Introducing Melitz-Style Firm Heterogeneity in CGE Models: Technical Aspects and Implications

Roberto Roson* Kazuhiko Oyamada†

Abstract

This paper discusses which changes in the architecture of a standard CGE model are needed in order to introduce effects of trade and firm heterogeneity à la Melitz. Starting from a simple specification with partial equilibrium, one primary production factor and one industry, the framework is progressively enriched by including multiple factors, intermediate inputs, multiple industries (with a mixture of differentiated and non-differentiated products), and a real general equilibrium closure. Therefore, the model structure is gradually made similar to a full-fledged CGE. Calibration techniques are discussed, and a number of changes from the original Melitz's assumptions are also proposed. It is argued that the inclusion of industries with heterogeneous firms in a CGE framework does not simply make the Melitz model "operational", but allows accounting for structural effects that may significantly affect the nature, meaning and implications of the model results.

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

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

1 Introduction and Motivation

Computable General Equilibrium (CGE) models have become part of the standard toolkit in applied economics. As such, they have been employed to conduct numerical simulations in a wide range of fields: from fiscal policy to international trade, from agriculture and resource economics to climate change.

*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.

It is known that these models are based on the neoclassical Walrasian paradigm: perfect competition, market clearing for both primary resources and produced goods, utility and profit maximization, budget constraints. What is perhaps less known is that CGE models are not really limited by the these, sometimes restrictive, assumptions. Price rigidities, market power, externalities, dynamics can actually be accounted for in most CGE models, whenever this is deemed to be necessary.

In a companion paper, Roson (2006) discusses how imperfect competition can be introduced and modeled in a CGE. The main message is that there is no single way of modeling imperfect competition; in fact, the methodology used and the implicit assumptions adopted greatly affect the model results. This work continues along the same line, this time by considering what it takes to include intra-industry firms' heterogeneity as specified by Melitz (2003) in a CGE setting.

Melitz (2003) is a seminal paper, which has triggered substantial interest and originated a stream of theoretical and empirical works in international economics. In the Melitz model, average productivity is endogenously determined and made dependent on the degree of trade openness. As a consequence, the model provides a "third explanation" for the benefits of trade, in addition to [1] the Ricardian comparative advantages and to [2] the economies of scale (and variety) associated with enlargements of the market size (Krugman, 1980).

Comparative advantages are obviously captured by any multi-country CGE model, because of its neoclassical nature. Monopolistic competition à la Krugman can be easily introduced in a CGE setting as well, by making industrial TFP productivity endogenous, in the appropriate way. Unfortunately, accounting for firms' heterogeneity à la Melitz it is not that easy. This is because the Melitz model is a rather stylized one and, although some empirical studies have been based on it (e.g., Santos Silva and Tenreyro (2009)), the model should be better regarded as a theoretical paradigm, susceptible of empirical validation by means of econometric techniques, rather than as a model that can be directly implemented.

Nonetheless, a number of authors (Zhai (2008); Oyamada (2013); Dixon, Jerie and Rimmer (2013); Itakura and Oyamada (2013)) have recently tried to get Melitz right into a CGE framework. In my opinion, these efforts have only partially succeeded, because a number of ad-hoc adjustments have been done along the way, in order to introduce "Melitz equations" into the CGE system. These adjustments have, on one hand, retained some of the unrealistic hypotheses of the original theoretical model and, on the other hand, may have affected the general equilibrium closure, possibly bringing about violations of the Walras law. By contrast, Balistreri and Rutherford (2013) propose an iterative method, in which a conventional CGE model is interfaced with a partial equilibrium Melitz model. The latter is used to get average industry productivity parameters on the basis of output volumes and prices, which are obtained from the CGE model.

In this paper, we start from a version of the original Melitz model, and we progressively relax some of its simplifying assumptions. A few changes in

Melitz's assumptions are also proposed. The framework is progressively enriched by including multiple factors, intermediate inputs, multiple industries (with a mixture of differentiated and non-differentiated products), and a real general equilibrium closure. Therefore, the model structure is gradually made similar to a full-fledged CGE, which could then be calibrated and implemented.

The paper is organized as follows. The following section illustrates a version of the basic Melitz model, proposed by Dixon, Jerie and Rimmer (2013). Section 3 introduces, step by step, a number of "improvements" in the model structure, bringing it towards a CGE formulation. Calibration techniques are discussed in Section 4. Some final comments conclude.

2 A Reference Industry Model

Our starting point is Dixon, Jerie and Rimmer (2013) [from now on, DJR], who elaborated the theoretical model of trade introduced by Melitz (2003), in order to make it implementable in a computational setting. We summarize here the main equations of this framework, providing only a brief description of every equation and a discussion of its meaning. The interested reader may get more details from the two papers above.

