Introducing Firm Heterogeneity into the GTAP Model with an Illustration in the Context of the Trans-Pacific Partnership Agreement

Zeynep Akgul[*], Nelson B. Villoria[†], and Thomas W. Hertel[‡]

April 2014

(Preliminary and Incomplete)

[*]Akgul is a Ph.D. candidate with the Department of Agricultural Economics, Purdue University, West Lafayette, IN 47907, USA.

[†]Villoria is a Research Assistant Professor of Agricultural Economics with the Center for Global Trade Analysis, Purdue University, West Lafayette, IN 47907, USA.

[‡]Hertel is a Professor of Agricultural Economics and the Executive Director of the Center for Global Trade Analysis, Purdue University, West Lafayette, IN 47907, USA.

Contents

1 Introduction

Changes in trade flows in traditional Computable General Equilibrium models (CGE) using the Armington specification (e.g. GTAP) are conditioned by pre-existing trade shares thus capturing trade adjustments only at the intensive margin. This is at odds with the empirical trade literature that highlights the role of extensive margin in explaining the expansion of trade following trade liberalization episodes (Hummels and Klenow, 2005; Chaney, 2008). The firm heterogeneity trade model proposed in the pioneering work of Melitz (2003) combines the intensive margin with the extensive margin effects of trade liberalization capturing the self-selection of firms into export markets based on their respective productivity levels. The resulting framework is solidly supported by empirical evidence (Eaton et al., 2004; Bernard et al., 2006). Including firm-level heterogeneity in a CGE model can improve the ability of the model to trace out the trade expansion and welfare implications of trade policy that were previously unexplored in traditional models. In fact, it is shown that incorporating firm heterogeneity into standard CGE models raises the gains from trade liberalization by a multiple of two in Zhai (2008) and by a multiple of four in Balistreri et al. (2011).

There have recently been some important efforts to incorporate Melitz (2003) into CGE modeling (Zhai, 2008; Balistreri and Rutherford, 2012; Dixon and Rimmer, 2011; Oyamada, 2013). However, a readily accessible GTAP implementation with firm heterogeneity has not yet become available. Our paper addresses this gap by incorporating Melitz-type firm heterogeneity into the GTAP model, calibrating it to the GTAP 8 Data Base and illustrating this framework with a stylized scenario based on the Trans-Pacific Partnership (TPP), a regional free trade agreement (FTA) currently under negotiation between the countries in the

Asia-Pacific region (Petri et al., 2012). A comparison with the Armington-based standard GTAP model, as well as a GTAP-based model of monopolistic competition allows us to shed light on the new elements which the Melitz model brings to bear on trade liberalization impacts.

One of the stylized facts shown by micro-level data is that within an industry firms vary by their productivity, size, profitability, the number of markets served and responses to trade shocks (Balistreri et al., 2011; Melitz and Trefler, 2012) which is supported by the micro-level data (Bernard et al., 2003; Eaton et al., 2004; Bernard et al., 2007). In particular, it is found that only some firms are able to engage in exporting and exporters tend to be larger and more productive than non-exporters (Balistreri et al., 2011; Bernard et al., 2003; Bernard and Jensen, 1999). These stylized facts are captured by the framework in Melitz (2003) who examines the intra-industry reallocation effects of international trade in the context of a model with monopolistic competition and heterogeneous firms. In his framework, opening the economy to trade or increasing the exposure to trade generates a reallocation of market power within the domestic and export markets based on the productivity differences of firms. In particular, firms with higher productivity levels are induced to enter the export market; firms with lower productivity levels continue to produce for the domestic market and the firms with the lowest productivity levels are forced to exit the industry. These inter-firm reallocations generate a growth in the aggregate industry productivity which then increases the welfare gains of trade. This channel is a unique feature of the firm heterogeneity model (Zhai, 2008). The main premise of the Melitz model is that aggregate productivity can change even though there is no change in a country's production technology. As opposed to the allocative efficiency gains in the firm heterogeneity model, aggregate productivity changes in traditional trade models with homogeneous firms and Armington assumption are brought about by changes in firm-level technology.

Melitz (2003) builds on Krugman's (1980) monopolistic competition framework to model trade; while it draws from Hopenhayn (1992) to model the endogenous self-selection of heterogeneous firms. Likewise, we build on Swaminathan and Hertel's (1996) monopolistically competitive GTAP model where variety effects (changes in the number of firms – and hence distinct varieties offered) and scale effects (changes in output per firm) are captured. We draw from the work of Zhai (2008) in modeling firm heterogeneity and parsing productivity thresholds to enter domestic and export markets and the calibration of fixed export costs, etc. This allows us to endogenize aggregate productivity in the monopolistically competitive sectors of the model, thereby capturing the intra-industry reallocation of resources in the wake of trade liberalization.

Our framework differs from Zhai (2008) in two dimensions: (i) firm entry and exit, and (ii) sunk entry costs. Specifically, Zhai (2008) assumes that the total mass of potential firms in each sector is fixed. This simplification eliminates the role of endogenous firm entry and exit; therefore, extensive margin effects pick up only the changes in the shares of firms. In contrast, our model extends his work by incorporating endogenous firm entry and exit behavior and tracing out the direct effect of changes in the productivity threshold on entry and survival in export markets. Another simplification of Zhai (2008) is the assumption of no sunk-entry costs of production in the monopolistically competitive industry. In contrast, our model incorporates fixed entry costs in the fashion of Swaminathan and Hertel (1996) by assuming that they are only comprised of value added inputs and calibrating them using the zero profits condition which arises from entry/exit of firms. An additional contribution of our model is the decomposition of the welfare implications of trade policy. This is an extension of the existing GTAP welfare decomposition (Huff and Hertel, 2000), which now includes, in addition to allocative efficiency and terms of trade effects, scale, variety, and endogenous productivity effects derived from the firm heterogeneity model.

This paper is organized as follows: In Section 2 the theoretical framework is laid out with details into the GTAP implementation. Section 3 describes the data requirement for the firm heterogeneity model. In Section 4 we illustrate this framework with a stylized trade liberalization scenario. Section 5 concludes the paper.

2 Monopolistic Competition with Heterogeneous Firms

This section describes the theoretical structure of the Melitz model and how it is implemented into the GTAP. We build on Swaminathan and Hertel's (1996) monopolistically competitive GTAP model and draw from Zhai (2008) for firm heterogeneity. In order to make this paper self-contained, we explain the significant theoretical concepts adopted from Swaminathan and Hertel (1996) and Zhai (2008) in the text. For details, we refer the reader to the associated papers or the appendix.

One of the aims of this paper is to lay out the direct link between the theory of Melitz-type CGE models and the firm heterogeneity GTAP model. In that sense the notation adopted in this paper is similar to the literature. We explicitly show how to bring the theory into GEMPACK. We hope that this approach makes it easier for the reader to follow the current literature and link it with our methodology such that comparisons with other Melitz-type CGE models is easier.

A quick summary of the notation that we adopt in this paper is warranted. In the sections that follow i denotes a commodity or an industry and r or s denote a country or a region depending on the particular aggregation used. It is important to highlight that we refer to three types of variables throughout the paper; firm-level variables, average variables, and industry-level variables. Firm-level variables are the variables associated with a different variety which are indexed by ω or φ. Average variables are associated with representative variety from the representative firm that operates at the average productivity level. This follows from Melitz (2003) and is explained in more detail in the sections that follow. At the industry-level we have the aggregate variables.

2.1 Demand

For the regional household we follow most of the standard GTAP model assumptions. The regional household collects and allocates all factor income and tax as regional income according to a per capita aggregate utility function of the Cobb-Douglas Form. The government utility function is also specified by the Cobb-Douglas function. We also retain the non-homothetic utility structure of private households via the Constant Difference of Elasticity (CDE) function. The utility tree structure, so far, is the same as the standard model with homogeneous goods produced in perfectly competitive industries. Differentiated varieties enter the utility tree at the third nest which is where we have the sub-utility function for composite commodities with a Constant Elasticity of Substitution (CES) form.

In the standard GTAP model, commodities are homogeneous and they are produced in perfectly competitive industries in each region. Moreover, the standard GTAP model adopts Armington approach to import demand according to which products are distinguished with respect to their country of origin. Thus, at the border we have 'composite imports' which are imperfect substitutes for domestic goods. In other words, composite imports compete with the domestically produced commodities independent of their source country. In the monopolistically competitive structure; however, firms within same industry of the same region produce differentiated products. What matters to the consumers is not where the variety originates, but what distinguishes it from the other products of the industry. As opposed to an import-domestic decision, agents make a variety decision for the monopolistically competitive industry products. As a result, sourced imports directly compete with the domestic varieties and agents face a set of sourced varieties.

The implications of allowing for competition at the variety level is discussed in more detail in Swaminathan and Hertel (1996). One important aspect they emphasize is the need for a change in the structure of the database. Even though the country of origin does not matter directly for consumers' decision about which variety to choose, they are affected by the number of varieties available in the source country. Following Swaminathan and Hertel (1996) we source imports to agents which requires the structural change in the data base. The transformation of the data base is described in Section 3.

