DISCUSSION PAPER SERIES

DP14157

THE SHRINKING ADVANTAGE OF MARKET POTENTIAL

Marius Brülhart, Klaus Desmet and Gian-Paolo Klinke

INTERNATIONAL TRADE AND REGIONAL ECONOMICS

ISSN 0265-8003

THE SHRINKING ADVANTAGE OF MARKET POTENTIAL

Marius Brülhart, Klaus Desmet and Gian-Paolo Klinke

Discussion Paper DP14157
Published 29 November 2019
Submitted 25 November 2019

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

THE SHRINKING ADVANTAGE OF MARKET POTENTIAL

Abstract

How does a country's economic geography evolve along the development path? This paper documents recent employment growth in 18,961 regions in eight of the world's main economies. Overall, market potential is losing importance, and local density is gaining importance, as correlates of local growth. In mature economies, growth is strongest in low-market-potential areas. In emerging economies, the opposite is true, though the association with market potential is also weakening there. Structural transformation away from agriculture can account for some of the observed changes. The part left unexplained by structural transformation is consistent with a standard economic geography model that yields a bell-shaped relation between trade costs and the growth of centrally located regions.

JEL Classification: O18, R11, R12

Keywords: regional growth, market potential, structural transformation, economic development

Marius Brülhart - marius.brulhart@unil.ch
University of Lausanne and CEPR

Klaus Desmet - kdesmet@smu.edu
Southern Methodist University and CEPR

Gian-Paolo Klinke - gian-paolo.klinke@bfs.admin.ch
Swiss Federal Statistical Office

Acknowledgements
We thank Da Li for excellent research assistance and Kentaro Nakajima, as well as seminar participants at the Philadelphia Fed, RIETI Tokyo, Groningen and Georgetown for helpful comments. Funding from the Swiss National Science Foundation (grant 147668) is gratefully acknowledged.

The Shrinking Advantage of Market Potential*

Marius Brülhart†
University of Lausanne

Klaus Desmet‡
Southern Methodist University

Gian-Paolo Klinke§
Swiss Federal Statistical Office

November 25, 2019

Abstract

How does a country's economic geography evolve along the development path? This paper documents recent employment growth in 18,961 regions in eight of the world's main economies. Overall, market potential is losing importance, and local density is gaining importance, as correlates of local growth. In mature economies, growth is strongest in low-market-potential areas. In emerging economies, the opposite is true, though the association with market potential is also weakening there. Structural transformation away from agriculture can account for some of the observed changes. The part left unexplained by structural transformation is consistent with a standard economic geography model that yields a bell-shaped relation between trade costs and the growth of centrally located regions.

Keywords: regional growth, market potential, structural transformation, economic development

JEL Codes: R11, R12, O18

*We thank Da Li for excellent research assistance and Kentaro Nakajima, as well as seminar participants at the Philadelphia Fed, RIETI Tokyo, Groningen and Georgetown for helpful comments. Funding from the Swiss National Science Foundation (grant 147668) is gratefully acknowledged.

†Department of Economics, Faculty of Business and Economics (HEC Lausanne), University of Lausanne, 1015 Lausanne, Switzerland; and CEPR. (`Marius.Brulhart@unil.ch`).

‡Department of Economics and Cox School of Business, Southern Methodist University, Dallas, Texas, United States; NBER and CEPR. (`kdesmet@smu.edu`).

§Federal Statistical Office, 2010 Neuchâtel, Switzerland. (`gian-paolo.klinke@bfs.admin.ch`).

1 Introduction

Geographic proximity to markets is often seen as an important locational advantage. Regions with access to large pools of nearby consumers save on transport costs and are hence attractive places for firms to locate. Examples of regions with high market potential include Northwest Europe, the U.S. Northeast and Midwest, China's eastern seaboard, and the coastal regions of Brazil and India.

The role of market potential in shaping the spatial distribution of economic activity changes as an economy develops. Early in this process, when agriculture is an economy's main activity and trade costs are high, a location's inherent productivity plays an outsized role and market potential is less relevant. As development proceeds, at least two important changes affect an economy's geography. First, structural transformation shifts employment from agriculture to manufacturing, and later to services. This diminishes the importance of local land suitability, and increases the role of connectivity to other places (Henderson *et al.*, 2018). As a result, we might expect central locations to grow faster. Second, falling transport costs first improve the fortunes of centrally located places, and later benefit more peripheral locations. In today's world of ever-lower trade costs, more isolated places experience greater relative improvements in access to markets. As a result, we might expect peripheral locations to grow faster.

To illustrate how the importance of market potential might change with economic development, compare the U.S. and Mexico. The five U.S. counties with the highest rates of job growth between 1990 and 2010 were Tunica, Miss., Douglas, Col., Forsyth, Ga., Dawson, Ga., and Williamson, Tex.[1] While all five are part of large metro areas (Memphis, Denver, Atlanta and Austin), they are relatively far from the high-market-potential areas of America's Northeast and Midwest. When focusing on counties with employment above 400,000, we observe a similar pattern: Maricopa (Phoenix) tops the employment growth ranking, despite its isolation (27th percentile in terms of market potential), while the lowest employment growth was recorded in Detroit (Wayne county), despite its centrality (99th percentile in terms of market potential). The picture is quite different in Mexico, where the five fastest-growing municipalities are in well-connected areas close to Monterrey, Mexico City and the Yucatan Coast (85th percentile in terms of market potential).[2] While we see opposing patterns when focusing on the importance of market potential, the same is not true for local density. In both countries growth concentrated in areas of relatively high employment density. This suggests that market potential and local density differ in their relevance for growth.

In this paper, we explore whether this comparison generalizes by empirically analyzing the changing importance of market potential and local density for employment growth across 18,961

[1] We exclude micro counties, defined as having 1990 employment below 300.

[2] See Figure 2 for the positions in the sample distribution of Monterrey, Mexico City, Cozumel (Yucatan, Mexico) as well as of Detroit and Maricopa county (United States).

regions over the period 1990-2010. These regions cover eight of the world's main economies and account for three-quarters of global GDP. In its scope and detail, our analysis constitutes the most comprehensive effort thus far to document local employment growth patterns across the globe. By including both more developed and less developed economies, we uncover meaningful differences in the spatial growth patterns of countries at different stages of development, and we relate those differences to structural transformation and declining trade costs.

Our empirical analysis identifies two novel stylized facts. First, for the world as a whole, when comparing the 2000s to the 1990s, market potential is becoming less important, and local density is becoming more important as a correlate of employment growth. Second, when comparing economies at different levels of development, we uncover a stark contrast: whereas in emerging economies growth tends to be greater in high-market-potential areas, in mature economies the opposite is true and growth concentrates in low-market-potential areas. This shows that the illustrative examples comparing Mexico to the U.S. hold more generally across many regions and countries. Taken together, the two stylized facts are consistent with a secular decline of market potential as a locational advantage.

What might account for the changing spatial distribution of employment, and in particular for the two stylized facts? A first possible determinant is structural transformation. In emerging economies, manufacturing and services are both disproportionately concentrated in high-density, high-market-potential areas. The same is true for services in mature economies. Hence, structural transformation, away from agriculture in emerging economies and towards services in mature economies, tends to shift employment to high-density, high-market-potential areas in both types of economies. Using a simple accounting approach in the spirit of Michaels, Rauch and Redding (2012), we show that structural transformation significantly contributes to explaining the overall move toward higher-density areas in the world, as well as for the shift to high-market-potential regions in emerging economies. However, it is unable to account for the weakening importance of market potential in the world as a whole, and in particular for the lower growth in the high-market-potential regions of mature economies.

A second possible determinant of the world's changing economic geography is the well documented decline in transport costs (World Trade Organization, 2008; Redding and Turner, 2015; Hummels, 2007). To form a prior on how we would expect falling trade costs to affect the location of employment, we propose a standard economic geography model with one central and two peripheral locations. When trade costs are prohibitive, there is no advantage to centrality, and employment is equally spread across the three locations. As trade costs drop, the central location initially gains employment at the expense of the peripheral locations, because it benefits disproportionately from improved access to the other locations. However, as trade costs continue to fall, the central location starts to lose its proximity advantage. The peripheral locations suffer increasingly less from worse market access, and become once again attractive because of their lower congestion. This yields a well known result in economic geography: a bell-shaped relation between the employment share of the central location and the level of trade costs.

While we do not have precise measures of the change in transport costs across space to directly test the model, we do know that mature economies started off at a lower level (World Bank, 2009, UNCTAD, 2012). Hence, to the extent that mature economies experienced a drop in already low transport costs, our simple model is consistent with the observed negative relation between initial market potential and subsequent growth in these economies. This does not imply that market potential is no longer an advantage. Instead, it reflects a shrinking advantage of centrality as market potential is growing relatively faster in peripheral places. We confirm this insight by analyzing the effect of a simulated uniform decline in transport costs in mature economies. We find that, as predicted by the model, low-market-potential regions experience greater improvements in market potential, and hence faster growth.

This paper contributes to the empirical literature that documents spatial economic trends across the world. Recent papers in that area include Gennaioli *et al.* (2014) who study income convergence across 1,528 sub-national regions in 83 countries, as well as Henderson, Storeygard and Weil (2012) and Henderson *et al.* (2018), who rely on satellite imagery of night lights to analyze local economic growth and the distribution of economic activity at a high spatial resolution across the globe. In contrast to our work, these papers do not draw on sector-level information, nor do they explore the changing role of market potential across different economies.

The sector-specific nature of our data allows us to assess the role of structural transformation in shaping the changing economic geography of the world. Previous work by Michaels, Rauch and Redding (2012, 2018) and Desmet and Rossi-Hansberg (2009, 2014) has analyzed the role of structural change in the long-run evolution of the U.S. economic geography. While our paper covers a shorter time span, it encompasses the major economies around the world. In distinguishing between emerging and mature economies, we further enhance our understanding of how development shapes an economy's geography (Desmet and Henderson, 2015).

Our paper also relates to theoretical and empirical work on the bell-shaped relation between transport costs and spatial concentration. Tabuchi (1998) generates this non-monotonicity result in a setting with mobile labor by combining urban agglomeration economies and congestion from intra-city commuting.[3] In contrast to our work, his two-city model cannot adequately disentangle market potential from local density, a crucial distinction in our analysis. The few quantitative or

[3] See also Krugman and Venables (1995) and Puga (1999) for models on the relation between transport costs and real wage differences across regions in models without geographic mobility of workers. Fujita and Thisse (2006) offer a corresponding analysis considering changes in communication costs in addition to changes in transport costs.

empirical papers that have explored the existence of a non-monotonic relation between trade costs and geographic concentration have focused on one country or region of the world. Examples include Kim (1995), Forslid, Haaland and Midelfart-Knarvik (2002) and Combes *et al.* (2011). Our paper is the first to provide both temporal and cross-sectional empirical evidence of this bell-shaped relation for a wide variety of regions across the globe.

There is an extensive literature on the importance of market potential for development. In a seminal paper, Redding and Venables (2004) estimate the role of market access in determining the cross-country variation in per capita income. In other cross-country studies, Head and Mayer (2011) and Jacks and Novy (2018) analyze the role of market potential as a determinant of growth in income per capita. The importance of market potential for development and growth is related to the role of distance for trade, and it features prominently in gravity models (Head and Mayer, 2004; Disdier and Head, 2008). Our paper also analyzes the association of market potential with growth, and we are first in documenting its weakening importance.

A key contribution of the paper is the construction of a sectoral dataset that spans the period 1990 to 2010 and covers sub-national regions in eight of the world's most important economies: Brazil, China, Europe (East and West), India, Japan, Mexico and the United States. The sectorally disaggregated employment data come from national censuses. In its spatial, sectoral and time coverage, the database we develop constitutes, to our knowledge, the largest source of census-based employment data currently available. In addition, given the importance of the world's densest metro areas for our analysis, we expend considerable effort in making those locations comparable across different countries. To that end, we combine information on nighttime lights and land usage to determine when different contiguous regions should be merged to form one urban area.

The rest of the paper is organized as followed. Section 2 describes the data sources for all countries, and explains the algorithm we use to define the world's densest areas in a consistent way across countries. Section 3 reports the empirical findings. Sections 4 and 5 relate our findings to structural transformation and the drop in transport costs. Section 6 concludes.

2 Data

This section discusses the data we use to investigate the effect of employment density and market potential on employment growth across 18,961 regions of the world over the period 1990-2010.[4] The analysis covers 34 countries, which together make up 55% of world population and 74% of global GDP in 2010.[5] For part of our analysis we split the sample into emerging and mature economies.

[4] Note that we use the term 'effect' is in a broad sense that need not imply direct causation.

[5] Data on population and GDP (at current market prices) are taken from http://data.worldbank.org/indicator, last accessed November 2019.

When doing so, we define the group of emerging economies to be Brazil, Central and Eastern Europe, China, India, and Mexico; and we define the group of mature economies to be Japan, the United States, and Western Europe.[6]

2.1 Geographic Units

Depending on the country, our geographic units of observation consist of second- or third-level administrative divisions. We use municipalities for Brazil, Japan and Mexico, counties for China and the U.S., subdistricts for India, and NUTS3 regions for Europe. To make these geographic units as comparable as possible across time and across space, we need to deal with mergers and break-ups of administrative regions, as well as with differences in the granularity of the data across countries.

Mergers and break-ups. Over a span of twenty years there are inevitably certain sub-national units that get merged and others that get divided. We aggregate sub-national units in such a way that the resulting geographic entities are consistent over the entire time period. Whenever a break-up occurs, we merge subdivisions in subsequent years, and whenever a merger occurs, we aggregate subdivisions in previous years. The numbers of actual and consolidated regions by economy are provided in columns (1) and (2) in Table 1. Starting with 22,677 regions, consolidation lowers this number to 19,725.

Table 1: Administrative Subdivisions: Number of Regions

	(1) 2010 Delineation	(2) Consolidated	(3) Urban Correction	(4) Final Sample
Brazil	5,512	4,263	4,207	4,204
Central & Eastern Europe	314	314	311	311
China	2,375	2,297	2,269	2,268
India	6,083	4,614	4,545	4,541
Japan	1,817	1,721	1,514	1,431
Mexico	2,434	2,404	2,326	2,200
United States	3,138	3,108	3,066	3,066
Western Europe	1,004	1,004	954	940
Total	22,677	19,725	19,192	18,961

Column (2) corrects for break-ups and mergers. Column (3) takes into account the urban area correction: administrative subdivisions that are part of a common urban area are combined. Column (4) presents the final sample of observations which is used in the analysis. The difference between columns (3) and (4) is due to missing values and the exclusion of consolidated regions with a surface area lower than 20 square kilometers.

[6] Central and Eastern European countries are Bulgaria, Czech Republic, Estonia, East Germany, Hungary, Latvia, Lithuania, Poland, Romania, Slovakia, and Slovenia. Western European countries are Austria, Belgium, Cyprus, Denmark, West Germany, Spain, Finland, France, Greece, Ireland, Italy, Luxembourg, Malta, Netherlands, Norway, Portugal, Sweden, and the United Kingdom.

Urban areas. Administrative regions do not always capture labor markets adequately. In particular, large urban areas often consist of multiple administrative regions that form one unified labor market. Because of this, we adopt a methodology that aggregates administrative units whenever they are part of the same urban area.[7] To do so, we rely on two datasets with worldwide coverage and use a procedure to group contiguous administrative regions into one urban area when their densities warrant us to do so. We combine nighttime lights data from the Defense Meteorological Satellites Program – Operational Linescan System (DMSP–OLS) with land cover information from ESA's Globcover project to identify high-density areas which cover several administrative subdivisions. Administrative subdivisions covered by a common urban area are then merged into a single region, provided a minimum share of their areas are covered by the high-density area. Appendix A.2 gives further details of the exact procedure we follow. Column (3) in Table 1 displays the number of regions resulting from consolidating urban areas. It shows a reduction in their total number from 19,725 to 19,192. After dropping consolidated regions with an area of less than 20 square kilometers, as well as a few regions with missing employment data, we are left with the 18,961 regions in column (4) that form the core database used in our empirical analysis.[8] In Sections 4 and 5, where we use sector-level employment data, we moreover drop all observations for which reported total employment differs by more than 15% from the sum of reported employment across the three sectors (agriculture, manufacturing, services). This adjustment mainly affects India, where we lose 22% of regions, and Brazil, where we lose 15% of regions. It brings the total number of observations to 17,418 (see Table 5).

Defining urban areas in a consistent and comparable manner across economies also helps to ensure that our results are not driven by systematic differences in the size of administrative units across economies. For instance, the metropolitan division of New York City covers 11 counties, whereas much of Beijing is covered by the single county of Beijing Shi.[9] Failing to aggregate regions which belong to the same urban area in countries with a finer level of administrative disaggregation could lead to artificial differences in our empirical findings across countries.

