Quantitative Spatial Economics I

Zibin Huang [1]

[1]College of Business, Shanghai University of Finance and Economics

March 15, 2026

Overview

Introduction

I will introduce basic methods for quantitative spatial models (QSM) in this course

- Static QSM: model construction, estimation, and solution (Week 1-4)
- An example of static QSM: Fang et al (2026) (Week 5)
- Model with Goods Trade (Week 6)
- Diamond style models in urban economics (Week 7-8)
- Dynamic QSM: model construction, estimation, and solution (Week 9-11)

Introduction

Theoretical model vs Structural model vs Design-based regression

- Theoretical model is based on the subjective thoughts of the researcher
- It is purely a deduction
- It is not related to data directly
- But usually it results in testable predictions $\Rightarrow$ Falsifiability (Karl Popper)
- Example: Traditional Keynesian Model, AMM Model, Principal-Agent Model...

Introduction

- Design-based regression is the main empirical approach
- Formalized by Angrist, Rubin, Card, and Imbens during 1990s
- It is a pure empirical method: let data speak
- It can be used as tests of theories
- Example: RCT, DID, RDD...
- Sometimes people call it "Reduced-form" analysis, although not so accurate

Introduction

- RF is a useful framework in tackling causal effects
- But the effect is a black-box
- Can hardly answer two questions
  - What and how important are the mechanisms (channel analysis)
  - What will happen if we impose a complicated new policy, or old policy in a new context (external validity)
- What is the effect of relaxing migrant children's enrollment restriction on the overall human capital in China?
- What is the effect of building a land quota trading system among cities on local and national outputs?

Introduction: Quantitative Spatial Equilibrium Model

- Structural model is a model directly and closely connected with data
- It is between pure theoretical model and design-based research
- We build a model, then connect to data by estimating model parameters and uncovering model unknowns
- Then we can simulate the counterfactual world in different proposed policies

Introduction: Quantitative Spatial Equilibrium Model

Quantitative Spatial Equilibrium Model is a powerful tool

- Developed from Eaton and Kortum (2002) model in trade (E-K Model)
- There are many locations, many workers, many goods
- Workers choose locations to live and work, s.t. wages, prices, and migration costs
- Goods' flows are determined by productions in each location and trade costs across regions
- There can be some other parts: amenity, land market, housing market etc. (Redding and Rossi-Hansberg, 2017)
- A spatial equilibrium is achieved when labor/goods supply=labor/goods demand in each location

Introduction: Quantitative Spatial Equilibrium Model

- A key component of spatial model is
  The cost of moving goods, people and ideas across regions
- That is why region/geography matters
- If people/goods can move totally freely across regions, then spatial structure does not matter at all
- If I have an Anywhere Door from Doraemon, why do I care where I live and where I work?

Introduction: Quantitative Spatial Equilibrium Model

- A dilemma: the complexity of the spatial model (structural model)
- If it is too complicated, it is not tractable
- Sometimes impossible to solve a complicated model with equilibria in hundreds of markets simultaneously
- If it is too simplified, it cannot incorporate data well
- Traditional spatial/urban models are only theoretical, e.g. AMM Model (monocentric city), Rosen-Roback Model

Introduction: Quantitative Spatial Equilibrium Model

- QSEM uses some specific distributional/structural assumptions and results in tractable solutions: Gravity Equations
- Gravity equations describe the spatial movement of goods and people
- This is an extension of simple discrete choice models
- Thus, QSEM can be taken to data!
- We can use it to simulate different policy counterfactuals

Introduction: Quantitative Spatial Equilibrium Model

- Highly recommend the following:
  - Redding and Rossi-Hansberg(2017) Quantitative Spatial Economics, *Annual Rev. Econ.*
  - 2020 UEA Lecture Series
    `https://urbaneconomics.org/workshops/lectures2020/`

Introduction: Quantitative Spatial Equilibrium Model

- In the first four weeks, we will introduce Fang et al (2024) to thoroughly investigate the implementation of a QSEM
- This is the baseline model with essential components of QSEM
- The steps of a study with QSEM:
  - 1. Build the model
  - 2. Estimate/calibrate model parameters given data
  - 3. Solve the model equilibrium and check the fitness
  - 4. Implement counterfactuals using the model
- Then in the fifth week, we will illustrate another example of QSEM with an extension in marriage market

Motivation

Place-based Land Policy and Spatial Misallocation: Theory and Evidence from China
Min Fang, Libin Han, Zibin Huang, Ming Lu, and Li Zhang

Motivation

Place-based policies are extensively used:

- Why: to promote balanced development across regions (Neumark and Simpson, 2015)
- How: land supply quotas, wage subsidies, tax subsidies, industrial zones, ...

Motivation

How effective are place-based policies at achieving their targets?

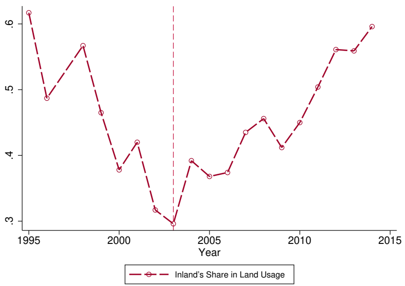
- What frictions are the policies alleviating (or amplifying)?
- Do they cause efficiency loss through spatial misallocation?
- Are the targets necessarily "place-based"? (versus "people-based")

In this paper, we aim to provide answers by studying a national large-scale place-based land policy implemented by the Chinese government around 2003.

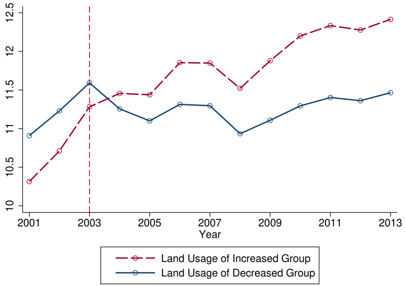
The Inland-favoring Land Supply Policy around 2003

- **Goal:** to promote convergence of development across regions (Non-East versus East)
- **Tool:** Land Quota System (controls land supply quota in each region)
- **Before 2003:** Mostly based on demand $\Rightarrow$ **After 2003:** More in Non-East regions

The Inland-favoring Land Supply Policy around 2003



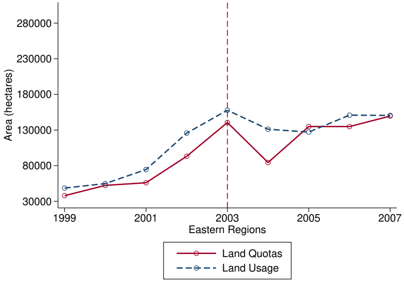
(a) Inland Provinces' Share



(b) Divergence between Prefecture Groups

Figure: New Urban Land Usage before and after 2003

The Inland-favoring Land Supply Policy around 2003



(a) Eastern Regions

(b) Inland Regions

Figure: Land Supply and Quota

What we do

We evaluate such a policy both causally and quantitatively:

- Using the change of policy in 2003 to find the causal effect on TFP
- Constructing a spatial equilibrium model to evaluate the mechanism and the implications

What are our findings?

- Empirically, it decreased TFP of Eastern cities relative to Inland cities
- It amplified floor space constraints in developed (East) regions, created spatial misallocation, and lowered national TFP, output, and welfare
- How about the targets?
  - It shrank the east-inland geographical output gap. $\rightarrow$ "place-based"
  - But actually decreases incomes of workers from poorer areas. $\rightarrow$ "people-based"
- Instead, regional transfer is both more equal and efficient.

Illustrating the mechanism

Consider two regions (East & West):

- East has high productivity and wages (Opposite for West)
- Massive workers migrate from West to East
- Land as a factor input is much more constrained in East
- Regional divergence (geographically) in total GDP and GDP per capita are observed

Now consider a West-favoring land policy to promote regional convergence

- Regardless of the migration inflow and constrained land supply in East
- Distribute much more new land quotas to West than East

Illustrating the mechanism

What happen then?

- More productive East is even more land-constrained:

  $$\text{Land Prices} \uparrow \Rightarrow \begin{cases} \text{Residential floor space cost } \uparrow \Rightarrow \text{Labor supply} \downarrow \\ \text{Production floor space cost } \uparrow \Rightarrow \text{Labor demand} \downarrow \end{cases} \Rightarrow$$

  Migrant to East $\downarrow$

  $$\Rightarrow \begin{cases} \text{Workers are locked in the West with lower income} \\ \text{Spatial misallocation in land and labor } \uparrow \text{ Agglomeration effects } \downarrow \\ \text{National TFP, output, and welfare } \downarrow \end{cases}$$

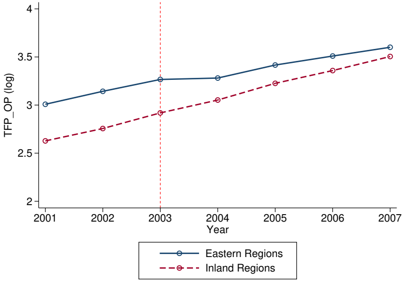
Empirical Analysis: DID Strategy

For prefecture j in year t, we have the following regression:

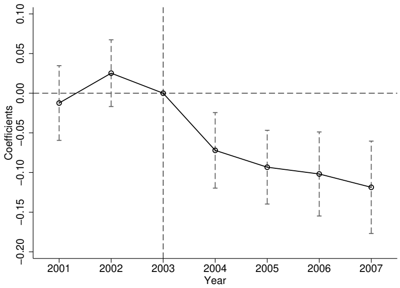
$$ln(Prod_{jt}) = \alpha + \delta_1 Post2003_t \times East_j + \phi_j + \gamma_t + \epsilon_{jt} \quad (1)$$

- $ln(Prod_{jt})$: the prefecture-level average firm TFP;
- $East_j$: eastern region dummy;
- $Post2003$: policy time dummy;
- ϕ_j: prefecture FE; γ_t: year FE.

