Geography Dictates, But How? Topography, Spatial Concentration and Sectoral Diversification

Mohammad Tarequl Hasan Chowdhury and Muhammad Habibur Rahman and Mehmet Ali Ulubasoglu

Chittagong University, Monash University Malaysia, Deakin University

31 May 2018

GEOGRAPHY DICTATES, BUT HOW?

TOPOGRAPHY, SPATIAL CONCENTRATION, AND SECTORAL DIVERSIFICATION

Mohammad Tarequl H. Chowdhury[a]

Muhammad Habibur Rahman[b]

Mehmet Ali Ulubaşoğlu[c]

[a]Chittagong University, Bangladesh

[b]Monash University, Malaysia

[c]Deakin University, Australia

Abstract

This study investigates the ways in which terrain ruggedness affects sectoral diversification. A cross-country analysis using data from 142 countries over the period 1970–2007 documents an inverted U-shaped link between terrain ruggedness and sectoral diversification, which mainly works through the extensive margin of diversification. A within-country analysis based on United States (US) state-level data over the period 1997–2011 confirms this non-monotonic relationship. The within-country analysis further reveals that an important mechanism through which terrain ruggedness affects sectoral diversification is the spatial concentration of economic activity, as measured by the concentration of satellite-based night lights.

Keywords: sectoral diversification, spatial concentration, extensive margin, intensive margin, terrain ruggedness.

JEL Codes: R12, O51, O57

1 Introduction

Imbs and Wacziarg's (2003) deservedly influential study demonstrates economies with low income levels experience rising levels of sectoral diversification; after reaching a certain level of development, they tend to concentrate (see also Koren and Tenreyro (2007)). While this curious finding has major implications for key stages of sectoral diversification, most countries face exogenous hurdles in their sectoral development trajectories—chiefly, non-economic barriers. This study examines the manner in which terrain ruggedness, a time-invariant topographic factor that constitutes a significant obstacle to economic activity, affects the course of sectoral diversification.

That rugged topography would be a major inhibiting factor for sectoral development is hardly a novel proposition. However, the ways in which this effect arises remain opaque, and require detailed examination given the widely held arguments and policy prescriptions around the world that emphasize the critical role of spreading risks across sectors by diversifying real economies. How could economies facing significant natural barriers adopt such prescribed policies? We investigate the topography–sectoral diversification relationship in both cross-country and single-country settings. The main focus of our cross-country analysis, which is based on an annual panel of 142 countries over the period 1970–2007, is to unpack whether topography affects sectoral diversification through the intensive or extensive margins of the latter. That is, does topography affect diversification by influencing the sizes of *existing* sectors (i.e., intensive margin of sectoral diversification), or by affecting the introduction of *newer* sectors and phasing out the *older* sectors (i.e., extensive margin of sectoral diversification)? Despite painting a useful global picture, the cross-country analysis nonetheless has limitations in terms of explaining the mechanism through which terrain ruggedness might impede or boost sectoral diversification. The key concern here is that a heterogeneous cross-country setting may not enable a comparable platform to tease out the mechanism of effect reliably. To address that gap, we extend our analysis to a within-country context using United States (US) state-level data for the period 1997–2011. The mechanism we conjecture is that rugged topography affects the spatial distribution of economic activity, which in turn may affect the levels of sectoral diversification. In particular, rugged topography may restrict geographic clusters from spatially (i.e., the extensive margin of spatial concentration), given that population growth requires residential and entrepreneurial areas to extend over space. Terrain ruggedness may also force economies to operate within certain spatial clusters characterized by relatively uneven nodal shares (i.e., intensive margin of spatial concentration). While counter-arguments in both veins are possible, spatial concentration induced by topographic barriers, according to our hypothesis, is likely to lead to sectoral concentration caused by the uneven relative sizes of existing sectors, or the prevention of new and/or phasing out of older sectors. Our within-country analysis has its own limitation of not enabling measurement of the intensive and extensive margins of sectoral diversification, because no manufacturing sector is introduced or phased out in the US in our sample period. Consequently, the cross-country and within-country analyses in this paper are complementary, with one addressing the other's shortcoming.

To investigate the effect of terrain ruggedness on sectoral diversification, an accurate quantification of topographic disturbances is vital. Crude indicators reflecting larger scale irregularities, such as the percentage of mountains in a country's surface area, are far from able to capture the proximate conditions that affect the sectoral distribution of economic activity. Thus, we adopt the terrain-ruggedness index developed by Nunn and Puga (2012), a precise metric that quantifies topographic irregularities in 30 arcseconds (926 meters).

To measure sectoral diversification, we exploit the Theil index, particularly its decomposability property (Cadot et al. 2011). The Theil index is an additive measure that usefully decomposes the extensive and intensive margins of the inequalities of sectors' sizes. Consequently, in a novel step, we decompose the sectoral diversification into its extensive and intensive margins (i.e., between- and within-components, respectively). For cross-country analysis, we use two-digit manufacturing data for 23 sectors of 142 countries obtained from the INDSTAT2 (2012) International Standard Industrial Classification (ISIC) Rev.3 database, whose employment and value-added data we use to compute the sectoral labor and value-added shares, respectively. Analogously, the Theil index of US states' sectoral diversification is computed using two-digit data on value-added shares of 18 manufacturing sectors classified according to the North American Industry Classification System (NAICS) for 50 US states for the period 1997–2011.

To measure the spatial concentration of economic activity for the US states, we again take advantage of the Theil index, but apply it to 30 arcsecond-gridded satellite night lights data (see Henderson, Storeygard and Weil (2012)). In contrast to most of the earlier studies that have used the night lights data only as a proxy for income, our approach adopts a more literal interpretation of the data and uses them as an indicator of the spatial concentration of economic activity. Further, we compute the spatial extensive and intensive margins of the distribution of night lights. To the best of our knowledge, no other study has hitherto considered the spatial extensive and intensive margins of the satellite night lights data.

Our cross-country results show that terrain ruggedness is significantly associated with sectoral diversification in a country, that this effect works through the sectoral diversification's extensive margin, and that the effect is inverse U-shaped. The effect's non-monotonicity may arise given the threshold effects, owing to the skewed distribution of surface irregularities across countries, and suggests that ruggedness is associated with higher (lower) levels of sectoral diversification (concentration) up to a certain level of ruggedness, after which it is associated with lower (higher) sectoral diversification. The within-country evidence also supports the inverted U-shaped effect of terrain ruggedness on sectoral diversification. In addition, it shows that rugged terrain increases spatial concentration, and this in turn reduces sectoral diversification. Further, there is some evidence that an extensive margin of spatial concentration (i.e., inability to extend the night lights into previously dark areas) drives this spatial mechanism. The intuition behind this finding is that if geographic clusters in a state cannot spatially extend, then they are likely to operate with a handful of sectors. The corollary is that if geographic decentralization occurs within a state (i.e., new night lights are introduced), then it is likely that new sectors have also been introduced. Although alternative mechanisms may not be ruled out, our ordinary least-squares (OLS) results are robust to Hausman–Taylor (HT) and instrumental variable (IV) estimations, which address the endogeneity in key control variables, such as income and trade openness.

This study thus makes two novel contributions to the literature by addressing topography's important role in sectoral diversification. First, in our cross-country analysis, we investigate not only the effect of terrain ruggedness on overall sectoral diversification, but also on the latter's extensive and intensive margins. Second, to investigate the mechanism of effect, we exploit the satellite night lights data in a single-country analysis to measure spatial concentration, which is predicated on a literal interpretation of what the night lights data actually display. Moreover, we exploit the intensive and extensive margins of the night lights data to delve deeper into the geographic concentration mechanism. Overall, our study stands at the crossroad of two strands of literature: one examining the determinants of sectoral diversification, and the other focusing on geography's effect on the real economy.

The remainder of the paper is organized as follows: section 2 provides a brief overview of topography and diversification relationship, section 3 explains the data used, and section 4 describes the estimation framework. Section 5 presents the results and section 6 concludes the study.

2 Topography, Spatial Concentration, and Sectoral Diversification: A Theoretical Discussion

In this sub-section, first we provide theoretical arguments as to how sectoral diversification and ruggedness may be related. Then, we explain how spatial concentration can work as a mechanism in this relationship.

2.1 Sectoral Diversification, Terrain Ruggedness, and the Extensive and Intensive Margins of Sectoral Economic Activity

Diversification has been at the center of a long-standing debate in economics. While several theories, such as the Ricardian and Heckscher-Ohlin theories of international trade, promote specialization, a number of others emphasize the importance of diversification. For example, diversification plays a critical role in spreading risks across sectors in economies facing external shocks and severe vulnerability (Imbs and Wacziarg 2003). Similarly, the International Monetary Fund (2014) observes that sectoral diversification has a "growth payoff" as well as a "stability payoff." Therefore, it is no wonder that international institutions such as the World Bank run field programs in developing countries to assist them with their diversification objectives.

The effect of physical geography on economic variables is widely acknowledged in the economic geography literature (see Malik and Temple (2009); Gallup, Sachs & Mellinger (1999)). The role of terrain ruggedness is well documented, not only in the spatial concentration of economic activity (Ramcharan 2009), but also in such diverse factors as exports (Radelet and Sachs, 1999), economic development (Mellinger, Sachs, and Gallup (2000)) civil conflict (Shaver, Carter and Shawa (2015)), social trust (Khalifa, (2016)), and state capacity (Jimenez-Ayora and Ulubasoglu (2015)). Nunn and Puga (2012) show that ruggedness was a 'blessing' for African countries in that it protected slave trade, and thus had an indirect bearing on the economic development of some African countries.

A priori, the manner in which physio-geography influences sectoral diversification is ambiguous. For example, settlements lying in different altitudes within a short distance are likely to face non-negligible costs of production. In addition, rugged terrain imposes serious constraints on delivering infrastructure and basic public services to different communities.[1] Transaction costs due to rugged terrain may thus impede private sector development and limit the economy from venturing into different sectors, inducing sectoral concentration. By contrast, some levels of ruggedness may provide advantages over a smoother terrain (e.g., by leading to higher productivity in certain industrial crops), or spatial segregation may motivate cooperation up to some level of topographic difficulty (e.g., building roads, railways, and other infrastructure for common use), which may thus lead to sectoral diversification.

It is important to note that negative consequences of rugged topography can be alleviated, if not reversed, through engineering. One example of a highly successful rugged country that won over its topography is Switzerland. Though not all rugged-terrain countries are as successful as Switzerland, it is reasonable to expect that an average country would attempt to overcome certain forms of ruggedness. Nonetheless, an overly rugged terrain would discourage investments into such infrastructure because of their non-negligible costs and uncertain returns—Nepal is a good example. These arguments suggest that ruggedness may initially assist sectoral diversification such that its adverse consequences may be observed beyond a certain point, but then sectoral concentration may win over. This would mean an inverted U-shaped relationship between ruggedness and sectoral diversification.

[1] Ulubasoglu and Cardak (2007) find that landlocked countries, generally mountainous, exhibit higher inequality in rural and urban schooling because of the difficulties associated with public service delivery to rural areas.

Whether and how the intensive and extensive margins of sectoral diversification may be affected by terrain irregularities constitutes an open, ambiguous, and critical empirical question. In terms of the extensive margin, rugged topography may prevent new sectors from being introduced or older ones from being phased out, and may thus result in a concentrated sectoral structure. A counter-argument is that terrain ruggedness may lead to pockets of clusters that host different sectors, which may underlie diversified sectoral production. Turning to the intensive margin, how sectoral shares in a given topographic cluster would be shaped is ambiguous, but, taking a cue from the extensive margin discussion, it is likely that constrained intersectoral linkages caused by having to operate in a narrow space may give rise to unequal size distribution for existing sectors, leading to sectoral concentration.

2.2. Spatial Concentration as a Mechanism

Critically, the above discussion suggests that spatial (i.e., geographic) concentration within an observational unit (i.e., state or country) is central to understanding the relationship between terrain ruggedness and sectoral diversification. To this end, Ramcharan (2009, 2010) provides very useful insights. Ramcharan (2010) reviews the road construction literature, and documents that the terrain grade variation—that is, the rise and fall of the surface area—as well as soil characteristics can exponentially affect the cost of building roadways and rail lines. This may result in spatial concentration and reduced sectoral diversification opportunities, given the increased cost of transporting goods across space. Conversely, ruggedness-induced spatial concentration may increase sectoral diversification through an agglomeration effect. Ruggedness may be responsible for concentrated economic activities in a narrow area, and thus create opportunities for a larger home market, boosting returns to scale and leading to higher productivity levels. The expanded domestic demand generated may facilitate scope for sectoral diversification (Ramcharan (2009)).

