Modeling Firm Heterogeneity in International Trade: Do Structural Effects Matter?

Roberto Roson*   Kazuhiko Oyamada†

Abstract

This paper analyzes the qualitative properties of a multisectoral, multiregional computable general equilibrium model where some industries include heterogeneous firms as in Melitz (2003). The model, formulated according to Roson and Oyamada (2014), adds endogenous productivity effects to a standard Walras-Ricardian framework. We argue that the inclusion of such effects changes the magnitude and distribution of welfare benefits obtainable by reductions in trade barriers, due to of comparative advantages.

We illustrate the point through a numerical example, in which alternative model formulations are assessed. A standard neoclassic GE model, a basic Melitz model and a hybrid model are then compared. The three model versions are all calibrated with the same data set and an identical simulation experiment (a 50% reduction of transport costs between two regions) is carried out in the three cases. The results show that the hybrid model displays the largest welfare gains, as it combines Ricardian comparative advantages with Melitz average productivity improvements. However, they also show that new effects, not present in the original Ricardo and Melitz frameworks, are at a work.

Keywords: Computable General Equilibrium Models, Melitz, Firm Heterogeneity, International Trade.

JEL CODES: C63, C68, D51, D58, F12, L11.

1   Introduction

A substantial body of literature has originated since the seminal work by Melitz (2003). One reason why such work has attracted so much interest is due to

---

*Dept. of Economics, Ca' Foscari University Venice and IEFE Bocconi University, Milan, Italy. E-mail: roson@unive.it.

†Institute of Developing Economies, Japan External Trade Organization. E-mail: kazuhiko_oyamada@ide.go.jp.

the fact that Melitz provides a third potential source of gains from trade, complementing the Ricardian comparative advantages and the intra-industry economies of scale analyzed by Krugman (1980).

In the Melitz model, firms within an industry are heterogeneous, in two dimensions. First, they produce horizontally differentiated goods, which allow them to benefit from some degree of market power, typical of monopolistic competition. Second, they vary in terms of productivity, that is in the amount of primary resource (labor) needed to produce the goods. Economies of scale are generated at the firm level because of the existence of two types of fixed costs: general ("headquarter") costs and trade link ("foreign subsidiary") costs.

Each firm enters the market by paying headquarter entry costs and subsequently "discovers" its productivity level. If the productivity level is too low, expected profits are negative and the firm is better exiting, thereby losing the entry sunk costs. Alternatively, the firm operates in the domestic market and possibly in some foreign markets. It is shown that those firms with the highest productivity levels are also larger than average and export abroad (the more so the more efficient they are), which is consistent with empirical facts.

A reduction in trade barriers (lower transport costs, import tariffs, etc.) makes it profitable for more firms to export. The increased competition drives smaller, relatively inefficient domestic firms out of the market. Welfare gains are obtained because of the increase in average productivity and in the number of available varieties.

A number of authors, (e.g., Zhai (2008); Dixon, Jerie and Rimmer (2013); Itakura and Oyamada (2013); Roson and Oyamada (2014)) have recently proposed alternative approaches and techniques to embody Melitz characteristics into computable general equilibrium models, which are the workhorses of applied trade economic analysis. We consider here Roson and Oyamada (2014), who propose a Melitz-like CGE model, differing from the basic Melitz model in several aspects:

- multiple primary resources (e.g., labor and capital) are considered;

- productivity is defined in terms of a scaling factor inside some cost functions;

- there can be different technologies for transportation, headquarter and trade fixed costs;

- conventional, perfectly competitive industries (e.g., services) can coexist alongside industries differentiated à la Melitz (e.g., manufacturing);

- there can be intermediate factors and input-output linkages among industries;

- transportation and fixed costs generate a demand for services;

- the establishment of a subsidiary branch abroad, necessary to start exporting, generates a demand for services in the host country.

Considering this extended CGE-Melitz setting, the question we want to address here is the following one: does a Melitz trade model with multiple factors and industries just provide more detailed results, or is the qualitative behavior of the model going to be different? Intuition may suggest that, once several primary factors and industries are considered, comparative advantages, relative endowments and terms of trade effects may operate alongside the endogenous productivity mechanism.

