DOCUMENTOS DE ECONOMÍA Y FINANZAS INTERNACIONALES

DEFI 07-09
December 2007

Trade growth in a heterogeneous firm model: Evidence from South Eastern Europe

d'Artis Kancs

Asociación Española de Economía y Finanzas Internacionales
www.aeefi.com
ISSN: 1696-6376

Trade Growth in a Heterogeneous Firm Model: Evidence from South Eastern Europe*

d'Artis Kancs†

2007

Abstract

In 2007 a Free Trade Area (BFTA) will be created in the Balkans. In this paper we study the potential impact of BFTA on trade growth in the SEE. Given that welfare impacts associated with trade growth depend on the growth channels, more goods and varieties exported or at higher price or higher volume of goods and varieties are exported, in this paper we investigate the structure of integration-induced export growth in the Balkans. The empirical implementation of our analysis is complicated by the fact that firm-level trade data is not available for the SEE economies. In order to cope with this data paucity, we adopt a heterogeneous firm framework, which allows us to decompose the aggregate trade growth in two parts: the intensive margin of trade and the extensive margin of trade using only aggregate trade data. The empirical findings of our study suggest that the BFTA would primarily trigger trade growth through a growing number of exported goods (the extensive margin of trade). Thus, the actual welfare gains from trade growth in the Balkans might be larger than predicted by previous trade studies. We also found that a variable trade cost reduction would lead to higher export growth rates compared to a fixed trade cost reduction. These results allow us to draw detailed policy conclusions.

Keywords: Balkans, export growth, regional integration, trade costs.
JEL classification: F12, F14, R12, R23.

1 Introduction

Since the early 1990s, the countries of South Eastern Europe (SEE)[1] have been reforming their centrally planned economies to be more market oriented. In contrast

*The author acknowledges helpful comments from Thomas Chaney, Ilze Kalnina and Stephen Redding as well as seminar participants at LSE and CEPR.

†London School of Economics and Catholic University of Leuven. E-mail: kancs@lse.ac.uk.

[1] Albania, Bosnia & Herzegovina, Bulgaria, Croatia, Macedonia, Moldova, Romania, Serbia & Montenegro.

to the Central and Eastern European (CEE) economies, the transition process in the Balkan peninsula has been considerably delayed and complicated by internal uprisings and the outbreak of civil wars in the beginning of the nineties. The civil wars in the Balkans began in 1991 and lasted for almost a decade.

The Balkan wars were characterised by armed conflicts between different ethnic groups of the former Yugoslavia (*Stubos* and *Tsikripis* 2007). According to *Stubos* and *Tsikripis*, the armed conflicts in the Balkans had their roots not only in the historical-cultural and religious tensions, but also in socio-economic problems. As a result, in addition to the humanitarian tragedy, the civil wars in the Balkans brought also a major deterioration in the SEE economic performance. The development of the SEE foreign trade was determined by collapse of the Yugoslav internal market and withering of the Socialist bloc external markets (*Stubos* and *Tsikripis* 2007). In addition, the political independence movements of the former Yugoslav republics set loose a wave of protectionism in the newly established states in the Balkans. All these factors together led to a sharp decline in foreign trade openness during the years of the Balkan wars, which resulted in rapidly declining foreign trade volumes. For example, the share of external trade in the SEE's GDP declined from 93% in 1990 to 54% in 1995 (*Eurostat* 2005).[2]

The Balkan wars ended in 1999 with much of the SEE economies reduced to poverty and economic disruption.[3] At the same time, the end of the civil wars created favourable circumstances for a new attempt of building social, economic and political stability in the Balkans. For example, the growing political stability, together with trade policy liberalisation measures, created favourable circumstances for foreign trade and foreign direct investment in the Balkans. As a result, in the post-war years the SEE external trade increased rapidly - 75% between 1994 and 2004 (*Eurostat* 2005).

Although, trade policy liberalisation was an important factor, which significantly contributed to foreign trade growth in the post-war period, these extraordinary high growth rates in Balkan foreign trade cannot solely be associated with the trade policy liberalisation measures. Other factors, such as the end of the civil wars in the Balkans, an improvement in the implementation and application of laws, a decrease in corruption, better management of basic public infrastructures and institutions, have also contributed to better functioning of markets and, hence, to trade growth. In this study we refer to all these factors together as trade freeness à la *Head* and *Mayer* (2004), which according to Figure 1 in section 4.3, have considerably increased since the end of the civil wars in the Balkans.

Considering the SEE's foreign trade policy, which significantly contributed to in-

[2]The external trade and foreign direct investment levels were also rather low relative to those of CEE economies. Whereas the share of external trade in SEE's GDP was 54% in 1995, the share in CEE's GDP was around 72% in 1995 (*Eurostat* 2005).

[3]Several armed conflicts have outbreak since the end of the civil wars in the Balkans, for example, the Presevo Rebellion (2000-2001) and the Albanian Uprising in Macedonia (2001).

creased trade freeness in the Balkans since the end of the civil wars, we can distinguish two phases: (i) bilateral trade agreements (1999-2006); and (ii) a single free trade area in the Balkans (2007-). In the context of the present study we can call the first phase the bilateral phase and the second phase the FTA phase.

The first (bilateral) phase was commenced in 1999, right after end of the civil wars in the Balkans, when eight SEE economies signed the Stability Pact for South Eastern Europe. The SEE Stability Pact was an expression of strengthening efforts to foster peace, democracy, respect for human rights and economic prosperity. Among other things, the Pact provides a framework for stimulating regional co-operation and economic integration. The SEE Stability Pact became operational in June 2001,[4] when the SEE countries agreed to implement bilateral Free Trade Agreements in order to develop their mutual trade and promote economic integration in the region. During the bilateral phase of trade policy liberalisation in the Balkans, the eight SEE economies have concluded a series of additional bilateral Free Trade Agreements with a goal of expanding regional trade and thereby promoting growth, investment and employment in the Balkan region. As a result, at the end of the bilateral phase there was a network of 31 bilateral Free Trade Agreements.

Based on these findings and the *Eurostat* (2005) data we can summarise the bilateral phase as follows: (i) since the end of the civil wars in the Balkans, trade freeness has increased considerably; (ii) foreign trade policy mainly consisted of bilateral free trade agreements; (iii) increased trade freeness induced sizeable trade growth in the Balkans; (iv) both types of trade flows are observed in the SEE trade data: positive for some products and some country pairs and zero trade flows for other products and other country pairs; (v) the empirical trade data for the SEE provide a strong evidence of intra-industry trade.

Preparations for the second (FTA) phase started in June 2005, when the SEE economies have agreed to work towards transforming the current network of bilateral Free Trade Agreements into a single regional free trade agreement. In July 2006 the SEE countries decided that the Balkan Free Trade Agreement will be fully implemented in 2007. The BFTA should simultaneously enlarge and amend the Central European Free Trade Agreement (CEFTA) to include all SEE economies and update it by including effective trade provisions.

Overall, the FTA phase is much less researched and its impacts are still largely unknown. With respect to the proposed BFTA, many questions arise. Will the SEE foreign trade continue to grow after the BFTA? What will be the impact of BFTA on trade growth rates? How exactly the trade growth will occur - will trade liberalisation increase the export volume and value of existing firms or give incentives for more firms to enter foreign markets and start exporting? Given that empirical knowledge about BFTA impact on trade in the Balkans is in an early stage, in this study we are interested in this second - FTA phase. The main goal of this paper is

[4] Albania, Bosnia & Herzegovina, Bulgaria, Croatia, Macedonia, Romania, Serbia & Montenegro. Moldova joined the SEE liberalisation process in 2002.

to analyse the BFTA-induced export growth in the SEE.

The paper is structured as follows. After providing an overview of historical development patterns and institutional settings in the SEE countries, we outline the evolution of foreign trade among SEE economies in the Balkans. Next, we review the most recent findings of international trade literature, where an increasing number of empirical studies are interested not only in the volume but also in the structure of trade growth. In section 3 we introduce the theoretical framework, which will be used in the empirical analysis. Our analysis is based on a monopolistic competition trade model, which allows to analyse the structure of trade growth based only on aggregate trade data. The second step in our analysis involves parameter estimation. This is done by deriving empirically estimable equations of trade flows and trade costs from the theoretical trade model and estimating the two equations. The resulting estimates provide numerical values of structural parameters for the underlying theoretical trade model. Finally, in section 5 we use the estimated parameters and statistical data to empirically implement the theoretical model and to make detailed predictions about the structure of export growth induced by the BFTA. Based on these results in section 6 we draw policy conclusions and sketch avenues for future research.

2 Empirical evidence

In this section we review previous empirical and theoretical trade literature, which offers useful insights for empirical analysis of trade growth. This allows us to establish several important findings about international trade flows. First, according to recent empirical research on the structure of commodities trade, aggregate trade flows are composed of several components and, with detailed enough data, changes in each of these components can be separately traced. Second, firm-level trade studies have identified several notable features of exporters that are overlooked in international macroeconomic literature, but might be relevant for the present study. These findings are relevant for selecting an appropriate theoretical framework for the present study.

2.1 Defining trade flows and trade growth

The trade literature uses a sizeable number of trade-specific terms, such as trade value, trade volume, trade flows, trade growth, intensive and extensive margins of trade. Sometimes, different terms are used to describe the same economic variables or growth processes, and vice-versa. In order to facilitate the understanding of our analysis and comparison with other studies, we start the literature review by defining key terms, which we consistently use throughout the paper. Given that in this study we focus on export flows and export growth, we define key terms used in export literature. The structure of import flows can be defined and analysed analogously.

Usually in empirical trade literature (e.g. *Anderson* and *Wincoop* 2003) the total value of exports, E_{od}, from exporting country o to importing country d is defined as the number of shipments, N_{od}, times the average value per shipment, e_{od}:

$$E_{od} = e_{od} \times N_{od} \tag{1}$$

Hummels and *Klenow* (2005) show that the decomposition of aggregate trade growth can be performed along the same lines, i.e., trade growth can be decomposed into two components. First, the incumbent firms adjust their volume of exports, i.e. the intensive margin measures trade growth within product lines. According to *Hummels* and *Klenow* (2005) terminology, the intensive margin, e_{od}, captures the trade growth within product lines.[5] Second, the number of traded varieties (traded goods) change (increase). Analogously, *Hummels* and *Klenow* label the trade growth that can be attributed to a larger number of traded varieties, N_{od}, as the extensive margin of trade.

In the context of empirical studies, which we review in the following section, we also need to define the ratios of firms. The exporter ratio is defined as the ratio of exporters over the number of all firms. The starter ratio is defined as the ratio of new firms, which enter foreign markets and start export to export, over the number of all exporting firms. Analogously, the stopper ratio is defined as the ratio of exporters that switch to non-exporter status to the number of last period exporters. Given that the starter and stopper ratios can be considered as the transition probabilities in export status, a higher starter ratio than stopper ratio does not necessarily mean that the exporter ratio is increasing all the time. Unless explicitly mentioned, all ratios are measured per year.

2.2 Empirical evidence: the dual margin

Early trade literature looked mainly at aggregate trade flows and trade growth. In more recent years the identification of the two components of aggregate trade flows (extensive and intensive margins) have attracted a considerable amount of research attention. Different strands of the international trade literature using different theoretical frameworks have all contributed. Identification of the intensive trade margin has most frequently been studied in plant-level trade studies. The extensive margin of trade has attracted a considerable amount of research attention in studies analysing differences between exporting firms (exporters) and local sellers (non-exporters). In this section we review main findings of the most recent international trade literature with a view of finding an appropriate analytical framework for the present study.