Weighting and standardizing. As can be seen in Table 2, there are important differences across economies in the fineness of administrative subdivisions. For example, the average U.S. county

[7] Our sector-level employment data being available at the level of administrative regions, we can only correct in the sense of combining regions that together form a functional urban area, but we cannot decompose regions that are "too large". For urban area definitions based entirely on remotely sensed data, see Baragwanath-Vogel, Goldblatt, Hanson and Khandelwal (2018).

[8] We drop consolidated regions with an area of less than 20 square kilometers because they typically exhibit exceptionally high density but due to their small size are nevertheless economically unimportant (e.g. military bases).

[9] Data on New York according to the U.S. Office of Management and Budget (OMB) delineation of MSAs (2013 delineation), https://www.census.gov/geographies/reference-files/time-series/demo/metro-micro/delineation-files.html, last accessed July 2017.

has an area of 2,953 square kilometers, whereas the average Mexican municipality has an area of 872 square kilometers. To the extent that these differences are especially relevant for the upper tail of the distribution, we have already addressed this issue by adopting a consistent definition of urban areas across economies. Of course, these differences may also matter for the lower tail of the distribution. For example, 94% of the Brazilian municipalities had a density lower than the first quartile of Chinese counties in 1990. This is mainly because employment density in Brazil is less than one-tenth that of China. To avoid a situation where different density intervals correspond to different economies, we express our independent variables in terms of their percentile within each economy. Recall that economies correspond to countries, except in Europe, where we aggregate countries into either Western Europe or Central and Eastern Europe.

Differences in the fineness of administrative subdivisions also pose another problem: without further adjustment, our results are likely going to be driven by economies that happen to be subdivided into a large number of regions, simply because these economies make up a larger share of the observations. To avoid this pitfall, our regressions weight each economy equally. That is, the weight of each region of economy c is $w_c = \frac{1}{N_c}$, where N_c denotes the number of regions in economy c. Also, since average employment growth varies considerably across economies, for reasons that are often unrelated to their geography, we mean-deviate growth rates by economy.

2.2 Employment and Density

We rely on multiple statistical sources to construct our employment database of 18,961 regions. For example, for China we use county-level employment data from the Population Census, distributed by the University of Michigan China Data Center; for Mexico we use municipal-level employment data from the General Census of Population and Housing, published by the Mexican National Institute of Statistics and Geography; and for the U.S. we use county-level employment data from the County Business Patterns, published by the U.S. Census Bureau. To enable us to study employment growth across time, our database focuses on three years: 1990, 2000 and 2010. In addition to total employment, we also consider sectoral employment in agriculture, manufacturing and services.[10] Table 2 reports summary statistics of the employment data, and Appendix A.1 provides further details about the different sources.

For our measure of land area, we use administrative boundaries from the Global Administrative Areas Database (GADM) for all countries, with the exception of China (Michigan's China Data Center), the U.S. (Siczewicz, 2011) and India (ORGI, 2011). Given that we are interested in land area to get a measure of employment density, we exclude areas unfit for economic activity. To that end,

[10] Strictly speaking, our data allow us to distinguish between primary, secondary and tertiary employment. Hence, "agriculture" also contains forestry and fisheries, and "manufacturing" includes energy, mining and construction. Given the relatively minor employment weights of those neighboring sectors, we prefer to use the simpler terminology.

Table 2: Summary Statistics Employment

	Year	Employment and Area Mean	Std. Dev.	Median	Sectoral Shares Agric.	Manuf.	Serv.
All economies	1990	73,448	232,263	16,509	49.9	19.1	30.9
(18,961 obs.)	2000	87,177	295,831	20,072	45.8	19.0	35.2
	2010	97,044	358,880	22,537	36.8	22.5	40.7
Administrative unit	Area	1,778	6,523	712			
Mature	1990	60,201	289,702	11,983	4.2	28.1	67.7
(5,437 obs.)	2000	66,253	304,685	13,704	2.7	23.6	73.7
	2010	67,254	302,517	13,185	2.2	19.0	78.8
Administrative unit	Area	2,345	9,023	1,229			
Emerging	1990	78,763	204,508	19,147	64.1	16.4	19.5
(13,524 obs.)	2000	95,572	291,794	24,650	57.7	17.7	24.6
	2010	108,997	378,497	28,541	45.4	23.3	31.3
Administrative unit	Area	1,551	5,174	558			
Brazil	1991	13,157	112,630	4,346	23.1	24.8	52.1
(4,204 obs.)	2000	15,616	125,059	4,992	18.7	21.8	59.5
	2010	22,248	178,019	6,419	15.2	22.0	62.8
Municipality	Area	1,990	8,406	466			
Central & Eastern Europe	1991	182,486	176,506	140,991	21.6	34.9	43.5
(311 obs.)	2000	163,980	164,224	127,113	22.2	28.6	49.2
	2010	164,297	181,282	118,022	12.9	28.8	58.3
NUTS3	Area	3,763	2,979	3,384			
China	1990	285,373	354,255	206,068	72.2	15.0	12.9
(2,268 obs.)	2000	330,156	567,629	237,874	64.2	16.9	18.9
	2010	360,498	759,377	250,530	48.1	24.1	27.9
County	Area	2,915	4,129	1,933			
India	1991	62,527	117,468	40,566	65.6	13.0	21.4
(4,541 obs.)	2001	87,003	168,980	57,902	60.0	16.2	23.8
	2011	103,751	221,944	69,123	52.4	21.6	26.0
Subdistrict	Area	643	672	425			
Japan	1990	42,646	403,704	11,855	7.2	34.0	58.8
(1,431 obs.)	2000	43,505	413,555	11,594	5.1	30.4	64.5
	2010	41,291	406,675	10,232	4.2	25.7	70.1
Subprefecture	area	255	293	160			
Mexico	1990	10,512	109,303	2,695	23.4	28.8	47.8
(2,200 obs.)	2000	15,140	151,738	3,532	16.3	28.7	55.0
	2010	19,165	179,822	4,125	13.2	25.1	61.7
Municipality	Area	872	2,224	264			
United States	1990	37,030	212,462	6,692	0.2	22.8	77.0
(3,066 obs.)	2000	44,152	230,477	8,303	0.1	18.2	81.6
	2010	43,425	219,688	8,116	0.1	13.0	86.9
County	Area	2,953	11,320	1,625			
Western Europe	1990	160,917	280,456	95,573	6.0	29.7	64.4
(940 obs.)	2000	171,271	301,316	99,349	4.0	25.4	70.6
	2010	182,623	323,343	104,599	3.1	21.4	75.4
NUTS3	Area	3,509	6,593	1,501			

Area is expressed in km². Uncategorized employment is excluded in the computation of sectoral shares.

we rely on the European Space Agency (ESA) Globcover project (Version 2.3, 2009) and its Land Cover gridded map which categorizes land cover into 22 classes at a spatial resolution of 300m. We define the relevant area as the land covered by all types of surface but water bodies, bare areas and permanent snow and ice.[11] Using data on employment and on area, Figure 1 panel (a) depicts employment density across our sample economies. The color-coding is based on economy-specific deciles.

2.3 Market Potential

A location's attraction as a place to produce depends partly on the market it can access in its own location and in all other locations, a concept Harris (1954) refers to as market potential. Intuitively, the market potential of location i depends (i) positively on the income of all locations, Y_j, and (ii) negatively on the cost of accessing all locations, d_{ij}. Based on this, we can define a location's market potential as the access-cost-adjusted sum of aggregate income in all other locations:

$$NMP_i = \sum_{j \in \mathcal{J}} Y_j d_{ij}^{-\gamma}, \tag{1}$$

where $\mathcal{J}$ is the set of all locations and γ is the rate at which the contribution of other locations to market potential decays with the cost of access. Head and Mayer (2004) refer to (17) as *nominal* market potential, hence our notation *NMP*. The related concept of *real* market potential adjusts for differences in the price index across locations. The simple economic geography model in Appendix B discusses this in further detail.

As for the value of the decay parameter, Disdier and Head (2008) carry out a systematic analysis of 1,467 estimates of γ in 103 papers, and find a range from -0.04 to 2.33, with a mean of 0.91 and a median of 0.87. We therefore set γ equal to 1, so that our expression of market potential simplifies to

$$NMP_i = \sum_j \frac{Y_j}{d_{ij}}. \tag{2}$$

This is equivalent to the market potential expression in Harris (1954). We now discuss how we get estimates of Y_j and d_{ij}.

When measuring market potential in the data, we use worldwide nighttime lights to proxy for income.[12] For the purpose of estimating the contribution of different locations to a region's market potential, the geographic units we have used so far are too coarse. Given the steep spatial decay

[11] The surface of bare areas is characterized by hardpans, gravels, bare rock, stones, boulders, sandy desert or salt hardpans. Considering the case of the USA, 97% of the territory is included. The excluded surface consists of water bodies (1.6%), permanent snow and ice (1.1%) and bare areas (0.3%)

[12] We approximate 1990 market potentials with lights data for 1996, the earliest available year. For 2000 and 2010, we can compute market potentials from lights measurements in the corresponding years.

in (2), the within-region geographic distribution of economic activity is relevant. For example, whether a neighboring region's economic activity is concentrated in the area closest to or farthest away from the region under consideration makes an important difference. The importance of the within-region distribution is even more obvious when we consider the market potential coming from the own region's economic activity. To address these concerns when estimating the market potential of region i in economy c, we discretize economy c into 6′ by 6′ cells, indexed by $h \in \mathcal{H}_c$, and we discretize the rest of the world in 1° by 1° grid cells, indexed by $g \in \mathcal{G}_c$.[13] The market potential of grid cell h in country c is then:

$$NMP_{hc} = \sum_{h' \in \mathcal{H}_c} \frac{Y_{h'}}{d_{hh'}} + \sum_{g \in \mathcal{G}_c} \frac{Y_g}{d_{hg}},$$

where the first term measures market potential originating in the own economy and the second term measures market potential from the rest of the world. For each region i in economy c, we then take the average market potential of all grid cells with centroids inside the region, weighted by their share of the nighttime lights of the entire region:

$$NMP_{ic} = \frac{\sum_{hc \in \mathcal{H}_{ic}} NMP_{hc} Y_{hc}}{\sum_{hc \in \mathcal{H}_{ic}} Y_{hc}} \tag{3}$$

where $\mathcal{H}_{ic}$ is the set of grid cells with centroids located in region i of economy c. This measure can be interpreted as the average market potential faced by a firm in region i of country c.

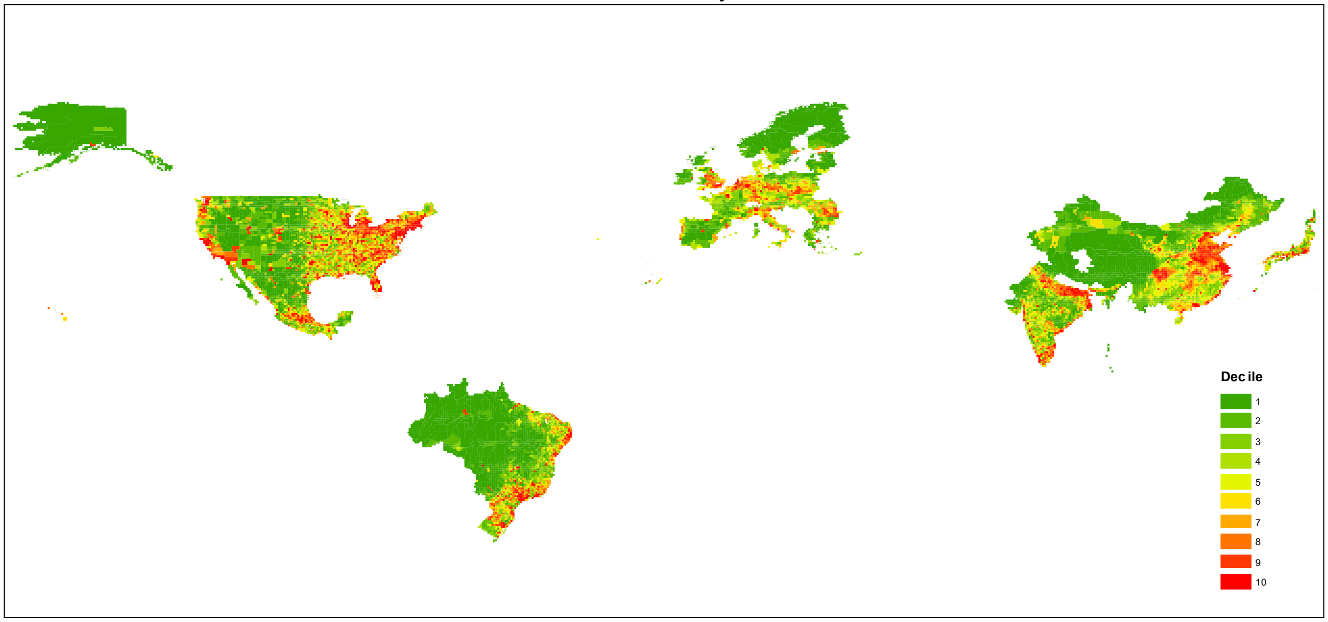
To estimate bilateral trade costs $d_{hh'}$, we use information on major roads, other roads, railroads and water to attach a cost to each grid cell. We then apply the method in Desmet, Nagy and Rossi-Hansberg (2018) and Allen and Arkolakis (2014) to create a cost surface, and use the 2D Fast Marching Algorithm to calculate $d_{hh'}$. The cost of a grid cell trading with itself is calculated as the average travel cost to the center of a disk with the same area. As shown by Head and Mayer (2000), this cost will be $0.67\sqrt{area/\pi}$ times the average cost per kilometer. Since the distance distortion in any projected coordinate system is quite large when we consider two locations that are far apart, we use the 3D Fast Marching Algorithm to calculate d_{hg}. As the cost surface in 2D and 3D Fast Marching exercises have different resolutions, the cost function in the 2D Fast Marching exercise is normalized to match the costs per kilometer in the 3D Fast Marching exercise, conditioning on the geographic features.

Figure 1 panel (b) depicts market potential across our sample economies. Once again, the color-coding is based on economy-specific deciles. The world's high-market-potential regions can be found in the Northeast and the Midwest in the U.S., northwestern Europe, the Gulf Coast of Mexico stretching to the nation's capital, Brazil's coastline area, China's eastern seaboard, India's coasts

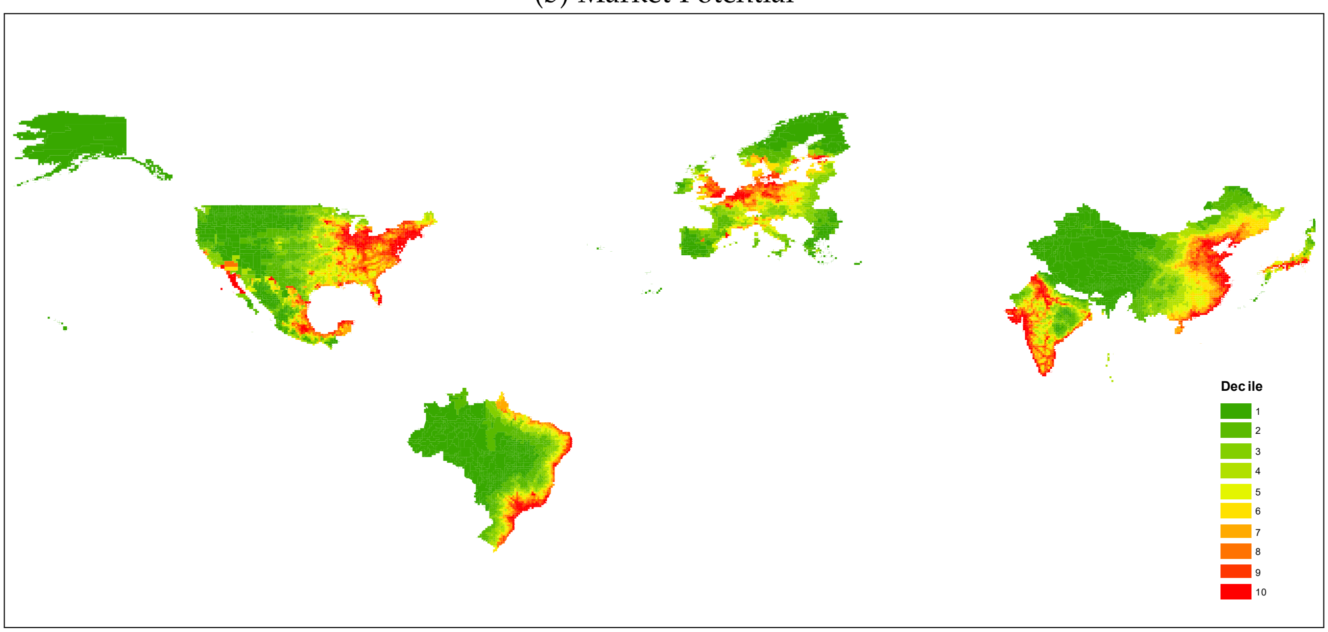
[13] We take a coarser grid cell for market potential from the rest of the world to keep computing time manageable.