Empirical Analysis: Parallel Trend



(a) Time Trends of Productivity



(b) Event Study - Productivity

Figure: Parallel Trend Test

Empirical Analysis: Regression Results on Firm TFP

Table: **DID Results on Productivity**

	(1)	(2)
Post2003×East	-0.0705***	-0.0749***
	(0.0267)	(0.0266)
Province × Time Trend	N	Y
GDP Per Capita × Time Trend	Y	Y
Industry Share × Time Trend	Y	Y
Year FE	Y	Y
Prefecture FE	Y	Y
Observations	1,792	1,792
R-squared	0.7529	0.7537

Empirical Analysis: Regression Results on Other Outcomes

Table: **Summary of Other Variables on Mechanism**

	Land Price	Housing Price	Average Wage	Migration Inflow
Post2003×East	$\Uparrow$	$\Uparrow$	$\Downarrow$	$\Downarrow$

Empirical Analysis: Takeaways

The Main Takeaway

- Inland-favoring land policy decreased relative productivity in the developed eastern region
- This is a causal evidence of distortion/misallocation
- We also find increased land and housing prices, and decreased wage and migration in the eastern region
- This offers preliminary empirical evidence for our model mechanism

Empirical Analysis: Takeaways

The Main Takeaway

- Two crucial questions remain
  - What is the national overall effect on economic efficiency and equality?
  - How important are these channels in contributing to the effect?
  - Can we find a better policy to balance efficiency and equality?
- These are questions cannot be answered by empirical regressions
- That is why we need structural model

Model: Overview

Spatial allocation of workers:

- K cities with two sectors: Urban v.s. Rural
- Workers choose city-sector s.t. wages, migration & housing costs
- Urban production combining H/L-skill workers & production floor space
- Agglomeration in urban productivity due to population density

Endogenous floor space market s.t. land supply constraints

- Floor space construction using fixed land supply (policy determined)
- Residential vs. Production floor space
- Endogenous floor space price due to production & residential demand
- Local residents gain all the returns from residential floor space market

▸ Model: Workers' Preferences ▸ Model: Migration ▸ Model: Production ▸ Model: Floor Space Market ▸ Model: Equilibrium

Model I: Worker Preferences

- Worker's Utility:

$$U^o_{in,jk} = \frac{z^o_{in,jk}}{\tau^s_{in,jk}} \left(\frac{c^o_{in,jk}}{\beta}\right)^{\beta} \left(\frac{s^o_{in,jk}}{1-\beta}\right)^{1-\beta} \tag{2}$$

- Shock ($z^o_{in,jk}$) follows Fréchet Distribution: $F(z^o_{in,jk}) = e^{-{z^o_{in,jk}}^{-\epsilon}}, \ \epsilon > 1$
- ϵ is called migration elasticity
- Income: (wage + hometown housing rent)

$$v^s_{in,jk} = w^s_{jk} + \frac{Q_{in} S^R_{in}}{H^R_{in}} \tag{3}$$

i,j: location (home, working); n,k: sector (rural, urban) c: goods consumption;
s: individual housing consumption; S^R: location total housing consumption;
H^R: hukou population τ: migration cost;
z: location preference shock; Q: housing rent; v: income

Model I: Worker Preferences

- We assume that migration cost can be decomposed into two parts:

$$\tau^s_{in,jk} = \bar{\tau}^s_{in} d_{in,jk} \tag{4}$$

- $d_{in,jk}$ captures cost specific for migrating from *in* to *jk*
- Physical distance and institutional costs due to the Hukou system
- $\bar{\tau}^s_{in}$ captures cost differences between individuals with different skills
- High skill people can get more public resources in non-Hukou cities

Model I: Worker Preferences

- We assume this timeline for workers:
  - 1. Observe location taste shock z
  - 2. Decide working location and sector j, k
  - 3. Decide consumption c, s
- Let's go to the bottom layer first: what is the optimal consumption choice given location and sector choices?

Model I: Worker Preferences

- Optimal consumptions from FOCs (given location choices):

$$c^o_{in,jk} = \beta v^s_{in,jk} \tag{5}$$

$$s^o_{in,jk} = (1-\beta)\frac{v^s_{in,jk}}{Q_{jk}} \tag{6}$$

- Property of CD utility function
- Workers spend β share of income on final goods and $1-\beta$ share on housing
- The property persists when you have more than two goods
- Plug (5) and (6) back to (2), we have indirect Utility:

$$U^o_{in,jk} = \frac{z^o_{in,jk} v^s_{in,jk} Q^{\beta-1}_{jk}}{\tau^s_{in,jk}} \tag{7}$$

Model I: Worker Preferences

- This kind of utility is called homothetic preference

$$f(ax_1, ax_2, ..., ax_n) = af(x_1, x_2, ..., x_n)$$

- Common utility functions such as Cobb-Douglas, CES are both homothetic
- When income is scaled up by λ:
  - the optimal consumption bundle is also scaled up by λ
  - Expenditure share on each good is constant
- We will discuss this in more details when introducing the CES demand system
- There are also non-homothetic preference
- They are used to investigate structural transformation (expansion of modern sector relative to agriculture)

Back to Main

Model I: Migration Flows

- The outer layer is the location choices
- The individual choice is uncertain due to taste shock z
- But we can get certain migration probabilities by aggregating over population
- This is what we call "Probabilistic migration"
- Key: the assumption of Fréchet distribution gives us a closed-form migration flow
- Analogously, we have probabilistic trade (Eaton and Kortum, 2002)
- The idea goes back to Daniel McFadden in 1970s on discrete choice models

Model I: Migration Flows

- Using the indirect utility equation, we can write the distribution of utility for a worker migrating from *in* to *jk* as:

$$G^s_{in,jk}(u) = Pr[U \leq u] = F\left(\frac{u\tau^s_{in,jk} Q^{1-\beta}_{jk}}{v^s_{in,jk}}\right) \tag{8}$$

- F is the cdf of a Fréchet distribution
- Using the Fréchet distribution cdf $F(x) = e^{-x^{-\epsilon}}$, we have:

$$G^s_{in,jk}(u) = e^{-\Phi^s_{in,jk} u^{-\epsilon}}, \ \Phi^s_{in,jk} = (\tau^s_{in,jk} Q^{1-\beta}_{jk})^{-\epsilon} (v^s_{in,jk})^{\epsilon} \tag{9}$$

$$g^s_{in,jk}(u) = \frac{dG^s_{in,jk}(u)}{du} = e^{-\Phi^s_{in,jk} u^{-\epsilon}} \cdot \Phi^s_{in,jk} \epsilon u^{-\epsilon-1} \tag{10}$$

Model I: Migration Flows

- Now we calculate the PDF of utility for individual coming from *in* with optimal destination choice
- For individuals from *in*, we can write the following equation:

$$G_{in}^{s}(u) = \prod_{jk=11}^{JK} e^{-\Phi_{in,jk}^{s} u^{-\epsilon}} \tag{11}$$

- The left-hand side is the PDF, defined as the probability that a worker from *in* has a utility smaller than u
- It means the worker from *in* has a utility less than u for all possible destinations *jk*, Which is just the right-hand side

Model I: Migration Flows

- Thus, we can simplify the notation as:

$$(11) \Rightarrow G^s_{in}(u) = e^{-\Phi^s_{in} u^{-\epsilon}}, \ \Phi^s_{in} = \sum_{jk=11}^{JK} \Phi^s_{in,jk} \tag{12}$$

- Watch out the difference!
- Equation (9) shows the CDF of utility for a specific location choice jk in a feasible choice set $\{j_1 k_1, j_1 k_2, j_2 k_1 ......\}$
- Equation (12) shows the CDF of utility for the optimal location choice in the set

Model I: Migration Flows

- Equation (12) shows a very important property of Fréchet distribution
- The maximum of a sequence of Fréchet distributed r.v. is itself Fréchet distributed
- This kind of distribution is called "extreme value distribution"
- Fréchet is "Type II Extreme Value Distribution"

Model I: Migration Flows

- Another common distribution is "Type I Extreme Value Distribution", which gives us Logit model
- Just as T1EV does for Logit, T2EV here gives us a closed-form migration flow
- T1EV (Logit model) is the log version of T2EV (EK style model)
- Logit model usually have a linear utility/production
- EK style model usually have a log-linear utility/production
- They are actually isomorphic

Model I: Migration Flows

- Now let's derive the core gravity equation for migration flows
- For people from *in* with skill *s*
- Proportion to migrate from *in* to *jk* is:

$$\pi^s_{in,jk} = P[u^s_{in,jk} \geq max\{u^s_{in,j'k'}, \forall j'k'\}] \tag{13}$$

- We then use Bayes's rule to decompose this probability

Model I: Migration Flows

- Proportion to migrate from *in* to *jk* is:

$$\pi^s_{in,jk} = P[u^s_{in,jk} \geq max\{u^s_{in,j'k'}, \forall j'k'\}]$$
$$= \int_u P[u^s_{in,jk} \geq max\{u^s_{in,j'k'}, \forall j'k'\}|u^s_{in,jk}] \cdot g[u^s_{in,jk}]du \quad (14)$$
$$= \int_0^\infty \prod_{j'k' \neq jk} G^s_{in,j'k'}(u) g^s_{in,jk}(u) du \quad (15)$$

- $\prod_{j'k' \neq jk} G^s_{in,j'k'}(u)$ means given a fixed utility value u of choosing jk, what is the probability to have utility of all other choices ($j'k'$) smaller than jk choice
- We then integrate over the domain of u with PDF $g^s_{in,jk}(u)$, the prob density of choice jk's utility to be u

Model I: Migration Flows

- Plug in (9) and (10), equation (15) can be written as:

$$
\begin{aligned}
\pi^s_{in,jk} &= \int_0^\infty \prod_{j'k' \neq jk} e^{-\Phi^s_{in,j'k'} u^{-\epsilon}} \cdot e^{-\Phi^s_{in,jk} u^{-\epsilon}} \cdot \Phi^s_{in,jk} \epsilon u^{-\epsilon-1} du \\
&= \int_0^\infty e^{-\Phi^s_{in} u^{-\epsilon}} \cdot \Phi^s_{in,jk} \epsilon u^{-\epsilon-1} du \qquad (16)
\end{aligned}
$$

- Notice that we have $\frac{d}{du}[\frac{1}{\Phi^s_{in}} e^{-\Phi^s_{in} u^{-\epsilon}}] = \epsilon u^{-\epsilon-1} e^{-\Phi^s_{in} u^{-\epsilon}}$
- We can transfer integral (16) to have:

$$
\pi^s_{in,jk} = \int_0^\infty \Phi^s_{in,jk} d[\frac{1}{\Phi^s_{in}} e^{-\Phi^s_{in} u^{-\epsilon}}] = \frac{\Phi^s_{in,jk}}{\Phi^s_{in}} \qquad (17)
$$