As indicated above, we investigate the spatial concentration mechanism as measured by US night lights data. To link the spatial distribution of economic activity with sectoral diversification, we assume the following: if sectoral diversification in a state takes place alongside spatial diversification, then it is likely that diversification of economic activity in space has contributed to the diversification of economic activity in sectors (or vice versa). This effect can arise through both the *spatial extensive* or *spatial intensive* margins of economic activity. Considering the spatial extensive margin, if new areas are lit up, then either new sectors are introduced/older sectors are phased out, or the relative shares of existing sectors have changed in a manner resulting in sectoral diversification. Of these possibilities, we rule out the introduction of new or phasing out of older sectors in the US, because we do not observe such development in the US manufacturing data for 1997–2011. We thus assume the only remaining possibility: that the introduction of night lights into previously dark areas has changed the relative shares of existing sectors in a manner that has led to more even sectoral shares. Considering the spatial intensive margin, we assume that if the relative strengths of existing night lights across different nodes have changed in a way that has reduced spatial concentration, then it is likely that the relative shares of existing sectors have become more even, such that some sectoral diversification has occurred. Taken together, both of our assumptions mean that changes in the extensive and intensive margins of spatial concentration are likely to affect sectoral structure by changing the relative shares of existing sectors.

The downside of our approach is that we are unable to geocode the new sectors or those existing sectors in a manner that fully connects sectoral diversification with spatial diversification, hence the above assumptions. Nonetheless, the effect of ruggedness on spatial concentration, and in turn the effect of spatial concentration (through both spatial extensive and spatial intensive margins) on sectoral diversification based on these assumptions, are exactly what we test in our empirical analysis.

3 Data and Descriptive Statistics

In this section, we discuss the construction of the main study variables. Appendix Tables A1 and A2 present a description of all variables used in this study.

3.1 Terrain Ruggedness

Our terrain-ruggedness measure, originally constructed by Riley et al. (1999) and later improved by Nunn and Puga (2012), precisely quantifies topographic irregularities in a country's land area.[2] A higher value of the terrain-ruggedness index implies a higher surface roughness or irregularity. Figure A1 shows the level of ruggedness across different countries of the world. In our sample, Dominica has the lowest ruggedness, with a value of 0.003, while Bhutan has the highest, with 6.74, and Barbados has the median ruggedness value of 0.963.

Turning to the within-country analysis, we construct the terrain-ruggedness index for the 51 US states/territories using the same procedure as Nunn and Puga (2012) (see Figure A2). Delaware has the least terrain irregularities, with a value of 0.06, while the state of Washington has the most, with a value of 2.6, and Arkansas has the median ruggedness value of 0.51.

3.2 Measure of Sectoral Diversification

In a recent innovative contribution, Cadot et al. (2011, pp. 594–596) take advantage of the Theil index's decomposability and compute the intensive and extensive margins of export diversification. In particular, they partition export diversification additively into within-group and between-group (product) components, referring, respectively, to diversification occurring within existing products, and diversification arising because of new products being introduced or older products being phased out.[3]

[2] Based on digital elevation data taken from GTOPO30, a global dataset led by the US Geological Survey Center, the terrain-ruggedness index accurately measures the topographic characteristics of a country's land surface. Elevation observations in GTOPO30 are gridded at 30 arcseconds (926 meters) across the entire surface of the Earth (Nunn & Puga 2012). These grid cells are subsequently averaged to compute the average ruggedness of the counry's lands that are free of water.

[3] They find that most of the variation between export diversification and income per capita is because of intensive trade margins (i.e., existing export products), but the non-monotonic relationship between the two is driven by extensive margins (new products, or new markets for existing products).

Following Cadot et al. (2011), we decompose the Theil index of sectoral *production* diversification in a country into its within- and between-components. Sectoral production diversification caused by changes in the within-component occurs because of sectoral employment or value-added shares within existing sectors becoming more even (which might occur as a result of reallocation and/or movement of labor and capital across the existing sectors). Sectoral diversification caused by changes in the between-component occurs because of new sectors being introduced or older sectors being phased out (which might occur as a result of switching the exising labor and capital to new sectors). A higher value of the overall Theil index for sectoral diversification indicates a lower level of sectoral diversification, or a higher level of concentration. In this setting, a higher between-component (extensive margin) of the Theil index means a higher concentration caused by the inability to introduce new sectors or phase out older sectors, while a higher within-component (intensive margin) indicates a higher concentration caused by some sectors dominating the rest in terms of size.

To calculate the Theil index of sectoral diversification at the cross-country level, we use a panel dataset for the period 1970–2007. This period is mainly dictated by the availability of the sectoral size data. We compute the index using the INDSTAT2 (2012) ISIC Rev.3 database, which provides manufacturing data for 23 sectors at the two-digit level of disaggregation. Our sectoral size measures are employment and value-added shares. Figure A3 presents the employment share-based sectoral diversification status of the countries averaged over the sample period.

Following Theil (1972), we compute a sectoral diversification index as follows:

$$S = \frac{1}{n}\sum_{i=1}^{n} \frac{Sectoral\ Size_i}{Average\ Sectoral\ Size} \ln\left(\frac{Sectoral\ Size_i}{Average\ Sectoral\ Size}\right)$$

where *i* stands for sectoral size, and *n* denotes the total number of sectors (omitting country and time subscripts for brevity).

Like Cadot et al. (2011), our sectoral diversification index can be calculated for groups of sectors and decomposed additively into within-groups and between-groups components. Mathematically, let us consider some partition of that total number of potential sectors (of a given country in a particular year) into $J+1$ groups denoted G_j, where $j = 0, \ldots, J$. Let n_j be the number of sectors in group j, and let S_j stand for the overall sectoral diversification index for group j. The between-groups component of sectoral diversification index is defined as follows:

$$S^B = \sum_{j=0}^{J} \frac{n_j \times Average\ Sectoral\ Size_j}{n \times Average\ Sectoral\ Size} \ln\left(\frac{Average\ Sectoral\ Size_j}{Average\ Sectoral\ Size}\right)$$

And its within-groups component is measured by the following:

$$S^W = \sum_{j=0}^{J} \frac{n_j \times Average\ Sectoral\ Size_j}{n \times Average\ Sectoral\ Size} S_j$$

Given the additivity property of Theil's index, S would be equal to S^B+S^W (for technical details, see Cadot et al., 2011).

To calculate sectoral diversification in the US, we use the state-level manufacturing data for the period 1997–2011. We now have 18 sectors (instead of 23 in the cross-country analysis), which are classified according to the NAICS. The NAICS classification is not directly comparable with ISIC data because of its different classification conventions. The measure of sectoral size in our within-country analysis is value-added. As indicated above, sufficient variation does not exist in the between-component of the US manufacturing value-added data, so we calculate only the overall Theil index for the US. Figure A4 presents the sectoral diversification status of the US states in 2011. On average, across the sample period, Alabama has the highest sectoral diversification, while Alaska has the lowest.

3.3 Measure of Spatial Concentration

Turning to spatial concentration, we compute the (spatial) Theil index across the US states using satellite night lights data for the period 1995–2013. Our unique contribution to the literature that uses satellite night lights data is to compute the intensive and extensive margins of night lights for the US states over the period 1997–2011[4]. The within-component (intensive margin) of the index would identify the changes in the lights' density within a localized geographic area, and the between-component (extensive margin) would capture the geographic areas with lights, or with no lights before (see below). The night lights raw data come with six-bit digital number S for every 30-arcsecond grid of the US; which are sourced from the National Oceanic and Atmospheric Administration's National Geophysical Data Center. Using this grid-level dataset, we obtain the Theil-based spatial concentration night lights index using the following formula:

$$L = \frac{1}{n}\sum_{i=1}^{n} \frac{Night\ Light_i}{Average\ Night\ Light} \ln\left(\frac{Night\ Light_i}{Average\ Night\ Light}\right)$$

in which *i* is the index for geographic nodal point at 30 arcseconds and *n* denotes the total number of nodes in a given state. In addition, using a similar analogy as sectoral diversification, we compute the within- and between-groups components of the spatial concentration index with the following formulas:

$$L^B = \sum_{j=0}^{J} \frac{n_j \times Average\ Night\ Light_j}{n \times Average\ Night\ Light} \ln\left(\frac{Average\ Night\ Light_j}{Average\ Night\ Light}\right)$$

$$L^W = \sum_{j=0}^{J} \frac{n_j \times Average\ Night\ Light_j}{n \times Average\ Night\ Light} S_j$$

where L^B+L^W is equal to L. A higher overall spatial concentration index points to a more concentrated geographic structure. Along these lines, a higher between-component of this index means greater geographic concentration caused by the inability to extend spatially and to light up new areas, while a higher within-component means a higher concentration caused by some nodes dominating the geographic (i.e., night lights) distribution within a spatial unit in terms of light density (*vis-à-vis* relatively equal strength of lights in that cluster).

[4] The computation of the between-component results in the loss of 1995–1996 and 2012–2013.

Figure A5 illustrates the between-group (extensive margin) and within-group (intensive margin) components of spatial concentration, displaying 30-by-30 arcsecond cells, where each cell is centered on a point from the night lights grid. The solid black circles represent the existing night lights, and the grey circles represent new night lights introduced in a previously dark cell. The empty cells show the absence of night lights. The size of a circle portrays the density of night lights in a given cell. Compared with the baseline case in 5.1, a simple exchange of the circles across two cells in 5.2 leads to no change in spatial concentration. However, with the previously smaller black circle becoming as large as the other black circle, 5.3 displays a reduction in spatial concentration caused by a change in the within-component. Compared with 5.1, 5.4 displays another case of no change in concentration, where the larger black circle disappears and simply shifts to another cell (with the grey color). Yet 5.5 represents a decrease in spatial concentration caused by the introduction of a new (grey) night light in a third cell, denoting geographic diversification at the spatial extensive margin (between-component); 5.6 represents a further reduction in spatial concentration, with the previously smaller black circle becoming larger, thus suggesting geographic diversification in the spatial intensive margin (within-component) compared with 5.5.

3.4 Descriptive Statistics

We present the descriptive statistics of our cross-country panel data in appendix Table A3. The Theil index statistics show that variations in sectoral diversification arise from both within- and between-components. Figure 1 displays the relationship between sectoral diversification (inversely) measured by the Theil index and terrain ruggedness.[5] The relationship is evidently non-monotonic—that is, ruggedness initially helps diversification, but at higher levels of ruggedness, countries tend to concentrate. This evidence supports the argument that the costs

[5] On average, across our sample, Slovak Republic, Canada, and the US have the highest sectoral diversification according to the employment-based metric, while Gambia, the Dominican Republic, and Rwanda have the lowest. For the value-added-based Theil index, the first three countries with the highest levels of sectoral diversification are France, the Czech Republic, and Canada, while the last three countries are Gambia, Rwanda, and Botswana.

associated with developing railways and roads and maintaining transport networks grows exponentially with terrain grade variation, as pointed out by Ramcharan (2010).

Appendix Figure A6 depicts the spatial distribution of economic activity across the US states based on 2011 satellite night lights data. According to our metric, the District of Columbia (DC) has the lowest spatial concentration of economic activity, given that almost all its surface is lit up, while Alaska has the highest concentration, given a handful of nodes with night light.[6] New Jersey follows DC in terms of (the lowest) spatial concentration.

Figure 2, constructed following Cadot et al. 2011, provides striking insights into the within- and between-components of nights lights as a function of average night lights for a given state. The figure demonstrates that as the average night light intensity increases, overall spatial inequality decreases, particularly up to the average night light level of 24. Importantly, the reduction is driven by the between-component up to the level of 15, meaning that new areas are lit up such that overall spatial concentration goes down. In our dataset, Indiana, Ohio, and North Dakota witnessed the highest drops in overall spatial concentration from 1997 to 2011. North Dakota is an especially instructive case—with an average night light value of 1.1 during 1997–2009, North Dakota attained the night light level of 5.1 in 2013; the between-component of its Theil index dropped from 1.02 in 1997 to 0.81 in 2011. North Dakota experienced reduced spatial concentration because of recently developed oil and gas fields driven by the groundbreaking extraction technology known as hydraulic fracturing, or 'fracking'; see Figure A7, which marks a patch of night lights that overlayed with the newly opened oil field.