There is an industry, in which several firms produce and sell (to geographically distinct markets) differentiated products. Each firm uses only one input (labour), and each one has a specific labour productivity parameter Φ. This expresses the units of output produced by one unit of labour in that firm.

The consumers have preferences determined by CES utility functions, with a parameter $\sigma > 1$ expressing the elasticity of substitution. Therefore, all goods (both domestic and imported) are regarded as imperfect substitutes.

Following DJR, we indicate with s the region of origin of trade flows, with d the destination market, and with the symbol $^\circ$ values referring to the "average" firm (in terms of productivity) among all those who are serving market d from region s.

The firms have some degree of market power and set their price on the basis of a mark-up rule over marginal cost, where the elasticity of substitution σ determines the price elasticity of individual demand functions. For the average, representative firm:

$$P_{^\circ sd} = \left(\frac{W_s T_{sd}}{\Phi_{^\circ sd}} \right) \left(\frac{\sigma}{\sigma - 1} \right) \tag{1}$$

where $T_{sd} > 1$ is a cost factor expressing "iceberg" transportation/trade costs in the sd link[1], and W_s is labour cost in region s.

In the destination market d, a CES price index is readily built by considering all goods flowing into that market:[2]

[1] In other words, $T_{sd} - 1$ are the units of product necessary to carry one unit of the produced good from s to d.

[2] DJR also include a parameter δ_{sd}, expressing preferences for the origin of the goods. This is omitted here for simplicity.

$$P_d = \left(\sum_s N_{sd} P_{\circ sd}^{1-\sigma} \right)^{\frac{1}{1-\sigma}} \tag{2}$$

where N_{sd} stands for the number of firms active in the link sd (a subset total firms N_s). The CES quantity index for sd can be computed on the basis of the output of the average firm:

$$Q_{sd} = N_{sd}^{\sigma/(\sigma-1)} Q_{\circ sd} \tag{3}$$

The demand for $Q_{\circ \mathrm{sd}}$ is, in turn, driven by aggregate demand in the destination market and relative prices:

$$Q_{\circ sd} = Q_d \left(\frac{P_d}{P_{\circ sd}} \right)^{\sigma} \tag{4}$$

Profits obtained by each firm active on the link sd are given by the difference between gross sale profits and fixed costs associated with the establishment of a foreign subsidiary in destination d, which requires F_{sd} units of labour. For the representative firm:

$$\Pi_{\circ sd} = \left(P_{\circ sd} - \frac{W_s T_{sd}}{\Phi_{\circ sd}} \right) Q_{\circ \mathrm{sd}} - F_{sd} W_s \tag{5}$$

In addition to link-related fixed costs, each firm has general "headquarters" fixed costs (H_s labour units). Like in a monopolistic competition setting, there is free entry in the industry in region s, driving total expected profits to zero:[3]

$$\sum_d N_{sd} \Pi_{\circ sd} - N_s H_s W_s = 0 \tag{6}$$

In the trade link sd, the marginal firm is the one having the minimum level of productivity Φ_{MINsd} compatible with non-negative profits on that link:

$$\Pi_{MINsd} = \left(P_{MINsd} - \frac{W_s T_{sd}}{\Phi_{MINsd}} \right) Q_{MINsd} - F_{sd} W_s = 0 \tag{7}$$

If the random productivity parameter has a Pareto distribution with parameter α [$p(\Phi) = \alpha \Phi^{-\alpha-1}, \ \Phi \geq 1$], it can be shown that the following relationships apply:

$$N_{sd} = N_s (\Phi_{MINsd})^{-\alpha} \tag{8}$$

$$\Phi_{\circ sd} = \beta \Phi_{MINsd} \tag{9}$$

$$\mathrm{Q}_{MINsd} = \mathrm{Q}_{\circ sd} / \beta^{\sigma} \tag{10}$$

[3] Profits are expected because each firm does not know its realization of the random variable Φ before entering the market. Timing is therefore as follows: (1) a enter/no enter decision in taken, (2) in case of entry, H_s units of labour are employed, (3) the random variable Φ is known, (4) the firms decides on which markets to operate, (5) prices/quantities are set.

where $\beta = \left(\frac{\alpha}{\alpha-(\sigma-1)}\right)^{1/(\sigma-1)}$.

Finally, total labour demand is given by:

$$L_s = \sum_d \frac{N_{sd} Q_{\circ sd} T_{sd}}{\Phi_{\circ sd}} + \sum_d N_{sd} F_{sd} + N_s H_s \tag{11}$$

The set of Equations (1)-(11) determines a system where, given cost, distribution and preference parameters, labour cost W_s and aggregate demand Q_d, the following endogenous variables can be computed:

1. The price $P_{\circ sd}$ of the average firm in link *sd*;
2. The price index in the destination market P_d;
3. The quantity $Q_{\circ sd}$ of the average firm in link *sd*;
4. The quantity index Q_{sd} in link *sd*;
5. The profit $\Pi_{\circ sd}$ of the average firm in link *sd*;
6. The number of active firms N_s in the home region s;
7. Demand for labour L_s in the home region s;
8. Number of firms in the *sd* link;
9. Productivity of the marginal firm in the *sd* link;
10. Productivity of the average firm in the *sd* link;
11. Quantity sold by the marginal firm in the *sd* link.

