Verti-zontal Differentiation in Export Markets*

Francesco Di Comite† Jacques-François Thisse‡ Hylke Vandenbussche§

September 29, 2012

We develop a new model of monopolistic competition that rationalizes the following empirical regularity: firm-level prices of products are strongly correlated across destinations, whereas product sales are not. In our setting, varieties enter preferences asymmetrically through horizontal and vertical differentiation. Equilibrium prices and quantities depend on new market aggregates which can be used to improve existing indicators of export performance. Last, we propose an identification strategy to separate quality and taste at the firm-product level.

JEL codes: D43 - F12 - F14 - L16

Keywords: Heterogeneous firms - Horizontal differentiation - Vertical differentiation - Monopolistic competition - Non-symmetric varieties

*We thank A. Bernard, G. Barlevy, T. Besedes, M. Chen, M. Crowley, E. Dhyne, J. Eaton, C. Eckel, M. Francesconi, G. Grossman, J. Harrigan, H. Kee, K. Krishna, M. Lovely, V. Lugovskyy, J. Martin, F. Mayneris, D. Mitra, J. Morrow, P. Neary, D. Nguyen, E. Oberfield, G. Ottaviano, M. Parenti, G. Peri, P. Picard, D. Richardson, B. Roberts, M. Roberts, A. Rodriguez-Clare, J. Tybout, E. Verhoogen, J. Vogel, X. Wauthy, and seminar participants at Columbia University, Chicago Federal Reserve, Georgia Tech, Emory, Penn State, Syracuse University, UCD Dublin, LSE- London, University of Virginia, the Worldbank and Aarhus for helpful discussions and suggestions. This research was financially supported by UCLouvain-ARC 09/14-019 and KULeuven-PF 10/003. We thank the National Bank of Belgium for the data provided.

†Department of Economics (IRES), Université Catholique de Louvain, and European Commission, DG ECFIN, Avenue de Beaulieu 1, 1150 Bruxelles, Belgium - Email: francesco.di-comite@ec.europa.eu

‡CORE, Université Catholique de Louvain, Voie du Roman Pays 34, 1348 Louvain-la-Neuve, Belgium - Email: jacques.thisse@uclouvain.be - Tel: +32 (0)10 474312

§National Bank of Belgium; CORE and Department of Economics (IRES), Université catholique de Louvain, Place Montesquieu 3, 1348 Louvain-la-Neuve and KULeuven, Licos, Belgium. Email: hylke.vandenbussche@uclouvain.be - Tel: + 32 (0)10 474137

1 Introduction

Firm efficiency is very important in explaining firms' entry into export markets.[1] However, this seems less the case for firm-level sales in different markets conditioning upon entry.[2] Using a large firm-product-country data for Belgian exporters, we establish a few robust stylized facts, which confirm this. First, when we compare export prices of Belgian firms selling the same variety to different destinations, we find that prices are highly correlated across markets. Highly-priced varieties in one market are also highly priced in other markets, although absolute price levels differ. Second, sales of varieties in different markets are weakly correlated. A particular variety that sells well in one destination market need not sell much in another. And third, in the same destination market, prices and quantities of exported varieties are also weakly correlated. This holds for all the varieties and geographical destinations.

These observations are hard to reconcile with the assumption that all varieties face the same demand in every market, with differences in performance depending only upon the cost at which each variety is produced. Heterogeneity in firm efficiency alone predicts a strong negative correlation between prices and quantities in a market and a strong positive correlation between prices and between quantities (sales) across markets. Instead, we find price-quantity correlations to be very weak in the same destination market whereas price correlations across markets are much stronger than quantity correlations.

Several papers analyzing the variability in firm-level prices and sales across a range of export destinations have reached the same conclusion: cost factors alone cannot account for all the variation in the data. Previous attempts to model additional heterogeneity allowing models to better fit the empirical evidence are those that augment firms' differences in productivity with quality differences.[3] While a combination of cost and quality heterogeneity can explain the low correlation within a market between prices and quantities, it would also predict quantity rankings to be just as regular as price rankings. This is not what we observe in the data since there appears to be a source of heterogeneity affecting quantities but not prices.

