Can Trade Networks Explain the Productivity-Compensation Gap?

Shishir Shakya[1,2*] and Alicia Plemmons[3†]

[1]John Chambers College of Business and Economics, West Virginia University, 1601 University Ave, Morgantown, WV, 26506, United States.
[2]Regional Research Institute (RRI), West Virginia University, 886 Chestnut Ridge Road, Morgantown, WV, 26505, United States.
[3]College of Business, Southern Illinois University Edwardsville, 1 Hairpin Dr, Edwardsville, IL, 62025, United States.

Abstract
Historically, compensation and productivity have steadily prospered together. But since the mid-1970s, these two measures have begun to diverge as increases in productivity have outpaced compensation rates. This paper approaches analyzing this productivity-compensation gap using two methods. First, we investigate a series of spatial contiguity matrices for industries between all U.S. states. Second, we determine the extent to which state-level trade networks of producers across industries explain the productivity-compensation gap. We incorporate the Commodity Flow Survey Public Use Microdata File (CFS PUMS) to determine the interregional industry trade. In conjunction with CFS PUMPS, we retrieve compensation, employment, and industrial output for 71 industries across all U.S. states using IO-SNAP software. This paper improves upon the current literature on spatial spillover and spatial contiguity matrices. We contribute by integrating an interregional industry trade network as the contiguity matrix. This contiguity matrix allows explaining how the labor compensation of several industries correlates across states. This paper reinforces previous results and contributes new insights on labor compensation effects within trade networks when compared to the spatial contiguity matrix. These results suggest the existence of productivity-compensation gaps, both spatially and within the regional trade networks.

1 INTRODUCTION

The presence and causes of the productivity-compensation gap in the United States has been debated through the economic literature for decades using spatial spillover models

*E-mail: ss0088@mix.wvu.edu
†E-mail: aplemmo@siue.edu

to depict the convergence or divergence of compensation rates over geographic space. We hypothesis that compensation is not only correlated over spatial networks, but over trade networks and that this is an important relationship to understand when designing empirical tests and developing labor policies. To empirically test this relationship, we use a series of network autoregression models using a variety of contiguity matrices developed from different attributes of interregional trade networks by industry.

The rate of growth between labor compensation and productivity has been steadily diverging in the United States since the 1970s (Brill et al., 2017; Fleck et al., 2011). Addressing the contiguity of states, which is the direct physical or trade connection between states, both geographically and within trade networks, is crucial since it encourages the process of convergence (Tavernier and Temel, 1997). To determine the intensity of the relationship of geography and trade networks, this study uses the theoretical and empirical foundations of gravity models and spatial autoregressive models, penned by Tavernier and Temel (1997) and Anselin (1988), respectively.

The purpose of this study is to use new measure of spatial and network spillover to discern the role of compensation correlations, derived from state-level interregional producer trade networks, within the compensation-productivity gap. To do so, we use IO-SNAP software to integrate industry data on employment, compensation, and gross industrial output with shipment value and flow data from the 2012 Community Flow Survey Public Use Microdata file. We then use a variety of spatial and network autoregressive models to analyze the convergence and divergence of wages within industries, as well as the rate of change in labor compensation with respect to productivity.

This study improves upon the current literature by addressing relationships among trade networks that influences labor compensation rather than only allowing the relationship to fluctuate spatially. We also provide ranges of estimates for a collection of industries that are able to determine if the productivity-compensation gap is random or strategically occurs.

We first confirm spatial correlation results from previous studies and find positive correlations between 20 industry groupings for compensation and spatial proximity. When allowing for a network autoregression model to integrate a row-standardizing binary contiguity network matrix, and later a shipment value weighted network matrix, we find evidence of large, and statistically significant correlations (above 0.40) between trade networks and labor compensation for Miscellaneous manufacturing, Motor vehicles, bodies and trailers, and parts, Nonmetallic mineral products, and Wood products. Interestingly, we also found that some industries have weakly negative network coefficients, implying labor compensation may diverge within a few distinct industries.

The paper proceeds as follows: Section 2 provides a comprehensive overview of the relevant literature. Section 3 provides details on the data that will be used within the analysis. Section 4 develops the methodology. Section 4 presents results of the empirical analysis. Section 6 concludes and discusses policy relevance.

2 LITERATURE REVIEW

The determinants of the gap between growth in compensation and growth in productivity over the last half century is an important unresolved question in regional and labor economics. Empirically, higher real compensation, as measured by the ratio of total compensation to total employment, is correlated with higher real productivity, or the real output per hour worked. But in the United States these measures have steadily diverged since the 1970s, a phenomenon that is referred to as the productivity-compensation gap and depicted in Figure 2 (Brill et al., 2017; Fleck et al., 2011; Sprague, 2017). Though this figure is intuitive, it highlights an issue in productivity-compensation gap research in that most studies are limited to total nonfarm business sectors and are aggregated at the national-level, disallowing for analysis at the regional, state, or metropolitan-levels (Bosworth et al., 1994). A recent report by Bureau of Labor Statistics illustrated the productivity-compensation gap in sectors and industries but does not account for the impact of trade network proximities on wages.

Addressing the contiguity of states and trade networks is crucial since the cost of mobility and technology transfers across state boundaries has reduced substantially over time, which encourages the process of spatial convergence (Tavernier and Temel, 1997). Using neoclassical growth theory, states operate similar to small open economies and increases in factor mobility rates diminish regional differences and over time the world economy will ultimately reach a steady state in employment, wages, and growth rates (Barro and Sala-i Martin, 1992; Mallick and Carayannis, 1994; McCombie, 1988; Smith, 1975). Therefore, if a state or regional does have a regional wage differential, this differential encourages the migration of labor from low to high wage states. This idea of regional convergence hinges on the assumption that there are diminishing returns to capital where high wage states will have slower growth in wages than the respective low wage states. In doing so, productivity increases should be compensated equally as states converge in wages.