Several recent empirical trade studies using firm-level trade data report that the sets of exporters, goods and sectors change over time and vary more than has traditionally been assumed. For example, using annual data of Colombian manufacturing

[5] The intensive margin can further be decomposed into quantity changes or changes in the number of units traded, x_{od}, and price changes or changes in the average price of the traded units, p_{od}. In such case the trade growth can be decomposed into three components ($E_{od} = x_{od} \times p_{od} \times N_{od}$). However, in the context of our study, only equation (1) is relevant.

Census 1981 - 1989 *Roberts* and *Tybout* (1997) find that on average, the starter and stopper ratios are about 3.3% and 11.5%, respectively with the average exporter ratio of 11.8%. Similarly, *Bernard* and *Wagner* (1998) use annual manufacturing plant-level data 1978 - 1992 in Lower Saxony, Germany and find that, on average, ratios of firms entering and exiting exporting were about 4.14% and 5.51%, respectively with the average exporter ratio of 41.2%.

Subsequent research by *Bernard* and *Jensen* (1999) suggests that exporters differ in the variety of goods that they trade and also in the range of countries they trade with. Using detailed data from individual plants for the entire US manufacturing sector *Bernard* and *Jensen* decompose sources of the US export boom in the late 1980s and early 1990s. They find that the preponderance of the increase in exports came from increasing export intensity at firms that were already exporting, but a non-negligible share came from firms that switched between only selling locally to selling both locally and abroad. This finding again indicates that a sizable share of new trade is in the form of new goods not previously traded.

Using the U.S. Census Bureau's Annual Survey of Manufacturers (ASM) 1986 - 1992 *Bernard et al* (2003) show that the ratio of non-exporters that transition to exporter status to the number of last period non-exporters, which they term the starter ratio, is about 14.4% per year. The ratio of exporters that switch to non-exporter status to the number of last period exporters, the stopper ratio, is about 12.2% on average per year. The dynamics of export status result in changes in the ratio of exporters among all firms over time. The starter ratio is slightly higher than stopper ratio and the average exporter ratio is about 51.8%. These results again indicate the prevalence of the extensive margin of trade.

Kehoe and *Ruhl* (2002) have obtained similar results in a somewhat different setting. *Kehoe* and *Ruhl* provide one of the most detailed analysis of the changing extensive margin in the wake of bilateral trade integration by studying trade integrations in 18 countries and show how substantial increases in the extensive margin coincide with trade integration. They investigate the importance of the extensive and intensive margins in six major trade integration periods: the accession of Greece, Spain and Portugal to the European Community, the Canada - USA Free Trade Agreement (CUSTA), the implementation of the Single Market Programme and the North American Free Trade Agreement (NAFTA). Using detailed data on international trade flows by commodity, they find significant evidence of trade adjustments through the extensive margin. *Kehoe* and *Ruhl* (2002) also find that the initially 'least traded' product categories experienced the largest increases in export shares following trade integration, which is a strong evidence of the extensive margin of trade.

Hillberry and *McDaniel* (2002), using an alternative measure, find evidence of a smaller, but still significant extensive margin growth for the United States following the implementation of the NAFTA. They estimate that Mexican exports to the US grew by $86 billion between 1993 and 2001, of which 12.5% is attributable to a larger extensive margin of trade. Correspondingly, Mexican imports from the US grew by $44 billion, of which 9.7 percent occurred at the extensive margin. Their lower estimates in comparison to *Kehoe* and *Ruhl* (2002) can likely be attributed to their different metric: *Hillberry* and *McDaniel* (2002) define a product category as traded even if actual exports, though positive, are virtually insignificant. In addition, they focus on the effects of NAFTA per se and not on trade integration undertaken by Mexico in earlier years.

Ruhl (2003) shows how permanent tariff reductions, as opposed to temporary business cycle shocks, affect firms' decisions to export. He finds that tariff reductions increase the extensive margin as new firms enter export markets. In a calibrated model *Ruhl* shows how the failure to account for these new goods produces upward biased aggregate elasticities of exports with respect to tariffs.

The empirical results for Europe are similar, although, scarcer. *Eaton et al* (2004) adopt the *Melitz*'s (2003) model to study French exports using firm level-trade data. They decompose French aggregate export flows based on data for individual shipments. While *Eaton et al* take the size of importers as given, they analyse how aggregate trade varies for a change in the importer's size and a change in trade costs. They find that a model with heterogeneous firms that gives rise to variable extensive and intensive margins is a reasonably accurate description of actual French trade patterns. The authors show that variations in aggregate French exports are mostly due to a change in the number of firms, which export to foreign markets. However, the dominance of the extensive margin is most visible, when the variation of aggregate trade flows is due to a change in trade costs, for given market sizes of destination countries. *Eaton et al* also analyse the decomposition of trade growth at the industry level and find that aggregate features emphasising the prevalence of the extensive margin do not differ significantly across sectors.

Hillberry and *Hummels* (2005) is one of the few studies which analyse the structure of intra-national trade flows. They investigate how U.S. domestic trade flows vary with distance, using the 1997 Commodity Flow Survey. *Hillberry* and *Hummels* show that distance reduces aggregate domestic flows mostly through a reduction in the number of trade flows: at the sample mean distance, the extensive margin represents 62% of the elasticity of aggregate trade flows with respect to distance.

Hummels and *Klenow* (2005) study the response of the intensive and extensive margins on country-level trade, relying on a definition of the margins based on the variation of exporting countries' sizes. They analyse exports in 1995 from 110 countries to 59 importers and decompose the greater trade of larger economies into contributions from the intensive and extensive margins of trade. In addition, they compare prices and quantities of exports by different countries to given market-categories and estimate quality differences across exporters. The main finding of *Hummels* and *Klenow* is that the extensive margin accounts for two-thirds of greater exports of larger economies, and one-third of greater imports of smaller economies. For both imports and exports, larger economies trade in more categories and trade with more partners. Richer countries export more units at higher prices, therefore producing a higher quality, and exporting mainly at the 'quality' (extensive) margin.

2.3 Empirical evidence: the heterogeneity of firms

The second key finding about the international trade flows which results from previous firm-level studies concerns heterogeneity of exporting firms. The main finding of these studies is that only few firms export, and among exporters, only few firms export to more than a few countries. Most exporters only sell a small fraction of their output abroad. These results are in sharp contrast to gravity models with homogenous firms, where every firm sells in every region/country.

Another finding of the firm-level studies is that exporters are different from non exporters, moreover, they are different in many respects. Usually, they are much larger and are more productive as well as more capital intensive than firms selling all output locally. Several studies have also found that having exported in the past significantly increases the probability of a firm exporting today. *Bernard* and *Jensen* (1999), for example, found that a firm exporting today is 36% more likely to export in the future than a firm not exporting today. This result imply that exporting and non-exporting firms are much more heterogeneous than used to assume in representative firm models.

The third finding, which is relevant for our study, is the evidence of export entry costs. Several firm-level studies found a significant evidence for the presence of sunk costs associated with exporting. *Bernard* and *Jensen* (1999) and *Bernard et al* (2003) all report substantial evidence of fixed entry costs into foreign markets. By accounting for fixed trade costs *Evenett* and *Venables* (2002) could explain the many zeros (non-traded varieties) in bilateral trade data. *Evenett* and *Venables* also document that the number of non-traded varieties has substantially dropped over time. This suggests that a reduction in fixed costs or growth of income can play an important role in accounting for the growth of world trade. *Evenett* and *Venables* (2002) find that the removal of zeros accounts for one third of developing countries' export growth since 1970. These empirical findings highlight the relevance of fixed costs associated with exporting.

Findings of previous trade studies discussed in sections 2.2 and 2.3 can be summarised as follows: (i) aggregate trade growth usually occurs through two channels: previously non-traded goods become traded and new firms start exporting; and existing exporters increase their export volume of goods and varieties already exported; (ii) those producers that export their goods abroad significantly differ from non-exporters (higher productivity, higher levels of output, and more capital intensive); and (iii) entering export markets is associated with entry costs, which are sunk after entry decision.

3 Theoretical framework

There are several methodological approaches for decomposing aggregate trade flows. The most straightforward is to use firm-level data for prices, quantities and the number of shipments. Unfortunately, such data is not available for the SEE post-war economies in the Balkans. Thus, this approach is not suitable for the SEE. Data limitations in the SEE require an analytical approach which would allow us to infer differential changes in the extensive and intensive margins in the pattern of trade growth using only aggregate trade volume data. In fact, given that we are interested in decomposing the SEE bilateral trade flows in to only two components (extensive and intensive margins), we need to identify either one. The other trade margin could then be calculated as a residual from aggregate trade flows, which are available in the SEE trade data.

Identification of the intensive trade margin is rather involved, as we would need data for both prices and quantities of exported goods. Identification of the extensive trade margin requires information about the number of traded varieties, which for most manufactured goods is equal to the number of exporting firms. If all firms in country o would export to country d, we could use the *Krugman*'s (1980) monopolistic competition trade model. However, according to empirical trade data for SEE, both types of trade flows are observed in the Balkans: positive trade flows for some products and some country pairs and zero trade flows for other products and other country pairs (*Eurostat* 2005). Thus, the *Krugman*'s (1980) model is not suitable for SEE. A data-undemanding analysis of trade structure in the Balkans requires an approach, where exporting firms is a subset of the total number of firms in country o, which is both sector and destination-specific.

Melitz (2003) extended the *Krugman*'s (1980) model by assuming that firms in country o are heterogenous according to their productivity and only the most productive ones export to country d. Separation of exporting firms from non-exporters à la *Melitz*'s (2003) requires statistical data for firm distribution and information about the threshold productivity above which firms export to d. The firm productivity distribution is set arbitrarily, usually it is assumed to be Pareto (*Melitz* 2003). The threshold productivity (exporting threshold) can be determined from the export entry cost.

The *Melitz*'s (2003) model would allow us to identify the number of exporting firms from country o to country d (extensive margin), from the aggregate trade data. The downside of the *Melitz* (2003) approach is that it requires more parameters and several additional assumptions about firm heterogeneity, firm distribution and export entry costs. The first assumption of productivity heterogeneity of firms finds indeed strong evidence in firm-level data, which we discussed in the previous two sections. In particular, firm-level empirical evidence suggests that those producers that export their goods abroad differ from non-exporters along several dimensions: exporters tend to have higher productivity, higher levels of output, and use more capital and labour inputs. The second assumption about distribution of firm heterogeneity is little researched and, therefore, is subject to sensitivity analysis. The third assumption of export entry cost finds strong support in the empirical trade data. All studies discussed in the previous two sections find strong evidence of fixed market entry costs associated with exporting abroad.

In this section we introduce the theoretical framework of our study, which is largely based on the *Melitz*'s (2003) model,[6] which in turn is an extension of the *Krugman*'s (1980) model of trade with monopolistic competition and increasing returns. We start by introducing the basic ingredients of *Melitz*'s model: defining preferences and technologies and characterising the optimal strategies of both firms and consumers in partial equilibrium. Next, by determining the selection of firms into local producers and exporters, we are able to compute the global general equilibrium. As in *Melitz* (2003), the selection among exporters and non-exporters is based on the assumptions that firms are heterogeneous and exporters face fixed costs associated with entering foreign markets, implying that less productive firms are not able to generate enough revenue abroad to cover the fixed costs of entering foreign markets. Thus, according to the *Melitz*'s (2003) model, exporters are only a subset of domestic firms and this subset of exporters varies with characteristics of foreign markets. This type of sorting mechanism is indeed in line with empirical findings, which we have established in the previous section, i.e. exporters are more profitable than non-exporters. We may conclude that the *Melitz*'s (2003) model can be applied for studying trade growth in the SEE transition economies under reasonable assumptions.