Another important aspect of allowing imported varieties for competing directly with domestic varieties is the increase in the model size. In the standard model agent demands are indexed by commodity and country. But, in the monopolistic competition since imports are sourced, agent demands are indexed by commodity, source and destination regions. Introducing this extra dimension increases the number of equations and the related unknown variables to solve for.

2.1.1 Price Linkage

Changes in the data structure due to sourcing of imports is also reflected in the price linkage. In the standard model, the prices that agents face are differentiated between import and domestic prices. However, in the monopolistically competitive model instead of discriminating between domestic and import prices, we have prices that are indexed by source and destination. The price linkage in the monopolistically competitive model is laid out in the Figure 2.1.1.

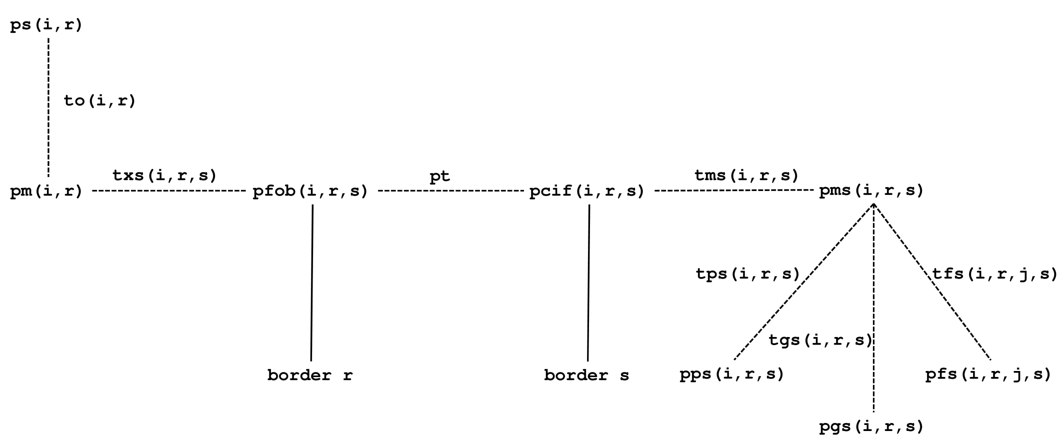

Figure 1: Price Linkage

2.1.2 Private Household Utility

Similar to the Krugman structure, we assume standard Dixit-Stiglitz preferences over a continuum of differentiated varieties. The sub-utility function for the representative household is given by:

$$Q_{is} = \left[\sum_r \alpha_{irs}^{\frac{1}{\sigma_i}} \int_{\omega \in \Omega} Q_{irs}(\omega)^{\frac{\sigma_i - 1}{\sigma_i}} \, d\omega \right]^{\frac{\sigma_i}{\sigma_i - 1}}, \tag{1}$$

where ω indexes each variety in the set of available varieties Ω, Q_{is} is a CES aggregate of all the varieties of commodity i demanded in region s, $Q_{irs}(\omega)$ is the demand for variety of commodity i produced in region r and sold in region s, α_{irs} is the Armington preference parameter reflecting consumers' tendency for home or imported products, and σ_i is the constant elasticity of substitution between different varieties in the monopolistically competitive sectors ($\sigma_i > 1$). Each variety is the product of a different firm. However, we assume that varieties are symmetric for simplification. Therefore, from now on i indexes a *representative product* of each commodity. This allows us to abstract from the continuum of varieties and work with the representative household's demand (see Appendix). The sub-utility function becomes:

$$Q_{is} = \left[\sum_r \alpha_{irs}^{\frac{1}{\sigma_i}} N_{irs} \widetilde{Q}_{irs}^{\frac{\sigma_i - 1}{\sigma_i}} \right]^{\frac{\sigma_i}{\sigma_i - 1}}, \tag{2}$$

where $\widetilde{Q}_{irs}$ is representative household's demand for a representative product i sourced in region r and sold in region s, and N_{irs} is the total number of varieties for product i in region s sourced from region r. In this framework there is a one-to-one correspondence between firms and varieties. Hence N_{irs} also represents the total number of firms in industry i that actively export from region r to s. This brings us to a key differences between the firm heterogeneity GTAP model and the monopolistically competitive GTAP model of Swaminathan and Hertel (1996). In the monopolistically competitive model, it is assumed that all firms that produce in the source country are active on the r-s export link. Therefore, the sub-utility of the representative consumer depends on the *total* number of varieties available for consumption in the monopolistically competitive GTAP model. However, we should highlight that when firms are heterogeneous with respect to their productivity levels, only a few firms that have high productivity levels afford to actively export on every bilateral trade market. Therefore, consumers in the destination region s only have access to the varieties of region r that are actually being exported to s. The set of varieties that are sold on the r-s link is determined endogenously in equilibrium. This will be explained in more detail in Section 2.3.

The average consumer in region s chooses $\widetilde{Q}_{irs}$ that minimizes his expenditure:

$$\min_{\widetilde{Q}_{irs}} \sum_r N_{irs}\widetilde{Q}_{irs}\widetilde{P}_{irs}$$
$$\text{s.to } Q_{is} = \left[\sum_r \alpha_{irs}^{\frac{1}{\sigma_i}} N_{irs}\widetilde{Q}_{irs}^{\frac{\sigma_i-1}{\sigma_i}}\right]^{\frac{\sigma_i}{\sigma_i-1}},$$

where $\widetilde{P}_{irs}$ is the unit price of product i set by the representative firm with average productivity level (from now on 'average firm') operating on the $r-s$ link. Note that this average price is defined as gross of taxes and transportation costs. The minimization problem yields the CES derived demand for the representative product i sourced from region r as:

$$\widetilde{Q}_{irs} = \alpha_{irs}Q_{is}\left[\frac{P_{is}}{\widetilde{P}_{irs}}\right]^{\sigma_i}. \tag{3}$$

By substituting the derived demand (equation 3) into the sub-utility function (equation 2) and rearranging we can obtain the dual Dixit-Stiglitz price index for product i in region s:

$$P_{is} = \left[\sum_r \alpha_{irs} N_{irs} \widetilde{P}_{irs}^{1-\sigma_i} \right]^{\frac{1}{1-\sigma_i}}. \tag{4}$$

In sectors with differentiated goods, the Armington share parameter, α_{irs}, is assumed to be one. This assumption follows from Zhai (2008) and accounts for the fact that bilateral trade patterns are determined by relative prices and the amount of available varieties in the monopolistically competitive industries. On the other hand, N_{irs} is assumed to be one in the standard GTAP model with Armington assumption of national product differentiation. Hence the variance in bilateral trade flows which cannot be explained by relative prices are contained in Armington share parameters (Hillberry et al., 2005). Hence incorporating monopolistic competition into the GTAP model improves the ability of the model to explain the variation in trade patterns by theory.

The rest of this section is devoted to the introduction of changes in the demand structure into the GTAP model and the GEMPACK TABLO file. For this purpose, we focus on equations (3) and (4).

Private Household's Demand for Differentiated Products

We start with derived demand. Equation (3) gives the demand in region s for the representative product i sourced from region r. However, it is only for an individual variety. To obtain the industry level demand, we need to aggregate across all available varieties. One of the reasons why we aggregate is to minimize the additional information required from the dataset. As mentioned in Swaminathan and Hertel (1996) the number, size and sales information in firm-level data is limited. We follow Melitz (2003) to aggregate the firm-level information into industry-level variables. The details of the Melitz-type aggregation is discussed later.

Aggregate derived demand in region s for all varieties of product i sourced from r is given by:

$$Q_{irs} = N_{irs}^{\frac{\sigma_i}{\sigma_i - 1}} \widetilde{Q}_{irs}. \tag{5}$$

Aggregate price index in region s for all varieties of product i sourced from r is given by:

$$P_{irs} = N_{irs}^{\frac{1}{1-\sigma_i}} \widetilde{P}_{irs}. \tag{6}$$

Using equations (3), and (6) in equation (5) we obtain the aggregate derived demand on the $r - s$ link as:

$$Q_{irs} = Q_{is} \left[\frac{P_{is}}{P_{irs}} \right]^{\sigma_i}, \tag{7}$$

where α_{irs} is assumed as one; thereby, is dropped for simplicity. Equation (7) is the levels form of the aggregate demand in region s for product i sourced from r. In GEMPACK we work with the linearized representations of equations; therefore, we totally differentiate the equations we will incorporate into the firm heterogeneity module.

Total differentiation of the aggregate demand yields:

$$q_{irs} = q_{is} - \sigma_i [p_{irs} - p_{is}], \tag{8}$$

where the lower-case letters denote percentage changes in the associated upper-case variables. Formally in TABLO:

```
EQUATION PHLDSRCDF
# Private HouseholD demand for SouRCed DiFFerentiated commodity #
# Equation names are identified by capitalized letters in comments
#
(all,i,MCOMP_COMM)(all,r,REG)(all,s,REG)
qdfps(i,r,s) = qp(i,s) - SIGMA(i) * [pps(i,r,s) - pp(i,s)];
```

This equation shows that percentage change in derived demand for differentiated products still depends on the standard expansion and substitution effects. However, we should highlight one significant difference. Imported varieties compete directly with domestic varieties in the monopolistically competitive market. Therefore, the substitution effect is the product of the constant elasticity of substitution between differentiated varieties, *SIGMA(i)*, and the price of sourced differentiated product relative to private household's unit expenditure (or composite price index), *[pps(i,r,s) - pp(i,s)]*.

