2020 Lectures on Urban Economics

Lecture 3: The Benefits and Costs of Cities

Gilles Duranton (Wharton)

25 June 2020

Introduction

- Cities are about one key tradeoff, dubbed the 'fundamental tradeoff of urban economics' by Fujita and Thisse (2013):
  - Cities offer higher efficiency and better opportunities for workers, firms, and consumers (a.k.a., agglomeration benefits)
  - Against those benefits, numerous costs: congestion, pollution, crowded and expensive housing, disease, etc (a.k.a., urban costs)
- Questions for today
  - What do we know about this tradeoff?
  - How does it get resolved in practice?
  - Focus on recent developments
- Main challenges
  - Cities are hard to measure
  - Cities are shaping the fundamental tradeoff as much as being shaped by it
  - Many benefits and costs of cities are externalities, which leave very few or no paper trails

Outline

- Measuring cities: delineation and metrics
- Agglomeration economies: identification and data
- Urban Costs: articulation and estimation
- Equilibrium: inefficiencies due to markets and local politics

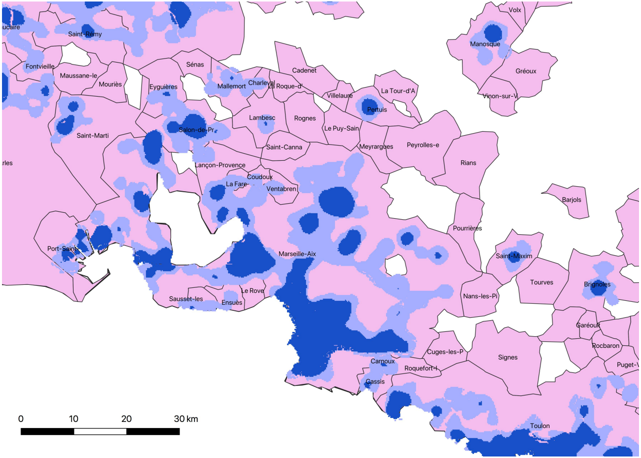
Defining our units

- The study of the fundamental tradeoff of urban economics is a study of how the 'local environment' affects individual outcomes and how that environment is shaped by individual actions
- Some characteristics of the local environment can potentially be measured in continuous space (eg, employment density)
- Others (eg, city population, measures of urban shape) require defining some units: cities
- Until quite recently, the question of using appropriate units was largely ignored
- Data advances have allowed us to define our units or, in some cases, work with continuous space directly
- Delineating cities is currently a vibrant area of research (eg, JUE special issue, 2020)

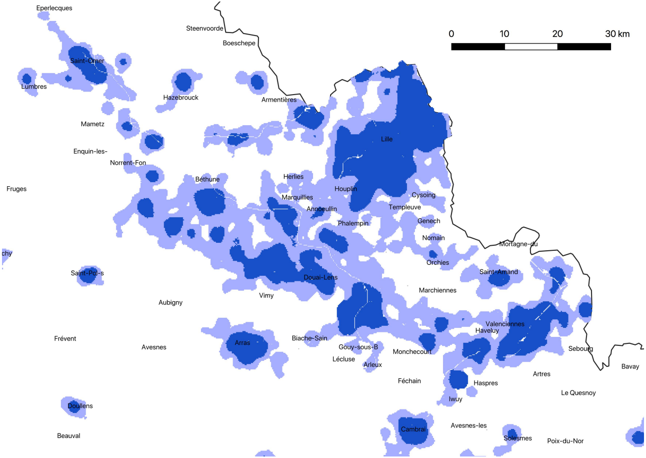
To delineate cities, first choose an approach...

- Two main approaches to urban delineation: functional (commuting flows) or morphological (continuous built-up)
- No approach is obviously better conceptually. It may depend on the question at hand (or worse, on the answer to the question at hand).
- Two problems with all existing approaches:
  - They rely on arbitrary thresholds (de Bellefon et al., 2020)
  - Delineations are sensitive to fine details of the approach (for area, for the population of specific cities, less for overall urban population)