A reduction in trade costs T_{sd} increases average productivity, therefore efficiency, in both the origin and destination markets. This is a source of trade-related welfare gains, supplementing the conventional sources based on Ricardian comparative advantages, and market-size economies of scale (à la Krugman).

3 From Single-Industry Partial Equilibrium to Multi-Industry General Equilibrium

The model described in the previous section is a partial equilibrium variant of the Melitz (2003) framework. The original Melitz model differs, however, in two main ways.

First, Melitz considers the industry dynamics, with entry and exit of firms. Instead, the DJR version above focuses on the steady-state distributions of firm productivity, which amounts to assume that all firms remain (potentially) active forever. This is a necessary shortcut, which does not affect the qualitative properties of the model.

Second, the original model has a general equilibrium nature. In the following, we shall show how a general equilibrium closure could be easily applied to the system (1)-(11). Nevertheless, the model structure remains very much different from the typical CGE framework, because of a number of simplifications: only one industry is considered, there is only one factor, no intermediate inputs, no taxes, no explicit transportation costs. The original model should therefore be regarded as a theoretical construct, not as a model designed for applied economic analysis and numerical simulations. For this reason, we consider in the following what changes in the basic model structure could be introduced to make it more similar to a standard CGE model[4].

3.1 Single-Industry General Equilibrium

Before considering the general equilibrium closure, let us discuss a simple but useful change in the specification of Equations (5) and (11).

In the Melitz model, two classes of fixed costs are considered: fixed costs associated with starting the business (H_s), and fixed costs associated with operating in a trade link (F_{sd}). In both cases, costs imply the consumption of primary resources (labour) in the *home* country. However, if we think about what kind of costs would be involved, in the real world, with the establishment of a foreign subsidiary, we can notice that most of them would generate demand in the *destination* country: general and legal services, construction, training, etc. To account for this different localization of link-related fixed costs, it would suffice to replace Equation (5) with:

$$\Pi_{\circ sd} = \left(P_{\circ sd} - \frac{W_s T_{sd}}{\Phi_{\circ sd}} \right) Q_{\circ sd} - F_{sd} W_d \tag{12}$$

and (11) with:

$$L_s = \sum_d \frac{N_{sd} Q_{\circ sd} T_{sd}}{\Phi_{\circ sd}} + \sum_d N_{ds} F_{ds} + N_s H_s \tag{13}$$

In the general equilibrium specification, the price of primary resources (wages) would be endogenously determined, on the basis of a market equilibrium in the "labour" market. Labour supply would then be given as a fixed parameter ($L_s = \overline{L_s}$), or as a function. Furthermore, aggregate demand would be endogenously determined on the basis of a budget constraint for the representative consumer, in each region:

$$L_d W_d = Q_d P_d \tag{14}$$

In the modified system, the Walras law applies. This means that the equations are not independent and the price in a market has to be chosen as the

[4]Numerical examples of the four model variants discussed below have been implemented in GAMS, using the PATH solver. The GAMS codes can be freely downloaded at http://venus.unive.it/roson/Soft.htm.

numeraire. A natural way would be setting to unity the wage in *one* country, by replacing the corresponding Equation (13) with $W_s = 1$.

3.2 Multiple Factors, Different Cost Structures

In the Melitz model, only one factor (called "labour") is taken into account. The model can be easily modified, however, to consider several factors, for example capital and labour. To achieve this, all parameters, variables and equations referring to primary resources should be indexed (with index i here). In addition, the intensity of use of each factor in the production process may differ. The latter may be captured by an input-output parameter like A^i_s, expressing the amount of factor i in production processes occurring in region s, for a unitary value of the productivity parameter.

Equation (1) would then become:

$$P_{\circ sd} = \left(\frac{\sum_i A^i_s T^i_{sd} W^i_s}{\Phi_{\circ sd}} \right) \left(\frac{\sigma}{\sigma - 1} \right) \tag{15}$$

Parameters A^i_s and T^i_{sd} could themselves be made endogenous if general production functions for processing and transportation are considered[5], thereby allowing for cost minimization and factor substitution. The same reasoning applies to parameters F^i_{sd} and H^i_s in the following[6].

Equation (5) or (12) would be replaced by:

$$\Pi_{\circ sd} = \left(P_{\circ sd} - \frac{\sum_i A^i_s T^i_{sd} W^i_s}{\Phi_{\circ sd}} \right) Q_{\circ sd} - \sum_i W^i_d F^i_{sd} \tag{16}$$

and Equation (7) in the same way.