We investigate the issue by conducting a simulation experiment with three alternative model formulations: a standard neoclassic GE model, a basic Melitz model and a hybrid model based on Roson and Oyamada (2014). To make the results comparable, parameters for the different models are calibrated with the same (fictitious) data and an identical simulation experiment, namely a 50% reduction of transport costs between two regions, is carried out in the three cases.[1] The results indeed show that the hybrid model exhibits the largest welfare gains, as it combines Ricardian comparative advantages with Melitz average productivity improvements. However, they also highlight new effects, not present in the original Ricardo and Melitz frameworks.

2 A Numerical Simulation Example

Let us consider a fictitious economy with two regions (A and B). Each region has an identical endowment of labor and capital, and two industries: manufacturing and services. Technology is the same in the two regions. Furthermore, services are needed to move (manufactured) goods from one region to the other one, but the unit service requirement is the same in both directions, which makes the two regions perfectly symmetric, initially at least.

We simulate the effects of halving transportation costs from A to B (not vice versa). This is, clearly, an efficiency improvement which could bring about welfare gains. However, the specific mechanisms which ultimately generate these gains actually differ, depending on the model formulation.

2.1 Trade Benefits in a Stylized Ricardian Model

In the simple Ricardian model we propose, all industries are perfectly competitive and prices equal production costs. The technology is Leontief, so there is no factor substitution and the price of industry j in region s is determined by:

$$p_s^j = \sum_i a_s^{ij} p_s^i + \sum_k a_s^{kj} w_s^k \qquad (1)$$

where a are input-output coefficients, w is the price of a primary resource (indexed by k). The final market price of traded goods include transport margins, given by the product between the unit transport services requirement t_{rs} and the price of services (in the origin region). On the basis of the standard

---

[1] GAMS software code and data can be dowloaded at http://venus.unive.it/roson/Soft.htm .

Armington assumption, manufactured goods produced in different regions are seen as imperfect substitutes, and aggregated into a CES composite when intermediate and final demand is computed.[2]

Labor and capital stocks are given. Primary resources are domestically mobile but cannot move from one region to the other. The price of labor in country A is taken as the numeraire and it is set to one.

Final demand is generated by a representative consumer, maximizing utility under budget constraint.[3] The utility is a simple linear logarithmic Cobb-Douglas function, and income is defined as the value of domestic primary resources. Under this specification, the expenditure shares $(0 \leq \varphi \leq 1)$ are fixed and the consumption level q of good j in region s can be expressed as:

$$q_s^j = \frac{\sum_k l_s^k w_s^k}{\widetilde{p}_s^j} \varphi_s^j \tag{2}$$

where l are primary resource endowments and $\widetilde{p}$ stands for final market price.

General equilibrium is achieved when supply equals demand in all markets (primary factors and produced goods/services in both regions).

Parameters for the illustrative model are obtained by calibration from a fictitious social accounting matrix. This means that, without changes in exogenous variables, the model replicates the values provided in the benchmark data set.

Table 1: Variations (%) in trade flows

	A	B
A	-4.9	14.8
B	9.4	-3.1

A counterfactual equilibrium is obtained when the unit transport coefficient t_{AB} is reduced by 50%. Table 1 displays the percentage changes in physical trade flows between the two regions. Not surprisingly, trade volume from A to B increases, because final market prices of goods produced in A and sold in B are now lower. Since a trade balance constraint is active, the higher competitiveness of the manufacturing industry in A must be compensated by a real devaluation occurring in country B, which fosters trade in the opposite direction, from B to A. Since primary resource stocks are unchanged, the higher external trade volume requires lower domestic trade, in both regions.

The smaller need for services to carry goods from A to B releases primary resources, which can then be devoted to production processes. Since services are labor intensive, the initial symmetry is broken and A gets a comparative advantage in labor intensive industries. This is reflected in the relative prices of primary factors, which move as shown in Table 2.