Returning to Figure 2, between the average night light levels of 15 and 24, the decrease in spatial inequality is mainly driven by the within-component, suggesting that the light densities among the existing lit nodes become relatively more equal. In this range of night lights, Delaware is among the top three (along with Ohio and Indiana) that experienced

[6] In our analysis, we drop DC given its very small size; it is also an outlier, with an index almost equal to zero.

intensive margin-driven spatial de-concentration. Its overall Theil index dropped from 0.60 to 0.42 from 1997 to 2011, which is almost wholly driven by the within-component. Such within-component domination of a spatial concentration decrease is depicted in Figure A8, where we observe no subtle change in the pattern of night lights in Delaware between 2012 and 2016, but some of the existing locations become more illuminated. This could be caused by the expansion of already-established corporate business in urban areas, where the financial services sector grew almost 21% between 2002 and 2012 (The News Journal, 2013).

In Figure 2, beyond the average night light level of 24, overall spatial concentration increases again, first by unequal relative shares of night light density among the extant lit nodes, but then chiefly because of the between-component. Expressed another way, some active night light nodes are darkened as the average night lights level goes beyond 30 (some inequality is also contributed by existing nodes that become more unequal). New Jersey serves as an ideal example in this context: it sparks the highest level of average night lights—32.11 in 1997—among all states (except DC). However, some parts of New Jersey have lost night lights over time because of economic activities in the state shifting to more vibrant areas. Figure A9 compares the night lights in New Jersey between 2012 and 2016—evidently, the night lights in Old Bridge Township of New Jersey (marked in red circles) have faded away. In Figure 2, it is also interesting to note that the average night lights do not correlate with the between-component except at either end of the spectrum. This is not surprising, because introducing new lights and phasing out older nodes with lights occurs at both ends. Taken together, the spatial distribution of economic activity as measured by satellite night lights seems to provide rich variation with which to explore the spatial concentration mechanism in this paper.

Table A4 presents descriptive statistics for all US state-level data. Figure A10 reinforces that the sectoral diversification–ruggedness relationship is non-monotonic within the US states: at low levels of ruggedness, states are highly diversified (i.e., lower Theil index), and at higher levels, they tend to become sectorally concentrated (i.e., higher Theil index). Finally, Figure A11 illustrates the relationship between ruggedness and spatial concentration, and shows that, on average, states with higher levels of ruggedness tend to exhibit higher spatial concentration. Unsurprisingly, this implies that entrepreneurial and residential areas are likely to cluster with more rugged terrain.

4 Estimation Method

For our empirical analysis, we estimate the following equation:

$$div_{it} = \beta_0 + \eta_t + \beta_1 rugged_i + \beta_2 rugged_{sq_i} + \boldsymbol{X}_{it}\boldsymbol{\delta} + \mathrm{Z}_\mathrm{i}\alpha + \varepsilon_{it} \qquad (1)$$

where *div* refers to sectoral diversification (measured by the Theil index, an inverse measure of diversification) for county/state *i* at time *t*, and *rugged* and *rugged_sq* are time-invariant terrain-ruggedness index and its quadratic form, respectively. The quadratic specification captures the possible non-monotonicity observed in the data presented in Figures 1, 2, and A10. $\mathbf{X}_{i,t}$ is a vector of controls that includes log per capita income, population log, and trade openness. Income level is correlated with diversification as demonstrated in Imbs and Wacziarg (2003), and we include this in the model along with its quadratic. Population is intended to control for country size, since a small economy may not be able to afford to diversify its production structure. Trade openness may create more scope for diversification because open countries are likely to be producing and consuming a variety of goods. Data for these controls come from the Penn World Tables. $\mathrm{Z_i}$ includes time-invariant country characteristics, including dummies for colonial origins and countries' oil-exporting status. Colonial origin captures the historical institutional factors, while oil exporters may have concentrated production structures. η_t is the aggregate time effects captured by year dummies. A fixed-effects estimation of equation (1) is not feasible because of the time-invariant nature of ruggedness. Therefore, our benchmark is a pooled panel approach, where standard errors are clustered at the country level.

We also subject our main results to a variety of estimation approaches, including random effects, HT estimation, and the instrumentation of log income per capita and trade openness in a two-stage least-squares estimation. A country's income is instrumented with the trade-weighted income of its neighbors (e.g., Acemoglu et. al (2008)), while trade is instrumented with the predicted changes in bilateral trade owing to foreign natural disasters, as adopted by Felbermayr and Groschl (2013).

As indicated previously, cross-country analysis is useful for examining the intensive and extensive margins of sectoral diversification with a global picture; however, it is limited in its ability to shed light on the possible mechanisms of effect, given that a large variety of countries is unlikely to form comparable units with which to tease out the mechanism. Hence, we turn to the single-country case to illuminate the potential channel. Thus, for within-country analysis, we specify an equation similar to equation (1), where we retain the quadratic form of ruggedness. The control variables are state size as measured by log population, trade openness as proxied by export share of gross state product, log per capita gross state product as a measure of income, and a dummy variable for major oil-producing states. The equation is estimated using OLS and by clustering standard errors at the state level. As above, unobserved state characteristics may result in omitted variable bias and thus could create a risk of endogeneity. While this is less of a problem in the single-country case, we nonetheless also estimate the equation using the HT estimation method, considering openness and income as endogenous.

5 Results

5.1 Cross-Country Evidence

Table 1 presents the pooled OLS results for the effect of ruggedness on sectoral diversification. Columns 1 to 6 report the results with employment as the sectoral size measure. Focusing on the overall Theil index, columns 1 and 2 demonstrate, without and with controls, respectively, that both ruggedness and its quadratic are statistically significant. Using controls in column 2, both coefficients are significant at the 1% and 5% levels, confirming that terrain ruggedness affects diversification in a non-monotonic fashion. In particular, terrain ruggedness exhibits a diversifying effect up to the ruggedness level of 2.06 (which corresponds to the 77th percentile) and has a concentrating effect beyond this. Conversely, the negative and positive signs of the coefficients of income and its quadratic, respectively, confirm the well-established non-monotonic relationship between income and sectoral diversification, as demonstrated by Imbs and Wacziarg (2003). Economic size, measured by log population, is estimated to be highly significant, with the negative coefficient implying that larger countries have a greater scope to diversify.[7]

We next focus on the intensive and extensive margins of sectoral diversification, which is one of our main contributions. Reporting the results for the extensive margin, columns 3 and 4, both with and without controls, respectively, document a statistically significant non-monotonic effect for ruggedness. By contrast, the coefficient estimates in columns 5 and 6 yield that intensive margins of sectoral diversification do not significantly correlate with a country's topographical structure. Thus, a key result here is that the relationship between the overall Theil index and terrain ruggedness is dominated by the between-component of sectoral diversification. Recall that between-component refers to the diversification arising because of new sectors being introduced or older sectors being phased out.[8] The implied turning point for the U-shaped effect of ruggedness on the between-component of the Theil index is approximately 1.7. This evidence offers important insights for economic geography literature, and should be further elaborated. It is well known that countries with rougher surfaces have less developed road and rail transport networks—for example, Ramcharan (2009) documents

[7] In unreported regressions, Herfindahl and Gini indices of sectoral diversification yield quite comparable findings to the overall Theil index regarding the effect of terrain ruggedness (results available upon request).

[8] The number of new sectors should be interpreted cautiously here, as these new products are not necessarily true entrepreneurial discoveries. In majority of cases, this corresponds to establishing firms and or producing products that already exist in other countries. True product innovation could have been measured at more detailed classifications of manufacturing firms; yet such data are not available.

that a 1% increase in surface roughness is associated with an approximately 1% decline in the number of kilometers of roadway within a country. Our result suggests that rugged terrain impedes new sectors beyond a certain level of surface roughness.

For time-varying controls, economic size proxied by the log of population is strongly related to both the between-component and within-component of the Theil index. In contrast to the case of the overall Theil, income and its quadratic do not always significantly vary with each component of diversification, but retain their negative and positive signs, respectively.

Columns 7 to 12 of Table 1 replicate the results using value-added as the sectoral size measure. The findings for terrain ruggedness are qualitatively similar, pointing to the extensive margin as the main driver of variation in sectoral diversification. As per the income effect—although the U-shaped effect is present in the overall Theil case—only the coefficient of the linear term is statistically significant. Inability to account for individual country heterogeneity and the extended timeframe in our analysis (compared with Imbs and Wacziarg (2003)) may explain the statistical insignificance of the quadratic term. Yet a significant relationship exists between income and the intensive margin of diversification, suggesting that the level of development correlates strongly with variations in the value-added shares of existing sectors.

5.2 Sensitivity Analysis of Cross-Country Results

Thus far, we have addressed individual country heterogeneity through clustered standard errors. Income and trade openness have also been assumed to be uncorrelated with the error term. To improve the econometric inference, we extend the analysis in several directions. First, we perform a random-effects estimation. As reported in Table 2, this method improves the standard errors, but Hausman tests indicate that time-varying controls are correlated with the error term (see columns 1, 4, and 7). Next, we adopt the HT estimation method, where time-varying controls are instrumented within the HT procedure.[9] Baltagi et al.'s (2003) version of the Hausman test demonstrates that compared with fixed-effects estimates, the HT estimates of time-varying controls are "acceptable." This estimation approach results in better standard errors compared with the benchmark pooled OLS estimation and reinforces the inverted U-shaped effect of topography on sectoral diversification through its extensive margin (see Table 2, columns 2, 5, and 8). Coefficients of ruggedness yield similar turning points, as in the pooled OLS estimation—that is, approximately two in the overall Theil case and approximately 1.7 in the between-Theil case. Columns 5 and 8 of Table 2 reinforce the key result that the extensive margin of diversification dominates the variations in the overall Theil index.

Nevertheless, income and trade openness may still be correlated with the regression error. We take an IV approach to address this correlation. Following Acemoglu et al. (2008), we adopt the trade-share-weighted average income of trading partners as an instrument for a country's own income. In addition, parallel to Felbermayr and Groschl (2013), we use the predicted changes in bilateral trade owing to foreign natural disasters as instruments for trade openness. Foreign natural disasters are clearly exogenous to a country's sectoral diversification, and are likely to influence diversification through their effects on trade links. However, it is not immediately clear, as also acknowledged by Acemoglu et al. (2008) in the context of their study, whether the trade-share-weighted income of trading partners can be excluded from the diversification equation. It is more reasonable to assume that this instrument provides only a partial correction to endogenous income, rather than a fully causal relationship. That said, both IVs have strong explanatory powers in the first stage, as can be seen in the F-statistics from Table 2 that are above the rule of thumb 10. Focusing on the sensitivity of ruggedness, the inefficient 2SLS estimator, as anticipated, produces higher standard errors

[9] Time-invariant controls are not instrumented. It does not appear that strong grounds exist for the endogeneity of ruggedness, colonial origins, and oil-exporter status.

compared with pooled OLS (see Table 2, columns 3, 6, and 9). Yet the results reinforce our main finding that ruggedness is related to sectoral diversification in an inverted U-shaped fashion through its extensive margin. Value-added-based regressions in panel B of Table 2 point to analogous variations, except that the standard errors are slightly higher.

5.3 Within-Country Evidence

We now turn to the US state-level data. We initially investigate whether the non-monotonic relationship obtained from cross-country data also holds for the US. Subsequently, and more importantly, we analyze the mechanism through which this relationship works. Recall that for the US, we are unable to compute the extensive and intensive margins of sectoral diversification, and hence consider only overall sectoral diversification—however, we are able to compute the extensive and intensive margins of spatial concentration.

Column 1 in Table 3, using the size metric of value-added to measure sectoral diversification, displays the inverted U-shaped link between ruggedness and sectoral diversification observed at the cross-country level. Although the coefficients of ruggedness and its quadratic are not significant at conventional levels, the *t*-statistic of the non-linear term is greater than 1.5. Further, the non-linear term becomes significant at the 5% level in the full model with controls (column 2), where the linear ruggedness term is negative with its t-statistic greater than 1.5. The HT estimation of the full model, which accounts for the possible endogeneity of income and trade openness, also confirms the non-linear relationship between ruggedness and sectoral diversification, albeit with high standard errors. The latter is most likely because the HT is an inefficient IV estimator. Among the controls, larger population creates more scope for diversification, while income has a concentrating effect.[10] There is also some evidence that major oil-producing states tend to concentrate.

[10] Quadratic of income is not included in this estimation because unlike cross-country data, income among the US states is less heterogeneous.

Our findings in Table 4 importantly document that the relationship between terrain ruggedness and spatial concentration is positive and significant. As the estimates with quadratic form of ruggedness are consistently insignificant, we surmise that the relationship is linear. The estimates, obtained using pooled OLS with clustered standard errors at the state level, are also robust to the HT estimation. Table 4 explores the effect of ruggedness on the between- and within-components of the *spatial* Theil index in columns 5 to 12 using both OLS and HT estimations. Columns 11 and 12 suggest that ruggedness contributes more to the within-component of the spatial concentration, given its statistically and economically more significant coefficients, than in the case of the between-component in columns 7 and 8.