Figure 1: The Sample Distribution of Density and Market Potential in 1990

(a) Density



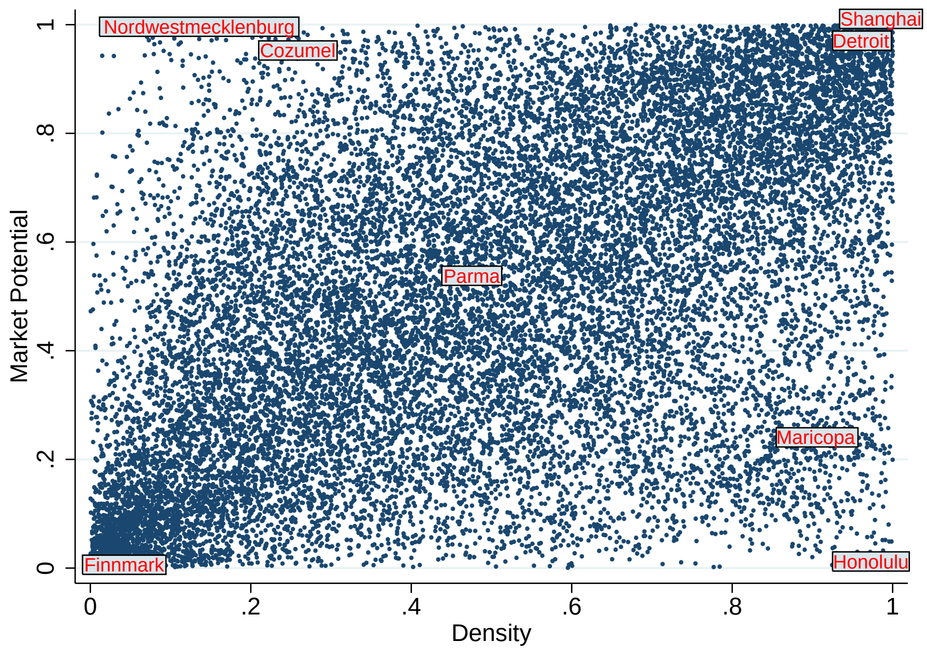
(b) Market Potential



Panel (a) shows regional population densities, binned in economy-specific deciles (data for 1990). Panel (b) shows regional market potential, binned in economy-specific deciles. Measures of market potential include own-region effects and are based on worldwide gridded nightlights data for 1996. Given the within-economy binning applied to both maps, color codes are comparable only within economies but not between them. We distinguish 8 economies: Brazil, Central and Eastern Europe, China, India, Japan, Mexico, United States, Western Europe.

and the Ganges Valley, and the Kyoto-Tokyo area in Japan. When comparing density with market potential, the world's high-density areas are often located within high-market-potential regions, but not always. Examples of high-density locations in areas of overall low market potential that are easily visible in the maps of Figure 1 include Denver and Salt Lake City in the U.S., Manaus in Brazil, Madrid in Spain, and Jiuquan in China.

Figure 2: The Sample Distribution of Density and Market Potential

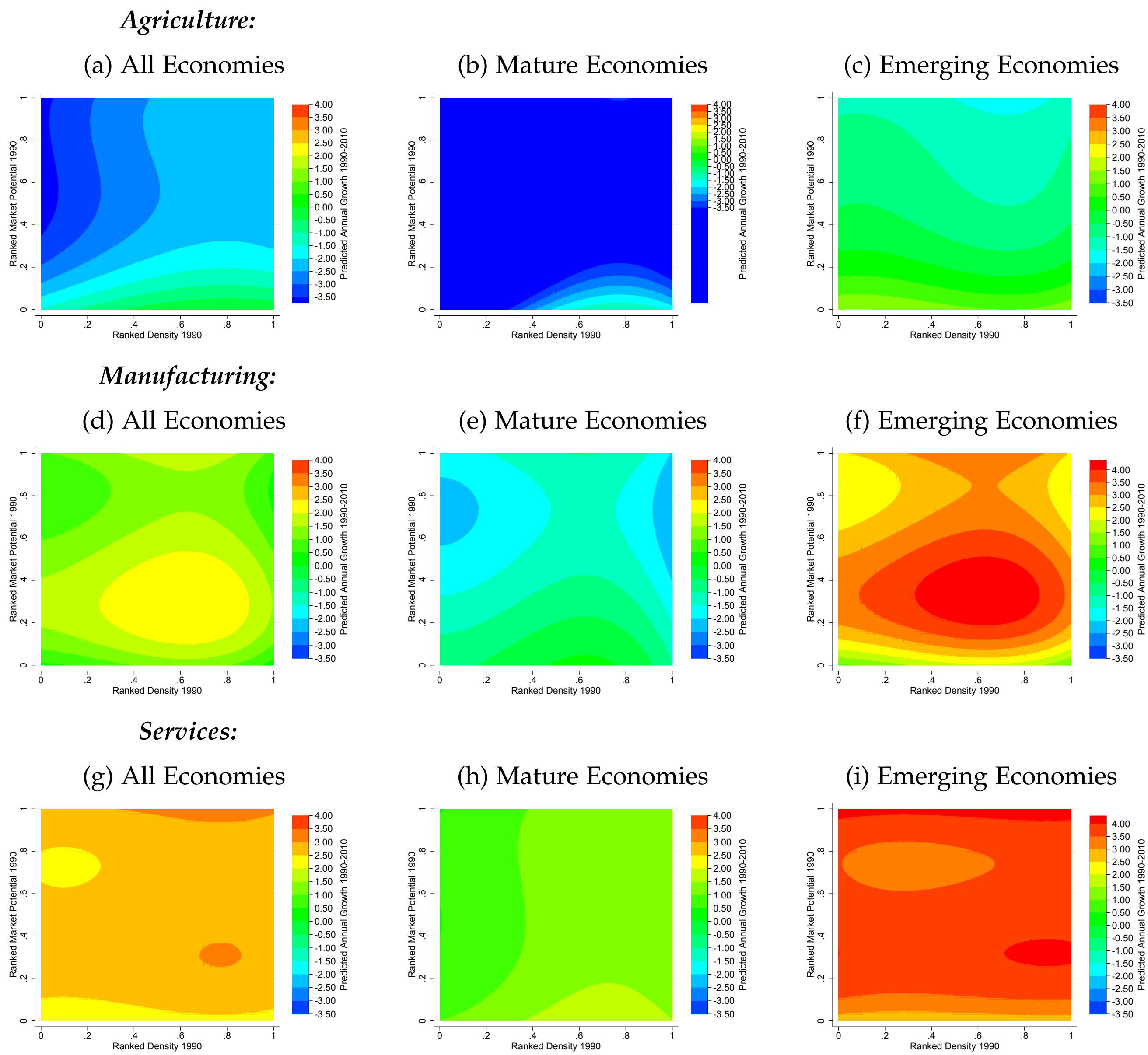


Density and market potential ranked within each economy and normalized to the range [0,1]. Data for 1990.

A further illustration of the joint distribution of density and market potential in our sample is given by the scatter plot of Figure 2. It shows that the two measures are correlated, as the scatter is at its densest in the vicinity of the diagonal. Indeed, the correlation between the (ranked and normalized) densities and market potentials is 0.54. While positive and significant, this correlation is far below unity. As can be seen in Figure 2, a large number of regions score relatively high on one variable but relatively low on the other. We therefore can exploit identifying variation that covers the entire density-market potential spectrum.

Finally, we illustrate some basic features of our data by documenting where sector-level growth took place over the 1990-2010 sample period. To this end, Figure 3, plots heat maps based on cubic regressions of sector-level employment growth against market potential and density. Unlike in our subsequent analysis, growth is not demeaned – allowing us to show differences in absolute growth rates across sectors and economies. In the remainder of the paper, we instead focus on reallocations across regions given an economy-sector average growth rate. The z-axis scale is held constant across the nine panels, making the color coding directly comparable.

Figure 3: Sector-Level Growth, 1990-2010: Third-Order Polynomial Prediction



Several findings emerge from these heat maps. First, employment growth was strongest in emerging economy service sectors and weakest (indeed negative) in mature economy agriculture. Second, in mature economies the growth patterns were similar across all sectors: employment grew relatively fast in low-market-potential, medium-high-density locations, not just in services, but also in agriculture and manufacturing. Third, the highest growth rates of all are observed for services in high-market-potential emerging-economy regions. Emerging-economy manufacturing employment, instead, appears to be shifting toward medium-market potential locations.

3 Market Potential, Local Density and Employment Growth

In this section we empirically analyze employment growth across 18,961 regions of the world for the decades spanning 1990 to 2010. We focus on the relation between market potential, local employment density and employment growth. We explore how the importance of market potential and local employment density has changed over time, and how this differs across emerging and mature economies.

The regressions of regional employment growth on regional market potential and regional employment density take the general form:

$$\frac{E_{i,c,t+1} - E_{i,c,t}}{E_{i,c,t}} - \frac{E_{c,t+1} - E_{c,t}}{E_{c,t}} = f(\widehat{E}_{i,c,t}) + g(\widehat{NMP}_{i,c,t}) + u_{i,c,t+1}, \tag{4}$$

where $\frac{E_{i,c,t+1}-E_{i,c,t}}{E_{i,c,t}} - \frac{E_{c,t+1}-E_{c,t}}{E_{c,t}}$ is the growth rate of employment in region i of economy c between years t and $t+1$ demeaned by the growth rate of employment in economy c between years t and $t+1$; $f(\widehat{E}_{i,c,t})$ is a linear, quadratic or cubic function of $\widehat{E}_{i,c,t}$, the employment density in region i of economy c in year t expressed as an economy-specific percentile; $g(\widehat{NMP}_{i,c,t})$ is a linear, quadratic or cubic function of $\widehat{NMP}_{i,c,t}$, the nominal market potential of region i in economy c in year t expressed as an economy-specific percentile; and $u_{i,c,t+1}$ is a mean-zero stochastic term.

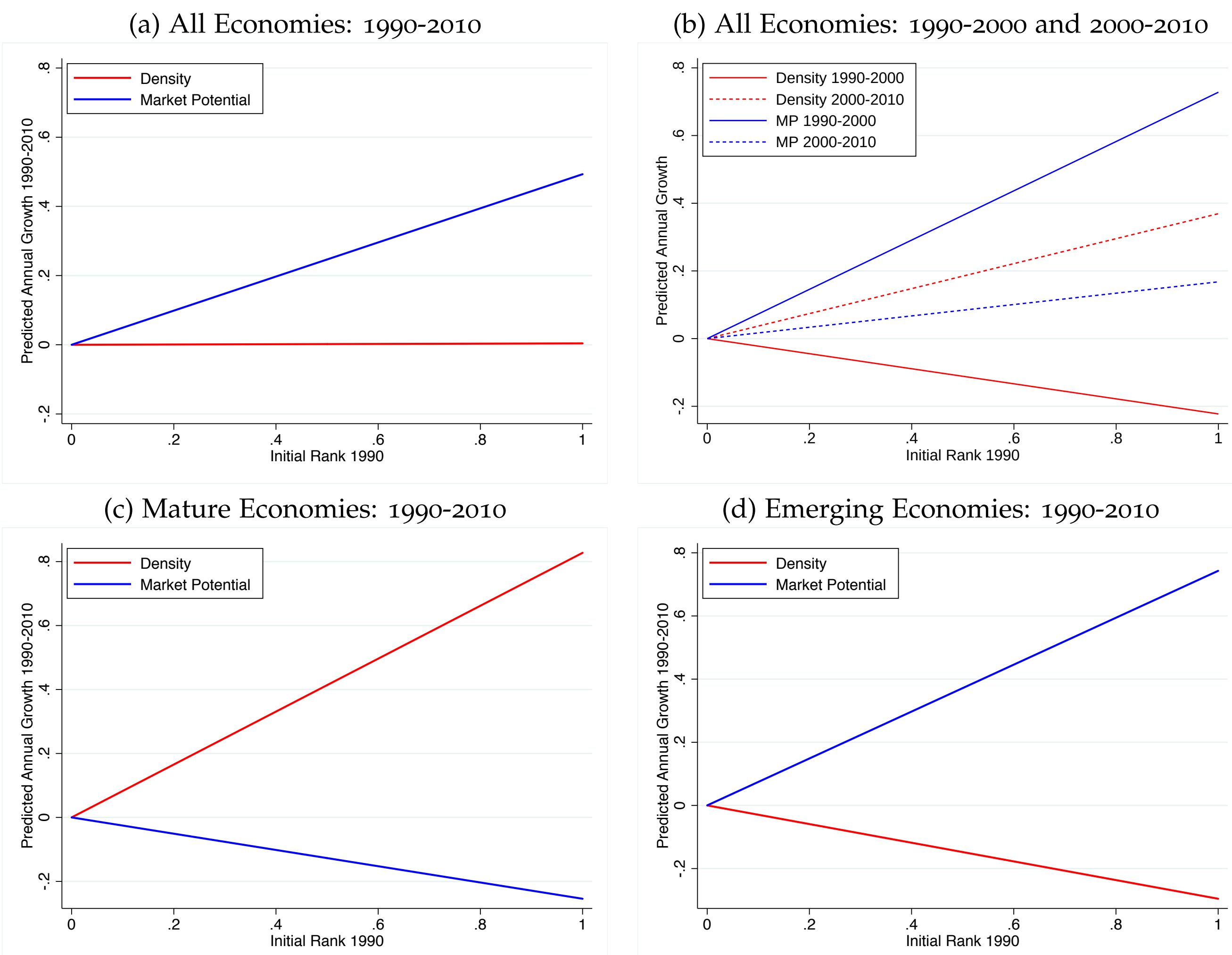
Table 3: Market Potential, Density and Growth: All Economies

Dependent Variable: Employment Growth							
	1990-2010			1990-2000		2000-2010	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Market Potential	0.475***	0.467***	-0.498	0.707***	0.032	0.137	-1.016***
	(0.055)	(0.081)	(0.303)	(0.095)	(0.413)	(0.094)	(0.341)
Market Potential (sq)			0.992***		0.698**		1.198***
			(0.256)		(0.354)		(0.289)
Density		0.015	-1.776***	-0.216***	-1.666***	0.386***	-2.136***
		(0.045)	(0.286)	(0.064)	(0.411)	(0.089)	(0.313)
Density (sq)			1.776***		1.438***		2.498***
			(0.251)		(0.357)		(0.274)
Constant	-0.238***	-0.241***	0.214***	-0.246***	0.105	-0.261***	0.344***
	(0.036)	(0.035)	(0.057)	(0.049)	(0.083)	(0.040)	(0.062)
Observations	18,961	18,961	18,961	18,961	18,961	18,961	18,961
R^2	0.009	0.009	0.022	0.007	0.011	0.006	0.023

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

All economies. We start by running a linear version of (4), and regress aggregate employment growth between 1990 and 2010 on market potential in 1990 for all 18,961 regions in our database, weighting economies by $w_c = \frac{1}{N_c}$. In column (1) of Table 3 we find market potential to have a positive effect on subsequent growth. When controlling for local density in column (2), the result is unchanged: market potential matters positively, whereas local density has essentially no effect. The estimated coefficient on market potential of 0.47 implies that annual employment growth in an economy's region with the highest market potential was on average about 0.47 percentage points higher than employment growth in that same economy's region with the lowest market potential. Given that the average annual employment growth rate in our sample was 0.8 percent, this is an economically sizable difference. To facilitate a visual interpretation of the results in column (2), Figure 4 panel (a) plots the linear predictions of the effect of market potential and density on aggregate employment growth .

Figure 4: Market Size, Density and Growth: Linear Prediction



When comparing the 1990s and the 2000s in columns (4) and (6) of Table 3, we find that the importance of market potential for employment growth has declined over time, becoming statistically insignificant in the latter decade. Density, however, goes from exhibiting a negative and statistically significant coefficient to exhibiting a positive and statistically significant effect. Once again, these changes are economically meaningful: whereas in the 1990s a region with the highest market potential is estimated to have benefitted from an annual growth rate 0.71 percentage points above that of a region with the lowest market potential, this advantage had dropped to 0.14 percentage points by the 2000s. In the case of density, the change goes in the opposite direction: in the 1990s a region with the highest density had an estimated growth rate 0.22 percentage points below that of a region with the lowest density; by the 2000s this disadvantage had turned to an advantage of 0.39 percentage points. The linear predictions of columns (4) and (6) are visually represented in Figure 4 panel (b). From this we can conclude that, for the world as a whole, market potential is losing importance, whereas density is gaining importance as a determinant of local growth. We refer to this finding as the first stylized fact we uncover in our analysis.