Equation (6) would become:

$$\sum_s N_{sd} \Pi_{\circ sd} - N_s \left(\sum_i H^i_s W^i_s \right) = 0 \tag{17}$$

Demand for primary factors would be given by:

$$L^i_s = \sum_d \frac{N_{sd} Q_{\circ sd} A^i_s T^i_{sd}}{\Phi_{\circ sd}} + \sum_d N_{ds} F^i_{ds} + N_s H^i_s \tag{18}$$

Finally, the budget constraint needs to be modified, to account for all primary factor endowments:

$$\sum_i L^i_d W^i_d = Q_d P_d \tag{19}$$

[5]By introducing coefficients T^i_{sd}, varying by factor, we depart from the iceberg transportation technology and we allow for general transportation cost structures.

[6]This means that different technologies are allowed for different fixed costs and different locations.

3.3 Multiple Industries, No Intermediate Inputs

So far, only one good has been considered. Suppose, now, that there are two industries (indexed j) in the economy: Manufacturing (m) and Services (s). For the sake of simplicity, further assume that:

- Only primary factors are used in production processes;
- Fixed costs (H_s, F_{sd}) only involve consumption of services[7];
- Only manufactured goods are consumed by households, and traded between regions.

Let us express with A_s^{ij} the amount of primary factor i used to produce one unit of output in industry j in region s, and with T_{sd}, F_{sd}, H_s the amount of *services* needed to: (1) carry one unit of manufactured good from s to d[8], (2) establish a trade link sd[9], (3) start a business in region s. The following equation, expressing industrial production costs, is added to the system:

$$C_s^j = \sum_i A_s^{ij} W_i^s \tag{20}$$

The parameters A_s^{ij} can be considered as endogenous variables, dependent on relative factor prices, if a general production function is assumed.

Equation (15) would be modified as such:

$$P_{\circ sd} = \left(\frac{C_s^m + T_{sd} C_s^s}{\Phi_{\circ sd}} \right) \left(\frac{\sigma}{\sigma - 1} \right) \tag{21}$$

Recall that only manufactured goods are traded and consumed. Equation (16) would change accordingly:

$$\Pi_{\circ sd} = \left(P_{\circ sd} - \frac{C_s^m + T_{sd} C_s^s}{\Phi_{\circ sd}} \right) Q_{\circ sd} - F_{sd} C_s^d \tag{22}$$

Similarly:

$$\sum_d N_{sd} \Pi_{\circ sd} - N_s H_s C_s^s = 0 \tag{23}$$

and also:

$$L_s^i = \sum_d \frac{N_{sd} Q_{\circ sd} \left[A_s^{im} + T_{sd} A_s^{is} \right]}{\Phi_{\circ sd}} + \sum_d N_{ds} F_{ds} A_s^{is} + N_s H_s A_s^{is} \tag{24}$$

[7] However, services in a region are produced with the same technology. This implies that transportation and all kind of fixed costs share the same cost structure.

[8] Notice that we are slightly changing the notation and now this parameter is no more a multiplicative factor greater than one.

[9] To be consistent with the previous setting, we shall keep assuming that the demand for services is generated in the destination country.

3.4 Hybrid Industrial Structure, Intermediate Inputs, Final Demand

In this last variant of the model, we allow for the existence of intermediate factors and for the consumption of "Services" (that is, products from conventional industries, without firm heterogeneity) by households. Two simplifying hypotheses will be introduced, though[10]:

- Intermediate factors are not substitutable among themselves (à la Leontief)[11];
- Services are domestically produced and consumed. They are not interregionally traded[12].

Let us indicate with a_d^{hj} the interregional input-output coefficients for intermediate inputs, that is the amount of factor goods produced by industry h, necessary to produce one unit of output in industry j located in d. There is an important difference here between services, which are an homogeneous industry, and manufacturing, 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, so that:

$$Z_d^{ms} + Z_d^{mm} + Q_d^m = \left(\sum_s N_{sd} Q_{\circ sd}^{(\sigma\text{-}1)/\sigma} \right)^{\sigma/(\sigma-1)} \tag{25}$$

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

$$Z_d^{mm} = a_d^{mm} \left(\sum_s N_{ds} Q_{\circ ds} / \Phi_{\circ ds} \right) \tag{26}$$

$$Z_d^{ms} = a_d^{ms} X_d^s \tag{27}$$

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

$$\begin{gathered} X_d^s = Q_d^s + a_d^{sm} \left(\sum_s N_{ds} Q_{\circ ds} / \Phi_{\circ ds} \right) + a_d^{ss} X_d^s + \\ + N_d H_d + \sum_s N_{sd} F_{sd} + \sum_s T_{ds} N_{ds} Q_{\circ ds} / \Phi_{\circ ds} \end{gathered} \tag{28}$$

[10]These assumptions are not essential. Results could be easily generalized.