- $\Phi^s_{in,jk}$ is not a function of u, which can be taken out of the integral

Model I: Migration Flows

- Then we have the Gravity Equation of Migration Flow:

$$\pi^s_{in,jk} = \frac{(\tau^s_{in,jk} Q_{jk}^{1-\beta})^{-\epsilon}(v^s_{in,jk})^{\epsilon}}{\sum_{j'k'=11}^{JK}(\tau^s_{in,j'k'} Q_{j'k'}^{1-\beta})^{-\epsilon}(v^s_{in,j'k'})^{\epsilon}} \tag{18}$$

- Proportion of people with skill s from in to migrate to jk
- Positively affected by destination income v
- Negatively affected by destination housing price Q and migration cost τ
- The denominator is a normalization
- This is the key to connect model with data

▸ Back to Main

Model I: Logit versus EK

- In a choice problem with individual n choosing alternatives j
- We have two ways of setting up the model
- Logit Model
  - Linear utility with T1EV error: $U_{nj} = V_{nj} + z_{nj}$
  - For T1EV error, we have: $F(z_{nj}) = e^{-e^{-z_{nj}}}$
  - Closed-form choice probability function: $P_{ni} = \frac{e^{V_{ni}}}{\sum_j e^{V_{nj}}}$
- EK Model
  - Log-linear utility with T2EV error: $U_{nj} = V_{nj} \cdot z_{nj}$
  - For T2EV error, we have: $F(z_{nj}) = e^{-z_{nj}}$
  - Closed-form choice probability function: $P_{ni} = \frac{V_{ni}}{\sum_j V_{nj}}$

Back to Main

Model I: Logit versus EK

- When to use Logit and when to use EK?
- They are isomorphic $\Rightarrow$ no difference in model
- Really depends on the convenience of calculation
- It is also a historical tradition that labor economists use Logit more but trade/urban economists use EK more
- Just an observation, I don't know why

Back to Main

Model II: Production

- In rural region r of prefecture j, we have rural production:

$$Y_{jr} = A_{jr} H_{jr} \tag{19}$$

- Depends only on fundamental productivity A
- And total labor $H_{jr} = H_{jr}^{h} + H_{jr}^{l}$

Model II: Production

- In urban region u of prefecture j, we have urban production:

$$Y_{ju} = (X_{ju})^{\alpha}(S_{ju}^{M})^{1-\alpha}, \text{ where } X_{ju} = [(A_{ju}^{h}H_{ju}^{h})^{\frac{\sigma-1}{\sigma}} + (A_{ju}^{l}H_{ju}^{l})^{\frac{\sigma-1}{\sigma}}]^{\frac{\sigma}{\sigma-1}} \tag{20}$$

- X_{ju} is a CES aggregated input of high skill labor H^{h} and low skill labor H^{l}
- S_{ju}^{M} is the production floor space
- A_{ju}^{h}, A_{ju}^{l} are skill-specific productivity
- H_{ju}^{h}, H_{ju}^{l} are high/low skill labors

Model II: Production

- First Order Conditions:

$$w^l_{ju} = \alpha X_{ju}^{\alpha-1} S_{ju}^{M\,1-\alpha} A^l_{ju}{}^{\frac{\sigma-1}{\sigma}} X_{ju}^{\frac{1}{\sigma}} H^l_{ju}{}^{-\frac{1}{\sigma}} \tag{21}$$

$$w^h_{ju} = \alpha X_{ju}^{\alpha-1} S_{ju}^{M\,1-\alpha} A^h_{ju}{}^{\frac{\sigma-1}{\sigma}} X_{ju}^{\frac{1}{\sigma}} H^h_{ju}{}^{-\frac{1}{\sigma}} \tag{22}$$

$$S^M_{ju} = \left(\frac{1-\alpha}{q_{ju}}\right)^{\frac{1}{\alpha}} X_{ju} \tag{23}$$

FOC gives us a measure of skill premium ω of city j:

$$\omega_{ju} = \frac{w^h_{ju}}{w^l_{ju}} = \left(\frac{A^h_{ju}}{A^l_{ju}}\right)^{\frac{\sigma-1}{\sigma}} \left(\frac{H^h_{ju}}{H^l_{ju}}\right)^{-\frac{1}{\sigma}} \tag{24}$$

▸ Back to Main

Model II: Production

- Assume that final good can be traded without cost and the market is perfectly competitive
- Price of final goods is normalized to 1: numeraire
- We have zero profit condition:

$$(X_{ju})^{\alpha}(S_{ju}^{M})^{1-\alpha} - W_{ju}X_{ju} - q_{ju}S_{ju}^{M} = 0 \tag{25}$$

  where $W_{ju}X_{ju} = w_{ju}^{l}H_{ju}^{l} + w_{ju}^{h}H_{ju}^{h}$
- w_{ju}^{h}, w_{ju}^{l} are wages, q_{ju} is production floor space price
- FOC (23) + Zero profit (25) gives us production floor space price:

$$q_{ju} = (1-\alpha)\left(\frac{\alpha}{W_{ju}}\right)^{\frac{\alpha}{1-\alpha}} \tag{26}$$

Model II: Production

- This is a very simple and standard CES production with representative firms
- We can normalize the goods price thanks to some simplifications
- Specifically, we have the following:
  - Single consumption goods (no sector difference in price)
  - Zero trade cost (no location difference in price)
- However, spatial difference in price level is very important
- We will show you how to deal with this issue in a very similar way as migration in the next lectures
- Local production $+$ Trade cost $\Rightarrow$ Prices differences by location

Model II: Production

- We assume to have economy of scale in urban region
- Agglomeration is very important in spatial economics
- For productivity of workers with skill s in urban region of prefecture j:

$$A_{ju}^{s} = a_{ju}^{s} \times (D_{ju})^{\gamma}, \quad D_{ju} = \frac{H_{ju}^{h} + H_{ju}^{l}}{\bar{L}_{j}} \tag{27}$$

- a_{ju}^{s} is the fundamental productivity
- D_{ju} is the urban population density, $\bar{L}_j$ is the total constructed land
- We will discuss the agglomeration issue in more details

Model III: Floor Space Market Clearing

- Housing market in rural area is simple
- In general, there is no commodity housing market in rural China
- Houses can only be leased informally from "zhai ji di"
- For simplicity, we assume rural floor space price is a proportion of τ of urban floor space price/rental rate in the same prefecture:

$$Q_{jr} = \tau Q_{ju}$$

Model III: Floor Space Market Clearing

- Floor space market in urban area is much more complicated
- We have two kinds of floor space: production vs residential
- There is a production & residential price difference

$$q_{ju} = \eta_j Q_{ju} \tag{28}$$

- η_j is city-level tax equivalent of land use regulations
- It is an exogenous parameter determined by the government
- This captures the fact that Chinese local governments are more willing to allocate land for manufacturing sector but not housing (Lu et al., 2024)

Model III: Floor Space Market Clearing

- Assume a simple linear technology to convert land L to floor space S for real estate firms:

$$S_{ju} = \phi_j L_j \tag{29}$$

- Urban land supply is monoplistic in China (only by government)
- Thus, L_j is exogenous and the main policy parameter
- Urban floor space market clearing: supply = demand

$$\text{Production: } S_{ju}^M = \theta_j S_{ju} = \left(\frac{(1-\alpha)}{q_{ju}}\right)^{\frac{1}{\alpha}} X_{ju} \tag{30}$$

$$\text{Residential: } S_{ju}^R = (1-\theta_j) S_{ju} = E[s_{ju}] H_{ju} = (1-\beta)\frac{E[v_{ju}]H_j}{Q_{ju}} \tag{31}$$

- θ is the share of floor space allocated to production

Model: Spatial Equilibrium

A **Spatial General Equilibrium** *for this economy is defined by a set of exogenous economic conditions* $\{\tau^s_{in,jk}, a^s_j, \eta_j, \phi_j, L_j, H^s_{in}\}$*, a list of endogenous prices* $\{Q_{ju}, q_{ju}, w^s_{jk}\}$*, quantities* $\{v^s_{in,jk}, Y_{jk}, H^s_{jk}, S_{ju}\}$*, and proportions* $\{\pi^s_{in,jk}, \theta_j\}$ *that solve firms' problem, workers' problem, floor space producers' problem, and market clearing such that:*

*(i).***[Worker Optimization]** *Taking the exogenous economic conditions* $\{\tau^s_{in,jk}, A^s_{jk}\}$ *and the aggregate prices* $\{Q_{ju}, w^s_{jk}\}$ *as given, workers' optimal choices of migration pin down the equilibrium labor supply in each city* H^s_{jk} *and the migration flow between each city pairs* $\pi^s_{in,jk}$.

*(ii).***[Firm Optimization]** *Taking the exogenous economic conditions* $\{A^s_{jk}\}$ *and the aggregate prices* $\{q_{ju}, w^s_{jk}\}$ *as given, firms' optimal choices of production pin down the equilibrium labor demand* H^s_j*, equilibrium production floor space demand* $\theta_j S_{ju}$ *in each city.*

*(iv).***[Market Clearing]** *For all prefectures, labor supply equals labor demand, floor space supply equals floor space demand, and final good supply equals final goods demand. This pins down the equilibrium aggregate prices* $\{Q_{ju}, q_{ju}, w^s_{jk}\}$*, equilibrium floor space* S_{ju}*, and equilibrium output* Y_{ju}.

