REGION
The Journal of ERSA
Powered by WU

Volume 9, Number 1, 2022, 1–27
DOI: 10.18335/region.v9i1.366

journal homepage: region.ersa.org
ISSN: 2409-5370

Estimating Determinants of Transportation and Warehousing Establishment Locations Using U.S. Administrative Data

Craig Carpenter[1], Rebekka Dudensing[2], Anders Van Sandt[3]

[1] Michigan State University, East Lansing, MI, USA
[2] Texas A&M University, College Station, TX, USA
[3] University of Wyoming, Laramie, WY, USA

Received: 26 April 2021/Accepted: 13 December 2021

Abstract. Interactions between transportation and warehousing and other industry clusters are not widely explored and the determinants of logistics locational determinants is limited in the U.S. context. These gaps in the literature, along with the U.S. transportation and warehousing sector's decentralization from urban areas and concentration in regions, highlight the importance of understanding the effects of place-based factors and inter-industry clusters on the locations and employment of transportation and warehousing industries. The analysis uses restricted-access U.S. Census Bureau data aggregated to the county level, along with secondary data sources, to estimate the locational determinants of transportation and warehousing (TW) industries based on transportation infrastructure as well as sociodemographic and institutional variables. The analysis takes a cross-sectional (non-causal) approach to focus on time-invariant location factors while testing and implementing zero-inflated count data distributions to model the data generation processes more accurately. Results indicate that subsectors are affected differently by infrastructure, sociodemographic, and institutional variables. Additionally, different factors are associated with industry presence versus size. Finally, we show that data using aggregated industries obscures locational factors' importance for individual sub-sectors and, further, that industrial aggregation obscures TW sectors' relationships to other clusters.

1 Introduction

Transportation and warehousing (TW) industries provide or support the transport of passengers and storage of cargo via roads, rails, water, air, and pipelines. Given these modes of transportation, and in addition to local labor force composition, certain locations may offer comparative advantages to local industries based on their economic and place-based assets, leading to clusters of transportation and warehousing establishments (Kang 2020, Ng, Gujar 2009)[1]. Porter (2000, 2001) popularized the idea that inter-related industries harness regional competitive advantage to strengthen innovation, productivity, and other economic outcomes. Transportation infrastructure is often an institutional factor used to predict or describe economic outcomes.

[1]This article necessarily uses many abbreviations and acronyms. In addition to in-text first-use definitions, we provide a list of acronyms in appendix Table A.1 for reference.

Synergistic effects between transportation and warehousing and other industry clusters are not widely explored, and in general, the literature on logistics locational determinants is fairly limited in the U.S. context (Rivera et al. 2014, 2016). These gaps in the literature, along with the U.S. transportation and warehousing sector's decentralization from urban areas and concentration in some U.S. regions (Cidell 2010), highlights the importance of understanding the effects of place-based factors and inter-industry clusters on the locations and employment of transportation and warehousing industries.

While transportation topics have been more explored in the European context, less attention has been given to the potentially different U.S. transportation context, likely due to prior data limitations. Fewer inter-regional differences in transportation policy exist between U.S. states than European countries (Rodrigue, Notteboom 2010), which may have provided richer research opportunities regarding both transportation and its relationship to economic competitiveness. Additionally, state ownership of some transportation networks (Clausen, Voll 2013) is more common in Europe, which may prompt published studies of both industries and infrastructure. Regardless, understanding the relationships between TW location decisions and local economic, demographic, and infrastructure measures is pertinent to both geographies as infrastructure spending has waned in recent years in both the US and Europe, but may be increasing in the near future. Using confidential U.S. administrative data aggregated to the county level, the focus of this article is to illustrate how more refined data and unique data generation processes may be utilized to explore the influence of local industry, socioeconomic, and place-based factors on the size of local transportation and logistics (including warehousing) industries. This illustration includes accounting for clustering by examining the association of various other industries on TW location. The analysis takes a cross-sectional (associative and non-causal) approach to focus on time-invariant location factors while testing and implementing zero-inflated count data distributions to most accurately model the data generation processes of TW industries.

The following section reviews the literature on TW locational determinants and transportation's effects on economic competitiveness. The novel data and methods underlying our approach are then presented, followed by study results for selected transportation sectors. The conclusion discusses general observations about TW cluster predictors and the propensity of TW sectors to support other clusters, opening the door to future U.S. TW locational choice research.

2 Literature Review

Economic studies of transportation and warehousing (TW) tend to focus on the costs and barriers of moving either goods or people. Analyses of the benefits and costs (including external costs) of infrastructure often account for both people and freight in traffic volume (Behiri et al. 2018, Li, Madanu 2009, Mayeres et al. 1996). Researchers accept that the people and freight require different treatment, but common models are readily adapted to address either passenger or freight concerns based on quantities of goods to be transported between specific locations, distribution flows between those locations, modal splits, and assignment to transportation networks (de Jong et al. 2004). In essence, these models orient networks and transportation infrastructure to minimize transportation distances and costs.

Despite its importance in moving goods as well as labor and consumers, transportation emerged as a topic within economic geography relatively recently (Hesse, Rodrigue 2004), and the literature on TW firm location decisions, in particular, remains relatively sparse (Holl, Mariotti 2018). A comprehensive review of the extensive firm location literature is beyond the scope of this article; however, this section discusses key advances in the literature since McFadden's (1973) discrete choice work.

Much of the literature pertaining to transportation in an economic development context is focused on transportation infrastructure and its role in supporting development (Chen et al. 2016, Maparu, Mazumder 2017, Melecky et al. 2019). Demand increasingly influences supply chains and TW logistics, particularly with the rising importance of e-commerce and its associated distribution (Bowen 2008, Hesse, Rodrigue 2004, Hughes, Jackson 2015, Rodrigue 2020). Beyond road densities, transportation hubs, flows, and networks are becoming increasingly important (Crang 2002). Inland hubs and intermodal facilities are becoming more important, and warehouses are moving to places with strong multi-modal networks, especially as technology promotes larger warehouses and distribution centers, which are increasingly independent firms rather than divisions of manufacturing or retail firms (Bowen 2008).

Cidell (2010) expands on this work, using Economic Census and County Business Patterns Data to measure the concentration of warehousing and distribution establishments across the U.S. by calculating Gini coefficients for each county, although they note that employment and payroll data are not reliably available from those sources for the warehousing sector, thus limiting her choice of dependent variable. Independent variables represent demand attributes, such as population and income, and physical infrastructure attributes such as interstate and railroad miles, enplanements, and distance. Infrastructure such as highways, rail, and inland waterways were not always significant predictors of freight establishments within a metropolitan area (Cidell 2010), although other researchers have continued to find these factors critical (Guerin et al. 2021, Holl, Mariotti 2018). Overall, warehouses and distribution centers decentralized and moved toward suburbs over time. Cities with strong warehousing sectors, as measured by the Gini coefficient, tended to have more concreated establishments near the city center. Concentration early in the study period was also associated with greater decentralization within the metropolitan area over time, although many cities added the largest number of freight establishments within the central county with the densest transportation infrastructure. Other researchers have found similar results related to suburbanization (Allen et al. 2012, Cidell 2010, Dablanc et al. 2014, Holl, Mariotti 2018) and inland hub development (Bowen 2008, Holl, Mariotti 2018, Monios, Wilmsmeier 2013, Rodrigue 2020).

While cost minimization drives TW location decisions, lower transportation costs often arise from agglomeration economies (Cidell 2010, Hesse, Rodrigue 2004). Further, the prevalence of railroad and highway miles, interstate highway and airport presence, and related transportation costs are often predictors of regional economic competitiveness, both for the TW sector itself (Cidell 2010, Guerin et al. 2021, Holl, Mariotti 2018) and for other sectors dependent upon the logistics (Belleflamme et al. 2000, Dudensing 2008). The range of variables included in recent logistics firm location models, such as gross domestic product (GDP) (Guerin et al. 2021), population density and accessibility to residential population and manufacturing (Sakai et al. 2020) and urban structure (Holl, Mariotti 2018), suggests that researchers are beginning to think beyond infrastructure density to broader factors that affect regional TW competitiveness. In the U.S. context, rural areas can be more attractive to certain types of firms because wages, property taxes, and land costs are all lower than in most metro areas (Parajuli, Haynes 2017, Wilkerson 2001).

Economic clustering – geographic concentrations of industries related through shared knowledge, skills, inputs, demand and/or other linkages – is a popular way of building economic competitiveness (Delgado et al. 2016). These geographic concentrations build upon regional competitive advantages (Porter 1990) and agglomeration factors (Marshall 1920). In Porter's (1990) framework, competitive advantage is influenced by four elements: firm strategy, structure, and rivalry; demand conditions; factor conditions; and related and supporting industries. Firm strategy, structure, and rivalry describes the size, number, ownership, and goals of firms. Demand conditions reflect needs of domestic (local) buyers, which can include the number of potential buyers, their purchasing needs, and their ability to pay. Factor conditions include the availability of factors of production, including land, labor, capital, and infrastructure. Related and supporting industries are those who require the goods or services of a given target industry, provide inputs, or use similar labor or technology. Each element is part of a system supporting the development of clusters that foster economic competitiveness within industries, and through those industries, of national and regional economies.

Clusters have been incorporated into a handful of TW location studies. Sakai et al. (2020) and Durmuş, Turk (2014) find the presence of industrial clusters has a positive effect on logistic establishment location decisions. Van den Heuvel et al. (2013) observes within-sector clustering as Belgian TW establishments tend to develop in proximity to existing logistics establishments, and they call for additional research into drivers and benefits of clustering within the TW sector.

3 Conceptual Framework

The focus on cost-distance minimization models in the transportation literature, gravity models in transportation and trade/regional economics literatures, and economic clusters within the regional economic literature seem to converge where the rubber meets the road. That is, for an entrepreneur or economic developer looking for opportunities to start or grow a transportation-based business, establishment and employment agglomeration (clustering) signal locations with existing advantages. Locations with similar demographic, transportation, and economic characteristics should be similarly suited to transportation business success. Durmuş, Turk (2014), Sakai et al. (2020), and Van den Heuvel et al. (2013) all incorporate clustering into their location choice models but fall short of placing their work within the broader clustering framework. They incorporate agglomerations of TW or other industry establishments with other elements of traditional cost-minimization location modeling. However, economic competitiveness signaled by clusters is a function of cost factors, including infrastructure, population density, and labor availability and quality. In fact, these factors are represented within the four elements of competitiveness in the Porter (1990) model. Thus, we embrace the broader framework of economic competitiveness and treat factor conditions (infrastructure, labor), demand conditions, industry structure, and related industries as part of the competitive locational choice decision. The competitiveness model also facilitates the exploration of how TW and other industry clusters interact. While interesting clusters themselves, the TW sectors in this study serve numerous other clusters and in fact may be considered parts of those clusters. In addition, subsectors within the broader TW sector often reinforce and strengthen each other. Figure A.1 represents our adaptation of the Porter framework, emphasizing the interactive relationship between firm structure and related industries, as well as government's direct influence on factor condition through infrastructure investment.

4 Transportation and Warehousing Data

We use restricted-access 2014 data from the Longitudinal Business Database (LBD) and Integrated Longitudinal Business Database (ILBD) for the continental U.S. along with 2014 secondary data sources to econometrically estimate the location decisions of five TW industries. In Europe, Eurostat offers a program similar to the U.S. Federal Statistical Research Data Center program (Eurostat 2020). As noted above, researchers often measure industry size in a location with public establishment counts, which is a non-negative integer count and includes numerous zeros that increase with more geographically refined units of observation. This measure is limited both because of disclosure issues and because an establishment count is a poor measure of regional industry size. Employment may serve as a better measure of industry size, while still using count data methods. We thus examine both non-employer establishments (i.e., establishments without any paid employees), employer establishments, and total employment as measures of industry size.