The main contribution of this paper is to provide a microfoundation for this source of heterogeneity. More generally, rather than assuming a *deus-ex-machina* type of shock, we propose an encompassing model that allows for a clear economic interpretation of all parameters.

We are not the first to point out the need for a more general demand setting. Based on French firm-level evidence, Eaton, Kortum and Kramarz (2011) report variation in the sales performance of the same firms in different markets, thus suggesting the existence of an additional source of variation on the demand side. They find that only half of the variation across firms can be attributed to efficiency. This suggests that conditional upon entry, firm efficiency, while important, is not the only determinant of sales variation across countries. Brooks (2006), using Colombian export data, makes a similar point. Also, Kee and Krishna (2008) find that the correlation between firm-level sales of Bangladesh firms in different destination markets is close

[1] See, for example, Aw, Chung and Roberts (2000), Bernard, Eaton, Jensen and Kortum (2003), Melitz (2003), and Arkolakis (2010).

[2] Examples are Eaton, Kortum and Kramarz (2011) and Manova and Zhang (2012).

[3] See Kugler and Verhoogen (2008), Hallak and Sivadasan (2009), Baldwin and Harrigan (2011), and Eckel, Iacovone, Javorcik and Neary (2011).

to zero and conclude that only demand shocks can explain these facts. These papers therefore stress the need for additional sources of variation but, unlike us, do not go as far as to offer a way to identify the demand heterogeneity.

Only a few theory models in the trade literature have accounted for demand factors. Either these models build on CES models of monopolistic competition,[4] or stem from discrete choice theory.[5] Such models are important to understand the patterns of trade and complement supply-side-oriented models. However, it is hard to disentangle empirically the different sources of variability at work in these models. Horizontal differentiation in CES models is associated with the elasticity of substitution and is constant across varieties. Therefore it cannot explain variation in sales for the same firm-product across countries. To remedy for the constant elasticity of substitution in the CES, one can introduce a firm-product specific demand shock per country that accounts for sales variation of the same firm-product across countries without affecting prices, as in Bernard, Redding and Schott (2010). Horizontal differentiation between products is then the combination of a constant parameter of substitution and a variable shock at the firm-product level. But since the parameter of substitution also enters the price equation, a clear separation of horizontal and vertical differentiation is difficult to attain with those functional forms.

What appears to be currently lacking is a model that allows for a clear distinction between horizontal and vertical differentiation. This can be particularly useful for researchers interested in identifying one or the other, or both. This is where we aim to contribute. The introduction of horizontal differentiation in our model differs in at least two ways from the current literature. First, it is captured by one single parameter that varies across each firm-product-country combination for which we provide a micro-foundation that goes back to spatial models of product differentiation à la Hotelling (1929). Second, we show how *taste* and *quality* can be separated in determining the market outcome and how they can be empirically identified by any researcher with access to data on firm characteristics. Following Hallak and Schott (2011), we define vertical differentiation to be *any tangible or intangible attribute of a good that increases consumers' willingness to pay for it.*[6] The indicator of horizontal differentiation that we propose here, the taste parameter is cast within a Lancasterian framework, which highlights how this parameter affects demands and sales asymmetrically.

We build on Melitz and Ottaviano (2008) in terms of consumer preferences. Our data on Belgian exporters confirm the prediction that fob export prices of the same firm-product vary across markets (see also Manova and Zhang, 2012; Martin, 2012). With CES preferences, fob export prices only vary by distance, if specific transport costs are assumed, as shown by Hummels and Klenow (2005) and Martin (2012). However, for the majority of shipments from Belgium, distances to destination markets are similar. This suggests that the variation in the fob export prices in our Belgian data is likely to be driven by other factors.[7] The quadratic preferences

[4]For example, see Crozet, Head and Mayer (2012).