Comparing mature and emerging economies. When comparing growth patterns in mature and emerging economies, we uncover a stark difference. As reported in column (1) of Table 4, market potential has opposing effects in the two economies: positive for emerging and negative for mature. We refer to this as the second stylized fact of our empirical analysis. Moreover, the positive effect in emerging economies has weakened in the most recent decade, whereas the negative effect in mature economies has strengthened between 1990-2000 and 2000-2010 – confirming the generality of our first stylized fact. To be precise, in emerging economies the estimated annual growth advantage of the highest-market potential relative to the lowest-market potential region dropped from 0.95 to 0.45 percentage points between the 1990s and the 2000s, whereas in mature economies the disadvantage increased from -0.03 to -0.59 percentage points over the same time period. Density also has opposing effects in both types of economies, but in the reverse sense: initial density hurts employment growth in emerging economies, whereas it helps employment growth in mature economies.

When plotting the predicted linear effects of market potential and density on growth, panels (c) and (d) of Figure 4 show that mature and emerging economies are mirror images of each other. However, our second stylized fact focuses only on the opposing effects of market potential, because the opposing effects of density are no longer present in the 2000s. In particular, columns (3) and (5) of Table 4 show how the coefficient on density in emerging economies turns from negative in the 1990s to positive in the 2000s.

Non-linear effects. There is no reason why the relations between market potential, density and employment growth should necessarily be linear. When allowing for quadratic terms in our estimating equation (4), Tables 3 and 4 show that these additional terms are statistically significant, and

Table 4: Market Potential, Density and Growth: Mature and Emerging Economies

Panel a: ***Mature Economies***

Dependent Variable: Employment Growth						
	1990-2010		1990-2000		2000-2010	
	(1)	(2)	(3)	(4)	(5)	(6)
Market Potential	-0.294***	-1.620***	-0.077	-1.565***	-0.627***	-1.185***
	(0.075)	(0.295)	(0.198)	(0.363)	(0.080)	(0.241)
Market Potential (sq)		1.280***		1.413***		0.557***
		(0.260)		(0.325)		(0.208)
Density	0.856***	1.892***	0.842***	2.605***	1.064***	1.210***
	(0.076)	(0.296)	(0.101)	(0.375)	(0.083)	(0.307)
Density (sq)		-0.996***		-1.706***		-0.140
		(0.262)		(0.336)		(0.266)
Constant	-0.281***	-0.229***	-0.383***	-0.423***	-0.217***	-0.150**
	(0.041)	(0.068)	(0.052)	(0.086)	(0.043)	(0.067)
Observations	5,437	5,437	5,437	5,437	5,437	5,437
R^2	0.029	0.036	0.021	0.028	0.030	0.031

Panel b: ***Emerging Economies***

Market Potential	0.720***	-0.325	0.938***	0.391	0.422***	-0.573
	(0.105)	(0.417)	(0.124)	(0.578)	(0.126)	(0.442)
Market Potential (sq)		1.069***		0.570		1.044***
		(0.358)		(0.504)		(0.368)
Density	-0.290***	-3.478***	-0.628***	-3.648***	0.157	-3.951***
	(0.104)	(0.395)	(0.129)	(0.585)	(0.117)	(0.422)
Density (sq)		3.171***		3.005***		4.084***
		(0.355)		(0.515)		(0.377)
Constant	-0.215***	0.488***	-0.155***	0.436***	-0.290***	0.552***
	(0.050)	(0.082)	(0.073)	(0.124)	(0.060)	(0.087)
Observations	13,524	13,524	13,524	13,524	13,524	13,524
R^2	0.012	0.040	0.010	0.020	0.006	0.035

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

the explanatory power of the model, while still low, rises by a factor of up to four. Figure 5 depicts the predicted quadratic relations based on column (2) of Table 4. In mature economies, market potential has an overall negative effect, and density an overall positive effect, whereas for emerging economies, market potential has a positive effect, whereas density has a nonlinear effect, with both high- and low-density places growing faster than middle-density places.

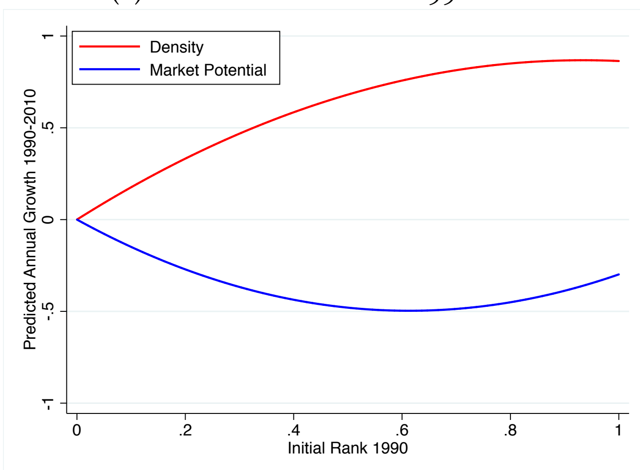
A more comprehensive way of visualizing our results is to use "heat maps" based on cubic regressions of growth on market potential and density.[14] Figure 6 plots, for all combinations of market potential and employment density, the predicted employment growth of regression (4) when

[14] For comparability, the color scale is held constant within sets of heat maps that illustrate predicted growth rates.

including linear, quadratic and cubic terms of the two explanatory variables. In the emerging economies, growth is concentrated in the high-market-potential regions, with both low- and high-density places doing well. In contrast, in mature economies, growth is highest in the low-market-potential, medium-high-density locations. This confirms our previous findings.

Figure 5: Market Size, Density and Growth: Quadratic Prediction

(a) Mature Economies: 1990-2010



(b) Emerging Economies: 1990-2010

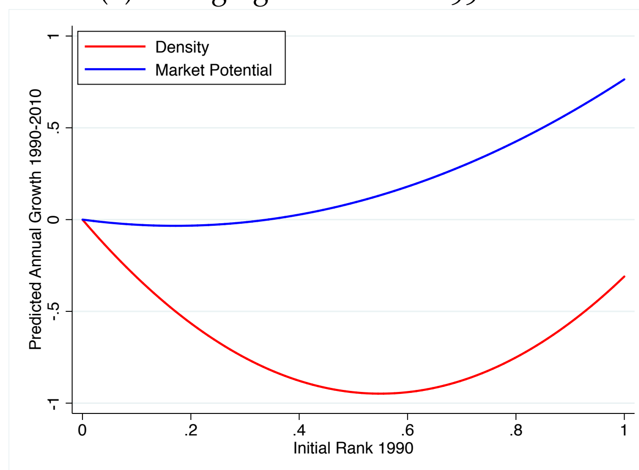
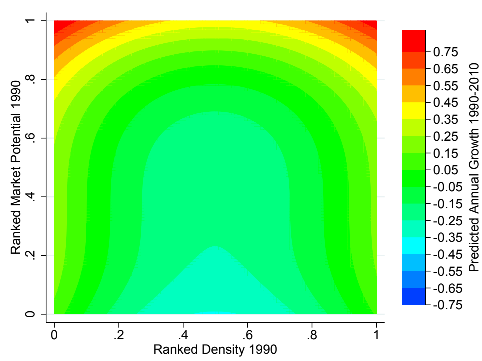
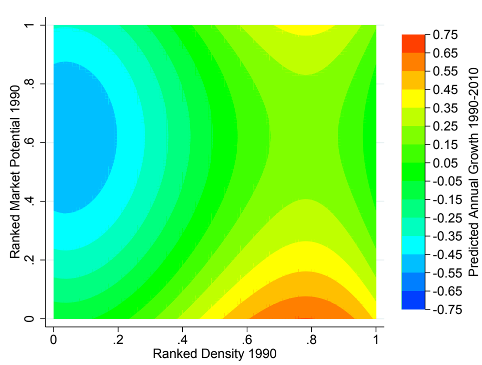


Figure 6: Market Size, Density and Growth, 1990-2010: Third-Order Polynomial Prediction

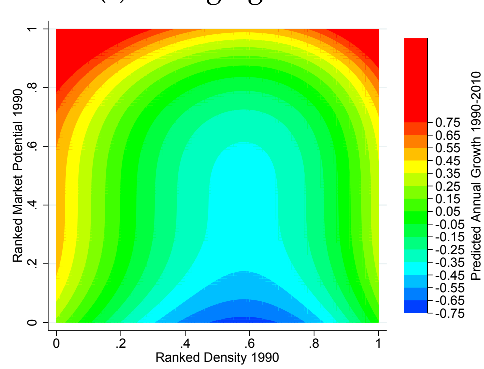
(a) All Economies



(b) Mature Economies



(c) Emerging Economies



4 The Spatial Implications of Structural Transformation

We now turn to exploring possible explanations for the spatial patterns we have identified. We start by assessing the role of structural transformation. Because sectors differ in where they tend to locate, the sectoral shift away from agriculture in emerging economies and towards services in mature economies is hence likely to change the spatial distribution of economic activity.

4.1 The Global Rise of Services

As can be seen in Table 2, the share of agricultural employment in our total sample fell from 49.9 percent in 1990 to 36.8 percent in 2010, whereas the corresponding employment share of services increased from 30.9 to 40.7 percent. The decline of agriculture and the rise of services can be observed in each of our eight sample economies. The share of manufacturing employment fell in six of our eight sample economies, the important exceptions being China and India.

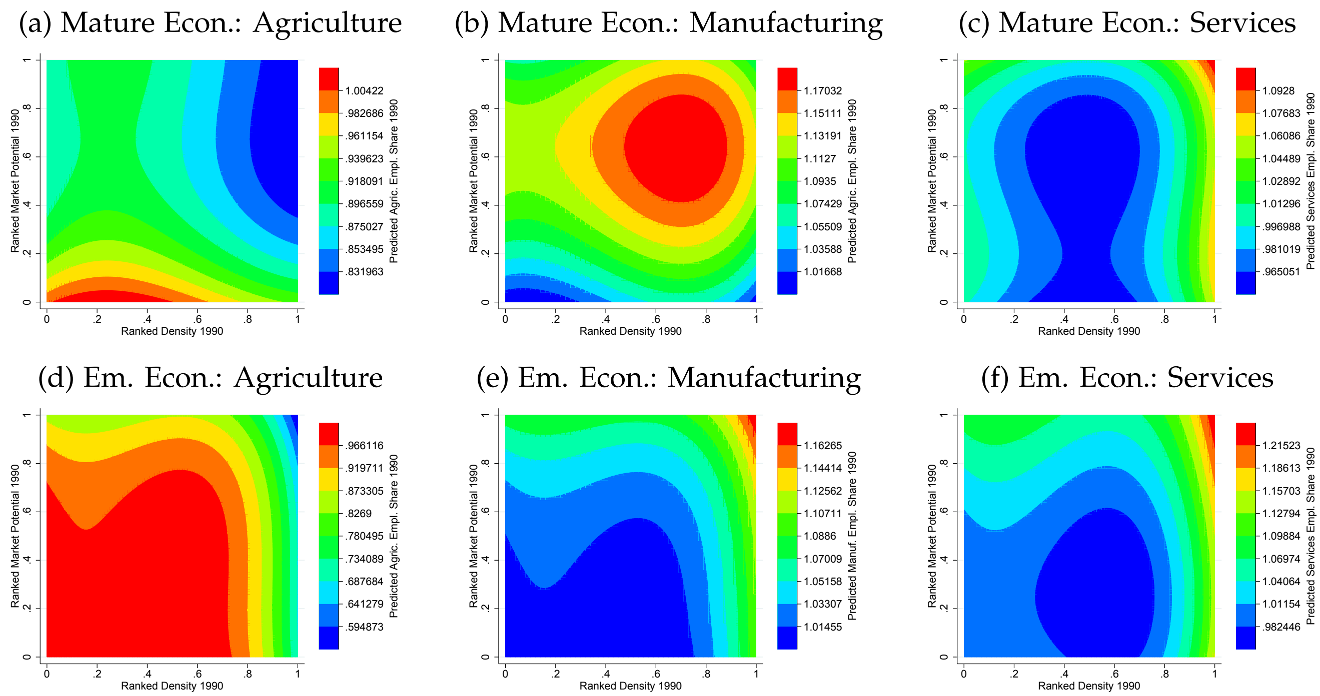
To understand how structural transformation might impact spatial growth patterns, we first document where different sectors located at the beginning of our sample period. Figure 7 depicts heat maps of predicted sectoral employment shares in 1990 as a function of third-order polynomials of market potential and employment density. In emerging economies, both manufacturing and service employment were strongly concentrated in high-density, high-market-potential locations. In mature economies, we see the same pattern for services, whereas manufacturing concentrates in medium-density, medium-market-potential places. In both economies, agriculture is concentrated in low-density, low-market-potential regions. Based on these initial sectoral shares, we would expect structural transformation, away from agriculture in emerging economies and towards services in mature economies, to especially benefit high-density, high-market-potential locations in both types of economies. In the next subsection, we analyze this prediction more formally.

4.2 Counterfactual Growth Due to Structural Change

To evaluate the role of structural transformation for the observed spatial growth patterns, we apply the accounting methodology used by Michaels, Rauch and Redding (2012). For each region, we calculate how fast growth would have been if each of its sectors had grown at its respective economy-wide sectoral growth rate. If actual growth patterns resemble these counterfactual growth patterns, we can conclude that structural transformation is an important driver of the relation between market potential, density and growth.

Formally, we take the employment level in sector s and region i of economy c in 1990, $E^s_{i,c,1990}$, and apply to it the economy-level employment growth rate of sector s between 1990 and 2010, $(E^s_{c,2010} - E^s_{c,1990})/E^s_{c,1990}$. Doing so for all three sectors – agriculture, manufacturing and services –

Figure 7: Sector Shares, 1990: Third-Order Polynomial Prediction

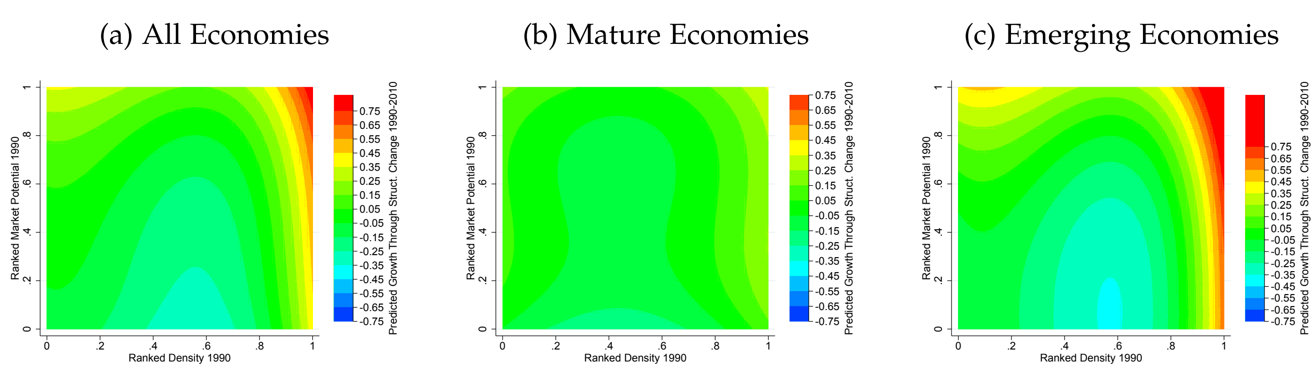


yields a counterfactual measure of aggregate employment in 2010 for each region i of economy c, $\widetilde{E}^{total}_{i,c,2010}$. To be precise,

$$\widetilde{E}^{total}_{i,c,2010} = \sum_{s} E^{s}_{i,c,1990} \frac{E^{s}_{c,2010}}{E^{s}_{c,1990}}.$$

We then use $\widetilde{E}^{total}_{i,c,2010}$ to compute a counterfactual growth rate between 1990 and 2010, to be included as dependent variable in regression (4). This counterfactual growth rate provides a measure of how much a particular region would have grown if the only force at work were nationally uniform structural transformation.

Figure 8: Structural Change: Third-Order Polynomial Prediction



Based on cubic regressions of the counterfactual growth rate on market potential and employment density, Figure 8 shows heat maps of the predicted counterfactual growth rates. The figures show that structural change correlates strongly with changing economic geographies in emerging economies, but not in mature economies. Panel (b), for mature economies, shows structural transformation leaving the spatial distribution of employment essentially unaffected. In contrast, structural change in emerging economies, away from agriculture and towards services, clearly favored employment growth in high-market-potential, high-density areas (Panel (c)).