What is urban? The Marseille region, France



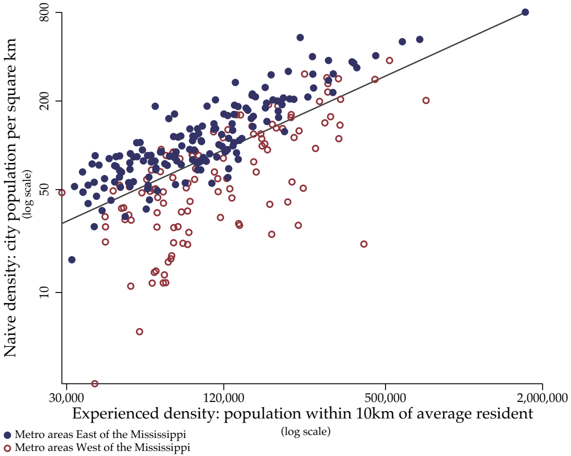
One or several separate cities? The Lille region, France



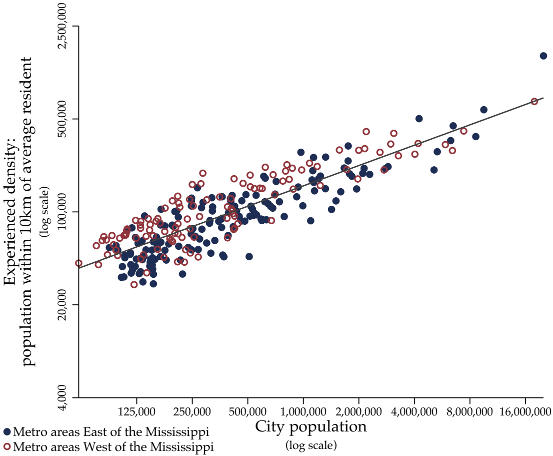
… then compute key metrics

- Population: delineation approaches do a relatively good job at counting population, except when cities are close to each other
- Density:
  - Because city area is not well measured, 'naive density' (=total population/area) should be avoided
  - Use 'experienced density' instead (eg, density within some distance perhaps discounted with distance and averaged across the population of the unit)
  - Experienced density can be parameterised in many ways
  - Unfortunately, I know of no good way to do that optimally

Raw density vs. experienced density for US metro areas



Experienced density vs. population for US metro areas



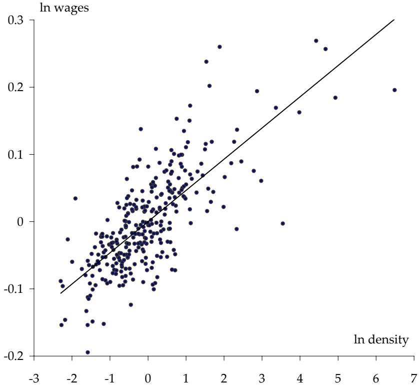
Which metrics to use?

- In the simplest urban models, both density and population, are sufficient statistics to describe a city
- Various forms of heterogeneity imply that this is not true in reality
- The high correlation between experienced density and population is useful in a first step but a hindrance when trying to disentangle different effects
- More city variables may affect the benefits and costs of cities (Harari, 2020)

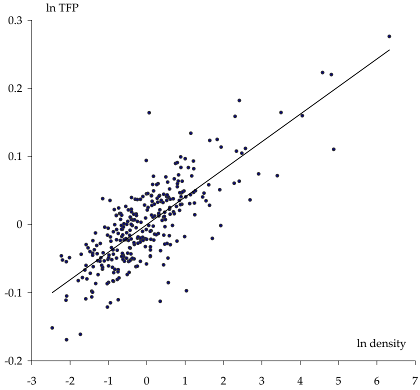
Productivity benefits

- Well-established literature (early contributions by Sveikaukas 1975, Carlino, 1979, Henderson, 1986)
- Many significant reviews and meta-analysis (Rosenthal and Strange, 2004, Puga, 2010, Combes and Gobillon 2015, Melo, Graham, and Noland 2009, Ahlfeldt and Pietrostefani 2019)
- Early studies: average output per worker, rents and other indirect outcomes
- More recent studies: earnings and firm TFP
- No attempt to be comprehensive - focus instead on measurement, identification, and recent contributions

Wage and TFP in French employment areas, Combes et al. (2012)

