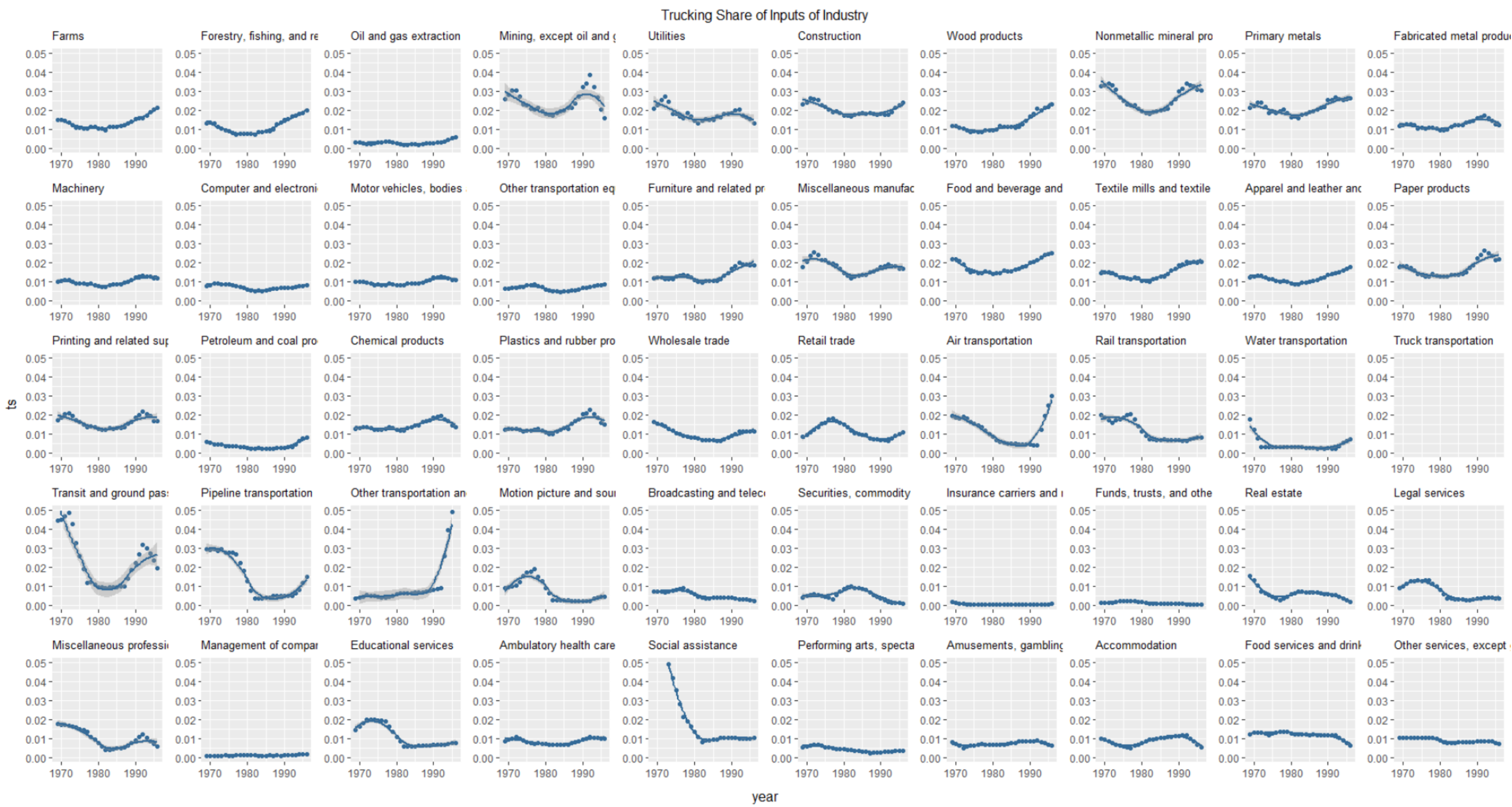

Figure 11: Trucking Share of Inputs by Industry

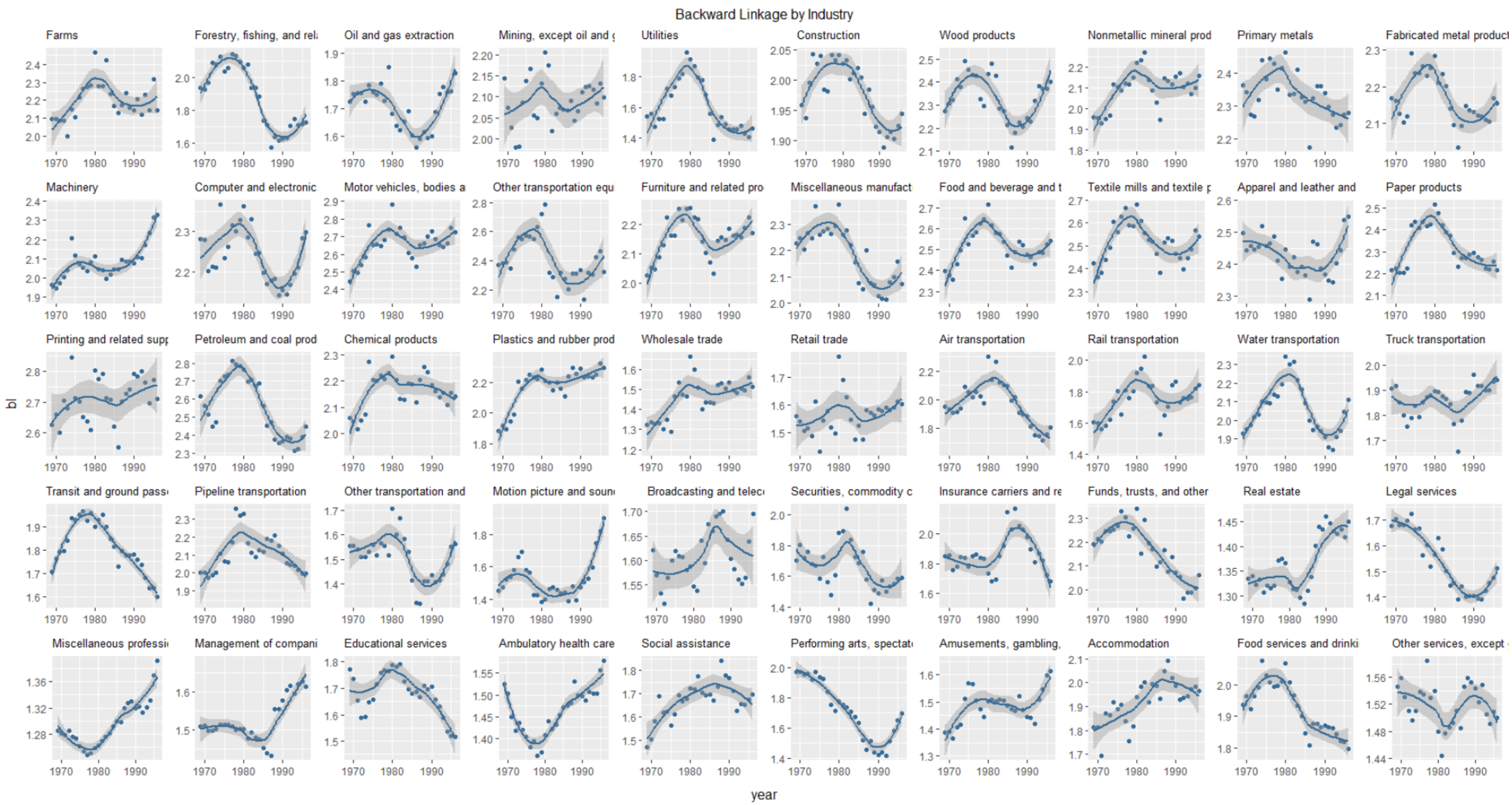

Figure 12: Rasmussen Measure of Backwards Linkages by Industry

Name(SIC)	mean(sGINI)	sGINI 69-96	mean(ts)	ts 69-96	mean(bl)	bl 69-96
Farmearnings	0.625	0.123	0.013	0.007	2.197	0.048
Agriculturalservicesforestryandfishing	0.757	0.054	0.012	0.007	1.88	-0.21
Oilandgasextraction	0.936	0.041	0.003	0.003	1.706	0.102
Mining	0.872	0.025	0.024	-0.01	2.089	-0.046
Electricgasandsanitaryservices	0.816	0.02	0.018	-0.008	1.606	-0.079
Construction	0.811	-0.018	0.02	0.001	1.978	-0.015
Lumberandwoodproducts	0.776	-0.064	0.013	0.012	2.332	0.129
Stoneclayandglassproducts	0.833	-0.032	0.026	-0.002	2.095	0.201
Primarymetalindustries	0.955	-0.028	0.021	0.005	2.334	-0.084
Fabricatedmetalproducts	0.892	-0.066	0.012	0	2.166	-0.014
Industrialmachineryandequipment	0.888	-0.046	0.01	0.002	2.085	0.364
Electronicandotherelectricequipment	0.945	-0.023	0.007	0	2.242	0.016
Motorvehiclesandequipment	0.977	-0.031	0.01	0.001	2.655	0.275
Othertransportationequipment	0.982	-0.011	0.007	0.002	2.406	-0.046