The Census Bureau's County Business Patterns (CBP) is among the most utilized county-level public data sources for establishment and employment counts for industries within the North American Industry Classification System (NAICS), including TW industries. Other federal data programs publish information about the economic activity in TW, including the Quarterly Census of Employment and Wages (QCEW), the Quarterly Workforce Indicators (QWI), and Non-employer Statistics (NS). Although public versions of these data are available, the exact counts of a particular NAICS code are also often suppressed. These limitations are particularly prevalent in some TW industries due to the small number of employers within some counties, especially rural counties. Given the extensiveness of the disclosure limitations, exact counts yield more precise estimates; specifically, Carpenter et al. (2021) show that the measurement error in public U.S. regional economic data implies substantially biased estimates. A Federal Statistical Research Data

Center (FSRDC) is thus a natural place to make improvements to existing research using the LBD and ILBD, which we aggregate to the county-level to develop unsuppressed non-employer establishment counts, employer establishment counts, and total county employment (as in Carpenter et al. 2021, Van Sandt et al. 2021), all within specific TW industries. The zero-inflation methods described below capitalize on this refined data by being able to accurately discern the zero-generating regime for unsuppressed counties with no TW industries.

The LBD is an annual series produced by the U.S. Census Bureau based on establishment records from the Business Register (Jarmin, Miranda 2002). The Business Register (BR) acts as the source of information for both the public CBP as well as the restricted LBD; however, the LBD undergoes more edits for longitudinal consistency and does not contain suppressed or noise infused values. Consequently, the LBD is a fundamental dataset for studying the determinants of firm strategy, structure, and rivalry, including entry, growth, and exit at the establishment, firm, industry, and economy-wide level (Carpenter, Loveridge 2018, 2019, 2020, Davis et al. 2006, Foster et al. 2006, Haltiwanger et al. 2013). Comprehensive longitudinal establishment-level data are similarly available in other countries, with economists often using them to examine firm establishment location decisions, though TW locational analysis remains understudied (Arauzo-Carod, Viladecans-Marsal 2009, Chen, Moore 2010, Devereux et al. 2007, Figueiredo et al. 2002, Holl 2004). Location quotients (LQs) of groups of non–TW industries explore the effects of related industries and the potential for inter-industry clustering.

Despite the LBD and the ILBD being at the establishment level, the Census Bureau requires us to conduct our analysis at an aggregated county level due to the sensitivity of the data. However, this unit of analysis is useful for the inclusion of important secondary data sources that capture the potential influence of factor conditions, including local internet access, travel infrastructure, water coverage, urban influence, and human capital, as well as demand conditions including other demographic and locational variables.

Further, county-level aggregation facilitates the implementation of the previously discussed count data methods. Table 1 provides descriptive statistics for the publicly available county-level data[2]. Variables are grouped by Porter Element with demographic variables representing either demand or factor conditions; infrastructure and institution variables reflect factor conditions. A handful of variables may cross elements, but each variable is listed only once. The TW establishment and employment data, of course, reflect firm strategy. The detailed nature of the data allows us to model the location and employment attributes for a range of TW industries, which fall within Transportation and Warehousing (NAICS 48-49) and Professional, Scientific, and Technical Services (NAICS 54). Specifically, we examine General Warehousing & Storage (NAICS 493110); Process, Physical Distribution, & Logistics Consulting (NAICS 541614); Pipeline Transportation (NAICS 486); Support Activities for Road Transportation (NAICS 488490); General Freight Trucking (NAICS 4841). Other sectors and industries are feasible and interesting, but we limit the presentation to example sectors here to cover a breadth of TW sectors, for which locational determinants have been shown to vary in international contexts (Holl, Mariotti 2018, Kang 2020, Rivera et al. 2014), while maintaining relative brevity[3]. Figure 1 shows a map of the distribution of Pipeline Transportation (NAICS 486) employment and data suppression issues. Maps of the other industries examined herein are available in the appendix (Figures A.2-A.5).

Motivation for the included determinants of transportation location draw on both past literature and virtual focus groups. Following a method established by Loveridge et al. (2013) the authors conducted virtual focus groups with U.S. economic development practitioners, small business support professionals, and entrepreneurs. Commonly found location determinants in published studies include local land and labor costs, taxes, infrastructure, market size, and agglomeration economies (Blair, Premus 1987). Proxy variables often include per capita income or wage rate, high school and bachelor degree

[2] Table A.1 in the appendix provides additional descriptive statistics.

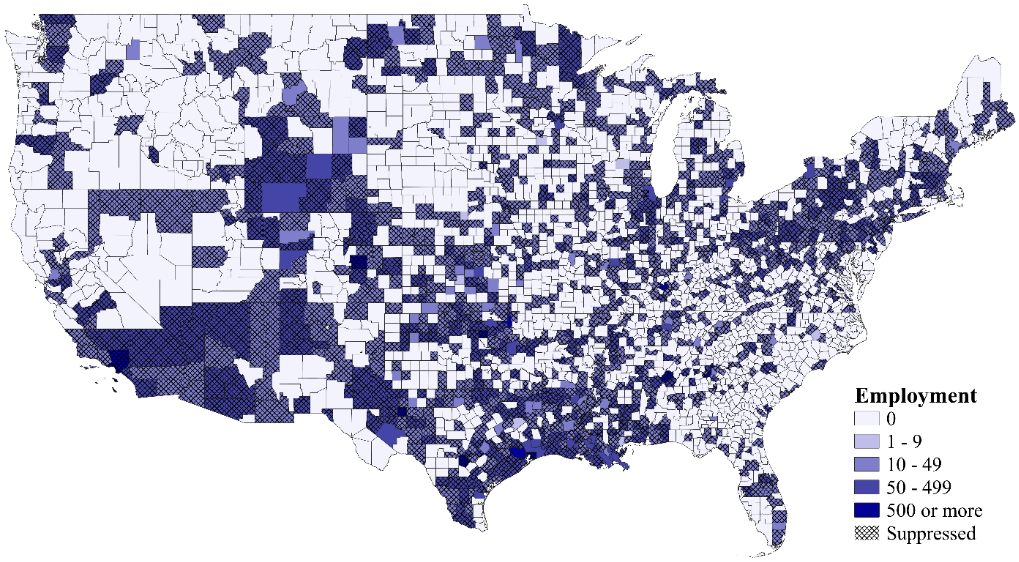
[3] The authors acknowledge that the choice of sectors could be criticized as ad hoc, rather than covering a breadth of industries. Nonetheless, these sectors are common to transportation and warehousing research, and as the results indicate, this breadth of sectors allows for interesting cross-sector comparisons of locational determinants and future researchers are encouraged to examine additional sectors.

Table 1: Descriptive statistics

Variable	Obs.	Mean	Std. Dev.	Source
Demographics – Demand Conditions				
Population	3,107	101,931.30	327,468.00	ACS
Population Density	3,107	0.27	1.80	ACS
Median Age	3,107	40.85	5.18	ACS
Per Capita Income (thousands $)	3,107	39.59	11.64	BEA
Percent of Residents in Poverty	3,107	16.84	6.55	ACS
Demographics – Factor Conditions				
Unemployment Rate (5yr avg.)	3,107	7.89	2.68	BLS
Social Capital Index	3,107	0.01	1.26	NERCRD
Opiates Prescribed / 100 People	2,944	85.72	49.37	CDC
Percent Work in Another County[1]	3,107	30.10	17.69	ACS
Percent Black	3,107	0.09	0.15	ACS
Percent Hispanic	3,107	0.09	0.14	ACS
Percent – Bachelors or Above	3,107	13.24	5.48	ACS
Infrastructure and Institutions – Factor Conditions				
Median Home Value ($1,000's)	3,107	135.77	79.20	ACS
Avg. Combined Sales Tax Rate[2]	3,107	7.01	1.68	Tax Foundation
Avg. Effective Property Tax Rate	3,107	1.06	0.51	SmartAsset.com
Internet Service Providers (ISPs)	3,106	5.19	1.12	FCC
Metro - Urban Influence Code	3,107	0.37	0.48	ERS, USDA
Micropolitan Metro Adjacent - UIC	3,107	0.33	0.47	ERS, USDA
Micropolitan Non-metro Adjacent - UIC	3,107	0.30	0.46	ERS, USDA
Interstate Density[3]	3,107	1.79	3.07	Census Shapefiles
Highway Density[3]	3,107	0.37	0.24	Census Shapefiles
Percent Covered by Water	3,107	4.50	11.15	ERS, USDA
Community Colleges	3,107	0.33	0.84	NCES
Universities or Colleges	3,107	0.72	2.40	NCES
Military Bases	3,107	0.04	0.22	US Census Bureau
Census Region Fixed Effects	3,107	N/A	N/A	US Census Bureau
TW Industry Establishments and Employment – Firm Strategy, Structure, and Rivalry				
General Warehousing & Storage (493110) Estab.	3,106	3.32	13.22	WholeData
General Warehousing & Storage (493110) Emp	3,106	209.12	881.42	WholeData
Warehousing and Storage (4631) Non-emp	1,681	5.21	24.31	NS
Management Consulting Services (54161) Estab.	3,106	41.78	198.01	WholeData
Management Consulting Services (54161) Emp.	3,106	293.25	1,705.54	WholeData
Management, Scientific, and Technical Consulting Services (5416) Non-emp.	3,072	228.31	924.11	NS
Pipeline Transportation (486) Estab.	3,106	1.33	3.93	WholeData
Pipeline Transportation (486) Emp.	3,106	16.89	187.20	WholeData
Pipeline Transportation (486) Non-emp.	511	1.23	5.91	NS
Support for Road Transportation (488490) Estab.	3,106	0.81	3.31	WholeData
Support for Road Transportation (488490) Emp.	3,106	10.89	59.73	WholeData
Support for Transportation (488) Non-emp.	2,924	44.28	229.92	NS
General Freight Trucking (4841) Estab.	3,106	22.18	85.38	WholeData
General Freight Trucking (4841) Emp.	3,106	293.68	1,014.61	WholeData
General Freight Trucking (4841) Non-emp.	3,100	160.87	781.12	NS

Notes: Due to disclosure prevention limitations, these descriptive statistics are based on public data sources, while the main regression results are based on the limit-access Longitudinal Business Database and Integrated Longitudinal Business Database. All data are based on 2014; 2014 is chosen for practical reasons related to the availability of many of the secondary data sources. The internal and unsuppressed data used in regressions differ slightly in addition to the inclusion of unsuppressed cells. The more refined 541614 NAICS was not publicly available. Industry size measure abbreviations: non-employer establishments (Non-Emp.); employer establishments (Est.); total employment (Emp.). [1] Percent of all residents age 16 and older. [2] State-level variable. [3] Miles of road per hundred square miles.

Abbreviations: ACS: American Community Survey, BLS: Bureau of Labor Statistics, BEA: Bureau of Economic Analysis, CDC: Center for Disease Control and Prevention, FCC: Federal Communications Commission, ERS: Economic Research Service, NS: Non-employer Statistics, NCES: National Center for Education Statistics, NERCRD: Northeast Regional Center for Rural Development, USDA: US Department of Agriculture



Notes: Map shows employment ranges in every county in the continental U.S. "Suppressed" indicates that the exact value of the data is suppressed to prevent improper disclosure of identifiable information and placed into bins (e.g., "20-99 employees"). For the sake of these maps, we replace these ranges with their midpoint.