We now turn to the mechanism through which topography may affect sectoral diversification. Recall that we measure the geographic concentration of firms and households by the Theil index of satellite night lights data along with its within- and between-components. We should reiterate here that the within-component refers to the density of light in a given localized geographical area. Conversely, the between-component captures how the spatial distribution of areas with light evolves over time. An increase in this component would imply a rise in spatial concentration. Table 5 reports our regression analysis, which includes variables of overall spatial Theil index (columns 1 and 2), spatial between-component (columns 3 and 4), and spatial within-component (columns 5 and 6), with OLS and HT estimations. The following key findings stand out. First, the coefficients of the overall spatial Theil index in columns 1 and 2 show that sectoral diversification tends to increase on average as economic activity becomes more spatially concentrated. This result is robust to HT estimation. Second, both the between- and within-components of the spatial Theil index (columns 3 and 5) are significant with OLS, and the between-component stands with a t-statistic of 1.56 using HT estimation (column 4). This implies that the more new residential or industrial areas are built by extending the geographic areas, the higher the sectoral diversification levels. Third, with spatial concentration measures added in the model, the coefficient of ruggedness is moderately augmented, especially when the between-component of (spatial) Theil index is included (column 3). This provides evidence for the extensive margin of spatial concentration as the mechanism of effect for the ruggedness–sectoral diversification relationship.

The HT estimates in columns 4 and 6 generally support the findings obtained with OLS, although the standard errors of the between- and within-components tend to be larger. This is most probably because of the HT estimator's inefficiency. Nonetheless, as a word of caution, the OLS coefficients obtained in columns 1, 3, and 5 are larger than the HT estimates, which may imply that our OLS coefficients for the Theil indices may be upwardly biased.

Taken together, there is strong evidence that the (adverse) effect of spatial concentration on sectoral diversification works through the extensive margin of spatial concentration. The broad conclusion is that rugged topography restricts the geographic clusters from spatially extending; therefore, their sectoral structure remains concentrated. The corollary of this finding is that if a state extends over space by introducing new lights, it may introduce new sectors. To exemplify this, consider chemical factories and food factories, which are unlikely to operate within proximity. Thus, operating in different (light) clusters would mean that more sectors are in operation. Our finding rules out the alternative hypothesis that ruggedness may enable opportunities for a larger home market by facilitating increasing returns to scale and higher productivity levels through concentrating economic activities in a narrow area, and thus generating scope for sectoral diversification (Ramcharan (2009)).

6 Conclusion

This paper investigates the role of physio-geography in sectoral diversification. Our investigation takes a two-pronged approach in which we first conduct a cross-country analysis of 142 countries over the period 1970–2007, then a within-country analysis using US state-level data for the period 1997–2011. While the cross-country data provide useful insights into the global relationship, the within-country data overcome the drawback of the former by offering a homogeneous platform for a meaningful examination of the mechanism.

We present strong evidence for a robust and highly significant non-monotonic effect of ruggedness in the cross-country setting. In particular, ruggedness initially assists sectoral diversification, but exhibits a concentrating effect beyond a certain level. In a unique contribution, we next explore the extensive and intensive margins of sectoral diversification. We demonstrate that the relationship between terrain ruggedness and patterns in the real economy mostly works through the extensive margin of sectoral diversification, and, to a lesser extent, through the intensive margin.

Moving forward, we verify the aforementioned non-monotonic effect with US state-level data. Then, we establish that terrain ruggedness increases spatial concentration, which in turn reduces sectoral diversification. We measure spatial concentration with the satellite night lights data of Henderson, Storeygard, and Weil (2012), where we adopt a more literal interpretation of data to represent the geographic distribution of night lights. Thus, our approach to using satellite night lights data contrasts with those of most previous studies, which have used the data only to proxy income. This approach, quite innovatively, also enables us to compute the spatial extensive and intensive margins of night lights for the US. In this vein, we document that as residential or industrial areas extend in space (i.e., spatial extensive margin), sectoral diversification tends to increase. This relationship is generally robust to alternative estimation approaches.

While geography's role in economic development is widely studied, to the best of our knowledge, this is the first paper that investigates its role in sectoral diversification, and, more importantly, in the extensive and intensive margins of the latter. Another novel aspect of this study is to shed light on the link between physio-geography and spatial concentration, as well as the extensive and intensive margins of spatial distribution of economic activity. Taken together, our findings could inform the global debate on spreading risks across sectors to mitigate the vulnerability of real economies. In particular, our results suggest that a combination of innovative policies that incorporate the efficient use of space, urban planning, and intersectoral linkages could help alleviate the adverse role of natural barriers that stand in the way of a healthy trajectory of sectoral development.

References

Acemoglu, D, Johnson, S, Robinson, J, & Yared, P 2008, 'Income and democracy', *American Economic Review*, vol. 98, no. 3, pp. 808–842.

Baltagi, B, Bresson, G, & Pirotte, A 2003, 'Fixed effects, random effects or Hausman–Taylor? A pretest estimator', *Economics Letters*, vol. 79, pp. 361–369.

Cadot, O, Carrère, C, & Strauss-Kahn, V 2011, 'Export diversification: what's behind the hump?', *Review of Economics and Statistics*, vol. 93, no. 2, pp. 590–605.

Cadot, O, Carrère, C, & Strauss-Kahn, V 2013, 'Trade diversification, income and growth: what do we know?' *Journal of Economic Surveys*, vol. 27, no. 4, pp. 790–812.

Felbermayr, G, & Groschl, J 2013, 'Natural Disasters and the effect of trade on income: a new panel IV approach', *European Economic Review*, vol. 58, pp. 18–30.

Gallup, J, Sachs, J D, & Mellinger, A 2009, 'Geography and Economic Development', *International Regional Science Review*, vol. 22, no. 2, pp. 179–232.

Henderson, J V, Storeygard, A, & Weil, D N 2012, 'Measuring Economic Growth from Outer Space', *American Economic Review*, vol. 102, no. 2, pp. 994–1028.

Hodler, R. & Raschky, P. 2014, 'Regional Favoritism', *Quarterly Journal of Economics*, vol. 129, no. 2, pp. 995 – 1033.

Imbs, J, & Wacziarg, R 2003, 'Stages of diversification', *American Economic Review*, vol. 93, no. 1, pp. 63–86.

IMF 2014, 'Sustaining Long-Run Growth and Macroeconomic Stability in Low-Income Countries—The Role of Structural Transformation and Diversification' IMF policy paper, International Monetary Fund, Washington.

INDSTAT2 2012 Rev.3, UNIDO Industrial Statistics Database at the 2-digit level of ISIC Code (Revision 3), United Nation's Industrial Development Organization.

Jimenez-Ayora, P & Ulubasoglu, M, 2015, 'What Underlies Weak States? The Role of Terrain Ruggedness', *European Journal of Political Economy*, vol. 39, pp. 167–183.

Khalifa, S 2016, 'Trust, Landscape, and Economic Development', *Journal of Economic Development*, vol. 41, no. 1, pp. 19–31.

Koren, M, & Tenreyro, S 2007, 'Volatility and development', *Quarterly Journal of Economics*, vol. 122, no. 1, pp. 243–287.

La Porta, R., Lopez-de- Silanes, F., Shleifer, A., Vishny, R., 1999, 'The quality of government', *Journal of Law and Economic Organisation*, vol. 15, pp. 222–279.

Malik, Adeel, & Temple, J 2009, 'The Geography of Output Volatility', *Journal of Development Economics*, vol. 90, no. 2, pp. 163–178.

Mellinger, A, Sachs, J D, & Gallup, J, "Climate, Coastal Proximity, and Development" (pp. 169–194), in Gordon L. Clark, Maryann P. Feldman, and Meric S. Gertler (Eds.), *Oxford Handbook of Economic Geography* (New York: Oxford University Press, 2000)

Nunn, N, & Puga, D 2012, 'Ruggedness: the blessings of bad geography in Africa', *Review of Economics and Statistics*, vol. 94, no. 1, pp. 20–36.

National Geophysical Data Center. 2016. "Version 4 DMSP-OLS Nighttime Lights Time Series." Earthquake Observation Group (EOG), National Oceanic and Atmospheric Administration. Website: https://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html (accessed July 15, 2016).

Radelet, S, & Sachs, J 1999, 'What have we learned, so far, from the Asian Financial Crises?', paper from the Consulting Assistance on Economic Reform (CAER) II Project.

Ramcharan, R 2009, 'Why an economic core: domestic transport costs', *Journal of Economic Geography*, vol. 9, pp. 559–581.

Ramcharan, R 2010, 'The link between the economic structure and financial development', *BE Journal of Macroeconomics*, vol. 10, no. 1, pp. 1–35.

Riley, S, DeGloria, S, & Elliot R 1999, 'A terrain ruggedness index that quantifies topographic heterogeneity', *Intermountain Journal of Sciences*, vol. 5, no. 1–4, pp. 23–27.

Shaver, A, Carter, D B, & Shawa, Tsering Wangyal 2015, 'Reassessing the Basic Determinants of Civil War with Improved Terrain Ruggedness and Land Cover Data', Forthcoming *at Conflict Management and Peace Science*.

The News Journal 2013, '*With right skills, Delaware's financial services industry is a land of plenty*', By Wade Malcolm. Website: http://www.delawareonline.com/story/money/business/2013/12/07/with-right-skills-delawares-financial-services-industry-is-a-land-of-plenty/3906739/ (accessed on 29 Aug 2017)

Ulubasoglu, M, & Cardak, B 2007, 'International comparisons of rural-urban educational attainment: data and determinants', *European Economic Review*, vol. 51, no.7, pp. 1828–1857.

Figures and Tables

Figure 1: Terrain Ruggedness and Sectoral Diversification – Cross-Country Data

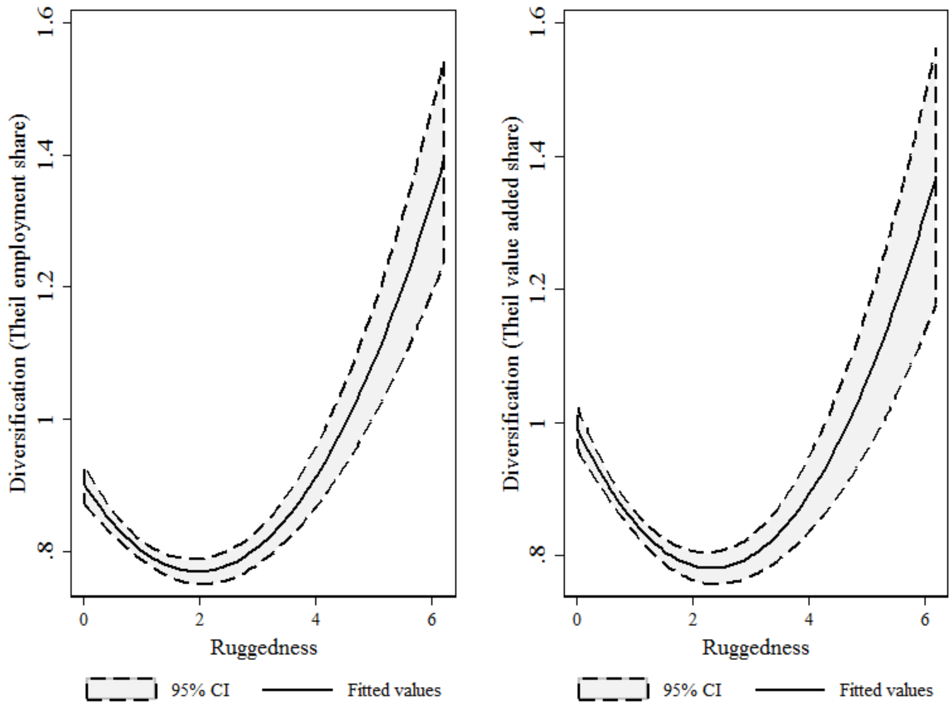


Figure 2: Average Night Lights and the Within- and Between Components of Night Lights

