

Keith Head, Thierry Mayer and Gianmarco Ottaviano
**A review of volume 5 of the handbook of
regional and urban economics.**
Article (Accepted version)
(Refereed)

Original citation:

Head, Keith, Mayer, Thierry and Ottaviano, Gianmarco. (2017). *A review of volume 5 of the handbook of regional and urban economics*. [Journal of Regional Science](#) ISSN 0022-4146

DOI: [10.1111/jors.12356](https://doi.org/10.1111/jors.12356)

© 2017 Wiley-Blackwell

This version available at: <http://eprints.lse.ac.uk/84052/>

Available in LSE Research Online: August 2107

LSE has developed LSE Research Online so that users may access research output of the School. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LSE Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain. You may freely distribute the URL (<http://eprints.lse.ac.uk>) of the LSE Research Online website.

This document is the author's final accepted version of the journal article. There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

A Review of Volume 5 of the *Handbook of Regional and Urban Economics*

Keith Head	Thierry Mayer	Gianmarco Ottaviano
UBC Vancouver	Sciences Po Paris	LSE and U Bologna

For the *Journal of Regional Science*

1. International trade and spatial economics: the great (re)divergence

The 5th volume of the Handbook of Regional and Urban Economics is a mammoth endeavour of more than 1600 pages. It reflects the vibrancy of a field that has evolved in many different directions since the more focused times when the previous volume was conceived. The table of contents sets up a menu whose richness may appear overwhelming to some readers. The two sections of Volume 5A are entitled Empirical Methods (**Chapters 1-3**) and Agglomeration and Urban Spatial Structure (**Chapters 4-10**). They are followed by the two additional sections of Volume 5B devoted to Housing and Real Estate (**Chapters 11-15**) and Applied Urban Economics (**Chapters 17-23**).

With twenty three chapters of around seventy pages each, this review will necessarily be very selective. Specifically, as contributors to the 4th volume of this Handbook, we have decided to frame it as a comparison between the state of the field of spatial economics circa 2003 (when the chapters of that volume were written) and the current state as reflected in the 5th volume. What questions have been of continued interest over the whole period? What new methods have been deployed to address these questions and to what extent have these methods

brought about revised answers? We observe that interest has waned in some of the questions that attracted serious interest in 2003 and speculate on why. On the flip side, we observe in the latest edition that new questions have surged to importance. Perhaps the most obvious case is the economics of housing markets, which forms the subject of seven chapters (11-16 and 19) in the present volume, whereas “housing” does not even appear in the table of contents of the 4th Volume. The 2008 collapses of housing prices and mortgage-based securities have probably contributed to drawing the attention of urban economists. The housing chapters are to be considered thoughtfully in the companion review by Thomas Davidoff so we will concentrate on the non-housing chapters in this review. We will conclude this review by offering our views on the questions that deserve renewed focus and also suggest some new questions that have not yet become part of mainstream research.

The 4th volume of this series was published at the high point for the strand of research known as the New Economic Geography (NEG). It was a period when, united by interest in research by Paul Krugman, trade economists and spatial economists associated closely with each other. We attended the same conferences and worked on similar topics. We debated what was new and what was valuable about the NEG—and whether the two sets overlapped. The Nobel prize received by Krugman in 2008 validated this line of research but also coincided with the time when it faded significantly from the priorities of urban economists.

Since then, with some prominent exceptions, trade and spatial economists have gone their separate ways. In this review we characterise with a broad brush the directions we perceive spatial economists have moved, as exemplified in chapters of this very valuable Handbook. Given the respective sizes of the Handbook (very long) and this review (very short), we will be highly selective in the material we discuss. We will also not shy away from expressing idiosyncratic rather than balanced views.

In some sense the 4th volume appeared at a very crucial time for both trade and urban economics. It was when both fields were hit by a major shock: the increased availability of

individual level data on firms and people as well as all sorts of GIS data. What this type of data made impossible to neglect any further was the vast heterogeneity observable at the individual level.

Heterogeneity itself was not at all new in both fields. Was space heterogeneous? Certainly different plots of land (whether countries, regions, cities or neighborhoods) had different characteristics, for example in terms of climatic conditions and access to natural resources. Were people heterogeneous? Of course, for example they differed in skills though we were happy with few broad skill groups. Were firms heterogeneous? Yes, they differed across sectors as different sectors implied different technologies. These dimensions of heterogeneity were indeed at the core of many studies in trade and spatial economics based on the concept of comparative advantage and its fundamental implication that, when people living in different plots of land are allowed to trade, they specialize in economic activities that are relatively cheap for them to perform given their skills. They then export the outputs of the economic activities they specialize in and import the others.