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

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

where Q_d^s is the quantity of services directly consumed by households in region d. Correspondingly, the demand for primary factors becomes:

$$L_s^i = \sum_d \frac{N_{sd} Q_{\circ sd} A_s^{im}}{\Phi_{\circ sd}} + X_s^s A_s^{is} \tag{29}$$

In this setting, 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 as specified in the left hand side of (19). For example, if the utility function is linear logarithmic (Cobb-Douglas), then budget shares (ψ^j) would be constant, and consumption levels would be implicitly set by:

$$\psi_d^j \left(\sum_i L_d^i W_d^i \right) = Q_d^j P_d^j \tag{30}$$

The inclusion of differentiated production factors adds a special feature to the model. Any increase in the number of trading manufacturing firms would not only bring about a welfare gain, because of the Dixit-Stiglitz "taste for variety" effect, but also an increase in productivity for intermediate imported factors, like in Fujita, Krugman and Venables (1999). Aggregate productivity effects therefore overlap firm-level productivity effects.

Furthermore, intermediate demand simply adds to final consumption. The quantity bundle on the right hand side of (25) refers to total demand in a region, implying that the internal composition of intermediate and final trade flows (and the associated price index) is the same.

4 Calibration

Calibration is the procedure which is followed to set parameter values in CGE models. The general equilibrium model can be seen as a system with n equations, determining *n-1* endogenous variables out of a total of $m>n$ variables and parameters. A CGE model can be calibrated when most "naturally endogenous" variables, like trade flows, are statistically observed at a specific time (calibration year). In this case, some endogenous variables can be "swapped" with an equivalent number of exogenous variables or parameters. That is, previously endogenous variables are fixed and the system is used to compute parameter values, which amounts to assuming that available economic data are describing a general equilibrium state.[13]

The extension of the CGE structure to include Melitz equations increases the total number of parameters in the system. The standard calibration method

[13] Not all parameters can be set this way. For example, elasticities of substitution are typically left out.

Table 1: An illustrative SAM matrix structure

	m_a	s_a	m_b	s_b	c_a	c_b
m_a	V_{aa}^{mm}	V_{aa}^{ms}	V_{ab}^{mm}	V_{ab}^{ms}	V_{aa}^{mc}	V_{ab}^{mc}
s_a	V_{aa}^{sm}	V_{aa}^{ss}	V_{ab}^{sm}		V_{aa}^{sc}	
m_b	V_{ba}^{mm}	V_{ba}^{ms}	V_{bb}^{mm}	V_{bb}^{ms}	V_{ba}^{mc}	V_{bb}^{mc}
s_b	V_{ba}^{sm}		V_{bb}^{sm}	V_{bb}^{ss}		V_{bb}^{sc}
l_a	V_a^{lm}	V_a^{ls}				
k_a	V_a^{km}	V_a^{ks}				
l_b			V_b^{lm}	V_b^{ls}		
k_b			V_b^{km}	V_b^{ks}		

for CGE models therefore falls short in the determination of all parameter values. In this section, we first reconsider the standard calibration procedure, to understand which (and how many) parameters can still be estimated. We shall do this by identifying the potentially observable endogenous variables and by associating them with specific parameters. Subsequently, we shall discuss how all remaining parameter values could be estimated with alternative methods.

The basic building block of a CGE calibration is a Social Accounting Matrix (SAM), which is a matrix displaying the flows of income among sectors of an economy (at a given year). The structure of a SAM consistent with the model presented in sub-section 3.4 is presented in Table 1. Here we are considering two regions/countries (a, b), two industries (m, s), one final consumption sector (c), two primary factors (l, k). Only manufactured goods produced in sector m are traded between regions. Primary factors are employed in the region where they are located.

The matrix show *values* flowing from row-sectors to column-sectors. Empty cells mean zero flows. Accounting balances ensure that:

- Costs (possibly including profits) equal revenues in production sectors

$$\sum_h \sum_d V_{ds}^{hj} + \sum_i V_s^{ij} = \sum_h \sum_d V_{sd}^{jh} + \sum_d V_{sd}^{jc}$$

- Final consumption expenditure equals income from primary factors

$$\sum_s \sum_j V_{sd}^{jc} = \sum_i \sum_j V_d^{ij}$$

We assume that a SAM having a structure like in Table 1 is available and we ask ourselves what parameter values for the model in sub-section 3.4 can be obtained from it.