Furnitureandfixtures	0.943	-0.02	0.013	0.007	2.157	0.145
Miscellaneousmanufacturingindustries	0.95	-0.036	0.018	-0.001	2.185	-0.154
Foodandkindredproducts	0.833	-0.01	0.018	0.003	2.518	0.143
Textilemillproducts	0.955	0.001	0.015	0.006	2.514	0.118
Apparelandothertextileproducts	0.9	-0.013	0.012	0.006	2.428	0.052
Paperandalliedproducts	0.939	-0.02	0.017	0.004	2.317	-0.002
Printingandpublishing	0.874	-0.012	0.016	-0.001	2.71	0.083
Petroleumandcoalproducts	0.971	-0.009	0.004	0.002	2.554	-0.167
Chemicalsandalliedproducts	0.945	-0.014	0.014	0.001	2.164	0.082
Rubberandmiscellaneousplasticsproducts	0.916	-0.094	0.014	0.002	2.175	0.413
Wholesaletrade	0.88	-0.019	0.01	-0.005	1.458	0.191
Retailtrade	0.804	0.01	0.011	0.002	1.564	0.043
Transportationbyair	0.979	-0.024	0.012	0.01	1.97	-0.14
Railroadtransportation	0.858	0.024	0.012	-0.012	1.753	0.237
Watertransportation	0.98	-0.016	0.004	-0.01	2.057	0.182
Truckingandwarehousing	0.803	-0.065	0.333	-0.195	1.861	0.034
Localandinterurbanpassengertransit	0.913	-0.031	0.022	-0.025	1.815	-0.104
