

New Frontiers in Urban Economics Research: A personal view

Gilles Duranton
University of Pennsylvania

World Bank, January 31, 2023

Introduction

The challenge of a "new frontier" talk

What I thought of doing:

- ~~My last paper~~
- ~~A convex combination of my last papers~~
- Explore a specific subtheme
 - ~~Land use~~
 - ~~Accessibility and urban mobility~~
 - ~~Costs and benefits of cities~~
 - ~~Regional disparities~~
 - ~~Infrastructure and development~~
- Instead, take a look at the overall landscape and present something I would have liked to listen to when younger
- Of course, this is very selective and a very personal vision

Modeling

Damned if you do,
damned if you don't

Modeling is making a comeback in urban economics

- Some papers don't need models because the model has already been written (e.g. agglomeration economies) or is perhaps self-evident (e.g. program evaluation)
- But we are now moving to novel and sometimes more complex (detailed) questions which require a clear conceptualization
- With more data, we can also go deeper into existing questions and explore specific mechanisms, which need to be clearly articulated
- But urban economics is also where everything interacts with everything else and we can never be sure we know the correct model
- Models are needed but model uncertainty reigns

Modeling

Data vs. models: Who should prevail?

People have very different opinions about this

- People who “bring their model to the data” have already decided that their model must be right and downplay (ignore) model uncertainty
- People who want to “estimate stuff” and let “the data talk” have essentially decided they did not care much (could not care less) about modeling
- I’m pushing typical approaches to the extreme to show neither data nor model should dominate
- Ideally, use an initial model to guide the empirics and revise it as progress is made in the data exploration, etc.
- Downside: this is slow and painful. (Diego Puga and I took seven years for our last paper...)
- Still, use the model to discipline the empirics and vice-versa

Modeling

**Work more at
the intensive
margin and less
at the extensive
margin**

Illustration: Quantitative Spatial Models

- Here is a cool new tool allowing its user to model any city and recover a lot of stuff (wages, amenities, etc) for every part of a city from very little (e.g. commuting data) (Heblich et al., 2019, Kreindler and Miyauchi, 2019)
- The temptation is to use the tool at the extensive margin and add more to it
- At the same time, the approach still suffers from (serious) flaws. Among others, implausible values of time or parameters that play different role and for which we would like very different values (Duranton and Puga, 2004, Ahlfeldt et al., 2015)
- Failure to improve the core of the key model(s) will make their flaws increasingly salient and eventually they will get abandoned and forgotten (eg, Hotelling type models, NEG, etc.)

Modeling

Some new modelling tools

Besides Ahfeldt et al., some recent progress:

- Dynamic location choices (Artuc et al., 2010, Caliendo et al., 2019)
- Dynamic spatial models with growth (Desmet and Rossi-Hansberg)
- Network modeling (Allen, Arkolakis and co-authors)
- "Reduced-form" modeling of cities in models of many cities (Henderson, Duranton, Puga, Diamond and others following on Henderson 1974)
- Modeling heterogeneity: differences across groups (sorting models following Bayer, Timmins, and co-authors, Epple, Sieg and co-authors, Behrens et al., 2014)
- Modeling heterogeneity: differences within group ("Frechet" or Gumbel-logit, BLP inc. Diamond, 2016)

Data

Urban economics: the first empirical turn

Historically, urban economics was born as a rationalization of obvious facts about cities and was a mostly theoretical enterprise (Alonso 1964 followed by Muth and Mills) and remain so for a long time until

- Ed Glaeser showed us that many urban questions could be explored with simple data (Glaeser et al, 1992, Ades and Glaeser, 1995, Cutler and Glaeser 1997, etc)
- Then, large scale administrative data became available and urban economics could follow in the footsteps of labor economics...
- ... and exploit their methods (Moretti 2004, Combes et al 2008, Holmes 1998, Bayer et al 2007, Baum-Snow 2007, Duranton and Turner, 2012 and many others)
- Administrative data are becoming more accessible allowing for interesting links (Raj Chetty and co-authors, data from Sweden)