This is a theory of spatial specialization and trade founded on heterogeneity that both spatial and trade economists were extremely familiar with. Since the 1970's to the eyes of both groups it had, however, two main limitations. First, from the viewpoint of trade economists, it had the prediction that countries exchange the output of different sectors and the more so the more different they are. This prediction had become increasingly counterfactual after the surge of intra-industry trade between rich similar countries in the aftermath of WWII. Second, from the viewpoint of spatial economists, heterogeneity in exogenous characteristics fell short of fully explaining the observed heterogeneity of endogenous outcomes across different places, especially at higher levels of spatial disaggregation. Trade economists were looking for mechanisms that would explain why similar countries would trade similar goods, and spatial economists were looking for mechanisms that would explain why cities and regions with similar exogenous attributes would endogenously develop different patterns of specialization and population density. It turned out that models of monopolistic competition in which goods

mobility or factors mobility are not free were able to provide the mechanisms the two fields were looking for. These models were adopted in trade economics in the 1980's and in spatial economics in the 1990's where they met the glorious tradition of location theory and spatial competition. This created a common language and a decade of cross-fertilization between the two fields that was dubbed "new economic geography" (NEG). The 4th Volume crystallized the results of that joint venture. By 2004, however, new approaches were about to rise to prominence.

The main objective of NEG models was to show how the geographical heterogeneity we observe around us can be generated endogenously by economic forces, even in the absence of any exogenous heterogeneity, due to localized technological and pecuniary externalities. Obstacles to the mobility of goods and factors affect the balance between positive externalities promoting the agglomeration of economic activities and negative externalities fostering their dispersion. When those obstacles change the balance between agglomeration and dispersion may change abruptly leading to catastrophic reorganization of the economic landscape. When agglomeration forces are strong enough, expectations can be self-fulfilling. Being able to explain the endogenous emergence of geographical heterogeneity was a great theoretical achievement.

The problem was that it had not been accompanied by comparable empirical success. As Duranton and Puga write in the 5th Volume, "To be useful and become more than a speculation, a conceptualization must confront the empirical reality" (**Chapter 8**, p.472). Unfortunately some the most intriguing conceptualizations of NEG largely remained speculations. While there had been progress in documenting spatial concentration, the empirical review by Head and Mayer in the 4th volume could not point to papers that clearly implicated plant-level increasing returns (the mechanism emphasized by Krugman) as a major source of agglomeration. It was also very hard to find unambiguous indications of the type of multiple equilibria that made NEG theory so intellectually exciting.

The desire of trade and spatial economists to confront empirical reality materialized at a time when new data (and new methods) were spilling over from other fields, mainly labour economics and industrial organization. Both fields had long been interested in individual heterogeneity and had been increasingly tapping worker- and firm-level datasets. Labour economists were mainly concerned with heterogeneous workers' outcomes while industrial economists focused on heterogeneous firms' outcomes. Spatial economists appear to us to have moved more in the direction of labour, both in terms of using similar worker-level data sets and in terms of greater focus on identification of treatment effects. Trade economists, on the other hand, have in some respects followed industrial organization, in terms of using firm-level data and in terms of tying in closely to theoretical models. Perhaps increased availability of micro data is a unified explanation for divergence as trade economists embraced firm level customs data sets at the same time as urban economists embraced labour (and housing) data sets.

Though there are still some points of contact, the 5th volume of the Handbook largely testifies this divergence since 2004. We would argue, however, that the stage is now set for renewed collaboration. Trade economists are increasingly using data on individual workers and urban economists have embraced structural models. Thus the current separation between trade and spatial economics is probably mainly attributable to focus on different questions. Towards the end of this review, we suggest some questions that seem likely to attract interest from researchers from both fields.