We see here the typical comparative advantage effect: labor gets relatively abundant and cheaper in country A, whereas capital gets relatively abundant

---

[2]The elasticity of substitution adopted for the simulation exercise is quite high: 4.
[3]Notice that the budget constraint implies trade balance.

Table 2: Variations (%) in the price of primary resources

	A	B
labor	-	64.6
capital	140	20.1

and cheaper in country B. In turn, this implies relatively cheaper services in A and relatively cheaper manufactured goods in B.

It is possible to compute changes in utility for the representative consumers in the two countries. There is a direct effect on consumption of (imported) manufactured goods in country B, raising welfare there by 1.8%. There is an indirect effect in region A, where consumption of manufactured goods decreases by -1.1%, but consumption of services increases by +10.7%, bringing about a net welfare gain of 0.8%. Therefore, the higher efficiency in transporting good from A to B benefits both countries, but the larger gains are obtained in B.

2.2 Trade Benefits in a Stylized Melitz Model

In our simplified version of the Melitz model, based on Dixon, Jerie and Rimmer (2013), there is a single industry and a single primary factor. As the model parameters are calibrated with the same data used for the Ricardian model, the primary factor can be regarded as a value added composite of labor and capital.

Each firm produces a different product and sets the price with a mark-up above marginal costs. Furthermore, each firm is characterized by a specific productivity parameter $\Phi_{\circ rs}$, so that the market price is determined by:

$$p_{\circ rs} = \left(\frac{w_r t_{rs}}{\Phi_{\circ rs}}\right)\left(\frac{\sigma}{\sigma - 1}\right) \tag{3}$$

where σ is a CES elasticity of substitution,[4] t_{rs} is a factor (>1) expressing "iceberg" transportation costs. In the destination market s, a CES price index can be built by considering all goods flowing into that market:

$$p_r = \left(\sum_r n_{rs} p_{\circ rs}^{1-\sigma}\right)^{\frac{1}{1-\sigma}} \tag{4}$$

where n_{rs} stands for the number of firms active in the link rs (a subset total firms n_r in the origin country) and $p_{\circ rs}$ can be interpreted as an "average" price. Because of the CES functional form, the demand for firm specific products $q_{\circ rs}$ is driven by aggregate demand in the destination market and relative prices:

$$q_{\circ rs} = q_s \left(\frac{p_s}{p_{\circ rs}}\right)^{\sigma} \tag{5}$$

---

[4]For our simulation exercise, we use the same elasticity value adopted in the Ricardian model (4).

Profits obtained by each firm active on the link rs are given by the difference between gross sale profits and fixed costs associated with the establishment of a foreign subsidiary in destination s, which requires F_{rs} units of labour. In addition to link-related fixed costs, each firm has general "headquarters" fixed costs (H_r labour units), to be paid by all firms, including those not active.[5] Like in a monopolistic competition setting, there is free entry in the industry, driving total expected profits to zero.

Income available to the representative consumer is the value of the primary resource, which is endogenously determined. Equilibrium in the system entails matching demand and supply for all differentiated goods, and between a given endowment of resource (l_r) and total resource demand. By expressing variables referring to the average firm in each trade link with a circle (°) the latter equilibrium condition can be stated as:

$$l_r = \sum_s \frac{n_{rs} q°_{rs} t_{rs}}{\Phi°_{rs}} + \sum_s n_{rs} F_{rs} + n_r H_r \tag{6}$$

In our two-regions case, the Φ parameters are distributed according to a Pareto law [$p(\Phi){=}\alpha\Phi^{-\alpha\text{-}1}$, $\Phi \geq 1$], where α is set to 4. In equilibrium, the set of all firms in a country is partitioned in three groups: non active firms, firms selling in the domestic market only, firms selling in both the domestic and foreign markets. Non active firms are those getting a too low productivity parameter Φ, below a given threshold. Exporters are firms getting a sufficiently high Φ, above another threshold. All thresholds are endogenously computed.

