Trade, Competition and Productivity

Gianmarco Ottaviano

LSE and U Bologna

22 June 2016
In the last couple of decades the field of international trade has become increasingly quantitative.

This is due to two major developments driven by easier accessibility of individual datasets and higher computing power:

1. Econometric works to study *ex post* the implications of firms’ and workers’ heterogeneity for the sources, the patterns and the gains from trade.

2. Calibration and simulation of statistical models to investigate *ex ante* the (welfare) implications of counterfactual scenarios for which data are necessarily unavailable (e.g. Brexit).

For lack of better name, call the latter models ”*new quantitative trade models*”:

- We should care about the detail of micro reallocations only if this changes our understanding of the aggregate gains from trade.
New Quantitative Trade Models (NQTM)

The idea of using mathematical or statistical models to simulate the effects of counterfactual scenarios has a long tradition (Baldwin and Venables, 1995).

In particular, ‘Computable general equilibrium’ (CGE) models remain a cornerstone of trade policy evaluation.

To this tradition NQTM contribute:

- A tighter connection between theory and data thanks to more appealing micro-theoretical foundations.
- A more careful estimation of the structural parameters necessary for counterfactual analysis.

The trailblazer NQTM is arguably the statistical model proposed and structurally estimated by Eaton and Kortum (2002) to quantify the effects of counterfactuals on trade liberalization and technological progress in 19 OECD countries.

However, by assuming perfect competition, the Eaton-Kortum model does not speak directly to the parallel research line based on individual heterogeneity, of which the main theoretical reference is Melitz (2003).

Introducing heterogeneous firms in Krugman (1980), Melitz (2003) provides a theoretical framework consistent with several stylized facts highlighted by the analysis of firm-level datasets.

But its initial applications did not include counterfactual simulations.
Early attempts at bridging the two lines of research can be found in Bernard, Eaton, Jensen and Kortum (2003) and Del Gatto, Mion and Ottaviano (2006) – see also Di Mauro, Ottaviano and Taglioni (2009)

Both papers apply the standard macroeconomic methodology of ‘calibration, validation and simulation’ for counterfactual analysis:

1. **Calibration** requires the values of the theoretical parameters to be set such that the model matches *some key moments* of the data
2. **Validation** requires the calibrated model to be able to match *other moments* of the data different from those used for calibrating
3. **Simulation** of counterfactual scenarios can be ‘reasonably’ performed only if the calibrated model passes the validation checks
In several respects, Eaton and Kortum (2002) and many variations of Melitz (2003) belong to the same family of models. All models in this family share the same predicted ‘gains from trade’ (defined as welfare with trade relative to welfare with autarky), conditional on the changes in two aggregate statistics:

- The observed share of domestic expenditure and an estimate of the trade elasticity

These models share four primitive assumptions: (a) Dixit-Stiglitz preferences; (b) one factor of production; (c) linear cost functions; (d) perfect or monopolistic competition.

They also share three common macro-level restrictions: (A) trade is balanced; (B) aggregate profits are a constant share of aggregate revenues; (C) the import demand system exhibits constant elasticity of substitution (CES).
As this set of assumptions is extremely restrictive, the finding by ACR could be dismissed as some sort of ‘impossibility theorem’ with very limited practical relevance.

What makes their finding interesting is that some of the most popular trade models do satisfy those restrictive assumptions such as Armington (1969), Krugman (1980), Eaton and Kortum (2002) and several variations of Melitz (2003).

In this respect, the main contribution of ACR is to theoretically define the main class of state-of-the-art NQTMs.
Do ACR show that the micro details of NQTMs are irrelevant for the quantification of the aggregate welfare effects of counterfactual shocks?