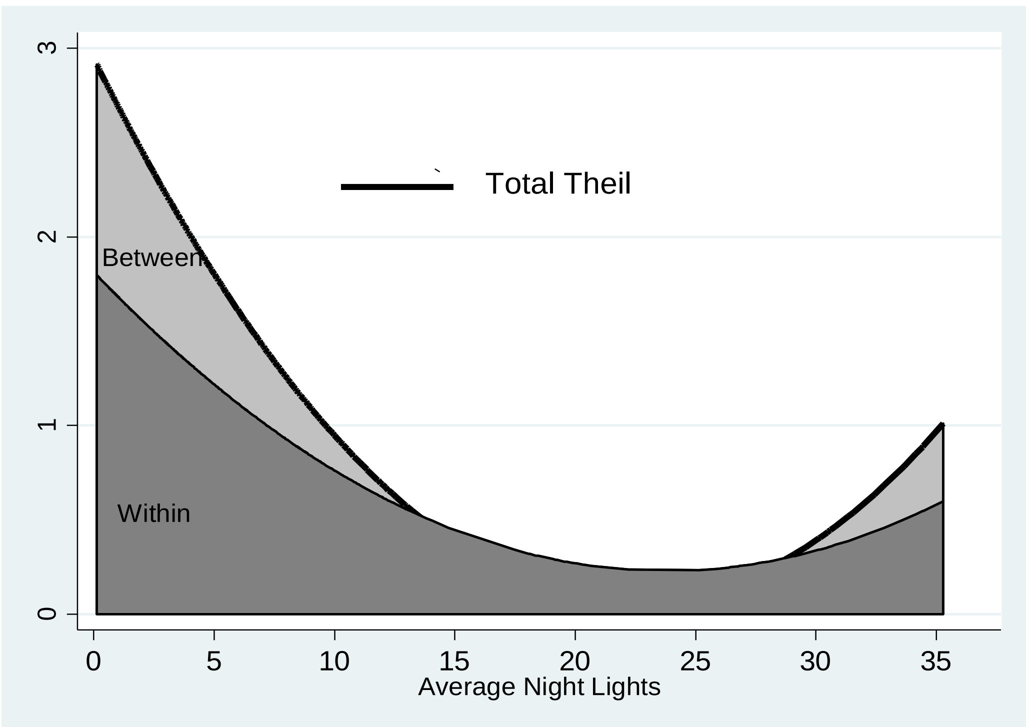


Table 1: Topography and Sectoral Diversification—Pooled OLS Results of Cross-Country Analysis

	Employment based diversification measure						Value-added-based diversification measure					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Theil overall	Theil overall	Theil between	Theil between	Theil within	Theil within	Theil overall	Theil overall	Theil between	Theil between	Theil within	Theil within
Ruggedness	-0.139*	-0.099**	-0.087**	-0.055*	-0.042	-0.046	-0.198**	-0.136**	-0.138**	-0.104*	-0.055	-0.043
	(0.078)	(0.044)	(0.041)	(0.033)	(0.056)	(0.040)	(0.092)	(0.056)	(0.060)	(0.055)	(0.058)	(0.035)
Ruggedness square	0.037**	0.024***	0.024***	0.016**	0.009	0.009	0.043**	0.031***	0.035***	0.028***	0.005	0.005
	(0.016)	(0.009)	(0.007)	(0.006)	(0.013)	(0.008)	(0.020)	(0.011)	(0.012)	(0.011)	(0.013)	(0.008)
Income (log)		-0.614***		-0.296		-0.172		-0.746*		0.043		-0.634**
		(0.236)		(0.231)		(0.192)		(0.391)		(0.332)		(0.265)
Income square (log)		0.022*		0.012		0.002		0.031		-0.007		0.029*
		(0.014)		(0.013)		(0.011)		(0.022)		(0.019)		(0.015)
Population (log)		-0.117***		-0.056***		-0.058***		-0.147***		-0.064***		-0.077***
		(0.012)		(0.008)		(0.011)		(0.014)		(0.012)		(0.009)
Trade openness		-0.020		-0.008		0.005		-0.060*		-0.038		-0.010
		(0.030)		(0.024)		(0.022)		(0.033)		(0.033)		(0.025)
Country	142	142	132	132	135	135	134	134	122	122	125	125
Observations	3,402	3,402	2,586	2,586	2,977	2,977	3,091	3,091	2,408	2,408	2,693	2,693
R-squared	0.049	0.544	0.057	0.319	0.008	0.373	0.040	0.590	0.075	0.319	0.039	0.489

Note: Dependent variable for each regression is the Theil index, which is the inverse measure of sectoral diversification. Robust standard errors clustered at the country level in parentheses. Columns 2, 4, 6, 8, 10 and 12 include a dummy for major oil-exporting countries and three dummies for colonial origins. All regressions include year dummies and a constant. *** p<0.01, ** p<0.05, * p<0.1.

Table 2: Topography and Sectoral Diversification—Random Effects, Hausman–Taylor and Two-Stage Least Squares Estimations of Cross-Country Analysis

	(1) RE	(2) HT	(3) 2SLS	(4) RE	(5) HT	(6) 2SLS	(7) RE	(8) HT	(9) 2SLS
	Theil overall	Theil overall	Theil overall	Theil Between	Theil between	Theil between	Theil within	Theil within	Theil within
Panel A: Employment-based diversification measure									
Ruggedness	-0.176***	-0.183*	-0.082	-0.094**	-0.094**	-0.065**	-0.084**	-0.091	-0.019
	(0.053)	(0.095)	(0.055)	(0.039)	(0.046)	(0.032)	(0.042)	(0.063)	(0.053)
Ruggedness square	0.040***	0.042**	0.021*	0.020***	0.021**	0.019**	0.018**	0.020	0.004
	(0.009)	(0.020)	(0.013)	(0.007)	(0.010)	(0.009)	(0.008)	(0.013)	(0.013)
Income (log)	-0.294	-0.213**	-0.213***	-0.122	-0.063	-0.087**	-0.293	-0.277***	-0.143***
	(0.264)	(0.098)	(0.049)	(0.229)	(0.110)	(0.037)	(0.256)	(0.091)	(0.044)
Income square (log)	0.010	0.008		0.002	-0.002		0.014	0.014***	
	(0.015)	(0.005)		(0.013)	(0.006)		(0.014)	(0.005)	
Population (log)	-0.086***	-0.050***	-0.121***	-0.057***	-0.056***	-0.048***	-0.050***	-0.046***	-0.069***
	(0.020)	(0.018)	(0.015)	(0.011)	(0.012)	(0.011)	(0.015)	(0.014)	(0.014)
Trade openness	-0.026	-0.025*	-0.063	0.006	0.014	-0.022	-0.017	-0.020	-0.012
	(0.039)	(0.015)	(0.056)	(0.039)	(0.016)	(0.042)	(0.028)	(0.013)	(0.042)
Country	142	142	111	132	132	103	135	135	104
Observations	3,402	3,402	2,848	2,586	2,586	2,226	2,977	2,977	2,507
R-squared	0.49	-	-	0.31	-	-	0.29	-	-
Hausman test p-value	0.03	0.99	-	0.00	0.99	-	0.03	-	-
F-stat-income, openness	-	-	12.7, 7.5	-	-	8.9, 5.9	-	-	11.1, 7.4
A-P F-Income, openness	-	-	20.6, 14.4	-	-	13.9, 10.4	-	-	18.6, 14.4
Panel B: Value-added-based diversification measure									
Ruggedness	-0.221***	-0.340	-0.107*	-0.169***	-0.219*	-0.050	-0.049	-0.107	-0.053
	(0.061)	(0.241)	(0.057)	(0.054)	(0.131)	(0.043)	(0.041)	(0.094)	(0.044)
Ruggedness square	0.049***	0.081	0.019	0.040***	0.053*	0.013	0.008	0.020	0.006
	(0.011)	(0.051)	(0.014)	(0.009)	(0.028)	(0.010)	(0.008)	(0.020)	(0.010)
Income (log)	-0.412	-0.656***	-0.200***	0.288	0.268**	-0.078*	-0.485	-0.447***	-0.135***
	(0.387)	(0.127)	(0.052)	(0.214)	(0.116)	(0.046)	(0.306)	(0.122)	(0.048)
Income square (log)	0.015	0.034***		-0.020*	-0.017***		0.022	0.022***	
	(0.022)	(0.007)		(0.012)	(0.006)		(0.018)	(0.007)	
Population (log)	-0.107***	0.188***	-0.149***	-0.043***	0.071***	-0.062***	-0.064***	-0.021	-0.081***
	(0.018)	(0.030)	(0.014)	(0.015)	(0.023)	(0.010)	(0.012)	(0.019)	(0.011)
Trade openness	-0.101**	-0.100***	-0.116*	-0.056	-0.053***	-0.089**	-0.016	-0.017	-0.007
	(0.051)	(0.018)	(0.063)	(0.034)	(0.016)	(0.043)	(0.033)	(0.017)	(0.046)
Country	134	134	107	122	122	100	125	125	100
Observations	3,091	3,091	2,709	2,408	2,408	2,132	2,693	2,693	2,374
R-squared	0.55	-	-	0.27	-	-	0.48	-	-
Hausman test p-value	0.00	0.68	-	0.02	0.96	-	0.26	-	-
F-stat-income, openness	-	-	14..4, 50.3	-	-	10.7, 50.3	-	-	12.4, 54.1
A-P F-Income, openness	-	-	17.4, 93.5	-	-	13.2, 90.6	-	-	15.3, 101.7

Note: Dependent variable for each regression is the Theil index, which is the inverse measure of sectoral diversification. RE, HT and 2SLS stand for random effects, Hausman–Taylor, and two-stage least square estimation, respectively. For RE and 2SLS, robust standard errors clustered at country level in parentheses, for HT standard errors in parentheses. All regressions include a dummy for major oil-exporting countries and three dummies for colonial origins. All regressions include year dummies and a constant. Hausman test for RE and HT are in comparison with fixed-effects estimation. Hausman test in column 8 is undefined. In 2SLS, only linear income is instrumented because the instrument cannot explain income and its quadratic. In each panel, F-stat and A-P F are the first stage F statistics and Angrist-Pischke F-statistics, respectively for both income and openness. The instruments are mostly strong when evaluated with respect to critical values reproduced in statistical software. 2SLS results are robust to LIML estimation. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Topography and Sectoral Diversification in the US – Pooled OLS and Hausman-Taylor Results

	(1) OLS	(2) OLS	(3) HT
Ruggedness	-0.152	-0.225	-0.100
	(0.195)	(0.146)	(0.429)
Ruggedness square	0.107	0.117**	0.091
	(0.070)	(0.053)	(0.179)
Population (log)		-0.114***	0.116*
		(0.031)	(0.061)
Income (log)		0.559***	0.261***
		(0.174)	(0.080)
Oil dummy		0.145*	0.208
		(0.085)	(0.200)
Export (% of GSP)		0.006	0.001
		(0.011)	(0.002)
Observations	750	750	750
R-squared	0.102	0.467	-
No. of State	50	50	50

Note: Dependent variable for each regression is the Theil index, which is the inverse measure of sectoral diversification. Robust standard errors clustered at the state level in parentheses. All regressions include year dummies and a constant. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Topography and Spatial Concentration of Economic Activity in the US - Pooled OLS and Hausman Taylor Estimation Results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	T_OLS	T_HT	T_OLS	T_HT	TB_OLS	TB_HT	TB_OLS	TB_HT	TW_OLS	TW_HT	TW_OLS	TW_HT
Ruggedness	0.327	0.482	0.408***	0.417***	0.242	0.280	0.144*	0.146*	0.085	0.172	0.264***	0.258***
	(0.504)	(0.532)	(0.130)	(0.144)	(0.280)	(0.286)	(0.081)	(0.078)	(0.282)	(0.287)	(0.068)	(0.078)
Ruggedness square	0.035	-0.028			-0.043	-0.058			0.078	0.038		
	(0.185)	(0.223)			(0.117)	(0.119)			(0.102)	(0.120)		
Population (log)	-0.250**	-0.190**	-0.249**	-0.191***	-0.147**	-0.131***	-0.148**	-0.132***	-0.102**	-0.105**	-0.102**	-0.104**
	(0.108)	(0.074)	(0.108)	(0.074)	(0.062)	(0.048)	(0.062)	(0.048)	(0.050)	(0.048)	(0.050)	(0.048)
Income (log)	-0.951	-0.369***	-0.946	-0.369***	-0.439	-0.304***	-0.445	-0.304***	-0.512	-0.071	-0.501	-0.071
	(0.612)	(0.095)	(0.607)	(0.095)	(0.338)	(0.096)	(0.341)	(0.096)	(0.311)	(0.090)	(0.307)	(0.090)
Oil dummy	0.939***	0.890***	0.937***	0.891***	0.366***	0.355***	0.369***	0.357***	0.574***	0.529***	0.568***	0.527***
	(0.211)	(0.248)	(0.207)	(0.247)	(0.117)	(0.133)	(0.119)	(0.133)	(0.111)	(0.133)	(0.105)	(0.133)
Export (% of GSP)	-0.026	-0.002	-0.025	-0.002	-0.007	-0.002	-0.009	-0.002	-0.019*	-0.001	-0.016	-0.001
	(0.029)	(0.002)	(0.027)	(0.002)	(0.021)	(0.002)	(0.018)	(0.002)	(0.010)	(0.002)	(0.010)	(0.002)
Observations	750	750	750	750	750	750	750	750	750	750	750	750
R-squared	0.466	-	0.465	-	0.365	-	0.363	-	0.500	-	0.496	-
No of State	50	50	50	50	50	50	50	50	50	50	50	50

Note: Dependent variable for each regression is the Theil index of spatial concentration of economic activities as measured by Satellite Night Lights data. T, TB and TW denote Theil overall, Theil between and Theil within, respectively. OLS and HT denote the underlying estimators and robust standard errors are in parentheses. Missing exports values are addressed with ‘missing dummy approach’. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Topography and Sectoral Diversification in the US: Is Spatial Concentration a Mechanism?