However, the contiguity of these states define the probability a worker will move between states. For example, a worker may be willing to move to a state with a shared geographic border for a nominal pay increase, but moving across the country over many states is a more costly investment and the worker would need a larger pay increase to be advantageous. These spatial spillovers, convergences, and divergences are important factors in regional analysis for understanding the motivations and spatial movements of individuals and businesses (Anselin, 1988; Burridge and Gordon, 1981; Manning, 1994; Molho, 1982; Plemmons, 2019). One way to visualize contiguity is spatially in terms of geographic proximity. Alternatively, contiguity of states can be evaluated on the degree to which states that they are close together within interregional industry trade networks rather than geographic distance. This is the operationalize measure of contiguity of states that we use in this study. To determine the intensity of the relationship of trade and geographical distance and location, this study utilizes the theoretical and empirical foundations of gravity models developed by Tinbergen (1962).

Gravity models of geographical analysis predict that the value of trade will decay as distance increases. However, by visualizing trade networks as neighborhoods and not requiring them to be located within a short physical distance but instead close in trade dependence, our model is able to provide richer supplemental results and information than limiting only to spatial proximity analysis. By evaluating how these trade proximities affect the spatial spillover of labor compensation within networks, this analysis provides new insight and understanding into the underlying determinants of the productivity-compensation gap.

3 DATA

Commodity Flow Survey Public Use Microdata File (CFS PUMS)

Inter-regional trade data was elicited through the 2012 Commodity Flow Survey Public Use Microdata File (CFS PUMS) . This data contains origin, destination, NAICS code of shipper, and the value of the shipment for approximately 4.5 million observable shipments. To protect individual business information, this data is aggregated by the Census at both the metropolitan and state-levels (ESMPD, 2012). This study aggregates this microdata to generate industry specific state-level shipment values for each NAICS code.

IO-Snap

Industry data regarding employment, compensation, and gross industrial output is collected through IO-Snap[1] for 71 industries within all U.S. states. From this, the compensation rate is calculated as the ratio of total compensation to total employees within each industry and state pair. Labor productivity is similarly calculated using gross industrial product and total employment for industry-state pairs. These state-level aggregations factor in corrections for scrap, second-hand, and used goods production within its impact assessment. These analyzes determine final demand changes of commodities to generate output, employment, and income impacts of final demand shocks on the industry and commodity space.

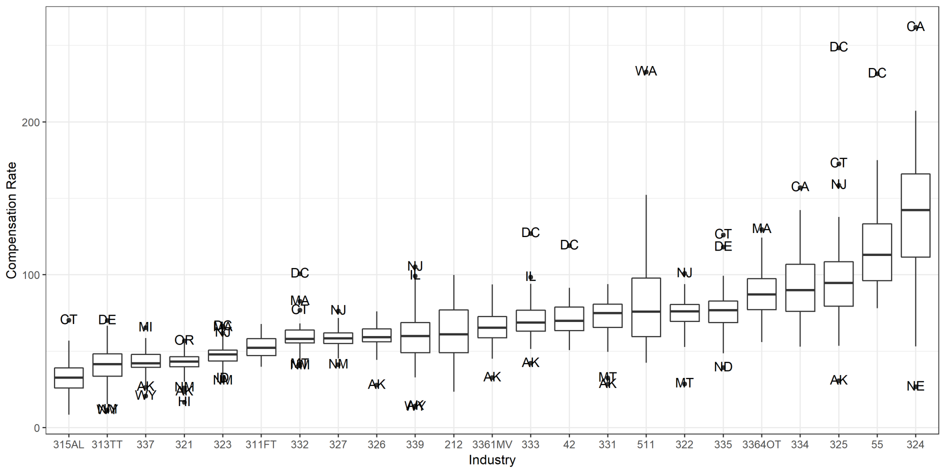
Summary Statistics

Figure 1 are provided for data visualization and discussion of the 71 industries included within this study. The boxplots for industries are ordered by representative median values. Figure 1, panel (a) depicts a boxplot of compensation rates adjusted for regional price parities in thousands of USD per employee, for a selection of industries. The petroleum and coal product industry (324) has the highest median compensation, while apparel and leather and allied products industry (315AL) compensates the least. Figure 1, panel (b) contains a series of boxplots of state-level productivities across industries. It is interesting to note, often the District of Columbia is an outlier with well above average compensation

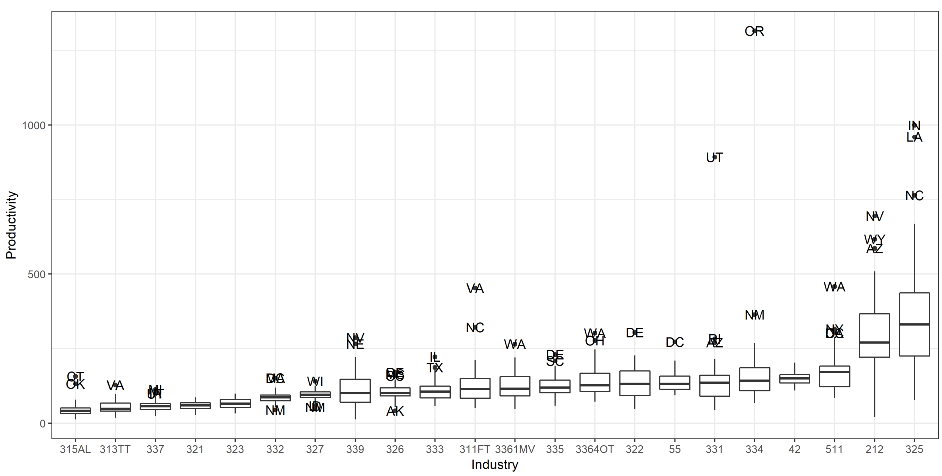
[1] IO-Snap uses national tables of input output table from the U.S. Bureau of Economic Analysis (BEA) to generate input output accounts and coefficients for state, national, and user customized regional economies using input-output impacts assessment tools and basic economic information on employment, compensation, and value-added data (RRI, 2017). The Regional Research Institute (RRI) at West Virginia University is the developer of the IO-Snap software.

in most industries. Similar to the compensation case, the petroleum and coal product industry (324) maintains the highest median productivity rates, followed closely by the chemical product industry (325). States with compensation rate and labor productivity above or below a 1.5 interquartile range of the mean are labeled as outliers.