Figure 1: Employment Distribution of Pipeline Transportation (NAICS 486)

shares, unemployment rate, highway and interstate coverage, property taxes, rurality spectrum codes, population density, climate, racial composition, and location quotients (Coughlin, Segev 2000, Guimarães et al. 2000, 2004). The virtual focus groups supported many of these previously found factors but added insights into the potential importance in the rural U.S. of broadband access and issues related to workforce turnover and opioid misuse, particularly in rural areas and particularly related to TW industries (Joudrey et al. 2019, Rigg et al. 2018). This article uses the opioid prescription rate as a proxy for opioid misuse; though imperfect, given regional variation in prescription rates and negative outcomes associated with opioid misuse (Quast 2018), we continue to include this variable given the relevance of the opioid epidemic in the U.S. noted in our qualitative focus groups and in past research (Rigg et al. 2018).

Finally, researchers define the transportation and logistics clusters based on research interests and purposes, and definitions may affect study outcomes. For example, the Kumar et al. (2017) transportation and logistics cluster includes both freight and passenger transportation, suggesting that passenger transportation might affect the prevalence of the cluster in metropolitan areas. For clarity in the presentation of results, this article relies on Location Quotients (LQs) of groups of non–TW industries to examining the potential for inter-industry clustering in this article[4].

5 Methods

Given the nature of geography-based establishment and employment data, and the ability to allow for two zero-generating processes through an added logit link function, count data models are useful for empirical estimation. Resultantly, count data estimators have become common in locational and threshold models (Carpenter et al. 2021), though researchers often fail to appropriately apply and compare the various models, as we will demonstrate. For non–TW establishments, there are numerous applications of Poisson (Arauzo-Carod, Viladecans-Marsal 2009, Papke 1991), Negative Binomial (NB) (Conroy et al. 2016, Holl 2004, Smith, Florida 1994), Hurdle Poisson (HP) (Chakraborty 2012,

[4]This article uses the entire U.S. as the reference for the location quotient because, as Figure 1 highlights, that TW industries tend to cluster in areas above alternatives, such as at the states. Additionally, many states are quite small and have few TW industries, confusing the interpretation of the LQs. As noted later, regressions include regional fixed effects.

Henderson et al. 2000), Zero-Inflated Poisson (ZIP) (Chakraborty 2012, List 2001, Reum, Harris 2006). We also consider the Zero-Inflated Negative Binomial (ZINB) here within TW when using a large geographic scope without data suppression.

The Poisson model is useful as a starting point for comparison to other estimation procedures. However, the Poisson model's assumption that dependent variable's conditional variance is equal to the conditional mean is often violated in practice due to multiple sources of overdispersion, in which case the NB is more efficient. Overdispersion is likely to exist in TW industries due to large infrastructure endowments in some geographical units motivating the need for many TW establishments and levels of employment[5]. Zero-inflated and hurdle versions of the Poisson model are often used to account for excess zeros in the data when it is believed that the zeros arise from two separate regimes. In the current context, the two zero-generating regimes in zero-inflated models may be interpreted as structural and non-structural zeros, i.e., counties that have zero establishments due to a lack of a requisite resource like water and counties that have zero establishments due to chance or extant economic conditions. On the other hand, HP only assumes one type of zero and truncates the Poisson distribution after the logit zero-generating, which is less flexible[6].

Finally, researchers can fail to account for overdispersion that remains after the zero-inflation. This is particularly important when examining TW sectors because remaining overdispersion (from unobserved heterogeneity or excessive concentration of firms) increases in economies of scale and agglomeration, which are common in TW sectors. If authors are interested in modeling both an inflation stage and accounting for these multiple sources of overdispersion, ZINB is underutilized; indeed, we find it to be preferred for many TW industries.

We compare results and diagnostics from the Poisson, NB, and their zero-inflated complements to provide results across five TW subsectors. This flexible approach takes advantage of the count data, while using the potential for overdispersion caused by (1) numerous zeros, which depend on industrial sector specificity, spatial monopsony, and economies of scale, (2) long-tailed distributions, resulting from spatial concentration, which varies depending on spatial resource dependence and economies of agglomeration, and (3) unobserved heterogeneity (Carpenter et al. 2021).

To choose among the various count data models, many researchers incorrectly use Vuong's Statistic to compare these various count data estimators (Wilson 2015). To provide a consistent comparison, we use graphical distribution and information criteria comparisons as suggested in Greene (1994). All models make use of White's robust standard errors, and variance inflation factors did not indicate any potential concern over loss of efficiency from multicollinearity.

6 Results

While TW includes dozens of industries spanning many sectors of the economy and NAICS sector codes, we focus on investigating a subset of these industries. To summarize, the purpose of this section is not to present an exhaustive analysis of all TW industries' locations, but instead to provide examples of different data generation processes and the advantages of different industry size measures. Through these example models, we illustrate how more refined data and unique data generation processes may be utilized to explore how local industry, socioeconomic, and place-based factors influence the size of

[5]The NB approach allows unobserved heterogeneity between subjects and implies overdispersion, but where the amount of overdispersion increases with $E(y_i|x_i)$ (Wooldridge 2010). Although this may hold when modeling TW employment (e.g., due to large employers), this relationship is unlikely to hold when measuring establishments due to excess zeros driving much of the overdispersion while *decreasing* $E(y_i|x_i)$. Additionally, we note that, though less efficient, Poisson would still produce consistent estimates with fewer assumptions than NB.

[6]The log-likelihood of the Zero-Inflated Poisson model may be written, $\ln L = \sum_{i\in S} \ln\{F(\gamma' Z_i) + [1-F(\gamma' Z_i)]\exp(-\lambda_i)\} + \sum_{i\notin S}\{\ln[1-F(\gamma' Z_i)] - \lambda_i + y_i\beta' X_i - \ln(y_i!)\}$, where S is the set of observations taking on a zero value ($y_i = 0$), F is the logit link function that determines the odds of a zero belonging to regime one or two, Z is a vector of covariates that describe the participation decision, X is a vector of covariates that describe the amount decision, and γ and β are the participation and amount parameters of interest to be estimated, respectively.

Table 2: T&W sector distributions across industry size measures

Industry (NAICS code)	Non-Employer Establishments	Employer Establishments	Employment
General Warehousing & Storage (493110)		ZINB	ZINB
Process, Physical Distribution, & Logistics Consulting (541614)		ZINB	ZINB
Pipeline Transportation (486)	ZINB	ZINB	ZINB
Support Activities for Road Transportation (488490)	ZIP	ZINB	ZINB
General Freight Trucking (4841)	NB	NB	NB

Notes: This article does not examine non-employer establishments for NAICS 493110 and 541614 (and there are empty cells in this table) due to no or very few non-employer establishments.

local TW industries. We also note that these results are associative, so we avoid causal interpretations of the results.

To facilitate comparisons across different measures of industry size (e.g., non-employer establishments, employer establishments, and employment), the model covariates are the same between each industry. Each model includes regional fixed effects, population demographics, sectoral location quotients (base = U.S.), employment in three transportation industries, rurality, and other place-based variables of interest through policy or industry perspectives.

We identified the data generation process underlying each industry by first testing for overdispersion in the data to select between the Poisson and negative binomial distributions[7]. After checking for zero-inflation, the significance of the alpha parameter was checked again to determine if overdispersion was still present in the data after accounting for a dual zero-generating process leading to excess zeros. Table 2 summarizes the distributional findings of a selection of TW industries for each industry size measure. The resulting distribution is the product of several economic and administrative features. The NB distribution indicates these establishments can operate in most locations and likely serve smaller, more rural communities where fewer agglomeration economies arise or fewer specialized resources are needed. Industry aggregation (i.e., industries with fewer NAICS digits) also influences the observed data generation process. Aggregating industries leads to fewer counties with zero establishments in the data, meaning that research questions specific to an industry subset of an aggregated NAICS code suffer not only from estimated coefficients values being confounded by noise from other industries, but also from the inability to identify the industry's true data generation process. In this study, we consider four- to six-digit NAICS. Specific codes are provided in Table 2. The rarity of Poisson data generation processes across TW establishments and employment informs our understanding of TW industries and implies that establishments in these industries frequently cluster together, leading to positive skewness and overdispersion.

The marginal effects for five TW models are presented in Tables 3.1–3.3 to demonstrate the effect on model coefficients from (1) different levels of industry aggregation and (2) different measures of industry size. We number these tables 3.1–3.3 to emphasize that the marginal effect estimates are derived from the same regression for each respective industry. In all zero-inflated models, the logistic link function in the ZI equation predicts the odds of a county being in the certainly zero category (i.e., a structural zero opposed to a sampling zero). Thus, the negative coefficients in the inflation stage of Table 3.1 indicate that an increase in the variable is associated with a decrease in the likelihood of the county being a structural zero.

Population and population density are most consistently associated with the probability of non-structural zeros in the inflation stage. Increasing home values and incomes are associated with reduced probability of structural zeros, indicating that economic forces are the dominant barriers to TW locating in areas with higher home values. However,

[7] In addition to the preferred models that use the LBD and ILBD shown in-text, the appendix provides results when using Poisson and the preferred models with publicly available data. This comparison is useful both to show the Poisson as a benchmark model and show the importance of using unsuppressed administrative data, especially in less-aggregated industries.

Table 3.1: Inflation Stage Marginal Effects for Locational Determinants of T&W Sectors

Industry (NAICS)	General Warehousing & Storage (493110)		Process, Physical Distribution, & Logistics Consulting (541614)		Pipeline Transportation (486)			Support Activities for Road Transportation (488490)		
Size Measure	Est.	Emp.	Est.	Emp.	Non-Emp.	Est.	Emp.	Non-Emp.	Est.	Emp.
Interstate Density	-0.005	-0.011***	-0.043***	-0.005	-0.010	-0.009	-0.004	-0.038	-0.039*	-0.002
Highway Density	-0.732	-12.120**	35.32**	1.667	25.04*	26.15***	23.77***	4.615	-5.244	4.249
Water coverage	-0.002	0.000	-0.025*	0.000	-0.027	0.004***	0.004***	0.004**	-0.002	0.000
Sales tax	0.007	0.004	-0.013	-0.006	0.005	-0.02**	-0.029***	-0.011	0.017	0.003
Property tax rate	0.005	-0.013	-0.001	0.011	-0.068	-0.009	-0.003	0.069	0.081	0.005
ISP count	0.029*	-0.012	-0.009	-0.003	-0.033	-0.002	0.015	0.020	-0.014	-0.013
Micro Metro-Adj.	-0.006	0.007	-0.046	0.006	-0.113**	0.043	0.05*	-0.422	-0.168**	0.039
Metropolitan	0.000	-0.006	-0.006	-0.005	-0.097	0.024	0.037	-0.172*	-74.140	0.015
ln(Population)	-0.074**	-0.212***	-0.052	-0.139***	-0.017	-0.129***	-0.159***	-0.035	-0.022	-0.140***
Population density	-4.485***	-0.068	-1.312**	-0.572	-0.050	-0.173*	-0.119**	-0.042	1.030	-0.059*
Per capita income	-0.006**	-0.003***	-0.001	-0.002*	0.002***	-0.007***	-0.009***	-0.009	-0.007**	-0.003**
Home Value	-0.001	0.000	-0.002**	-0.001***	-0.003***	0.001***	0.002***	0.000	0.000	0.000

Significance levels: *** $p<1\%$, ** $p<5\%$, * $p<10\%$

Notes: Table presents the inflation stage for the regressions specified in Table 2 (either Zero-Inflated Poisson or Zero-Inflated Negative Binomial). Industry size measure abbreviations: non-employer establishments (Non-Emp.); employer establishments (Est.); total employment (Emp.). There is no inflation stage for the General Freight Trucking (4841) industry because it did not need to be zero-inflated (it is modeled as NB). Regressors in Table 3.1 are included in same respective regressions as Tables 3.2 and 3.3.

a county with a higher per-capita income has a greater chance of observing structural barriers preventing the location of pipeline non-employer establishments in the county. Similarly, higher home values are associated with more structural zeros for pipeline employer establishments and employment. Highway density is also associated with greater chance of structural zeros in the Process, Physical Distribution & Logistics Consulting and pipeline sectors, perhaps due to the opportunity costs imposed by the existing infrastructure already present.