Pipelinesexceptnaturalgas	0.991	0.003	0.014	-0.015	2.106	-0.004
Transportationservices	0.941	-0.03	0.011	0.053	1.506	0.007
Motionpictures	0.971	-0.009	0.007	-0.005	1.527	0.444
Communications	0.877	0.028	0.005	-0.005	1.606	0.074
Securityandcommoditybrokers	0.982	-0.018	0.005	-0.003	1.659	-0.111
Insurancecarriers	0.928	-0.019	0	-0.001	1.862	-0.167
Depositoryandnondepositoryinstitutions	0.848	0.03	0.001	-0.001	2.165	-0.129
Realestate	0.919	-0.008	0.006	-0.014	1.365	0.115
Legalservices	0.908	0.039	0.007	-0.005	1.538	-0.166
Engineeringandmanagementservices11	0.923	X	0.01	-0.012	1.296	0.107
Businessservices	0.918	-0.012	0.001	0.001	1.524	0.104
Educationalservices	0.939	-0.01	0.011	-0.007	1.68	-0.256
Healthservices	0.851	0.009	0.008	0.001	1.461	0.054
Socialservices10	0.833	X	0.021	-0.041	1.672	0.229
Museumsbotanicalzoologicalgardens	0.987	-0.02	0.004	-0.002	1.709	-0.275
Amusementandrecreationservices	0.881	0.026	0.007	-0.002	1.484	0.225
Hotelsandotherlodgingplaces	0.878	0.022	0.009	-0.005	1.922	0.15
Foodstores	0.787	-0.017	0.012	-0.006	1.925	-0.141
Miscellaneousservices	0.875	-0.082	0.009	-0.003	1.519	-0.046