Private Household's Composite Price Index for Differentiated Products

We follow the same method in incorporating the composite price index. Total differentiation of equation (4) yields:

$$p_{is} = \sum_r \alpha_{irs} N_{irs} \left[\frac{\widetilde{P}_{irs}}{P_{is}} \right]^{1-\sigma_i} \widetilde{p}_{irs} - \frac{1}{\sigma_i - 1} \sum_r \alpha_{irs} N_{irs} \left[\frac{\widetilde{P}_{irs}}{P_{is}} \right]^{1-\sigma_i} n_{irs}. \tag{9}$$

Using equation (3) in equation (9) and rearranging we obtain:

$$p_{is} = \sum_r \frac{N_{irs} \widetilde{Q}_{irs} \widetilde{P}_{irs}}{Q_{is} P_{is}} \widetilde{p}_{irs} - \frac{1}{\sigma_i - 1} \sum_r \frac{N_{irs} \widetilde{Q}_{irs} \widetilde{P}_{irs}}{Q_{is} P_{is}} n_{irs}, \tag{10}$$

which can be simplified by defining the expenditure share as:

$$\theta_{irs} = \frac{N_{irs}\widetilde{Q}_{irs}\widetilde{P}_{irs}}{Q_{is}P_{is}}, \tag{11}$$

where θ_{irs} is the expenditure share of all varieties of product i originating from source r in total expenditure of all varieties from all sources in region s. In TABLO:

```
FORMULA (all,i,MCOMP_COMM)(all,r,REG)(all,s,REG)
PTHETA(i,r,s) = VPAS(i,r,s) / sum{k,REG, VPAS(i,k,s)};
```

where *VPAS(i,r,s)* is the value of private household expenditure in region s at agent's price by source. The same applies to the computation of θ_{irs} for the government and intermediate inputs. See Swaminathan and Hertel (1996) for more details.

Substituting θ_{irs} (equation (11))into equation (10), the linearized representation of the composite price index equation becomes:

$$p_{is} = \sum_r \theta_{irs}\widetilde{p}_{irs} - \frac{1}{\sigma_i - 1}\sum_r \theta_{irs}n_{irs}. \tag{12}$$

In TABLO:

```
EQUATION PHLDDFCOMPR
# Private HouseholD PRice for DiFferentiated COMposite commodity #
(all,i,MCOMP_COMM)(all,s,REG)
  pp(i,s) = sum{r,REG, PTHETA(i,r,s) * ppsf(i,r,s)}
          - {1/[SIGMA(i) - 1]} * vp(i,r,s) ;
```

Equation PHLDDFCOMPR shows that there are two things that distinguish price index for differentiated products from that of homogeneous products. First of all, data is sourced by agent. Hence agents face the average price of varieties from each source region, $ppsf(i,r,s)$. That applies to the expenditure shares $PTHETA(i,r,s)$, as well. In contrast, for homogeneous products, agents face a composite of imported products. Secondly, we see the effect of variety on price index via a variety index measure, $vp(i,r,s)$. For simplicity, variety index is defined in TABLO as a separate equation:

```
EQUATION PHLDVARIN
# Private HouseholD VARiety INdex #
(all,i,MCOMP_COMM)(all,s,REG)
vp(i,r,s) = sum{r,REG, PTHETA(i,r,s) * nx(i,r,s)} + vpslack(i,r,s);
```

Equation PHLDVARIN shows that additional varieties on the $r-s$ link raises the variety index given the budget share. We introduce a slack variable, $vpslack(i,r,s)$, into this equation to allow for alternative closure options when there is no variety effect (no entry and exit of new firms into the market).

As described by equation PHLDDFCOMPR, we see that an increase in the variety index lowers the price index ($SIGMA(i) > 1$). As more and more varieties are available for consumption, the amount of expenditure necessary to obtain a unit of utility declines given prices.

Other Demand for Differentiated Products

The same changes apply to the government and intermediate input demand equations. The new equations introduced for the government module are GOVSRCDF, GOVDFCOMPR, and GOVVARIN and for intermediate input are INDSRCDF, INDDFCOMPR, and INDVARIN.

There are several issues related to intermediate input demands that need to be clarified. As previously explained by Swaminathan and Hertel (1996), industry derived demand equations have an extra dimension indexing the industry which demands the intermediate input. With the extra dimension, sourced intermediate input demand equations become four dimensional.

Firms are among the sources of demand and they, too, demand differentiated varieties as intermediate inputs. We assume that firms are symmetric in their demand for differentiated intermediate inputs. What matters for derived demand equations is the nature of the intermediate input, not the nature of the industry that demands the input. For instance, if a firm from a perfectly competitive market demands differentiated intermediate inputs, then its derived demand equation incorporates the effect of available varieties. Conversely, if a firm from a monopolistically competitive market demands homogeneous intermediate inputs, then the derived demand equation it faces will have the standard form with Armington assumption and composite import commodity formed at the border.

2.2 Production

Industry is characterized by a continuum of firms each producing a unique variety. There is free entry and exit. Entry into an industry requires paying a fixed entry cost which is thereafter sunk. This entry cost allows for increasing returns to scale. Before entering the industry, firms are identical. Only after entering the market firms draw their initial productivity parameter from a common distribution and their productivity levels are revealed. Firms are heterogenous in their productivity levels so that productivity is inversely related to marginal costs. This firm-level heterogeneity means that production is carried out only by firms that are productive enough to afford staying in the market given sunk-entry costs. Compared to low-productivity firms, high-productivity firms produce more output and earn higher profits by charging a lower price. Therefore, while low-productivity firms are forced to exit the market, high-productivity firms expand their market shares. Surviving firms have the choice to supply foreign markets as well as satisfying home demand. To enter export markets, firms face another set of fixed costs which are destination specific. Just as firms self-select into the domestic market, they self-select into export markets based on their respective productivity levels. Hence the distribution of firms is such that while the most productive firms serve in the export markets, firms with lower productivity levels supply only the domestic market, and the lowest-productivity firms exit the industry.

This section discusses the production structure that characterize the monopolistically competitive industry with firm-level heterogeneity. We modify the approach of Swaminathan and Hertel (1996) in modeling mark-up pricing and cost-structure. We follow Zhai(2008) to introduce endogenous productivity changes into the framework.

2.2.1 Fixed Costs

Fixed costs are associated with the fact that differentiating a product within a particular market is costly. Each firm targeting a niche has to invest in research and development (R & D) to adapt its product to customer needs. Each new market has different legislations and standards. Firms need to do market research to learn about these rules and comply with them. These are fixed costs each firm incurs to produce and sell its unique variety in a monopolistically competitive industry. Upon entering the industry/market, these costs are sunk and do not change with the production level. In this paper, we refer to the domestic market entry cost as 'sunk-entry cost' and export market entry-cost as 'fixed export cost'.

Following Swaminathan and Hertel (1996) we assume that all fixed costs (sunk-entry and fixed export costs) are made up of primary input costs. The rationale is that firms need capital to build the R & D labs, pay salaries to the employees that work in R & D labs, doing market research, and marketing the new product. These are all assumed to be primary factor costs. All of the primary factor costs are referred to as Value Added Costs, VA(i,r), in the GTAP model. Hence VA(i,r) is split into three parts: variable value-added, fixed value-added for the domestic market and fixed value-added for the export market. We will examine the cost structure in more detail later.

2.2.2 Value Added

Since differentiated industries require resources devoted to fixed costs to adapt their products for new markets, demand for fixed value added is directly proportional to the number of successful firms in the monopolistically competitive market. This applies to both sunk-entry costs and fixed export costs. Formally (in TABLO code):

```
EQUATION VAFDDEMAND
# monopolistically competitive industry DEMAND for Fixed Value-
Added #
(all,j,MCOMP_COMM)(all,s,REG)
qvafd(j,s) = n(j,s) ;
```

Sunk entry-costs are faced by all firms that enter the industry. We state that as more firms enter the industry (as $n(j,s)$ increases), the need for primary factors increase, and demand for fixed value-added to cover sunk-entry costs, $qvafd(j,s)$, rises.

Some of the successfully producing firms in the industry self-select into export markets based on their productivity levels and the fixed export costs they face. Like with sunk-entry costs, demand for value-added used in fixed export costs is directly proportional to the number of exporting firms on the specific bilateral link. In TABLO code:

```
EQUATION VAFXDEMAND
# Value-Added DEMAND for Fixed eXport Costs #
(all,i,MCOMP_COMM)(all,r,REG)(all,s,REG)
    qvafx(i,r,s) = nx(i,r,s) ;
```

Sunk-entry costs, and fixed export costs are the two components that make up total demand for fixed value-added. To obtain the total demand for fixed value-added we aggregate $qvafd(j,s)$, and $qvafx(i,r,s)$ based on their respective shares in total fixed costs. In TABLO:

```
EQUATION VAFIX
# monopolistically competitive industry demand for FIXed Value-
Added #
(all,j,MCOMP_COMM)(all,s,REG)
qvaf(j,s) = [VAFD(j,s)/VAF(j,s)] * qvafd(j,s)
          + sum(k,REG,[VAFX(j,s,k)/VAF(j,s)] * qvafx(j,s,k)) ;
```

where $VAFD(j,s)/VAF(j,s)$ is the share of sunk-entry cost in total fixed cost and $[VAFX(j,s,k)/VAF(j,s)]$ is the share of fixed export cost in total fixed cost.