2. New methods

A critical difference between the empirical literature in spatial economics circa 2004 and nowadays is the insistence on what could be called "meaningful" empirical analysis. The era of reduced form analysis purely intended to uncover an interesting spatial fact can be considered as over (at least for publication in general interest or top field journals). Part of the reason might be because most facts are now well established, but the bulk of the explanation is that

the field has matured and adopted the general norm of economics: researchers should convince the reader either through a rigorous causal analysis, or structure its empirics with strong theoretical background. In both cases, the empirical finding has deep meaning and usefulness since it points to variables that can truly be considered exogenous and potentially acted upon to achieve a certain objective in terms of spatial activity. The last methodological improvement is common to other fields, in that access to individual level data (based on firm or consumer/worker level information) has become much easier. This has allowed for new identification strategies used both for the causal and the structural approaches.

Chapter 1 (by Baum-Snow and Ferreira) of this new volume is called “Causal Inference in Urban and Regional Economics.” This summarizes well the new importance given to causality issues in the field. A very interesting point of departure is the quantification of the share of “causal” methods (simply going beyond cross-sectional OLS, as defined in Table 1 of this chapter). If the slope seems promising, the level is still quite low in terms of usage of advanced methods in urban economics. The big revolution seems to be mostly in the use of panel data econometrics over the last 30 years, with instrumental variables still present in less than half of empirical papers, and randomization and matching being extremely rare.

A particularly illustrative example of the increased importance of causality concerns the estimation of agglomeration economies. In the Rosenthal and Strange chapter of 2004, only two pages strictly talked about endogeneity, while at least 15 pages concern the same issue in the new chapter by Combes and Gobillon (**Chapter 5**). Another very good example is the treatment of place-based policies in this handbook, in the chapter by Neumark and Simpson (**Chapter 18**). One of the prime justifications of spatially targeted policies designed to attract activity is the existence of some form of local external gain from the policy measure. Most notable are spillovers to the productivity of neighboring firms when the policy successfully attracts a foreign investor or a large plant. This makes it at least possible that public policy is needed to promote more agglomeration than the market would do on its own. Therefore a correct quantification of agglomeration economies is central to motivate those policies.

Neumark and Simpson devote a whole section to this issue, starting with the premise that estimating agglomeration economies raises important omitted variable bias and pointing to studies using geological and historical instruments of density. It seems unlikely, however, that economists will ever be able to manipulate the density of economic activity in a truly experimental way. Moreover, merely showing a link between density and productivity (or some other urban performance outcome) in a more causal way does not uncover the mechanisms underlying agglomeration.

Neumark and Simpson (**Chapter 18**) point to important selection issues in measuring productivity gains from density. If higher productivity firms self-select into more dense areas, the causal impact of density will be over-estimated. A key usage of micro data is to be able to control for firm fixed effects that deal with all unobservables that might be confounded with agglomeration economies. Naturally, this does not solve all endogeneity problems, but conditional on firms not changing location, this should control for most of the endogenous sorting effects. Another important use of such data is to document heterogeneity in agglomeration economies. Combes et al. (2012) document in particular that the most productive firms are the ones benefitting most from spatial externalities. This has important consequences for the design of place-based policies.

One area where the identification of causal effects has proven especially thorny is the estimation of how proximate individuals (or firms) influence each other. While Volume 4 contained a chapter on neighbourhood effects, Volume 5 contains *two* chapters, largely devoted to the issue. Again we see urban economists taking inspiration from labour economists, this time drawing on the literature on peer effects. **Chapter 3**, “Spatial Methods” (by Gibbons, Overman, and Patacchini) focusses on the problems of measuring how one entity’s behaviour (e.g. R&D spending or criminal activity) is influenced by the corresponding average behaviour in the same sub-space or cluster. **Chapter 9**, “Neighborhood and Network Effects” (by Topa and Zenou) also investigates the question of estimating non-market interactions. As with Chapter 3, they use criminal activity as a motivating example. Both chapters contain

subsections about the “reflection” problem (expressed concisely in Chapter 9 as “I affect my social contact and she simultaneously affects me”). Topa and Zenou’s chapter is distinguished by its focus on “social spaces” which can be very different from physical spaces like classrooms or industrial districts. In particular, they think of social spaces in term of networks of linked agents.

While Chapters 3 and 9 overlap in their coverage, they differ considerably in the overall message conveyed. Topa and Zenou are optimistic regarding the promise of empirical work on network effects. They highlight results showing neighbourhood effects, for example on crime in Sweden. They advocate a tractable linear-quadratic model that provides comparative statics and a guide to estimation. In addition to various types of experimental evidence, Topa and Zhenou advocate identification based on the partial overlap characterizing real-world networks: “my friends’ friends may not ... be my friends.” Gibbons et al emphasize the obstacles to achieving persuasive identification. They demonstrate that cluster randomization will not solve the reflection problem. Moreover, techniques involving spatial differencing tend to require strong restrictions on the structure of interactions, often with ad hoc justifications. As data on networks becomes more widely available—in part because it is being created by online social networks—both chapters should become required reading.