	(1)	(2)	(3)	(4)	(5)	(6)
	Theil Sectoral	Theil Sectoral	Theil Sectoral	Theil Sectoral	Theil Sectoral	Theil Sectoral
	OLS	HT	OLS	HT	OLS	HT
Ruggedness	-0.254*	-0.134	-0.265*	-0.116	-0.236*	-0.103
	(0.133)	(0.420)	(0.133)	(0.424)	(0.138)	(0.429)
Ruggedness square	0.114**	0.094	0.124**	0.094	0.107**	0.090
	(0.048)	(0.176)	(0.050)	(0.178)	(0.049)	(0.179)
Population (log)	-0.092***	0.123**	-0.090***	0.118*	-0.101***	0.117*
	(0.027)	(0.060)	(0.028)	(0.060)	(0.028)	(0.061)
Income (log)	0.642***	0.284***	0.631***	0.275***	0.623***	0.262***
	(0.158)	(0.080)	(0.142)	(0.080)	(0.173)	(0.080)
Oil dummy	0.063	0.150	0.085	0.190	0.073	0.200
	(0.092)	(0.198)	(0.089)	(0.198)	(0.090)	(0.200)
Export (% of GSP)	0.008	0.002	0.007	0.001	0.008	0.001
	(0.010)	(0.002)	(0.010)	(0.002)	(0.010)	(0.002)
Spatial concentration (Theil overall)	0.088**	0.065**				
	(0.039)	(0.032)				
Spatial concentration (Theil between)			0.165**	0.050		
			(0.075)	(0.032)		
Spatial concentration (Theil within)					0.126**	0.015
					(0.060)	(0.034)
Observations	750	750	750	750	750	750
R-squared	0.511		0.514		0.495	
No. of State	50	50	50	50	50	50

Note: Dependent variable for each regression is the Theil index, which is the inverse measure of sectoral diversification. OLS and HT denote the underlying estimators. Robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

ONLINE APPENDICES

Figure A1: Terrain Ruggedness Across Countries

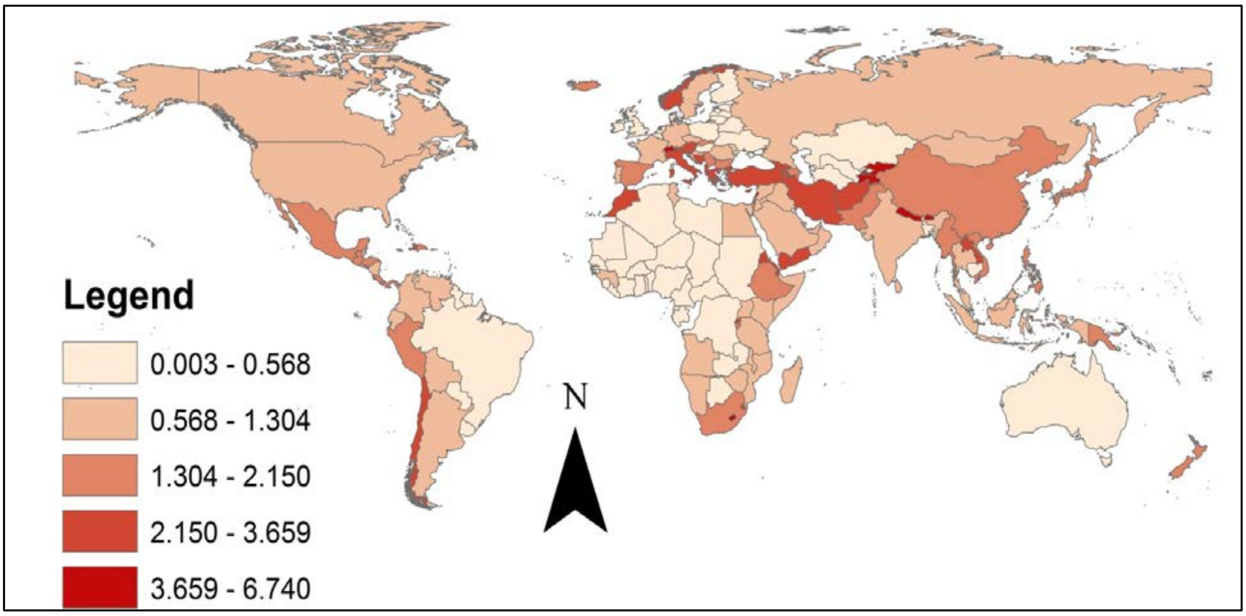


Figure A2: Terrain Ruggedness Across the US States

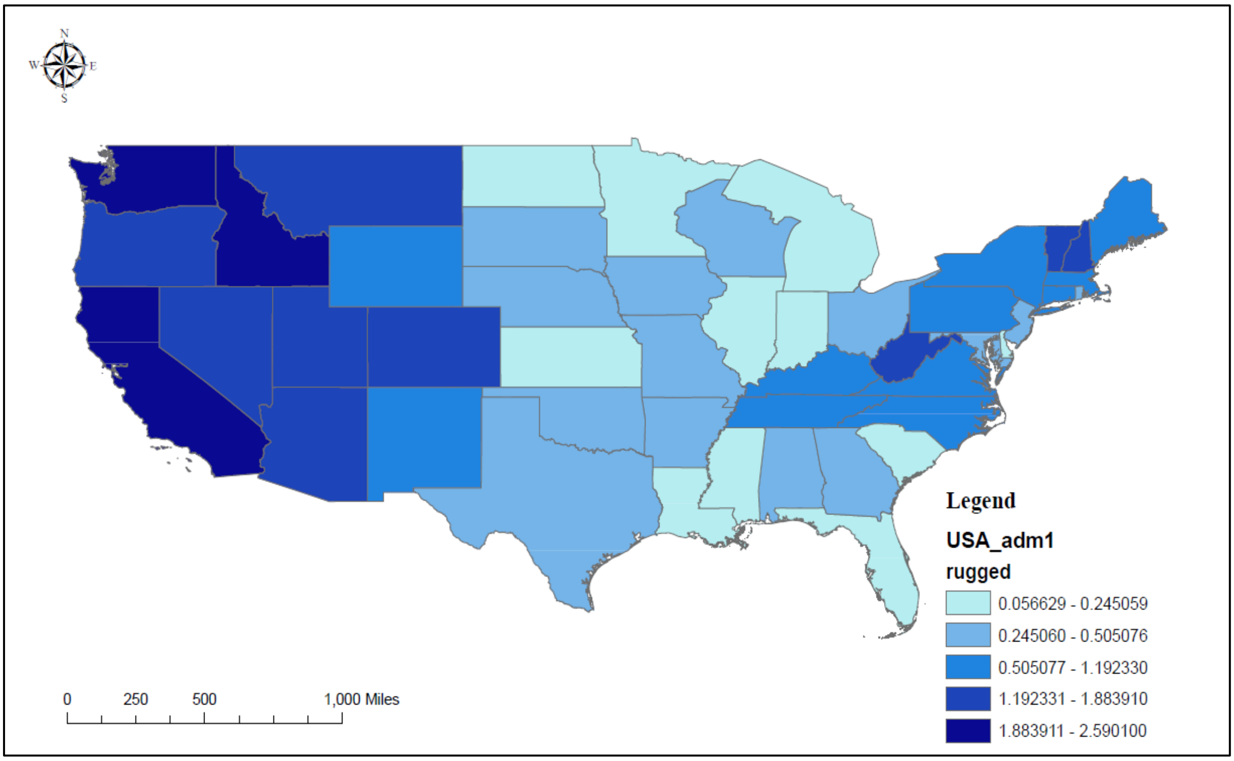


Figure A3: Sectoral Diversification Across Countries (Employment Share-Based Theil index)

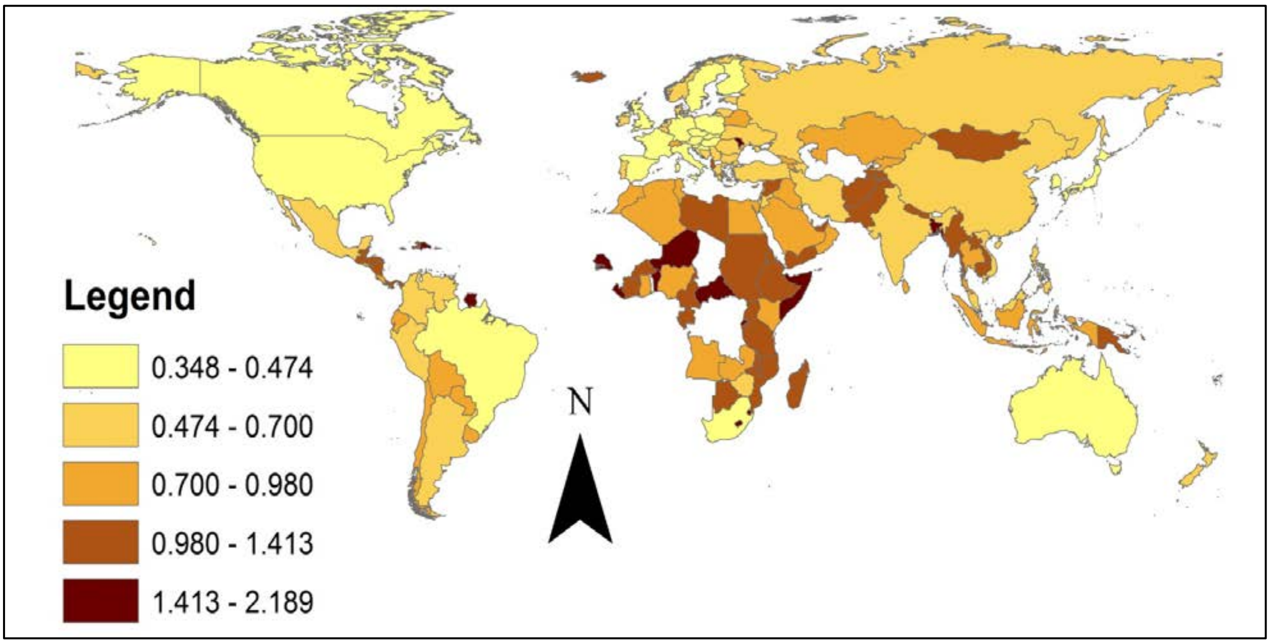


Figure A4: Sectoral Diversification Across US states (Value Added-Based Overall Theil Index)