As in the *Melitz*'s (2003) model there are R countries that produce goods using only labour. Country r has a total labour force L_r. All countries have access to the same technologies. There are two types of sectors: one traditional sector, A, and one manufacturing industry, X. Given the trade focus of our analysis we assume that all manufacturing goods can be traded among all countries. The 'traditional' sector produces a homogenous 'traditional' good under perfect competition, constant returns to scale with unit labour requirement. As usual, the 'traditional' sector is immobile, and the 'traditional' good is assumed to be traded freely at zero trade cost. It serves as a numeraire in our model, therefore, its price is normalised to 1. Given that every country produces the homogenous good and the homogenous good is set as a numeraire, wages are equalised to unity in every country.

The manufacturing industry supplies a continuum of differentiated goods and, as usual in monopolistic competition models, each firm is a monopolist for the variety it produces. Manufacturing goods face positive trade costs. As in *Melitz* (2003), we assume two types of trade costs: variable trade cost and fixed trade cost. In contrast to *Melitz*'s model, which assumes that a firm has first to pay a fixed cost to survive at home and then it has to pay a fixed cost for entering export markets, we assume that all firms have to pay only one fixed cost for entering any market. This adjustment, which considerably reduces the fixed cost data requirements, is required to make

[6] The present model also incorporates features of *Chaney* (2007) and *Helpman et al* (2007).

the empirical implementation of the model feasible in the SEE transition economies, where no comparable fixed cost data is available. Although, the two different entry costs in the *Melitz*'s (2003) model might more precisely describe firms dynamics, they offer little additional insights in the behaviour of exporters, which is the main focus of the present study.

3.1 Preferences and technology

We start the formal description of the model with consumer preferences. We assume that the produced goods are consumed by workers, which are the only consumers. All consumers have identical CES preferences over traditional and manufacturing goods. A consumer that consumes C_A units of the homogenous good, x_j units of each variety j of the manufacturing good, and N varieties of the differentiated manufacturing good achieves total utility U:

$$U = C_A^{\alpha_A} \left(\int_0^N (x_j)^{\frac{\sigma-1}{\sigma}} dx \right)^{\frac{\sigma}{\sigma-1}\alpha_x} \tag{2}$$

where σ_x is the elasticity of substitution between manufacturing varieties and α is consumer demand parameter determining expenditure shares, with $\sigma_x > 1$ and $\alpha_A + \alpha_x = 1$.

There are two types of trade costs for shipping manufacturing goods from origin country o and selling in destination country d: variable trade cost, τ_{od} and fixed trade cost, FC_{od}. The variable trade cost are of 'iceberg' form: if one unit of the differentiated manufacturing good is shipped from origin country o to destination country d, only fraction $\frac{1}{\tau_{od}}$ arrives at d. Following *Samuelson* (1954), we assume that the rest melts on the way. The higher is τ, the higher is the variable trade cost. The second type of trade cost manufacturing firms face are export entry costs, which do not depend on the quantity sold abroad. If a firm in country o exports to country d, it must pay a fixed cost FC_{od}. These costs include foreign marketing and distribution costs, bureaucratic procedures on the border, and required changes in product characteristics to match up to the tastes of foreign consumers and government regulations. The presence of fixed cost in the differentiated manufacturing sector gives rise to increasing returns to scale production technology. By abstracting from additional domestic production entry costs allows us to focus on export entry and exit decisions of firms.

Assuming that each manufacturing firm draws a random unit labour productivity φ, a firm from country o with productivity φ has the following cost of producing x units of manufacturing good x and selling it in country d: $c(x) = \frac{x}{\varphi} + FC_{od}$.

As usual in the monopolistic competition framework, firms are price setters. Given that demand functions are iso-elastic, the optimal price charged in country d by firm j from country o is a constant mark-up over the unit cost (including the transportation cost):

$$p_{od}(\varphi) = \frac{\sigma}{\sigma - 1}\frac{\tau_{od}}{\varphi} \tag{3}$$

where p_{od} is price of the manufacturing variety produced in region o and sold in region d. The restriction $\sigma > 1$ ensures that the output price, p_o, is always positive.

Furthermore, we assume that the total mass of firms is proportional to country's endowment with labour force, L_r.[7] As in *Melitz* (2003), we assume that firms draw the productivity from a Pareto distribution with scaling parameter γ and that firm productivity is distributed according to $P(\tilde{\varphi} < \varphi) = F(\varphi) = 1 - \varphi^{-\gamma}$, with $dF(\varphi) = \gamma^{-\gamma}\varphi d\varphi$ for $\varphi \geq 1$. Variable γ is an inverse measure of firm heterogeneity in the manufacturing sector, with $\gamma > 2$ and $\gamma > \sigma - 1$.[8] Sectors with lower γ are more heterogeneous, in the sense that more output is concentrated among the largest and most productive firms.

3.2 Equilibrium

As in *Melitz* (2003), we assume that each firm in every country chooses a strategy, taking strategies of all other firms and all consumers as given. A strategy for a firm is both a subset of countries, where to sell its output and prices to set for its goods in each market. A strategy for a consumer is the quantity to consume of each variety of every good available domestically, given its price. From the optimal strategies of firms and consumers in every country, we can subsequently compute the global general equilibrium. The global trade equilibrium is characterised by a set of prices and quantities that correspond to a fixed point of the best response graph of each agent.

Given the optimal pricing strategy of firms and the optimal demand strategy of consumers, we can derive firm exports, e_{od}, from origin country o to destination country d:

$$e_{od}(\varphi) = p_{od}(\varphi)\, x_{od}(\varphi) = \alpha L_d \left(\frac{p_{od}(\varphi)}{P_d}\right)^{1-\sigma} \tag{4}$$

where φ is firm-specific productivity and P_d is price index of horizontally differentiated manufacturing goods in destination country d. If only those firms above the

[7] Implicitly, we assume that there is a group of firms proportional to the size of the country. We could remove this assumption, and allow for the free entry of firms, with an infinite set of potential firms. According to *Chaney* (2007), we would obtain qualitatively the same results, if trade barriers are not negligible.

[8] $\ln\varphi$ has a standard deviation equal to $\frac{1}{\gamma}$. The assumption $\gamma > \sigma - 1$ ensures that, in equilibrium, the size distribution of firms has a finite mean. If this assumption were violated, firms with an arbitrarily high productivity would represent an arbitrarily large fraction of all firms, and they would overshadow less productive firms. Results on selection into export markets would be degenerate. This assumption is satisfied in the data for all countries in our sample.

productivity threshold $\bar{\varphi}_{rd}$ from country o would export to country d, then the ideal price index, P_d, in country d can be defined as follows:

$$P_d = \left(\sum_{r=1}^{R} L_r \int_{\bar{\varphi}_{rd}}^{\infty} \left(\frac{\sigma - 1}{\sigma} \frac{\varphi}{\tau_{rd}} \right)^{\sigma-1} dF(\varphi) \right)^{\frac{-1}{\sigma-1}} \tag{5}$$

As long as net profits generated by exports to country d are sufficient to cover fixed entry cost, FC_{od}, firms will be willing to export to country d. The profits earned by firm n in country o from exporting to country d are then given by:

$$\pi_{od}(\varphi) = \frac{r_{od}(\varphi)}{\sigma} - FC_{od} \tag{6}$$

where $r_{od}(\varphi)$ is firm revenue from selling in country d. As in *Melitz* (2003), the productivity threshold, $\bar{\varphi}_{od}$, corresponds to productivity of the least productive firm in country o, for which gross profits earned in country d are just enough to cover the fixed costs of entering market d:

$$\pi_{od}(\bar{\varphi}_{od}) = FC_{od} \tag{7}$$

$$\bar{\varphi}_{od} = \lambda_1 FC_{od}^{\frac{\sigma}{\sigma-1}} \left(P_d^{\sigma-1} L_d \right)^{\frac{-1}{\sigma-1}} \tau_{rd} \tag{8}$$

with λ_1 a constant.[9] We assume that trade barriers are always high enough to ensure that $\forall\ j,\ r,\ \bar{\varphi}_{od} > 1$.[10]

Consumer prices in destination country d depend on country characteristics. More precisely, they are increasing in trade costs decreasing in market size. From equation (8) we can calculate the set of firms that export to country d. Because of the selection that takes place among exporters, this set (firms exporting to country d) only depends on country d's characteristics and trade costs.

By definition, the price index in country d is given by:

$$P_d^{1-\sigma} = \sum_{r=1}^{R} L_r \int_{\bar{\varphi}_{rd}}^{\infty} \left(\frac{\sigma - 1}{\sigma} \frac{\tau_{rd}}{\varphi} \right)^{1-\sigma} dF(\varphi) \tag{9}$$

Plugging the productivity threshold from equation (8) into price index (9), we can solve for the general equilibrium price index, P_d:

$$P_d = \lambda_2 \left(\frac{L_d}{L} \right)^{\frac{1}{\gamma}} L_d^{\frac{-1}{\sigma-1}} \theta_d \tag{10}$$

[9] $\lambda_1 = \left(\frac{\sigma}{\alpha}\right)^{\frac{1}{\sigma-1}} \left(\frac{\sigma}{\sigma-1}\right)$.

[10] This assumption is well supported by the empirical research using firm-level data, e.g. from the U.S. Census Bureau's Annual Survey of Manufacturers.

where λ_2 is constant[11] and $\theta_d^{-\gamma} \equiv \sum_{r=1}^{R} \frac{L_r}{L} \tau_{rd}^{-\gamma} FC_{rd}^{1-\frac{\gamma}{\sigma-1}}$, with $L \equiv \sum_{r=1}^{R} L_r$. Variable θ_d is an aggregate index of d's remoteness from the rest of the world.[12] It is similar to the 'multilateral resistance variable' introduced by *Anderson* and *Wincoop* (2003). In addition to their measure, it takes into account the impact of fixed costs and the impact of firm heterogeneity on prices.

Firm heterogeneity has a direct impact on the average productivity of exporters in the destination market, d. Larger and more integrated markets attract more firms, and the new firms that enter foreign markets are typically less productive. These new entrants lower the average productivity of suppliers. Equations (7) and (10) allow us to derive the average productivity of firms exporting to country d as a function of export entry cost into, FC_{od}, country d's remoteness from its trading partners, θ_d, country d's market size, L_d, and the bilateral unit trade cost, τ_{rd}:

$$\tilde{\varphi}_{od} = \lambda_{\tilde{\varphi}} \left(\frac{L}{L_d} \right)^{\frac{1}{\gamma}} \left(\frac{\tau_{od}}{\theta_d} \right) FC_{od}^{\frac{1}{\sigma-1}} \tag{11}$$

where $\lambda_{\tilde{\varphi}}$ is constant.[13] According to equation (11), countries that are expensive for exporting firms to enter (FC_{od} large), far away (τ_{od} large), or which have a small market (L_d low), attract only the most productive exporters. If country d is far away from its trading partners (θ_d large), it is harder for exporting firms to compete, implying that only the most productive firms from country o are able to enter country d. According to equation (11) variable trade cost, τ_{od}, with elasticity one is the major determinant of the average productivity of exporters in the destination market, d.