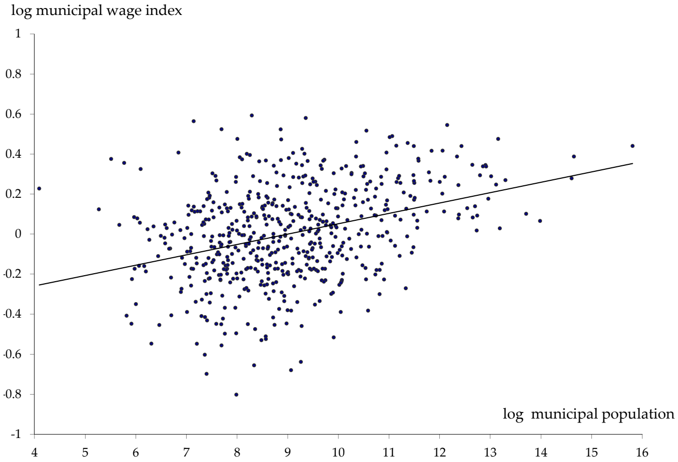


(a) Wages and employment density
(306 employment areas, 1976-1996 average)



(b) TFP (Olley-Pakes) and employment density
(306 employment areas, 1994-2002 average)

Municipal wages and municipal population in Colombia, 2008-2012



Estimating equation

With agglomeration economies on the production side, simple models of agglomeration typically imply:

$$\log w_{ic} = \alpha \log pop_c + X_c \psi + u_i + \varepsilon_{ic} \tag{1}$$

where c is a city and i an individual (and we ignore any subscript t for now)

- Unit of observation: individuals in metropolitan areas
- Wage as dependent variable but can use TFP with firms
- log-log specification?

Estimating equation: complications

- Population is used as a measure of scale but perhaps density matters too. Ideally, we should use both:

$$\log w_{ic} = \alpha \log den_c + \beta \log area_c + X_c \psi + u_i + \varepsilon_{ic} \tag{2}$$

  where α measures the elasticity wrt to density and β measures the elasticity wrt to aggregate population (ie, an increase in area keeping density constant)

- To think about different scales at which agglomeration effects take place, we can enrich the specification further:

$$\log w_{ic} = \alpha \log den_c + \beta \log area_c + \gamma \log ExtMA_c + X_c \psi + u_i + \varepsilon_{ic} \tag{3}$$

  where *ExtMA* measures external market access

- Overall effects of scale or sector specific effects?
  This may call for enriching the regression further...

Productivity benefits: Density and/or population

- Tradeoff between specification error/completeness and complexity of the regression
- Typical (OLS) finding: $\alpha > \beta$ (3-4% vs. 1-2% - Combes et al., 2008)
- But these coefficients may be sensitive to where we draw the line:
  - Briant et al. (2010): agglomeration effects are not sensitive to the shape of units in France and moderately sensitive to their spatial scale
  - Bosker et al. (2020): confirm the scale results for Indonesia
- Attempts to measure agglomeration effects spatially:
  - Arghazi and Henderson (2008): strong effects for advertising in Manhattan
  - Rosenthal and Strange (2003, 2008, 2020)

Identification

Let us now concentrate on identification issues. In the regression:

$$\log w_{ic} = \alpha \log pop_c + X_c \psi + u_i + \varepsilon_{ic} \quad (4)$$

We face the following problems:

- Some city confounding factors X_c are unobserved
- Reverse causation
- Individual effect u_i is unobserved
- All with possible correlations with the measure(s) of agglomeration of interest

To see this more clearly, rewrite the estimating equation as:

$$\log w_{ic(i)} = \alpha \log pop_{c(i)} + X_{c(i)} \psi + u_i + \varepsilon_{ic(i)} \quad (5)$$

Two main sources of bias

- Endogenous quantity of labour: Missing variables in X_c or reverse causation
- Endogenous quality of labour: Sorting on u_i

Identification: Fixed effects

- Naive OLS

$$\log w_{ic(i)} = \alpha \log pop_{c(i)} + \xi_{ic(i)} \tag{6}$$

  where $\xi_{ic(i)} = X_{c(i)}\psi + u_i + \varepsilon_{ic(i)}$

  Fall fouls of the biases described

- With fixed effects (worker and city)

$$\log w_{ic(it)t} = \alpha \log pop_{c(it)t} + \eta_{c(it)} + u_i + \varepsilon_{ic(it)t} \tag{7}$$

  This regression is...
  - ...identified through the movers (selected sample? not randomly moving?)
  - ...identified from changes in population within cities (making endogeneity worse? E.g., a new road increases productivity and attracts new residents)