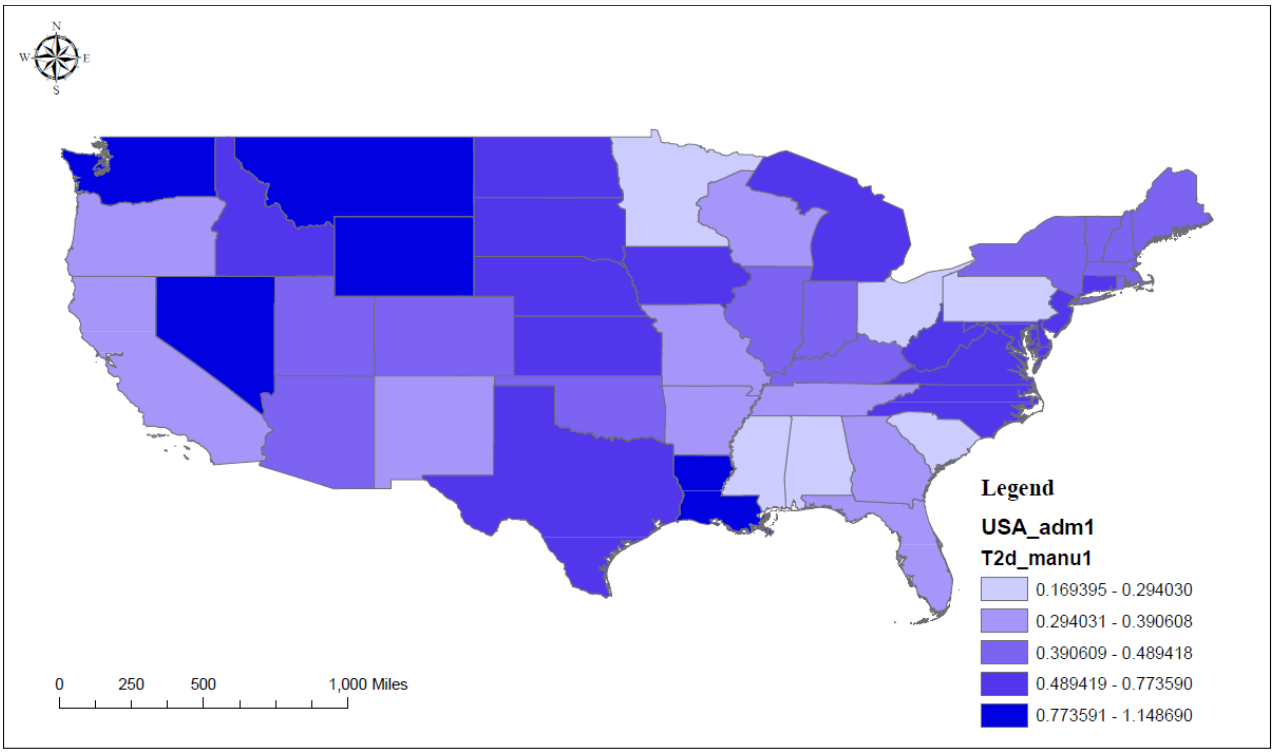


Figure A5: Spatial Concentration of Night Lights: Schematic of the Within- and Between-Components

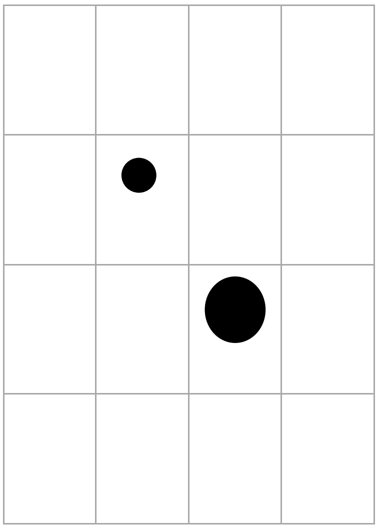


Figure A5.1: Baseline

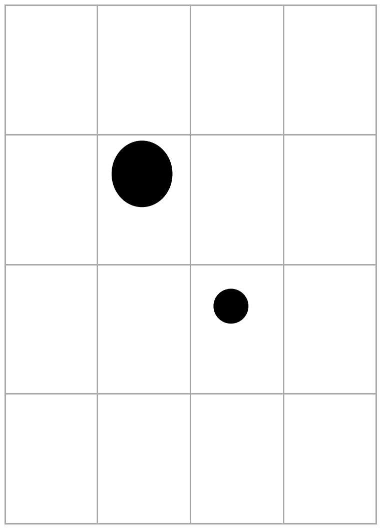


Figure A5.2: No Change in Spatial Concentration

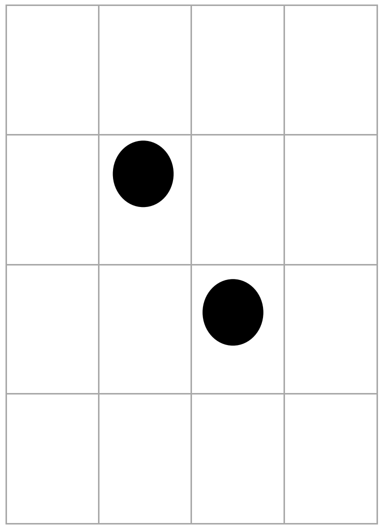


Figure A5.3: A Decrease in the Within-Component of Spatial Concentration

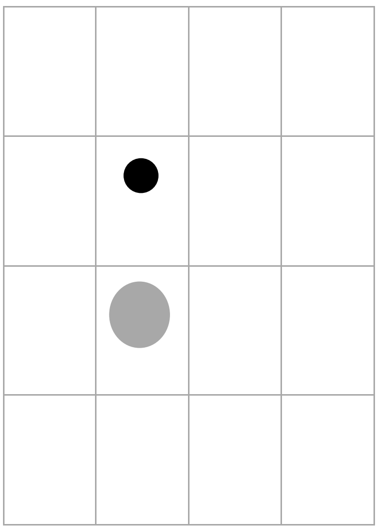


Figure A5.4: No Change in Spatial Concentration

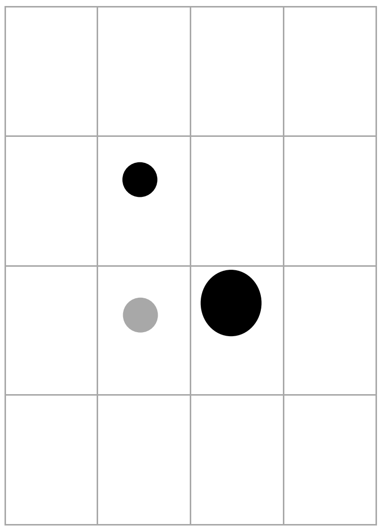


Figure A5.5: A Decrease in the Between-Component of Spatial Concentration

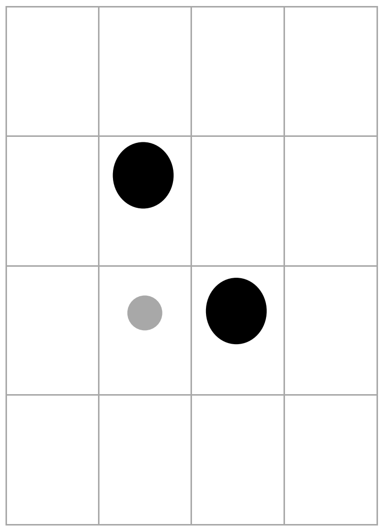


Figure A5.6: A Decrease in both the Between- and Within-Components of Spatial Concentration

Note: Figure 5 illustrates 30-by-30 arcsecond cells, with each cell centered on a point from the night lights grid. The solid circles in black present the size of existing night lights, and the grey colored circles represent new night lights introduced in cells that were dark previously. The empty cells show the absence of night lights. The size of the circle shows the density of night lights in a given cell.

Figure A6: Spatial Concentration of Economic Activity Across the US States (Overall Theil Index Based on Satellite Night Lights Data)

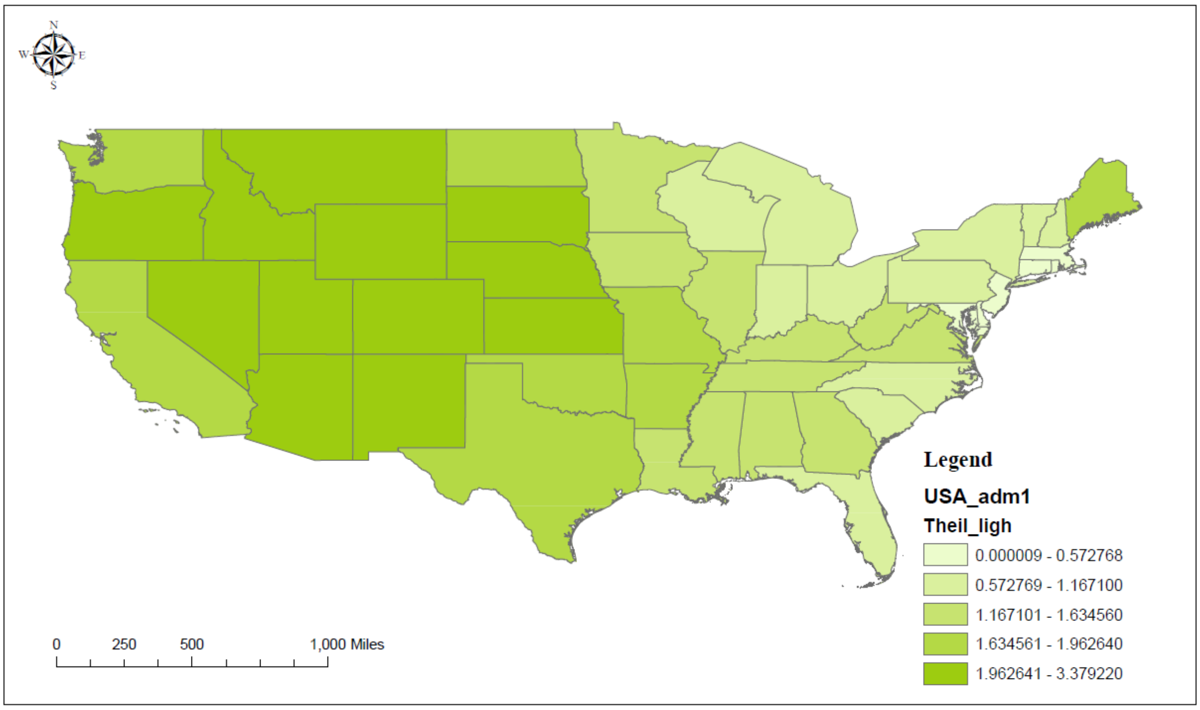
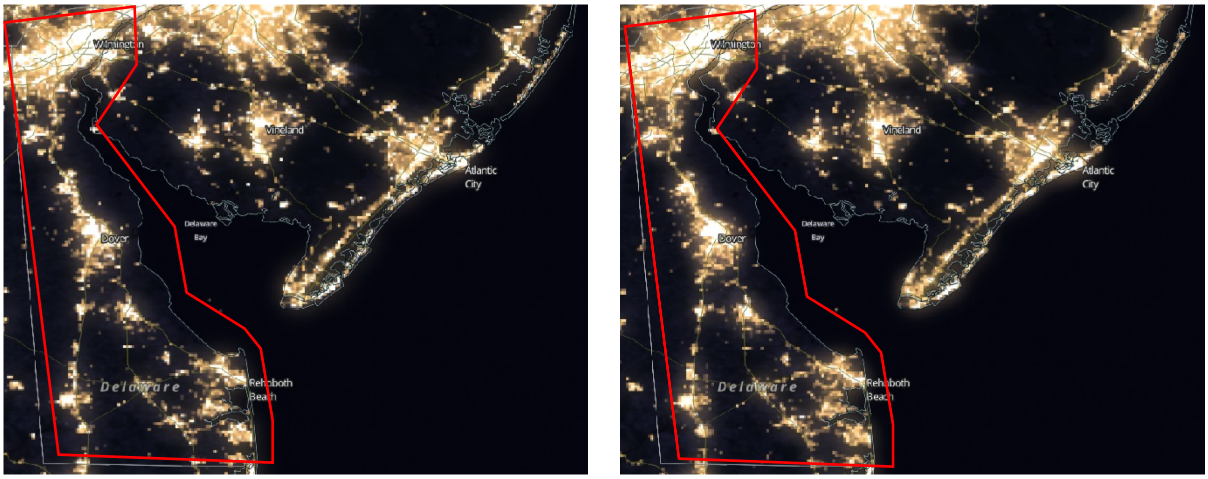


Figure A7: Oil Fields in North Dakota, 2012

Note: This picture shows the 2012 night lights in some parts of the United States. The area marked with a red circle shows the location of oil fields in North Dakota; this area was not illuminated until 2006.