Model: Conclusion

Now let's briefly conclude the building blocks of the model

- This is a Quantitative Spatial General Equilibrium model
- Two markets: labor + land
- N prefectures, each has two regions, urban + rural
- Labor supply is determined by migration flows
- Workers choose working locations based on wages, housing prices, migration costs, and taste shocks
- With a Fréchet distributed taste shock, we derive closed-form migration flows $\Rightarrow$ Gravity Equation (Probabilistic Migration)

Model: Conclusion

- Labor demand is determined by firms
- They maximize their profits by choosing optimal labor and production floor space inputs
- Urban land is exogenously controlled by government and can be used to produce floor space
- Floor space is used as either production or residential
- A Spatial GE is achieved by a series of wages and floor space prices when
  - Workers maximize utility; Firms maximize profit
  - In each location, labor supply = labor demand
  - In each location, floor space supply = floor space demand

Model: Conclusion

- Next, we introduce how to incorporate data to this model
- Given data, we need to
  - Recover the parameters (Calibration + Estimation)
  - Recover unknown variables (Solve unobservables)
- Then, we introduce the algorithm to solve model equilibrium
- Last, we discuss the implementation of policy counterfactuals

Incorporate Data to Model

- Now we start to incorporate data to the model
- To transfer a theoretical model to a quantitative one
- We will separate all parameters and variables to three parts:
  - Observed variables
  - Estimable (or calibrated) parameters from data
  - Unobserved variables: productivity, migration cost, floor space
- We need to estimate/calibrate parameters (except agglomeration) given data
- Then recover the unobserved variables given data and parameters
- Then estimate the agglomeration parameter using data and recovered variables (productivity)

Incorporate Data to Model: Data

Data Used (233 cities with 2 sectors in both 2005 and 2010)

1. City-sector-level Hukou/working population and city-sector-pair migration flow from *Census*: $\pi, H_{in}^{s}, H_{jk}^{s}$
2. City-sector-level average residential housing cost from *Census*: Q_{ju}
3. City-sector-level high/low-skill wages from various *City Statistic Yearbooks* of each city: w_{jk}^{s}
4. Land usage and other aggregate city-sector-level data from *Urban Statistic Yearbooks*: L_j
5. Land price gap between production and residence from China Land Market Website: η_j

Incorporate Data to Model: Targeted Parameters and Unobservables

Unobservables and Parameters to be Solved or Estimated

1. Preference, Production, Friction Parameters: $\{ \beta, \alpha, \sigma, \epsilon, \tau, \eta_j, \phi_j \}$
2. Unobserved Productivities and Agglomeration Parameter: $\{ A^h_{ju}, A^l_{ju}, \gamma \}$
3. Unobserved Floor Space Market Variables: $\{ S^R_{ju}, S^M_{ju}, q_{ju} \}$
4. Migration Costs: $\tau^s_{in,jk}$

Incorporate Data to Model: Calibration and Estimation

- Let's consider parameters first
- Two ways to incorporate data: Calibration, Estimation
- Calibration is simple, using widely accepted values
- The basic idea is that you have a one-to-one mapping from model parameter to data/literature without uncertainty
- Then just match it, done
- This is a standard process to make your model comparable to other models
- Especially for more macro models and parameters

Incorporate Data to Model: Calibration and Estimation

- Estimation is more complicated, requiring you to directly use your data
- The main idea relies on how to add uncertainty to your model to capture data
- A model without uncertainty will give deterministic results
- All firms choose same FOC, all workers choose same consumption
- However, this is definitely not true in data
- Because models are limited, there are always something you cannot capture
- Also, there are always measurement errors in data

Incorporate Data to Model: Calibration and Estimation

- Thus, the key to connect data to model is to know how to add uncertainty in model
- Then with uncertainty, we "estimate" parameters
- Like in regression, you always have ϵ_{it} as the error
- On the contrary, calibration does not care about uncertainty at all
- It just simply matches a parameter with a single aggregated moment from data or previous literature
- This aggregation ignores all uncertainty in micro data

Incorporate Data to Model: Calibration and Estimation

- For instance, consider when you have a C-D utility function:

$$U = x_1^{\beta} x_2^{1-\beta} \tag{32}$$

- You want to pin down the C-D utility parameter β
- In calibration, you directly equal it as the average final consumption share of good 1 in UHS data
- Though you know there are heterogeneity and uncertainty here:
  - Each family may have different consumption shares
  - UHS is a small sample of Chinese households
  - There are measurement errors for each family's consumption composition
- You just ignore them to make your life easier

Incorporate Data to Model: Calibration and Estimation

- But if you want to consider these heterogeneity and uncertainty, you have to make utility parameters to be random coefficients
- What are random coefficients? Coefficients that are random variables
- By replacing constant β to be a random variable

$$U = x_1^{\beta_i} x_2^{1-\beta_i} \tag{33}$$

$$\beta_i \sim N(\mu_\beta, \sigma_\beta^2) \tag{34}$$

- Now β_i is different across families
- It can capture heterogeneity, sampling errors, and measurement errors we just mentioned
- But then, you have to estimate μ_β and σ_β^2 using methods like simulated method of moments

Incorporate Data to Model: Calibration and Estimation

- Except for parameters, we need to pin down other unobserved variables
- In our case, floor space, migration cost, and local productivity
- In many spatial models, we have 1-1 mapping from model to unobservables
- Then we can simply invert the model to solve them
- You can consider it as a process of calibration
- If not, we have to estimate them like parameters using the methods we will introduce

Incorporate Data to Model: Calibration and Estimation

- Now let's go through the process in our model
- To see how to estimate and pin down parameters and unobservables

Calibration

Step 1: Calibrated Fixed Parameters: $\{ \beta, \alpha, \eta_j, \sigma, \epsilon, \tau \}$

Table: **Fixed Parameters**

Parameter	Description	Value	Source
β	share of consumption in utility	0.77	*Urban Household Survey*
α	share of labor in production	0.88	*Enterprise Surveys*
η_j	relative cost of production to residential land	city-specific	China Land Market Website
σ	elasticity of substitution between H/L-skills	1.4	Katz and Murphy (1992)
ϵ	migration elasticity	1.9	Fang and Huang (2022)
τ	relative cost of rural housing	0.34	*Census*

Calibration

- When can we calibrate parameters?
  - Very common parameters widely used in other literature σ
  - Parameters that have been estimated in almost the same context in other paper ϵ
  - Parameters that clearly and exactly match one specific data pattern $\alpha, \beta, \eta, \tau$
  - Parameters that are not central to your model
- Use calibration to make your life much easier
- High-dimensional optimization in estimation is computationally intensive

Solve for Unobservables

- Now we start to solve the unobserved variables
- That is, given observed variables listed in page 61 and our estimated parameters
- We invert the model to solve for the unobserved variables listed in page 62

Solve for Unobservables

Step 2: Solve Unobserved Productivities, Floor Space, and Migration Costs from "Data Used".

- **Unobserved Productivities:** (from the FOCs of the firm)

$$A^l_{ju} = \frac{q_{ju}^{\frac{1-\alpha}{\alpha}} w^l_{ju} (\Xi^l_{ju})^{\frac{1}{\sigma-1}}}{\alpha(1-\alpha)^{\frac{1-\alpha}{\alpha}}}, \quad A^h_{ju} = \frac{q_{ju}^{\frac{1-\alpha}{\alpha}} w^h_{ju} (\Xi^h_{ju})^{\frac{1}{\sigma-1}}}{\alpha(1-\alpha)^{\frac{1-\alpha}{\alpha}}} \tag{35}$$

  where $\Xi^s_{ju} = \frac{w^s_{ju} H^s_{ju}}{w^h_{ju} H^h_{ju} + w^l_{ju} H^l_{ju}}$ is the share of labor income distributed to low skill workers.

- Intuitively, observed higher production floor prices, higher wages, and a higher share of skill s in total payroll in "Data Used" all correlates with higher skill s productivity at equilibrium.

Solve for Unobservables

Step 2: Solve Unobserved Productivities, Floor Space, and Migration Costs from "Data Used".

- **Land Market Clearing:** (from the FOCs of firm and HHs)

$$S_{ju}^R = \frac{1-\beta}{\beta Q_{ju}}\left[w_{ju}^l H_{ju}^l + w_{ju}^h H_{ju}^h\right], \quad S_{ju}^M = \left(\frac{(1-\alpha)}{q_{ju}}\right)^{\frac{1}{\alpha}} X_{ju}, \quad S_{jr}^R = \frac{1-\beta}{\beta Q_{jr}}\left[w_{jr} H_{jr}\right]$$

  We are then able to calculate the total amount of urban floor space $S_{ju} = S_{ju}^R + S_{ju}^M$ and finally back out the implied construction intensity $\phi_j = S_{ju}/L_j$.

- S_{ju}^R, S_{jr}^R come from consumers' housing demand
- S_{ju}^M comes from the firm FOC w.r.t. production floor space

Solve for Unobservables

Step 2: Solve Unobserved Productivities, Floor Space, and Migration Costs from "Data Used".

- **Migration Costs:** (from gravity equation) We have the migration flow as:

$$\pi^s_{in,jk} = \frac{(\tau^s_{in,jk} Q^{1-\beta}_{jk})^{-\epsilon}(v^s_{in,jk})^{\epsilon}}{\sum_{j'k'=11}^{JK}(\tau^s_{in,j'k'} Q^{1-\beta}_{j'k'})^{-\epsilon}(v^s_{in,j'k'})^{\epsilon}} \tag{36}$$

We also assume that $\pi^s_{in,in} = 1$. Therefore, we can solve the denominator $\Phi^s_{in} = \sum_{j'k'=11}^{JK}(\tau^s_{in,j'k'} Q^{1-\beta}_{j'k'})^{-\epsilon}(v^s_{in,j'k'})^{\epsilon}$ as:

$$\Phi^s_{in} = (\tau^s_{in,in} Q^{1-\beta}_{in})^{-\epsilon}(v^s_{in,in})^{\epsilon} \tag{37}$$

Plug this back, we have:

$$\tau^s_{in,jk} = \frac{v^s_{in,jk}}{Q^{1-\beta}_{jk}(\pi^s_{in,jk}\Phi^s_{in})^{1/\epsilon}} \text{ , for } i \neq j \tag{38}$$

Estimation

- Agglomeration parameter γ is one of the most important parameters we have in our model
- What is the effect of population density on productivity?
- It controls the strength of the main channel
- It is also not estimated in other literature in China
- We will use a method called indirect inference

Estimation

- Before we go to the estimation of this paper
- Let's take a detour to introduce four main methods of estimation commonly used in structural works:
  - Regression derived from model
  - Generalized Method of Moments/Simulated Method of Moments
  - Maximum Likelihood Estimation/Simulated Maximum Likelihood
  - Indirect Inference
- GMM/SMM and MLE/SML are not used in this paper
- We will not discuss GMM/SMM and MLE/SML in very details
- Rather, we will focus on the basic ideas
- A full and long structural course is required to learn it:
  Professor Junjian Yi has a great course