There is an interesting bridge between causal inference and structural analysis in **Chapter 1** by Baum-Snow and Ferreira (p.19): “While perhaps not ideal, there are many contexts in which neither randomization nor credible strategies for controlling for unobservables are available to recover treatment effects of interest. The main alternative viable strategy is to explicitly model the heterogeneity and sorting equilibrium and recover treatment effects through model simulation.” Holmes and Sieg discuss such structural options at length in **Chapter 2**. They conclude the chapter by urging urban and regional economists to rise to the challenge of developing and estimating quantitative models, arguing that such models have already proven valuable in other fields of economics.

A good example of how structural analysis and causal inference can complement each other can be found in the chapter by Redding and Turner (**Chapter 20**) where they survey the theoretical and empirical literature on the relationship between the spatial distribution of economic activity and transportation costs. They first develop a multi-region model of economic geography that they use to understand the general equilibrium implications of transportation infrastructure improvements within and between locations for wages, population, trade, and industry composition. They then use the predictions of their model to guide the subsequent review of the empirical literature on the effects of transportation infrastructure improvements on economic development. In doing so, they emphasize the use of exogenous sources of variation in the construction of transportation infrastructure. Redding and Turner point out four main advantages of a structural approach (p. 1388). First, this approach enables general equilibrium effects to be captured. Second, a structural approach allows for the estimation or testing of specific economic mechanisms. Third, the estimated model can be used to quantify aggregate welfare effects. Fourth, the estimated model can be used to undertake counterfactuals and generate ex ante predictions for the effects of policies that have not yet been implemented.

3. Questions

The first thing that comes to mind when going through the table of contents of the 5th volume is that spatial economics has become much more *urban* than regional. It has also become essentially intra-national with virtually no international trade dimension. The key chapters in which the current conceptual framework of spatial economics is spelled out are those by Behrens and Robert-Nicoud (**Chapter 4**), Duranton and Puga (**Chapter 8**) and Redding and Turner (**Chapter 20**). Taken together, these chapters conceive the spatial economy as a national system of cities (sometimes with a surrounding rural area as, for example, in **Chapter 22** by Desmet and Henderson on developing countries) between which goods and factors flow freely (**Chapter 4**), not at all (**Chapter 8**) or with frictions (**Chapter 20**). Cities themselves may have a relevant internal geographical dimension (**Chapters 8 and 20**) or not (**Chapter 4**) associated

with land use and commuting costs. However, no chapter really provides a complete framework in which space matters both within and between cities, and the national system of cities interacts with the rest of the world (except through immigration as in **Chapter 10** by Lewis and Peri).

The three chapters spelling out the current conceptual framework of spatial economics share a common logical structure. They all start with a list of generally accepted “stylized facts” looking for an explanation. They then discuss the various explanations proposed in the literature. Moreover, **Chapters 8** and **20** look also at the empirical evidence supporting those explanations.

As all three chapters define the questions of interest starting from stylized facts, the facts list will tend to define the questions that are vital for current spatial economics. In **Chapter 4** on “Agglomeration Theory with Heterogeneous Agents”, Behrens and Robert-Nicoud use a cross section of US cities to document the following six stylized facts: (1) the population size and density of a city are positively correlated with the quality of its fundamentals (size and fundamentals); (2) the unconditional elasticity of mean earnings and city size is about 8%, and the unconditional elasticity of median housing rents and city size is about 9% (urban premiums); (3) the share of workers with at least a college degree increases with city size (sorting); (4) the share of self-employed is negatively correlated with urban density and with net entry rates of new firms (selection); (5) the Gini coefficient of urban earnings is positively correlated with city size and the urban productivity premium increases with the education level (inequality); (6) the size distribution of US places follows closely a log-normal distribution; and that of US metropolitan statistical areas (MSAs) can be closely approximated with a power law with parameter near one (Zipf’s law). We think that fact (5) deserves special emphasis since it could be compatible with a three-way linkage between trade, urbanization, and rising inequality.