The simulation experiment considers again a 50% reduction in the t_{AB} parameter. Because of the reduction in transport costs, more firms from A can exports to B. Consumers in B spend more on imported items and less on domestic products. Nonetheless, a trade balance constraint is active, implying (as in the Ricardian model) a real devaluation, which in this case takes the form of -4% reduction in the relative price of the primary resource in country B.

Goods traded from A to B get cheaper because of the lower transport costs, whereas goods traded from B to A get also cheaper, because of the lower value of production factors in B. As shown in Table 3, more inter-region trade and less intra-region trade is the outcome. This results is somewhat similar to the one observed for the Ricardian model (Table 1), but the causal mechanism is different.

Table 3: Variations (%) in trade flows

	A	B
A	-3.2	29.1
B	21.6	-4.1

Welfare of the representative consumers in both countries improves because

[5]In the original Melitz (2003) formulation, firms incurring in losses can exit after a period. In the steady state, the inflow and outflow of firms in the industry is balanced.

of a higher purchasing power, but also because of more product varieties available. Utility in country A increases by about 0.8% and in country B by 1%. Even in this case, then, all countries benefit from the reduction in transport cost in only one direction. However, benefits are more equally distributed than in the Ricardian model.

2.3   Trade Benefits in a Stylized Hybrid Model

We repeat the simulation experiment with a model, whose structure is defined as in Roson and Oyamada (2014). The key characteristics of this alternative formulation are:

- Manufacturing is an industry composed of heterogenous firms à la Melitz. Services is a conventional industry, with perfect competition and constant returns to scale;

- Local services are needed to supply headquarters services (fixed costs), as well as to transport manufactured goods. The amount of services required for the different purposes is determined by technological requirements, and depend on firm productivity. The services needed to export in foreign markets, as trade link fixed costs, generate a demand for *foreign* resources;

- There are two primary factors (labor and capital) with given endowments. Manufacturing is capital intensive, services are labor intensive.

Furthermore, it assumed that: (a) intermediate factors are not substitutable among themselves (à la Leontief)[6]; (b) services are domestically produced and consumed. They are not inter-regionally traded[7].

Let us indicate with a_s^{ij} the input coefficients for intermediate inputs, that is the amount of factor goods produced by industry i, necessary to produce one unit of output in industry j located in s. There is an important difference here between services (s), which are an homogeneous industry, and manufacturing (m), which is a differentiated one. "Inputs" and "outputs" refer to physical quantities in homogeneous industries but, actually, to CES quantity composites in differentiated industries.

The demand for differentiated intermediate factors adds to final consumption demand to determine the overall regional demand for manufactured goods:

$$z_s^{ms} + z_s^{mm} + q_s^m = \left(\sum_r n_{rs} q_{\circ rs}^{(\sigma-1)/\sigma} \right)^{\sigma/(\sigma-1)} \tag{7}$$

where z_s^{ms} stands for intermediate demand for manufactured goods generated by services, and z_s^{mm} for intra-manufacturing intermediate demand. In particular:

---

[6]However, manufactured factors are differentiated and substitutable inside the CES aggregate.

[7]Nonetheless, foreign services are needed to establish subsidiary branches abroad.

$$z_s^{mm} = a_d^{mm} \left(\sum_r n_{sr} q_{°sr} / \Phi_{°sr} \right) \tag{8}$$

$$z_s^{ms} = a_s^{ms} x_s^s \tag{9}$$

where x_s^s is the output level of the services industry in s, given by:

$$
\begin{aligned}
x_s^s = q_s^s + a_d^{sm} &\left(\sum_r n_{sr} q_{°sr} / \Phi_{°sr} \right) + a_s^{ss} x_s^s + \\
&+ n_s H_s + \sum_r n_{rs} F_{rs} + \sum_r t_{sr} n_{sr} q_{°sr} / \Phi_{°sr}
\end{aligned} \tag{10}
$$

where q_s^s is the quantity of services directly consumed by households in region s. T_{sr}, F_{rs}, H_s express the amount of *services* needed to: (1) carry one unit of manufactured good from s to r[8], (2) establish a trade link rs[9], (3) start a business in region s. The demand for primary factors is given by:

$$l_r^k = \sum_s \frac{n_{rs} q_{°rs} a_r^{km}}{\Phi_{°rs}} + x_r^s a_r^{ks} \tag{11}$$

where a_r^{kj} stands for the amount of primary factor k used to produce one unit of output in industry j in region r.