In addition to the fixed component, there is also a variable component of the value-added demand. The derived demand equation for variable value-added in monopolistically competitive industry is similar to that of the perfectly competitive industry and follows from the cost minimization problem in Gohin and Hertel (2003). In TABLO:

```
EQUATION VAVDEMAND
# monopolistically competitive industry DEMAND for Variable Value-
Added #
(all,j,MCOMP_COMM)(all,s,REG)
qvav(j,s) + avav(j,s) = qo(j,s) - ao(j,s) ;
```

The variable value-added demand in industry j in region s, $qvav(j, s)$, is proportional to the industry output given productivity. If firms in the industry become more productive, they use less variable value-added to produce a given level of output. Hence demand for variable value-added declines.

Demand for total value added in industry j in region s, qva(j,s), is a weighted summation of variable and fixed value-added demand based on respective weights of variable and fixed value-added in total value-added. In TABLO:

```
EQUATION VATOT
# monopolistically competitive industry demand for TOTal Value-
Added #
(all,j,MCOMP_COMM)(all,s,REG)
qva(j,s) = [VAV(j,s)/VA(j,s)] * qvav(j,s)
        + [VAF(j,s)/VA(j,s)] * qvaf(j,s) ;
```

2.2.3 Variable Costs

Each firm minimizes its cost according to the following cost minimization problem:

$$\min_{X_{ijr}} \sum_j W_{ijr} X_{ijr}$$

$$\text{s.t.} Q_{jr} = \varphi_{jr} \left[\sum_i \delta_{ijr} X_{ijr}^{\frac{\sigma_j - 1}{\sigma_j}} \right]^{\frac{\sigma_j}{\sigma_j - 1}},$$

where W_{ijr} is the price of input i employed in industry j of region r and X_{ijr} is the industry j demand for input i in region r, φ_{jr} is the productivity of industry j in region r, and δ_{ijr} is a distribution parameter that is used as input-augmenting technical change in the GTAP model.

In contrast to Melitz (2003), production in our model requires multiple factors (land, labor, capital, intermediate input). Therefore, we define productivity as output per composite input according to equation (**??**).

The cost minimization problem yields the following unit cost, C_{jr} for product j in region r:

$$C_{jr} = \frac{1}{\varphi_{jr}} \left[\sum_i \delta_{ijr}^{\sigma_j} W_{ijr}^{1-\sigma_j} \right]^{\frac{1}{1-\sigma_j}}. \tag{13}$$

Total differentiation of equation (13) yields:

$$c_{jr} = \sum_i \beta_{ijr} \left[w_{ijr} - \frac{\sigma_j}{\sigma_j - 1} \widehat{\delta}_{ijr} \right] - \widehat{\varphi}_{jr}, \tag{14}$$

where β_{ijr} is the cost share of factor i in industry j,and lowercase letters and variables with hats denote percentage changes. We talk about what cost shares entail later in this section. For details about this linearization see Gohin and Hertel (2003).

Equation (14) is used to determine two types of costs: average variable cost and average total cost. As explained before, average variable cost is a function of the variable input prices such as intermediate inputs, and variable value-added. In TABLO:

```
EQUATION AVERAGEVC
# AVERAGE Variable Cost #
(all,j,MCOMP_COMM)(all,r,REG)
  VC(j,r) * avc(j,r)
    = sum{i,TRAD_COMM, VFA(i,j,r) * [pf(i,j,r) - af(i,j,r)]}
    + VAV(j,r) * [pva(j,r) - avav(j,r)] - VC(j,r) * ao(j,r) ;
```

Note that the cost shares, β_{ijr}, in equation (14) correspond to $\frac{VFA(i,j,r)}{VC(j,r)}$ for intermediate inputs, and $\frac{VAV(i,j,r)}{VC(j,r)}$ for the variable value-added in the code. The distribution parameters, $\frac{\sigma_j}{\sigma_j - 1}\widehat{\delta}_{ijr}$, in equation (14) correspond to input-augmenting technical change parameters, $af(i,j,r)$ and $avav(j,r)$, in the code.

As equation AVERAGEVC shows average variable cost does not vary with the output level. On the other hand, average total cost of each firm change as a result of two sources: (i) changes in firm output at constant input prices, or (ii) changes in input prices at constant firm-level output. Following Swaminathan and Hertel (1996) we focus on the latter and calculate changes in average total cost at constant firm-level output. Average total cost in TABLO follows from equation (14) as:

```
EQUATION SCLCONATC
# Average Total Cost at CONstant SCale #
(all,j,MCOMP_COMM)(all,r,REG)
  VOA(j,r) * scatc(j,r)
        = sum{i,TRAD_COMM, VFA(i,j,r) * [pf(i,j,r) - af(i,j,r)]}
        + VA(j,r) * [pva(j,r) - ava(j,r)] - VOA(j,r) * ao(j,r) ;
```

Since equation SCLCONATC determines average total cost at constant firm-level output (constant scale), we adopt the convention of Swaminathan and Hertel (1996) and refer to $scatc(j,r)$ as the ‘scale constant average total cost’.

2.2.4 Productivity Draw

Firms are assumed to draw their productivity level, φ, from a Pareto distribution with the lower bound φ_{min} and shape parameter γ. The cumulative distribution function of the Pareto distribution, $G(\varphi)$, is:

$$G(\varphi) = 1 - \varphi^{-\gamma}, \tag{15}$$

for $\varphi \in [1, \infty)$ where the lower bound φ_{min} is assumed to be equal to one following Zhai(2008). The corresponding density function, $g(\varphi)$, is:

$$g(\varphi) = \gamma\varphi^{-\gamma-1}. \tag{16}$$

The shape parameter, γ, is an inverse measure of the firm heterogeneity. If it is high, it means that the firms are more homogeneous. It is also assumed that $\gamma > \sigma - 1$. This assumption is important in aggregation and it ensures that the size distribution of firms has a finite mean (Zhai, 2008). The ex-ante probability of successful entry is captured by:

$$1 - G(\varphi^*) = (\varphi^*)^{-\gamma}, \tag{17}$$

where φ^* is the threshold productivity level of producing in the market. This definition of the ex-ante probability is used in all the sectoral aggregations over individual varieties which we focus on in the following sections.

2.2.5 Markup Pricing

In the standard GTAP model each firm produces a homogeneous product under constant returns to scale. The optimal pricing rule is that price equals marginal cost. This is still the case for the perfectly competitive sectors in our model. However, firms in the monopolistically competitive industry are price setters for their particular varieties. Therefore, the optimal pricing rule for such a firm is to charge a constant markup over marginal cost which is referred to as the mark-up pricing rule given by:

$$P_{ir} = \frac{\sigma_i}{\sigma_i - 1} \frac{C_{ir}}{\widetilde{\varphi_{ir}}}, \tag{18}$$

where P_{ir} is the supply price of product i in the monopolistically competitive sector in region r , $\frac{\sigma_i}{\sigma_i - 1}$ is the mark-up in industry i, C_{ir} is the unit cost of product i in region r, and $\widetilde{\varphi_{ir}}$ is the average productivity of industry i in region r. Firms in the perfectly competitive sectors do not have mark-ups and the industry average productivity level is one in those sectors, , $\widetilde{\varphi_{ir}} = 1$.

Simplifying the mark-up rule in equation (18) we obtain:

$$PS_{ir} = MARKUP_{ir} MC_{ir}, \tag{19}$$

where PS_{ir} is the supply price (excluding taxes and transportation costs), and MC_{ir} is the marginal cost which correspond to $\frac{C_{ir}}{\widetilde{\varphi_{ir}}}$ in equation (18). Since we assume that production occurs under constant returns to scale, average variable cost equals the constant marginal cost of production. Substituting the average variable cost, AVC_{ir} for MC_{ir} in equation (19) we obtain:

$$PS_{ir} = MARKUP_{ir} AVC_{ir}. \tag{20}$$

Total differentiation of (20) yields:

$$ps_{ir} - avc_{ir} = markup_{ir} = 0 \tag{21}$$

According to equation ((21)), in the monopolistically competitive industry, changes in the producer price is directly proportional to changes in average variable cost at constant mark-up. In TABLO:

```
EQUATION MKUPRICE
# MarKup pricing (with constant markup) #
(all,j,MCOMP_COMM)(all,r,REG)
psf(j,r) = avc(j,r) + mkupslack(j,r) ;
```

where $psf(j,r)$ is the price received by average firm j in the monopolistically competitive industry in region r. Industry level price requires aggregation over all the firms in the industry. Equation MKUPRICE determines firm-level output, $qof(j,r)$, in the monopolistically competitive industry in region r. We include a slack variable, mkupslack(j,r), in order to allow for alternative closures for different trade policy applications where firm-level output is fixed. In such a scenario, firm-level price would no longer equal average variable cost and the slack variable, mkupslack(j,r), would absorb the difference between them.