The marginal effects of the locational factors in the amount stage (of the same ZIP and ZINB model) are presented in Table 3.2. Population, income, and housing factors also influence the amount stage for several transportation sectors. Logged population is universally associated with more establishments and income; it is logical that more businesses and employees are needed to serve a larger population. Population density also has a strong positive relationship with general warehousing establishments and employment – suggesting a particular need to serve denser populations by storing goods – and consulting establishments, which likely respond to urbanization economies and flourish where a large number of client firms exist. However, while denser populations are associated with both extensive and intensive expansion for the general warehousing and storage industry, they are only associated with extensive growth for consulting establishments. Income has a positive relationship with consulting and pipeline industry establishments as well as pipeline employment, while higher home values have a negative relationship with transportation sector establishment and job counts. As Calafati et al. (2021) note, housing, transportation, and utility costs have confounding effects with income, but it also makes sense that warehouses and trucking yards are constructed in areas with lower real estate costs. Racial and ethnic diversity also generally related to more TW employer and non-employer establishments and sometimes more employees. The share of the Hispanic population, in particular, was generally associated with more non-employer and employer establishments.

The density of transportation infrastructure including interstates and highways, as well as sales tax rates and commuting, are associated with general warehousing and storage and general freight trucking. While more interstate and highway miles support increased establishments and employees, the quadratic relationship of these determinants indicate both establishments and employment experience diminishing returns from higher densities. Nonetheless, the magnitude of these coefficients is quite large, emphasizing the importance of infrastructure. Sales tax usually has a negative relationship but bolsters the number of pipeline employer establishments. Interestingly, the share of out-commuters has a negative relationship with employer establishments but a positive relationship with employment levels within the warehousing industry, while indicating near-opposite relationships for non-employer establishments and employment within the general freight trucking industry. We posit that interactions with other industries affect the community relationships, which is addressed with locational determinants. Somewhat similar factors are associated with the number of pipeline employer establishments, which may reflect overlapping agricultural and oil- and gas-mining regions. Most variables are associated with general freight trucking, although interstate and highway density are only weakly significant factors.

Share of the population with a bachelor's degree is positively associated with general warehousing and storage and consulting establishments but negatively associated with all three measures of general freight trucking, which demonstrates how sociodemographic factors affect TW subsectors differently. Consulting is usually considered to include higher education levels, and warehousing is increasingly computerized and requires special skills. In contrast, general freight trucking continues to have relatively low education requirements and low barriers to entry. The effects of unemployment were felt most acutely in warehousing and general freight trucking while other sectors, such as pipeline transportation, were more insulated. Processes generating establishments and employment also differ, as evidenced by conflicting coefficient signs on the share of workers out-commuting for general warehousing and storage, as well as the opioid prescription rate in general freight trucking. This negative association is surprising, given the literature on the opioid epidemic and discussions in our focus groups, which

Table 3.2: Amount Stage Marginal Effects for Locational Determinants of T&W Sectors

Industry (NAICS)	General Warehousing & Storage (493110)		Process, Physical Distribution, & Logistics Consulting (541614)		Pipeline Transportation (486)			Support Activities for Road Transportation (488490)			General Freight Trucking (4841)		
Size Measure	Est.	Emp.	Est.	Emp.	Non-Emp.	Est.	Emp.	Non-Emp.	Est.	Emp.	Non-Emp.	Est.	Emp.
Interstate Density	0.146***	15.85***	-0.021	-0.220	-0.014	0.044	1.699**	0.010	0.019*	-227.3	0.485	0.237*	10.49***
Interstate Density2	-0.005**	-0.662***	0.001	-0.019	0.000	-0.003	-0.119**	0.000	0.000	13.140	0.001	-0.003	-0.296***
Highway Density	228.5***	28680***	70.54*	765.1	10.83	-2.003	-1955*	-7.309	11.900	244500	3093*	298.2	8426
Highway Density2	-9867***	-1267000***	-3298**	-37070	-218.5	-2852	14900	-34.45	-293.5	-1371000	-1473*	-23680	-578900
Water coverage	-0.022***	-4.113***	-0.002	0.115	0.007	-0.003	-0.005	0.001	-0.004**	-41.37	-1.062***	-0.115***	-2.405***
Sales tax	-0.126**	-16.93**	-0.086***	-2.069*	-0.022	0.092***	0.563	-0.021	-0.015	-73.600	-2.261*	-0.397**	-3.209
Property tax rate	-0.042	-20.25	0.213*	-1.831	-0.110*	-0.025	-4.810*	0.082	0.001	-329.00	-15.67***	-1.545**	-49.300**
Military bases	-0.346	-26.76	-0.003	6.641	-0.029	-0.168	0.216	-0.079*	-0.101	820.00	-18.96**	-2.243**	-21.31
Community colleges	0.057	11.76	-0.008	-0.550	0.004	-0.036	2.040	-0.032**	-0.010	-161.2	0.139	-0.186	-0.977
Universities	-0.016	-4.585	-0.004	-0.231	-0.009	-0.024	-0.123	0.018***	0.013*	116.5	-3.969**	-0.150	2.815
ISP count	-0.022	-5.013	-0.124**	-0.951	-0.011	-0.139***	-0.284	-0.036	-0.007	-182.0	-0.230	0.569**	7.898
Social capital	-0.123	-29.32*	-0.092	4.871	0.015	0.133***	2.878*	-0.082	0.011	-517.1	0.695	0.932	21.65*
Micro Metro-Adj.	-0.465	-66.47	0.224	9.495	-0.033	-0.112	-0.469	-0.383*	-0.396**	-124.5	-10.25**	-3.383***	-41.47**
Metropolitan	-0.236	-98.35**	0.309	15.490*	-0.045	0.02	0.348	-0.501***	17.320*	126.9	-23.89***	-5.000***	-19.16
Out-commute %	-0.019***	1.676**	-0.003	-0.111	0.001	-0.009***	0.016	0.000	-0.002	11.16	0.365**	-0.029	-2.561***
Opioid RX rate	0.001	-0.410	0.003	-0.016	-0.001	0.001	0.031	-0.001	-0.001	2.197	-0.189***	-0.013*	0.438**
Poverty rate	-0.087***	-5.063*	-0.035	-1.342**	0.009	-0.016	0.143	-0.010	-0.008	-75.81	-0.728	-0.370***	-6.646***
Median age	-0.039	-3.415	0.044**	-0.430	0.005	-0.048***	-1.337***	0.000	-0.004	139.0	-0.394	-0.260***	-10.19***
Unemployment rate	0.014	-23.51***	-0.028	-0.927	-0.022	-0.048*	1.191	-0.017	-0.030	-60.38	-0.681	-0.733***	-16.18***
Bachelors degree %	0.066**	2.341	0.129***	0.698	-0.004	-0.038**	0.172	-0.005	-0.011	72.29	-5.455***	-0.429***	-6.615**
ln(Population)	3.593***	225.2***	2.521***	33.20***	0.503***	0.917***	15.32***	0.580***	0.767***	2542	164.0***	23.07**	362.3***
Population density	2.395***	12.22**	0.560***	7.782	0.019	0.313*	1.046	0.011	-0.263*	633.6	2.799	0.152	-3.847
Per capita income	-0.01	-1.163	0.009*	0.200	0.007***	0.029***	0.430***	-0.003	-0.001	51.02	1.120***	0.050	0.898
Hispanic %	0.029***	-0.381	0.034***	0.661***	0.004**	0.002	0.041	0.004*	0.006*	38.31	1.119***	0.093*	0.337
Black %	0.039***	2.181**	0.008	0.286	-0.003	0.008	-0.031	0.000	0.006**	61.07	0.158	0.008	1.782*
Home Value	-0.002	-0.427*	-0.001	-0.109***	-0.001	-0.008***	-0.091***	0.000	0.000	-5.060	-0.272***	-0.025***	-0.218

Significance levels: *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$

Notes: Table presents the inflation stage for the regressions specified in Table 2 (either Zero-Inflated Poisson or Zero-Inflated Negative Binomial). Industry size measure abbreviations: non-employer establishments (Non-Emp.); employer establishments (Est.); total employment (Emp.). There is no inflation stage for the General Freight Trucking (4841) industry because it did not need to be zero-inflated (it is modeled as NB). Regressors in Table 3.2 are included in same respective regressions as Tables 3.1 and 3.3.

emphasized the likely association of opioid misuse and rural TW. However, again, these relationships may reflect the need for freight trucking services by industries (e.g., meat processing, resource extraction) prevalent in areas with high levels of opioid use, or it may be an artifact of opioid prescription rates being an imperfect measure of opioid misuse and the regional variation thereof (Quast 2018). These differences between subsectors and between establishment and employment generation may pose challenges in building or strengthening local transportation clusters. Developers may need to tailor development strategies to specific TW subsectors.

To the authors' knowledge, this is among the first attempts to model non-employer establishments and employment using demand threshold modeling. In considering the associations of other industries on transportation, the TW location quotient (LQ) was generally associated with more establishments across transportation subsectors[8]. General warehousing and storage; process, physical distribution, and logistics consulting; and general freight trucking had negative associations with most LQs in other clusters. However, these clusters are often service-based and require less transportation services. More transportation-dependent sectors, such as agriculture and manufacturing, may have positive relationships with TW, but these sectors are very broad and include a large number of clusters, which are outside the scope of this analysis. Positive associations between pipeline transportation establishments and employment and the LQs for mining and gas and construction likely reflect pipeline transportation's role within oil and gas clusters, which also require various types of construction activities.

The different data generation processes found between non-employer and employer establishments and the differences in significance and magnitude across model covariates within each industry demonstrate that these establishment types behave very differently. This suggests that a given policy designed to support TW employer establishments may have insignificant or opposite effects on non-employers, some of whom may be entrepreneurs responsible for tomorrow's innovations in the TW sector.

7 Conclusions

The chief contributions of this manuscript lie in its use of a fully disclosed dataset and its elaboration of a methodology for modelling industry spatial distributions characterized by overdispersion (in this instance few US counties with many firms in an industry and many counties with none). Specifically, the data set and model allow researchers to carefully consider the presence of structural and non-structural zeros and understand the reasons a county has no establishments in a given industry. Different factors influence industry presence (inflation) and employment levels (amount) for TW sectors. From there, planners and policymakers can use the locational determinants to strategize ways to strengthen transportation and related sectors in regional economies. Population, population density, and sometimes highway density tended to influence sector formation in a county. Population, income, and housing factors consistently influenced amount-stage results, but the effects of other sociodemographic and institutional factors varied by TW sector. Results presented in Table 3.2 suggest that aggregating industries decreases or obscures locational factors' importance for individual sectors.

Similarly, aggregation obscures TW sectors' relationships to other clusters. Most TW sectors had a positive association with wholesale LQ. Warehousing and general freight trucking tended to be negatively associated with LQs in other industries included in the analysis. These results were consistent using both establishment and employment data.

Several studies consider the effects of transportation infrastructure on the economy in general and specific economic sectors. This study furthers that research by considering the impact of sociodemographic and institutional variables on transportation business establishments while also controlling for infrastructure. The study also shows that other industries influence transportation subsectors in different ways. For example, a larger professional and technological LQ is associated with few warehousing and general freight establishments and lower employment in those industries. Future research could use

[8] TW LQs are calculated on all NAICS 48, 49, and 541614, except for the respective NAICS under consideration.