3.3 Trade

In the previous section we have solved for price indices of tradable goods in every country. In this section we use the general equilibrium price index to solve for firm level exports and the exporting productivity threshold. The expression for firm level exports, which we obtain in this section differs from neoclassical models of trade in homogeneous goods. Moreover, because of the two simplifying assumptions which we have made at the beginning, they also differ from *Melitz* (2003) implying that the results we obtain in this section are not directly comparable to the *Melitz*'s model. While in *Melitz*'s model firm level exports depend on both domestic and exporting productivities, in our model, which is similar to *Chaney* 2007, firm level exports only depend on exporting productivity. As a result, we are able to derive closed

---

[11] $\lambda_2 = \left(\frac{\gamma-(\sigma-1)}{\gamma} \right)^{\frac{1}{\gamma}} \left(\frac{\sigma}{\alpha} \right)^{\frac{1}{\sigma-1}-\frac{1}{\gamma}} \left(\frac{\sigma}{\sigma-1} \right)$.

[12] A simple way to interpret this aggregate index is to look at a symmetrical case: when $\tau_{rd} = \tau_d$ and $FC_{od} = FC_d$ for all d's, $\theta_d \equiv \tau_d FC_d^{\frac{1}{\sigma-1}-\frac{1}{\gamma}}$. In asymmetric cases, θ_d is a weighted average of bilateral trade costs.

[13] $\lambda_{\tilde{\varphi}} = \left(\frac{\sigma}{\alpha(\gamma-(\sigma-1))} \right)^{\frac{1}{\gamma}}$.

form solutions of intensive and extensive margins, which are important for empirical analysis of the present study.

By plugging the general equilibrium price index from equation (10) into the demand function and into the productivity threshold (8), we obtain general equilibrium exports, $e_{od}(\varphi)$, from origin country o to destination country d:

$$e_{od}(\varphi \mid \varphi > \bar{\varphi}_{od}) = \lambda_3 \left(\frac{L_d}{L}\right)^{\frac{\sigma-1}{\gamma}} \left(\frac{\tau_{od}}{\theta_d}\right)^{1-\sigma} \varphi^{\sigma-1} \tag{12}$$

the productivity threshold $\bar{\varphi}_{od}$ above which firms from o export to d, is given by

$$\bar{\varphi}_{od} = \lambda_4 \left(\frac{L}{L_d}\right)^{\frac{1}{\gamma}} \left(\frac{\tau_{od}}{\theta_d}\right)^{\sigma-1} FC_{od}^{\frac{1}{\sigma-1}} \tag{13}$$

where λ_3 and λ_4 are constants.[14] According to equation (12), firm exports are determined by the countries' relative size, L_d, bilateral trade barriers, FC_{od} and τ_{od}, and the d's remoteness from the rest of the world, θ_d. Individual firm exports depend on the transportation cost, τ_{od}, with elasticity $1-\sigma$ and on the size of the destination market, L_d, with elasticity $\frac{\sigma-1}{\gamma}$, which is less than one, because of the impact of market size and because of the impact of price competition. Both these elasticities are smaller than the corresponding elasticities of aggregate trade, because aggregate trade volume depends also on the number of exporters, N_{od}, which is defined as follows:

$$N_{od} = L_o P_d (\varphi > \varphi^*) = \lambda_E \frac{L_o L_d}{L} \left(\frac{\tau_{od}}{\theta_d}\right)^{-\gamma} FC_{od}^{-\frac{\gamma}{\sigma-1}} \tag{14}$$

where λ_E is a constant.[15] According to equation (14), the number of firms, N_{od}, reacts to changes in unit trade costs, τ_{od}, with an elasticity of γ, and to changes in the size of origin and destination countries, L_r, with elasticity 1, which is close to the values recovered from the firm-level trade data (see section 4).

According to the definition of E_{od}, which is given in equation (1), aggregate exports (*f.o.b.*) from origin country o to destination country d can be decomposed as the number of exporters times the average exports per firm with an average productivity above $\tilde{\varphi}_{od}$:

$$E_{od} = \underbrace{e_{od}(\tilde{\varphi}_{od})}_{\text{Intensive}} \underbrace{N_{od}(\cdot)}_{\text{Extensive}} \tag{15}$$

where N_{od} is the number of exporting firms (the extensive margin of trade) and e_{od} is the average value per shipment (the intensive margin of trade). Adopting *Hummels* and *Klenow* (2005) terminology, the extensive margin is defined by the number of

[14] $\lambda_3 = \alpha \left(\frac{\gamma-(\sigma-1)}{\gamma}\right)^{\frac{\sigma-1}{\gamma}} \left(\frac{\sigma}{\alpha}\right)^{1-\frac{\sigma-1}{\gamma}}$ and $\lambda_4 = \left(\frac{\sigma}{\alpha} \frac{\gamma}{\gamma-(\sigma-1)}\right)^{\frac{1}{\gamma}}$.

[15] $\lambda_E = \alpha \frac{\gamma-(\sigma-1)}{\gamma\sigma}$.

exporting firms, which in our model is equal to the number of goods/varieties traded. The intensive margin is accordingly defined by the size of exporters, which in our model is equal to the size of firm exports.

Substituting equations (12) and (14) into equation (15), total exports, E_{od}, in manufacturing sector x from origin country o to destination country d can be expressed as:

$$E_{od} = \underbrace{\lambda_E^{\frac{\sigma-1}{\gamma}} \sigma \bar{\varphi}^{\sigma-1} \left(\frac{L_d}{L}\right)^{\frac{\sigma-1}{\gamma}} \left(\frac{\tau_{od}}{\theta_d}\right)^{1-\sigma}}_{\text{Intensive margin}} \underbrace{\lambda_E \frac{L_o L_d}{L} \left(\frac{\tau_{od}}{\theta_d}\right)^{-\gamma} FC_{od}^{-\frac{\gamma}{\sigma-1}}}_{\text{Extensive margin}} \qquad (16)$$

According to equation (16), the aggregate exports from origin country o to destination country d depend on the relative size of countries, $\frac{L_r}{L}$, on destination country d's multilateral resistance, θ_d, and on bilateral transport costs (both fixed and variable) among the trading partners. As in equation (15), aggregate exports, E_{od}, may vary due to changes in average value, e_{od}, per shipment (intensive margin of trade) or due to changes in the number of shipments, N_{od}, (extensive margin of trade) both of which in turn may vary across destinations and co-vary with trade costs.

3.4 Discussion of the model

The right hand side explanatory variables in equation (16) are similar to the traditional explanatory variables of conventional gravity models of trade with representative firms. Despite the underlying common gravity structure, the trade model derived in equation (16) differs from gravity models with representative firms in several respects. In this section we identify these features. In light of these differences we then discuss those assumptions, which have led to these differences.

First, given that in our model firms have to pay export entry cost, trade costs may reduce quantities exported to the point that firms can no longer cover fixed costs of exporting. Thus, our model allows for firms from country o to choose not to export to country d, because, it is possible that no firm in country o has productivity above the threshold, $\bar{\varphi}_{od}$, that makes exports to d profitable, if trade barriers are sufficiently high. The model is therefore able to predict zero exports from o to d for some country pairs and some products. As a result, our model is consistent with zero trade flows in both directions between some Balkan countries, as well as zero exports from o to d but positive exports from o to r for some other country pairs. Both types of trade patterns exist in the SEE trade data.

Second, our model can predict positive trade flows in both directions for some country pairs. The two-way trade flows are important if we wish our model to be consistent with the two-way trade observed in trade data for the SEE economies. According to the *Eurostat* (2005) data, positive two-way trade flows are observed for all SEE countries in our sample. Although, one-way international trade flows prevail in the Balkans underlying the inter-sectoral trade pattern and the complementarity of factor endowments, the share of two-way trade significant (and increasing). Considering the 1999-2004 period, similar dynamics are observed in different SEE countries, although with different intensities, pointing to the reduction of one-way trade and the increase of two-way trade, especially in vertically differentiated goods (*Eurostat* 2005).

Third, the elasticities with respect to trade costs are different. In gravity models with representative firms the elasticity of exports with respect to trade costs are equal to $\sigma - 1$. In contrast, in our model the elasticity of exports with respect to variable costs depends on the degree of firm heterogeneity, γ, but not on the elasticity of substitution between manufacturing varieties, σ. Given that $\gamma > \sigma - 1$, in our model the elasticity of exports with respect to variable trade barriers, τ_{od}, is larger than in the absence of firm heterogeneity. The elasticity of exports with respect to fixed trade costs is negatively related to the elasticity of substitution σ. This prediction is in stark contrast to gravity models with representative firms. Moreover, in our model the elasticity of exports with respect to trade costs depends on the degree of firm heterogeneity, γ. In more homogeneous sectors (γ high) large productive firms represent a smaller fraction of firms. The productivity threshold moves in a region where most of the mass of firms lies. In those sectors, aggregate exports are sensitive to changes in trade costs because many firms exit and enter when variable trade costs fluctuate.

Finally, in the context of trade growth in the SEE transition economies, we are particularly interested in data requirements. The advantage of our approach is that, despite the fact that the theoretical model assumes firm-level heterogeneity, the underlying theoretical model does not require firm-level data to study the structure of trade growth. This stems from the fact that exporter features can be identified from variations in the characteristics of trading countries. Given that for every origin country o, its exports to different destination countries vary by characteristics of importing countries, there exist sufficient statistics, which can be computed from aggregate data that can decompose the aggregate export volume between the SEE Balkan economies. The downside of our approach is that introducing firm heterogeneity and fixed trade costs requires more parameters for the empirical implementation of the theoretical trade model. In particular, an additional parameter describing firm heterogeneity is required and a parameter capturing export entry cost is required. In order to deal with increased parameter requirements, we estimate model parameters econometrically, which is done in the next section.

4 Parameter estimation

In the previous section we have presented the underlying trade model, which forms the theoretical basis for the empirical analysis. Before the theoretical trade model can be empirically implemented, it needs to be parameterised. Two types of parameters are required by the theoretical trade model: trade freeness and behavioural parameters.

In order to obtain numerical values of model parameters, we estimate two equations: a trade freeness equation and an alternative gravity equation of trade flows. The estimated coefficients of the former will provide trade cost estimates, while the latter will provide estimates for behavioural parameters.

We proceed as follows. First, we derive an empirically estimable trade freeness equation. Next, we estimate the freeness of trade for selected SEE economies. Second, we use the theoretical trade model, which we have presented in the previous section, to derive an empirically estimable gravity model of bilateral trade flows. The gravity model of trade is estimated econometrically and the results are presented in section 4.4.

4.1 Trade cost specification

One of the key explanatory variables in the underlying theoretical trade model are inter-regional trade costs. The theoretical trade model distinguishes between two types of trade costs: variable trade cost, τ_{od}, and fixed trade cost, FC_{od}. Given that the true trade costs are unobservable in the SEE transition economies, we follow *Head* and *Mayer* (2004), which proposed an alternative measure of trade costs. They propose that trade costs can be proxied by the trade freeness. According to *Head* and *Mayer*, the index of trade freeness, ϕ_{od}, captures the easiness with which two countries participate in reciprocal trade and is defined as $\phi_{od} = \tau_{od}^{1-\sigma}$.[16] Given that $\sigma > 1$, trade freeness is inversely related to trade costs.

