Quantitative International Economics

Heitor S. Pellegrina

University of Notre Dame

Weeks 1 and 2 - The Basic Structure of Quantitative Trade Models - Part 2
Spring 2025

Outline

- EK model with 1 sector and no intermediate inputs
  - Solving the model in levels (DONE)
  - Policy counterfactuals (DONE)
  - Solving the model in changes (hat-algebra) (DONE)
  - Inverting the model
  - Estimating trade costs
  - Estimating trade elasticities

Model Inversion

From Previous Lecture

- So far, we have learned how to solve the EK model **given parameters**
  - What do we need to simulate the model in levels?
    - Productivity shifters T_i
    - Trade costs d_{ni}
    - Population N_i (observable)
    - Trade elasticity θ
  - What do we need to simulate the model in differences?
    - Baseline trade shares π_{ni}^0
    - Trade elasticity θ
- This set of lectures
  - How to recover T_i, d_{ni}, and θ
- I will not cover all possible methods, but will give references

Calibrating T_i with Model Inversion

- **<u>GOAL</u>**: Calibrate T_n given d_{ni} and data on X_n and N_n
  - d_{ni} could come from a first step in which we estimate trade costs
  - Set T_n so that gross output in the model matches the data

- Allen and Arkolakis (2014) prove that T_n is identified "up-to-scale", which means that we need a normalization of T_n. To see this point, notice that identification of T_n comes from the following equation

$$X_n = \sum_i \frac{T_n (w_n d_{ni})^{-\theta}}{\sum_s T_s (w_s d_{is})^{-\theta}} X_i$$

  - If T_n is too low, then county n's gross output will be too low relative to the data
  - X_n is homogeneous of degree zero in T_n (what does that mean?)
  - *Relative* gross output identifies *relative* T_n

Algorithm

- Step 1: Guess T_n^g (pick a normalization)
- Step 2: Solve the model given T_n^g
- Step 3: Obtain model-implied gross output X_n^{model} (pick a normalization)
- Step 4: Check gross output
  - If $\max_n\{|X_n^{data} - X_n^{model}|\} < \epsilon$, then stop algorithm
  - If $\max_n\{|X_n^{data} - X_n^{model}|\} > \epsilon$, then update T_n^g and go back to step 2
    - Increase T_n^g if $X_n^{data} - X_n^{model} > 0$
    - Decrease T_n^g if $X_n^{data} - X_n^{model} < 0$
    - Remember to re-normalize T_n^g before starting the algorithm again

Estimation of Trade Costs

Estimation of Trade Costs

- A (very) large literature on the estimation of trade costs
- We will discuss two approaches, each one of them is based on

  1. Price data
  2. Trade data

- The goal here is not to be exhaustive. I'll provide what I believe are the main intuitions and approaches used in the literature

Estimation of Trade Costs based on Prices

- Fundamental challenge
  - If we look at the price gap between two regions. Does it tell us something about the trade costs?

Estimation of Trade Costs based on Prices

- Define d_{in} as the trade cost to take a good from i to n
  - For example, if the good costs 2 at i and d_{in} is 1.5, then price at n is 3
- If a product is not exported from i to n, then it must mean that producers at i found it not profitable to sell it to n, therefore
$$p_n < p_i d_{in}$$
Similarly, we get
$$p_i < p_n d_{ni}$$
- That puts a bound on price differentials
$$\frac{1}{d_{ni}} < \frac{p_n}{p_i} < d_{in}$$
- Price ratios are bounded above and below by trade costs. If we only look at price ratios, we will understate trade costs!

Estimation of Trade Costs based on Prices

- Problem 1
  - We do not observe whether regions are actually trading
- Problem 2
  - Even if they are, we might not know if the product is exactly the same
- Problem 3
  - Even if products are exactly the same, price ratios could also reflect regional differences in markups