Name(SIC)	Name(NCIS)
Farmearnings	Farms
Agriculturalservicesforestryandfishing	Forestryfishingandrelatedactivities
Oilandgasextraction	Oilandgasextraction
Mining	Miningexceptoilandgas
Electricgasandsanitaryservices	Utilities
Construction	Construction
Lumberandwoodproducts	Woodproducts
Stoneclayandglassproducts	Nonmetallicmineralproducts
Primarymetalindustries	Primarymetals
Fabricatedmetalproducts	Fabricatedmetalproducts
Industrialmachineryandequipment	Machinery
Electronicandotherelectricequipment	Computerandelectronicproducts
Motorvehiclesandequipment	Motorvehiclesbodiesandtrailersandparts
Othertransportationequipment	Othertransportationequipment
Furnitureandfixtures	Furnitureandrelatedproducts
Miscellaneousmanufacturingindustries	Miscellaneousmanufacturing
Foodandkindredproducts	Foodandbeverageandtobaccoproducts
Textilemillproducts	Textilemillsandtextileproductmills
Apparelandothertextileproducts	Apparelandleatherandalliedproducts
Paperandalliedproducts	Paperproducts
Printingandpublishing	Printingandrelatedsupportactivities
Petroleumandcoalproducts	Petroleumandcoalproducts
Chemicalsandalliedproducts	Chemicalproducts
Rubberandmiscellaneousplasticsproducts	Plasticsandrubberproducts
Wholesaletrade	Wholesaletrade
Retailtrade	Retailtrade
Transportationbyair	Airtransportation
Railroadtransportation	Railtransportation
Watertransportation	Watertransportation
Truckingandwarehousing	Trucktransportation
Localandinterurbanpassengertransit	Transitandgroundpassengertransportation
Pipelinesexceptnaturalgas	Pipelinetransportation
Transportationservices	Othertransportationandsupportactivities
Motionpictures	Motionpictureandsoundrecordingindustries
Communications	Broadcastingandtelecommunications
Securityandcommoditybrokers	Securitiescommoditycontractsandinvestments
Insurancecarriers	Insurancecarriersandrelatedactivities
Depositoryandnondepositoryinstitutions	Fundstrustsandotherfinancialvehicles
Realestate	Realestate
Legalservices	Legalservices
Engineeringandmanagementservices11	Miscellaneousprofessionalscientificandtechnicalservices
Businessservices	Managementofcompaniesandenterprises
Educationalservices	Educationalservices
Healthservices	Ambulatoryhealthcareservices
Socialservices10	Socialassistance
Museumsbotanicalzoologicalgardens	Performingartsspectatorsportsmuseumsandrelatedactivities
Amusementandrecreationservices	Amusementsgamblingandrecreationindustries
Hotelsandotherlodgingplaces	Accommodation
Foodstores	Foodservicesanddrinkingplaces
Miscellaneousservices	Otherservicesexceptgovernment