However, given that in the underlying theoretical trade model manufacturing firms face two different trade costs, we cannot straightforwardly apply the *Head* and *Mayer* (2004) definition of trade freeness. Instead, we need to derive a measure of trade freeness, which would be consistent with the underlying theoretical framework. According to our theoretical trade model, the composite measure of trade freeness is defined as $\phi_{od} = \tau_{od}^{-\gamma} FC_{od}^{1-\frac{\gamma}{\sigma-1}}$. Thus, our measure is different from the *Head* and *Mayer* trade freeness measure in two respects. First, our measure, ϕ_{od}, accounts for both fixed and variable trade costs. Second, the elasticities which relate the trade freeness to trade costs are different.

A structural estimation of the index of trade freeness, ϕ_{od}, is extremely data demanding and cannot be performed even for the old EU member states, where the statistical data base is considerably more developed than for transition economies in Eastern Europe. In order to cope with data limitations, *Head* and *Mayer* (2004) suggested that calculation of the trade freeness index, ϕ_{od}, can be facilitated by making two simplifying assumptions: symmetric trade costs for external trade ($\phi_{od} = \phi_{do}$) and zero trade costs for trade within countries ($\phi_{rr} = 1$).

Because of data limitations in the SEE transition economies, in this study we can only calculate the 'reduced form' of trade freeness, ϕ_{od}. In view of SEE, the first

[16] This trade cost measure, which *Baldwin et al* (2003) cunningly refer to as the 'phi-ness' of trade, is often employed in economic geography models as a proxy for trade costs.

assumption is not critical for countries in our sample, because none of the included Balkan economies has a significant geographical advantage or disadvantage, which could asymmetrically affect bilateral trade flows. The second assumption might become critical under certain circumstances. In particular, internal trade costs usually increase with size of the country. Consequently, when trade costs arise only for cross-border transactions, trade freeness might be underestimated, suggesting lower levels of trade integration. Thus, from the economic geography's perspective, the estimated trade freeness might potentially be upward biased for geographically large countries, such as Bulgaria and Romania.

According to *Head* and *Mayer* (2004), assuming frictionless intranational trade and symmetric trade costs for bilateral trade, the index of country trade freeness, ϕ_{od}, can be calculated as follows:

$$\phi_{od} = \sqrt{\frac{E_{od}E_{do}}{E_{oo}E_{dd}}} \tag{17}$$

where ϕ_{od} is the trade freeness index, E_{od} is value of goods and services exports from origin country o to destination country d and E_{do} captures exports from d to o. Denominator factors E_{oo} and E_{dd} are exporting and importing countries domestic sales. They are calculated as the value of all shipments of an industry minus the sum of shipments to all other countries (exports).

Two-way parameter restrictions need to be imposed, when estimating equation (17): the trade freeness estimates, $\hat{\phi}_{od}$, need to be bounded both from above and from below. These restrictions imply that the estimated trade freeness can only take values between zero and one, $0 < \phi_{od} < 1$, with 0 denoting prohibitive trade costs and 1 denoting free trade.[17]

4.2 Trade flow specification

In this section we specify an econometrically estimable gravity model of bilateral trade flows. The departure point of the empirical specification is equation (16), which suggests a trade model with exports, E_{od}, as the dependent variable and, the relative market sizes, bilateral trade costs and country multilateral resistance as explanatory variables. Plugging equation (13) into equation (16) and collecting terms we obtain the following equation for export flows:

$$E_{od} = \alpha \frac{L_o L_d}{L} \left(\frac{\theta_d}{\tau_{od}} \right)^{\gamma} FC_{od}^{1-\frac{\gamma}{\sigma-1}} \tag{18}$$

According to equation (18), aggregate exports from origin country o to destination country d is determined by the relative size of countries, $\frac{L_r}{L}$, the destination country

[17] Theoretically, the trade freeness index could be larger than 1, if the external trade of both trading partners is larger than internal trade. However, this is not an issue in the SEE trade.

d's multilateral resistance, θ_d, the variable trade costs, τ_{od}, and by the fixed trade costs, FC_{od}.

The empirical estimation of equation (18) faces several complications. In particular, we identify two issues: the potential endogeneity of the right-hand side explanatory variables and the omitted variables bias. In the following we discuss these two estimation issues and propose solutions how do we deal with them.

We start with the potential endogeneity problems. According to the underlying theoretical trade model, the endogeneity of explanatory variables might be caused in at least two ways. First, export flows might potentially give rise to adjustments in the explanatory variables, i.e. reverse causality. For example, labour demand in country o is an increasing function of exports from country o. Second, in the SEE transition economies there may exist confounding factors, such as macroeconomic shocks and structural adjustments, which might contemporaneously affect both sectoral employment and export flows. For instance, a negative income shock through the Balkan wars may induce emigration and, at the same time, reduce export demand within SEE.

The potential endogeneity of explanatory variables implies that equation (18), will likely yield biased and inconsistent estimates. We use two different approaches to get around the endogeneity problems. First, we consider relative export flows instead of gross exports, i.e., we estimate the ratio of gross exports from country o to country d with respect to to exports from country d to country o. This transformation allows us to substitute out the sectoral labour demand, L_o, which is a major source of endogeneity in equation (18).[18] Second, we follow *Honoré* and *Kyriazidou* (2000) and use instrumental variables with lagged values of right-hand side explanatory variables as 'instruments'.[19] Thus, we implicitly assume that exporting decision at date t are determined from a comparison of potential profits and costs at date $t-1$. We restrict the number of lags to one in order not to loose further time-series observations. For the instrumental variables estimation we need to assume that instruments are predetermined, and export flows and confounding factors in residuals only affect contemporaneous and future labour force supply in exporting and in importing countries.

After these two transformations we obtain the following equation of export flows:[20]

$$\triangle E_{od} = (\triangle FC_{od})^{1-\frac{\gamma}{\sigma-1}} (\triangle\theta_{do})^{\gamma} \tag{19}$$

where $\triangle E_{od} \equiv \frac{E_{od}}{E_{do}}$ are relative exports from country o to country d, $\triangle FC_{od} \equiv \frac{FC_{od}}{FC_{do}}$ is the ratio of export entry costs and $\triangle\theta_{do} \equiv \frac{\theta_d}{\theta_o}$ is the ratio of multilateral resistance between origin country o and destination country d. Although, none of the right hand side variables in (19) are directly observable in the data, they all can be

[18] Due to symmetric per-unit trade costs, τ_{od}, cancels out too.

[19] Although, properly taken, these are not instruments but lagged values of explanatory variables, given that in the context of our analysis this does not cause a confusion, we call them instruments.

[20] The time notation will be introduced in the econometric model.

calculated from statistical data which is available. The dependent variable, $\triangle E_{od}$, can be straightforwardly calculated from the bilateral exports between o and d, and d and o. The two explanatory variables are unobservable, but can be calculated on the basis of observable variables. Given that the ratio of fixed trade costs, $\triangle FC_{od}$, is equal to the profit ratio, $\triangle\pi_{od}$, it can be calculated from the firm profit data, which is available in the SEE data.[21] The other explanatory variable in equation (19) is the ratio of multilateral resistance between origin region o and destination region d. It can be calculated according its definition ($\theta_r^{-\gamma} \equiv \sum_{r=1}^{R} \frac{L_r}{L}\phi_{rd}$) using data for the regional labour endowment and the trade freeness estimates.

Obviously, beyond the included explanatory variables, unobservable economic and non-economic characteristics of regions, such as amenities, also play an important role in exporting decisions of firms. According to previous research (e.g. *Mátyás* 1998, *Egger* 2000), there are several reasons to assume that country-specific fixed effects are relevant when geographical, political or historical determinants that could drive or hamper trade flows are present. These factors are deterministically linked to country-specific characteristics and cannot consequently be considered as random. Failing to account for the unobserved cross-section heterogeneity would yield biased estimates.

Following these findings, we explicitly account for country pair-specific effects, which – as emphasised by *Arellano* and *Honoré* (2001) – should reduce the heterogeneity bias. More precisely, in order to to avoid potential misspecification problems due to omitted variables, we include a constant term of country-specific characteristics and use the fixed effects estimator. According to previous panel data studies of trade flows, which estimate gravity model using panel data estimators, in order to obtain an efficient estimator, instead of using one dummy variable per country, individual country pair dummies (fixed effects) should preferentially be included in the econometric model (*Anderson* and *Wincoop* 2003). Following these findings, we include a time invariant constant term, ν_{od}, which captures fixed effects between exporting country o and importing country d, among the right hand side explanatory variables in equation (19).

Applying a logarithmic transformation to equation (19) and introducing the time reference, we obtain the following linearly estimable gravity equation of export flows from origin country o to destination country d:

$$\log \triangle E_{odt} = \beta_1 + \beta_2 \log \triangle FC_{odt-1} + \beta_3 \log \triangle\theta_{dot-1} + \nu_{od} + \epsilon_{odt} \qquad (20)$$

where β_1 is intercept and β_2 and β_3 are the coefficients to estimate, ν_{od} is a time invariant constant term capturing country-pair specific fixed effects and ϵ_{od} is a random prediction error. According to equation (20), exports from country o to

[21] Using equations (6), (10) and (11), we can express average profits of exporting from country o to country d, as a constant mark-up over fixed trade costs: $\pi_{od} = \lambda_\pi FC_{od}$ where $\lambda_\pi = \left(\frac{\sigma}{\alpha}\right)^{\frac{2-\sigma}{\sigma-1}} \left(\frac{\sigma-1}{\sigma}\right)^{\sigma-2} \left(\frac{1}{\gamma}\right)^{\frac{1}{\gamma}}$ is the mark-up. The mark-up cancels out, when we take the ratios of fixed costs and profits.

country d are determined by fixed trade costs and the multilateral resistance in importing country o and exporting country d. These effects are amplified by the degree of firm concentration (heterogeneity) and by the degree of product differentiation (substitutability).

The inclusion of fixed effects does not allow estimation of time-invariant explanatory variables, which enter also into the fixed effects.[22] This implies that we will not be able to identify the time-invariant explanatory variables in equation (20). Analysing the right hand side explanatory variables in equation (20) we note that the first term capturing fixed trade costs, $\triangle FC_{od}$, is calculated on the basis of firms profits, $\triangle \pi_{od}$. According to the firm-level tax data, the profits of exporting firms vary considerably over time without a clear trend in the Balkans. The index of trade freeness, ϕ_{od}, which we use as a proxy for trade costs is also time-variant (see Figure 1 in section 4.3); according to our estimates, the SEE trade freeness has almost doubled since the end of the Balkan wars. Given that the multilateral trade resistance, θ_d, is calculated on the basis of two time-variant variables, it is time-variant too.[23] We may conclude that the inclusion of fixed effects is not an issue for our model, because all right-hand side explanatory variables in equation (20) are time-variant.

In order the fixed effects estimator to be unbiased, we need to assume that explanatory variables, $\triangle FC_{odt-1}$ and $\triangle \theta_{dot-1}$, are strictly exogenous conditional on time invariant constant, ν_{od}. In order to test for endogeneity we add next year's explanatory variables (in logarithmic form) and run a robust t-test. In doing so we lose the last year of the data. Both coefficients are very small, -0.059 and -0.0471 and the t-statistic is only -0.018. From the robust t-test results, we may conclude there is no evidence against the strict exogeneity assumption.

In order to ensure that the fixed effects estimator is well behaved asymptotically, we need a standard rank condition on the matrix of time-demeaned explanatory variables. In order to ensure that the fixed effects estimator is efficient, we need to assume that the conditional variances are constant and the conditional covariances are zero. While the heteroscedasticity in time-demeaned errors, $\ddot{\epsilon}_{odt}$, might be a potential problem, serial correlation is likely to be less important. Because of the time demeaning, the serial correlation in the time-demeaned errors, $\ddot{\epsilon}_{odt}$, under the latter assumption causes only minor complications.