In **Chapter 8** Duranton and Puga on “Urban Land Use” look at the distribution of land across uses in Paris to motivate why the study of urban land use should focus on the following four issues: the differences in land and property prices across locations; the patterns of location choices by types and subgroups of users; the patterns of land conversion across uses, and the patterns of residential and business location changes within cities.

In **Chapter 20** on “Transportation and the Spatial Organization of Economic Activity”, Redding and Turner use data on several countries to highlight the following stylized facts on transportation. Both a secular decline in transportation costs for goods and a change in the relative importance of different transportation modes over time (e.g., rail versus road versus air) and for value versus weight are observed between cities. Transportation costs for people continue to be important. In particular, within cities commuting costs remain substantial, both in terms of time and money. The effects of transportation infrastructure on the organization and growth of the spatial economy remain a core issue of spatial economics.

Which models are used to address all these questions? In **Chapter 4**, Behrens and Robert-Nicoud start with a simple model of a system of cities in the spirit of the canonical model of Henderson (1974). They then show how that model can be progressively enriched to arrive at an internally consistent explanation of all their stylized fact within the same theoretical framework. A key extension is to introduce heterogeneous agents in order to be able to talk about sorting; i.e., the heterogeneous location choices of heterogeneous workers and firms. Agent heterogeneity also allows for selection effects, either via occupational choice by heterogeneous workers as in Lucas (1978) or via an entry choice by heterogeneous firms as in Melitz (2003). Trade economists have also devoted considerable energy to the issue of selection during the last decade. While trade economists have mostly focused on firms’ entry choices, Behrens and Robert-Nicoud only deal with workers’ occupational choices. On the other hand, they rely on the venerated monopolistic competition model for market structure. In this respect, they use a key ingredient of NEG borrowed from the trade literature to extend the canonical model of a system of cities.

The main model utilized by Duranton and Puga in **Chapter 8** is an old acquaintance of spatial economics: the monocentric city model. They show that this workhorse of urban economics provides a surprisingly rich and flexible conceptual framework that can be extended to deal with multiple business centres, various dimensions of heterogeneity, and durable housing. The augmented model allows one to understand the main features of the complex distribution of land across uses in cities.

Both **Chapter 4** and **Chapter 8** discuss how theoretical modelling can help identification when one gets to the data. The importance of dialogue between theory and empirics is promoted even more strongly in **Chapter 20**. In this chapter Redding and Turner look at the literature that has tried to measure the effects of infrastructure on the spatial economy. They conclude that the current empirical literature provides credible causal estimates of the effects of transportation infrastructure. However, “it is impossible for the reduced-form regressions conducted in almost all of the empirical studies that we survey to separately identify the effect of transportation infrastructure on the growth and reorganization of economic activity” (i.e. whether changes in transportation costs affect the amount of economic activity or reorganize existing economic activity). They, therefore, suggest two approaches to this problem: a simple extension of the existing reduced-form literature, and the implementation of a structural model. Their model is very interesting and opens a window on what a large part of spatial economics may look like in the near future. As explained by the authors (see Footnote 13, p. 1356), it builds on the NEG model by Helpman (1998), parameterized so as to avoid multiple equilibria. While NEG models assume firm product differentiation and monopolistic competition, the model by Redding and Turner shares many properties with perfectly competitive stochastic trade models of “discrete choice” à la Eaton and Kortum (2002), which are the pillars of the recent wave of new quantitative models that are changing the way trade economists look *ex ante* at the possible implications of alternative policy scenarios. This shows once more that, whereas the questions of interest may have largely diverged between trade and spatial economics, methods have not.

There are questions that were central in the 4th volume and have virtually vanished in the 5th volume. As already noted, the international dimension appears to have fallen into oblivion. There is no consideration of the ramifications of different accessibility of cities to international transactions for the national distribution of economic activities. There is little interest in the interactions among national economies, as these are generally assumed to be “closed” to any foreign shock. A notable exception is **Chapter 10** on “Immigration and the Economy of Cities and Regions” by Lewis and Peri. This chapter surveys studies that have tried to measure the effects of immigration on the receiving economy. It shows how a more careful consideration of how immigrants’ skills differ from natives’ and the adoption of model-based ways of testing local (vs national) responses to immigration have been crucial parts of the recent developments in this area of research.