Figure 1: Box Plots of Compensation Rates Adjusted for Regional Price Parities



(a) Compensation Rates Adjusted for Regional Price Parities



(b) Labor Productivity

4 METHODOLOGY

To determine a range of potential effects of networks on compensation and productivity, we develop four spatial autoregressive models utilizing unique variations of contiguity. The first is a spatial contiguity weight matrix focuses on geographic borders (W_s); followed by the row-standardized binary contiguity network $(W_{N1,i})$; the shipment value weighted network matrix $(W_{N2,i})$; and finally a scale-free random trade network incorporating 1000 counterfactuals $(W_{N3,i})$. The theoretical arguments and empirical outlines for these methodologies are discussed below.

Spatial Contiguity Weight Matrix

To develop a base for comparison, we first begin by presenting a spatial model with a binary contiguity weight matrix known as the spatial autoregressive model (Anselin, 1988). The purpose of this model is to provide insight into how labor compensation varies spatially, with the assumption that the labor compensation of each industry depends upon labor productivity. The model is structured as follows:

$$y_i = \rho_0 W_s y_i + \gamma_0 + \beta_0 x_i + \varepsilon_i \tag{1}$$

where y_i is the compensation rate in a industry, i. The ρ_0 is the spatial autoregressive coefficient, W_s represents the row-standardized binary contiguity spatial weight matrix for each individual state. The β_0 coefficient represents the rate of change of the compensation rate with respect to the labor productivity of each industry. A positive (negative) spatial autoregressive coefficient of an industry indicated that the change of labor compensation within that industry tends to converge (diverge) among the states.

Row-Standardizing Binary Contiguity Network Matrix

Deviating from the spatial autoregressive model, the row-standardizing binary contiguity network matrix utilizes a network autoregressive model. These models weigh distance between two states in a trade network within industries, rather than over spatial distance by utilizing a similar approach to (Anselin, 1988), but, replacing the binary contiguity spatial weight matrix with a binary contiguity network matrix. The logic behind this model is that with the rise of technology and the decreased cost of transporting good and services, how states relate through trade could be a more important factor to their relative productivity and compensation rates than how close states are in a physical proximity. The model is similar to Model (1), except, ρ_1 represents the network autoregressive coefficient and W_s is replaced by $(W_{N1,i})$ to represent the interregional trade network.

$$y_i = \rho_1 W_{N1,i} y_i + \gamma_1 + \beta_1 x_i + \eta_i \tag{2}$$

A positive (negative) network autoregressive coefficient of an industry indicates that labor compensation within that industry tends to coverage (diverge) among states that connected within the trade network supply chain.

For a given industry, the row-standardizing binary contiguity matrix shows which region ships how much value of goods to other region. These matrices represent direct graphs of trade networks, where nodes of regions are linked if there are commodity shipments between the two regions greater than an arbitrary threshold. Using the common approach for binary contiguity matrices a network is equal to one if they are above the value threshold where all regions have at least one neighbor in the trade network. This threshold method results in a non-symmetric matrix of the first-order neighbors, which represent the shortest path between nodes or regions.

Shipment Value Weighted Network Matrix

The third model also follows the functional form of the network autoregressive model. The theoretical basis for this model is the assumption that labor compensation of an industry can be potentially correlated through their trading patterns among regions or states based on the value of the shipments through those trade networks. This means that an industry operating in multiple states may see a convergence of wage compensation based upon the value of goods shipped that are transported through established trade networks between states. This model is similar to Model (2) and differs in that the weighted matrix is replaced by the shipment flow contiguity network matrix, $(W_{N2,i})$

$$y_i = \rho_2 W_{N2,i} y_i + \gamma_2 + \beta_2 x_i + \mu_i \tag{3}$$

Scale-Free Random Trade Network Matrix

The scale-free random trade network matrix within the network autoregressive models are developed by utilizing a system of counterfactual analyzes of the trade network relationships. The theoretical foundation for these network relationships are discussed further in Appendix A.

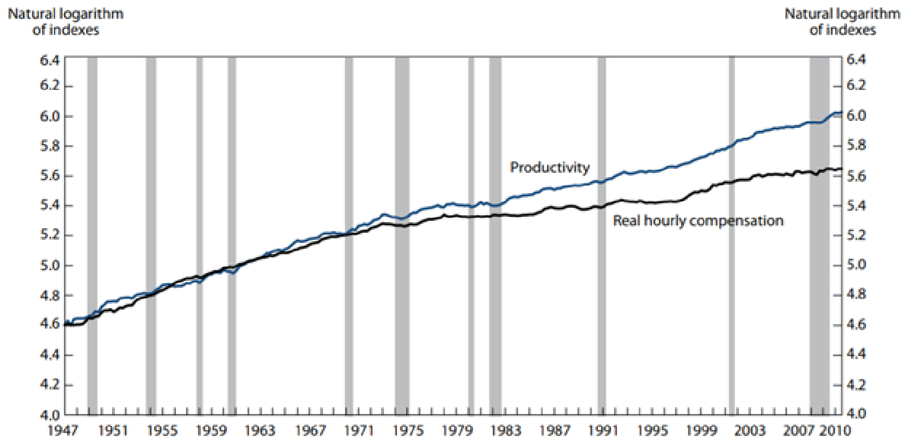
For a particular value of x_{min}, the MLE can be directly estimated the standard error of $\widehat{\alpha}$. Bootstrapping, following the Efron and Tibshirani (1993) procedure, is incorporated to account for additional uncertainty of x_{min} and to quantify the plausibility that the distribution of data follows a power distribution. The bootstrap method generates a series of goodness-of-fit tests and generates a p-value representing the plausibility that the data follows a power law distribution. Once x_{min} and α are estimated, we use these parameters to simulate 1000 counterfactual scale-free random trade networks $(W_{N3,i})$ and run the following counter network autoregressive model. Then, we record the value of ρ_3 to show it's distribution.