Appendix C. Robustness

	Dependent variable:			
	g			
	(1)	(2)	(3)	(4)
tt	-0.817***	-1.189***	-2.533***	
	(0.119)	(0.185)	(0.731)	
ts	-1.268***	-1.294***	-0.853	0.135***
	(0.230)	(0.230)	(1.237)	(0.041)
bl	-0.155***	-0.151***	-0.432***	0.014***
	(0.023)	(0.023)	(0.147)	(0.005)
Year:ts				-0.009***
				(0.001)
Year:bl				-0.001***
				(0.0001)
Year		-0.001***	-0.004**	0.002***
		(0.0003)	(0.002)	(0.0003)
tt:ts	3.502***	3.554***	2.462	
	(0.560)	(0.559)	(3.074)	
tt:bl	0.421***	0.420***	1.116***	
	(0.060)	(0.060)	(0.365)	
ts:Year			-0.003	
			(0.008)	
bl:Year			0.002*	
			(0.001)	
Constant	0.926***	1.068***	1.611***	0.597***
	(0.045)	(0.070)	(0.294)	(0.010)
Observations	1,375	1,375	1,375	1,375
R^2	0.960	0.960	0.960	0.960
Adjusted R^2	0.958	0.959	0.959	0.958
Residual Std. Error	0.015 (df = 1320)	0.015 (df = 1319)	0.015 (df = 1317)	0.015 (df = 1320)

Figure 13: Table 2: Regressions with a Time Trend

Table 2: Estimation Results

	Controls		Without bl		Theil		80:40	
Coef.	RE	FE1	RE	FE1	RE	FE1	RE	FE1
α	1.09	-	.89	-	6.55	-	226	-
	(.05)		(.04)		(.85)		(13.6)	
tt	-.57	-.56	-.04***	-.05***	-13.9	-13.59	-621	-617
	(.13)	(.13)	(.04)	(.04)	(2.25)	(2.24)	(36.3)	(36.3)
ts	-.75	-.63	-.70	-.58*	-10.0*	-7.77**	-108**	-81.0***
	(.24)	(.24)	(.24)	(.24)	(4.22)	(4.35)	(64.7)	(69.4)
bs	7.89*	7.96*	8.58*	8.65*	84.0***	80.4***	-4060	-3940
	(3.56)	(3.56)	(3.58)	(3.58)	(63.9)	(63.9)	(1000)	(1010)
rs	3.18	3.61	2.64*	3.10	31.4***	42.7*	-850	-750*
	(1.12)	(1.12)	(1.12)	(1.12)	(19.85	(20.1)	(308)	(319)
as	1.07***	1.19***	1.96***	2.05**	104	105	-21.3***	-31.9***
	(1.22)	(1.22)	(1.21)	(1.21)	(20.9)	(20.8)	(347)	(347)
bl	-.10	-.10	-	-	-3.09	-2.93	-72.6	-68.9
	(.02)	(.02)			(.43)	(.43)	(6.84)	(6.92)
*tt*ts*	2.27	2.02	2.10	1.86	24.5*	19.6**	295**	239***
	(.59)	(.59)	(.59)	(.59)	(10.4)	(10.6)	(161)	(168)
*tt*bs*	-21.0*	-21.2*	-23.0*	-23.1*	-248***	-239***	10600	10400
	(9.03)	(9.02)	(9.1)	(9.06)	(162)	(162)	(2540)	(2560)
*tt*rs*	-4.35***	-5.06**	-2.30***	-3.10***	-93.2**	-111*	2150	2040
	(2.75)	(2.75)	(2.72)	(2.72)	(49.1)	(49.2)	(771)	(781)
*tt*as*	-2.75***	-3.09***	-5.1***	-5.36***	-291	-296	37.4***	49.5***
	(3.29)	(3.29)	(3.26)	(3.25)	(56.3)	(56.1)	(933)	(935)
*tt*bl*	.28	.27	-	-	9.30	9.00	211	206
	(.06)	(.06)			(1.13)	(1.13)	(18.2)	(18.3)
R^2_{adj}	.13	.09	.12	.08	.13	.11	.44	.44