Estimation of Trade Costs based on Prices

- Solution 1: EK (2002) approach

$$\widehat{d}_{in} = \max_{\omega \in \Omega} \frac{p_n(\omega)}{p_i(\omega)}$$

  - where ω is a good from country n and country i

- Solution 2: Donaldson (2018) approach
  - Pick an homogeneous variety of salt $\Rightarrow$ price differentials do not reflect quality and perfect competition is reasonable assumption for salt
  - Data shows origin and destination $\Rightarrow$ location of production is unique!
  - Make the following parametric assumption on d^v_{in} and take logs

$$p^v_n = p^v_i \underbrace{exp\left(\beta \text{dist}_{in}\right) exp\left(\epsilon^v_{in}\right)}_{d^v_{in}}$$

$$\log \frac{p^v_n}{p^v_i} = \beta \text{dist}_{in} + \epsilon^v_{in}$$

  - Simulate model using fitted value $\widehat{d}_{ni}$ as trade cost, for example

$$\widehat{d}_{ni} = \widehat{\beta} \text{dist}_{in}$$

Estimation of Trade Costs based on Prices

- Previous approaches are based on perfectly competitive markets
- The following papers address the issue of market power when estimating trade costs
  - See Atkin and Donaldson (2015)
  - See Dominguez-Iino (2024)
- A key insight here is that pass-throughs contain information about market power that can be used as additional information to recover trade costs

Estimation of Trade Costs based on Trade Data

- Trade data contains information about trade costs. To see why, imagine a world with two countries and the following data structure.

$$\begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}$$

where X_{in} are the sales of country i to country n. We also define $X_1 = X_{11} + X_{12}$ and $X_2 = X_{22} + X_{21}$.

- The corresponding equations in the EK model are

$$\begin{bmatrix} \frac{T_1 (w_1 d_{11})^{-\theta}}{T_1 (w_1 d_{11})^{-\theta} + T_2 (w_2 d_{21})^{-\theta}} X_1 & \frac{T_1 (w_1 d_{12})^{-\theta}}{T_1 (w_1 d_{12})^{-\theta} + T_2 (w_2 d_{22})^{-\theta}} X_2 \\ \frac{T_2 (w_2 d_{21})^{-\theta}}{T_2 (w_2 d_{21})^{-\theta} + T_1 (w_1 d_{11})^{-\theta}} X_1 & \frac{T_2 (w_2 d_{22})^{-\theta}}{T_2 (w_2 d_{22})^{-\theta} + T_1 (w_1 d_{12})^{-\theta}} X_2 \end{bmatrix}$$

Estimation of Trade Costs based on Trade Data

- Say we already know θ, then we have 6 parameters $T_1, T_2, d_{11}, d_{12}, d_{21}, d_{22}$, to rationalize 6 observations $X_{11}, X_{12}, X_{21}, X_{22}, X_1$ and X_2.
  - However, we already have 2 restrictions $\Rightarrow X_1 = X_{11} + X_{12}$ and $X_2 = X_{22} + X_{21}$
  - Recall that we can only recover *relative* gross output (we have to pick a numeraire for the economy)
  - We are therefore down to 6 observations and 3 restrictions, with 6 parameters.
  - Normalize $d_{11} = d_{22} = 1$, now we have 4 parameters.
  - Normalize $T_1 = 1$, now we have 3 parameters and 3 degrees of freedom.
- We can perfectly rationalize the data if we pick T_2, d_{12} and d_{21}. This is for the 2 by 2 case. Here, we are perfectly identified and these 3 parameters fully saturate the data. For $N > 2$ countries, however, we will be under-identified, so that there is more than one combination of TFPs and trade costs that can perfectly rationalize the data. How should we proceed with N countries?

Estimation of Trade Costs based on Trade Data

- Estimate

$$\log X_{ni} = \alpha_n + \gamma_i + \underbrace{\epsilon_{ni}}_{\log\left(\tilde{d}_{ni}^{-\theta}\right)}$$

  recover residuals from the regression and obtain

$$\tilde{d}_{ni} = \exp\left(\widehat{\epsilon_{ni}}\right)^{-\frac{1}{\theta}}$$

- Ensure that $\widehat{d}_{ii} = 1$ by constructing

$$\widehat{d}_{ni} = \frac{\tilde{d}_{ni}}{\tilde{d}_{ii}}$$

- Use $\widehat{d}_{ni}$ as your estimates of trade costs to simulate the model

Estimation of Trade Costs based on Trade Data

- Why don't we need price data here? Let's take the Armington trade expression

$$X_{in} = \frac{(p_i d_{in})^{1-\sigma}}{P_n^{1-\sigma}} E_n$$

  - p_i is the price at origin country i, P_n the price index, X_{in} sales of country i to n and E_n total expenditure