$$y_i = \rho_3 W_{N2,i} y_i + \gamma_3 + \beta_3 x_i + \nu_i \tag{4}$$

5 RESULTS

The results are presented for each of the four autoregressive models. The first uses the standard spatial contiguity model determining how sharing geographic proximity affects the distribution of compensation rates. The second model expands and focuses on regions that are connected within trade networks rather than with physical borders using a binary identification strategy. The third model allows for variation in the trade networks by accounting for shipment values of commodities between regions. Finally, the last model uses counterfactual trade networks within the autoregressive model to determine the range of effects of these trade network connections on compensation rates.

Figure 2: Quarterly Productivity and Real Hourly Compensation, Non-Farm Business Sector (1947-2010)



Source: Brill et al. (2017)

Spatial Contiguity Weight Matrix

Table 1 contains the coefficient estimates of the spatial autoregressive model. The purpose of this model is to establish a baseline of the effects of geographic proximity on labor compensation differentials as a point of comparison for the trade network models. The results are presented in an increasing order based on the spatial correlation of labor compensation. The first column of results contains the spatial correlation estimate, and its corresponding p-value are listed in the second column. The third and fourth columns contain the β_0 coefficient estimates, or the rate of change of the wage compensation with respect to labor productivity within that industry, and the corresponding p-value.

Within the results, we find that the labor compensation of Chemical products, Primary metals, and Wood products industry are highly spatially correlated with coefficient estimates above 0.30. Of these three categories with the largest rates of spatial correlation, Wood products appears to have the most substantial change in compensation with respect to labor productivity, with a 100% increase in labor productivity resulting in an increase in wage compensation of approximately 33%. These estimates are much smaller for chemical production and primary metals, with the rate of change in compensation with respect to labor productivity maintaining a value of 0.08 and 0.03, respectively. 17 industries have a positive estimate of spatial correlation between 0.01 and 0.30, but these relationships are only statistically significant at the 5%-level for Miscellaneous manufacturing, Furniture and related products, Printing and related support activities, Electrical equipment, appliances, and components, Machinery, and Food and beverage and tobacco products. The labor compensation coefficient estimate is also slightly negative and statistically significant for Management of companies and enterprises. The coefficient estimates for the change in

TABLE 1: Spatial Autoregressive Model using a Spatial Contiguity Weight Matrix

Industries	ρ_0	p-value	β_0	p-value	Pseudo R^2
Chemical products (325)	0.41	0.00	0.08	0.00	0.32
Primary metals (331)	0.38	0.00	0.03	0.01	0.16
Wood products (321)	0.31	0.00	0.33	0.00	0.75
Miscellaneous manufacturing (339)	0.28	0.01	0.24	0.00	0.65
Textile mills and textile product mills (313TT)	0.25	0.06	0.32	0.00	0.43
Furniture and related products (337)	0.25	0.00	0.28	0.00	0.52
Printing and related support activities (323)	0.24	0.01	0.25	0.00	0.37
Electrical equipment, appliances, and components (335)	0.24	0.01	0.23	0.00	0.45
Computer and electronic products (334)	0.24	0.06	0.06	0.00	0.26
Machinery (333)	0.23	0.03	0.20	0.00	0.29
Plastics and rubber products (326)	0.22	0.00	0.18	0.00	0.52
Apparel and leather and allied products (315AL)	0.18	0.18	0.34	0.00	0.53
Other transportation equipment (3364OT)	0.16	0.12	0.20	0.00	0.36
Wholesale trade (42)	0.15	0.02	0.49	0.00	0.71
Mining, except oil and gas (212)	0.15	0.28	0.09	0.00	0.38
Paper products (322)	0.13	0.15	0.11	0.00	0.28
Nonmetallic mineral products (327)	0.12	0.08	0.17	0.00	0.27
Motor vehicles, bodies and trailers, and parts (3361MV)	0.11	0.27	0.13	0.00	0.33
Fabricated metal products (332)	0.07	0.30	0.37	0.00	0.66
Food and beverage and tobacco products (311FT)	0.04	0.66	0.05	0.00	0.25
Publishing industries, except internet (includes software) (511)	0.00	0.98	0.48	0.00	0.95
Management of companies and enterprises (55)	-0.02	0.02	0.84	0.00	1.00
Petroleum and coal products (324)	-0.06	0.70	0.04	0.00	0.36

labor compensation with respect to labor productivity, β_0, and positive and significant for all industries, implying that compensation rates increase as productivity increases, though none of these are a one-to-one ratio, and instead this change in labor compensation rates range from 0.30 for Primary metals to 0.84 for Management of companies and enterprises. This suggests that industries vary drastically in how their compensation rates change with respect to productivity rates, even when accounting for the spatial nature of these regions relative to other regions using physical proximity.