2.2.6 Firm Profits (Productivity Threshold)

Each firm with productivity φ_{irs} makes the following profit from selling variety i on the $r-s$ link:

$$\pi_{irs}(\varphi) = Q_{irs}(\varphi)\frac{P_{irs}(\varphi)}{(1+t_{irs})} - Q_{irs}(\varphi)\frac{C_{ir}}{\varphi_{irs}} - W_{ir}F_{irs}, \tag{22}$$

for all r where the first component, $Q_{irs}(\varphi)\frac{P_{irs}(\varphi)}{(1+t_{irs})}$, gives the total revenue, the second component, $Q_{irs}(\varphi)\frac{C_{ir}}{\varphi_{irs}}$, gives the variable cost and the third component, $W_{ir}F_{irs}$, gives the fixed cost of exporting on the $r-s$ link. Substituting the optimal demand and price for each variety, we obtain the maximized profit for each firm as follows:

$$\pi_{irs} = \frac{Q_{irs}P_{irs}^{\sigma_i}}{\sigma_i(1+t_{irs})}\left[\frac{\sigma_i}{\sigma_i - 1}\frac{(1+t_{irs})C_{ir}}{\varphi_{irs}}\right]^{1-\sigma_i} - W_{ir}F_{irs}. \tag{23}$$

Firms in industry i export on the $r-s$ link as long as the variable profit they make cover the fixed cost of exporting. The firms with high productivity levels set a lower price with a higher markup, produce more output; thereby, earn positive profits. The only firm that exports on the $r-s$ link and makes zero profits is the marginal firm which produces at the threshold productivity level. At that threshold variable profit only covers the export costs so the firm makes zero economic profit. Thus, the zero-cutoff level of productivity for exporting on the $r-s$ link is where:

$$\pi_{irs}(\varphi_{irs}^{*}) = 0.$$

Solving equation (2.2.6) for the threshold productivity level yields:

$$\varphi_{irs}^{*} = \frac{C_{ir}}{\sigma_i - 1}\left[\frac{P_{irs}}{\sigma_i(1+t_{irs})}\right]^{\frac{\sigma_i}{1-\sigma_i}}\left[\frac{W_{ir}F_{irs}}{Q_{irs}}\right]^{\frac{1}{\sigma_i - 1}}. \tag{24}$$

Any firm that has a productivity level below φ_{irs}^{*} cannot afford to produce in that market, and therefore exits. On the other hand, any firm that has a productivity level above φ^*_{irs} stays in the market. Total differentiation of equation (24) yields:

$$\widehat{\varphi}^*_{irs} = c_{ir} + \frac{\sigma_i}{1-\sigma_i}[p_{irs} - t_{irs}] + \frac{1}{\sigma_i - 1}[w_{ir} + f_{irs} - q_{irs}]. \tag{25}$$

Equation (25) shows that the change in cutoff productivity level depends on the unit cost of production, c_{ir} , price net of taxes and transportation costs, $[p_{irs} - t_{irs}]$, and the fixed cost per sale, $[w_{ir} + f_{irs} - q_{irs}]$. The same equation is used to determine the productivity threshold for the export market and for the domestic market. The only difference is with respect to the fixed costs. For the domestic threshold, sunk-entry costs per domestic sale is used, while for the export threshold, fixed export costs per bilateral sale is used.

For the domestic market ($r = s$), equation (25) reduces to:

```
EQUATION PRODTRESHOLDD
# PRODuctivity THRESHOLD for the Domestic market#
(all,i,MCOMP_COMM)(all,r,REG)
  aodf(i,r) = avc(i,r) - MARKUP(i,r) * ps(i,r)
                       + [MARKUP(i,r) - 1] * [fdc(i,r) - qds(i,r)] ;
```

Note that $fdc(i,r)$ is the sunk-entry cost which is a product of value added price and fixed value-added inputs. In TABLO:

```
EQUATION FIXEDD
# FIXED Domestic costs (sunk-entry costs) #
(all,i,MCOMP_COMM)(all,r,REG)
fdc(i,r) = pva(i,r) + qvafd(i,r);
```

where $qvafd(i,r)$ is the quantity of fixed value-added demanded by firms in industry i in region r. Productivity threshold for the export market ($r \neq s$) according to equation (25) is:

```
EQUATION PRODTRESHOLDX
# PRODuctivity THRESHOLD for the eXport market#
(all,i,MCOMP_COMM)(all,r,REG)(all,s,REG)
aoxf(i,r,s) = [1 - DELTA(r,s)]
              * {avc(i,r) - MARKUP(i,r) * ps(i,r)
              + [MARKUP(i,r) - 1] * [fxc(i,r,s) - qs(i,r,s)]};
```

where $DELTA(r,s)$ is called the Kronecker delta which is equal to one when $r = s$. It is used in order to calculate the productivity threshold for just the export market. Similar to the domestic sunk-entry cost, fixed export cost, $fxc(i,r,s)$, is a product of value added price and fixed value-added inputs. In TABLO:

```
EQUATION FIXEDX
# FIXED Export costs #
(all,i,MCOMP_COMM)(all,r,REG)(all,s,REG)
fxc(i,r,s) = pva(i,r) + qvafx(i,r,s);
```

where $qvafx(i,r)$ is the quantity of fixed value-added demanded by firms in industry i that export on the $r-s$ link. Equation PRODTRESHOLDD and PRODTRESHOLDX give us productivity thresholds at the firm-level for the domestic and export markets, respectively. An increase in average variable cost makes it more costly to enter a new market which in turn raises the productivity threshold. This causes less productive firms to exit the market. If fixed export cost per sale declines, the productivity threshold for the export market also declines, given $[MARKUP(i,r) - 1] > 0$. Fixed costs are now spread over more output such that it is less costly for lower productivity firms to engage in exports on the r-s link.

2.2.7 Average Productivity

In equilibrium, firms that have productivity levels above the threshold, φ^*_{irs}, produce for the market. Thus, the productivity of the industry is a weighted average of the productivity levels of the firms that make the cut. The distribution of productivity in equilibrium is given by

$$\mu(\varphi) = \{ \begin{array}{ll} \frac{g(\varphi)}{1-G(\varphi^*)} & \text{if } \varphi \geq \varphi^* \\ 0 & \text{otherwise} \end{array} \tag{26}$$

The distribution can be thought of as a conditional distribution of $g(\varphi)$ on $[\varphi^*, \infty)$ (meaning successful entry). This makes sense since the average productivity of the sector will only be affected by the firms that are successful entrants. Now, we can define the weighted average productivity level of the active firms on the r-s link are as a function of the cutoff level of productivity as follows:

$$\widetilde{\varphi}_{irs}(\varphi^*_{irs}) = \left[\int_{\varphi^*_{irs}}^{\infty} \varphi^{\sigma_i - 1} \mu(\varphi) \mathrm{d}(\varphi) \right]^{\frac{1}{\sigma - 1}}, \tag{27}$$

$$= \left[\frac{1}{1 - G(\varphi^*_{irs})} \int_{\varphi^*_{irs}}^{\infty} \varphi^{\sigma_i - 1} g(\varphi) \mathrm{d}(\varphi) \right]^{\frac{1}{\sigma_i - 1}}, \tag{28}$$

where $\widetilde{\varphi}_{irs}$ is a CES weighted average of the firm productivity levels and the weights reflect the relative output shares of firms with different productivity levels. Note that it is also independent of the number of successfully exporting firms. Substituting φ^*_{irs} in and applying the Pareto Distribution, the average productivity is found as:

$$\widetilde{\varphi}_{irs}(\varphi^*_{irs}) = \varphi^*_{irs} \left[\frac{\gamma_i}{\gamma_i - \sigma_i + 1} \right]^{\frac{1}{\sigma_i - 1}}, \tag{29}$$

where $\gamma_i > \sigma_i - 1$. Total differentiation of equation (29) yields:

$$\widehat{\tilde{\varphi}}_{irs} = \widehat{\varphi}^*_{irs}, \tag{30}$$

where hat represents percentage change. Equation (30) shows thatchanges in average productivity is directly proportional to changes in the productivity threshold. For the domestic market ($r = s$):

```
EQUATION AVEPRODD
# AVErage PRODuctivity for the Domestic market#
(all,i,MCOMP_COMM)(all,r,REG)
    aods(i,r) = aodf(i,r);
```

For the export market ($r \neq s$),

```
EQUATION AVEPRODX
# AVErage PRODuctivity for the Export market #
(all,i,MCOMP_COMM)(all,r,REG)(all,s,REG)
    aoxs(i,r,s) = aoxf(i,r,s);
```

Now that we have the average productivity for the domestic market and the average productivity for export markets, we can calculate the average productivity for the industry as a whole. It is a weighted average of productivity in the domestic and export markets based on their relative market shares. In TABLO:

```
EQUATION AVEPRODS
# aggregate AVErage PRODuctivity of the whole Sector #
(all,i,MCOMP_COMM)(all,r,REG)
    ao(i,r) = SHRDM(i,r)* aods(i,r)
               + sum(s,REG, SHRXMD(i,r,s) * aoxs(i,r,s))
               + prodslack(i,r);
```

where SHRDM(i,r) is the share of domestic sales of industry i in region r, and SHRXMD(i,r,s) is the share of sales of industry i on the $r - s$ link with respect to total sales of the industry. According to equation AVEPRODS, aggregate productivity rises with an increase in average productivity in the domestic market, $aods(i,r)$, average productivity in the export market, $aoxs(i,r,s)$, or the respective shares of domestic and/or export markets in total sales.