First, notice that the value flows in Table 1 are not endogenous variables in the model but can, however, be derived using the following set of auxiliary equations:

$$V_{ss}^{ss} = A_s^{ss} X_s^s P_s^s \tag{31}$$

$$V_{sd}^{mm} = \theta_{sd} A_d^{mm} \left(\sum_z N_{sz} Q_{\circ sz} / \Phi_{\circ sz} \right) P_s^m \tag{32}$$

$$V_{sd}^{ms} = \theta_{sd} A_d^{ms} X_s^s P_s^m \tag{33}$$

$$V_s^{is} = A_s^{is} X_s^s W_s^i \tag{34}$$

$$V_s^{im} = A_s^{im} \left(\sum_d N_{sd} Q_{\circ sd} / \Phi_{\circ sd} \right) W_s^i \tag{35}$$

$$V_{ss}^{sm} = \left(N_s H_s + N_{ss} F_{ss} + \sum_d \left(A_s^{sm} + T_{sd} \right) N_{sd} Q_{\circ \mathrm{sd}} / \Phi_{\circ sd} \right) P_s^s \tag{36}$$

$$V_{ds}^{sm} = N_{sd} F_{sd} P_d^s \tag{37}$$

$$V_{sd}^{mc} = \theta_{sd} Q_d^m P_d^m \tag{38}$$

$$V_{ss}^{sc} = Q_s^s P_s^s \tag{39}$$

where θ_{sd} is the value share of manufactured goods consumed in region d and supplied by region s:

$$\theta_{sd} = \frac{N_{sd} Q_{\circ sd} P_{\circ sd}}{\sum_z N_{zd} Q_{\circ zd} P_{\circ zd}} \tag{40}$$

The subscripts z also denote regions like s.

The relationships above can be used to set values for a number of parameters corresponding to the number of equations. To this end, prices of all goods, services and primary factors can be set to one. This is a standard assumption in CGE and IO models, and it is legitimate because it amounts to choose convenient (normalized) units of measure for the quantity flows. This methodology applies equally well here to homogeneous products and to CES aggregates of manufactured goods. However, notice that the price of the differentiated manufactured bundle would be set to one in the *destination* market, which implies that the origin price at the firm level would typically differ from unity.

To see this point, suppose that units of measure for produced quantities are chosen in such a way that the price set by each firm, including the average firm in any trade link, has the same value $P_{\circ s}$. Recall that:

$$P_d^m = \left(N_{dd} P_{\circ d}^{1-\sigma} + N_{sd} P_{\circ \mathrm{sd}}^{1-\sigma} \right)^{\frac{1}{1-\sigma}} \tag{41}$$

and, since $P_s^s = 1$:

$$P_{\circ sd} = P_{\circ s} + \frac{T_{sd}}{\Phi_{\circ sd}} \frac{\sigma}{\sigma - 1} \tag{42}$$

For the case of two regions (a and b), when $P_a^m = P_b^m = 1$ and assuming $T_{aa} = T_{bb} = 0$, the following system must hold:

$$\begin{cases} 1 = N_{aa} P_{\circ a}^{1-\sigma} + N_{ba} \left(P_{\circ b} + \frac{T_{ba}}{\Phi_{\circ ba}} \frac{\sigma}{\sigma-1} \right)^{1-\sigma} \\ 1 = N_{ab} \left(P_{\circ a} + \frac{T_{ab}}{\Phi_{\circ ab}} \frac{\sigma}{\sigma-1} \right)^{1-\sigma} + N_{bb} P_{\circ b}^{1-\sigma} \end{cases} \tag{43}$$

Solving the system (43) allows determining firm-level prices $P_{\circ s}$.[14]

Notice that, by setting all $W_s^i = 1$, the exogenous endowments of primary resources L_s^i needed in (29) are simply given by:

$$L_s^i = V_s^{im} + V_s^{is} \tag{44}$$

Equations (31), (34), (35) and (36), referring to intermediate purchases of primary factors or homogeneous goods (services), can be employed to set specific values for the input-output parameters A_s^{ij}. In principle, this would also apply to (37) if imported services directly enter into the production processes. In this example, however, we are assuming that foreign services are only required to establish subsidiary branches abroad. Therefore, there are no input-output parameters in (37) and that equation could be used, instead, to set values for the F_{sd} parameters (for $s \neq d$).

In the same vein (39) could be used to set preference parameters for services in the utility function of the representative consumer, as it is normally done in CGE models. Here we are assuming that there are no imported services in final consumption. If we relax this hypothesis, preference parameters could be estimated on the basis of the Armington assumption, which means considering goods or services produced in different locations as materially different.

Not surprisingly, the trickiest part of the calibration process has to do with the treatment of differentiated products (manufacturing). According to (25) total demand for manufactured goods is generated for intermediate input, by all domestic industries, and final consumption. This demand pool is satisfied by a combination of manufactured goods produced by different firms in different regions (including the domestic region), that is a CES bundle, as expressed in the right hand side of (25).

The parameter θ_{sd} in (40) can be computed to identify the contribution (in value terms) of region s inside the consumption bundle of region d. Most SAM data bases include information about the region of origin of all trade flows, both intermediate and final. This means that the regional structure of final consumption (or, alternatively, the one of total consumption) would be sufficient to get the value shares θ_{sd}. However, this regional structure would normally be inconsistent with the one that can be observed in the purchases of intermediate factors.