- Retain fixed effects for workers but not for cities? Combes et al. (2008) find that the agglomeration elasticity is about half the elasticity in the naive OLS

Identification: : Instrumental variables

Individual fixed effects help mitigate the endogenous quality of labour bias but not the endogenous quantity of labour bias. For this, use IV:

- Ciccone and Hall (1996) and followers: long population lags
- Rosenthal and Strange (2008) and Combes et al. (2010): land suitability/soil fertility
- Rationale: past population and determinants of past population explain current population
- Defense of the exclusion restriction:
  - Avoids reverse causation and biases caused by contemporaneous shocks
  - Sources of natural advantage are different today
  - More generally unobserved city productivity is not expected to be serially correlated over long lags

Identification: : Instrumental variables

But:

- This type of instrument does not solve for endogenous quality of labour (fixed effects still needed)
- The exclusion restriction can be questioned:
  - More accumulation of public capital in historically more populated places
  - Other potential sources of persistence: local institutions, culture, etc
- Main results: using population lags and soils typically lowers the population/density elasticity modestly

Identification: Quasi/natural experiments

- Greenstone, Hornbeck, and Moretti (2010)
  - Use contests for 'million dollar plants'
  - Winning counties subsequently did much better than runner-ups
  - Surviving manufacturing plants experienced a much lower TFP decline in winning counties
  - Overall, they find large gains from agglomeration
  - More recent research finds smaller but still substantial effects (Patrick, 2016)
- Other creative ideas
  - Redding and Sturm (2008)
  - Zhang (2019)
- Dilemma: Internal vs. external validity

Identification: Structural/quantitative approaches

- To substitute for data limitations, put more structure on the model
- Baum-Snow and Pavan (2012)
  - Structural labour model with 3 types of locations
  - Very detailed modelling of workers and labour search
  - Estimate static and dynamic agglomeration effects using NLSY data
- Ahlfeldt, Redding, Sturm, and Wolf (2015)
  - Embed Eaton-Kortum (allowing for a lot of heterogeneity) in a model of a city
  - Rely on commuting patterns to estimate a cost of distance and heterogeneity
  - Together with other moments, they can recover all key parameters
  - Perform counterfactuals on important policies (infrastructure)
- Duranton and Puga (2019)
- Some problems:
  - Sensitivity to functional forms
  - Modelling too much or too little? How much complexity?
  - Focus on which moments?
- Looking forward: Improve modelling at the intensive or extensive margin?

One channel approaches

- Marshall (1890): thick local labour markets, input-output linkages, and knowledge spillovers
- Duranton and Puga (2004): sharing, matching, and learning
- This suggests regressions like

$$\log w_{ic(i)} = \alpha \log pop_{c(i)} + \beta Mech_{c(i)} + X_{c(i)}\gamma + u_i + \varepsilon_{ic(i)} \tag{8}$$

and

$$Mech_c = \delta \log pop_c + X_c\zeta + \eta_c \tag{9}$$

where $Mech_c$ is the mechanism/market under scrutiny

One channel approaches

- This type of approach (e.g. Charlot and Duranton 2004) raises hard identification questions:
  - The wage equation suffers from the same issues as previously...
  - ... Plus separating the effects of the specific mechanism being examined and other agglomeration effects from population
  - ... Plus the endogeneity of city in the second equation
- It is also very data demanding

One channel approaches: Recent work

- Dauth, Findeisen, Moretti and Suedekum (2018) on labour market matching
- Buechel and Ehrlich (2016) on proximity and interactions
- Atkin, Chen, and Popov (2020) on random meeting and innovation

Horse races: Sorting or learning?

- ‘Unobserved’ worker ability play an important role: is that natural ability/upbringing or learning?
- De La Roca and Puga (2017):
  - Estimate both static and dynamic (experience) effects of living in big cities
  - Find some wage effects of large city experience in both Spain and the US
  - Dynamic/learning agglomeration effects similar to static effects

Horse races: Agglomeration or market selection or sorting?