If variable trade costs decline, profits of exporting firms increase and the productivity threshold for the export market falls. This enables new and less productive firms to enter the particular export market. Also, the existing exporting firms will increase their sales and thereby their market shares as a result of lower variable trade costs. The import competition in the domestic market forces less productive domestic firms to shrink or exit. Therefore, while the most productive firms increase their market shares, the least productive firms exit the market. This raises the aggregate productivity of the industry.

2.2.8 Aggregation

As in Melitz (2003), the definition of average productivity is used to obtain the aggregate variables of the model. Let us start with the aggregate price which is given by

$$P_{irs} = \left[\int_0^\infty P_{irs}(\varphi)^{1-\sigma_i} N_{irs} \mu(\varphi) \mathrm{d}(\varphi) \right]^{\frac{1}{1-\sigma}}. \tag{31}$$

This is simplified by using the relationship between the firm with productivity level φ and the average firm (see Appendix). Hence it reduces to

$$P_{irs} = N_{irs}^{\frac{1}{1-\sigma_i}} \widetilde{P}_{irs}, \tag{32}$$

where $\widetilde{P}_{irs}$ is actually the price of the representative firm that produces at the average productivity level. Similar aggregations for demand and profit yields,

$$Q_{irs} = N_{irs}^{\frac{\sigma_i}{\sigma_i - 1}} \widetilde{Q}_{irs}, \tag{33}$$

$$\Pi_{irs} = N_{irs} \widetilde{\Pi}_{irs}. \tag{34}$$

2.3 Endogenous Entry and Exit

In this section, we examine the zero profit condition and the endogenous entry and exit of firms. All the firms that produce and export on the r-s link except for the marginal firm earn positive profits. Thus, the average firm in the market makes positive profits. Each firm stays in the business as long as all the profit they make cover the sunk-entry cost of entering the market. In fact, Melitz (2003) states that the firms forgo the entry cost due to the expectation of future positive profits. Firms will enter the particular sector as long as the expected profits from each potential market exceeds the firm level entry payment flow. At the industry level the free entry/exit of firms fully exhaust the total profits such that the industry makes zero profits.

Our framework for firm entry/exit differs from Zhai (2008). Specifically, Zhai(2008) assumes that the total mass of potential firms in each sector is fixed. This simplification eliminates the role of endogenous firm entry and exit; therefore, extensive margin effects pick up only the changes in the shares of firms. In contrast, our model extends his work by incorporating endogenous firm entry and exit behavior and tracing out the direct effect of changes in the productivity threshold on entry and survival in export markets.

2.3.1 Industry Profit (Zero Profits)

Heterogeneous firms can make individual profits based on their respective productivities (marginal costs). However, at the industry level there is zero profits due to entry/exit of firms. Therefore, industry total revenue is fully exhausted by total costs. An average firm makes the following total profits:

$$\sum_s \widetilde{\Pi}_{irs} = \sum_s \left[\frac{\widetilde{Q}_{irs}\widetilde{P}_{irs}}{(1+t_{irs})} - \frac{\widetilde{Q}_{irs}C_{ir}}{\widetilde{\varphi}_{irs}} - W_{ir}F_{irs} \right] - W_{ir}H_{ir}, \tag{35}$$

where H_{ir} is the sunk-entry costs. [1] The total industry profit is just the average profit adjusted by the number of successful entrants. Using the aggregation rule in equation(34), the total industry profit is found as:

$$\Pi_{irs} = \sum_s \frac{N_{irs}\widetilde{Q}_{irs}\widetilde{P}_{irs}}{(1+t_{irs})} - \sum_s \frac{N_{irs}\widetilde{Q}_{irs}C_{ir}}{\widetilde{\varphi}_{irs}} - \sum_s N_{irs}W_{ir}F_{irs} - N_{ir}W_{ir}H_{ir}. \tag{36}$$

Hence the zero profit condition is:

$$\sum_s \frac{N_{irs}\widetilde{Q}_{irs}\widetilde{P}_{irs}}{(1+t_{irs})} = \sum_s \frac{N_{irs}\widetilde{Q}_{irs}C_{ir}}{\widetilde{\varphi}_{irs}} + \sum_s N_{irs}W_{ir}F_{irs} + N_{ir}W_{ir}H_{ir}. \tag{37}$$

Using the GTAP notation, equation (37) is corresponds to:

$$\begin{aligned} VOA(j,r) = \sum_{i \in TRAD} & VFA(i,j,r) \ + \ VAV(j,r) \\ & + \sum_{s \in REG} VAFX(j,r,s) \ + \ VAFD(j,r), \end{aligned} \tag{38}$$

[1] Note that the unit cost for the sunk-entry costs and fixed export costs is the same, W_{ir}, which is the composite price of value-added inputs. This follows from our assumption of all fixed costs are made up of primary factor costs.

where $VOA(j,r)$ is the total revenue in industry j, $\sum_{i \in TRAD} VFA(i,j,r) + VAV(j,r)$ is the total variable cost of production, $\sum_{s \in REG} VAFX(j,r,s)$ is the total fixed cost of exporting, and $VAFD(j,r)$ is the total sunk-entry costs in the industry. Following Swaminathan and Hertel (1996), we totally differentiate equation (38) and use the Envelope Theorem which yields:

$$
\begin{aligned}
VOA(j,r)\, ps(j,r) = \sum_{i \in TRAD} & VFA(i,j,r)\, pf(i,j,r) \\
& + VA(j,r)\, pva(j,r) - VAF(j,r)\, qof(j,r), \quad (39)
\end{aligned}
$$

In TABLO:

```
EQUATION MZEROPROFITS
# ZERO pure PROFIT condition for firms in the Monopolistically comp
 industry #
(all,j,MCOMP_COMM)(all,r,REG)
VOA(j,r) * ps(j,r) = VOA(j,r) * scatc(j,r)
                 - VAF(j,r) * qof(j,r) + zpislack(j,r) ;
```

The zero-profit equation captures the free-entry condition and determines the endogenous number of firms in the industry. Note that the slack variable, $zpislack(j,r)$, is introduced in this equation to allow for alternative closures. For instance, if there is no entry/exit in the industry, the number of firms in the industry is fixed. In that case, the industry profit may be positive in the short-run. This is captured by allowing the slack variable to be non-zero by endogenizing $zpislack(j,r)$ in the closure.

2.3.2 Number of Firms

This section focuses on two different free entry conditions; for the domestic market and for the export market. As mentioned in section (2.3.1) domestic free entry and exit is determined by the zero-profit condition. In fact, the zero-profit condition dictates the change in output per firm, $qof(j, s)$, which then determines the change in the number of firms in the industry through the industry output equation. This follows from Swaminathan and Hertel (1996). The total output in the industry is a product of the successfully producing firms and the output of the individual firm in the industry. It is given by the

$$Q_{ir} = N_{ir}\tilde{Q}_{ir}, \tag{40}$$

where N_{ir} is the total number of firms in the industry, $\tilde{Q}_{ir}$ is the output of the representative firm in the monopolistically competitive industry. Note that we assume symmetry in firm output. Total differentiation of 40 yields:

$$q_{ir} = n_{ir} + \tilde{q}_{ir}. \tag{41}$$

In TABLO:

```
EQUATION INDOUTPUT
# INDustry OUTPUT in the monopolistically competitive industry #
(all,j,MCOMP_COMM)(all,r,REG)
qo(j,r) = qof(j,r) + n(j,r) ;
```

According to equation $INDOUTPUT$ if output per firm rises less than the industry output, then it means that new firms enter the industry. On the other hand, if output per firm rises more than the industry output, then some firms must exit the industry.

The entry and exit of firms in the domestic market is based on the interaction between the industry and the average firm. The export market is a little different. It directly depends on what happens in the productivity threshold of the export market. The number of firms that successfully export is assumed to be given by

$$N_{irs} = N_{ir}[1 - G(\varphi^*_{irs})], \tag{42}$$

where N_{irs} is the number of firms that export product i from region r to s, and $[1-G(\varphi^*_{irs})]$ is the ex-ante probability of successfully exporting on the r-s link. This representation recognizes that not all firms in industry i are able to export on the particular r-s link. Among all the firms in the industry only the firms that pass the threshold productivity level of exporting are able to enter the export market, given the productivity distribution. This representation follows Zhai (2008). However, in our framework the number of exporting firms is endogenously determined by this equation, while Zhai (2008) assumes that the total mass of potential firms in each sector is fixed

Assuming Pareto distribution, equation (42) becomes:

$$N_{irs} = N_{ir}(\varphi^*_{irs})^{-\gamma_i}. \tag{43}$$

Total differentiation yields

$$n_{irs} = n_{ir} - \gamma_i(\varphi^*_{irs}), \tag{44}$$

where γ_i denotes the shape parameter of Pareto distribution. In TABLO:

```
EQUATION XFIRM
# number of eXporting FIRMs #
(all,i,MCOMP_COMM)(all,r,REG)(all,s,REG)
    nx(i,r,s) = n(i,r) - SHAPE(i) * aoxf(i,r,s);
```

According to equation XFIRM, if the productivity threshold for the marginal firm in the export market increases, some of the firms that did not make the cut will be forced to exit the market. This is, of course, based on the heterogeneity of the particular industry which is captured by the shape parameter of Pareto Productivity. Recall that γ_i is an inverse measure of productivity. Therefore, as γ_i increases, productivity becomes more uniform and firms become more homogeneous. In a more homogeneous industry, some firms must exit the export market given constant productivity threshold and constant mass of firms since there are more firms with similar productivity levels.