Table 3.3: Amount Stage Marginal Effects for Locational Determinants of T&W Sectors

Industry (NAICS)	General Warehousing & Storage (493110)		Process, Physical Distribution, & Logistics Consulting (541614)		Pipeline Transportation (486)			Support Activities for Road Transportation (488490)			General Freight Trucking (4841)		
Size Measure	Est.	Emp.	Est.	Emp.	Non-Emp.	Est.	Emp.	Non-Emp.	Est.	Emp.	Non-Emp.	Est.	Emp.
Location Quotients													
TW	0.290***	8.728	0.066***	0.507	0.026***	0.003	0.258	0.024***	0.043***	286.9	3.618***	0.767***	4.348*
Finance	-0.066**	-8.276	-0.006	-0.801	0.006	-0.032	-0.291	0.003	-0.006	13.43	-0.023	-0.178	-0.107
Real Estate	-0.258***	-29.36***	-0.011	-1.515	0.009	0.002	-0.424	0.039***	0.015	-268.2	-1.234	-0.109	-17.96***
Professional and tech	-0.274***	-23.58***	0.056	3.842**	-0.015	0.021	-0.665	-0.02*	-0.026*	-24.51	-4.500***	-1.307***	-24.84***
Education	-0.130***	-9.738**	-0.029	-3.868***	0.000	-0.019	-1.226***	-0.019*	-0.033***	12.23	-2.117***	-0.714***	-19.1***
Health services	-0.234***	-9.108	-0.186***	-3.732***	0.011	0.003	-0.197	-0.039**	-0.032*	-44.80	-4.844***	-0.968***	-39.18***
Art and recreation	-0.006	-0.347	0.020	0.458	0.001	-0.004	0.456	-0.018**	0.001	-41.75	-1.727***	-0.282***	-9.524***
Accommodation & Food	0.013	-13.64*	-0.067*	0.08	-0.002	-0.023	-0.878	0.004	0.013	194.7	-4.427***	-0.875***	-15.58***
Mining and gas	-0.014***	-1.559***	0.000	0.162*	0.002***	0.002***	0.108***	0.001	0.001*	10.26	-0.281***	-0.011*	-0.265**
Construction	-0.197***	-27.73***	-0.094***	-2.725***	0.020***	0.067***	0.457	-0.005	-0.008	-68.78	1.026	-0.174*	-7.705***
Retail	-0.182**	-5.665	-0.163***	-6.563***	-0.017	-0.059	-1.512*	-0.012	0.016	-7.510	1.817	0.206	-19.15***
Regional FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
n	3,100	3,100	3,100	3,100	3,100	3,100	3,100	3,100	3,100	3,100	3,100	3,100	3,100

Significance levels: *** p <1%, ** p <5%, * p <10%

Notes: The exact observation count is suppressed by U.S. Census Bureau disclosure review process. We use all counties in the continental U.S. in 2014. Industry size measure abbreviations: non-employer establishments (Non-Emp.); employer establishments (Est.); total employment (Emp.). TW LQs are calculated on all of NAICS 48, 49, and 541614, except for the respective NAICS under consideration. No inflation stage for the General Freight Trucking (4841) industry. Regressors in Table 3.3 are included in same respective regressions as Tables 3.1 and 3.2. Regional fixed effects (FE) based on Census regions. Number of observations (counties) in the sample (n) rounded due to disclosure limitations.

spatial econometric techniques to measure effects of neighboring counties. The increasing availability of disclosed data through multiple U.S. agencies also presents opportunities to further explore differences in locational determinants of TW in the U.S. and European contexts. Finally, future research could expand the study of inter-cluster dependence, which could reduce hazards associated with cluster targeting in economic development (Barkley, Henry 2009) and improve overall competitiveness opportunities (Porter 2001). Such research could support efforts to capture inter-industry linkages and define industries within local clusters (Delgado et al. 2016).

References

Allen J, Browne M, Cherrett T (2012) Investigating relationships between road freight transport, facility location, logistics management and urban form. *Journal of Transport Geography* 24: 45–57. CrossRef

Arauzo-Carod JM, Viladecans-Marsal E (2009) Industrial location at the intra-metropolitan level: The role of agglomeration economies. *Regional Studies* 43: 545–558. CrossRef

Barkley DL, Henry MS (2009) Targeting industry clusters for reginal economic development: The REDRL approach. In: Goetz S, Deller S, Harris T (eds), *Targeting Regional Economic Development*. Routledge, London, 183–197

Bartik TJ, Biddle SCY, Hershbein BJ, Sotherland ND (2018) Wholedata: Unsuppressed county business patterns data: Version 1.0 [dataset]. Kalamazoo. Retrieved from https://upjohn.org

Behiri W, Belmokhtar-Berraf S, Chu C (2018) Urban freight transport using passenger rail network: Scientific issues and quantitative analysis. *Transportation Research Part E: Logistics and Transportation Review* 115: 227–245. CrossRef

Belleflamme P, Picard P, Thisse JF (2000) An economic theory of regional clusters. *Journal of Urban Economics* 48: 158–184. CrossRef

Blair JP, Premus R (1987) Major factors in industrial location: A review. *Economic Development Quarterly* 1: 72–85. CrossRef

Bowen JT (2008) Moving places: the geography of warehousing in the US. *Journal of Transport Geography* 16: 379–387. CrossRef

Calafati L, Froud J, Haslam C, Johal S, Williams K (2021) Diversity in leading and laggard regions: Living standards, residual income and regional policy. *Cambridge Journal of Regions, Economy and Society* 14: 117–139. CrossRef

Carpenter CW, Loveridge S (2018) Differences between latino-owned businesses and white-, black-, or asian-owned businesses: Evidence from census microdata. *Economic Development Quarterly* 32: 225–241. CrossRef

Carpenter CW, Loveridge S (2019) Factors associated with latino-owned business survival in the United States. *The Review of Regional Studies* 49: 73–97. CrossRef

Carpenter CW, Loveridge S (2020) Business, owner, and regional characteristics in latino-owned business growth: An empirical analysis using confidential census microdata. *International Regional Science Review* 43: 254–285. CrossRef

Carpenter CW, Van Sandt AT, Dudensing R, Loveridge S (2021) Profit pools and determinants of potential county-level manufacturing growth. *International Regional Science Review*. CrossRef

Carpenter CW, Van Sandt AT, Loveridge S (2021) Empirical methods in business location research. *Regional Studies, Regional Science* 8: 344–361

Chakraborty K (2012) Estimation of minimum market threshold for retail commercial sectors. *International Advances in Economic Research* 18: 271–286. CrossRef

Chen MX, Moore MO (2010) Location decision of heterogeneous multinational firms. *Journal of International Economics* 80: 188–199. CrossRef

Chen Y, Salike N, Luan F, He M (2016) Heterogeneous effects of inter- and intra-city transportation infrastructure on economic growth: Evidence from Chinese cities. *Cambridge Journal of Regions, Economy and Society* 9: 571–587. CrossRef

Cidell J (2010) Concentration and decentralization: The new geography of freight distribution in US metropolitan areas. *Journal of Transport Geography* 18: 363–371. CrossRef

Clausen U, Voll R (2013) A comparison of North American and European railway systems. *European Transport Research Review* 5: 129–133. CrossRef

Conroy T, Deller S, Tsvetkova A (2016) Regional business climate and interstate manufacturing relocation decisions. *Regional Science and Urban Economics* 60: 155–168. CrossRef

Coughlin CC, Segev E (2000) Location determinants of new foreign-owned manufacturing plants. *Journal of Regional Science* 40: 323–351. CrossRef

Crang M (2002) Between places: Producing hubs, flows, and networks. *Environment and Planning A: Economy and Space* 34: 569–574. CrossRef

Dablanc L, Ogilvie S, Goodchild A (2014) Logistics sprawl. *Transportation Research Record* 2410: 105–112

Davis SJ, Faberman RJ, Haltiwanger J (2006) The flow approach to labor markets: New data sources and micro-macro links. *Journal of Economic Perspectives* 20: 3–26. CrossRef

de Jong G, Gunn H, Walker W (2004) National and international freight transport models: An overview and ideas for future development. *Transport Reviews* 24: 103–124. CrossRef

Delgado M, Porter ME, Stern S (2016) Defining clusters of related industries. *Journal of Economic Geography* 16: 1–38. CrossRef

Devereux MP, Griffith R, Simpson H (2007) Firm location decisions, regional grants and agglomeration externalities. *Journal of Public Economics* 91: 413–435. CrossRef

Dudensing R (2008) Biomedical devices clusters in South Carolina and the United States. Clemson, SC

Durmuş A, Turk SS (2014) Factors influencing location selection of warehouses at the intra-urban level: Istanbul case. *European Planning Studies* 22: 268–292. CrossRef

Eurostat (2020) Access to microdata. Retrieved november 1, 2020, from https://ec.europa.eu/eurostat/web/microdata/overview

Figueiredo O, Guimarães P, Woodward D (2002) Home-field advantage: Location decisions of Portuguese entrepreneurs. *Journal of Urban Economics* 52: 341–361. CrossRef

Foster L, Haltiwanger J, Krizan CJ (2006) Market selection, reallocation, and restructuring in the U.S. retail trade sector in the 1990s. *Review of Economics and Statistics* 88: 748–758

Greene WH (1994) Accounting for excess zeros and sample selection in poisson and negative binomial regression models. NYU Working Paper No. EC-94-10. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?.abstract_id=1293115

Guerin L, Vieira JGV, de Oliveira RLM, de Oliveira LK, de Miranda Vieira HE, Dablanc L (2021) The geography of warehouses in the São Paulo metropolitan region and contributing factors to this spatial distribution. *Journal of Transport Geography* 91: 102976

Guimarães P, Figueiredo O, Woodward D (2000) Agglomeration and the location of foreign direct investment in Portugal. *Journal of Urban Economics* 47: 115–135. CrossRef

Guimarães P, Figueiredo O, Woodward D (2004) Industrial location modeling: Extending the random utility framework. *Journal of Regional Science* 44: 1–20. CrossRef

Haltiwanger J, Jarmin RS, Miranda J (2013) Who creates jobs? Small versus large versus young. *Review of Economics and Statistics* 90: 347–361

Henderson JW, Kelly TM, Taylor BA (2000) The impact of agglomeration economies on estimated demand thresholds: An extension of Wensley and Stabler. *Journal of Regional Science* 40: 719–733. CrossRef

Hesse M, Rodrigue JP (2004) The transport geography of logistics and freight distribution. *Journal of Transport Geography* 12: 171–184. CrossRef

Holl A (2004) Transport infrastructure, agglomeration economies, and firm birth: Empirical evidence from Portugal. *Journal of Regional Science* 44: 693–712. CrossRef

Holl A, Mariotti I (2018) The geography of logistics firm location: The role of accessibility. *Networks and Spatial Economics* 18: 337–361

Hughes C, Jackson C (2015) Death of the high street: Identification, prevention, reinvention. *Regional Studies, Regional Science* 2: 237–256. CrossRef

Jarmin RS, Miranda J (2002) The longitudinal business database (No. 02-17). Retrieved from https://www.census.gov/ces/pdf/CES-WP-02-17.pdf

Joudrey PJ, Edelman EJ, Wang EA (2019) Drive times to opioid treatment programs in urban and rural counties in 5 US states. *JAMA - Journal of the American Medical Association, American Medical Association* 322: 1310–1312

Kang S (2020) Warehouse location choice: A case study in Los Angeles, CA. *Journal of Transport Geography* 88: 102297