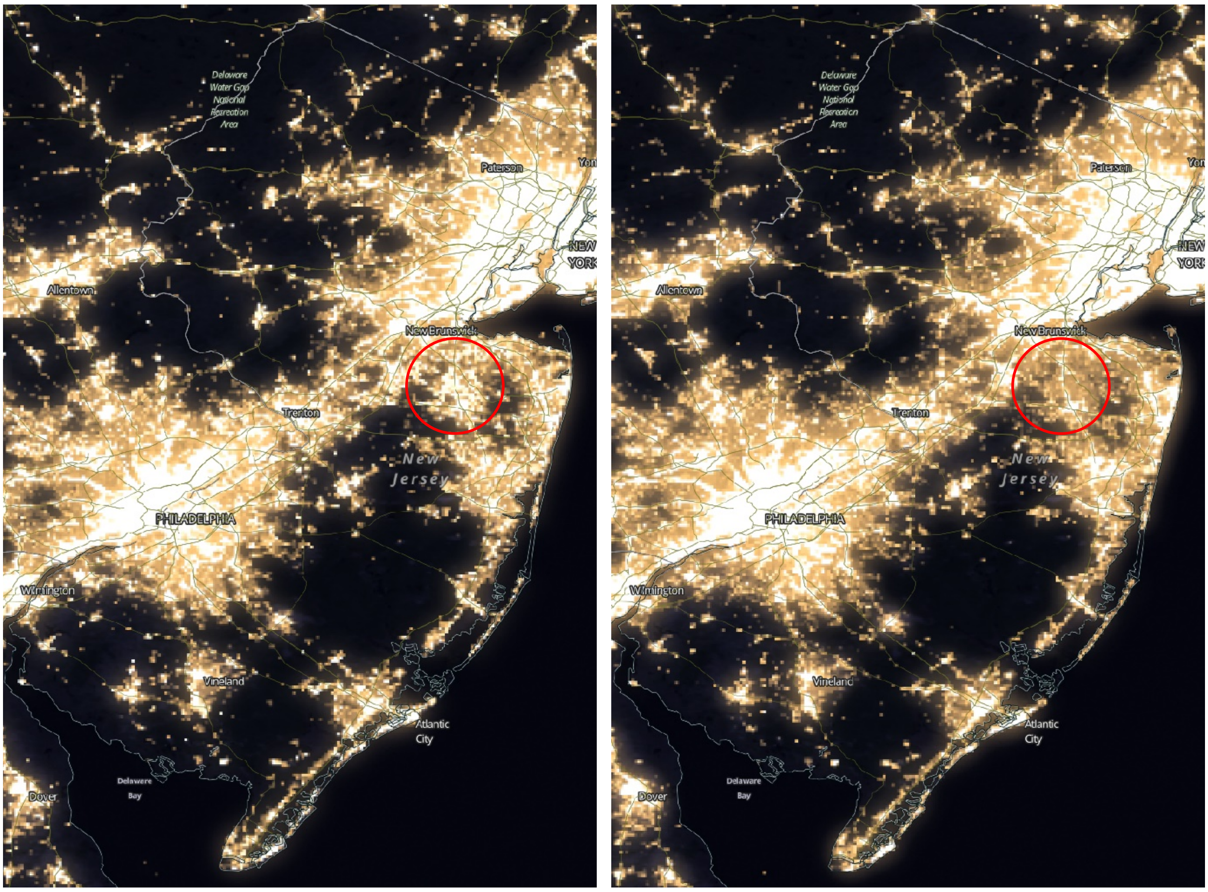
Figure A8: Night Lights in Delaware, 2012 and 2016



Night Lights, Delaware 2012

Night Lights, Delaware 2016

Figure A9: Night Lights in New Jersey, 2012 and 2016



Night Lights, New Jersey 2012

Night Lights, New Jersey 2016

Figure A10: Terrain Ruggedness and Sectoral Diversification in the US

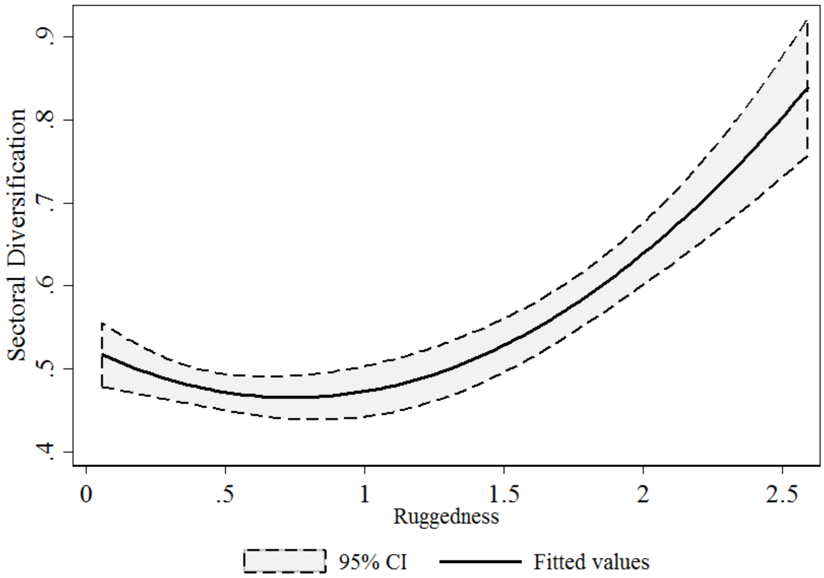


Figure A11: Terrain Ruggedness and Spatial Concentration of Economic Activity in the US

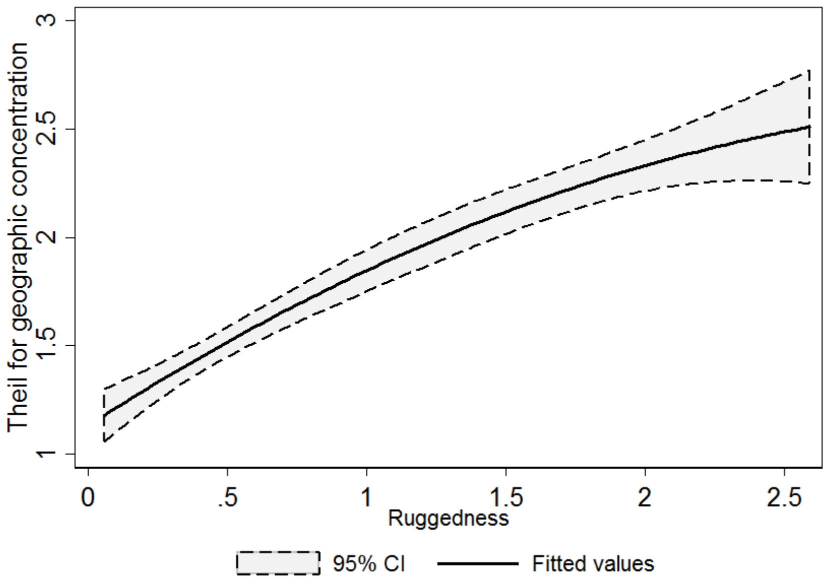


Table A1: Variables and Data Sources – Cross-Country Analysis

Variable	Variable description	Data source	Availability
Ruggedness	Terrain ruggedness	Nunn & Puga (2012)	1970 - 2007
Sectoral diversification - Theil index overall (value added)	Value added share based overall Theil index, a measure of overall sectoral diversification	Constructed by the authors using INDSTAT2 (2012) sectoral share data and following Cadot et al. (2011) methodology	1970 - 2007
Sectoral diversification - Theil index within (value added)	Value added share based within component of Theil index, a measure of intensive margin of sectoral diversification	Constructed by the authors using INDSTAT2 (2012) sectoral share data and following Cadot et al. (2011) methodology	1970 - 2007
Sectoral diversification - Theil index between (value added)	Value added share based between component of Theil index, a measure of extensive margin of sectoral diversification	Constructed by the authors using INDSTAT2 (2012) sectoral share data and following Cadot et al. (2011) methodology	1970 - 2007
Sectoral diversification - Theil index overall (employment)	Employment share based overall Theil index, a measure of overall sectoral diversification	Constructed by the authors using INDSTAT2 (2012) sectoral share data and following Cadot et al. (2011) methodology	1970 - 2007
Sectoral diversification - Theil index within (employment)	Employment share based within component of Theil index, a measure of intensive margin of sectoral diversification	Constructed by the authors using INDSTAT2 (2012) sectoral share data and following Cadot et al. (2011) methodology	1970 - 2007
Sectoral diversification - Theil index between (Employment)	Employment share based between component of Theil index, a measure of extensive margin of sectoral diversification	Constructed by the authors using INDSTAT2 (2012) sectoral share data and following Cadot et al. (2011) methodology	1970 - 2007
Income (log)	Log of per capita income	Penn World Table version 7.1	1970 - 2007
Trade Openness	Sum of exports and imports as percentage of GDP	Penn World Table version 7.1	1970 - 2007
Population (log)	Log of population, measures country size	Penn World Table version 7.1	1970 - 2007
Oil dummy	A dummy representing twenty major oil exporting country	The World Bank	1970 - 2007
Colonial origin dummy	Dummies for colonial origin	La Porta et al. (1999)	1970 - 2007
Neighbor's income	Trade share weighted income of neighbors of a country; an instrument for income of the country a la Acemoglu et al. (2008).	Constructed by the authors following Acemoglu et al. (2008) methodology	1970 - 2007

Predicted trade	Predicted changes in bilateral trade owing to foreign natural disasters. An instrument for trade openness as proposed by Felbermayr & Groschl (2013).	Felbermayr & Groschl (2013)	1970 - 2007

Table A2: Variables and Data Sources – Within-Country Analysis for the US

Variable	Variable description	Data source	Availability
Ruggedness	Terrain ruggedness of the US states	Constructed by the authors following Nunn & Puga (2012) methodology	1997 - 2011
Sectoral diversification – Theil index overall	Manufacturing sector diversification index as measured by Theil index for the US states	Constructed by the authors using sectoral value added shares of 18 manufacturing sectors, and Cadot et al. (2011) methodology	1997 - 2011
Geographic concentration – Theil index overall	Measure of overall spatial concentration of economic activities across the US states	Constructed by the authors using night-time satellite light data and, Cadot et al. (2011) methodology	1997 -2011
Geographic concentration – Theil index within	Measure of within component (intensive) of spatial concentration of economic activities across the US states	Constructed by the authors using night-time satellite light data, Cadot et al. (2011) methodology	1997 -2011
Geographic concentration – Theil index between	Measure of between (extensive) component of spatial concentration of economic activities across the US states	Constructed by the authors using night-time satellite light data, and Cadot et al. (2011) methodology	1997 -2011
Income (log)	Per capita Gross State Product (GSP)	Sourced from the Regional Product Division, Bureau of Economic Analysis (BEA), the U.S. Department of Commerce (June 6, 2013). Available on BEA's website at https://www.bea.gov/regional/	1997 - 2011
Export (% of GSP)	A proxy for trade openness of the US states	Sourced from the Foreign Trade Division, the U.S. Census Bureau. Available at http://tse.export.gov/TSE/TSEReports.aspx?DATA=SED	1999 - 2011
Population (log)	Log population measuring size of the US states	The Census U.S. Intercensal County Population Data, 1970-2014, The National Bureau of Economic Research (NBER). Available at http://www.nber.org/data/census-intercensal-county-population.html	1997 - 2011

Oil dummy	A dummy representing ten largest oil producing states	US Energy Information Administration (EIA), available at https://www.eia.gov/state/rankings/#/series/46	1997 - 2011

Table A3: Descriptive Statistics of Cross-Country Data, 1970-2007

Variable	Country	Obs.	Mean	Median	S. D.	Max	Min
Ruggedness	189	7182	1.35	0.96	1.27	6.74	0.003
Sectoral diversification - Theil index overall (value-added based)	139	3192	0.86	0.76	0.43	2.86	0.24
Sectoral diversification - Theil index between (value-added based)	127	2493	0.33	0.25	0.26	2.04	0.04
Sectoral diversification - Theil index within (value-added based)	130	2778	0.55	0.47	0.30	1.88	0.14
Sectoral diversification - Theil index overall (employment based)	150	3558	0.82	0.72	0.40	2.95	0.22
Sectoral diversification - Theil index between (employment based)	140	2709	0.33	0.25	0.24	2.04	0.04
Sectoral diversification - Theil index within (employment based)	143	3100	0.51	0.46	0.27	2.37	0.06
Income (log)	184	6403	8.53	8.53	1.16	11.62	5.03
Trade openness	184	6403	79.68	68.88	50.46	456.56	1.09
Population (log)	182	6916	8.42	8.56	2.00	14.09	3.08

Table A4: Descriptive Statistics of the US State-Level Data, 1997-2011

Variable	Obs.	Mean	Median	S.D.	Min	Max
Ruggedness	765	0.83	0.51	0.69	0.06	2.59
Sectoral diversification (Theil index overall)	765	0.52	0.45	0.27	0.12	1.62
Geographic Concentration (Theil index overall)	764	1.69	1.60	0.92	0.00	5.24
Geographic Concentration (Theil index within)	764	1.14	1.06	0.53	0.00	3.34
Geographic Concentration (Theil index between)	764	0.55	0.50	0.45	0.00	2.97
Export (% gross state product)	663	6.62	5.87	3.87	0.58	28.10
Population (log)	765	15.06	15.19	1.04	13.08	17.44
Per capita Gross State Product (log)	765	10.60	10.57	0.25	10.15	11.91