3 Data Transformation

In the monopolistic competition model imports are sourced by agent as mentioned in the previous sections. The structure of the standard GTAP database is not compatible with sourced imports. Therefore, in order to allow for sourced imports and work with the monopolistically competitive GTAP model, we need to transform the standard GTAP database. This section outlines the steps in making this transformation following Swaminathan and Hertel(1996). Given the transformed database, we turn to the additional data requirement of the firm heterogeneity model such as fixed export costs.

3.1 Data

There are three steps to generate the monopolistically competitive data base:

- Sourcing of imports valued at importer's market prices by agents
- Deriving the values of the sourced demands at agent's prices
- Deriving the trade data

We summarize each step in this section for completeness purposes. For more details, we refer the reader to Swaminathan and Hertel (1996).

3.1.1 Sourced Imports at Market Prices

In the standard GTAP database, consumption expenditure on domestic and imported goods are given separately. For instance, the private household consumption expenditure is VDPM(i,s) (on domestic goods) and VIPM(i,s) (on imported goods). The first step is to transform agents' domestic and import demands into sourced demands valued at market prices. The transformation is outlined in Figure 3.1.1. Sourcing out of aggregate imports consumed by the agent is done

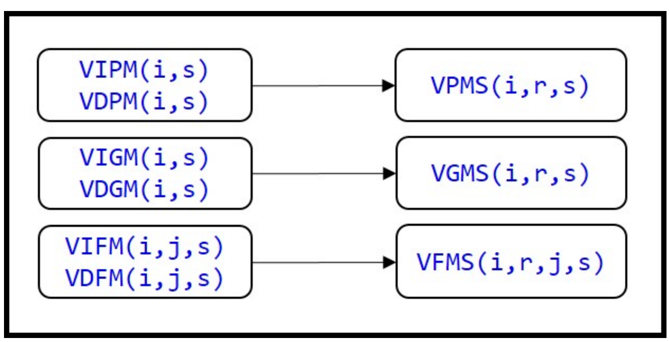

Figure 2: Sourced Imports at Market Prices

through their market share. MSHRS(i,r,s) is defined as the market share of source r in total imports of i by region s. The formula to calculate this share is as follows:

$$MSHRS(i,r,s) = \frac{VIMS(i,r,s)}{\sum_k VIMS(i,k,s)}, \tag{45}$$

where VIMS(i,r,s) is the value of imports of i by source. DATAHETV1.TAB uses the expenditure on imports and the market share to generate the expenditure for sourced imports. If the source region, r, is the same as the destination region, s, agents' purchases of domestically produced goods are also added. The same method is used for private households, government and firm intermediate input demands. An example for private household is given as follows: for $r \neq s$

$$VPMS(i,r,s) = MSHRS(i,r,s) * VIPM(i,s), \tag{46}$$

and for $r = s$

$$VPMS(i,r,s) = MSHRS(i,r,s) * VIPM(i,s) + VDPM(i,r,s). \tag{47}$$

3.1.2 Sourced Imports at Agent's Prices

The second step is to generate the sourced import demands valued at agents' prices. The transformation we need is outlined in Figure 3.1.2.

We have already obtained the sourced imports at market prices in section 3.1.1 which together with the power of the tax generates sourced import demands value at agent's prices. We define the power of the average (ad volarem) tax on total demand by an agent (TP(i,s), TG(i,s) and TF(i,j,s)). For private households it is

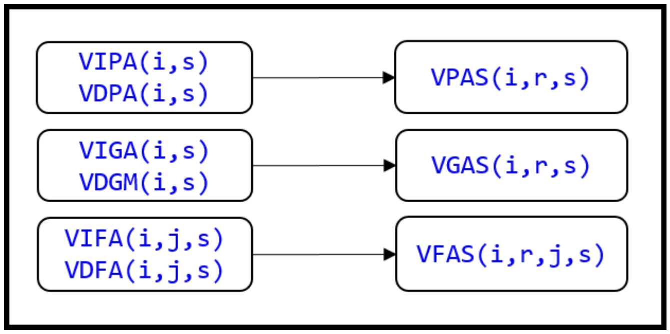

Figure 3: Sourced Imports at Agent's Prices

calculated as follows:

$$TP(i,s) = \frac{VIPA(i,s) + VDPA(i,s)}{VIPM(i,s) + VDPM(i,s)}. \tag{48}$$

The same method is used for private households, government and firm intermediate input demands. Sourced purchases at agents' prices is, then, just a product:

$$VPAS(i,r,s) = TP(i,s) * VPMS(i,r,s) \tag{49}$$

3.1.3 Trade Data

Trade data does not go through a complete change since it is already sourced. The transformation is outlined in Figure 3.1.3. There are just two differences compared to the standard GTAP trade data. The first one is the change in the notation. Exports and Imports are renamed as "sales" and "demands", respectively. The second one is the fact that for $r = s$, aggregate domestic sales are taken into account as well as the import sales due to market equilibrium. For instance, the transformation of the value of sales by destination at market prices is as follows: for $r \neq s$

$$VSMD(i,r,s) = VXMD(i,r,s), \tag{50}$$

and for $r = s$

$$VSMD(i, r, s) = VXMD(i, r, s) + VDM(i, s), \tag{51}$$

where VDM(i,s) is the value of aggregate domestic sales of i in s at market prices:

$$VDM(i, s) = VDPM(i, s) + VDGM(i, s) + \sum_j VDFM(i, j, s). \tag{52}$$

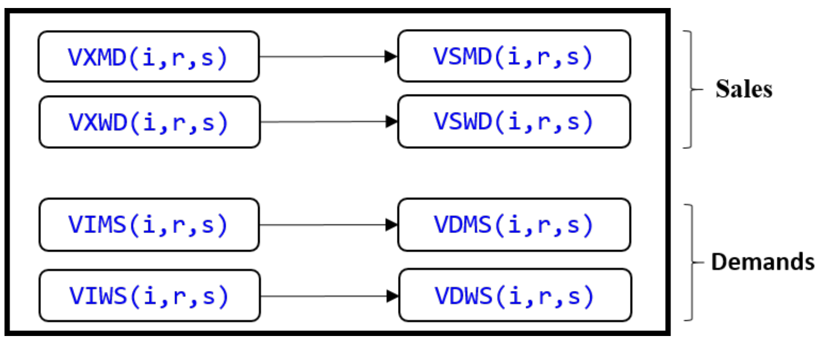

Figure 4: Trade Data

3.2 Calibration

Some of the information required in the firm heterogeneity model is not available in the GTAP database such as the elasticity of substitution between varieties, the shape parameter of Pareto distribution, data for sunk-entry costs and fixed exporting cost. We take the elasticity of substitution between varieties (σ_i), and the shape parameter of Pareto distribution (γ_i) from the literature. On the other hand, we calibrate total fixed costs and fixed export costs. For the calibration of total fixed costs we follow Swaminathan and Hertel (1996), while for the calibration of fixed export costs we follow Zhai (2008).

As explained before value added costs are composed of a fixed, VAF(i,r), and a variable, VAV(i,r), part. Initial value for the fixed value-added is calibrated by using mark-up pricing and market clearing equation. It follows that the share of fixed costs in total costs is given by the following formula:

```
formula (all,i,MCOMP_COMM)(all,r,REG)
VAF(i,r) = VOA(i,r) * {1 - [1 / MARKUP(i,r)]};
```

where $\frac{1}{\sigma_i}$ portion of total costs is fixed costs. The rest of the value-added costs are attributed to variable value-added in the following fashion:

```
formula (all,i,MCOMP_COMM)(all,r,REG)
VAV(i,r) = VA(i,r) - VAF(i,r);
```

This follows directly from Swaminathan and Hertel (1996). Recall that in the firm heterogeneity model fixed value-added costs, VAF(i,r), is split into the sunk-entry costs, VAFD(i,r), and fixed export costs, VAFX(i,r,s). The initial value of the fixed export costs is calibrated to the base year bilateral trade flows following Zhai (2008). Solving the bilateral trade flow equation for fixed export costs yields:

$$N_{irs}F_{irs} = \frac{P_{irs}Q_{irs}}{(1+t_{irs})}\frac{\gamma_i - \sigma_i + 1}{\sigma_i\gamma_i}. \tag{53}$$

This caibration appears in the code as:

```
FORMULA (initial)(all,i,MCOMP_COMM)(all,r,REG)(all,s,REG)
VAFX(i,r,s) = [1 - DELTA(r,s)]
            * VSMD(i,r,s) * {1 - [1 / (MARKUP(i,r)-1)]
            * [1/SHAPE(i)]} ;
```

where DELTA(r,s) is, again, the Kronecker delta which is equal to one when $r = s$. Once fixed export costs are calibrated, sunk-entry cost is a residual in total fixed costs. It appears in the code as follows:

```
FORMULA (all,i,MCOMP_COMM)(all,r,REG)
VAFD(i,r) = VAF(i,r) - sum(s,REG,VAFX(i,r,s));
```

4 Policy Application

In order to illustrate the usefulness of this model, we aggregate the GTAP v.8 data base to 2 commodities (manufacturing, and non-manufacturing) and 3 regions (USA, Japan, and ROW), wherein the manufacturing sector is treated as monopolistically competitive with heterogeneous firms, and the non-manufacturing sector is treated as perfectly competitive/Armington. The aggregation adopted is the same as in Swaminathan and Hertel(1996) for comparison purposes. The program requires minor changes to run with a different aggregation. We eliminate Japanese tariffs on the import of US manufacturing goods to illustrate the workings of our model.