Kumar I, Zhalnin A, Kim A, Beaulieu LJ (2017) Transportation and logistics cluster competitive advantages in the U.S. regions: A cross-sectional and spatio-temporal analysis. *Research in Transportation Economics* 61: 25–36. CrossRef

Li Z, Madanu S (2009) Highway project level life-cycle benefit/cost analysis under certainty, risk, and uncertainty: Methodology with case study. *Journal of Transportation Engineering* 135: 516–526. CrossRef

List JA (2001) US county-level determinants of inbound FDI: Evidence from a two-step modified count data model. *International Journal of Industrial Organization* 19: 953–973. CrossRef

Loveridge S, Nawyn S, Szmecko L (2013) Conducting virtual facilitated discussions. Community development practice. Retrieved from http://www.comm-dev.org/images/pdf/Conducting-virtual-facilitated-dicussions template-new 1.pdf

Maparu TS, Mazumder TN (2017) Transport infrastructure, economic development and urbanization in India (1990-2011): Is there any causal relationship? *Transportation Research Part A: Policy and Practice* 100: 319–336. CrossRef

Marshall A (1920) *Principles of Economics*. Palgrave MacMillan, London. CrossRef

Mayeres I, Ochelen S, Proost S (1996) The marginal external costs of urban transport. *Transportation Research Part D: Transport and Environment* 1: 111–130. CrossRef

McFadden D (1973) Conditional logit analysis of qualitative choice behaviour. In: Zarembka P (ed), *Frontiers in Econometrics*. Academic Press, New York, NY, 105–142

Melecky M, Roberts M, Sharma S (2019) The wider economic benefits of transport corridors: A policy framework and illustrative application to the China-Pakistan economic corridor. *Cambridge Journal of Regions, Economy and Society* 12: 17–44. CrossRef

Monios J, Wilmsmeier G (2013) The role of intermodal transport in port regionalisation. *Transport Policy* 30: 161–172. CrossRef

Ng KYA, Gujar GC (2009) The spatial characteristics of inland transport hubs: Evidences from Southern India. *Journal of Transport Geography* 17: 346–356. CrossRef

Papke LE (1991) Interstate business tax differentials and new firm location: Evidence from panel data. *Journal of Public Economics* 45: 47–68. CrossRef

Parajuli J, Haynes KE (2017) Panel data models of new firm formation in New England. *REGION* 4: 65–76. CrossRef

Porter ME (1990) Competitive advantage of nations. *Harvard Business Review* 68: 15–34. CrossRef

Porter ME (2000) Location, competition, and economic development: Local clusters in a global economy. *Economic Development Quarterly* 14: 15–34. CrossRef

Porter ME (2001) *Clusters of Innovation: Regional Foundations of U.S. Competitiveness*. Council of Competitiveness, Washington, D.C

Quast T (2018) State-level variation in the relationship between child removals and opioid prescriptions. *Child Abuse and Neglect* 86: 306–313. CrossRef

Reum AD, Harris TR (2006) Exploring firm location beyond simple growth models: A double hurdle application. *The Journal of Regional Analysis & Policy* 36: 45–67

Rigg KK, Monnat SM, Chavez MN (2018) Opioid-related mortality in rural America: Geographic heterogeneity and intervention strategies. *International Journal of Drug Policy* 57: 119–129. CrossRef

Rivera L, Gligor D, Sheffi Y (2016) The benefits of logistics clustering. *International Journal of Physical Distribution and Logistics Management* 46: 242–268. CrossRef

Rivera L, Sheffi Y, Welsch R (2014) Logistics agglomeration in the US. *Transportation Research Part A: Policy and Practice* 59: 222–238. CrossRef

Rodrigue JP (2020) The distribution network of Amazon and the footprint of freight digitalization. *Journal of Transport Geography* 88: 102825

Rodrigue JP, Notteboom T (2010) Comparative North American and European gateway logistics: The regionalism of freight distribution. *Journal of Transport Geography* 18: 497–507. CrossRef

Sakai T, Beziat A, Heitz A (2020) Location factors for logistics facilities: Location choice modeling considering activity categories. *Journal of Transport Geography* 85: 102710. CrossRef

Smith DF, Florida R (1994) Agglomeration and industrial location: An econometric analysis of Japanese-affiliated manufacturing establishments in automotive-related industries. *Journal of Urban Economics* 36: 23–41. CrossRef

Van den Heuvel FP, de Langen PW, van Donselaar KH, Fransoo JC (2013) Spatial concentration and location dynamics in logistics: The case of a Dutch province. *Journal of Transport Geography* 28: 39–48. CrossRef

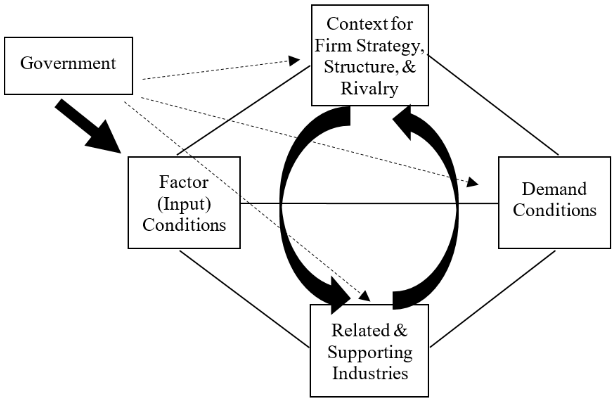
Van Sandt AT, Carpenter CW, Dudensing RM, Loveridge S (2021) Estimating determinants of health care establishment locations with restricted federal administrative data. *Health Economics* 30: 1328–1346. CrossRef

Wilkerson C (2001) Trends in rural manufacturing. Federal Reserve Bank of Kansas City. Retrieved from https://www.kansascityfed.org/documents/363/mse-Trends in Rural Manufacturing .pdf

Wilson P (2015) The misuse of the Vuong test for non-nested models to test for zero-inflation. *Economics Letters* 127: 51–53. CrossRef

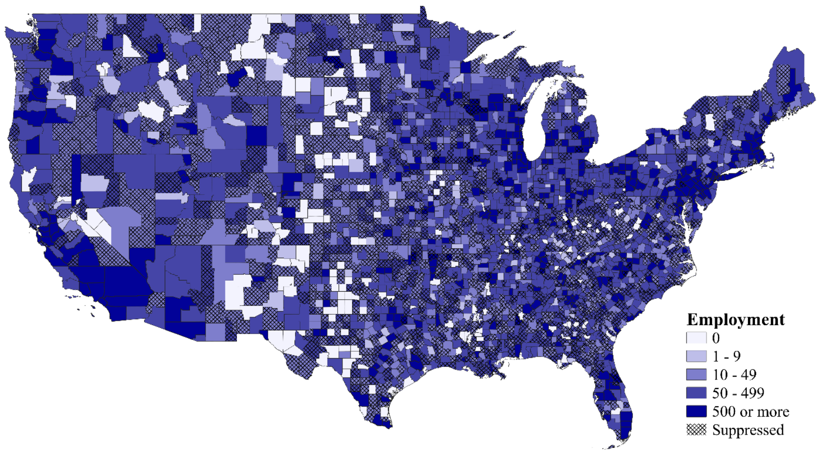
Wooldridge JM (2010) *Econometric Analysis of Cross Section and Panel Data* (2nd ed.). MIT Press, Cambridge, MA

A Appendix:



Notes: This figure adapts the Porter Diamond model to reflect the specific interaction of TW sector firm structure and related industries. This interaction recognizes that TW is often considered a support sector to other clusters and that various subsectors of TW reinforce the broader TW sector. The adapted model also shows the substantial direct effect of government on factor conditions through infrastructure investment. The bold arrow indicates the substantial effect that government investment (typically) has on factor (input) conditions.

Figure A.1: Porter diamond model adapted for TW industry clustering



Notes: Map shows employment ranges in every county in the continental U.S. "Suppressed" indicates that the exact value of the data is suppressed to prevent improper disclosure of identifiable information and placed into bins (e.g., "20–99 employees"). For the sake of these maps, we replace these ranges with their midpoint.

Figure A.2: Employment Distribution of General Freight Trucking (NAICS 4841)

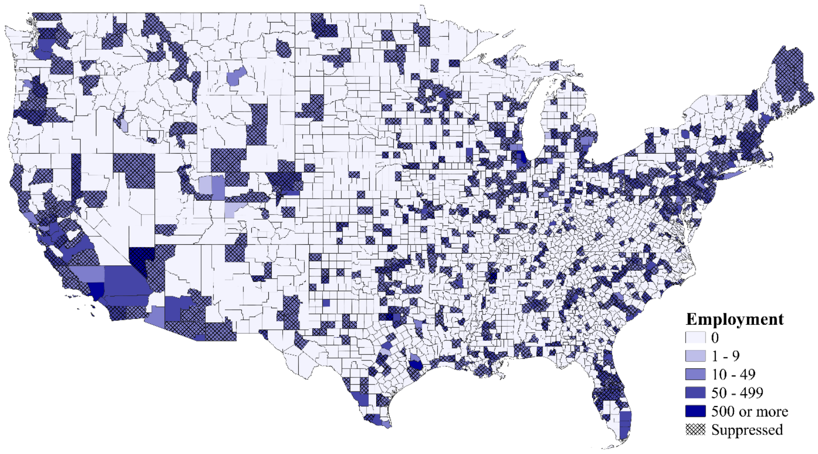


Figure A.3: Employment Distribution of Support Activities for Road Transportation (NAICS 488490)

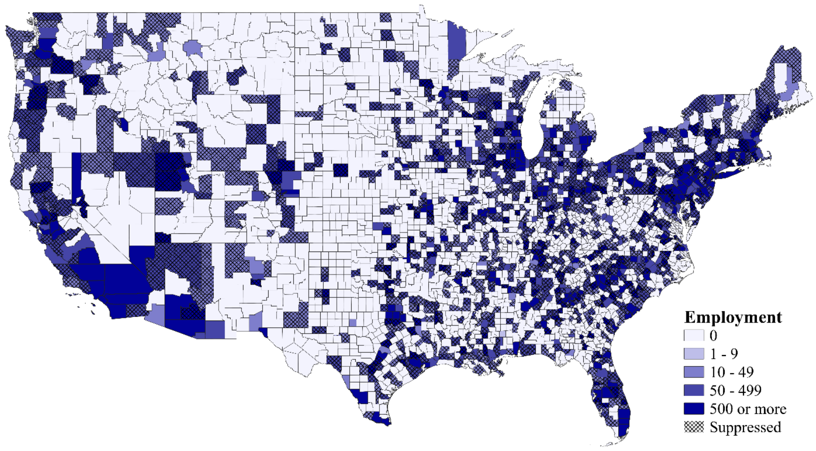


Figure A.4: Employment Distribution of General Warehousing & Storage (NAICS 493110)

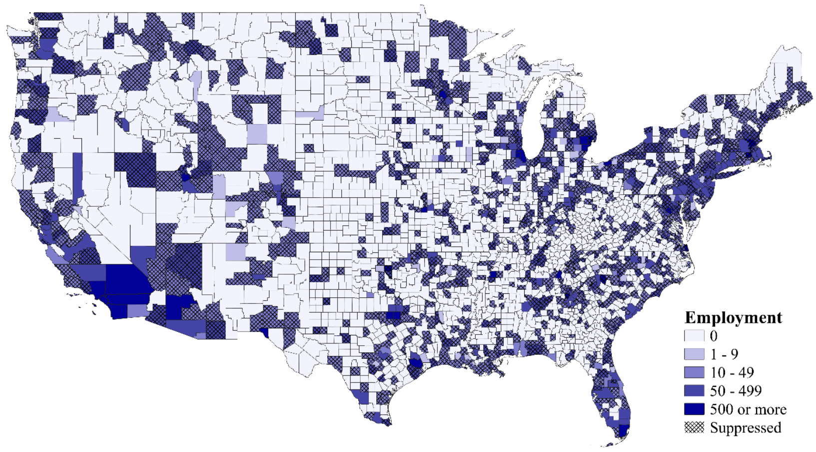


Figure A.5: Employment Distribution of Process, Distribution, & Logistics Consulting (NAICS 541614)