Row-Standardizing Binary Contiguity Network Matrix

Table 2 presents the results for the network autoregressive model utilizing a row-standardizing binary contiguity network matrix. The structure of Table 2 is similar in format to Table 1, except instead of using the spatial contiguity matrix, the network correlation coefficient estimates, ρ_1, are depicted for a row-standardizing binary contiguity network matrix, $W_{N1,i}$. We find that there are highly significant and large network autoregressive coefficients, ranging from 0.60 to 0.74, for Wood products, Motor vehicles, bodies and trailers, and parts, Miscellaneous manufacturing, and Nonmetallic mineral products. This suggest that within these industries, labor compensation is highly positively correlated among the trade network. There is also evidence of a positive correlation, though not significant, of magnitudes between 0.03 and 0.37 for 11 other industries. Some industries may instead diverge in their labor compensation rates within trade networks, which is found as a negative and insignificant relationship for 7 industries. Electrical equipment, appliances, and components maintains the only statistically significant, large, negative coefficient estimate for the effect of network contiguity on labor compensation of -0.76, indicating that within this industry wages may instead diverge as regions become connected in trade which may represent specialization or competition.

The β_1 coefficient representing the compensation rate changes with respect to productivity are positive and significant for all industries. These estimates range from 0.04 to 0.84 across industries and show that as productivity increases, compensation also increases but not at a one-to-one ratio and instead gains in productivity are associated with smaller percentage increases in wages, which depicts the compensation-productivity gap persists even within this specification of trade network measures.

TABLE 2: Network Autoregressive Model using a Row-Standardizing Binary Contiguity Network Matrix

Industries	ρ_2	p-value	β_2	p-value	Pseudo R^2
Wood products (321)	0.74	0.00	0.37	0.00	0.70
Motor vehicles, bodies and trailers, and parts (3361MV)	0.69	0.00	0.29	0.00	0.33
Miscellaneous manufacturing (339)	0.66	0.00	0.22	0.00	0.50
Nonmetallic mineral products (327)	0.60	0.00	0.32	0.00	0.48
Computer and electronic products (334)	0.37	0.11	0.34	0.00	0.57
Chemical products (325)	0.36	0.16	0.14	0.00	0.33
Paper products (322)	0.31	0.42	0.23	0.00	0.25
Wholesale trade (42)	0.24	0.21	0.50	0.00	0.69
Plastics and rubber products (326)	0.23	0.41	0.20	0.00	0.36
Primary metals (331)	0.21	0.37	0.25	0.00	0.62
Mining, except oil and gas (212)	0.19	0.51	0.09	0.00	0.39
Fabricated metal products (332)	0.13	0.26	0.06	0.00	0.28
Apparel and leather and allied products (315AL)	0.12	0.79	0.04	0.02	0.10
Furniture and related products (337)	0.07	0.84	0.09	0.00	0.28
Management of companies and enterprises (55)	0.03	0.42	0.84	0.00	1.00
Textile mills and textile product mills (313TT)	-0.03	0.93	0.17	0.00	0.21
Petroleum and coal products (324)	-0.05	0.89	0.26	0.00	0.45
Publishing industries, except internet (includes software) (511)	-0.05	0.52	0.48	0.00	0.95
Other transportation equipment (3364OT)	-0.07	0.82	0.04	0.00	0.35
Food and beverage and tobacco products (311FT)	-0.08	0.75	0.42	0.00	0.64
Machinery (333)	-0.17	0.54	0.38	0.00	0.39
Printing and related support activities (323)	-0.49	0.11	0.07	0.00	0.26
Electrical equipment, appliances, and components (335)	-0.76	0.02	0.11	0.00	0.34

Shipment Value Weighted Network Matrix

Table 3 contains the coefficient estimates, in a similar structure to Table 1 and 2, with the exception that the Table 3 results utilize the Shipment Value Weighted Network Matrix to represent trade networks. In this case, the network autoregressive model uses a row-standardized flow matrices, $W_{N2,i}$, that is not limited to binary identifiers of trade relationships. We find positive, large, and statistically significant coefficient estimates for the network effect, ranging between 0.42 to 0.82, for Miscellaneous manufacturing, Motor vehicles, bodies and trailers, and parts, Nonmetallic mineral products, Wood products, and Textile mills and textile product mills, Plastics and rubber products. This means that for these industries, there is evidence that labor compensation correlates positively within trade networks, meaning that compensation tends to converge. There are also positive, yet insignificant, estimates for this network coefficient for an additional 11 industries, as well as insignificant negative estimates for an additional six industries.

The β_2 coefficient representing the compensation rate changes with respect to productivity are positive and significant for all industries. These estimates range from 0.03 to 0.84 across industries and show that as productivity increases, compensation also increases but not at a one-to-one ratio and instead gains in productivity are associated with smaller percentage increases in wages, which depicts the compensation-productivity gap persists even when allowing the matrix to vary based upon trade network relationships and opposed to spatial relationship.

TABLE 3: Network Autoregressive Model using a Shipment Value Weighted Network Matrix

Industries					
Miscellaneous manufacturing (339)	0.82	0.00	0.21	0.00	0.52
Motor vehicles, bodies and trailers, and parts (3361MV)	0.74	0.00	0.27	0.00	0.35
Nonmetallic mineral products (327)	0.73	0.00	0.33	0.00	0.51
Wood products (321)	0.62	0.00	0.37	0.00	0.68
Textile mills and textile product mills (313TT)	0.42	0.07	0.17	0.00	0.21
Plastics and rubber products (326)	0.42	0.03	0.21	0.00	0.33
Apparel and leather and allied products (315AL)	0.38	0.16	0.03	0.02	0.11
Paper products (322)	0.37	0.15	0.23	0.00	0.25
Chemical products (325)	0.33	0.12	0.14	0.00	0.32
Wholesale trade (42)	0.32	0.12	0.50	0.00	0.68
Computer and electronic products (334)	0.30	0.15	0.34	0.00	0.54
Petroleum and coal products (324)	0.27	0.35	0.27	0.00	0.44
Electrical equipment, appliances, and components (335)	0.26	0.37	0.12	0.00	0.26
Machinery (333)	0.22	0.40	0.37	0.00	0.41
Fabricated metal products (332)	0.18	0.57	0.06	0.00	0.25
Primary metals (331)	0.11	0.58	0.26	0.00	0.61
Other transportation equipment (3364OT)	0.03	0.87	0.04	0.00	0.35
Furniture and related products (337)	-0.01	0.98	0.09	0.00	0.28
Management of companies and enterprises (55)	-0.01	0.54	0.84	0.00	1.00
Publishing industries, except internet (includes software) (511)	-0.02	0.74	0.48	0.00	0.95
Mining, except oil and gas (212)	-0.17	0.28	0.09	0.00	0.37
Printing and related support activities (323)	-0.21	0.46	0.07	0.00	0.24
Food and beverage and tobacco products (311FT)	-0.29	0.28	0.42	0.00	0.63