Figure 9 illustrates the role of structural transformation in a different way. It compares the actual growth rates and the counterfactual growth rates as predicted by quadratic regressions on market potential and density. When considering all economies for the period 1990-2010, we see very little difference between actual growth and counterfactual growth. Hence, if the only force at work had been structural transformation, it would have let to a similar relation between market potential, density and growth as the one observed in the data. This suggests that structural transformation accounts well for the relationship between market potential, density and growth. If we look over time, structural transformation is unable to pick up the drop in importance in market potential between the two decades. If anything, it predicts market potential having an increasing effect, while the opposite is true in the data.

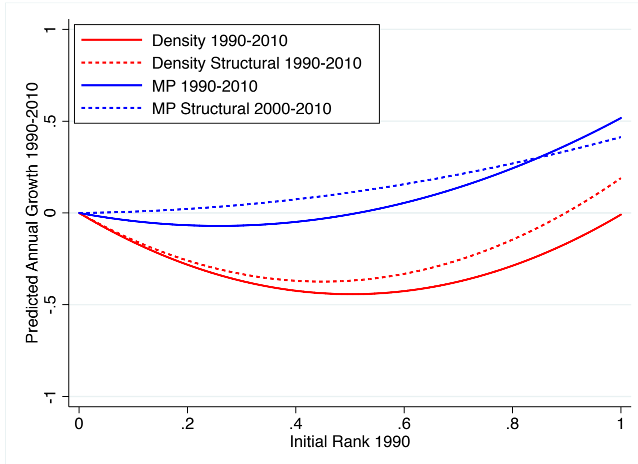
In Figure 10, we illustrate the role of structural change separately for mature and emerging economies. It again becomes clear that structural transformation is able to account for the high growth of high-market-potential regions in emerging economies. Panel (b) shows that if the only force at work were structural transformation, it would have generated essentially the same relation between growth and market potential as the one observed in the data. In contrast, structural transformation is unable to explain the negative effect of market potential on growth in mature economies. In fact, Panel (a) shows that structural transformation by itself would have generated growth rates that bear little relation to density and market potential.

4.3 Residual Growth

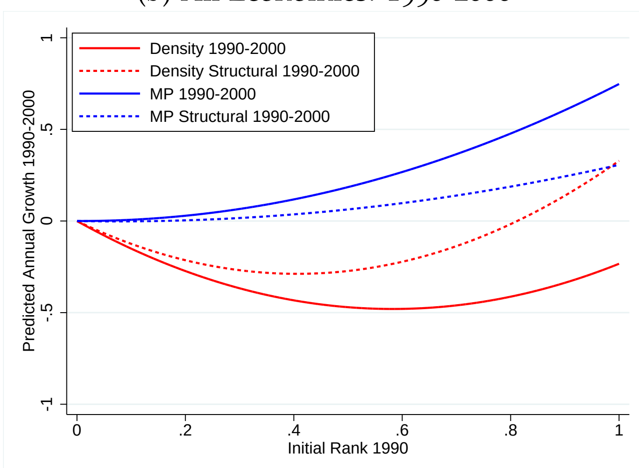
An alternative way to assess the role of structural transformation is to analyze what it leaves unexplained. We refer to this unexplained part as residual growth, and define it as the difference between actual growth and counterfactual growth. Figure 11 shows heat maps of residual growth on market potential and density. These heat maps can be compared to Figure 6 for total employment growth and to Figure 8 for counterfactual employment growth due to structural change. For all economies taken together, residual growth patterns are largely unrelated to market potential and density, confirming that structural transformation accounts for the observed patterns in the data. Likewise, for emerging economies, structural transformation is able to explain the relatively fast

Figure 9: Market Potential, Density and Growth: Actual vs. Structural Transformation

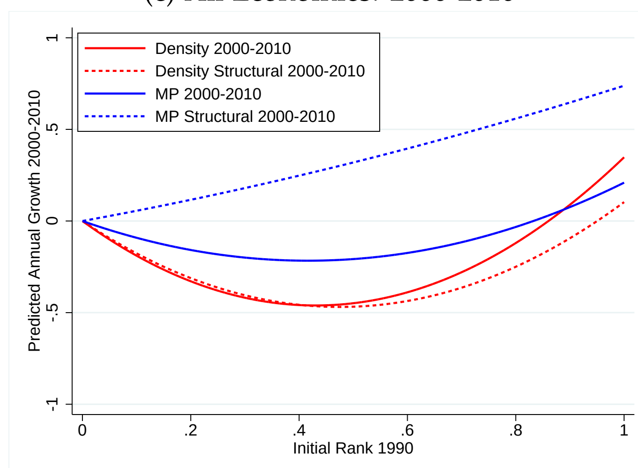
(a) All Economies: 1990-2010



(b) All Economies: 1990-2000



(c) All Economies: 2000-2010



growth of high-market-potential locations, though it underpredicts growth if those locations have low density – a fact that is likely due to population and economic activity "spreading out" into hitherto sparsely populated areas. In contrast, for mature economies, the residual growth patterns look very similar to the actual growth patterns, indicating that structural change plays no discernible role in the rise of low-market-potential locations.[15]

In addition to visually inspecting heat maps, we also compare regressions of residual employment growth (Table 5) to regressions of actual employment growth (Table 3). By comparing column (1) of Table 5 to column (2) of Table 3, we can infer that, for the world as a whole, structural transformation accounts for about two-thirds of the positive effect of market potential on growth. To see this, in the actual growth regression the coefficient on market potential is 0.47, whereas in the residual growth regression that same coefficient is 0.15, so that a share $(0.47 - 0.15)/0.47 = 0.68$ of

[15] This conclusion is consistent with Figure 3, where we find that in mature economies all three sectors grew most strongly in these low-market-potential regions.

Figure 10: Market Potential, Density and Growth: Actual vs. Structural Transformation

(a) Mature Economies: 1990-2010 (b) Emerging Economies: 1990-2010

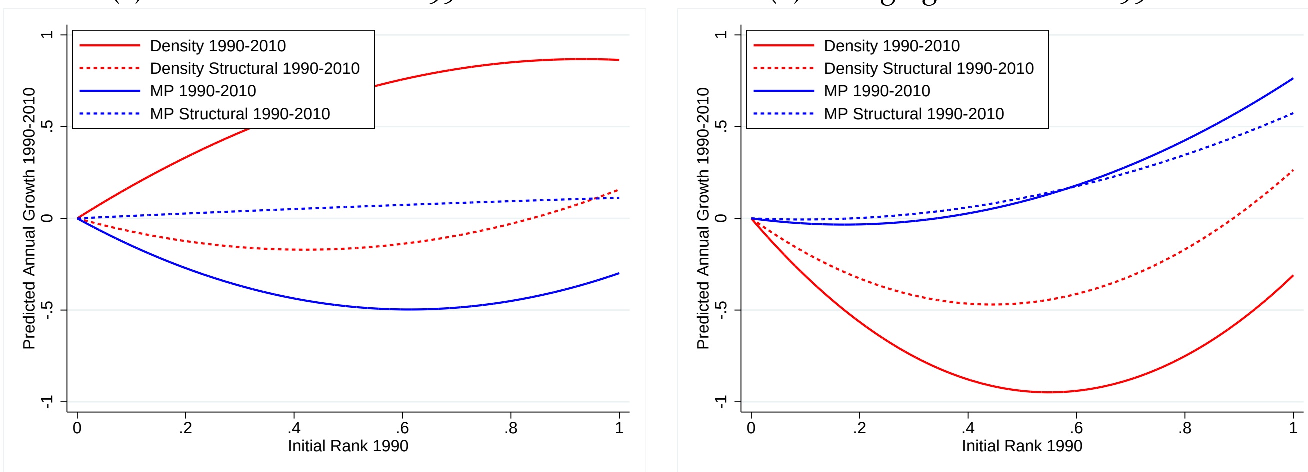
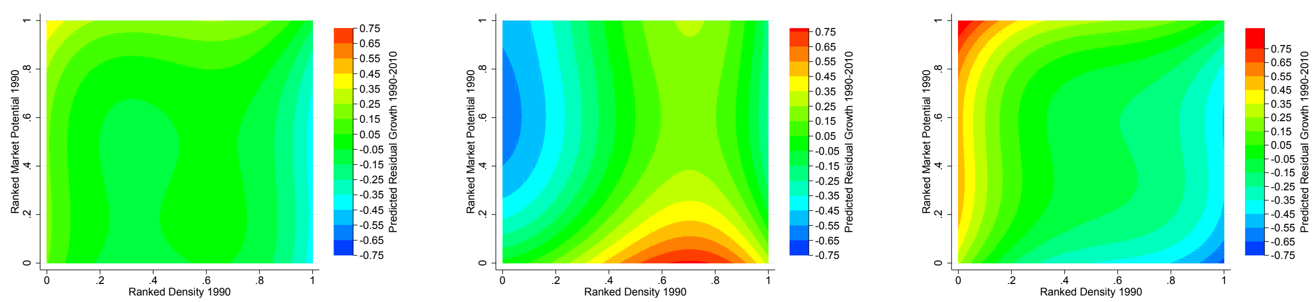


Figure 11: Residual Growth, 1990-2010: Third-Order Polynomial Prediction

(a) All Economies (b) Mature Economies (c) Emerging Economies



the overall effect can be accounted for by structural transformation, with the remaining one third still to be explained.

We now do the same comparisons for emerging and mature economies. In emerging economies, comparing column (3) of Table 5 to column (1) of Table 4 reveals that the positive effect of market potential on residual growth is only about 40 percent as large as the effect on total growth – some 60 percent of the observed market potential effect is therefore accounted for by structural transformation. In contrast, a comparison of column (2) of Table 5 to column (1) of Table 4 shows that structural transformation does not correlate significantly with changes in the economic geography of mature economies. If anything, the result goes the other way: the coefficient on market potential in the residual growth regression is negative and of a larger magnitude than its corresponding coefficient in the actual growth regression. If the only force at work had been structural transformation, it would have benefited high-market-potential places in mature economies.

In columns (4) and (5) of Table 5, we subdivide our sample period into its two constituent

Table 5: Market Potential, Density and Residual Employment Growth

Dependent Variable: Residual Employment Growth (= *Empl. Growth - Structural Empl. Growth*)		1990-2010		1990-2000	2000-2010
	All	Mature	Emerg.	All	All
	(1)	(2)	(3)	(4)	(5)
Market Potential	0.144**	-0.367***	0.282***	0.513***	-0.446***
	(0.073)	(0.071)	(0.098)	(0.100)	(0.085)
Density	-0.219***	0.676***	-0.634***	-0.557***	0.0824
	(0.074)	(0.074)	(0.100)	(0.101)	(0.083)
Constant	0.012	-0.149***	0.130***	0.008	0.147***
	(0.034)	(0.040)	(0.050)	(0.050)	(0.040)
Observations	17,418	5,432	11,986	17,418	17,418
R^2	0.001	0.017	0.011	0.005	0.005

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

decades. A comparison with Table 3 confirms that also within each decade the effect of market potential is consistently less positive for residual employment growth than for total employment growth. The most striking finding of the between-decade comparison is the strong decline in the residual employment growth effect of market potential across the two subperiods, turning from significantly positive in the 1990s to significantly negative in the 2000s. Hence, both the cross-sectional comparison of mature and emerging economies and the intertemporal comparison across our two sample decades suggest that structural change acted as a brake on the relative loss of dynamism of high-market-potential regions.

Summarizing the evidence in the last two subsections, we can conclude that structural transformation accounts rather well for the spatial growth patterns in the world as a whole over the period 1990-2010. This is almost entirely due to its high explanatory power in emerging economies. It does less well in accounting for the weakening importance of market potential over time in the world economy, and for the negative effect of market potential on growth in mature economies. That is, the shift away from market potential, especially in mature economies, remains to be explained.

5 Transport Costs and Market Potential

In this section we offer a possible explanation for the negative effect of market potential on residual growth in mature economies and in the world as a whole in the more recent decade. In addition to structural transformation, a second force that is likely to have shaped the world's changing economic geography is the systematic decline in transport and trade costs (World Trade Organization, 2008; Redding and Turner, 2015). For example, the cost of air shipping dropped by 92 percent between 1955 and 2004, and the cost of maritime transport has been steadily declining since the mid-1980s (Hummels, 2007). Several factors have contributed to this trend, including technological improvements, scale economies and market liberalization.

To understand how the drop in transport costs might have affected the relation between market potential and growth, consider a standard economic geography model with one central location and two peripheral locations. The full details of the model are given in Appendix B; here we limit ourselves to a brief description. In an otherwise symmetric setup, when trade costs are prohibitive, free mobility implies that employment will be equally spread across the three locations. As trade costs gradually drop, the high-market-potential central location initially gains employment at the expense of the peripheral locations, because it benefits disproportionately from improved access to the other locations. However, as trade costs continue to fall, the proximity advantage of the high-market-potential central location starts to erode. The peripheral locations benefit from a larger improvement in market access, and they once again become attractive because of their lower congestion. This yields a bell-shaped relation between the employment share of the high-market-potential location and the level of trade costs.

In this model, when transport costs are already relatively low, any further drop generates a negative relation between initial market potential and growth: the low-market-potential (peripheral) locations gain relative to the high-market-potential (central) location. The lower growth of the high-market-potential region does not imply that there is a disadvantage to market access. Rather, it reflects peripheral locations gaining increasingly equal market access, leading to a weakening of the relative advantage of centrality. That is, the improvement in market potential is larger in locations that start off with lower market potential. This explains why a drop in transport costs causes peripheral locations to grow faster.

Might this account for the negative relation between market potential and residual growth in mature economies? We do not have precise measures of the improvement in transport costs across space for the period 1990-2010, and we are therefore unable to test the model's predictions directly. However, if mature economies experienced a drop in already low transport costs, our empirical findings are consistent with the model. Moreover, to the extent that the fall in transport costs was uniform across space in mature economies, we can gauge the relevance of the model's predictions by simulating the effect of lower trade costs. Since transport infrastructure networks in mature economies were already largely in place in 1990, most cost savings in interregional transport came from technological and regulatory changes. As a first approximation, it is therefore reasonable to assume that such changes affected all locations in a similar way. Following this argument, we consider a uniform drop in the distance decay parameter γ in equation (1) to generate simulated estimates of the changes in market potential in all locations. This approach yields negative (but not perfect) correlations between simulated increases in market potentials and initial market potentials. Hence, the gain in market potential is larger in locations with initially lower market potential. Table 6 reports regressions of mature-economy residual growth on simulated changes of market potential, assuming different reductions in the decay parameter.[16] We observe that in all specifications employment growth correlates positively and significantly with simulated increases in market potential. This is consistent with the mechanism of our geography model. Both in the model and in the data, locations with lower initial market potential benefit from greater gains in market potential. This explain why the negative relation between initial market potential and growth implies a positive relation between the change in market potential and growth.

Table 6: Residual Growth and Simulated Changes in Market Potential, Mature Economies

Dependent Variable: Residual Employment Growth				
	(= Empl. Growth - Structural Empl. Growth)			
	(1)	(2)	(3)	(4)
Δ Market Potential 1	0.315***			
	(0.073)			
Δ Market Potential 2		0.237***		
		(0.075)		
Δ Market Potential 3			0.314***	
			(0.073)	
Δ Market Potential 4				0.236***
				(0.075)
Density	0.643***	0.591***	0.643***	0.591***
	(0.076)	(0.078)	(0.076)	(0.078)
Constant	-0.474***	-0.409***	-0.473***	-0.408***
	(0.069)	(0.072)	(0.069)	(0.072)
Observations	5,436	5,436	5,436	5,436
R^2	0.015	0.014	0.015	0.014

Robust standard errors in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Δ Market Potential 1: distance decay parameter γ set to 0.5;

Δ Market Potential 2: distance decay parameter γ set to 0.8;

Δ Market Potential 3: distance decay parameter γ set to 0.5 (except w.r.t. own region);

Δ Market Potential 4: distance decay parameter γ set to 0.8 (except w.r.t. own region).

Why did we not see this same relation between market potential and growth in emerging economies? We can think of at least three potential reasons. First, the model's bell-shaped prediction says that a drop in transport costs will hurt high-market-potential locations if those costs are relatively low, but it will benefit high-market-potential locations if those costs are relatively high. Hence, if transport was still relatively expensive in emerging economies, our findings of central lo-

[16] In some specification, we assume transport costs fall only between regions but not within regions.

cations experiencing faster growth is consistent with the theory.[17] Second, transport infrastructure investment has been a big driver of falling transport costs in emerging economies. To the extent that these improvements did not occur uniformly across the board but were mostly concentrated in locations with already high market potential, we are less likely to see a negative relation between market potential and growth.[18] Third, in emerging economies structural transformation is the main determinant of growth in high-market-potential locations over the full sample period. As is obvious from Figure 10, Panel (b), counterfactual growth based on structural transformation lines up nearly perfectly with the actual growth rate in emerging economies. Hence, residual growth patterns are less revelatory of the dominant forces shaping the location of overall employment in emerging economies than in mature economies.