Finally, in order to ensure that the obtained parameters are consistent with the theoretical trade model, we need to impose parameter restrictions implied by the theoretical trade model. Equation (20) contains two key parameters of the underlying theoretical trade model ($\beta_2 = 1 - \frac{\gamma}{\sigma-1}$ and $\beta_3 = \gamma$). In particular, the theoretical trade model requires that $\sigma > 1$ and $\gamma > \sigma - 1$.

[22] In panel data analysis the term 'time-varying explanatory variables' means that each explanatory variable varies over time for some cross section units. There might be explanatory variables that are constant across time for a subset of the cross section, but this is irrelevant.

[23] This might result in non-stationary multilateral resistance, $\triangle \theta_{do}$, which we discuss below.

4.3 Estimation results: trade costs

We begin by presenting the estimation results for trade costs between the SEE Balkan economies. Before presenting the estimation results, we briefly discuss data, which we use for estimating the trade freeness.

Our sample consists of eight SEE economies - Albania, Bosnia & Herzegovina, Bulgaria, Croatia, Macedonia, Moldova, Romania, Serbia & Montenegro.[24] The time period covered spans 1999 to 2004. Given that no reliable international bilateral trade statistics exist for such geographic coverage and time period, we had to draw on national trade statistics data on bilateral export flows.[25] Although cumbersome, mapping of trade data provided by national statistical offices is the only way to obtain complete statistical information for the SEE bilateral trade. Given that the SEE economies use their own national currencies, the obtained export values need to be converted in one currency. We calculate all flows in Euros, as since 2003 the SEE national statistical resources report all international statistics not only in national currencies, but also in Euros.

As a result, we obtain eight equally sized panels each containing 48 observations (8 countries $\times$ 6 years). Using this data we estimate the index of trade freeness, ϕ_{od}, according to equation (17) for each SEE country. The obtained trade freeness estimates are reported in Figure 1, where the trade freeness estimates, ϕ_{od}, are on the vertical axis and time span is on the horizontal axis.

Given that the measure of trade freeness, ϕ_{od}, is negatively related to trade costs with 0 denoting prohibitive trade costs and 1 free trade, estimates in Figure 1 suggest that the overall level of trade freeness is very low in the SEE economies. Although, these countries are known for their high levels of the formal trade integration since the end of the Balkan wars in 1999 (there exist a network of 31 bilateral FTAs), the estimated trade freeness is lower than 0.1 (SEE average). This indicates that less that 10% of the total trade in the SEE crosses national borders.

These estimates are very low compared to the *Head* and *Mayer* (2004) trade freeness estimates of the EU internal trade. Depending on the period covered and countries included, *Head* and *Mayer* estimates range from 0.315 to 0.478. As a robustness test of our estimates we also estimate trade freeness for the SEE trade with the EU. The obtained trade freeness estimates for the SEE trade with the EU range from 0.125 to 0.176 in 2004 and are between our estimates for intr-SEE trade and *Head* and *Mayer* estimates for intra-EU trade. These results suggest that our results are consistent with *Head* and *Mayer* (2004) estimates of trade freeness.

The second attribute, which can be taken from Figure 1, is that the SEE trade freeness has increased between 1999 and 2004. On average, trade freeness has increased by almost one third from 0.066 to 0.084. Moreover, the estimates reported in

[24]Despite the recently re-established independence of Montenegro, Serbia and Montenegro is considered as one country in the empirical analysis.

[25]For example, Eurostat's *Comext* trade data does not cover bilateral trade flows among third countries. It only contains SEE trade with the EU.

Figure 1 suggest that the trade freeness has increased at different growth rates within SEE. The most sizeable increase in the regional trade freeness we have estimated for Albania (+86.4%), Romania (+85.3%) and Moldova (+85.0%). According to the same estimates, the bilateral trade freeness has increased least rapidly in Serbia & Montenegro. Trade costs might have decreased slower in Serbia & Montenegro because of two reasons: relatively large internal market and continuing armed conflicts, such as Presevo Rebellion in 2000 to 2001.

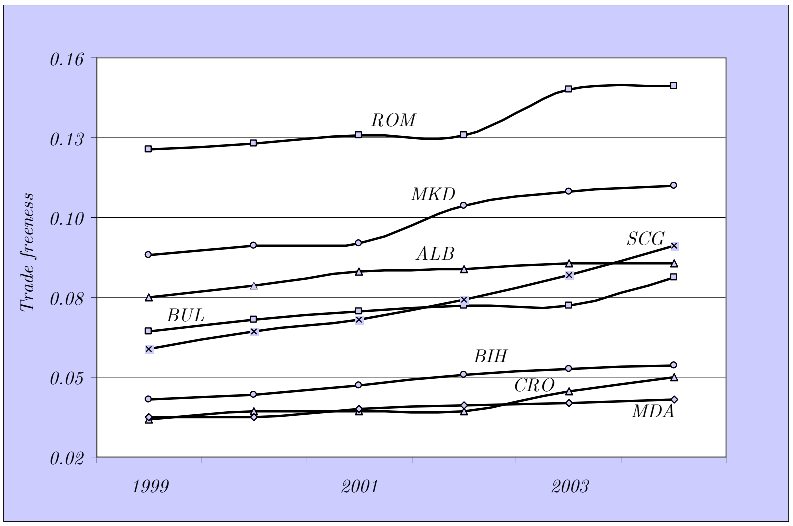

Figure 1: Trade freeness of the SEE bilateral trade, 1999-2004

The obtained trade freeness estimates can be used for both to estimate the gravity model of trade flows and to empirically implement the theoretical trade model for policy simulations. These estimates also allow us to draw several conclusions, which are relevant for both applications: (i) compared to the EU internal trade, trade freeness is still very low in the SEE countries; (ii) trade freeness is increasing rapidly (SEE average +77.5% in the period 1999 to 2004) and increasing with an increasing rate, which implies that the inclusion of the trade freeness estimates among the explanatory variables in the gravity model of trade might lead to non-stationarity problems; and (iii) because of (i) and (ii), the proposed BFTA has large potential in increasing trade openness and facilitating regional trade in the Balkans.

4.4 Estimation results: trade flows

In this section we estimate the gravity equation of trade (20) using panel data for eight SEE countries - Albania, Bosnia and Herzegovina, Bulgaria, Croatia, Macedonia, Moldova, Romania, Serbia and Montenegro. As above, the time series in our data goes from 1999 to 2004. Given that we only consider bilateral trade flows among the SEE countries, the cross-section dimension of our data is equal to eight. Similarly to the trade freeness estimations reported in the previous section, our data allows to build eight equally sized panels with 48 observations in each panel. Lagging explanatory variables by one year reduces the number of observations per panel to 40 (8 countries $\times$ 5 years).

Estimation of equation (20) requires time series cross section data of bilateral trade flows, E_{od}, firm profit data and multilateral resistance variable, θ_d. Calculation of the multilateral resistance requires data for trade freeness, ϕ_{od}, supply of labour force in each country, L_r, and the total labour force, L. Data sources for export flows have already been detailed in the previous section. Firm profit data are drawn from national tax registers, which are available on yearly basis for all SEE countries in our sample. Numerical values for importer and exporter multilateral resistance, θ_r, are calculated by drawing the supply of labour force in each country, L_r, and the total labour force, L, from the *Vienna Institute for International Economic Studies* (*WIIW*) (2005 and 2006) Handbook of Statistics and the above presented estimates of trade freeness, $\hat{\phi}_{od}$.

Regression results for the fixed effects model are presented in Table 1. According to Table 1, bilateral export flows are negatively affected by fixed trade costs - coefficient β_2 estimates are negative for all countries. The magnitude of these estimates are around one (except Serbia and Montenegro -2.118) and are of the same order across the SEE countries in our sample. The largest coefficient have been estimated for Serbia and Montenegro and the smallest for Albania (-0.726) which is in line with previous studies (*Messerlin* and *Miroudot* 2004, *Bussiere*, *Fidrmuc* and *Schnatz* 2004). The estimates of coefficient β_2 are statistically significant for three SEE countries: Albania, Bosnia and Herzegovina, and Macedonia.

Are these results in line with our expectations and previous trade studies on Balkans? Given that the relationship between explanatory and dependent variables in equation (20) is non-linear, and coefficient β_2 is non-linear in structural parameters, is not straightforward to answer the consistency question. According to the underlying theoretical model, exports are decreasing in fixed trade costs, FC_{od}. This implies that the ratio of export flows, $\triangle E_{odt}$, is also decreasing in the ratio of fixed costs, $\triangle FC_{od}$. I.e., the lower are fixed export costs from origin country o to destination country d in terms of fixed costs from d to o, the higher are exports from o to d and vice versa. Thus, for those countries, where coefficient β_2 estimates are negative, the total impact of the first right-hand side term on exports flows is consistent with the underlying theory. This is true for all countries in our sample. We conclude that our estimates are in line with the underlying theoretical trade model.

The other explanatory variable, which has been regressed on export flows, is the multilateral resistance, θ_r. According to Table 1, bilateral export flows are positively affected by the multilateral resistance - coefficient β_3 estimates are positive for all countries in our sample. Given that all estimated β_3 coefficients are larger than one, the multilateral trade resistance raises trade at an increasing rate. The cross-section variation of coefficient β_3 estimates is higher compared to β_2. Signs of the estimated impact of the multilateral resistance are in line with the underlying theoretical trade model and with previous gravity studies (e.g. *Anderson* and *Wincoop* 2003).

Table 1: Fixed effects estimates of bilateral exports

	ALB	BIH	BUL	CRO	MKD	MDA	ROM	SCG
β_2	-0.726††	-1.264†	-0.749	-1.392	-0.758†	-1.065	-1.212	-2.118
	(0.223)	(0.563)	(0.428)	(0.741)	(0.320)	(0.754)	(1.179)	(3.026)
β_3	4.009††	7.602	3.391†	3.849†	6.445	3.015††	3.346†	4.174†
	(1.210)	(5.483)	(1.496)	(1.606)	(3.558)	(0.853)	(1.354)	(1.915)
N	40	40	40	40	40	40	40	40
R^2	0.533	0.491	0.667	0.534	0.608	0.525	0.528	0.479

Dependent variable: log of bilateral exports, $\ln E_{od}$, (equation 20). Standard errors in parenthesis. † significant at 95% level, †† significant at 99% level.

As usual, we test the robustness with respect to the choice of estimator and the underlying assumptions. First, we estimate equation (20) using contemporaneous values of explanatory variables. On average, this reduces the numerical values of coefficients by one third, but does not change signs of the estimated coefficients. Testing the idiosyncratic errors for serial correlation is more tricky, as we cannot estimate the ϵ_{odt}. Because of the time demeaning used in fixed effects, we can only estimate the time-demeaned errors, $\ddot{\epsilon}_{odt}$. Given the relatively short time dimension of our panel, we neglect this issue in the empirical analysis.

From the estimated coefficients we can calculate parameter values for the theoretical trade model. More precisely, from β_2 estimates we obtain values for the elasticity of substitution, σ_r, where $\sigma_r = \frac{\gamma}{1-\beta_2} + 1$, with $\gamma \neq 0 \wedge \frac{1}{\gamma}(-\beta_2 + 1) \neq 0$. Numerical values of the firm heterogeneity parameter, γ_r, are obtained from the coefficient β_3 estimates, where $\gamma_r = \beta_3$. The bilateral trade cost values are obtained from the trade freeness estimates, where $\phi_{od} = \tau_{od}^{-\gamma} FC_{od}^{1-\frac{\gamma}{\sigma-1}}$.