The direct effect of the tariff liberalization policy is a reduction in the price of US manufactures in the Japanese market, an increase in the quantity of US manufactures sales to Japan, a decrease in domestic sales and in the sales to the ROW – results which are similar to those from the monopolistic competition and firm heterogeneity models. A unique aspect of the firm heterogeneity model is its ability to capture the impact of trade liberalization on marginal firms for each source-destination link. In our example, the Japanese tariff cut reduces the productivity threshold for successfully exporting on the US-Japan link (-1.6%). With the lower tariff, US exporters obtain a larger market in Japan and sales to Japan rise (22.6%). This reduces fixed export costs per unit of exports, capturing scale effects in the export market and allowing the firms with lower productivity to profitably export on the US-Japan link. Therefore, new US varieties enter the Japanese market (13%). The productivity threshold for exports on the Japan-US link also declines (-0.1%), but mostly due to the reduction in average variable costs in Japan, as the cost of intermediate inputs falls.

The within industry firm reallocation extends to the domestic markets. The productivity threshold of producing for the domestic market increases for the US (0.03%) and Japan (0.1%). The rise in average variable cost and the sunk entry cost per domestic sale makes it costly for the US firms to produce in the manufacturing sector and reduces their profits; therefore, low-productivity firms exit the industry (-0.04%). In contrast, the domestic productivity threshold increase in Japan is driven by the decline in Japanese prices due to intensified competition. Trade liberalization, therefore, reallocates market share by shifting resources towards more productive firms improving the aggregate productivity in the US (0.02%) and Japan (0.04%). We also compare the change in the terms of trade (TOT) across models. Standard GTAP model results demonstrate that there is a shift in favor of the US, whereas TOT decline for Japan. But, the experiments in monopolistic competition and firm heterogeneity models result in a smaller improvement in the US TOT and magnify the decline in that of Japan.

5 Conclusions and Future Work

In this paper we discuss how to implement monopolistic competition with firm heterogeneity into the GTAP model. Different from the standard GTAP model with Armington specification, the firm heterogeneity module includes the effect of new varieties in markets (extensive margin), the effect of scale economies, and the effect of endogenous productivity. We build on Zhai (2008) for firm heterogeneity; however, compared to his approach we incorporate endogenous firm entry/exit, and we distinguish between sunk-entry costs, and fixed export costs.

The model is calibrated to GTAP (version 8.0) database. There are three pieces of information not contained in the GTAP database that are needed in firm heterogeneity approach: (i) the elasticity of substitution between varieties, (ii) the shape parameter of the Pareto productivity distributions, and (iii) the magnitude of fixed export costs. We adopt Zhai's (2008) approach to calibrate fixed export costs based on a gravity model of trade using bilateral trade flows. The elasticity of substitution between varieties, and shape parameter of the productivity distribution is taken from the literature. Model results in firm heterogeneity module depends on the choice of substitution elasticity and shape parameter. For future work, we aim to combine econometric work on model parameters with policy analysis to obtain more robust results.

To illustrate the behavioral characteristics of the model, we analyze the effects of eliminating Japanese tariffs on the import of US manufacturing goods under a three region - two sector aggregation. This is a highly stylized FTA scenario in TPP with the aim of laying out the mechanics of this Melitz-type GTAP model. We observe that productivity threshold for the US-Japan export market reduces mostly due to the reduction in fixed export costs per sale. This scale effect is the dominant factor in threshold reduction and a subsequent increase in the number of US manufacturing firms exporting in Japanese markets. This firm reallocation in US-Japan link is in favor of lower-productivity firms. On the other hand, the within firm reallocation in the domestic market is such that low-productivity firms are forced to exit due to higher average variable costs. As a result of exit of firms in the domestic market, the productivity of US manufacturing sector rises.

By incorporating monopolistic competition and firm heterogeneity, we are able to capture and analyze the previously unobserved effects of trade agreements. The question to ask at this point is whether these effects matter for trade policy implications. In other words, do we expand our insight by using a Melitz-type GTAP framework over the traditional model with Armington assumption? An initial comparison of model responses to tariff elimination across GTAP models with Armington, Krugman, and Melitz specifications do not show significant variation when the policy instrument is the tariff rate.

The main premise of the TPP agreement is to develop comprehensive, high-quality rules in trade that harmonize standards and thereby reduce barriers to trade. The variation in trade standards across regions force firms to incur significant fixed export costs. Reduction in these costs are expected to generate huge gains for the member countries. As a future work we aim to analyze a more comprehensive TPP scenario with fixed export costs as the policy instrument. The GTAP model with firm heterogeneity responds to fixed export cost reductions by changing industry productivity as a result of shifts in the export market productivity thresholds which is unique to the firm heterogeneity framework.

A second goal in our future work is to discuss the welfare implications of TPP under this new framework. There are three additional channels through which trade liberalization yields welfare gains in the firm heterogeneity GTAP model; namely, variety, scale, and productivity effects. We will focus on the welfare response of the GTAP model to different policy instruments such as tariff, and fixed export costs under Armington, Krugman and Melitz specifications.

References

Balistreri, E. J., R. H. Hillberry, and T. F. Rutherford (2011): "Structural estimation and solution of international trade models with heterogeneous firms," *Journal of International Economics*, 83, 95–108.

Balistreri, E. J. and T. F. Rutherford (2012): "Computing General Equilibrium Theories of Monopolistic Competition and Heterogeneous Firms," in *Handbook of Computable General Equilibrium Modeling*, ed. by P. B. Dixon and D. W. Jorgenson.

Bernard, A. B., J. Eaton, J. B. Jensen, and S. Kortum (2003): "Plants and productivity in international trade," *American Economic Review*, 93, 1268–1290.

Bernard, A. B. and J. B. Jensen (1999): "Exceptional exporter performance: cause, effect, or both?" *Journal of International Economics*, 47, 1–25.

Bernard, A. B., J. B. Jensen, S. J. Redding, and P. K. Schott (2007): "Firms in international trade," *Journal of Economic Perspectives*, 21, 105–130.

Bernard, A. B., J. B. Jensen, and P. K. Schott (2006): "Trade costs, firms and productivity," *Journal of Monetary Economics*, 53, 917–937.

Chaney, T. (2008): "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *American Economic Review*, 98, 1707–1721.

Dixon, P. B. and M. T. Rimmer (2011): "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, 150–154.

Gohin, A. and T. W. Hertel (2003): "A Note on the CES Functional Form and Its Use in the GTAP Model," GTAP Research Memorandum.

Hillberry, R. H., M. A. Anderson, E. J. Balistreri, and A. K. Fox (2005): "Taste Parameters as Model Residuals: Assessing the 'Fit' of an Armington Trade Model," *Review of International Economics*, 13, 973–984.

Hopenhayn, H. (1992): "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica*, 60, 1127–1150.

Huff, K. M. and T. W. Hertel (2000): "Decomposing Welfare CChange in the GTAP Model," GTAP Technical Paper No. 5.

Hummels, D. and P. J. Klenow (2005): "The variety and quality of a nation's exports," *American Economic Review*, 95, 704–723.

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, 1695–1725.

Melitz, M. J. and D. Trefler (2012): "Gains from Trade when Firms Matter," *Journal of Economic Perspectives*, 26, 91–118.

Oyamada, K. (2013): "Calibration and Behavioral Characteristics of Applied General Equilibrium Models with Flexible Trade Specifications Based on the Armington, Krugman, and Melitz Models," .

Petri, P. A., M. G. Plummer, and F. Zhai (2012): *The Trans-Pacific Partnership and Asia-Pacific Integration: A Quantitative Assesment*, Peterson Institute for International Economics.

Swaminathan, P. and T. W. Hertel (1996): "Introducing Monopolistic Competition into the GTAP Model," GTAP Technical Paper No. 6.

Zhai, F. (2008): "Armington Meets Melitz: Introducing Firm Heterogeneity in a Global CGE Model of Trade," *Journal of Economic Integration*, 23, 575–604.