- Higher firm productivity in large cities could be due to the elimination of less productive firms (Syverson 2004)
- Combes et al. (2012) propose a model of agglomeration and firm selection
- Key feature: firm selection should left truncate the productivity distribution
- No evidence of this on French data. Instead, the distribution of firm TFP in large cities is shifted to the right and dilated
- In turn, stronger agglomeration effects in larger cities for more productive firms should induce the sorting of firms (Gaubert 2018)

'Partial' approaches

- Jaffe, Trajtenberg and Innovation (1993) on proximity and citations
- Holmes (1999): firms outsource more in larger cities
- Duranton and Jayet (2011): workers are more specialised in larger cities
- Audretsch and Feldman (1996) or Rosenthal and Strange on the determinant of industry concentration
- Ellison, Glaeser, and Kerr (2010) on co-agglomeration
- Costa and Kahn (2000) on the location of 'power couples'

Accessibility and consumption benefits

- Greater density will reduce expected trip distances in cities
- This may encourage more trips and lead to slower trips
- Duranton and Turner (2018): Elasticity of (i) mileage to density: -0.13, (ii) travel speed to density: -0.11, and (iii) the number of trip to density: 0.01 in the US
- Agarwal, Jensen and Monte (2019): Decline in shopping travel distance with density of stores
- Couture (2016): Similar results for restaurants
- Handbury and Weinstein (2015): The price of consumption goods is no higher in bigger cities...
- ... and available variety is much greater
- Couture (2016) confirms this for restaurants and shows that restaurant goers become much more choosy in denser urban settings

Accessibility: new directions

- Linking with accessibility measures
  - Couture (2016) uses a discrete choice framework in the spirit of Ben-Akiva and Lehman (1985)
  - This type of approach delivers a value of accessibility if we know about both the chosen and potential destinations. These can now be measured with new sources of data
  - Other empirical challenge: the origins and destinations of trips are endogenous
- Quantitative modelling
  - Current quantitative models consider the choice of both residential and workplace locations
  - But ignore other errands (80+% of trips and 75% of the mileage)
  - Challenge for the future research: fuller modelling of travel behaviour while retaining the tractability of current approaches
- Link with productivity
  - Manning and Petrongolo (2017): accessibility and labour search
  - Xiao and Wu (2020): commuting and propensity to innovate

Other agglomeration benefits

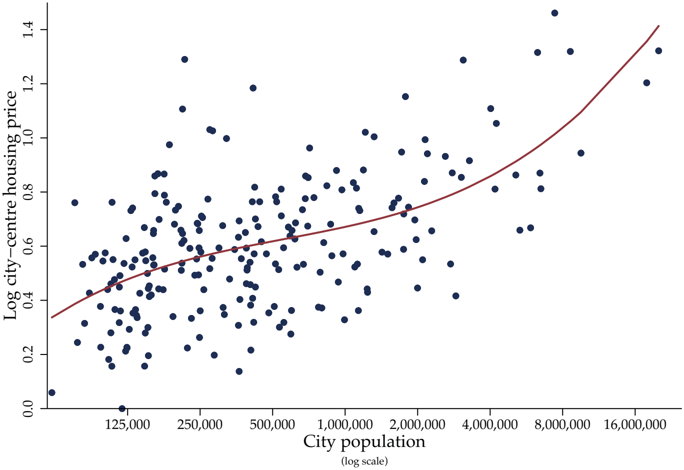
Many other potential benefits of agglomeration have been explored, let me highlight only three:

- Innovation: Carlino, Chatterjee, and Hunt (2007), Moretti (2019)
- Pollution: Glaeser and Kahn (2008), Carozzi and Roth (2019)
- Endogenous Amenities: Diamond (2016), Couture and Handbury (2019), Davis, Dingel, Monras, and Morales (2019), Couture, Handbury, Hurst, and Gaubert (2020)