Scale-Free Random Trade Network Matrix

The final empirical approach deviates from the spatial and trade network matrices of the first three models, and instead tests for preferential attachment on trade networks using a trade network weight matrix generated from 1000 counterfactual simulation distributions. Likewise, Table 4 presents the distribution parameters for each industry, x_{min} and γ, from which the scale-free random network is generated. The $p-$value shows the probability to fail to reject the null hypothesis that trade networks follow a power law distribution, as discussed in Appendix A. We find that all but three industries (Management of companies and enterprises, Printing and related support activities, and Plastics and rubber product industries) show significant evidence of following a power distribution, which is indicative of existing preferential attachment between some regions within the trade networks of industries.

Based upon these distribution parameter estimates for each industry; we generate a

TABLE 4: Test of Preferential Attachment on Trade Networks

Industries	γ	x_{min}	p-value
Wholesale trade (42)	3.49	7	0.85
Management of companies and enterprises (55)	2.42	8	0.01
Mining, except oil and gas (212)	3.03	12	0.45
Wood products (321)	2.22	4	0.23
Paper products (322)	2.35	5	0.37
Printing and related support activities (323)	2.03	4	0.00
Petroleum and coal products (324)	6.94	28	0.61
Chemical products (325)	6	17	0.48
Plastics and rubber products (326)	2.48	4	0.06
Nonmetallic mineral products (327)	2.25	4	0.26
Primary metals (331)	8.77	14	0.37
Fabricated metal products (332)	5.84	19	0.69
Machinery (333)	6.25	21	0.72
Computer and electronic products (334)	2.41	5	0.36
Electrical equipment, appliances, and components (335)	4.59	7	0.37
Furniture and related products (337)	2.27	2	0.51
Miscellaneous manufacturing (339)	6.04	17	0.31
Publishing industries, except internet (includes software) (511)	9.51	16	0.93
Food and beverage and tobacco products (311FT)	7.37	24	0.38
Textile mills and textile product mills (313TT)	4.53	12	0.50
Apparel and leather and allied products (315AL)	3.31	6	0.44
Motor vehicles, bodies and trailers, and parts (3361MV)	2.38	3	0.69
Other transportation equipment (3364OT)	2.38	3	0.69

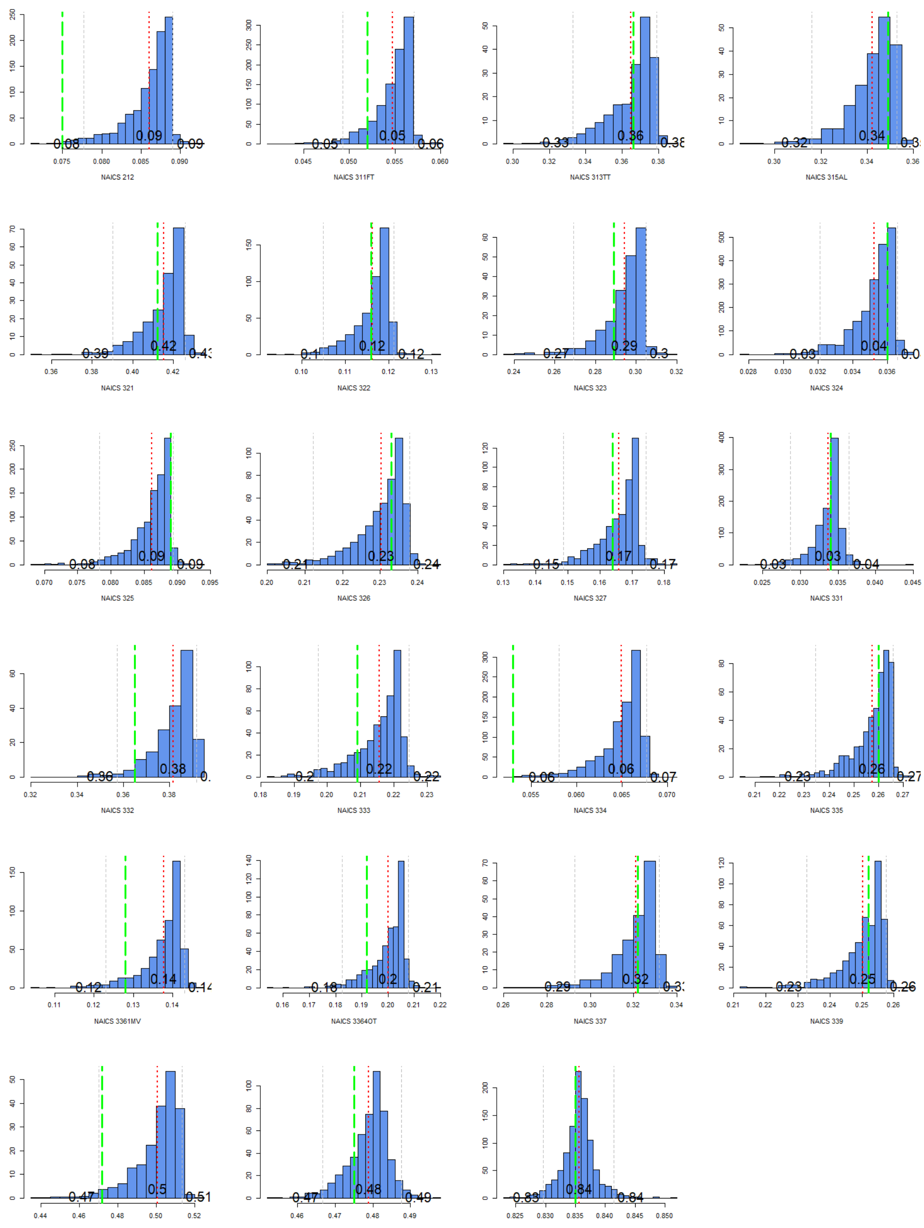
scale-free random network from 1000 generations of the network autoregressive model for each industry. These results represent the distribution of estimates for the rate of change of compensation with respect to productivity within the synthetically generated network matrix. Figure 3 depicts the distribution of estimates of the rate of change of compensation with respect to productivity, which can be thought of as the range of the productivity-compensation gap within industries, if the trade network was scale-free random. If actual estimates of the rate of change of compensation with respect to productivity do not lie within the 95% confidence intervals of the estimated synthetic rate of change of compensation, then the productivity-compensation gap is strategic and not plausibly random.