6 Conclusions

This paper has documented employment growth patterns in 18,961 regions across the world for the period 1990 to 2010. In doing so, we have identified two stylized facts. First, for the world as a whole, market potential is becoming less important, and local density is becoming more important for economic growth. Second, market potential has a positive effect on growth in emerging economies, while the opposite is true in mature economies. We have shown that structural transformation – especially the growing share of employment in services – goes some way in accounting for the world's changing economic geography. It offers an explanation for the increasing importance of local density in the world as a whole, and it accounts well for local growth patterns in emerging economies. However, it fails to provide an explanation for the shrinking advantage of market potential, especially in mature economies. We have then shown that this latter finding is consistent with a standard economic geography model that yields a bell-shaped relation between trade costs and the growth of high-market-potential areas.

[17] We indeed know that transport costs are higher in emerging economies than in mature economies (World Bank, 2009). As an example, over the period 1995-2014 maritime freight costs were between 20 and 50 percent higher in the developing countries of Asia and America than in the developed world (UNCTAD, 2012).

[18] This also means that simulations of uniform changes in transport costs across region pairs analogous to Table 6 would not be appropriate in the case of our emerging-economy sample.

References

[1] Allen, T., and Arkolakis, C., 2014. “Trade and the Topography of the Spatial Economy,” *Quarterly Journal of Economics*, 129, 1085-1140.

[2] Baragwanath-Vogel, K., Goldblatt, R., Hanson, G., and Khandelwal, A., 2018. “Detecting Urban Markets with Satellite Imagery: An Application to India,” *NBER Working Papers*, #24796.

[3] Behrens, K., Duranton, G., and Robert-Nicoud, F., 2014. “Productive Cities: Sorting, Selection, and Agglomeration,” *Journal of Political Economy*, 122, 507-553.

[4] Combes, P.-P., Lafourcade, M., Thisse, J.-F. and Toutain, J.-C., 2011. “The Rise and Fall of Spatial Inequalities in France: A Long-Run Perspective,” *Explorations in Economic History*, 48, 243-271.

[5] Autor, D., Dorn, D. and Hanson, G., 2013, “The China Syndrome: Local Labor Market Effects of Import Competition in the United States," *American Economic Review*, 103, 2121-2168.

[6] Daumal, M. and Zignago, S., 2010. “Measure and Determinants of Border Effects of Brazilian States,” *Papers in Regional Science*, 89, 735-758.

[7] Desmet, K. and Henderson, J. V., 2015. “The Geography of Development within Countries,” In: Duranton, G., Henderson, J. V. and Strange, W. (eds.), *Handbook of Urban and Regional Economics*, Volume 5, 1457-1517, Elsevier.

[8] Desmet, K., Nagy, D. and Rossi-Hansberg, E., 2018. “The Geography of Development,” *Journal of Political Economy*, 126, 903-983.

[9] Desmet, K. and Rossi-Hansberg, E., 2009. “Spatial Growth and Industry Age,” *Journal of Economic Theory*, 144, 2477-2502.

[10] Desmet, K. and Rossi-Hansberg, E., 2014. “Spatial Development,” *American Economic Review*, 104, 1211-1243.

[11] Disdier, A.-C. and Head, K., 2008. “The Puzzling Persistence of the Distance Effect on Bilateral Trade,” *Review of Economics and Statistics*, 90, 37-48.

[12] Forslid, R., Haaland, J. I. and Midelfart Knarvik, K. H., 2002. “A U-Shaped Europe? A Simulation Study of Industrial Location,” *Journal of International Economics*, 57, 273-297.

[13] Fujita, M., Krugman, P. and Venables, A., 1999. *The Spatial Economy: Cities, Regions, and International Trade*, MIT Press: Cambridge.

[14] Fujita, M. and Thisse, J.-F., 2006. "Globalization and the Evolution of the Supply Chain: Who Gains and Who Loses?" *International Economic Review*, 47, 811-836.

[15] Gennaioli, N., La Porta, R., Lopez-de-Silanes, F. and Shleifer, A., 2014. "Growth in Regions," *Journal of Economic Growth*, 19, 259-309.

[16] Harris, C. 1954. "The Market as a Factor in the Localization of Industry in the United States," *Annals of the Association of American Geographers*, 64, 315-348.

[17] Head, K. and Mayer, T., 2000. "Non-Europe : The Magnitude and Causes of Market Fragmentation in Europe," *Weltwirtschaftliches Archiv*, 136, 285-314.

[18] Head, K. and Mayer, T., 2004. "The Empirics of Agglomeration and Trade," in: J. V. Henderson and J. F. Thisse (eds.), *Handbook of Regional and Urban Economics*, Edition 1, Volume 4, Chapter 59, 2609-2669, Elsevier.

[19] Head, K. and Mayer, T., 2011. "Gravity, Market Potential and Economic Development," *Journal of Economic Geography*, 11, 281-294.

[20] Henderson, J. V., Squires, T., Storeygard, A. and Weil, D., 2018. "The Global Distribution of Economic Activity: Nature, History and the Role of Trade," *Quarterly Journal of Economics*, 133, 357-406.

[21] Henderson, J. V., Storeygard, A. and Weil, D., 2012. "Measuring Economic Growth from Outer Space," *American Economic Review*, 102, 994-1028.

[22] Hummels, D., 2007. "Transportation Costs and International Trade in the Second Era of Globalization," *Journal of Economic Perspectives*, 21, 3, 131-154.

[23] Jacks, D. S. and Novy, D., 2018. "Market Potential and Global Growth over the Long Twentieth Century," *Journal of International Economics*, 114, 221-237.

[24] Kim, S. 1995. "Expansion of Markets and the Geographic Distribution of Economic Activities: The Trends in US Regional Manufacturing Structure, 1860-1987," *Quarterly Journal of Economics*, 110, 881-908.

[25] Krugman, P. and Venables, A. J., 1995. "Globalization and the Inequality of Nations," *Quarterly Journal of Economics*, 110, 857-880.

[26] Michaels, G., Rauch, F. and Redding, S. J., 2012. "Urbanization and Structural Transformation," *Quarterly Journal of Economics*, 127, 535-586.

[27] Michaels, G., Rauch, F. and Redding, S. J., 2018. "Task Specialization in U.S. Cities from 1880 to 2000," *Journal of the European Economic Association*, 17, 754-798.

[28] ORGI, 2011. *Administrative Atlas of India,* Map Division, Office of the Registrar General & Census Commissioner, India Ministry of Home Affairs, Government of India.

[29] Pinkovskiy, Maxim L., 2017. "Growth Discontinuities at Borders," *Journal of Economic Growth*, 22, 145-192.

[30] Puga, D., 1999. "The Rise and Fall of Regional Inequalities," *European Economic Review*, 43, 303-334.

[31] Redding, S. and Turner, M., 2015. "Transportation Costs and the Spatial Organization of Economic Activity," in: (eds) G. Duranton, J. V. Henderson and W. Strange, *Handbook of Urban and Regional Economics*, Chapter 20, 1339-1398.

[32] Redding, S. and Venables, A. J., 2004. "Economic Geography and International Inequality," *Journal of International Economics*, 62, 53-82.

[33] Siczewicz, P., 2011. U.S. Historical Counties Dataset. Emily Kelley, Digital Comp. Atlas of Historical County Boundaries, ed. by John H. Long. Chicago: The Newberry Library. Available online at http://publications.newberry.org/ahcbp.

[34] Simonovska, I. and Waugh, M., 2014. "Trade Models, Trade Elasticities, and the Gains from Trade," NBER Working Paper #20495.

[35] Tabuchi, T., 1998. "Urban Agglomeration and Dispersion: A Synthesis of Alonso and Krugman," *Journal of Urban Economics*, 44, 333-351.

[36] UNCTAD, 2012. "Freight Rates and Maritime Transport Costs," *Review of Maritime Transport 2012*, Geneva: UNCTAD.

[37] World Bank, 2009. *World Development Report 2009: Reshaping Economic Geography,* Washington, D.C.

[38] World Trade Organization, 2008. *World Trade Report 2008: Trade in a Globalizing World*, Geneva: World Trade Organization.

A Data Construction

A.1 Country Data

Brazil. Brazilian employment data come from the Demographic Censuses for the years 1991, 2000 and 2010. These microlevel data are published by the Brazilian Institute for Statistics and Geography (IBGE). We aggregate the data to the municipal level. This corresponds to the second level of administrative divisions, beneath the states. We use concordance tables from IBGE to convert the 1991 industrial classification, specific to the 1991 census, to the *CNAE-Domiciliar* classification of the 2000 census, which is used for both the 2000 and 2010 censuses.

China. Employment data for China come from the 1990, 2000 and 2010 Population Census. We use data at the county level, the third division level after the province and the prefecture. These data are distributed by the University of Michigan China Data Center.

Europe. European regional employment comes from Cambridge Econometrics. It covers the 27 member states of the European Union and Norway. Data are at the third administrative division level (NUTS3) and are available at an annual frequency. For part of our analysis we split Europe into Western Europe and Central and Eastern Europe, with the divide being what used to be the Iron Curtain. Due to data availability, the initial year of analysis for Central and Eastern European countries is 1991.

India. The main data source for India is the Primary Census Abstract published by the Office of the Registrar General of India and Census Commissioner of India. We use data for the years 1991, 2001 and 2011. The data are available at the subdistrict level, which corresponds to the third-level administrative division, beneath the states and the districts. We ignore the state of Jammu & Kashmir, due to the unavailability of data for the year 1991. The 1991 census provides employment information for 10 sectors, while the censuses of 2001 and 2011 only distinguishes four sectors. To address this shortcoming, we augment data for 2000 and 2011 with sectoral details from the Employment & Unemployment Survey published by the National Sample Survey Office (NSSO).[19]

Japan. Japanese employment data come from the Population Censuses for the years 1990, 2000 and 2010. They are distributed by the Statistical Information Institute for Consulting and Analysis (Sinfonica). The data are available at the municipality (Shi, Machi and Mura) level, which is the second

[19] The 55th (2000) and 66th (2010) rounds of the NSS provide employment at the district level for 5-digit industries. To get estimates of employment at the subdistrict level, we regress district-level sectoral employment shares from NSS on different variables available in the census: 4 sectors (cultivators, agricultural, household industry, others), the share of urban population, the literacy rate and state fixed effects. Since the dependent variable is a share $\in [0, 1]$, we use GLM estimation with logit link and binomial family. We then use the parameters estimates from the regression to predict sectoral shares at the subdistrict level, making sure that the estimated share corresponds to the actual district aggregates from the NSS data.

degree of administrative divisions, below the prefecture.

Mexico. Data for Mexico are taken from the census *(Censo General de Población y Vivienda)* for the years 1990, 2000 and 2010. These data are published by the Mexican National Institute of Statistics and Geography (INEGI) and are available at the municipality level, which corresponds to the third-degree administrative divisions, below the states and districts.[20]

United States. U.S. employment data comes from the County Business Patterns (CBP) dataset published by the U.S. Census Bureau. The data cover workers in the private sector at the county level (second-level administrative division, below the states).[21] We complement private sector employment from the CBP with government employment from the Regional Economic Accounts of the U.S. Bureau of Economic Analysis. The classification of the CBP data is based on the 1988-1997 Standard Industrial Classification (SIC) nomenclature in 1990, the 1998-2001 North American Industry Classification System (NAICS) in 2000 and the 2007-2008 NAICS in 2010. We convert 1990 and 2000 data into the NAICS 2007 classification using concordance tables from the U.S. Census Bureau.

A.1.1 Sectoral Disaggregation.

The available sectoral detail for regional employment varies across countries and over time. For the sake of comparability, we aggregate employment data to the following three broad sectors: agriculture (primary), manufacturing (secondary) and services (tertiary). All details of how this is done is available in the paper's Online Appendix.

A.2 Urban Areas

To construct comparable urban areas across economies, we start by using the Defense Meteorological Satellites Program – Operational Linescan System (DMSP–OLS) nighttime lights data of 2010. We correct the DMSP–OLS data for over-glow following Pinkovskiy (2017). To capture high-density areas, we form polygons based on the top most luminous grids of night light. The yellow polygons in Figure A6 panel (a) illustrate this for the case of the U.S. Eastern Seaboard. For a polygon to qualify as an urban area, we assume a minimum size, and for an administrative region to be included in an urban area, we assume a minimum share of its area must be covered by a polygon.

To determine the values of this minimum size and this minimum share, we perform a calibration exercise that targets the high-density part of large Metropolitan Statistical Areas (MSAs) in

[20] We augment the data for Oaxaca in 2000, available only at the district-level in the census, with a 10.6% micro-data municipality-level sample from the census obtained through the Integrated Public Use Microdata Series (IPUMS). We obtain total municipality level data by using weights from the IPUMS data that correspond to district employment shares. Likewise, due to limited sectoral information for 2010 in the state of Oaxaca, we also rely on a 10% micro-data sample of the 2010 census obtained from IPUMS. In that case, we multiply municipality-level sectoral shares from IPUMS by the total employment level in the census.

[21] Since CBP employment by industry is sometimes reported as intervals, we use the fixed-point algorithm of Autor, Dorn and Hanson (2013) to estimate employment numbers within those intervals.

the U.S. We obtain reasonable results by considering counties with at least 20% of their surface covered by night light polygons of at least 200 square kilometers.[22] One issue is that some night light polygons generate continuous urban areas that are unreasonably large. For example, Figure A6 panel (a) shows one of the night light polygons covering, without discontinuity, two MSAs, Washington-Arlington-Alexandria and Baltimore-Towson. To avoid such cases, we use additional data on *Artificial surfaces and associated areas (Urban areas >50%)* from ESA's Globcover project. The polygons formed by artificial surfaces aim to capture the cores of the different urban areas which otherwise cannot be segregated using the night light data. Such artificial surfaces are represented in red in Figure A6 panel (a). When a polygon of night light covers several urban cores, we assign the county to an urban area based on its proximity to those cores.

Applying this algorithm, we identify in the U.S. 37 urban areas covering 77 counties, all of which are assigned to the correct MSAs as delineated by the OMB. Because our aim is to capture the high-density parts of large MSAs, we exclude certain types of counties that would otherwise enter into the definition of an MSA. An example would be a county that is divided into a large low-density hinterland and a small high-density cluster close to an urban core. Although most of its population is functionally part of the urban area, including such a county would introduce a downward bias on the density of urban areas. Figure A6 panel (b) provides a comparison between our methodology and the OMB classification of MSAs.[23] If we consider MSAs with a population of at least one million, our procedure captures the different MSAs reasonably well. From north to south on the U.S. Eastern Seaboard, we can distinguish Boston-Cambridge-Quincy, New York-Northern New Jersey-Long Island, Philadelphia-Camden-Wilmington, Baltimore-Towson, Washington-Arlington-Alexandria, Richmond, and Virginia Beach-Norfolk-Newport News.

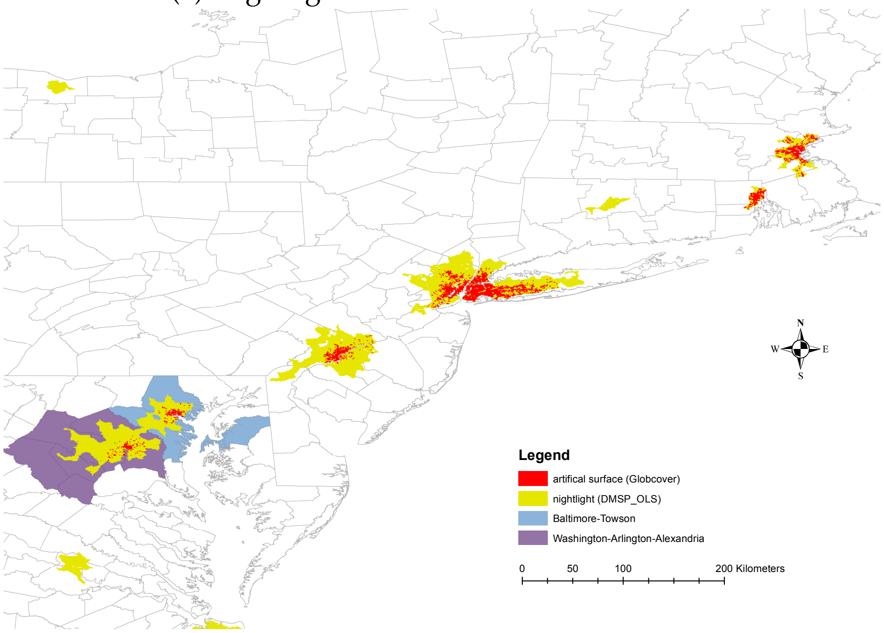
However, with the exception of Baltimore-Towson, in general our procedure does not capture the full extent of the MSAs. The correlation between the counties we identify as urban and the counties that belong to MSAs with a population of at least one million is 0.51. Although a lower parameter for the share of a county covered by night light would yield a higher correlation, this would come at the price of including certain low-density and other intruder counties that are not categorized as belonging to any MSA. Since our main aim is to identify high-density urban areas, we avoid this by setting parameter values that exclude both low-density and intruder counties.