Data

Empirical urban economics: we keep winning the data lottery

In the last 20 years the field has kept receiving new waves of data

- Remote-sensing data allowing to observe physical development (Burchfield et al. 2006, Harari, 2020, Baragwanath et al. 2021) and generate proxies for economic activity where data is missing (Henderson et al. 2012, Jean et al. 2016, Storeygard 2016, Kachiyan et al. 2022)
- Cellphones data and pings from other GPS devices to track fine features of the travel behavior of people (Couture et al. 2023) or vehicles (Allen et al. 2023), uncover migration decisions and social networks (Buchel et al. 2020), innovation (Atkin et al., 2020), etc.

Data

Empirical urban economics: we really keep winning the data lottery

In the last 20 years the field has kept receiving new waves of data (continued)

- Scrapped/web-collected data has often a location tag and can be used to study a variety of issues including transportation (Akbar and Duranton, 2019, Akbar et al. 2023, etc.), housing (Calder-Wang 2022, Almagro-Dominguez-Iino, 2022), urban segregation (Davis et al. 2019), urban change (Naik et al. 2017) and more to come!
- Firm/corporate/marketing data including price scanner data (Handbury and Weinstein, 2014, Handbury, 2021) banking data (Relihan 2022 and work by Gannong and Noel) and a lot more to come!

Data

Empirical urban economics: we really keep winning the data lottery (continued!)

- Old maps are rich sources of information which can now be processed with computer vision (Gorin et al., 2023, Combes et al. in progress) and much historical data can be geographically tagged (Heblich et al. in progress)
- The statistical apparatus of some countries is increasingly organized around a main map (France, Spain)
- And I am sure I am forgetting some other important novel data sources -- everything any economist may be interested in is located somewhere and we are increasingly able to access this information

Data

The challenges of having too much

- "Big data" is hard and costly to manipulate (and this leads to bigger research teams...)
- Many projects require merging several (large) datasets around a common geography
- Cool data is not always synonymous with useful data about important issues we care about
- Exploiting data to generate facts is useful and important but as economists we want to leverage data to learn about key economic mechanisms (and as young economists you need to show your skills)
- => Novel data is not a substitute for an important and well-posed research question, insightful modeling, and a good research design

Methods

Three types of methods

Urban economic research is about

- i. Acquiring/obtaining data from a four-dimensional world (inc. time)
- ii. Projecting/processing them onto a dataset with rows and columns of with numbers
- iii. Estimating quantities of interest out of this table/dataset

We have seen progress in the last few years on all three steps, but the second remains grossly neglected

Note: Having a lot of data with a lot of spatial detail also raises a data management challenge and we (I) need to get better at this

Methods

Data creation: predicting and classifying

Creating data is now a (big) "thing". A variety of predictive methods can be used to transform complex objects in the world into exploitable data with a location tag

- Textual analysis and language processing (Gentzkow et al. 2019) for anything from categorizing occupations (Michaels et al. 2019) to parsing out realtors' blurbs (Levitt and Syverson, 2008)
- Computer vision to recognize anything from symbols on a map (Combes et al. 2023) to objects on a picture (Naik et al. 2022)
- Computer audition someone?

Machine-learning approaches have made their splashy entrance into our toolkit both to create data (here) and extract patterns out of them. We need to learn to use appropriate approaches and use them appropriately

Methods

Delineating and choosing spatial units

Once we have data and observations with lats and longs, we need to choose which spatial units to use

- Working in continuous space is fun (Duranton and Overman, 2005) but often difficult or even impossible if the object of interest is a spatial construct like a city or a neighborhood
- "Official" and "semi-official" spatial units are often poorly defined for a variety of reasons (Duranton, 2021)
- There are many ways to define spatial units (JUE 2021) and research on this is just getting started

Much extent research uses the spatial units it uses "just because". This leads to a variety of problems (measurement error, MAUP, etc). The cost of delineating appropriate units for a given research project is no longer prohibitive and will increasingly become a requirement