6 CONCLUSION

The purpose of this study is to investigate if the productivity-compensation gap can be partially explained through labor compensation spillovers over physical distance as well as through shared trade networks, and to in turn present estimates of the effects of these spatial and network relationships for comparison and discussion. To determine spatial and network relationships, we utilized IO-Snap software to collect compensation, employment, and industrial output for a variety of industry sectors and integrating this information with inter-regional U.S. trade networks data from the 2012 Commodity Flow Survey Public Use Microdata File. This paper successfully shows the existence of labor compensation rates being correlated both spatially and within inter-regional trade networks for many industries.

This study develops and tests four distinct empirical hypothesis of labor compensation

Figure 3: Counterfactual Distributions of Productivity-Compensation Gap by Industry



relations. First, a spatial contiguity matrix is used to run a standard spatial autoregressive model to determine how compensation rates between regions are affected by physical proximity. We find that Chemical products, Primary metals, and Wood product industries are highly, positively, spatially correlated (estimates between 0.30-0.42), are weakly positively correlated for 17 other industries, and are significantly negatively correlated for

Management of companies and productivity.

After conducting this spatial autoregression, we shift the focus on trade network relationships as the cost of transportation has decreased over time and technology has improved making physical proximity a less important determinant of trade relationships. First, we develop a network autoregressive model using the row-standardizing binary contiguity network matrix. Second, we integrate volumes of trade and implied network dependence and develop a network autoregressive model using a row-standardizing flow contiguity network matrix. The results from both empirical results are similar finding highly significant and large estimates for the correlation of labor compensation among trade networks for Wood products, Motor vehicles, bodies and trailers, and parts, Nonmetallic mineral products, and Miscellaneous manufacturing industries. The responsiveness of compensation rates with respect to labor productivity are all positive and significant, suggesting that compensation changes with respect to productivity, but since these values are less than one, it also implies that compensation increases at only a percentage value compared to productivity increases, which depicts the compensation-productivity gap.

Finally, we develop a test of preferential attachment of trade networks using a scale-free random trade network matrix developed from synthetically performing 1000 counterfactual empirical assessments of the industry parameters. This final strategy provides a range of coefficient estimates to determine if productivity-compensation gaps are random or strategic for each industry within our sample.

Overall, using a variety of empirical strategies, we present an argument for the importcon of considering trade networks when discussing the productivity-compensation gap. Compensation rates within many industries are highly correlated across these trade networks and discussions of this gap are inappropriate when limited to discussions of spatial contiguity over physical proximity. This is crucially important when developing labor practices and policies as compensation is not only driven over physical space, but between trading partners. This study is limited to a small selection of industries and further research must be conducted to include a wider variety of industries over longer time horizons before causal relationships can be determined. It is also important to note that our discussion of trade treats these relationships as symmetric, when trade is often unidirectional. Further study is needed to determine the extent that unilateral trade affects productivity and compensation. This study also abstracted away from potential spillovers in the compensation rates of related fields that could exacerbate the productivity-compensation gap.

REFERENCES

Abbasi, A., Hossain, L., and Leydesdorff, L. (2012). Betweenness centrality as a driver of preferential attachment in the evolution of research collaboration networks. *Journal of Informetrics*, 6(3):403–412.

Albert, R., Jeong, H., and Barabási, A. L. (1999). Diameter of the world-wide web. *Nature*, 401(6749):130–131.

Anselin, L. (1988). *Spatial Econometrics: Methods and Models*, volume 4.

Barabási, A. L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439):509–512.

Barro, R. J. and Sala-i Martin, X. (1992). Convergence. *Journal of Political Economy*, 100(2):223–251.

Bohorquez, J. C., Gourley, S., Dixon, A. R., Spagat, M., and Johnson, N. F. (2009). Common ecology quantifies human insurgency. *Nature*, 462(7275):911–914.

Borgs, C., Chayes, J., Daskalakis, C., and Roch, S. (2007). First to market is not everything: an analysis of preferential attachment with fitness. *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, pages 135–144.

Bosworth, B., Perry, G. L., and Shapiro, M. D. (1994). Productivity and Real Wages : Is There a Puzzle ? *Brookings Paper on Economic Activity*, 1994(1):317–344.

Brill, B. M., Holman, C., Morris, C., Raichoudhary, R., and Yosif, N. (2017). Understanding the labor productivity and compensation gap. Technical Report 6, U.S. Bureau of Labor Statistics.

Burridge, P. and Gordon, I. (1981). Unemployment in the British Metropolitan Labour Areas. *Oxford Economic Papers*, 33(2):274–297.

de Benedictis, L. and Tajoli, L. (2011). The World Trade Network. *World Economy*, 34(8):1417–1454.

Efron, B. and Tibshirani, R. (1993). *An Introduction to the Bootstrap*, volume 57. Chapman & Hall/CRC.