Table A.1: List of Abbreviations and Acronyms

ACS	American Community Survey
BEA	Bureau of Economic Analysis
BLS	Bureau of Labor Statistics
BR	Business Register
CBP	County Business Patterns
CDC	Center for Disease Control and Prevention
Emp.	Total employment
ERS	Economic Research Service
Est.	Employer establishments
FE	Fixed Effects
FCC	Federal Communications Commission
FSRDC	Federal Statistical Research Data Center
GDP	Gross Domestic Product
HP	Hurdle Poisson
ILBD	Integrated Longitudinal Business Database
ISP	Internet Service Provider
LBD	Longitudinal Business Database
LQ	Location Quotients
NAICS	North American Industry Classification System
NB	Negative Binomial
NCES	National Center for Education Statistics
Non-Emp.	Non-employer establishments
NERCRD	Northeast Regional Center for Rural Development
NS	Nonemployer Statistics
QCEW	Quarterly Census of Employment and Wages
QWI	Quarterly Workforce Indicators
TW	transportation and warehousing
USDA	United States Department of Agriculture
ZINB	Zero-Inflated Negative Binomial
ZIP	Zero-Inflated Poisson

Table A.2: Additional TW Industry Summary Statistics

Industry and size measure	Min.	Median	Mean	Max.	Source
General Warehousing & Storage (493110) Estab.	0	2	8	411	CBP
General Warehousing & Storage (493110) Emp	0	0	209	17,576	WholeData
Warehousing and Storage (4931) Non-emp	0	0	3	572	NS
Management Consulting Services (54161) Estab.	0	5	49	5,380	CBP
Management Consulting Services (54161) Emp.	0	6	293	37,724	WholeData
Management, Scientific, and Technical Consulting Services (5416) Non-emp.	0	25	226	23,562	NS
Pipeline Transportation (486) Estab.	0	2	3	155	CBP
Pipeline Transportation (486) Emp.	0	0	17	10,067	WholeData
Pipeline Transportation (486) Non-emp.	0	0	0.2	110	NS
Support for Road Transportation (488490) Estab.	0	1	3	110	CBP
Support for Road Transportation (488490) Emp.	0	0	11	1,282	WholeData
Support for Transportation (488) Non-emp.	0	9	42	6,786	NS
General Freight Trucking (4841) Estab.	0	9	23	3,929	CBP
General Freight Trucking (4841) Emp.	0	45	294	20,782	WholeData
General Freight Trucking (4841) Non-emp.	0	50	161	23,074	NS

Notes: Due to disclosure prevention limitations, these descriptive statistics are based on public data sources, while the main regression results are based on the limit-access Longitudinal Business Database and Integrated Longitudinal Business Database. All data are based on 2014; 2014 is chosen for practical reasons related to the availability of many of the secondary data sources. The internal and unsuppressed data used in regressions differ slightly in addition to the inclusion of unsuppressed cells. The more refined 541614 NAICS was not publicly available. For details on WholeData, see Bartik et al. (2018).
Abbreviations: Estab: employer establishments; Non-Emp: non-employer establishments; Emp: total employment; CBP: County Business Patterns; NS: Non-employer Statistics.

Table A.3: Poisson Marginal Effects for Locational Determinants of T&W Sectors

Industry (NAICS)	General Warehousing & Storage (493110)		Process, Physical Distribution, & Logistics Consulting (541614)		Pipeline Transportation (486)			Support Activities for Road Transportation (488490)			General Freight Trucking (4841)		
Size Measure	Est.	Emp.	Est.	Emp.	Non-Emp.	Est.	Emp.	Non-Emp.	Est.	Emp.	Non-Emp.	Est.	Emp.
Interstate Density	0.076**	16.08***	-0.006	0.125	0.059***	2.324***	0.061***	0.262***	-1.396***	0.008	-2.273***	0.887***	0.070
Interstate Density2	0.002	-0.208***	0.0009	-0.039***	-0.014***	-0.145***	-0.003	0.004	0.074***	-0.000004	0.342***	0.839***	0.022***
Highway Density	3.332***	291.8***	1.320***	5.340***	-0.224*	8.134***	-0.075	-1.537	19.29***	0.094	46.41***	280.4***	7.885***
Highway Density2	-1.578***	-154.5***	-0.418***	2.447***	0.305***	-17.58***	-0.226	-0.952*	-11.25***	-0.100	-22.60***	-131.6***	-3.697***
Water coverage	-0.027***	-4.546***	-0.007***	0.054***	-0.001	-0.111***	-0.004*	-0.067***	-0.091***	-0.002	-0.590***	-2.298***	-0.037***
Sales tax	-0.058**	0.205	-0.038	-0.949***	-0.024	-0.708***	0.034*	-0.324***	0.036	0.0005	2.937***	-9.601***	-0.192***
Property tax rate	0.014	-19.64***	-0.141	-2.433***	0.004	-0.522**	-0.064	-2.973***	1.391***	-0.003	9.717***	-27.31***	2.530***
Military bases	-0.113*	-8.909***	0.044	2.803***	0.083***	0.519***	-0.156**	-2.770***	0.412***	-0.147***	-18.22***	-21.12***	-2.162***
Community colleges	0.122***	20.69***	-0.061***	-1.023***	0.016**	-1.835***	0.056***	-0.467***	-0.734***	-0.009	-0.117	10.27***	-0.200***
Universities	-0.066***	-8.727***	0.017**	0.110***	-0.003	1.136***	0.013	0.390***	0.266***	0.005	2.317***	-1.468***	0.508***
ISP count	0.049	2.117***	-0.153***	-0.485***	-0.005	-3.371***	-0.116***	-0.499***	0.143*	0.020	-0.009	1.454***	0.691***
Social capital	0.316***	23.97***	-0.115	0.928***	0.036*	2.302***	0.169***	-2.677***	0.585***	0.037	2.218***	21.96***	1.004***
Micro Metro-Adj.	0.191	4.955***	0.268	-3.723***	-0.089	-2.853***	-0.115	-1.896***	-2.750***	-0.151	-24.87***	-19.72***	-4.514***
Metropolitan	0.509**	-13.37***	0.631***	8.756***	-0.128**	-3.609***	-0.103	-8.930***	3.582***	0.063	-61.42***	36.12***	-6.539***
Out-commute %	-0.008***	0.263***	-0.006**	-0.074***	0.002*	-0.161***	-0.006***	0.223***	-0.064***	-0.003**	0.816***	-1.424***	-0.034***
Opioid RX rate	0.013***	0.269***	0.001	0.161***	0.001**	-0.017***	0.001	0.004	0.026***	0.009	-0.485***	0.793***	-0.022***
Poverty rate	-0.001***	-0.107***	-0.0002*	-0.009***	-0.00004	-0.003***	-0.0002***	0.003***	-0.003***	-0.0001	-0.007***	-0.042***	-0.005***
Median age	-0.100***	-13.27***	0.053***	-0.181***	-0.011**	-0.602***	-0.046***	0.585***	0.396***	-0.004	0.319***	-9.345***	-0.278***
Unemployment rate	0.021	-3.166***	-0.005	0.569***	-0.003	-0.244***	-0.036**	-0.533***	-0.592***	-0.017	0.527***	-22.54***	-0.088
Bachelor's degree %	0.086***	-4.197***	0.146***	0.978***	-0.00001	-1.101***	-0.048***	-0.662***	0.450***	0.004	-4.168***	1.563***	-0.183***
ln(Population)	3.644***	259.6***	2.688***	34.50***	0.317***	16.60***	0.902***	45.85***	11.42***	0.810***	165.7***	294.2***	20.01***
Population density	-0.0002	-0.106	-0.018**	0.274***	-0.034***	-0.429***	-0.037*	-0.346***	-0.150***	-0.026***	-0.179*	-5.858***	-0.018
Per capita income	0.003	0.136***	0.008***	-0.013	0.004***	0.322***	0.016***	0.051***	0.022***	-0.002	0.649***	-0.884***	-0.018
Hispanic %	0.048***	0.325***	0.019***	0.111***	0.004***	-0.080***	0.001	0.487***	0.229***	0.010***	1.981***	3.276***	0.205***
Black %	0.020***	-0.203***	0.005	0.015	-0.0009	0.084***	0.017***	0.073***	0.075***	0.003	0.151***	1.576***	-0.019
Home Value	-0.005***	-0.554***	-0.001**	-0.030***	-0.000003	-0.036***	-0.004***	0.032***	-0.011***	0.0004	-0.339***	-0.642***	-0.024***

Significance levels: *** $p<1\%$, ** $p<5\%$, * $p<10\%$
Note: Industry size measure abbreviations: non-employer establishments (Non-Emp.); employer establishments (Est.); total employment (Emp.). Regressors in Table A.3 are included in same respective regressions as Table A.4.

Table A.4: Poisson Marginal Effects for Locational Determinants of T&W Sectors

Industry (NAICS)	General Warehousing & Storage (493110)		Process, Physical Distribution, & Logistics Consulting (541614)		Pipeline Transportation (486)			Support Activities for Road Transportation (488490)			General Freight Trucking (4841)		
Size Measure	Est.	Emp.	Est.	Emp.	Non-Emp.	Est.	Emp.	Non-Emp.	Est.	Emp.	Non-Emp.	Est.	Emp.
Location Quotients													
TW	1.102***	66.41***	-2.497***	-72.76***	0.049***	1.195***	0.035	4.571***	5.500***	0.247***	15.78***	82.31***	3.767***
Finance	-0.305***	3.089***	-0.172***	-4.355***	-0.263***	-1.430***	0.001	-0.991***	-0.966***	0.033	4.882***	14.45***	-0.394**
Real Estate	-0.283***	-48.51***	0.009	-1.288***	0.051***	2.287***	0.223***	2.329***	-0.282*	0.047	-4.488***	5.991***	0.117
Professional and tech	-0.552***	-68.38***	1.850***	55.11***	-0.073**	2.016***	-0.056	-1.453***	-3.574***	-0.037	-4.804***	-80.04***	-3.565***
Education	-0.181***	-4.207***	-0.077*	-2.064***	-0.055**	-2.163***	-0.117***	-1.191***	0.219***	-0.012	-8.694***	-9.805***	-1.545***
Health services	-0.776***	-33.54***	-0.825***	-13.27***	0.126***	1.005***	0.111**	-6.506***	-3.118***	-0.031	-19.85***	-73.12***	-2.183***
Art and recreation	-0.087	9.509***	0.032	-3.366***	0.017	-0.987***	-0.052**	0.213	-0.663***	0.010	-3.518***	-12.62***	-0.548***
Accommodation & Food	-0.237**	-24.41***	-0.190**	-2.526***	-0.146***	-3.186***	-0.098*	-1.369***	2.705***	0.017	-20.45***	-118.9***	-5.207***
Mining and gas	-0.082***	-6.507***	0.013**	0.167***	0.003**	0.230***	0.021***	0.087***	0.103***	0.010***	-0.542***	-3.890***	-0.034**
Construction	-0.625***	-45.65***	-0.268***	-3.843***	0.022	2.388***	0.173***	-1.003***	-2.726***	-0.045	-0.880**	-11.95***	0.235
Retail	-0.683***	10.30***	0.061	-0.839	0.056	-5.767***	-0.075	1.937***	-1.395***	0.074	12.20***	-45.02***	1.097***
Regional FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
n	3,063	3,063	3,063	3,063	3,063	3,063	3,063	3,063	3,063	3,063	3,063	3,063	3,063

Significance levels: *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$

Note: The exact observation count is suppressed by U.S. Census Bureau disclosure review process. We use all counties in the continental U.S. in 2014. Industry size measure abbreviations: non-employer establishments (Non-Emp.); employer establishments (Est.); total employment (Emp.). TW LQs are calculated on all of NAICS 48, 49, and 541614, except for the respective NAICS under consideration. TW LQs are calculated on all of NAICS 48, 49, and 541614, minus the respective NAICS under consideration. Regressors in Table A.4 are included in same respective regressions as Table A.3. Regional fixed effects (FE) based on Census regions.