Not really:

- Different models of the ACR family often produce very different predictions for the same counterfactuals (Costinot and Rodriguez-Clare, 2014)

Current debate has mostly focused on first moment of aggregate welfare changes:

- How much countries gain/lose

Another interesting way to check robustness is to look at higher moments:

- How gains/losses are distributed across countries
A Simple Example from CR

- Welfare losses of a 40% increase in worldwide import tariffs for 20 European countries
- Let’s look at correlations of losses across countries generated by different NQTMs

While considering intermediates mostly affects the average losses, the choice of market structure also affects the cross-country distribution of losses.
The predictions of NQTMs on the average welfare effects seem to be quite sensitive to considering or not intermediate goods

⇒ More attention to I-O linkages and GVCs, seller-buyer relations

The predictions of NQTMs on the distribution of welfare effects seem to be very sensitive to the choice of market structure

⇒ More attention to the actual market structures that characterize different sectors

NQTMs are mostly silent on the ‘dynamic’ effects that policy intervention may have on economic growth

⇒ More attention to competition, innovation and technology adoption
Validation has increasingly gone missing in NQTMs (‘exactly identified’ instead of ‘overidentified’ models)

- Micro data are a mine of additional moments for validation

But even the ‘four primitive assumptions’ have implications that are clearly at odds with key features of firm-level data

- More attention to demand characteristics, markup behavior, passthrough from input prices to output price, intensive margin reallocations

To sum up:

- Simulated macro models are needed to quantify the aggregate implications of counterfactual scenarios for which data are by definition not available
- Micro data can be used to discipline the structure of macro models and to validate their calibration
- Micro and macro data/models are complementary
Beyond ACR/CR?

- Models in the ACR family were developed under those restrictive assumptions to get as far as possible with analytical solutions
  - This loses relevance when in the end NQTMs are anyway solved numerically
- Which assumptions could be dropped more easily?
  - (a) Dixit-Stiglitz preferences? YES
  - (b) one factor of production? YES
  - (c) linear cost functions? YES
  - (d) perfect or monopolistic competition
- What about (d) perfect or monopolistic competition?
  - General equilibrium models of oligopoly with heterogeneous firms and asymmetric countries are plagued by problems of existence, uniqueness and stability of equilibria
  - This is a NO-NO for numerical analysis!
When one is interested in general equilibrium, monopolistic competition with adequate demand properties is able to "mimick" crucial oligopolistic outcomes while avoiding problems of existence, uniqueness and stability.


In particular, Mayer, Melitz and Ottaviano (2015) shows how dispensing with CES demand allows monopolistic competition to:

- Capture *ex post* the observed behavior of markup, passthrough and intensive margin reallocations
- Characterize the "right properties" of demand to be used for *ex ante* counterfactual analysis through NQTMs
- Show that neglecting these properties prevents identifying a key channel through which *trade shocks affect aggregate productivity by altering firms’ competitive environment*
The Model: Utility and Profit Maximization

- \( L^c \) identical consumers with individual expenditure normalized to 1:
  - In partial equilibrium (PE) labor supply is perfectly elastic at wage 1
  - In general equilibrium (GE) labor supply is \( L^w = L^c \) and wage equals 1 by choice of units

- \( M \) horizontally differentiated products indexed \( i \in [0, M] \)

- Utility maximization problem:

\[
\max_{x_i \geq 0} \int_0^M u(x_i) \, di \quad \text{s.t.} \quad \int_0^M p_i x_i \, di = 1
\]

- The FOC determine the inverse demand function:

\[
p_i = \frac{u'(x_i)}{\lambda}, \quad \text{with} \quad \lambda = \int_0^M u'(x_i) x_i \, di,
\]

where \( \lambda > 0 \) is the marginal utility of income.
Production has linear cost with marginal cost \( v \) varying across products

Product-level profit maximization problem:

\[
\max_{q_i \geq 0} \pi(q_i) = p_iq_i - vq_i - f
\]

The optimal level of output \( q_v = x_vL^c \) satisfies the first order condition:

\[
u'(x_v) + u''(x_v)x_v = \lambda v,
\]

where \( r(x_v) = \phi(x_v)/\lambda \) with \( \phi(x_v) \equiv u'(x_v) + u''(x_v)x_v \) is the marginal revenue associated with a given variety.
Necessary and sufficient conditions for these max problems are:

- **(A1)** $u(x_i) \geq 0$ with $u(0) = 0$; $u'(x_i) > 0$ and $u''(x_i) < 0$ for $x_i \geq 0$
- **(A2)** elasticity of inverse demand $\varepsilon_p(x_v) < 1$
- **(A3)** elasticity of marginal revenue $\varepsilon_r(x_v) > 0$

IFF. (A1), (A2) and (A3) hold there exists a unique output and price level for all varieties $x_v > 0$ and $p(x_v) > 0$, and for any given lagrangean multiplier $\lambda > 0$
De Loecker, Goldberg, Pavcnik and Khandelwal (NBER 2012) find that lower costs are associated with larger markups so that cost advantages are not fully passed through to prices.

\[ \epsilon_p'(x_v) > 0 \] 

Berman, Martin and Mayer (QJE 2012) find that high-performance firms react to a real exchange depreciation by increasing significantly more their markup.

\[ \frac{\epsilon_p'(x_v)x_v}{\epsilon_p(x_v)} < \frac{\epsilon_r'(x_v)x_v}{\epsilon_r(x_v)} \]

Empirically lower cost firms/products are associated with larger employment.

\[ \epsilon_r(x_v) < 1 \]
The Model: Endogenous Heterogeneity

- Products are supplied by firms that may be single- or multi-product.
- Market structure is monopolistically competitive:
  - Each product is supplied by only one firm and each firm supplies a countable number of the continuum of products.
- Technology exhibits increasing returns to scale:
  - Fixed cost $f$ is the same for all products, marginal cost $v$ differs across them.
  - For a given firm, products are indexed in increasing order $m$ of marginal cost from a ‘core product’ indexed by $m = 0$.
  - $M(c)$ denotes the number of products supplied by a firm with core marginal cost $c$ and $v(m,c)$ to denote the marginal cost of its $m^{th}$ product such that $v(m,c) = cz(m)$ with $z(0) = 1$ and $z'(m) > 0$.
- Firm entry incurs a sunk cost $f^e$. Only after this cost is incurred, entrants randomly draw their marginal cost levels for their core products from a common continuous differentiable distribution defined over the support $[0, \infty)$, with density $\gamma(c)$ and cumulative density $\Gamma(c)$. 
The Model: Equilibrium Conditions

- **Zero Cutoff Profit (ZCP)**

\[ \pi^* (\hat{c}, \lambda) L^c = f \]

- **Free Entry Condition (FE):**

\[ \sum_{m=0}^{\infty} \left[ \int_{0}^{\hat{c}/z(m)} \left[ \pi^* (cz(m), \lambda) L^c - f \right] \gamma(c) dc \right] = f^e \]

- **Budget Constraint (BC)**

\[ N_e \left( \sum_{m=0}^{\infty} \int_{0}^{\hat{c}/z(m)} p^* (cz(m), \lambda) x^* (cz(m), \lambda) \gamma(c) dc \right) = 1 \]

where \( N_e \) is the number of entrants and * signifies profit, price and quantity that solve the max problems.
Define a ‘positive demand shock’ as more consumers ($dL^c > 0$)

These comparative statics hold in both PE and GE:

- **Lemma 1.** A positive demand shock increases the marginal utility of income.
- **Proposition 1 - Extensive margin adjustment.** (B1) is necessary and sufficient for a positive demand shock to reduce the cost cutoff, thus increasing multi-product firm productivity through extensive margin adjustment. (B1) is also necessary and sufficient for a positive demand shock to increase (decrease) profit for low (high) cost products.
- **Proposition 2 - Intensive margin adjustment.** (B1) and (B2) are sufficient for a positive demand shock to reallocate output and revenue from higher to lower cost products. As long as (B3) holds, assumptions (B1) and (B2) are also sufficient for a positive demand shock to reallocate employment from higher to lower cost products, thus increasing multi-product firm productivity through intensive margin adjustment.
The model predicts that a positive demand shock increases multi-product firm productivity by shifting resources from high to low cost products (higher ‘skewness’).

In particular, the chain of causation is:

Demand shock → Skewness → Productivity

This can be tested exploiting detailed export data looking at the impact of demand shocks in export destinations on the skewness of export sales and productivity of multi-product firms.