Final consumption includes manufactured goods as well as services. Manufactured goods are differentiated goods produced by both domestic and foreign firms. Services are domestically produced and are homogeneous.

For both industries, final consumption levels are determined on the basis of utility maximization of the representative consumer, given the budget constraint:

$$\sum_k l_s^k w_s^k = \sum_j q_s^j p_s^j \tag{12}$$

The utility function adopted here is linear logarithmic (Cobb-Douglas), giving raise to demand function as in (2).

When transportation costs from A to B are lowered, the model generates endogenous productivity improvements, like in the Melitz model, but also changes in relative prices, which are typical of a Ricardian neoclassic framework. The pattern of trade flows varies as displayed in Table 4.

Even in this case foreign trade expands and domestic trade shrinks. However, when comparing Tables 3 and 4 one can notice that the trade expansion is smaller than in the Melitz model.

When transport costs are reduced, labor and capital resources are released, creating a comparative advantage in labor intensive productions for A. Imported manufactured goods becomes more convenient in B. However, when new firms

---

[8]Contrary to the basic Melitz setting, this parameter is no more a multiplicative factor greater than one.

[9]Notice that the demand for services is generated in the destination country.

Table 4: Variations (%) in trade flows

	A	B
A	-7.6	24.1
B	18.3	-5.8

start exporting from A to B, they need to establish foreign branches, which creates a demand for services in B. The first order effect is, therefore, the creation of additional demand for manufacturing in A and for services in B. This effect works to the opposite direction of comparative advantages, since changes in relative prices would make A relatively more efficient in services and B in manufacturing. This ultimately dampens the impact of trade creation. The phenomenon rests on some specific assumptions of the model, in particular the hypotheses that transportation from A to B requires services produced in A, whereas the establishment of a foreign subsidiary in B requires services produced in B.

Table 5: Variations (%) in the price of primary resources

	A	B
labor	-	18.1
capital	37.3	-2.8

Another way of considering this aspect is by analyzing changes in relative prices of primary factors (Table 5) which are not as wide as in the Ricardian model (Table 2).

Results in terms of welfare are a 1.1% utility gain for consumers in country A and a 2.4% utility gain for consumers in country B.

2.4   A Qualitative Comparison of Alternative Models

In order to compare the key characteristics of the three model versions, we focus here on the impact of the 50% reduction in transport costs on the welfare (utility) of the representative consumers in the two regions. Table 6 summarize the findings and provides some additional information.

Our results show that all models highlight welfare improvements in both countries. The Ricardo version displays larger gains than the Melitz one, accruing primarily to country B. The Hybrid version generates much larger gains than the Melitz one, because comparative advantages effects operate, in addition to productivity improvements.

One could wonder if welfare gains in the Hybrid model are just the sum of the gains detected in Melitz and Ricardo. Numbers in the last row of Table 6 indicates that this is not the case. Welfare gains in the hybrid version are smaller than the sum, and this is due to two reasons.

Table 6: A comparison of welfare impacts

	A	**B**
Melitz	0.76	0.99
Ricardo	0.80	1.77
Ricardo vs. Melitz	+5.2%	+79.4%
Hybrid	1.11	2.43
Hybrid vs. Melitz	+45%	+146.4%
M+R-H	0.46	0.33

First, both Melitz and Ricardo models account for a direct efficiency effect, that is the release of resources that were previously employed to carry goods from A to B, which would be equivalent to an increase in labor and capital stock for country A. Summing welfare gains of the two models, therefore, would implying double counting, overestimating gains especially for country A.

Second, the generation of demand for services in country B, as discussed in the previous section, lessen the gains due to changes in relative prices and comparative advantages. These gains are therefore smaller in the Hybrid model than in the Ricardo one. Disregarding this point would entail again an overestimation, but this time particularly for country B.