Estimation: Method Introduction - Regression

- The first estimation method is simple
- We can linearize equations in the model, add error terms, and run regressions to estimate the parameters
- Specifically, this is widely used in estimating gravity equations
- Now let's see how to implement it in this model

Estimation: Method Introduction - Regression

- Let's implement log linearization for equation (37):

$$ln\pi^s_{in,jk} = -\epsilon ln(\tau^s_{in,jk}) - \epsilon(1-\beta)ln(Q_{jk}) + \epsilon ln(v^s_{in,jk}) - ln(\sum_{j'k'=11}^{JK}(\tau^s_{in,j'k'}Q^{1-\beta}_{j'k'})^{-\epsilon}(v^s_{in,j'k'})^{\epsilon}) \quad (39)$$

- The denominator is a constant for each $s - in$, which is exogenous
- We can replace it by FE of $s - in$
- Migration cost can be decomposed by equation (4): $\tau^s_{in,jk} = \bar{\tau}^s_{in} d_{in,jk}$
- Which can be absorbed in a $s - in$ level FE and $in - jk$ level FE

Estimation: Method Introduction - Regression

- Thus, we can simplify the linear regression as:

$$ln\pi^s_{in,jk} = \epsilon ln(v^s_{in,jk}) + FE_{in,jk} + FE^s_{in} + error^s_{in,jk} \tag{40}$$

- $error^s_{in,jk}$ is the measurement error we add to capture the uncertainty in data
- This regression can be easily estimated using OLS to get ϵ
- We can see from the regression that the meaning of ϵ is the elasticity of migration on income
- If considering endogeneity of income v, we can also use IV
- We will discuss more about using shift-share IV to estimate regressions like this in later classes

Estimation: Method Introduction - GMM/SMM

- The second method is GMM/SMM
- This is also widely used since it is simple and clear
- The basic idea is to match moments from data to moments from model
  - Step 1: Calculate some data moments (say, average test scores of different groups of students)
  - Step 2: Given a set of parameters, simulate the model moments many times
  - Step 3: Find the parameters that can generate the most similar model moments compared with data moments

Estimation: Method Introduction - GMM/SMM

- Assume that we want to estimate θ in a model $y_i = r(err_i; \theta_0)$
- err is error, θ_0 is the true value of the parameter in DGP
- i is individual, we have n people in the sample
- Define a set of moments in data:

$$E[K(y_i)] = E[K(r(err_i; \theta_0))]$$

- $K(y_i)$ can be some endogenous outcome you can find in data
- RHS of this equation means that data comes from the "true model"
- $E[K(y_i)]$ is the mean outcome in the real world
- If you know θ_0, you can exactly match $E[K(y_i)]$ by your model results $E[K(r(err_i; \theta_0))]$

Estimation: Method Introduction - GMM/SMM

- Given a guess of parameter $\tilde{\theta}$, define a set of moments simulated from your model:

$$\tilde{k}(err_i^s; \tilde{\theta}) = K(r(err_i^s; \tilde{\theta}))$$

- Superscript s means this is the s-th simulation
- We simulate this outcome $\tilde{k}$ for each individual i for S times
- Then for one guess of θ, the simulated outcome for individual i is $\frac{1}{S}\sum_s \tilde{k}(err_i^s; \theta)$

Estimation: Method Introduction - GMM/SMM

- At last, we minimize the following function to find the best θ:

$$\Phi(\theta) = \left\{ \sum_i^N \left[K(y_i) - \frac{1}{S} \sum_s \tilde{k}(err_i^s;\theta) \right] \right\}' \Omega \left\{ \sum_i^N \left[K(y_i) - \frac{1}{S} \sum_s \tilde{k}(err_i^s;\theta) \right] \right\}$$

- Ω is some weighting matrix
- $\Phi(\theta)$ is a weighted euclidean distance between data moments and simulated moments

Estimation: Method Introduction - GMM/SMM

- Let's give an example in this model
- Say we want to estimate migration elasticity ϵ in this model
- We can choose the set of data moments as migrants' shares in each city
- Then we simulate our model for different guesses of ϵ and get the simulated migration choices for each worker
- Then we average over these simulated choices to have a series of model moments $\Rightarrow$ simulated migration shares from the model for each guess
- We minimize the distance between real and simulated migrants' shares to get an estimation of ϵ^*

Estimation: Method Introduction - GMM/SMM

- GMM/SMM is still the most commonly used method
- The advantage of GMM/SMM is that it is relatively simple to implement
  - Sometimes no complicated distributional assumptions
  - No need to write likelihood function: in complicated models, usually we cannot derive a closed-form likelihood function
- The disadvantage of GMM/SMM is that we do not use full information
- We match several moments, rather than the whole distribution
- Moreover, the identification is always a question
- Do we have enough moments for our parameters?
- Hard to prove it rigorously

Estimation: Method Introduction - MLE/SML

- The third common estimation method is MLE/SML
- This is the most efficient estimation method
- Because it requires you to write the likelihood function of the endogenous variables and make full use of the information
- GMM/SMM matches only some moments; MLE/SML matches the full distribution
- There are three steps:
  - We make full distributional assumptions for all uncertainty in a model
  - Then we directly calculate the likelihood for this model to generate the observed endogenous variables
  - We maximize this likelihood function by optimizing parameters

Estimation: Method Introduction - MLE/SML

- Although MLE is efficient
- It is the most complicated estimation method of all these four
- When you have a complicated model, say Dynamic GE
- The likelihood function can be very hard to write
- Even if you write it down, the simulation process can still be a disaster
- It may take a whole semester of classes to learn this

Estimation: Method Introduction - Indirect Inference

- The last method we introduce is called Indirect Inference
- The basic idea is as simple as follows
- First, we create an "auxiliary model" from the main one
- This auxiliary model can be a regression (like DID) with parameter β
- Then we implement the following three steps:
  - Run this regression with real data to get $\hat{\beta}^{data}$
  - Given some guess of θ, simulate all needed variables from main model, then run this regression with simulated (S times) data and get $\hat{\beta}^{s}$
  - Choosing θ in main model to minimize the distance between $\hat{\beta}^{s}$ and $\hat{\beta}^{data}$

Estimation: Method Introduction - Indirect Inference

- The key point is to differentiate between θ and β
- θ is the target structural parameter we want to estimate in the main model
- β is the parameter generated in an anxiliary model
- $\beta \neq \theta$, but β gives us information of θ
- That is, in our model, with some θ, we can generate the simulated data, which can give us same β as using the real data
- To match model simulated "coefficient" with data "coefficient"
- This is used to estimate the agglomeration parameter in this paper

Estimation: Method Introduction - Indirect Inference

- Therefore, indirect inference looks like GMM/SMM
- What you do is to match something generated from model to the corresponding thing from data
- In indirect inference, you usually match regression coefficients
- Something like a treatment effect, a policy effect
- It is a great way to combine design-based and structural approaches
- In GMM/SMM, you match moments
- But if you consider regression coefficients as some special moments, then they are the same thing

Estimation: Agglomeration Parameter

- OK we have already learned four important estimation methods
- Now let's go back to the model
- Let me show you what methods do we use to estimate the agglomeration parameter in this paper

Estimation: Agglomeration Parameter

- We have the decomposition of productivity as:

$$A^s_{ju} = a^s_{ju} \times (D_{ju})^{\gamma}, \quad D_{ju} = \frac{H^h_{ju} + H^l_{ju}}{\bar{L}_j} \tag{41}$$

- a^s_{ju} is the fundamental productivity
- D_{ju} is the urban population density
- Now, what is the simplest way to estimate γ?
- Log linearize it and run a reg, right?

Estimation: Agglomeration Parameter

- We can have a regression as follows:

$$log(A^s_{ju}) = \gamma log(D_{ju}) + log(a^s_{ju})$$

- We already calculate A^s_{ju}, and know D_{ju} from data
- Can we estimate γ by running A^s_{ju} on D_{ju} and treat a^s_{ju} as error term?
- No! Because $log(a^s_{ju})$ is unobserved and endogenous
- Based on our model, fundamental productivity affects wages and housing prices, thus, migration flows
- Which is surely correlated with $log(D_{ju})$

Estimation: Agglomeration Parameter

- This goes back to a very traditional identification issue in urban economics
- How to distinguish agglomeration from fundamental productivity?
- Several methods are available (Combes and Gobillon, 2015)
  - Find IV, such as geographic conditions or population in ancient times (Ciccone and Hall, 1996)
  - Use a natural experiment (Ahlfeldt et al., 2015)
  - Model the mechanism of agglomeration and estimate (Baum-Snow and Pavan, 2012)
- IV is very hard to find in China due to data limitation
- Fortunately, we have a natural experiment: Inland-favoring land policy

Estimation: Agglomeration Parameter

- The identification idea is that the reallocation of land quota will not affect fundamental productivity in different locations
- a^s_{ju} is exogenous in the model
- Thus, given γ, we can simulate the model with/without inland-favoring policy
- Then, we use simulated data to run the same DID regression as in our empirical part
- We find the best γ to match these two coefficients

Estimation: Agglomeration Parameter

Step 3: Estimate Agglomeration Elasticity using Indirect Inference

- First calculate real world equilibrium city-urban TFP: $ln(\widetilde{TFP}_{ju}) = ln\left(\frac{Y_{ju}}{(H^h_{ju}+H^l_{ju})^\alpha}\right)$
- Second, choose agglomeration elasticity γ_0 (and correspondingly, $a_j^{s,0}$), simulate a counterfactual equilibrium of 2005 without inland-favoring policy, then calculate: $ln(\widetilde{TFP}^0_{ju})$
- Third, run pooled reg. of "real world" ($Post2003 = 1$) and counterfactual ($Post2003 = 0$):

$$ln(\widetilde{TFP}^0_{ju}) = \alpha + \delta_1^0 Post2003 \times East_{ju} + \phi_j + \gamma_t + \epsilon_{jut}$$

Estimation: Agglomeration Parameter

- We have the following empirical results from real data:

Data Estimation of $\hat{\delta}_1$		
	(1) OP	(2) LP
Post2003×East	-0.0749***	-0.0516*
	(0.0241)	(0.0268)
Trend Variables	Y	Y
Year FE	Y	Y
Prefecture FE	Y	Y
Observations	1,788	1,788
R-squared	0.7537	0.6351

Estimation: Agglomeration Parameter

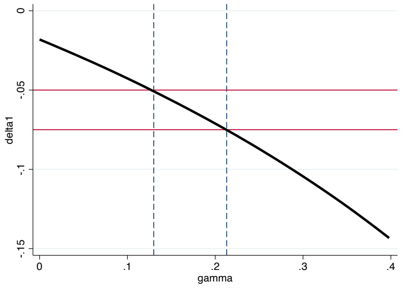


Figure: Relation between γ and δ

Estimation: Agglomeration Parameter

- γ and δ are positively correlated
- Why?
  Because agglomeration effect amplifies the impact of the inland-favoring policy
- The inland-favoring land policy $\Rightarrow$ migration from inland to east $\downarrow$ $\Rightarrow$ Agglomeration in east $\downarrow$

Estimation: Agglomeration Parameter

- We have a range of coefficient $\hat{\delta}_1^* \in [-0.075, -0.05]$
- It corresponds to $\gamma \in [0.13, 0.21]$
- This is larger than the agglomeration effect estimated in developed countries
- There can be several reasons:
  - Supply chain integration is more profitable when trade cost is high in China
  - Knowledge spillover is strong within regions relative to across regions in China

Solve Model Equilibrium: Overfitting

- We have all parameters and variables in our hands
- It is time for us to calculate the equilibrium using our model
- We solve the model Eq separately in 2005 and 2010
- Usually equilibrium fitness is very good in QSGE models
- Since we impose a large set of free parameters for migration cost and productivity
- There are criticism about over-fitting here (Dingel and Tintelnot, 2020)

Solve Model Equilibrium: Overfitting

- Especially when the geographical setting is very "granular"
- When we have one migration cost for each pair of cities, we are trying to match migration flows for all city pairs
- The geographical dimension can be very high
- If we have 3000 counties in China, then we have $3,000 \times 3,000 = 9,000,000$ county pairs

Solve Model Equilibrium: Overfitting

- Even if we have Census data, it usually covers only 1 million people
- But we have 9 million pairs!! There will be a lot of pairs with no people at all
- Then, you will have infinitely large migration cost for these pairs
- However, is there really no people in these pairs? Not necessarily. We have only a small sample of all people in China
- The small sample issue naturally creates a measurement error problem

Solve Model Equilibrium: Overfitting

- If we have too many migration cost parameters, we are not only fitting signals, but also these measurement errors
- That is why you may have many infinitely large migration costs
- This is a very typical bias-variance tradeoff
- We are making this model too complicated to capture all data pattern, including noises
- Perfect within sample prediction means poor out-of-sample prediction

Solve Model Equilibrium: Overfitting

Remember this?

Solve Model Equilibrium: Overfitting

- A fix to this problem is to parameterize migration cost
- We assume that migration cost is a function of distance
- Rather than consider it as a series of fixed effects
- We will further discuss this issue later in the dynamic model part

Solve Model Equilibrium: Algorithm

- Now we show how to solve the equilibrium
- The main target of the equilibrium solving process is to calculate endogenous variables using parameters and exogenous variables
- It is used to calculate: Original Eq + Counterfactual Eq
- This is a reverse process of estimation/calibration
- Estimation/calibration:
  Endogenous and exogenous variables in data $\Rightarrow$ parameter
- Solving model equilibrium:
  Parameter and exogenous variables in data $\Rightarrow$ Endogenous variables from model
- Then compare endogenous variables solved from model to data/facts (fitness)

Solve Model Equilibrium: Algorithm

- The existence and the uniqueness of the model equilibrium is always a problem
- Allen and Arkolakis (2014) shows that in a spatial GE model, as long as the agglomeration (positive externality) effect is not too strong compared with the congestion (negative externality), the equilibrium exists and is unique
- The intuition is clear: if positive externality is too high, there can be at least two equilibria you can imagine
  - Everyone stay in Shanghai, reaping off spillovers from each other
  - People distribute in different places, with no incentive to move to Shanghai when others don't
- That is why agglomeration and congestion forces are important

Solve Model Equilibrium: Algorithm

- Exogenous variables: $\{\tau_{ij}^s, a_j^s, \eta_j, \phi_j, L_j, H_{in}^s\}$
- Endogenous variables: $\{Q_{ju}, q_{ju}, w_{jk}^s, S_{ju}^R, S_{ju}^M, v_{in,jk}^s, \pi_{in,jk}^s, Y_{jk}\}$
- Three blocks of this model:
  - Migration Block: worker income and gravity equations
  - Production Block: production, wage, and floor space price equations
  - Housing Block: construction equations and market clearing equations

Solve Model Equilibrium: Algorithm

- There are two ways to solve the equilibrium in QSGE models
- The first is the "hat-algebra", which solves the model in "change" but not level
- Anyone who is interested in it can go to Prof Deng's course
- We will discuss it a little bit in the dynamic model part
- Today, we introduce a more traditional and general method
- Using contraction algorithm, we solve the model in level
- This method can be used in any structural model

Solve Model Equilibrium: Algorithm

- Before we go to the algorithm, we need to understand the theory behind it
- Let's take a review of the fixed-point theorem
- What are metric and metric space?
- This goes back to your introduction to analysis or introductory math econ course

Solve Model Equilibrium: Algorithm

Definition: Metric Space

A metric space is an ordered pair (M, d) where M is a set of points and d is a metric on M, that is, a function $d : M \times M \to \mathbb{R}$ satisfying the following axioms for all points $x, y, z \in M$:
(1) $d(x, x) = 0$;
(2) Positivity: If $x \neq y$, then $d(x, y) > 0$;
(3) Symmetry: $d(x, y) = d(y, x)$;
(4) Triangle inequality: $d(x, z) \leq d(x, y) + d(y, z)$.

- d is a distance function defined in the space of M
- Like Euclidean distance in the Euclidean space

Solve Model Equilibrium: Algorithm

Definition: Contraction Mapping

Let (X, d) be a metric space. Then a map $T : X \to X$ is called a contraction mapping on X if there exists $q \in [0, 1)$ such that:

$$d(T(x), T(y)) \leq qd(x, y), \forall x, y \in X$$

- The contraction mapping means a function that squeezes points closer together in a space
- If you have such a function, then you can apply it again and again and squeeze everything to one point

Solve Model Equilibrium: Algorithm

- Then we have the famous Banach fixed-point theorem

Banach Fixed-point Theorem

Let (X, d) be a non-empty complete metric space with a contraction mapping $T : X \rightarrow X$. Then T admits a unique fixed-point x^* in X, that is, $T(x^*) = x^*$. Furthermore, x^* can be found as follows: start with an arbitrary element $x_0 \in X$ and define a sequence $(x_n)_{n \in N}$ by $x_n = T(x_{n-1})$ for $n \geq 1$. Then, $lim_{n \rightarrow \infty} x_n = x^*$

- The existence of a contraction mapping $T \Rightarrow$ Unique fixed point $T(x^*) = x^*$
- We can find x^* by iterate some arbitrary initial x_0 with T
- $x_0, T(x_0), T(T(x_0)), T(T(T(x_0)))...)$
- Therefore, we can use this method to find the equilibrium (fixed-point in the model)

Solve Model Equilibrium: Algorithm

- There are two reasons why we need to solve the equilibrium
- First, we want to evaluate the fitness of the model
- We want to see how similar it is for equilibrium values in the model, compared with those in data
- Here is the process:
  - We start from one block using some variable value (data value in the first iteration)
  - We solve the variables in other blocks using model equations and update them with these values
  - We solve the variables in the starting block by model equations and *updated* values of other variables
  - We iterate this process until all endogenous variables converge

Solve Model Equilibrium: Algorithm

- Second, we need to solve the variables in the model under counterfactuals
- To calculate the counterfactuals following policy changes
- The process is similar
- We start with the block where the changes happen
- Then iterate block by block to update the endogenous variables
- The iteration ends until all endogenous variables converge
- Example: An increase in land supply

Solve Model Equilibrium: Algorithm

- Let's denote x^0 as the initial value
- $\hat{x}^n$ as the updated value in the n-th iteration,
- x^n as the final updated value of the n-th iteration, which will be used in the next iteration

Solve Model Equilibrium: Algorithm

Step 1: Initiation (ensuring non-zero floor space supply)
Change in land supply $\Rightarrow$ Change in urban floor space supply

$$\{\hat{S}_{ju}\}^1 = \phi_j \hat{L}_j^0$$

$$\{\hat{S}_{ju}^R\}^1 = S_{ju}^R \times (\{\hat{S}_{ju}\}^1 / S_{ju}^0)$$

$$\{\hat{S}_{ju}^M\}^1 = S_{ju}^M \times (\{\hat{S}_{ju}\}^1 / S_{ju}^0)$$

Solve Model Equilibrium: Algorithm

Step 2: Update floor space price

- Given updated values S and L, update residential and production floor space prices

$$\{\hat{Q}_{ju}\}^{\mathbf{1}} = \frac{1-\beta}{\beta} \frac{\{w_{ju}^{l} H_{ju}^{l} + w_{ju}^{h} H_{ju}^{h}\}^{\mathbf{0}}}{\{\hat{S}_{ju}^{R}\}^{1}}$$

$$\{\hat{q}_{ju}\}^{\mathbf{1}} = (1-\alpha)\left(\frac{\alpha}{\{\hat{W}_{ju}\}^{\mathbf{0}}}\right)^{\frac{\alpha}{1-\alpha}}$$