Table A.5: Inflation Stage Marginal Effects for Locational Determinants of T&W Sectors

Industry (NAICS)	General Warehousing & Storage (493110)		Process, Physical Distribution, & Logistics Consulting (541614)		Pipeline Transportation (486)			Support Activities for Road Transportation (488490)		
Size Measure	Est.	Emp.	Est.	Emp.	Non-Emp.	Est.	Emp.	Non-Emp.	Est.	Emp.
Interstate Density	-0.011	-0.027***	-0.053*	-0.02	0.002	-0.020*	0.037	-0.015**	-0.026***	-0.066*
Highway Density	-0.0006	0.002*	0.004	0.003	-0.0001	0.002	-0.022	0.002**	0.003**	0.007*
Water coverage	-0.066	-0.137	1.752***	0.761***	0.265***	0.260*	-0.223	0.238*	0.495***	0.911
Sales tax	0.232	0.119	-1.870**	-0.812***	-0.205***	0.025	0.443	-0.265	-0.430**	-1.536
Property tax rate	-0.0002	-0.00001	-0.031*	-0.0004	-0.0005	0.005***	0.001	-0.001*	0.001	-0.006
ISP count	0.011	0.004	-0.00003	0.0009	-0.032***	-0.029***	-0.010*	-0.008**	-0.002	-0.013
Micro Metro-Adj.	-0.016	0.0005	0.030	0.025	-0.013	0.006	-0.012	0.067***	-0.005	-0.111
Metropolitan	0.018	-0.005	-0.005	-0.005	-0.003	0.025**	0.022***	0.002	-0.010	0.022
ln(Population)	-0.016	0.007	0.019	-0.034	0.023	0.036	-0.982	-0.019	0.007	-0.159*
Population density	-0.047	0.022	-0.003	-0.025	0.042**	0.010	-0.040	-0.048***	-0.014	-0.147
Per capita income	-0.131***	-0.195***	-0.031	-0.111***	-0.052***	-0.159***	-0.064***	-0.203***	-0.147***	-0.080*
Home Value	-2.568**	-0.582**	-0.985**	-1.037***	0.015**	-0.154*	-0.131	0.007	-0.164	-0.187

Significance levels: *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$

Note: Industry size measure abbreviations: non-employer establishments (Non-Emp.); employer establishments (Est.); total employment (Emp.). No inflation stage for the General Freight Trucking (4841) industry. Regressors in Table A.5 are included in same respective regressions as Tables A.6 and A.7.

Table A.6: Amount Stage Marginal Effects for Locational Determinants of T&W Sectors

Industry (NAICS)	General Warehousing & Storage (493110)		Process, Physical Distribution, & Logistics Consulting (541614)		Pipeline Transportation (486)			Support Activities for Road Transportation (488490)			General Freight Trucking (4841)		
Size Measure	Est.	Emp.	Est.	Emp.	Non-Emp.	Est.	Emp.	Non-Emp.	Est.	Emp.	Non-Emp.	Est.	Emp.
Interstate Density	0.194***	31.01***	-0.014		0.067*	0.413	-0.005	0.156	-1.086	0.012	-1.350	6.508	0.052
Interstate Density2	-0.014**	-2.739***	0.0003		-0.015**	0.047	0.019	0.004	-0.038	-0.001	0.181	-0.133	0.007
Highway Density	3.740***	368.0***	0.911		-0.381	-2.822	0.16	-1.714	10.98	-0.040	10.93	174.1**	2.485
Highway Density2	-1.786***	-176.5***	0.249		0.315	-3.607	-0.555	0.149	-0.646	0.220	-15.30*	-97.97**	-2.447
Water coverage	-0.026***	-5.346***	0.006		-0.002	-0.115	-0.009**	-0.066***	-0.138*	-0.001	-0.831***	-3.033***	-0.097***
Sales tax	-0.051	4.84	-0.038		-0.157***	0.032	0.026	-0.209*	-0.991	-0.002	-2.731***	-11.10**	-0.513***
Property tax rate	-0.17	-10.13	-0.17		-0.112*	-0.902	0.017	-2.605***	-2.032	-0.014	-0.065	-23.88	0.206
Military bases	-0.248	22.59	0.094		0.060**	3.857	-0.029	-2.331***	5.225*	-0.108**	-7.913	-40.29	-1.690*
Community colleges	0.047	6.841	-0.043		0.002	2.224	0.205***	-0.362***	-0.060	-0.0003	5.200**	-2.626	0.124
Universities	-0.031	1.958	0.019		0.002	0.785	-0.022	0.400***	0.739	0.005	0.472	-2.849	0.208
ISP count	0.027	5.501	-0.144***		-0.053**	-0.535	-0.053	-0.703***	-0.559	0.023	-0.944	-8.057	0.506**
Social capital	0.108	7.375	-0.122		0.031	3.832**	0.194***	-2.504***	-0.894	0.047	5.065***	31.12***	1.630***
Micro Metro-Adj.	-0.468	-67.24*	0.363		-0.012	-1.479	0.416	-5.784***	-4.664	-0.373**	-8.256**	-24.41	-3.122***
Metropolitan	-0.552	-117.5***	0.275		-0.061	-1.312	-0.243*	-12.66***	-0.566	-0.172	-21.39***	-22.99	-4.500***
Out-commute %	-0.017***	1.018	-0.005		-0.001	-0.106	-0.004	0.204***	0.051	-0.003*	0.523***	-2.661***	-0.018
Opioid RX rate	0.006**	-0.917**	-0.0004		-0.001	0.003	0.0007	-0.036***	0.023	0.0008	-0.200***	0.144	-0.015**
Poverty rate	-0.001***	-0.096***	-0.0002		0.00004	-0.006*	-0.0002**	0.002***	-0.002	-0.00004	-0.008**	-0.080***	-0.004***
Median age	-0.052**	-3.915	0.053***		0.009	-1.189**	-0.055***	0.764***	0.687**	-0.0007	-0.089	-10.76***	-0.237***
Unemployment rate	0.105*	-22.52**	-0.021		0.003	0.961	0.019	-0.638***	-0.586	-0.011	-1.903**	-28.53***	-0.790***
Bachelor's degree %	0.061**	-4.458	0.134***		0.017*	-0.144	-0.033**	-0.684***	0.028	0.002	-5.332***	-8.055***	-0.435***
ln(Population)	3.699***	236.2***	2.593***		0.260***	11.99***	0.867***	45.38***	9.421***	0.796***	138.4***	369.2***	20.44***
Population density	1.279**	91.58***	0.421*		0.006	1.14	0.100	-0.286	1.904	0.019	2.563**	0.712	0.201
Per capita income	-0.003	-0.146	0.008		0.0004	0.227**	0.015***	0.046***	0.005	-0.003	0.947***	0.047	0.056*
Hispanic %	0.039***	1.038	0.019***		0.002	-0.119	-0.002	0.486***	0.271**	0.009***	1.123***	0.524	0.093***
Black %	0.037***	2.380*	0.005		-0.004	-0.086	0.012***	0.099***	0.245*	0.004	0.528***	2.291***	0.02
Home Value	-0.003*	-0.643**	-0.0002		-0.001	-0.064**	-0.004***	0.027***	-0.002	0.001	-0.185***	-0.479**	-0.021***

Significance levels: *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$

Note: Industry size measure abbreviations: non-employer establishments (Non-Emp.); employer establishments (Est.); total employment (Emp.). There is no inflation stage for the General Freight Trucking (4841) industry because it did not need to be zero-inflated (it is modeled as NB). Regressors in Table A.6 are included in same respective regressions as Tables A.5 and A.7.

Table A.7: Amount Stage Marginal Effects for Locational Determinants of T&W Sectors

Industry (NAICS)	General Warehousing & Storage (493110)		Process, Physical Distribution, & Logistics Consulting (541614)		Pipeline Transportation (486)			Support Activities for Road Transportation (488490)			General Freight Trucking (4841)		
Size Measure	Est.	Emp.	Est.	Emp.	Non-Emp.	Est.	Emp.	Non-Emp.	Est.	Emp.	Non-Emp.	Est.	Emp.
Location Quotients													
TW	1.315***	101.0***	-3.332***		0.047*	1.518	0.020	4.572***	8.460***	0.247***	6.303***	69.34***	2.594***
Finance	-0.207*	-0.032	-0.116		-0.135**	-1.94	0.140*	-0.664***	2.174	0.053	6.334***	10.69	-0.483
Real Estate	-0.396**	-81.95***	0.007		0.021	1.268	0.212***	2.399***	2.695	0.064	-1.527	-9.462	0.345
Professional and tech	-0.589***	-49.66**	2.560***		-0.026	1.363	-0.049	-1.286***	-0.283	-0.013	-11.18***	-58.93***	-3.313***
Education	-0.116	-1.873	-0.087		0.007	-2.107**	-0.093**	-1.254***	0.492	-0.029	-3.335***	-20.54***	-1.155***
Health services	-0.566***	36.09	-0.661***		-0.070	3.703*	0.199**	-5.707***	-4.684*	-0.030	-8.907***	-64.94***	-1.977***
Art and recreation	0.020	12.61	0.017		-0.010	-0.834	-0.045	0.484***	-1.752*	0.003	-2.238***	-18.93***	-0.745***
Accommodation & Food	-0.227	-60.12**	-0.215		-0.081	-2.495	-0.072	-1.788***	-1.174	-0.006	-14.20***	-45.68***	-3.201***
Mining and gas	-0.075***	-4.745***	0.015*		0.004*	0.394**	0.030***	0.103***	-0.081	0.010***	-0.882***	-2.506***	-0.096***
Construction	-0.540***	-55.36***	-0.201**		0.006	2.133*	0.168***	-0.886***	-1.958	-0.050	-3.695***	-13.39**	-0.229
Retail	-0.922***	49.48	-0.153		0.017	-4.057*	-0.131	0.595	-0.31	0.031	3.422	-120.8***	-1.468**
Regional FE	YES	YES	YES		YES	YES	YES	YES	YES	YES	YES	YES	YES
n	3,063	3,063	3,063		3,063	3,063	3,063	3,063	3,063	3,063	3,063	3,063	3,063

Significance levels: *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$

Note: The exact observation count is suppressed by U.S. Census Bureau disclosure review process. We use all counties in the continental U.S. in 2014. Industry size measure abbreviations: non-employer establishments (Non-Emp.); employer establishments (Est.); total employment (Emp.). TW LQs are calculated on all of NAICS 48, 49, and 541614, except for the respective NAICS under consideration. No inflation stage for the General Freight Trucking (4841) industry. TW LQs are calculated on all of NAICS 48, 49, and 541614, minus the respective NAICS under consideration. Regressors in Table A.7 are included in same respective regressions as Tables A.5 and A.6. Regional fixed effects (FE) based on Census regions. The employment model for 541614 did not converge – likely due to lack of variation in public data.