To illustrate the point, consider Equation (33). Suppose that parameters θ_{sd} have already been obtained from the structure of final consumption. The variable X_s^s can also be endogenously computed once all parameters of the system are known. The total number of equations like (33) is equal to the square of the number n of regions (in our example, four), but the only remaining unknowns are the A_s^{ms}, which are n (in our example, two). This means that the SAM has to have a specific structure, to be fully consistent with the model. In other words, although a SAM could be constructed after computing an equilibrium,

[14]However, this is not needed to calibrate the model. The system (43) would be automatically solved.

the reverse may not be true: a given SAM could not correspond to any equilibrium state. In the latter case, the model can still be calibrated by dropping a certain number of equations like (33) , which are in excess. The same reasoning applies to (32). Of course, a model calibrated in this way will not be able to exactly reproduce the initial SAM.

The transport cost factors T_{sd} and the fixed costs H_s and F_{ss} cannot be estimated on the basis of SAM table alone. However, they could be estimated by exploiting other informational sources. For example, many SAMs and IO tables are provided in two versions: with value flows expressed at market (*cif*) prices, as it assumed in Table 1:

$$V_{sd}^{mm} + V_{sd}^{ms} + V_{sd}^{mc} = V_{sd}^{m(cif)} \tag{45}$$

and with value flows expressed using out-of-the factory, "free on board" (*fob*) prices ($V_{sd}^{m(fob)}$). The difference between *cif* and *fob* prices is given by trade margins on the specific link. Using (21) one can notice that:

$$\frac{V_{sd}^{m(cif)}}{V_{sd}^{m(fob)}} = \frac{P_{\circ sd}}{P_{\circ s}} = \frac{C_s^m + T_{sd}C_s^s}{C_s^m} \tag{46}$$

where C_s^m is the production or marginal cost of one unit of (differentiated) manufactured good in region s for a firm with unitary productivity ($\Phi = 1$), whereas C_s^s expresses the same concept applied to the production of services[15]. Since all market prices are normalized to one, $C_s^s = 1$ and $C_s^m = \sum_i A_s^{im} + \sum_j A_s^{jm}$.[16] Therefore, the following condition can be readily applied for the calibration of the transport cost factors T_{sd} :

$$T_{sd} = \left(\frac{V_{sd}^{m(cif)}}{V_{sd}^{m(fob)}} - 1 \right) \left(\sum_i A_s^{im} + \sum_j A_s^{jm} \right) \tag{47}$$

For the fixed cost H_s, which applies to all firms, including those not active in any market, Balistreri and Rutherford (2013) suggest to implicitly calibrate these parameters by linking the mass of firms N_s to the number of active domestic firms N_{ss}, that is by imposing:

$$N_{ss} = \nu_s N_s \tag{48}$$

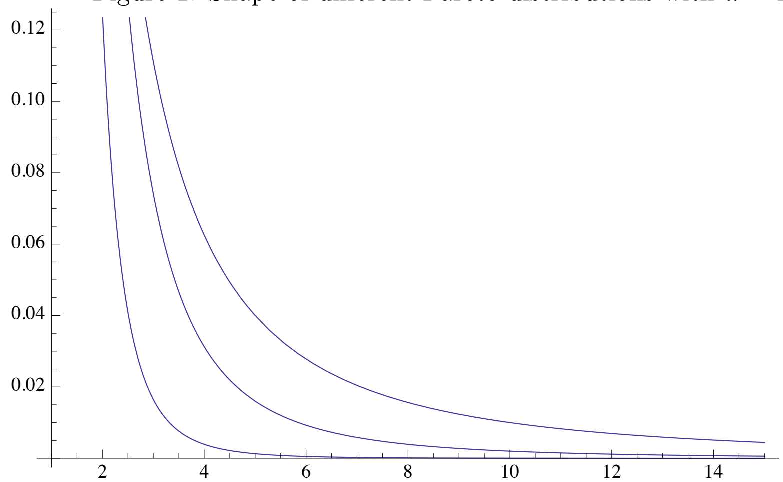
where $\nu_s < 1$ is a chosen parameter. As one can see, H_s does not appear in (48), but it will be automatically set if this condition is imposed, since there is a direct relationship between H_s and N_s, due to the zero-profit equation (23).

The fixed cost F_{ss} applies to all firms active in the domestic market. Its value can be inferred if information is available about the share ϑ_s of exporting firms to total domestic firms. In the Melitz model, the set of exporting firms

[15]In this case, productivity is assumed to be one for all firms.

[16]Notice that, contrary to standard CGE and IO models, the latter may differ from one.