- Take logs

$$\log X_{in} = \underbrace{\log p_i^{1-\sigma}}_{\text{origin FE}} + \underbrace{\log \frac{E_n}{P_n^{1-\sigma}}}_{\text{destination DE}} + \underbrace{\log d_{in}^{1-\sigma}}_{\text{trade cost}}$$

- Price is fully absorbed by the origin FE
  - After taking logs, it is additive to trade costs!
  - If I increase price by 1 percent, **every country** will decrease their imports by $1-\sigma$ percent

Estimation of Trade Costs based on Trade Data

- Head and Ries approach

$$
\begin{aligned}
\frac{X_{21}}{X_{22}} \times \frac{X_{12}}{X_{11}} &= \frac{T_2\left(w_2 d_{21}\right)^{-\theta}}{T_2\left(w_2 d_{22}\right)^{-\theta}} \times \frac{T_1\left(w_1 d_{12}\right)^{-\theta}}{T_1\left(w_1 d_{11}\right)^{-\theta}} \\
&= \left(d_{21}\right)^{-\theta} \times \left(d_{12}\right)^{-\theta}
\end{aligned}
$$

- Assume trade costs are symmetric, then

$$
\widehat{d}_{12} = \left(\frac{X_{21}}{X_{22}} \times \frac{X_{12}}{X_{11}}\right)^{-\frac{1}{2\theta}}
$$

- Useful statistic for counterfactuals
  - Farrokhi and Pellegrina (2023) recover the symmetric component between 1980 and 2010
  - Jacks, Meissner and Novy (2011) study trade costs changes between 1870 and 2000

Estimation of Trade Elasticity

Estimation of Trade Elasticity

- Approach 1: Once with estimates of trade costs, we can estimate

$$\log X_{ni} = \alpha_n + \gamma_i - \theta \log \widehat{d_{ni}} + \epsilon_{ni}$$

  via OLS and use $\widehat{\theta}^{OLS}$ as the trade elasticity.
  - Donaldson (2018), Pellegrina (2022), EK (2002)

- Approach 2: Double ratio estimator. Define

$$d_{in} = \tau_{in} \kappa_{in}^{sym} \mu_n \gamma_i \epsilon_{in}$$

  - where κ_{in}^{sym} is a symmetric component of trade cost, μ_n is a destination specific cost, γ_i is an origin specific cost, and τ_{in} is 1 plus the tariff
  - Take the triple ratio

$$\log \frac{X_{ni}}{X_{in}} \frac{X_{ih}}{X_{hi}} \frac{X_{hn}}{X_{nh}} = -\theta \log \frac{\tau_{ni}}{\tau_{in}} \frac{\tau_{ih}}{\tau_{hi}} \frac{\tau_{hn}}{\tau_{nh}} + \tilde{e}$$

  - It removes κ_{in}^{sym}, μ_n and γ_i. Here, the tariffs τ_{ni} are observable, so we can simply estimate the equation above by OLS

Estimation of Trade Elasticity

- What if trade costs or tariffs are measured with error?
  - If we understate their variance, we would overstate the trade elasticity
- Literature finds a large range of values
  - Eaton and Kortum (2002) finds 4.17 to 9.6 (single sector)
  - Simonovska and Waugh (2014) $\rightarrow \approx$ 4.17 (single sector)
  - Caliendo and Parro (2015) $\rightarrow$0.37 to 51.08 (multiple sector)
- Many papers pick values from the literature

Conclusion

Recent Developments on these Tools

- Eaton and Kortum (2002) microfoundations
  - EK (2002) formulation can be microfounded with a model in which firms draw ideas and there is a probability that the ideas are better than previous ones – see the excellent JEP from EK (2012). A few papers have explored these microfoundations to study innovation and diffusion.
  - Lind and Ramondo (2022), Buera and Ozberfield (2019)
- Non-parametric approach
  - Large scale CGE models contain thousands of parameters and regions to simulate counterfatuals. EK has 1, but it comes at a cost. Can we simulate counterfactuals with fewer parametric assumptions?
  - Adao, Costinot and Donaldson (2016) $\Rightarrow$ It turns out that the hat-algebra approach does not depend on parametric assumptions such as CES
- Can we test these models?
  - Because we saturate the data, it is challenging to test the model.
  - Adao, Costinot and Donaldson (2024) $\Rightarrow$ Develop new tests for these types of model.

Next class

- Welfare gains from trade
- Multi-sectoral models and input-output structures
- Sectoral comparative advantage