Note: Thinking about spatial units is also thinking about the fundamentals of a question (esp. its scale)

Methods

Estimating

We have seen incremental progress on many fronts in empirical methods

- Spatial RDD (Cattaneo and co-authors, Imbens and co-authors), spatial "kinks" (Turner et al., 2014)
- "Constructed IVs" (Adao et al., 2019, Goldsmith-Pinkham et al. 2020, Borusyak et al., 2022a)
- Event studies (Borusyak et al., 2022b)
- Diff in diff (Rambachan and Roth, 2019, De Chaisemartin and d'Haultfoeuille, 2022)
- Heterogenous treatment effects (De Chaisemartin and d'Haultfoeuille, 2020)
- IO tools for discrete choice estimation (Anagol et al., 2022, Hsiao, 2023)

A lot of that research is (rightfully) trying to instill rigor in what were arguably adhoc heuristics and, perhaps more importantly, to handle heterogeneity

Topics

"New Frontiers": a word of caution

Let me now explore recent research on a few topics. Note:

- This is a personal choice. Some topics are related to what I have done. Others are related to things I would have liked to do or that I might consider doing on the future
- Some topics have seen a lot of recent work. Others not so much.
- Some of these topics relate closely to what has been presented earlier today and yesterday. I think I would have flagged them in other contexts as well

A word from an over-solicited tenure-letter writer: There is a huge premium associated with (i) having a body of work on a given topic and (ii) show a clear sense of direction and progression. At the same time, don't keep working on subjects that bore you or which are no longer promising

Topics

Endogenous amenities

- The literature has mostly thought about productivity benefits from agglomeration, but cities also offer consumption benefits (Glaeser et al., 2001). To avoid endogeneity, much of the literature for a long time focused on exogenous amenities in the Rosen-Roback literature (e.g. Rappaport 2007)
- Starting with Handbury and Weinstein (2014) and Couture (2019), the initial focus was on prices and accessibility.
- Subsequent work has sought to put this in broader GE frameworks (Ahlfeldt et al., 2015, Diamond, 2016), explore the underlying heterogeneity (Handbury, 2021), understand the retail supply response (Alcott et al., 2019) or explore further questions (Gorback, 2021, Oh and Seo, 2022, Relihan, 2022)

Topics

Endogenous amenities

Key questions to be answered:

- Lower prices vs. diversity benefits (urban accessibility as possibly core to the organizing framework)
- Spatial reach of urban consumption benefits (especially post-Covid)
- The relative role of fixed vs. endogenous amenities
- Income effects and sorting

Key challenges

- Be able to track fine patterns of consumer behavior in space (cell-phone, banking data?)
- Consumption most often does not take place in isolation (social networks)

IMHO, this set of issues is bound to become even more central to urban economics and attract interest from many other fields (PF, IO, etc.)

Topics

Channels of agglomeration economies

- Estimating agglomeration economies was central to empirical urban economies for many years. While we have found a relative consensus on their magnitudes, isolating the channels through which they percolate has vexed many of us for many years
- Recent developments include better evidence about learning (De La Roca and Puga, 2017), knowledge spillovers (Atkin et al., 2022), and labor market matching (Moretti and Yi, 2022) or IO linkages (Miyachi, 2022), the role firms (Giroud et al., 2022), within-industry effects (Moretti 2021)
- Areas to explore include, job to job mobility, more about firms and establishments, understanding how different sources lead to different aggregate effects, etc

After some years of neglect, agglomeration is becoming a big issue again and will likely stay so for many years

Topics

The costs of cities

- The fundamental tradeoff of urban economics is between agglomeration benefits and urban costs. The second term of this tradeoff was badly neglected for many years but there are signs this is changing
- This issue underlies problems associated with housing affordability, restrictive land use regulations, urban transportation, congestion, costs of construction, etc.
- Settings have been proposed to estimate overall urban costs or parts of them (Combes et al., 2019, Duranton and Puga, 2022). Progress has also been made on a range of sub-issues with more to come (see next)