FE1 is individual 'within' fixed effects

Removing the trucking industry makes ts large significant in 80:40 and Theil (ts outlier)

Signif. codes: .01 ' ' .05 '*' .1 '**' 1 '***'

References

[1] `https://cdn.vanderbilt.edu/vu-my/wp-content/uploads/sites/2318/2019/07/29170757/ddtiming_7_29_2019.pdf`

[2] Panzer and Postiglione (2019)

[3] Rey and Smith (2012)

[4] Rothenberg, Alexander. *Transport Infrastructure and Firm Location Choice in Equilibrium: Evidence from Indonesia's Highways*. Manuscript, 2013.

[5] Eriksson, Katherine, Katheryn Russ, Jay C. Shambaugh and Minfei Xu. *Trade Shocks and the Shifting Landscape of U.S. Manufacturing*. NBER Working Paper, 2019.

[6] Frye, Dustin. *Transportation Networks and the Geographic Concentration of Industry*. Working Paper, 2016.

[7] Jaworski, Taylor, Carl T. Kitchens and Sergey Nigai. *The Interstate Highway System and the Development of the American Economy*. Preliminary draft, 2018.

[8] Faber, Benjamin. *Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System* Review of Economic Studies, 2014.

[9] Redding, Stephen and Matthew Turner. *Transportation Costs and the Spatial Organization of Economic Activity*. NBER Working Paper, 2014.

[10] Ellison, G. and E. L. Glaeser (1997): "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach," *Journal of Political Economy*, 105, 889–927

[11] Alder, Simon. *Chinese Roads in India: The Effect of Transport Infrastructure on Economic Development*. 2016.

[12] Donaldson, Dave and Richard Hornbeck. *Railroads and American Economic Growth: A "Market Access" Approach*. The Quarterly Journal of Economics, 2016.

[13] Baum-Snow, Nathaniel. *Did Highways Cause Suburbanization?*. The Quarterly Journal of Economics, 2007.

[14] Brambor, Clark, and Golder (2006)

[15] Bell and Jones (2015)

[16] Fujita, Masahisa, Paul Krugman, Anthony J. Venables. *The Spatial Economy*. MIT Press, 1999.

[17] von Thunen, Johann Heinrich. *The Isolated State*. Hamburg: Perthes, 1829. English translation. Oxford: Pergamon, 1966.

[18] Marshall, Alfred. *Principles of Economics*. Macmillan, London, 1890.

[19] Alonso, William. *Location and Land Use*. Harvard University Press, Cambridge MA, 1960.

[20] Fernald, John. *Roads to Prosperity? Assessing the Link between Public Capital and*

Productivity American Economic Review, 1999.

[21] Stole, Lars *Price Discrimination and Competition.* Handbook of Industrial Organization, 2007.

[22] Michaels, Guy, Ferdinand Rauch, and Stephen Redding. *Task Specialization in U.S. Cities from 1880-2000.* NBER Working Paper, 2013.

[23] Allen, Treb and Costas Arkolakis. *Trade and the Topography of the Spatial Economy.* The Quarterly Journal of Economics, 2014

[24] Duranton, Gilles, Peter Morrow, and Matthew Turner. *Roads and Trade: Evidence from the US* The Review of Economic Studies, 2014.

[25] Eaton, Jonathon and Samuel Kortum. *Technology, Geography, and Trade.* Econometrica, 2002.

[26] Redding, Stephen and Anthony J. Venables. *Economic Geography and international inequality.* Journal of International Economics, 2004.

note this list is incomplete