Figure 1: Shape of different Pareto distributions with α = 1, 2, 4



is always a subset of total domestic firms, that is $N_{ss} > N_{sd}$. Therefore, the condition:

$$\vartheta_s = \frac{\max\limits_{d \neq s} N_{sd}}{N_{ss}} \tag{49}$$

is sufficient to determine F_{ss}. This is because, like in the case discussed above, F_{ss} can be computed by Equation (36) once N_{ss} is known, and vice versa.

Another parameter that cannot be obtained from a SAM matrix, in addition to the elasticity σ, is α in the Pareto distribution of productivity [$p(\Phi)$=α$\Phi^{-\alpha\text{-}1}$, $\Phi \geq 1$]. A Pareto distribution is tail shaped. The lower the α, the thicker the tail of the distribution (Figure 1). Industrial, firm-level data on productivity could be used to infer reasonable values for α by means of non-linear regressions. Balistreri, Hillberry and Rutherford (2011) provide estimates for α in the range 3.9-5.2. Estimates by Bernard, Redding and Schott (2007) and Eaton, Kortum and Kramarz (2004) are 4.2 and 3.4, respectively.

5 Concluding Remarks

In this paper, we discussed how a CGE model should be designed, in order to capture some productivity effects due to firms' intra-industry heterogeneity, like in Melitz (2003). The original Melitz model is not suited to conduct numerical simulation experiments, because it is a stylized one and therefore lacks the wealth of realistic details, which is typical of applied general equilibrium models. Nonetheless, in this paper we showed how a CGE model with "Melitz characteristics" can be built, thereby demonstrating that the two models can actually be merged into a single one.

This result can be achieved at a cost, though. First, the new model is more complex than a standard CGE. Computing the solution may be difficult and, because of potential non-convexities, the choice of starting values for some endogenous variables may be critical, as well as the magnitude of some simulated shocks. Second, calibrating the model parameters involves solving a fairly large and complex non-linear system, which may itself pose computational challenges.

Alongside the costs of introducing the richer model structure there are potentially substantial benefits. New issues and new aspects, which could not be considered in the original Melitz's framework, would now be addressed. For example, the typical experiment of lowering trade barriers, leading to firm selection and aggregate productivity gains in Melitz (2003), now also triggers a reallocation of production among industries and a change in relative returns of primary factors, as it is typical for multi-sectoral general equilibrium models. An increase in the number of exporting firms, for instance, generates an additional demand for services in both the origin and destination countries, because of the presence of variable and fixed trade costs.

Arkolakis, Costinot and Rodriguez-Clare (2012) obtain an equivalence result according to which, despite the fact that new theories have identified additional sources of trade gains, from an empirical perspective and conditional on observed trade data, the total size of the gains from trade may turn out to be the same as that predicted by old-style models. Balistreri, Hillberry and Rutherford (2011) argue that this equivalence may hold in one-good one-factor environments, but does not hold anymore with multiple industries, regions and factors. This debate substantiates our assertion that CGE models with heterogenous firms, like the ones analyzed in this paper, do not only broaden the scope of applied general equilibrium analysis, but can also highlight key qualitative properties of some underlying theoretical models, which cannot be noticed in simpler settings.

References

Arkolakis, C., A. Costinot and A. Rodriguez-Clare. 2012. "New Trade Models, Same Old Gains?" *American Economic Review* 102(1):94–130.

Balistreri, E. and T. Rutherford. 2013. Computing general equilibrium theories of monopolistic competition and heterogeneous firms. In *Handbook of Computable General Equilibrium Modeling*, ed. Dixon P.B. and D.W. Jorgenson. Elsevier chapter 23, pp. 1513–1570.

Balistreri, E.J, R.H. Hillberry and T.E Rutherford. 2011. "Structural Estimation and Solution of International Trade Models with Heterogeneous Firms." *Journal of International Economics* 83(2):95–108.

Bernard, A.B., S. Redding and P.K Schott. 2007. "Comparative advantage and heterogeneous firms." *Review of Economic Studies* 74:31–66.

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

Eaton, J., S. Kortum and F. Kramarz. 2004. "Dissecting trade: Firms, industries and export destinations." *American Economic Review* 94(2):150–154.

Fujita, M., P. Krugman and A.J. Venables. 1999. *The Spatial Economy: Cities, Regions and International Trade*. MIT Press.

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.

Oyamada, K. 2013. Parameterization of Applied General Equilibrium Models with Flexible Trade Specifications Based on the Armington, Krugman, and Melitz Models. IDE discussion paper 380 Institute of Developing Economies.

Roson, R. 2006. "Introducing Imperfect Competition in CGE Models: Technical Aspects and Implications." *Computational Economics* 28:29–49.

Santos Silva, J. M. C. and S. Tenreyro. 2009. Trading Partners and Trading Volumes: Implementing the Helpman-Melitz-Rubinstein Model Empirically. CEP Discussion Paper 935 Centre for Economic Performance.

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