5 BFTA impact on trade in the Balkans

In the previous two sections we have presented the theoretical trade model and estimated parameters which are required for empirical implementation of the theoretical trade model. In this section we substitute the estimated parameters into the theoretical trade model and drawing on statistical data for the base year we apply the model for assessing impacts of the proposed trade integration in the Balkans. More precisely, we perform simulation experiments of the proposed Balkan Free Trade Agreement by simulating three hypothetical policy scenarios. Beyond quantifying the aggregate impact on trade flows, our trade model also allows for decomposing the aggregate trade growth into two separate components - the intensive margin of trade and the extensive margin of trade growth.

5.1 Empirical implementation

Empirical implementation of the general equilibrium trade model requires two types of data: model parameters and numerical values of exogenous variables. Parameter values have already been estimated in the previous section. The only two parameters left, which could neither be estimated nor could be drawn from statistical data are the two types of trade costs. In particular, the theoretical trade model requires separate values for variable trade cost, τ_{od}, and fixed trade cost, FC_{od}. Numerical values of these two parameters are obtained by combining the survey-based shares for differentiated trade costs in the SEE countries with the estimated values of trade freeness, ϕ_{od}.

Numerical values of exogenous variables have to be drawn from statistical data. For the present study we require numerical values for regional employment, L_r, and total employment, L. This data is available for all SEE economies in both primary and secondary statistics. Given that the *WIIW*'s (2005) data does not reveal significant deviations from the national statistics data in 2004, we draw regional employment data and total labour force endowment in SEE from the *WIIW*'s (2005 and 2006) Handbook of Statistics. The base year, to which we fit the theoretical trade model, is 2004. This is the most recent year for which the required statistical data is available for all eight SEE economies.

Using the base year data for regional employment, L_r, total employment, L, bilateral trade costs (variable trade costs, τ_{od}, and fixed trade costs, FC_{od}), multilateral trade resistance, θ_d, and the estimated model parameters, we are able to empirically implement and solve the model for the general trade equilibrium. In the context of the present study we are particularly interested in export flows, E_{od}, and its components N_{od} and e_{od}, which are calculated according to equations (12), (14) and (16). Given that these equations do not contain any endogenous variables, we can straightforwardly plug equation (13) into equations (12), (14) and (16) and solve the model for the general trade equilibrium.

By implementing the theoretical trade model empirically and solving for the long-run trade equilibrium, we obtain a set of endogenous variables, which we call the base run equilibrium. In order to assess robustness of these results, we compare the obtained base run values of endogenous variables with those observed in the base year data. Comparing the obtained results with statistical data suggests that our model has not been able to exactly replicate the statistically observed trade in 2004.[26] However, the simulated trade flows are of the same order of magnitude as the

[26] Given that there are many other aspects that determine trade flows in the Balkans (e.g., histor-

corresponding values recorded in the SEE statistical data.[27]

5.2 Impact of declining trade barriers

In order to study the impacts of the proposed BFTA, in this section we perform ex ante simulations of alternative trade policy scenarios. Given that the exact magnitude of the effective reduction of trade barriers is not known a priori, we set up hypothetical trade policy liberalisation scenarios, in which the potential impacts of the proposed trade integration in the Balkans can be studied. In order to facilitate the identification of changes in the two different trade margins (the intensive margin and the extensive margin of trade), three alternative trade policy liberalisation scenarios are constructed. First, we reduce both the variable and the fixed trade costs in 10% steps up to 30%, which corresponds to 70% of the initial trade cost values. Second, we simulate fixed trade cost reduction in the same order of magnitude. In a third scenario we reduce the per unit (variable) trade costs in 10% steps up to 30%. The obtained simulation results are reported in Tables 2-4.

We start with presenting the aggregate impacts on export flows. Simulation results reported in Table 2 suggest that a 10% reduction in both variable and fixed trade costs have a positive and large impact on trade flows. In terms of the total export value (row E_{od} in Table 2) our model predicts sizable trade gains. These results are in line with previous studies of trade liberalisation in South Eastern Europe (*Stubos* and *Tsikripis* 2007). According to our simulations, if both fixed and variable trade costs were reduced by the same percentage, then the largest gainers from trade integration were Bosnia and Herzegovina, Croatia, and Romania. The estimated trade gains are different across the SEE countries ranging from +149.8% in Moldova to +208.2% in Bosnia and Herzegovina.

The aggregate impact of declining trade barriers on the export value is not only sizable, it is also larger than the impact of the same trade barriers estimated in trade models with representative firms (*Messerlin* and *Miroudot* 2004, *Bussiere*, *Fidrmuc* and *Schnatz* 2004). In our study the impact of trade integration is larger because, in addition to adjustments in the intensive margin of trade, which is well captured in trade models with representative firms, our model also captures adjustments in the extensive margin of trade. According to simulation results reported in Table 2, when both fixed and variable trade costs are reduced by the same percentage, adjustments in the extensive margin of trade are even more significant than adjustments in the intensive margin of trade in the SEE economies. In trade models with representative firms, when trade barriers decline, each firm exports more. In our model, however, in addition to larger average shipments per each existing firm, new firms would enter foreign markets and start exporting. Our simulation results suggest that the entry

ical ties, cultural preferences), which are not captured in our model, our predictions deviate from the base year statistical data.

[27] For the base year (2004) the correlation is rather high ($R^2 = 0.871$).

Table 2: Variable and fixed trade cost impact on export flows

	ALB	BIH	BUL	CRO	MKD	MDA	ROM	SCG
$E_{od}^{\dagger}$	168.8	208.2	199.4	207.4	195.4	149.8	206.6	185.5
e_{od}	33.2	51.1	35.2	39.1	47.2	28.0	41.4	29.3
N_{od}	101.8	104.0	121.5	121.1	100.7	95.2	116.8	120.8

$^{\dagger}E_{od}$-% change in total trade flows, e_{od}-% change in the intensive margin of trade, N_{od}-% change in the extensive margin of trade.

margin of exporting is both statistically and economically significant (compare row N_{od} to row e_{od} in Table 2).

According to previous studies, the aggregate trade growth might be more or less valuable for trading partners because the induced welfare impacts depend on how exactly does the trade growth occur - along the intensive or the extensive margin of trade (*Hillberry* and *Hummels* 2005, *Hummels* and *Klenow* 2005). While the ability to account for both these aspects is one of the main strengths of our approach, data limitations do not allow us to perform more detailed welfare analysis along sectoral, regional and socio-economic trade components, which is a promising avenue for future research.

Our results suggest that if trade growth along the extensive margin is more valuable than trade growth along the intensive margin, then the largest gainers from the trade integration in the Balkans were Serbia and Montenegro. In these countries the export volume grows more than four times faster along the extensive margin than the intensive margin of trade. The export growth along the extensive margin is about twice as high as export growth along the intensive margin in Bosnia and Herzegovina and Macedonia. Considering both these aspects, the aggregate trade growth and growth share of extensive margin, the largest gainers might turn out to be countries, which do not head any of the two categories.

5.3 Impact of variable trade barriers

In the previous section we have assessed the impact of changes in aggregate trade costs, i.e. both fixed and variable trade costs were reduced simultaneously by the same percentage. This type of analysis is done in most empirical ex-ante trade policy studies of SEE. Compared to previous trade studies our model has the advantage that, in addition to assessing the aggregate impact of trade policy integration, it allows us to investigate how separate parts of total trade costs (fixed and variable trade costs) affect export growth. In this section the aggregate impact of declining trade barriers is decomposed into two separate parts: changes in export value due to declining variable trade costs and changes in export value due to declining fixed trade costs. We are interested in decomposing the aggregate impact of trade costs, because if these impacts turn out to be significant and asymmetric, then trade policy implications will be different too.

In this section we perform scenario simulations of reducing the per unit trade costs in 10% steps up to 30%. The obtained results are reported in Table 3. As above, the other two rows in Table 3 (e_{od} and N_{od}) separate out impacts of the two trade margins. The intensive margin (row e_{od} in Table 3) reports how much each existing exporter changes the size of its exports. The extensive margin (row N_{od} in Table 3) reports how much new entrants export (in the case of a reduction in trade barriers).

Table 3: Variable trade cost impact on export flows in the SEE

	ALB	BIH	BUL	CRO	MKD	MDA	ROM	SCG
$E_{od}^{\dagger}$	122.5	166.6	143.7	153.7	154.5	104.4	156.0	126.8
e_{od}	33.2	51.1	35.2	39.1	47.2	28.0	41.4	29.3
N_{od}	67.1	76.5	80.3	82.4	72.9	59.8	81.0	75.4

$^{\dagger}E_{od}$-% change in total trade flows, e_{od}-% change in the intensive margin of trade, N_{od}-% change in the extensive margin of trade.

According to simulation results reported in Table 3, the impact of variable trade costs reduction on aggregate export value in SEE is always positive, but smaller than the impact of reducing both variable and fixed trade costs. Decomposing the aggregate impact on exports suggests that in addition to adjusting the average export size of existing exporting firms, the set of exporters adjusts in the SEE economies too, which amplifies the aggregate impact of changes in variable trade costs. According to simulation results reported in Table 3, declining variable trade costs induce not only an increase in the average size of exporters (row e_{od} in Table 3), but also attract new firms to enter foreign markets and start exporting (row N_{od} in Table 3).

The simulated trade gains are different across the SEE countries ranging from +104.4% in Moldova to +166.6 in Bosnia and Herzegovina (row E_{od} in Table 3) suggesting that in some SEE countries (e.g. Bosnia and Herzegovina) the responsiveness of exports with respect to variable trade barriers is considerably higher than in other Balkan economies (e.g. Moldova). The cross-country trade cost differences alone would not lead to such sizeable differences in trade gains (i.e. export growth). According to the underlying theoretical model, the export responsiveness with respect to variable trade barriers depends on the elasticity of substitution, as well as on the degree of firm heterogeneity, γ.[28] In more homogeneous sectors (high γ) large productive firms represent a smaller fraction in the total set of firms (see equation 16) implying that the productivity threshold of exporting moves to a region, where the main mass of firms lies. As a result, in those countries with more homogenous manufacturing sectors, exports are rather sensitive to changes in variable trade costs because many firms enter, when variable trade costs decline. These results are in line with previous studies on industry concentration in the SEE, which suggest that manufacturing industries are less concentrated in the EU accession countries (Bulgaria, Croatia and Romania) than in other SEE economies (*Astrov* 2001).

5.4 Impact of fixed trade costs

According to *Hunting Technical Services* (*HTSPE* 2005), non-tariff trade barriers are extremely high in SEE. The proposed BFTA should significantly contribute to reducing non-tariff barriers of bilateral SEE trade. Given that in empirical trade cost literature (e.g. *Anderson* and *Wincoop* 2003) non-tariff trade barriers are usually associated with fixed trade costs, in this section we investigate how export flows in the SEE economies would be affected if the BFTA would mainly reduce fixed trade costs. As in the previous two sections, fixed trade costs are reduced in 10% steps up to 30%, which corresponds to 70% of the initial trade cost. The obtained simulation results are reported in Table 4.

Table 4: Fixed trade cost impact on export flows in the SEE

	ALB	BIH	BUL	CRO	MKD	MDA	ROM	SCG
$E_{od}^{\dagger}$	20.8	15.6	22.9	21.2	16.1	22.2	19.8	25.9
e_{od}	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
N_{od}	20.8	15.6	22.9	21.2	16.1	22.2	19.8	25.9

$^{\dagger}E_{od}$-% change in total trade flows, e_{od}-% change in the intensive margin of trade, N_{od}-% change in the extensive margin of trade.