[22] Including smaller polygons results in identifying too many urban areas which do not exist in the U.S. Office of Management and Budget (OMB) delineation of MSAs. It also results in splitting actual MSAs into several metropolitan areas. For instance, considering polygons of night light with a surface greater or equal to 100 km splits the MSA of Virginia Beach-Norfolk-Newport News into two distinct urban regions. This is due to the agglomerations of Hampton and Newport News (night light surface of 167 km^2) being separated from the larger agglomeration of Virginia Beach, Norfolk and Portsmouth (night light surface of 512 km^2).

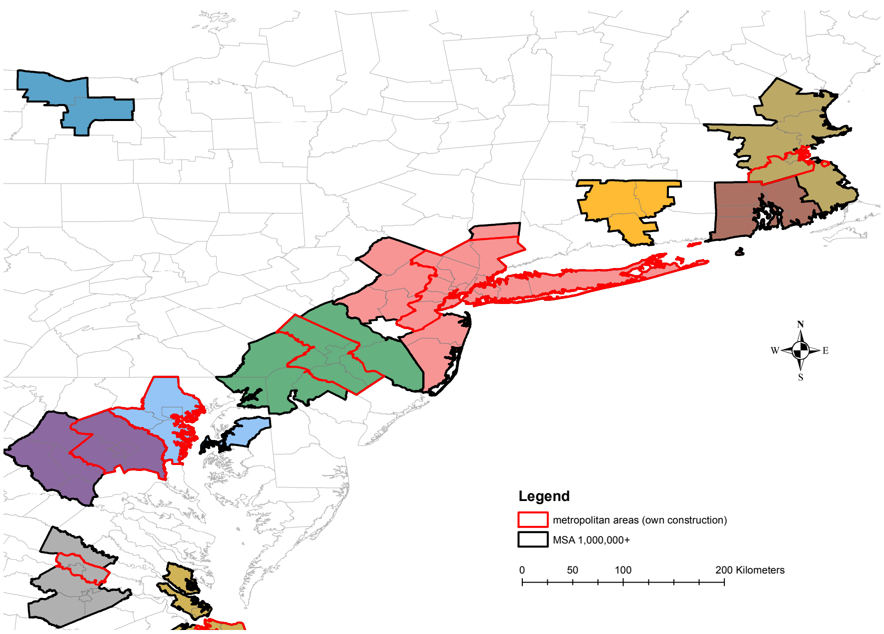
[23] We only include the core counties of MSAs.

Figure A6: High Density Areas on the U.S. Eastern Seaboard

(a) Nightlight Data and Artificial Surfaces



(b) Regional Allocation



B Model

In this section we provide a simple three-location economic geography model to analyze the relation between trade costs, market potential and population growth. We analytically show that the effect of market potential on growth is non-monotonic in trade costs. Starting at a high level, falling trade costs makes market potential more important, favoring growth in the location with better market access. However, if trade costs continue to drop, eventually growth shifts to the peripheral locations, with worse market access. We further illustrate the model's implications with a series of numerical examples.

B.1 Setup and Equilibrium

Endowments. The economy consists of three locations on a line, a central location, k, and two peripheral locations, ℓ_1 and ℓ_2. Locations are indexed by either i or j, where $i, j \in \{k, \ell_1, \ell_2\}$. The distance between location i and j is denoted by d_{ij}, where $d_{ij} = d_{ji}$. We refer to k as central and to ℓ_1 and ℓ_2 as peripheral, because $d_{\ell_1 k} < d_{\ell_1 \ell_2}$ and $d_{\ell_2 k} < d_{\ell_1 \ell_2}$. Each location has one unit of land, collectively owned by the local population. The economy has a labor endowment L which is freely mobile across locations.

Technology and supply. Each location produces a different good. Firms are perfectly competitive. The production technology is Cobb-Douglas in land and labor. Given the normalization of land, output in location i can be written as:

$$Y_i = \left(\bar{A}_i L_i^{\varepsilon}\right) L_i^{\alpha} \tag{5}$$

where L_i is the labor input in location i, and $\bar{A}_i L_i^{\varepsilon}$ is TFP in location i. TFP depends on a location's exogenous productivity, $\bar{A}_i$, and on local density, L_i^{ε}, which captures agglomeration economies. Congestion costs come from the use of land. Their strength depends on the importance of land, $1 - \alpha$. We assume that agglomeration economies are weaker than congestion costs, $\varepsilon < 1 - \alpha$.

Profit maximization implies that agents in location i earn a wage

$$w_i = \alpha p_i \bar{A}_i L_i^{\varepsilon + \alpha - 1} \tag{6}$$

where p_i is the price of good i in location i. When one good of location i is shipped to location j, $(1 + d_{ij})^{-\gamma}$ units arrive. Hence, the price of good i in location j is $p_i(1 + d_{ij})^{\gamma}$. Each agent in location i earns land income equal to

$$r_i = (1 - \alpha) p_i \bar{A}_i L_i^{\varepsilon + \alpha - 1} \tag{7}$$

so that his total income is

$$y_i = p_i \bar{A}_i L_i^{\varepsilon+\alpha-1}. \tag{8}$$

Preferences and demand. Consumers in location i have CES preferences over the three different goods:

$$u_i = \left(\sum_j c_j^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \tag{9}$$

where c_j is the consumption of good j. An agent in location i earning income y_i has the following demand for good j:

$$c_j = \frac{y_i (p_j (1+d_{ij})^\gamma)^{-\sigma}}{\sum_j (p_j (1+d_{ij})^\gamma)^{1-\sigma}}. \tag{10}$$

Her indirect utility will then be:

$$u_i = \frac{y_i}{\left(\sum_j (p_j (1+d_{ij})^\gamma)^{1-\sigma} \right)^{\frac{1}{1-\sigma}}} \tag{11}$$

where the denominator corresponds to the price index of location i, P_i:

$$P_i = \left(\sum_j (p_j (1+d_{ij})^\gamma)^{1-\sigma} \right)^{\frac{1}{1-\sigma}}. \tag{12}$$

Aggregate demand for goods produced in location j is:

$$C_j = \sum_i \frac{y_i L_i (p_j (1+d_{ij})^\gamma)^{-\sigma}}{P_i^{1-\sigma}}. \tag{13}$$

Equilibrium. Workers are freely mobile, so utility equalizes across locations. We are now ready to define an equilibrium.

For a given total population L, an equilibrium is a collection of variables $\{p_i, L_i, \bar{u}\}$ where $i \in \{k, \ell_1, \ell_2\}$, that satisfy equations (8), (12) and

1. *Goods market clearing:*

$$y_i L_i = \sum_j \frac{y_j L_j (p_i (1+d_{ij})^\gamma)^{1-\sigma}}{P_j^{1-\sigma}} \tag{14}$$

2. *Labor market clearing:*

$$L = \sum_j L_j \tag{15}$$

3. *Labor mobility:*

$$\frac{y_i}{P_i} = \bar{u} \tag{16}$$

Recall that congestion costs from land are stronger than agglomeration economies from density, $1 - \alpha > \varepsilon$. This implies that any equilibrium is stable: if utility equalizes across locations, no individual has an incentive to deviate and change location. To see this, recall that the utility of an agent in location i is $u_i = \left(p_i \bar{A}_i L_i^{\varepsilon+\alpha-1}\right) / P_i$. Any individual deviation leaves p_i and P_i unchanged, so u_i is decreasing in the size of the local population L_i. As such, any individual deviation from an equilibrium where utility equalizes across locations results in a decrease in the utility of the deviating agent. Hence, that agent prefers not to deviate, which proves that the equilibrium is stable.

Concept of market potential. A location's attraction depends on the market it can access in its own location and in other locations. This is the concept Harris (1954) refers to as market potential. Intuitively, the market potential of location i depends (i) positively on the income of all other locations, $y_j L_j$, and (ii) negatively on the difficulty of accessing all other locations, $(1 + d_{ij})^\gamma$. Based on this, we can define a location's market potential as the distance-adjusted sum of aggregate demand in different locations:

$$NMP_i = \sum_j y_j L_j ((1 + d_{ij})^\gamma)^{1-\sigma} \tag{17}$$

Head and Mayer (2004) call this concept nominal market potential to distinguish it from the related concept of real market potential that adjusts for the variation in the price index across locations. The rationale for the concept of real market potential is that if a firm faces greater competition in a given market, as captured by a lower price index in that market, this limits the effective demand it faces. To see this, re-write the total revenue of location i in (14) as:

$$\sum_j \frac{y_j L_j (p_i (1 + d_{ij})^\gamma)^{1-\sigma}}{P_j^{1-\sigma}} = RMP_i p_i^{1-\sigma} \tag{18}$$

where real market potential of location i, RMP_i, is defined as

$$RMP_i = \sum_j \frac{y_j L_j ((1 + d_{ij})^\gamma)^{1-\sigma}}{P_j^{1-\sigma}}. \tag{19}$$

This is also the concept of market potential used in Fujita *et al.* (1999) and Redding and Venables (2004). For measurement reasons, in the empirical section we focus on NMP_i, but the difference between both concepts is inconsequential for the insights we are interested in.

Market potential and trade costs. In a world with transportation costs, the central location has an advantage in terms of market potential, because of the relatively shorter distance to the other locations. The strength of that centrality advantage is non-monotonic in transport costs. If transport costs are prohibitively high, being in the center, rather than on the periphery, yields no additional benefit. Likewise, if transport costs are zero, then all locations have the same market potential, so there is no benefit to being centrally located. However, when transport costs are at an intermediate level, the central location enjoys better market potential, and as a result attracts a greater population.

It is straightforward to formally show this non-monotonicity result in a symmetric setting. Suppose the central location is equidistant from both peripheral locations, so that $d_{\ell_1 k} = d_{\ell_2 k} = d_{\ell k}$ and $d_{\ell_1 \ell_2} > d_{\ell k}$. All locations have the same exogenous productivity, so that $\bar{A}_i = \bar{A}$ for all i. Consider the case where γ goes to infinity, so trade is prohibitively expensive. It is easy to see that an allocation where each location has an equal share of the population and where the mill prices of all three goods are the same satisfies all equilibrium conditions (8), (12), (14), (15) and (16). Now consider the case where γ goes to zero, so trade is free. Again, it is straightforward to see that a symmetric allocation of population, with equal prices across locations, satisfies (8), (12), (14), (15) and (16).

If γ is strictly in between zero and infinity, the symmetric allocation is no longer an equilibrium. To see this, start from symmetry, with labor equally distributed across locations and the mill prices of all three goods the same. In that case, the price index of the central location, P_k, will be lower than the price index of either one of the peripheral regions, $P_{\ell_1} = P_{\ell_2}$. Because $d_{\ell_1 k} = d_{\ell_2 k} < d_{\ell_1 \ell_2}$, $\left(\sum_j (p_j(1 + d_{kj})^{\gamma})^{1-\sigma}\right)^{\frac{1}{1-\sigma}} < \left(\sum_j (p_j(1 + d_{\ell j})^{\gamma})^{1-\sigma}\right)^{\frac{1}{1-\sigma}}$. As a result, $u_k > u_{\ell_1} = u_{\ell_2}$, so people will have an incentive to move from the peripheral locations to the central location. Hence, in equilibrium the population of the central location will be larger than that of the peripheral locations.

What does this imply in terms of the relation between market potential, transport costs and population growth? The above discussion says that if transport costs drop, the population share of the central location first increases and then decreases. We also know that the market potential of the central location is weakly greater than that of the peripheral locations. To see this, notice that in a symmetric allocation $NMP_k > NMP_{\ell_1} = NMP_{\ell_2}$, as long as $\gamma \in]0, \infty[$. We already know that in that case, free mobility will make people move to the central location, so that $L_k > L_{\ell_1} = L_{\ell_2}$. Because the elasticity of substitution is greater than one, this implies that $y_k L_k > y_{\ell_1} L_{\ell_1} = y_{\ell_2} L_{\ell_2}$, so that a fortiori $NMP_k > NMP_{\ell_1} = NMP_{\ell_2}$. As a result, the central location has greater market potential. Hence, we can conclude that if transport costs drop, the population of the high market potential location first increases and then decreases. In what follows we summarize this result.

Result 1. *Consider a symmetric setting, with $\bar{A}_i = \bar{A}$ for all $i \in \{k, \ell_1, \ell_2\}$ and $d_{\ell_1 k} = d_{\ell_2 k} = d_{\ell k}$ and $d_{\ell_1 \ell_2} > d_{\ell k}$. Starting at prohibitively high transport costs, a gradual drop in transport costs first increases the population share of the high market potential (central) location k, and later decreases the population share of the high market potential (central) location k.*

The second stage of this evolution, when the population share of the central location declines, is what we refer to as the shrinking advantage of market potential.

B.2 Numerical Examples

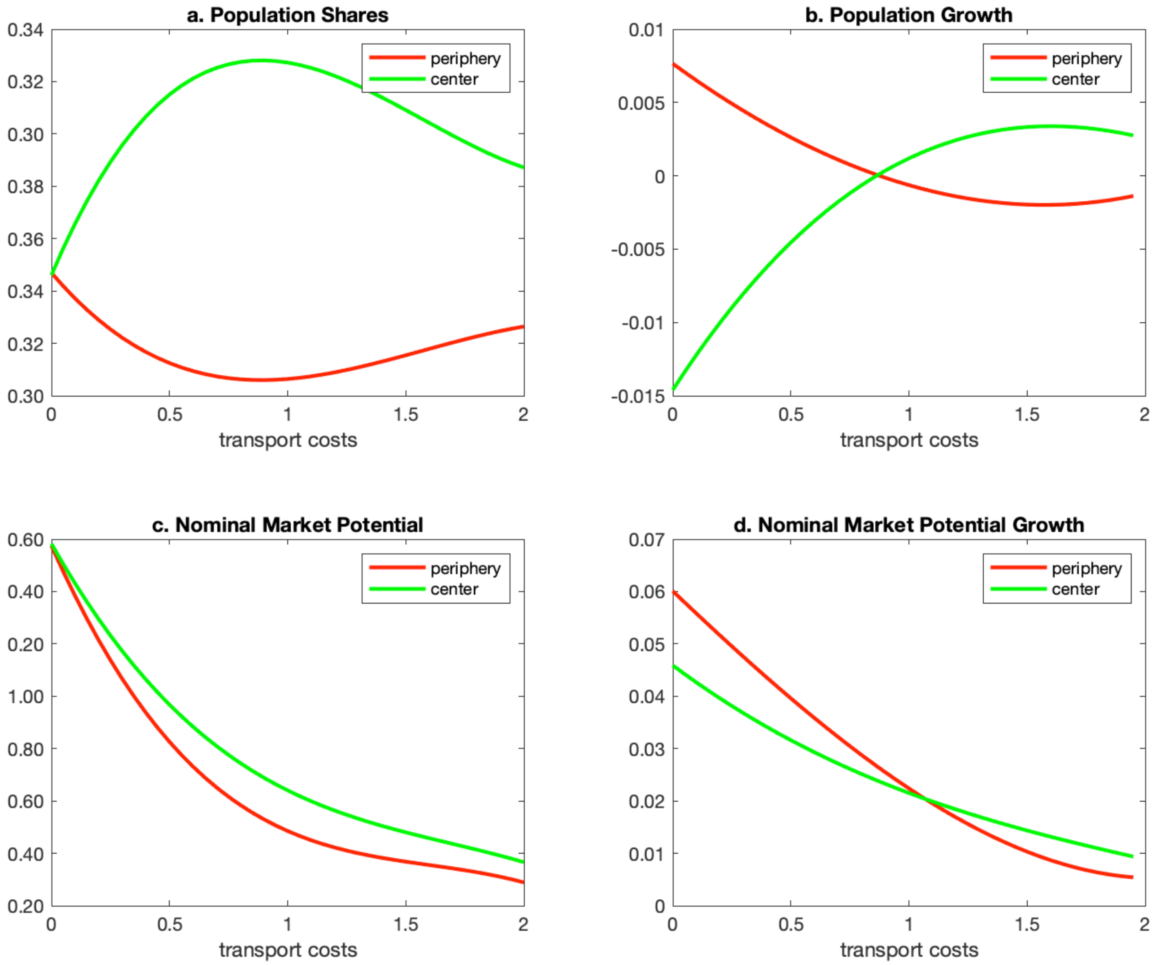
In this section we use numerical examples to further illustrate the different insights we can gain from the model. We start by discussing some of the parameter values. For the elasticity of substitution, we set σ equal to 3, a value within the range of those estimated by Simonovska and Waugh (2014). For the labor (or non-land share) we choose a value for α of 0.7. This is slightly high if interpreted as the labor share, but slightly low if interpreted as the non-land share. For agglomeration economies, the parameter that measures the elasticity of TFP to population, ε, is set to 0.05, as estimated in Behrens, Duranton and Robert-Nicoud (2014). The total population is set to 9, but this choice is inconsequential, as the entire problem can easily be rewritten in terms of population shares, rather than absolute levels.