Solve Model Equilibrium: Algorithm

Step 3: Update wage and income

- Given updated values S and Q, update wage and income

$$\{\hat{w}^l_{ju}\}^1 = \alpha\{X_{ju}\}^{0\,\alpha-1}\{\hat{S}^M_{ju}\}^{1\,1-\alpha}\{A_{ju}\}^{0\,l\,\frac{\sigma-1}{\sigma}}\{X_{ju}\}^{0\,\frac{1}{\sigma}}\{H^l_{ju}\}^{0\,-\frac{1}{\sigma}}$$

$$\{\hat{w}^h_{ju}\}^1 = \alpha\{X_{ju}\}^{0\,\alpha-1}\{\hat{S}^M_{ju}\}^{1\,1-\alpha}\{A_{ju}\}^{0\,h\,\frac{\sigma-1}{\sigma}}\{X_{ju}\}^{0\,\frac{1}{\sigma}}\{H^h_{ju}\}^{0\,-\frac{1}{\sigma}}$$

$$\{v\}^1 = \{\hat{w}\}^1 + \frac{\hat{Q}^1\{\hat{S}^R\}^1}{\{H^R\}^0}$$

Solve Model Equilibrium: Algorithm

Step 4: Update migration and population distribution

- Given updated values v and Q, update migration probability

$$\{\hat{\pi}^s_{in,jk}\}^1 = \frac{(\tau^s_{in,jk}\{\hat{Q}_{jk}\}^{1^{1-\beta}})^{-\epsilon}(\{\hat{v}^s_{in,jk}\}^1)^{\epsilon}}{\sum_{j'k'=11}^{JK}(\tau^s_{in,j'k'}\{\hat{Q}_{j'k'}\}^{1^{1-\beta}})^{-\epsilon}(\{\hat{v}^s_{in,j'k'}\}^1)^{\epsilon}}$$

- Using updated migration probability, we can update population distribution across locations H

$$\{\hat{H}^s_{jk}\}^1 = \sum_{in}\{\hat{\pi}^s_{in,jk}\}^1 \cdot H^s_{in}$$

Solve Model Equilibrium: Algorithm

Step 5: Step length and damping factor

- Up until now, we have updated all necessary endogenous variables in the model
- Before we iterate, we need a damping factor ξ:
  $\{x\}^1 = (1-\xi)x^0 + \xi\{\hat{x}\}^1$
- For endogenous variables in the next loop $\{x\}^1$, we do not directly use $\{\hat{x}\}^1$, but a weighted average of the old and the new values
- This technique is used to control the pace of the iteration
- In some cases, if the updating speed is too fast, we may fail the convergence
- Usually we can choose 0.5 (just my habit)

Solve Model Equilibrium

- Implement 2-5 until converge
- That is, until $|\{x\}^n - \{x\}^{n-1}| \leq 1 \times 10^8$
- For the metric evaluating the distance between vectors $\{x\}^n$ and $\{x\}^{n-1}$, we can use L1 norm or L2 norm
- The convergence tolerance 1×10^8 can also be changed to some other small number

Solve Model Equilibrium

- Using this algorithm, we can calculate model responses when some policy is implemented
- We compare the original Eq with the changed Eq
- We can then evaluate the policy effect

Solve Model Equilibrium: Analysis

- Before we go to the results of the counterfactuals
- Let's first analyze the original equilibrium solved from the data
- We solve the model separately for data in 2005 and 2010
- Let's see what is going on in reality in China in 2005 and 2010

Solve Model Equilibrium: Analysis

- First, we investigate the spatial distribution of the productivity
- The measured productivity in our model is $ln(\widetilde{Prod}_{ju}) = ln\left(\frac{Y_{ju}}{(H^h_{ju}+H^l_{ju})^\alpha}\right)$
- We can decompose the labor productivity as follows:

$$\begin{aligned} ln(\widetilde{Prod}_{ju}) =& (1-\alpha)ln(S^M_{ju}) + \alpha ln\left(\frac{[(A^h_{ju}H^h_{ju})^{\frac{\sigma-1}{\sigma}} + (A^l_{ju}H^l_{ju})^{\frac{\sigma-1}{\sigma}}]^{\frac{\sigma}{\sigma-1}}}{H^h_{ju}+H^l_{ju}}\right) \\ =& \underbrace{(1-\alpha)ln(S^M_{ju})}_{\text{land scale premium}} + \frac{\alpha\sigma}{\sigma-1}\left(\underbrace{ln\left((\frac{A^h_{ju}}{A^l_{ju}}\Gamma^h_{ju})^{\frac{\sigma-1}{\sigma}} + (\Gamma^l_{ju})^{\frac{\sigma-1}{\sigma}}\right)}_{\text{skill premium}} + \underbrace{ln(a^l_{ju})}_{\text{fundamental}} + \underbrace{\gamma ln(D_{ju})}_{\text{agglomeration}}\right) \end{aligned} \tag{42}$$

Solve Model Equilibrium: Analysis

- We calculate each component of this productivity for six groups of prefectures in 2005 and 2010
- The six groups are categorized as follows:
  - Divide by development level $\{high, mid, low\}$
    Based on $\{10\%, 45\%, 45\%\}$ of the distribution of GDP per capita.
  - Divide by region: East vs Inland

Solve Model Equilibrium: Analysis

Table: Spatial Distribution of Measured Productivity and Land Tightness

		Measured Productivity										Land Tightness	
Regions	No. of	2005					2010					2005	2010
(loc., dev.)	prefectures	Total	LSP	SP	Fund	Agg	Total	LSP	SP	Fund	Agg	Land/Worker	
National	225	33.84	2.19	0.59	31.06	-0.01	35.86	2.22	0.62	32.92	0.11	0.093	0.083
(east, high)	21	35.21	2.24	0.67	32.07	0.22	36.81	2.29	0.67	33.51	0.33	0.077	0.068
(east, mid)	51	33.84	2.25	0.49	31.06	0.04	35.75	2.24	0.57	32.76	0.17	0.084	0.082
(east, low)	25	32.61	2.13	0.50	30.00	-0.02	34.84	2.06	0.50	32.57	-0.30	0.080	0.108
(inland, high)	2	33.69	2.06	0.59	31.44	-0.40	35.24	2.13	0.77	32.65	-0.33	0.127	0.130
(inland, mid)	50	32.97	2.11	0.69	30.34	-0.17	35.35	2.17	0.69	32.40	0.09	0.140	0.101
(inland, low)	76	32.50	2.09	0.56	30.21	-0.37	35.10	2.14	0.52	32.74	-0.30	0.104	0.086

Solve Model Equilibrium: Analysis

- Fundamentals and agglomeration effects drive the spatial dispersion of productivity
- Fundamentals and agglomeration effects also drive the growth of productivity
- Eastern and more developed prefectures have 30% to 50% less land per worker
- Land tightness in Eastern and more developed prefectures is worsening from 2005 to 2010

Solve Model Equilibrium: Analysis

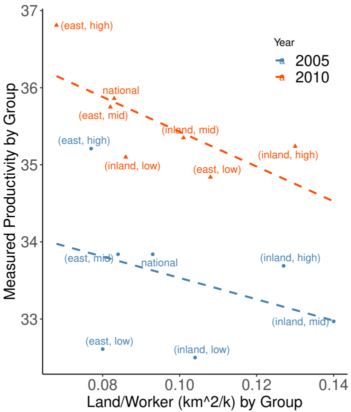


Figure: Correlation between Productivity and Land Tightness

Counterfactual Analysis: No Inland-favoring Policy

- Now let's go to the main counterfactual
- What will happen if we get rid of this inland-favoring land supply policy in 2003?
- In the model, we keep the land growth rate before 2003 unchanged
- For prefecture j in year t, we have the following allocation rule:

$$\widehat{L_j(t)} = L_j(2003) + \underbrace{\sum_j [L_j(t) - L_j(2003)]}_{\text{actual total increment of land}} \times \underbrace{\frac{L_j(2003)(1+g_{L_j})^{t-2003}}{\sum_j L_j(2003)(1+g_{L_j})^{t-2003}}}_{\text{prefecture j's share if no inland-favoring}}$$

- $L_j(2003)$: Urban land stock in 2003
- $\sum_j [L_j(t) - L_j(2003)]$: Actual national total increment of land
- g_{L_j}: Average land supply growth rate before 2003

Counterfactual Analysis: No Inland-favoring Policy

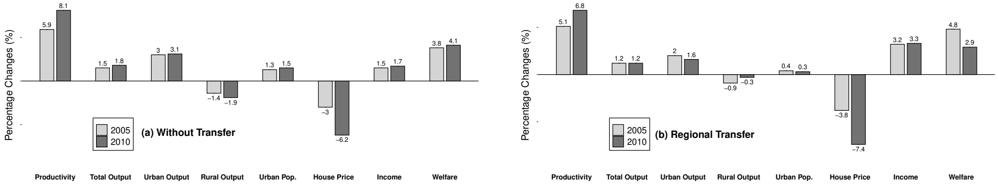
Table: Counterfactual Total Land Supply (km^2)

Regions (loc., dev.)	No. of prefectures	Reality 2005	Reality 2010	Counterfactual $\widehat{2005}$	Counterfactual $\widehat{2010}$
National	225	22268	28336	22268	28336
(east, high)	21	5838	7272	6597	10958
(east, mid)	51	5875	7832	5734	6551
(east, low)	25	1418	1681	1472	1596
(inland, high)	2	169	206	169	169
(inland, mid)	50	5131	6578	4537	4819
(inland, low)	76	3837	4767	3760	4244

Counterfactual Analysis: No Inland-favoring Policy

- We know that the target of inland-favoring land policy is to encourage the development of underdeveloped regions
- Can we have other options to achieve it?
- A policy of replacement is a regional transfer program
- We allocate land as before 2003, but transfer additional income in developed cities to underdeveloped ones
- It mimics a regional land quota trading system
- Inland cities can sell land quota to eastern cities

Counterfactual Analysis: No Inland-favoring Policy



(a) No Transfer (b) With Transfer

Figure: National Results of Main Counterfactual

Counterfactual Analysis: No Inland-favoring Policy

Conclusions at national level

- Removing inland-favoring land policy can
  - Increase national productivity by 5.9% (8.1%)
  - Increase total output by 1.5% (1.8%)
  - Increase urban population by 1.3% (1.5%)
  - Increase Welfare by 3.8% (4.1%)
- Adding regional transfer may distort the market a little
- But much smaller than the original inland-favoring policy

Counterfactual Analysis: No Inland-favoring Policy

- Inland-favoring land policy affected China's economy by distorting both land and labor markets
- We further decompose the impact of the inland-favoring policy in three channels:
  - Direct effect from production floor space input changes
  - Indirect effect from induced labor demand and supply changes
  - Agglomeration effect from induced population density changes
- We can shut down channels by fixing different variables