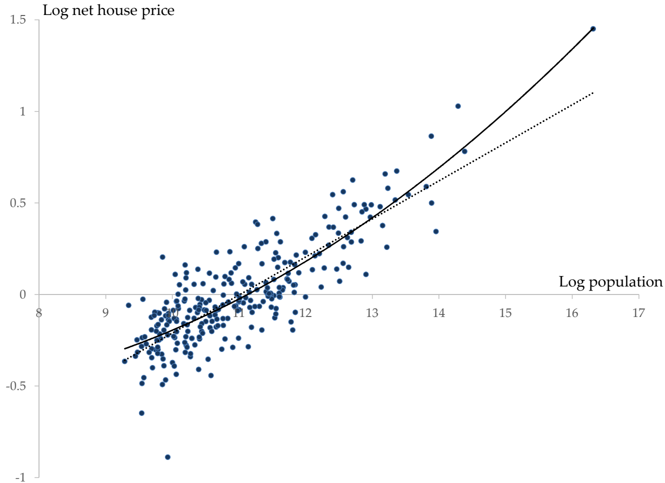
Costs of cities: Land scarcity

- A greater population/density makes land scarce and in turn housing more expensive
- To compare likes with likes, we compare land or house prices at the centre
- Because the center may be 'special' (and observations sparse), better to estimate land or house prices at the centre using data for the entire city
- For France and US: elasticity of central house prices wrt to city population: ≈ 0.1
- For France: elasticity of central house prices wrt to density about $\approx 0.2 - 0.3$
- For France: elasticity of central land prices wrt to density $\approx 0.6 - 0.8$
- Mild evidence that these elasticities may be increasing with size or density

Housing prices at the centre and city population, US MSAs c. 2015

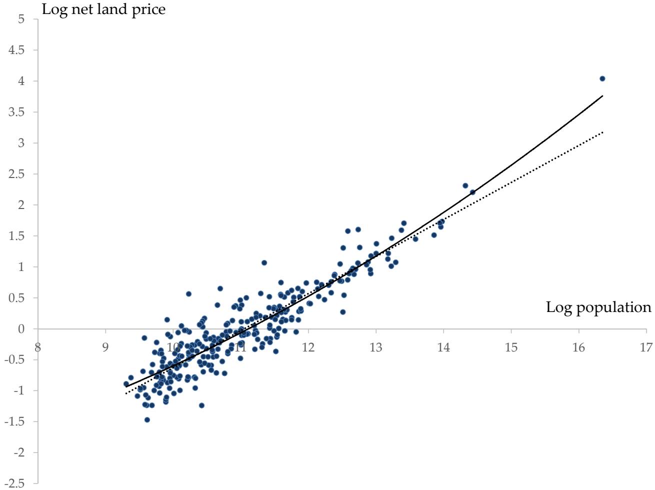


Housing prices at the centre and city 'density', France 2006-2012



Note: Residual plus component plot conditioning for city area. Population measures a density effect

Land prices at the centre and city 'density', France 2006-2012



Note: Residual plus component plot conditioning for city area. Population measures a density effect

Costs of cities: Land scarcity

- But higher land prices are not true economic costs, only transfers
- However, higher land prices lead to (and reflect) real costs of city population and density

Costs of cities: Building higher

A first response to higher land prices: build higher

- Chicago c. 2000: Elasticity of building height wrt land prices: 0.30 for residential buildings and 0.45 for commercial buildings (Ahlfeldt and McMillen, 2018)
- Two interesting wrinkles:
  - This elasticity is higher today
  - Elasticity of floorspace is only about 0.1
- Elasticity of unit building costs wrt to building size varies with building size:
  - 0.25 for small buildings
  - Well above 1 for skyscrapers
- There are other costs associated with tall buildings (Liu, Rosenthal, and Strange, 2018)
- Building more lessens the pressure on the price of floorspace but does not eliminate it
- As a result, consumption of floorspace per capita declines (and density increases)

Costs of cities: Building further away

A Second response to higher central land prices: go further away

- For US MSAs, elasticity of the distance between home and the center for the median resident wrt to city population: 0.3
- Elasticity of distance traveled wrt to distance to the centre: 0.07 in the US (Duranton and Puga, 2019)
- This elasticity is equal in magnitude but opposite in sign to the elasticity of house prices

  This is consistent with the Alonso-Muth condition of monocentric models

Costs of cities: Congestion

Greater density and urban expansion following population growth are expected to put a strain on the transportation infrastructure

- Elasticity of travel speed wrt to population in the US: 0.04-0.05 (Couture, Duranton, and Turner, 2018, Duranton and Puga, 2019)
- Same elasticity for India (Akbar, Couture, Duranton, and Storeygard, 2019)
- Akbar et al. (2019) also estimate a much higher elasticity wrt population density (≈ 0.20)

Costs of cities: Worse amenities?

- Crime
- Pollution
- Disease

Estimating costs: the aggregation problem

Challenges:

- Many costs are hard to assess, eg how should we value time in travel?
- How do we aggregate to avoid double counting?