ESMPD, U. C. B. (2012). US Census Bureau Commodity Flow Survey - About the Survey.

Fleck, S., Glaser, J., and Sprague, S. (2011). The Compensation-Productivity Gap: A Visual Essay. *Monthly Labor Review*, (January):57–69.

Jeong, H., Néda, Z., and Barabási, A. L. (2003). Measuring preferential attachment in evolving networks. *Europhysics Letters*, 61(4):567–572.

Mallick, R. and Carayannis, E. G. (1994). Regional Economic Convergence in Mexico: An Analysis by Industry. *Growth and Change*, 25(3):325–334.

Manning, N. (1994). Earnings, Unemployment and Contiguity: Evidence from British Counties 1976–1992. *Scottish Journal of Political Economy*, 41(1):43–68.

McCombie, J. S. (1988). A Synoptic View of Regional Growth and Unemployment: I - The Neoclassical Theory. *Urban Studies*, 25(4):267–281.

Molho, I. (1982). Contiguity and Inter-Regional Migration Flows in Great Britain. *Scottish Journal of Political Economy*, 29(3):283–297.

Newman, M. E. (2001). Clustering and preferential attachment in growing networks. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, 64(2):4.

Newman, M. E. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, 46(5):323–351.

Plemmons, A. (2019). Occupational licensing effects on firm entry and employment. CGO Working Paper No. 2019.008.

Poncela, J., Gómez-Gardeñes, J., Floría, L. M., Sánchez, A., and Moreno, Y. (2008). Complex cooperative networks from evolutionary preferential attachment. *PLoS ONE*, 3(6).

RRI (2017). IO-SNAP.

Serrano, M. A. and Boguna, M. (2003). Topology of the World Trade Web.

Smith, D. M. (1975). NEOCLASSICAL GROWTH MODELS AND REGIONAL GROWTH IN THE U.S. *Journal of Regional Science*, 15(2):165–182.

Sprague, S. (2017). Below trend: the U.S. productivity slowdown since the Great Recession : Beyond the Numbers: U.S. Bureau of Labor Statistics. Technical Report 2, U.S. Bureau of Labor Statistics.

Stumpf, M. P. and Porter, M. A. (2012). Critical truths about power laws.

Tavernier, E. and Temel, T. (1997). National and regional analysis of convergence of real wages in the US agricultural sector.

Tinbergen, J. (1962). *Shaping the World Economy; Suggestions for an International Economic Policy*. Twentieth Century Fund, New York.

A COUNTERFACTUAL TRADE NETWORKS

Understanding the nature of network matrices is crucial for developing counterfactual trade network models. Within network theory, there is an important recorded phenomenon of an "80/20 rule" first penned by Italian economist Vilfredo Pareto, in that within networks seemingly "80% of all effects result from 20% of all causes". This relationship has been observed all manner of networks (Albert et al., 1999; Barabási and Albert, 1999), for example, the nature of connectivity between webpages in the form of hyperlinks (Albert et al., 1999), world trade data (de Benedictis and Tajoli, 2011; Serrano and Boguna, 2003) unique words in novels (Newman, 2005), and casualty numbers in armed conflicts (Bohorquez et al., 2009). The trade networks analyzed within this paper are similar to world trade networks, and in that most trades result from a small selection of regions.

The 80/20 rule is a heavy-tailed distribution, which is modeled with the power exponent of a log-log plot, commonly referred to as the power law (Stumpf and Porter, 2012). The existence of this power law within networks is also referred to preferential attachment theory (Abbasi et al., 2012; Borgs et al., 2007; Jeong et al., 2003; Newman, 2001; Poncela et al., 2008). The use of the power law within the counterfactual trade network accounts for the observation that within the trade network, a majority of first-order trade network nodes may be accounted for by a small number of regions.

A power law distribution exhibits the preferential attachment. The probability density function of a power law distribution is:

$$p\left(x\right) = \frac{\alpha - 1}{x_{\min}} \left(\frac{x}{x_{\min}}\right)^{-\alpha}$$

where, scaling parameter $\alpha > 1$ and $x_{\min} > 0$. The cumulative density function is:

$$P\left(X \leq x\right) = 1 - \left(\frac{x}{x_{\min}}\right)^{-\alpha+1}$$

moreover, the moments are:

$$E\left[X^m\right] = \int_{x_{\min}}^{\infty} x^m p\left(x\right) dx = \frac{\alpha - 1}{\alpha - 1 - m} x_{\min}^m$$

such that $1 < \alpha \leq 2$ all moments diverges, i.e., $E\left[X\right] = \infty$; $2 < \alpha \leq 3$, all second and higher-order moments diverge, i.e., $E\left[X^2\right] = \infty$ and $3 < \alpha \leq m + 1$ all m and higher-order moments diverges, i.e.,$E\left[X^m\right] = \infty$. The scaling parameter α is estimated with the maximum likelihood estimator (MLE) as:

$$\hat{\alpha} = 1 + n \left[\sum_{i=1}^{n} \ln \frac{x_i}{x_{\min}}\right]^{-1}$$

The estimation of the scaling parameter αmust is conditioned with the value of $x_{\min}$. However, as $x_{\min}$increases, the amount of data discarded also increases. So, some care is taken when choosing this parameter. A common approach is a visual inspection of the data on the log-log plot, but such technique is highly subjective and error-prone. Following Clauset, Shalizi, Newman (2009) the lower threshold can be identified using the Kolmogorov-Smirnov approach. This statistic is merely the maximum distance between the data and fitted model CDFs and given as:

$$D = \max_{x \geq x_{\min}} |S(x) - P(x)|$$

where $S(x)$ and $P(x)$ are the CDFs of the data and model respectively for $x \geq x_{\min}$. The estimate $x_{\min}$ is the value of $x_{\min}$ that minimizes the D. This is an entirely a general approach and can also be implemented in conjunction with other distribution.