The objective of the numerical examples is to show how the population shares of the central and the peripheral locations vary with trade costs. To that end, we will conduct comparative statics over a range of values of the trade cost parameter γ. Disdier and Head (2008) carry out a systematic analysis of 1,467 estimates of γ in 103 papers, and find a range from -0.04 to 2.33, with a mean of 0.91 and a median of 0.87. We will therefore run our model for a range of $\gamma \in [0.0, 2.0]$, keeping in mind that on average the world is a bit below 1.0. For less developed countries the number is bound to be larger than for more developed countries. For example, for the case of Brazil, Daumal and Zignago (2010) estimate an elasticity of 1.9.

Symmetry. We start by illustrating the non-monotonic relation between the population share of the high market potential central location and trade costs. To focus on the sole effect of centrality, we choose a symmetric setup, where each location has the same exogenous TFP level, $\bar{A}_i = 1$, $i \in \{\ell_1, \ell_2, k\}$, and where the distance between the central location and the two peripheral locations is identical, $d_{\ell_1 k} = d_{\ell_2 k} = 1$ and $d_{\ell_1 \ell_2} = 2$.

The results are shown in Figure A6. Consistent with Result 1, in panel a. we see that lowering trade costs initially benefits the central high-market-potential location in terms of population share, but eventually that benefit is eroded as trade costs drop to zero. As a result, panel b. shows population growth being higher in the central location when transport costs are high, but then shifting to the peripheral locations when transport costs drop enough. As can be seen in panel c., this means that with high enough transport costs, population growth concentrates in the high market potential region, whereas for low enough transport costs, population growth shifts to the low market potential regions. The reason for this shift can be seen in panel d. Starting off at high transport costs, a lowering of these costs increases the market potential of the central region

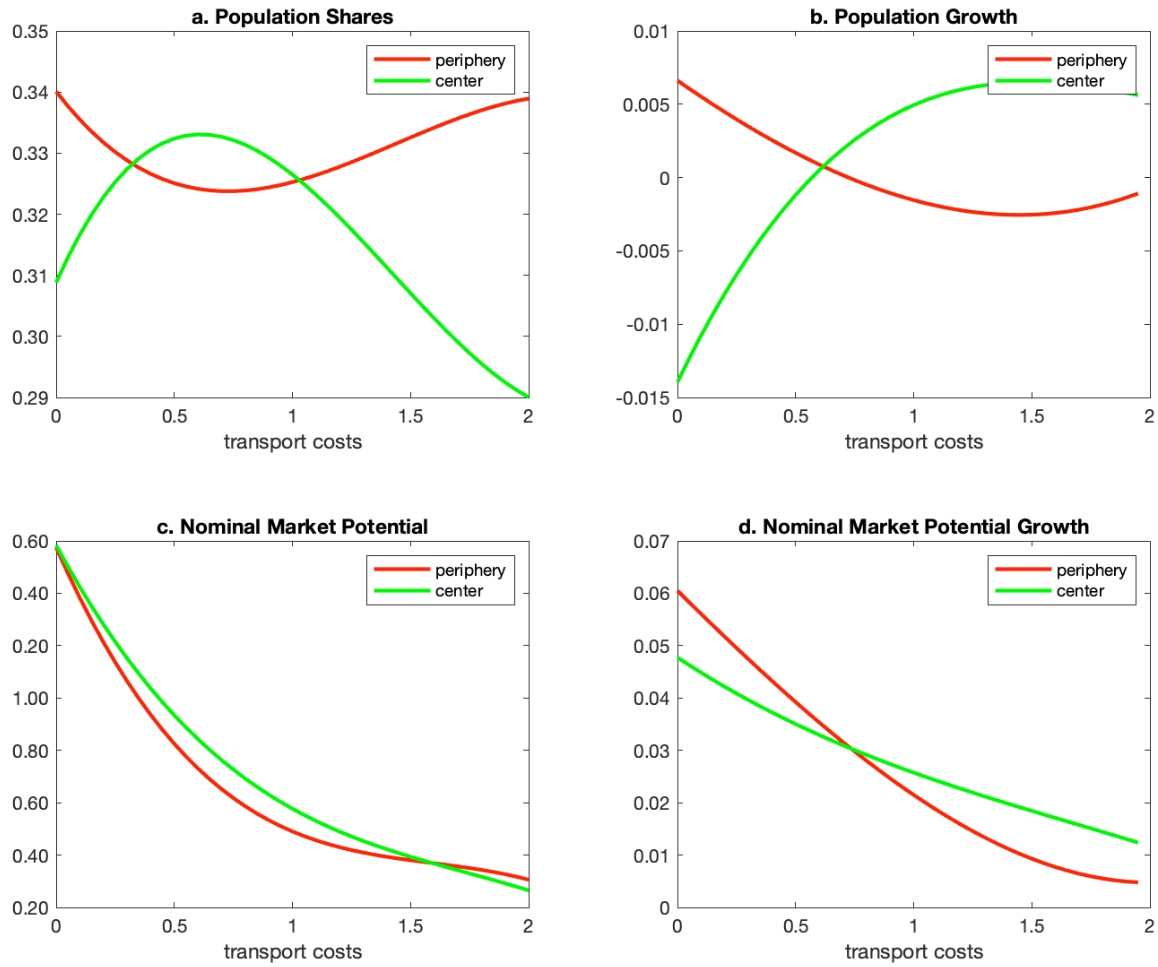
Figure A6: Market Potential and Trade Costs: Symmetric Case



more than that of the peripheral regions. However, if transport costs keep falling, eventually the peripheral regions catch up in terms of market potential. Hence, with low enough transport costs, the market potential of the peripheral regions grows faster than that of the central region.

Low-productivity central location. In an alternative exercise, we give the peripheral locations a higher exogenous level of TFP than the central location. In particular, we set $\bar{A}_{\ell_1} = \bar{A}_{\ell_2} = 1.075 > \bar{A}_k = 1$. Figure A6 shows the results. The dynamics do not change qualitatively: as transport costs drop, initially the central location gains in population share, but eventually population growth shifts to the peripheral locations. The only difference is that the central location no longer always coincides with the high market potential location. With high enough transport costs, the greater market access of the central region no longer compensates for its lower TFP, so that its market potential is lower.

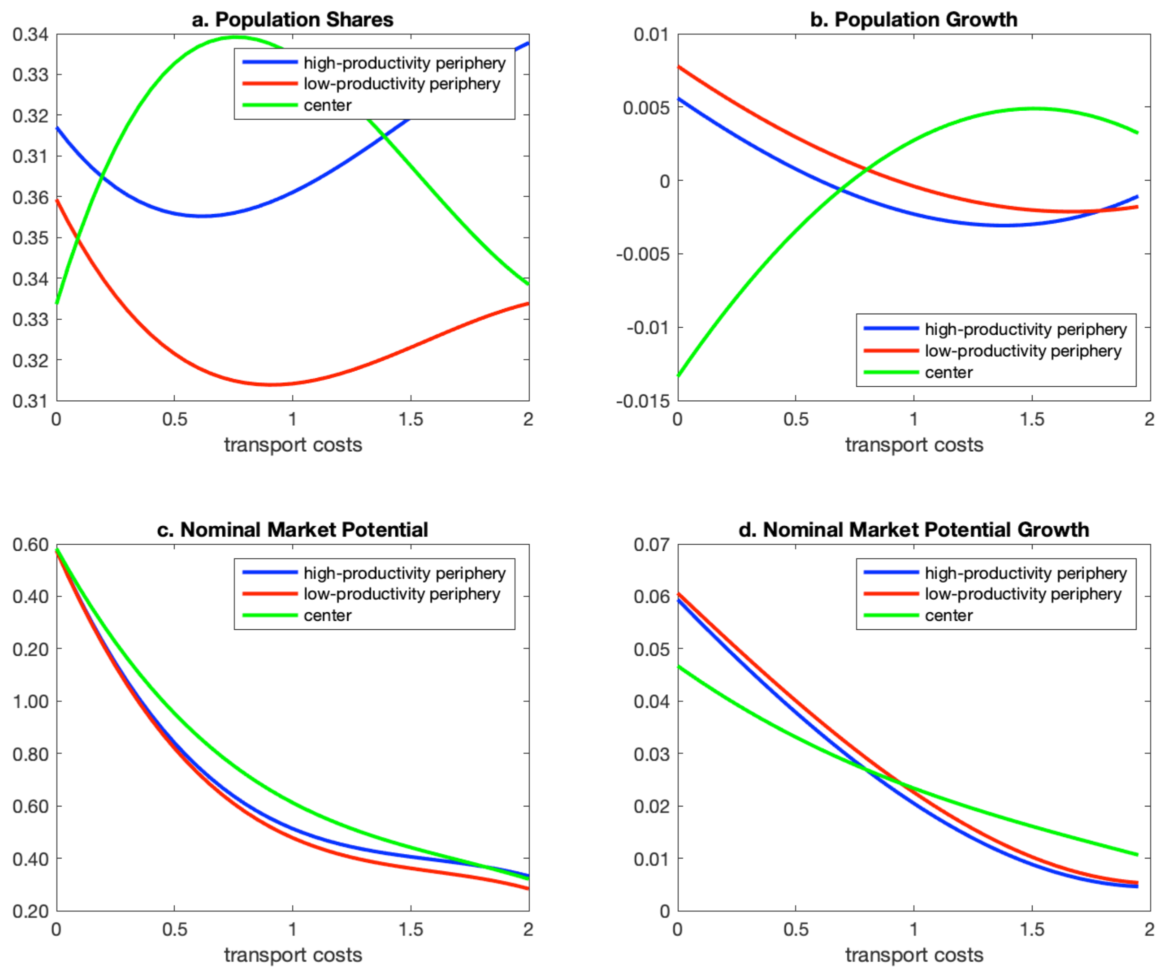
Figure A6: Market Potential and Trade Costs: Low-Productivity Center



Asymmetry in productivity of peripheral locations. We now introduce an asymmetry between the exogenous productivity levels of the two peripheral locations. In particular, we set $\bar{A}_{\ell_1} = 1.05$, $\bar{A}_{\ell_2} = 1.025$ and $\bar{A}_k = 1.0$. Figure A6 plots our findings. As before, when transport costs drop, the central region first gains in population, but later growth shifts to the periphery. When it does so, the less populous and less productive peripheral location experiences faster population growth than the more populous and more productive peripheral location. As can be seen in panel d., this happens because the drop in transport costs benefits the least productive peripheral region relatively more. This is what we would expect: relative to the less productive peripheral region, the more productive peripheral region gets more of its market potential from its own demand, and hence relies less on the other two locations.

Asymmetry in distance and productivity. In this last simulation, we simultaneously introduce asymmetries in the exogenous productivity levels and in the distances to the peripheral regions.

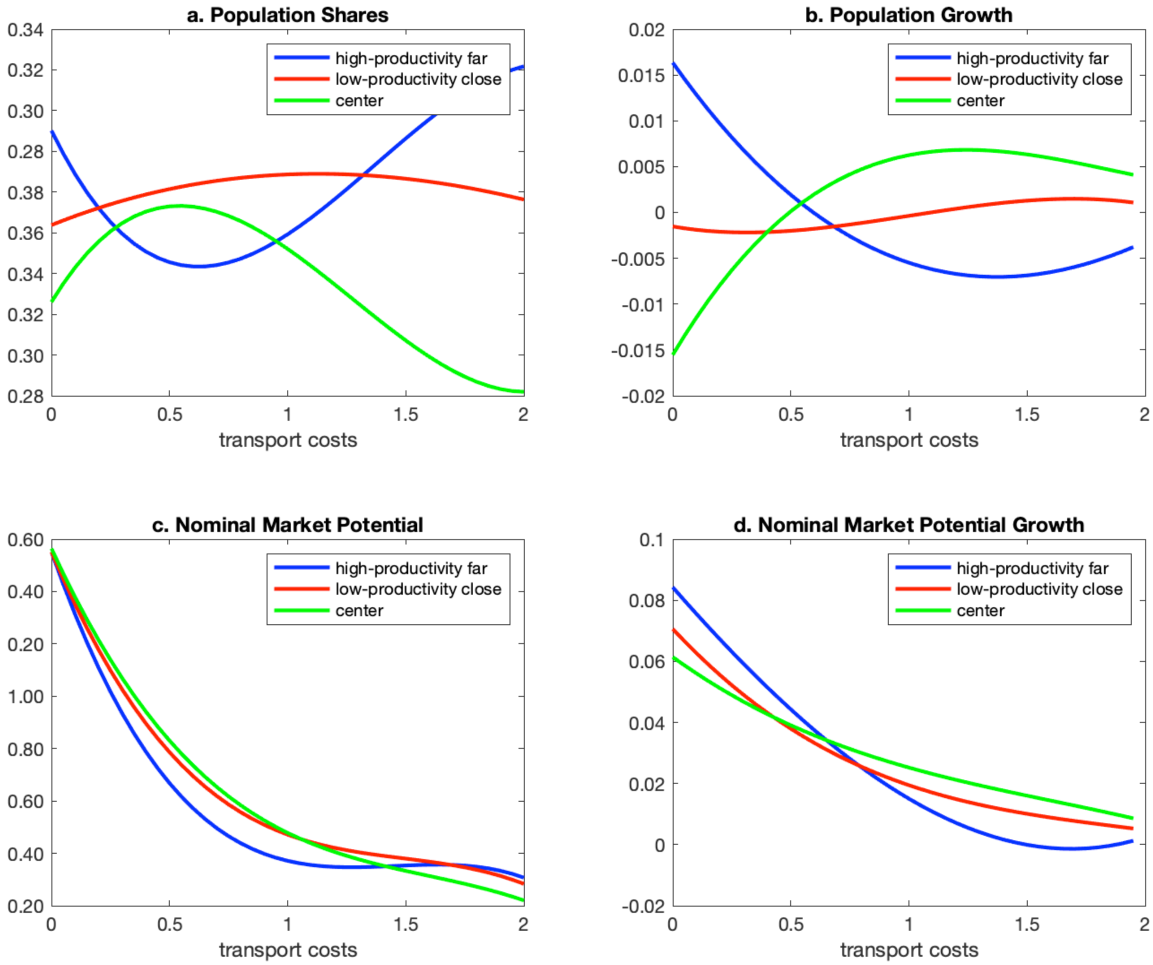
Figure A6: Market Potential and Trade Costs: Asymmetry in Productivity of Periphery



In particular, we set $\bar{A}_{\ell_1} = 1.1$, $\bar{A}_{\ell_2} = 1.05$ and $\bar{A}_k = 1.0$, and we set $\bar{d}_{\ell_1 k} = 2$, $\bar{d}_{\ell_2 k} = 1$ and $\bar{d}_{\ell_1 \ell_2} = 3$. Now, as transport costs drop, population growth shifts from the central region to the high-productivity peripheral region, rather than to the low-productivity peripheral region (Figure A6). With low enough transportation costs, growth is concentrated in the high-density, low-market-potential location. The difference with the previous exercise is that the high-productivity peripheral location is further removed from the center, so a drop in trade costs benefits it more than the other, lower-productivity, peripheral location.

Summary of results. The different numerical examples show that growth tends to concentrate in the locations where market potential improves relatively more. This insight is useful to further explain two important findings. First, when a drop in trade costs yields faster growth in the low-market-potential peripheral locations than in the high-market-potential central location, this is not because market potential does no longer contribute to a location's attractiveness. Instead, it is

Figure A6: Market Potential and Trade Costs: Asymmetry in Distance to Periphery



because the drop in transport costs improves the market potential of the peripheral regions more than it improves the market potential of the central region. As such, it shifts growth from the center to the periphery. Second, when growth moves away from the center, it benefits the less dense of the two peripheral locations if both are at equal distance from the center, but it benefits the more dense peripheral region if it is further away from the center. Once again, this difference depends on which of the two peripheral regions gains relatively more market potential when trade costs drop.