[28] In trade models with representative firms the amount exported to a given country depends on how competitive a firm is against other foreign exporters. Differences in competitiveness due to trade costs have a greater or lesser impact on trade flows depending on whether the exported goods are more or less substitutable. If the exported goods are more substitutable (σ_k is high), the intensive margin of trade is strongly affected by even tiny changes in trade barriers, implying that the elasticity of substitution, σ_k, is the only parameter determining the adjustment speed in trade models with representative firms.

The simulation results reported in Table 4 suggest that fixed trade cost reduction has a positive, but a rather moderate impact on export flows. On average, reducing fixed trade costs by 10 % increases the aggregate export value by only 19%. Only in Serbia and Montenegro the induced trade growth is more sizable (+25.9%).

Why do changes in fixed trade costs have a considerably smaller impact on export flows compared to changes in variable trade costs? The answer can be found by considering the underlying theoretical trade model. According to equation (16), the intensive trade margin does not respond to changes in fixed trade costs (row e_{od} in Table 4). When fixed trade costs change, all adjustments work solely through the extensive margin (row N_{od} in Table 4).

Although, according to *HTSPE* (2005), non-tariff trade barriers are extremely high in SEE, the resulting trade gains from reducing these barriers (i.e. fixed costs) seem to be rather limited. These results suggest that either the payoff of reducing non-tariff barriers in the SEE trade would be rather limited or that we cannot straightforwardly associate fixed trade costs with non-tariff trade barriers, as is usually done in empirical trade literature. In order to obtain more insights about the impact of fixed trade costs on trade flows, these results urge for more empirical research on fixed trade costs.

6 Conclusions

In this paper we study how trade integration would affect the structure of export growth in the SEE economies. By decomposing aggregate trade flows into extensive and intensive margins of trade, we found that in contrast to previous SEE studies (*Messerlin* and *Miroudot* 2004, *Bussiere*, *Fidrmuc* and *Schnatz* 2004), trade policy liberalisation in the Balkans would primarily increase bilateral trade flows through a larger number of exported goods and varieties, but not by increasing the average value per shipment. These results complement the results of previous studies reporting that the extensive margin of trade is more significant for trade growth than the intensive margin (*Kehoe* and *Ruhl* 2002, *Hillberry* and *McDaniel* 2002, *Eaton et al* 2004, *Hummels* and *Klenow* 2005).

Our empirical findings are summarised in Table 5, which report average export growth rates for the SEE economies, when variable trade cost, τ_{od}, and fixed trade cost, FC_{od}, would decline. The results reported in Table 5 allows us to draw clear-cut policy conclusions. First, if the main policy objective is to increase the aggregate value of exports, then the most effective trade liberalisation policy in the Balkans would be to reduce variable trade costs, τ_{od}, by extending the proposed BFTA to areas such as improving rail and road infrastructure, with the goal to reduce per-unit shipping costs. Second, reducing solely fixed trade costs, such as non-tariff trade barriers and border-crossing bureaucracy, would induce only limited export growth. Therefore, fixed trade cost reduction is the second best policy option, if the main policy objective is to increase the aggregate value of exports. Third, if the main policy objective is to maintain and extend market share of a few large internationally competitive enterprises, then, depending on the relative reduction costs (governments may face different budget costs of reducing 10% variable vs. 10% fixed trade costs), the most effective trade policy might be reducing fixed trade costs, FC_{od}.

We also found that reducing variable trade costs has a quantitatively larger impact on export growth than reducing fixed trade costs by the same percentage. The largest gainers from reducing variable trade costs were Bosnia and Herzegovina, Bulgaria, Croatia, Macedonia and Romania. However, if the proposed BFTA would mainly reduce fixed costs of trade, such as non-tariff trade barriers, the largest gainers were Albania, Bulgaria, Croatia, Moldova, Romania, Serbia and Montenegro. These results suggest that Bulgaria, Croatia and Romania would be among the largest winners of free trade area in the Balkans under either scenario.

Table 5: Average impact of fixed and variable trade cost reduction on export growth in the SEE

	Trade cost decline (-10%)	
	variable, τ_{od}	fixed, FC_{od}
Export growth (%)		
Intensive margin, e_{od}	+38.0	±0.00
Extensive margin, N_{od}	+74.4	+20.5

Trade costs – exogenous, export growth – endogenous.
Source: own calculations based on Tables 2−4.

Turning to potential deficiencies of our approach, we identify two issues which could be further developed in future research. First, the assumptions of identical technologies across countries and proxying country size by the employed labour force need to be reconsidered. Indeed, relaxing these two assumptions in a heterogenous firm framework would be a promising area for future research and, in the context of our study, potentially lead to different welfare implications of trade liberalisation. In fact, some progress has already been made in this area since this study was completed. For example, *Bernard et al* (2007) assume that in addition to heterogeneous productivity of firms and country differences in factor endowment, countries also differ in terms of relative factor abundance and industries vary in terms of relative factor intensity. Findings of *Bernard et al* suggest that falling trade costs induce reallocation of resources both within and across industries and countries, which magnifies the comparative advantage and creates additional welfare gains from trade. Given that these additional welfare gains of resource reallocation found by *Bernard et al* (2007) may be unevenly distributed between the intensive and extensive margins of trade, the *Bernard et al* framework might potentially lead to different welfare and policy conclusions.

The second issue which need to be addressed in future research is the consistency of econometric specification with the theoretical model. Although, the reduced form model, which we estimated in section 4, is largely derived from the theoretical trade model (section 3) and most of the estimated parameter values are significant and robust, the consistency of the two specifications might still be a critical issue. The consistency problem of the two specifications could be circumvented, for example, if either the reduced form gravity model would be used for the ex-ante trade policy impact analysis; or if model parameters would be estimated (calibrated) within the general equilibrium trade model. Although, both techniques are widely used in ex-ante trade policy impact assessment studies, they suffer from other limitations, which we discussed in above.

References

[1] Anderson, J. E. and E. van Wincoop (2003), “Gravity with Gravitas: A Solution to the Border Puzzle,” American Economic Review, 93: 170-192.

[2] Anderson J. E., E. van Wincoop, (2004), "Trade Costs," Journal of Economic Literature, vol. 42 (3): 691-751.

[3] Arellano, M. and B. Honore (2001), "Panel Data Models: Some Recent Developments," Handbook of Econometrics, Volume 5, Chapter 53.

[4] Astrov V. (2001), Structure of trade in manufactured products between Southeast European countries and the EU, WIIW Report No 8-9.

[5] Baldwin R., R. Forslid , Ph. Martin, G. Ottaviano and F. Robert-Nicoud. (2003), Economic geography and public policy. Princeton, N.J., Princeton University.

[6] Bergin, P. R., R. Glick (2005), Tradability, Productivity, and Understanding International Economic Integration, NBER Working Paper No. 11637.

[7] Bernard A.B., and J. Wagner, (1998), "Export Entry and Exit by German Firms," NBER Working Papers 6538, National Bureau of Economic Research.

[8] Bernard, A. B. and J. B. Jensen (1999), ”Exceptional exporter performance: cause, effect, or both?”, Journal of International Economics, 47: 1-25.

[9] Bernard, A., J. Eaton, J. B. Jensen and S. Kortum (2003), “Plants and Productivity in International Trade”, American Economic Review, 93: 1269-1290.

[10] Bernard, A. B., S. J. Redding and P. K. Schott (2007), "Comparative Advantage and Heterogeneous Firms", Review of Economic Studies, 74 (1): 31-66.

[11] Bussiere, M., J. Fidrmuc and B. Schnatz (2004), Trade Integration of the New EU Member States and Selected South Eastern European Countries: Lessons from a Gravity Model, Austrian National Bank, Discussion Paper.

[12] Chaney, T. (2007), "Distorted Gravity: The Intensive and Extensive Margins of International Trade," University of Chicago, mimeo.

[13] Eaton, J., S. Kortum and F. Kramarz (2004), "An Anatomy of International Trade: Evidence from French Firms" mimeo, NYU.

[14] Egger, P. (2000), "A note on the proper econometric specification of the gravity equation," Economics Letters, vol. 66 (1): 25-31.

[15] European Commission (2005), The Western Balkans in transition, Directorate-General for Economic and Financial Affairs, Occassional Paper No 5.

[16] Eurostat (2005), External Trade of Western Balkan Countries, Statistical Office of the European Communities, Luxembourg.

[17] Evenett, Simon J. and Anthony J. Venables (2002), "Export Growth in Developing Countries: Market Entry and Bilateral Trade Flows," Mimeo.

[18] Hunting Technical Services PE (HTSPE) (2005), Helping to Tackle Non-tariff Barriers in the Western Balkans, The European Union's CARDS Programme for the Western Balkans.

[19] Head, K. and T. Mayer (2004), The Empirics of Agglomeration and Trade, in J. V. Henderson & J. F. Thisse (ed.), Handbook of Regional and Urban Economics, 1 (4): 2609-2669.

[20] Helpman, E., M. Melitz and Y. Rubinstein (2007), "Trading Partners and Trading Volumes," NBER Working Paper 12927.

[21] Hillberry, R., and Ch. McDaniel (2002), "A Decomposition of North American Trade Growth since NAFTA." International Economic Review, 43 (2): 1-6.

[22] Hillberry, R. and D. Hummels (2005), Trade Responses to Geographic Frictions: A Decomposition Using Micro-Data, NBER Working Paper 11339.

[23] Hummels, D. and P. Klenow (2005), "The Variety and Quality of a Nation's Exports," American Economic Review, 95: 704-723.

[24] Kehoe, T. J. and K. J. Ruhl (2003), "How Important is the New Goods Margin in International Trade?" Federal Reserve Bank of Minneapolis Staff Report 324.

[25] Krugman, P. (1980), "Scale Economies, Product Differentiation, and the Pattern of Trade", American Economic Review, 70, 950-959.

[26] Mátyás, L. (1998), “The gravity model: Some econometric considerations.” The World Economy, 21 (3): 397-401.

[27] Melitz, M. J. (2003), "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," Econometrica, 71: 1695-1725.

[28] Melitz, M.J. and G.I.P. Ottaviano (2005), Market Size, Trade, and Productivity. NBER Working. Paper no. 11393. Cambridge, Mass.: National Bureau of Economic Research.

[29] Messerlin P.A. and Miroudot S. (2004), "Trade Liberalisation in South East Europe: Review of Conformity of 23 FTAs with the MoU", Stability Pact of South Eastern Europe.

[30] Roberts, M. J. and J. R. Tybout (1997), "The Decision to Export in Colombia: An Empirical Model of Entry with Sunk Costs," American Economic Review, 87 (4): 545-564.

[31] Ruhl, K. (2003), "Solving the Elasticity Puzzle in International Economics", Mimeo, University of Minnesota.

[32] Samuelson, P. (1954), “The Transfer Problem and the Transport Costs, II: Analysis of Effects of Trade Impediments,” Economic Journal, 64 (2): 264–289.

[33] Stubos, G., Tsikripis I. (2007), Regional Integration Challenges in South East Europe, Southeast European and Black Sea Studies, 7 (1): 57-81.

[34] WIIW (2005), Handbook of Statistics: Countries in Transition 2005, The Vienna Institute for International Economic Studies, Vienna.

[35] WIIW (2006), Handbook of Statistics: Countries in Transition 2005, The Vienna Institute for International Economic Studies, Vienna.