Another set of questions that have somewhat surprisingly dropped off the radar screen are those on regional growth, at least in the case of developed economies. **Chapter 22** by Desmet and Henderson, entitled “The Geography of Development Within Countries”, reviews what is known about the spatial distribution of economic activities and development looking first at aggregate population growth across different locations, then at the incentives of different industries to concentrate or disperse, and finally at the evolution of the urban hierarchy. The emphasis is on how the spatial distribution of economic activity changes as a country grows and develops. The chapter shows how our understanding in this area is changing thanks to the rapidly increasing availability of geographically disaggregated data for developing countries. On a different note, **Chapter 6** by Carlino and Kerr on “Agglomeration and Innovation” surveys empirical studies on clusters as the engines of invention and innovation. They conclude (p. 397) that, while some progress has been achieved in the last decade, “we still have not opened the black box of how clusters operate” and better insight is needed into the long-term relation between agglomeration and innovation, in particular on the “life cycles of innovative places”.

4. The road ahead

By this point it should be clear that we are eager to see renewed linkages between international trade and urban economics. Forecasting future research topics must always be suspect. After all, if the idea is so good why hasn't it already been written? With that caveat in place, we suggest three areas where we see fertile soil.

In keeping with our theme about "supply" side influences of availability of new data sets, we think there is valuable work to be done using firm-employee matched data sets. Firms are the entities identified in trade statistics so if we want to see the impact of trade on worker outcomes or vice versa, using such matched data seems promising. The work by Helpman, Itskhoki, Muendler and Redding (2016) is one prominent paper to advance this approach.

A second area of active interest is a new line of economic geography combining gravity equations within countries as delivered by the Eaton and Kortum (2002) model. We are thinking of the recent paper by Arkolakis and Allen (2014).

A third area where regional and trade issues interact is in the response of migration to trade shocks, such as the emergence of China as an export powerhouse. There appear to be intra-national migration asymmetries. In recent work Autor, Dorn and Hanson (2016; ADH) point to the "migration of 250 million workers from farms to cities" as a factor contributing to rising Chinese exports of labour-intensive goods. It seems likely there was also reverse causation. Namely, increased export success in China's coastal cities motivated the migration of farm workers. On the flip side, in the countries experiencing rising Chinese imports, ADH report that "there is little geographic migration in response to the trade shock." Can it really be the case that workers only move *to* boom areas but do not exit *from* areas where labour markets are depressed by trade shocks? Since functioning adjustment paths are an important condition for gains from trade, and such adjustment may need to take the form of migration across cities, this seems like a promising area for the combined efforts of trade and urban economists.

A related area where trade and urban and regional studies stand to gain from collaboration relates to spatial specialization. Both trade and labour economists have become interested in task-oriented views of production, partly to understand offshoring and partly to investigate the causes of labour market polarization (see for example, Goos et al, 2014). There would seem to be interesting trade/geography research to be done on fragmentation of production tasks across space, *between and within nations*.

In closing we commend the authors and editors of the 5th volume for a Handbook that showcases the advances of urban economics, both in terms of method and substantive contributions. The toolkit now available is both powerful and versatile. Our hope, and indeed our expectation based on work we have been seeing at seminars and conferences, is that future applications will contain a larger international dimension.

References

Arkolakis, C., Allen, T., 2014. Trade and the topography of the spatial economy. *Quarterly Journal of Economics* 129, 1085-1140.

Autor, D., Dorn, D., Hanson, G., 2016. The China Shock: learning about labor market adjustment to large changes in trade. *Annual Review of Economics*, forthcoming.

Eaton, J., Kortum, S., 2002. Technology, geography, and trade. *Econometrica* 70, 1741–1779.

Goos, M., Manning, A. Salomons, A. 2014. "Explaining Job Polarization: Routine-Biased Technological Change and Offshoring." *American Economic Review*, 104(8): 2509-26.

Helpman, E., 1998. The size of regions. In: Pines, D., Sadka, E., Zilcha, I. (Eds.), *Topics in Public Economics*. Cambridge University Press, Cambridge, UK, pp. 33–54.

Helpman, E., Itskhoki, O., Muendler, M., Redding, S., 2016. Trade and inequality: from theory to estimation. *Review of Economic Studies*, forthcoming.

Henderson, J.V., 1974. The sizes and types of cities. *Am. Econ. Rev.* 64, 640–656.

Lucas Jr., R.E., 1978. On the size distribution of business firms. *Bell J. Econ.* 9, 508–523.

Melitz, M.J., 2003. The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica* 71, 1695–1725.