Counterfactual Analysis: No Inland-favoring Policy

Table: Aggregate Effects Decomposition

Decomp.	Δ Productivity		Δ Urban Output		Δ Rural Output		Δ Urban Pop.	
	2005	2010	2005	2010	2005	2010	2005	2010
	(a) Without Transfer							
Total	5.9%	8.1%	3.0%	3.1%	-1.4%	-1.9%	1.3%	1.5%
Direct	0.3%	-0.8%	0.3%	-0.8%	0.0%	0.0%	0.0%	0.0%
Indirect	3.2%	5.8%	1.6%	3.2%	-1.4%	-1.7%	1.3%	1.2%
Agglomeration	2.4%	3.1%	1.0%	0.8%	0.0%	-0.3%	0.0%	0.3%
	(b) Regional Transfer							
Total	5.1%	6.8%	2.0%	1.6%	-0.9%	-0.3%	0.4%	0.3%
Direct	0.3%	-0.8%	0.3%	-0.8%	0.0%	0.0%	0.0%	0.0%
Indirect	2.9%	5.3%	1.1%	2.4%	-0.9%	-0.6%	0.4%	0.3%
Agglomeration	1.9%	2.3%	0.6%	0.0%	0.0%	0.3%	0.0%	0.0%

Counterfactual Analysis: No Inland-favoring Policy

- Indirect channel and agglomeration effects accounts for most productivity/output reductions
- The distortion effect of the inland-favoring land policy on labor markets is very important
- The general equilibrium effect is very important

Counterfactual Analysis: No Inland-favoring Policy

- We have investigated the national effect
- Now let's go to the spatial effect
- This is crucial since it directly points to the main target of the policy
- Can inland-favoring land policy achieve its original goal to promote regional balanced development?

Counterfactual Analysis: No Inland-favoring Policy

Table: Spatial Effects on Economic Development

Regions (loc., dev.)	No. of prefectures	Δ Productivity $\widehat{2005}$	Δ Productivity $\widehat{2010}$	Δ Urban Output $\widehat{2005}$	Δ Urban Output $\widehat{2010}$	Δ Rural Output $\widehat{2005}$	Δ Rural Output $\widehat{2010}$	Δ Urban Pop. $\widehat{2005}$	Δ Urban Pop. $\widehat{2010}$	Δ House Price $\widehat{2005}$	Δ House Price $\widehat{2010}$
National	225	5.9%	8.1%	3.0%	3.1%	-1.4%	-1.9%	1.3%	1.5%	-3.0%	-6.2%
(east, high)	21	7.4%	14.9%	8.1%	17.8%	0.0%	3.3%	6.9%	13.9%	-17.4%	-32.4%
(east, mid)	51	-0.3%	-2.3%	-0.7%	-4.4%	-0.4%	0.0%	-0.7%	-3.0%	1.4%	11.9%
(east, low)	25	-0.6%	-2.7%	-0.8%	-4.6%	-1.4%	-3.5%	-0.6%	-3.2%	-3.3%	2.8%
(inland, high)	2	-0.1%	-2.6%	0.0%	-3.2%	0.0%	1.7%	0.1%	-1.0%	1.6%	18.5%
(inland, mid)	50	-0.7%	-7.8%	-2.6%	-11.8%	-1.5%	-2.9%	-1.9%	-7.5%	1.6%	9.6%
(inland, low)	76	-0.4%	-4.9%	-1.7%	-6.7%	-1.9%	-3.2%	-1.6%	-5.1%	-3.8%	-1.7%

Counterfactual Analysis: No Inland-favoring Policy

Table: Spatial Effects on Income and Welfare

Regions (loc., dev.)	No. of prefectures	Without Transfer: Δ Income $\widehat{2005}$	Without Transfer: Δ Income $\widehat{2010}$	Without Transfer: Δ Welfare $\widehat{2005}$	Without Transfer: Δ Welfare $\widehat{2010}$	Regional Transfer: Δ Income $\widehat{2005}$	Regional Transfer: Δ Income $\widehat{2010}$	Regional Transfer: Δ Welfare $\widehat{2005}$	Regional Transfer: Δ Welfare $\widehat{2010}$
National	225	1.46%	1.74%	3.8%	4.1%	3.18%	3.26%	4.8%	2.9%
(east, high)	21	2.69%	7.43%	10.8%	14.5%	-10.3%	-10.9%	7.7%	2.5%
(east, mid)	51	0.28%	-0.08%	-0.2%	-4.0%	0.49%	5.03%	1.2%	2.1%
(east, low)	25	1.10%	1.92%	-1.5%	1.2%	0.72%	6.49%	1.9%	6.3%
(inland, high)	2	0.01%	-1.61%	-0.6%	-5.3%	2.30%	5.63%	2.0%	3.1%
(inland, mid)	50	0.95%	-0.91%	-0.1%	-5.1%	20.0%	6.94%	5.8%	4.2%
(inland, low)	76	2.24%	1.92%	2.7%	-3.5%	6.49%	7.05%	5.0%	4.0%

Counterfactual Analysis: No Inland-favoring Policy

- By removing the inland-favoring land policy, we can
  - Increase population, productivity and output in developed regions
  - But decrease population, productivity and output in underdeveloped regions
  - Housing prices are dramatically reduced in developed regions
- Thus, the inland-favoring land policy shrank the regional development gap
- But did it really help people from there?

Counterfactual Analysis: No Inland-favoring Policy

- We increase incomes for workers from underdeveloped regions by removing the policy
- The welfare effect is at best mixed
- Thus, inland-favoring policy helped region, but not necessarily people there
- Since it prevented people from migrating to developed regions
- Replace it with a regional transfer can help people from poor areas with minimal spacial misallocation

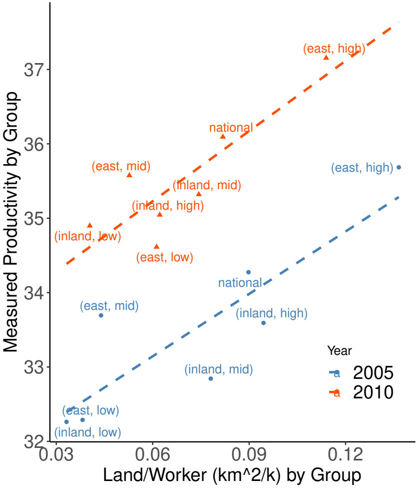
Counterfactual Analysis: Optimal Policy

- In the main counterfactual, we investigate removing the inland-favoring policy in 2003
- But land allocation before 2003 is not necessarily efficient
- What will happen if we eliminate all frictions in land market?
- How much did the inland-favoring land policy account for in the overall misallocation?
- It then goes to finding an optimal land allocation

Counterfactual Analysis: Optimal Policy

- In total, there are three layers of misallocation
  - National total land supply cap
  - Land supply allocation across prefectures
  - Production & residential land allocation within prefecture
- We ignore the first one and focus on the second and the third ones
- Optimal policy construction: We find the land allocation rule $\{L_j, \eta_j\}$ such that:
  - (1) Marginal production output of land is equalized across regions
    $\partial Y_{ju}/\partial S_{ju}^M = \partial Y_{iu}/\partial S_{iu}^M$, $(q_{ju} = q_{iu})$, for any i, j
  - (2) Price gap between production and residential floor space is eliminated
    $\eta_j = 1$, $(q_{ju} = Q_{ju})$, for any j

Counterfactual Analysis: Optimal Policy



Counterfactual Analysis: Optimal Policy

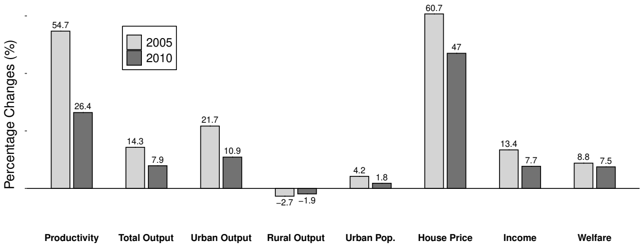


Figure: Results of the Optimal Policy

Counterfactual Analysis: Optimal Policy

- The optimal policy can increase
  - Productivity by 55% (26%)
  - Output by 14% (8%)
  - Welfare by 8.8% (7.5%)
- The overall spatial misallocation in China is large
- But it reduced across time from 2005 to 2010
- Inland-favoring land policy contributed a sizable part of it

Final Conclusion

- QSGE Model is interesting and important
- The crucial part is how to combine data with model
- This is the key to all structural methods
- We have introduced it in details using one paper as the example
- Let's continue our journey with more contents in the following weeks!

References

Ahlfeldt, Gabriel M, Stephen J Redding, Daniel M Sturm, and Nikolaus Wolf. 2015. "The Economics of Density: Evidence from the Berlin Wall." *Econometrica* 83 (6):2127–2189.

Allen, Treb and Costas Arkolakis. 2014. "Trade and the Topography of the Spatial Economy." *The Quarterly Journal of Economics* 129 (3):1085–1140.

Baum-Snow, Nathaniel and Ronni Pavan. 2012. "Understanding the City Size Wage Gap." *The Review of Economic Studies* 79 (1):88–127.

Ciccone, Antonio and Robert Hall. 1996. "Productivity and the Density of Economic Activity." *American Economic Review* 86 (1):54–70.

Combes, Pierre-Philippe and Laurent Gobillon. 2015. "The Empirics of Agglomeration Economies." In *Handbook of Regional and Urban Economics*, vol. 5. Elsevier, 247–348.

Dingel, Jonathan I and Felix Tintelnot. 2020. "Spatial Economics for Granular Settings." Tech. rep., National Bureau of Economic Research.

Eaton, Jonathan and Samuel Kortum. 2002. "Technology, Geography, and Trade." *Econometrica* 70 (5):1741–1779.

Lu, Ming, Yuta Suzuki, Xican Xi, Le Xu, and Yuejun Zhong. 2024. "Pro-Manufacturing Land Policies of Competing Local Governments: A Quantitative Analysis of China." *Working Paper* .

Neumark, David and Helen Simpson. 2015. "Place-based Policies." In *Handbook of Regional and Urban Economics*, vol. 5. Elsevier, 1197–1287.

Redding, Stephen J and Esteban Rossi-Hansberg. 2017. "Quantitative Spatial Economics." *Annual Review of Economics* 9:21–58.