3   Conclusion

Is a computable general equilibrium model featuring firm heterogeneity in some industries qualitatively different from a standard CGE model? Our analysis suggests that the answer is yes. Benefits arising from a reduction of trade barriers, for example, tend to be significantly larger, when endogenous productivity effects à la Melitz are accounted for. Since the structure of a Melitz-CGE model is more complex than that of a standard CGE framework, and there are additional parameters to estimate, this outcome indicates that the development of such model may be worth the effort, especially when effects of trade liberalization are investigated.

Is a Melitz model with multiple industries and factors qualitatively different from the original Melitz setting, which considers a single factor and a single industry? Our results show that, with multiple factors and industries, effects related to changes in relative prices, terms of trade and comparative advantages are at work, ultimately boosting the welfare gains obtained through endogenous productivity, but also changing their regional distribution.

Finally, does a hybrid Melitz-CGE simply combine Ricardian comparative advantages and Melitz productivity enhancements? The answer is no. Although such a model does displays characteristics which are typical of the two settings, there are also some novel, peculiar features embedded in its structure.

Appendix

The three models used in this study are calibrated to an identical fictitious data set. The set for the Ricardian model consists of an input-output (I-O) table for the symmetric two regions, trade-flow tables at two price levels, free on board (FOB) and cost, insurance, and freight (CIF), and the elasticity of substitution between varieties. The assumed values are shown in Tables 7. In these tables, *a, b, m, s, c, l, k, e*, and *t* respectively denote: countries, manufacturing sector, services sector, final consumption, labor input, capital inuput, net exports, and trade cost. The value (5) appearing in the cell corresponding to net exports of services is the interregional shipping supply. The value (-5) appearing in the cell corresponding to the trade cost in the manufacturing sector is subtracted to make the row total to meet the sales of the manufactured commodity at the producer prices.

In turn, the set for the Melitz model consists of trade-flow tables at two price levels, fixed cost to establish a firm, fixed cost to activate a firm, and the Pareto shape parameter. These are shown in Tables 8. In the trade-flow tables, the flows of services are added to the intraregional trade part. The set for the Hybrid model consists of the information included in both Tables 7 and 8, without the trade-flow tables included in Tables 8.

The calibration procedure can be found in Roson and Oyamada (2014).

Table 7: Data for the Ricardian Model

Trade Flow (FOB)

	a	b
a	90	30
b	30	90

Trade Flow (CIF)

	a	b
a	90	35
b	35	90

Input-Output Table

	m	s	c	e	t (-)	total
m	25	20	80		-5	120
s	45	30	20	5		100
l	25	40				65
k	25	10				35
total	120	100				

Elasiticity of Substitution

4

Table 8: Data for the Melitz Model

Trade Flow (FOB)	a	b		Trade Flow (CIF)	a	b
a	185	30		a	185	35
b	30	185		b	35	185

Number of Firms	a	b		Number of Active Firms	a	b
	1	1		a	0.6	0.4
				b	0.4	0.6

Pareto Shape Parameter
4.25

References

Dixon, P. B., M. Jerie and M.T. Rimmer. 2013. "Deriving the Armington, Krugman and Melitz models of trade.".

Itakura, K. and K. Oyamada. 2013. Incorporating firm heterogeneity into the GTAP Model. In *XVI Conference on Global Economic Analysis*.

Krugman, P. 1980. "Scale Economies, Product Differentiation, and the Pattern of Trade." *American Economic Review* 70:950–959.

Melitz, M.J. 2003. "The impact of trade on intra-industry reallocations and aggregate industry productivity." *Econometrica* 71(3):1695–1725.

Roson, R. and K. Oyamada. 2014. Introducing Melitz-Style Firm Heterogeneity in CGE Models: Technical Aspects and Implications. Working Paper 63 IEFE Bocconi University, Milan.

Zhai, Fan. 2008. "Armington meets Melitz: Introducing firm heterogeneity in a global CGE model of trade." *Journal of Economic Integration* 23(3):575–604.