Topics

The costs of cities

- Land use regulations (Anagol et al. 2022, Song 2022, Cui 2023, Gyourko et al., 2019, Hsieh and Moretti, 2019, Duranton and Puga, 2022)
- Housing costs (Combes et al., 2021, Peng 2022, Ahfeldt et al., 2023, Ahfeldt and McMillen 2018)
- Congestion (e.g., Akbar and Duranton, 2019/in progress and many others)
- Pollution (Adelco et al., 2021)
- "Local costs": Qian and Tan (2022)

Still a lot to do before we know as much about costs as we know about agglomeration

Topics

Urbanization and development

- The urbanization of the developing world raises a broad set of issues associated with
 - Weak property and protection rights
 - The four informalities (labor, housing, transportation, and credit) and how they interact
 - “Partially-durable” construction
- Recent developments have been many, let me flag the following:
 - Abebe et al. (2021)’s RCT on labor market access
 - Henderson et al. (2021), Michaels et al. (2021) and Harari and Wong (2023) on “partially-durable” built-up
 - Conwell (2023) on informal transit
 - Gollin et al. (2016) on the aggregate drivers of urbanization

Topics

Urbanization and development

- Opportunities are nearly infinite for program evaluation type approaches given the number of quirks that have been generated in several thousand cities. The main issue here is pick up a really interesting one.
- We will also need broader (i.e., macro) approaches to think about urban development
- Some areas are still terra incognita, e.g. housing finance, the "Raj Chetty" agenda of mapping opportunity
- Data are a key challenge. Although they are becoming more available, it is difficult / impossible to measure things back in time. Panels of household/worker data are seriously scarce

Urban economics is increasingly interested in developing cities and development economists are working in cities more often - will the "twins" ever meet?

Topics

Urban political economy

- Urban political economy was viewed as the topic of the future for many years (Helsley 2004) but it has remained so and is yet to take off
- These gaps persist despite the importance of politics in urban management and development. Recent contributions have focused on:
 - Local voting behavior (Ferreira and Gyourko, 2009), local political contributions (Yu, 2022), and local political pressures (D'Amico, 2022)
 - Evaluating mayoral quality (Sieg and Yoon, 2021)
 - The architecture of decisions (Whalley, 2013, Boffa et al., 2016, Tricaud 2022, sitting on a large literature after Tiebout, 1956, and work by Alesina and co-authors)
 - Effects of local taxation (e.g. Brulhart et al., 2012) and place-based policies (reviewed by Slattery and Zidar, 2019)

Perhaps, I am lacking knowledge and vision here but this feels piecemeal and lacks integration with the rest of the field

Topics

Resilience, persistence and path dependence

- As discussed at length yesterday and by the papers earlier this morning, we want cities to be more resilient/sustainable/durable and, at the same time, more adaptable to change (cakeism?)
- The urban resilience literature is slowly defining a set of core issues around values at risk and changes (or lack thereof) to avoid risks (Balboni, 2022, Lin et al., 2022, Mulder, 2022, Ostriker and Russo, 2023) beyond the standard evaluation of the effects of shocks (e.g., Boustan et al., 2020)
- Closely related but different: the much older literature on urban persistence (Davis and Weinstein, 2002) with a broad variety of outcomes (e.g., Hornbeck and Keniston 2017)
- This literature is in clear need of an organizing framework (important recent progress by Allen and Donaldson, 2022, and Yoo, 2022)

These are fundamental issues but we still need clarity about the basic facts (as per Mushfiq yesterday)

Conclusion

**The best of times
but the most
demanding of
times to be an
urban economist**

- Better models and better methods
- A LOT more data
- Abundant data does not mean easy data. It also means that the premium goes to its main complementary inputs: important questions with convincing research designs and careful execution
- I flagged a small number of research topics but, by no means, I meant to be exclusive. Other topics that I like include: microfoundations of the housing market, declining internal migrations in rich countries, spatial statistics, urban economic history, the geography of innovations, joining with networks economics, etc.

These are great times, truly but standards are going up, and pretty fast