Estimating costs: Using the curvature of the agglomeration curve

Au and Henderson (2006)

- Urban costs dampen agglomeration effects
- Main result: Chinese cities are grossly undersized
- Great advantage: very little data required
- Drawbacks:
  - This approach tries to estimate a bell shape directly. Perhaps OK in China where cities are far from the optimum. Hard otherwise when we expect a large flat region around the optimum
  - How much of urban costs is reflected in value added per worker? Congestion can potentially raise GDP per worker while reducing welfare

Estimating costs: Using consumer theory

Combes, Duranton, and Gobillon (2019):

- Consumer theory approach (characterize the expenditure function as city population increase)
- Use theory to simplify the estimation
- Eventually: urban cost elasticity = central house price elasticity X share of income in housing
- Main result: urban cost elasticity ranging from 0.03 in small cities to 0.08 in cities the size of Paris
- Main weaknesses: heavy reliance on free mobility and no endogenous amenities

Estimating costs: Using parameters of a specific model

Duranton and Puga (2019):

- Fully specify a model with parameters governing travel distance, the income elasticity of travel costs, agglomeration, etc
- The approach relies on 'standard' parameters and ...
- ... uses equilibrium relationships as overidentifying restrictions
- Main result: urban cost elasticity of 0.07 amplified by congestion elasticity of 0.04 and travel time is valued with an income elasticity just slightly below one
- Main drawback: highly data demanding and model dependent

Estimating costs: Some preliminary conclusions

- Support for a hill-shaped relationship between the net benefits of cities and their population / density
- The top of the hill is fairly flat: small costs from being moderately under- or oversized
- While the costs of population growth are fairly low when the city can expand, the costs of density growth are much higher
- Some evidence of convexity in these costs, they may increase steeply beyond a certain size
- A lot remains to be done, from a better handling of endogenous amenities to looking more at developing countries

Wedges: at the margin vs. on average

- Wedge 1: Workers choose their location on the basis of average agglomeration spillovers to be received, not marginal
- Wedge 2: Unless the revenue of land is entirely redistributed locally, some benefits from from agglomeration are transferred away
- Wedge 3: Residents choose their location based on average travel costs not marginal congestion
- As these wedges push in opposite directions, the overall effect is unknown
- The wedges may affect city population and density differently
- There is also a coordination failure in urban development
- The market is unlikely to provide an optimum but it is unclear in which direction the bias goes

Pervasive regulations

- Land use regulations are pervasive
- Take many forms: minimum lot size, FAR, setbacks, irregular shaped lots, protections for wetlands, community vetoes, long approval delays, etc
- Often extreme: 94% of residential land in San Jose is zoned single-family homes only
- Become more stringent over time (Gyourko, Hartley, and Krimmel, 2019)
- Locally decided in the US, presumably in favor of incumbent residents

Why regulations?

- Avoid free-riding of local public goods by the poor (Tiebout 1956, Fischel 1987)
- Risk aversion (Fischel 2001)
- Localized costs vs. more diffuse benefits of new developments
- Exploitation of monopoly power of unique locations (Ortalo-Magné and Prat 2014; Hilber and Robert-Nicoud 2013)
- Misalignment between local and global welfare maximization (Duranton and Puga 2019)

Consequences

- Equity: Major intergenerational redistribution following the rise of land use regulations
- Efficiency: Costly misallocation
  - 'Traditional view': Harberger's deadweight loss triangle is at most 2% of welfare (Glaeser and Gyourko, 2018)
  - 'New view': misallocating workers across cities with different productivity levels is much more costly (Hsieh and Moretti, 2019, Duranton and Puga, 2019, Rossi-Hansberg, Sarte, and Schwartzman, 2020)

Conclusions

- Some areas of progress:
  - New data tracking locations, mobility, interactions, land use, etc
  - A new generation of models able to handle heterogeneous location choices
  - New models to explore a range of issues: sorting, urban costs
  - Less novelty on causal identification in the last 10 years
- My wish list going forward:
  - Even more data
  - Better modelling of mobility and location choices
  - Higher standards for the identification of causal effects
  - Bring data, models, and empirical care closer together

Covid conclusions

- A big overhaul of costs and benefits of cities in the short-run
- How much of that is going to last?
  - Will social distancing leave a permanent footprint like tuberculosis or the cholera did before?
    More outdoor public space for humans? Death of transit?
  - Permanently fewer face-to-face interactions? (But more valuable ones?)
  - Stronger sorting patterns?