[5]See Fajgelbaum, Grossman and Helpman (2011); Katayama, Lu and Tybout (2009); Khandelwal (2010) or Verhoogen (2008).

[6]While Hallak and Schott (2011) use sector-level data, the method we propose here is suited for those with access to firm-level data.

[7]Baldwin and Harrigan (2011) for the US also observe that other factors are at work, as distances below 4000 km do not affect fob prices.

allow for country-specific competition effects, which offers a more plausible explanation for the observed variation in price levels.

The objective of this paper is to bring together different sources of heterogeneity in an encompassing model that can better explain the stylized facts mentioned above. The sources of heterogeneity currently identified by the literature and included in the model we propose are: vertical differentiation (Baldwin and Harrigan, 2011; Eckel, Iacovone, Javorcik and Neary, 2011; Khandelwal, 2010; Kugler and Verhoogen, 2008), horizontal differentiation (Bernard, Redding and Schott, 2010; Eaton, Kortum and Kramarz, 2011; Kee and Krishna, 2008), and cost heterogeneity (Melitz, 2003), which altogether generate new country-specific competition effects.

It is worth stressing that our model is different from Melitz and Ottaviano (2008) in a few important ways. Melitz and Ottaviano assume that consumers are endowed with the same and symmetric demand across countries. The only source of variation comes from the supply side where firm-products differ in marginal costs. Our model assumes that demand for a firm-product is country-specific. This allows for a richer parameterization on the demand side through the introduction of *asymmetric preferences*, while continuing to assume that firm-products differ in efficiency. The predictions arising from our model are thus quite different. One difference lies in the prediction on prices and quantities of firm-products and how they are correlated within destination markets. The model proposed by Melitz and Ottaviano implies a strong negative price-quantity correlation within destination markets, where the model we propose here can rationalize the weak correlation observed. In addition, their model implies strong quantity and price correlations across markets. Instead, our model predicts a strong price correlation across markets but a weaker quantity correlation, stemming from the fact that consumers in each country are allowed to have different tastes and evaluate the horizontal attributes of a product differently, hence buying different amounts.

The low quantity correlation across markets is a prediction that our model has in common with Bernard, Redding and Schott (2010), Eaton, Kortum and Kramarz (2011), and Kee and Krishna (2008). However, our model is different from these models in allowing fob prices of the same firm-product to vary across markets, a feature prominently present in firm-level datasets. Our model also differs from the others in allowing products to be vertically differentiated as in Baldwin and Harrigan (2011), Eckel, Iacovone, Javorcik and Neary (2011), Kugler and Verhoogen (2008) and Feenstra and Romalis (2012), among others. However, in contrast to these models, we do not require any prior relation between marginal costs and willingness to pay. As such we allow for the possibility of quality resulting from fixed investments in research and development or advertising. In the same spirit of most models of monopolistic competition we treat taste, quality and cost as exogenous parameters, but endogenizing these parameters would not overturn our results.

The identification of horizontal versus vertical differentiation is linked to the functional form that we consider. The use of quadratic preferences appears more suited for this than the CES because it allows for a clear separation of the two dimensions of differentiation. The parameter of horizontal differentiation (*taste*) which we identify from a quadratic utility framework has the appealing feature, from an empirical point of view, that it does not affect prices directly but only affects market shares.

Fajgelbaum, Grossman and Helpman (2011) combine horizontal and vertical differentiation within a new and rich framework which encompasses two quality levels, high and low, with idiosyncratic tastes captured through the use of a random term. The functional form they work with is the logit, which is symmetric in nature and as such does not allow for asymmetric preferences as we pursue here. Given that the logit is a close relative of the CES (Anderson, de Palma and Thisse, 1992), the parameter capturing horizontal differentiation enters the price equation derived under a logit specification. However, this is not an attractive property for those researchers who want to separate price from non-price attributes of products. The functional form we work with allows us to do so. Another major contribution of Fajgelbaum, Grossman and Helpman (2011) is that they are able to cope with consumer/income heterogeneity within each country. This is more difficult to achieve in the model we develop, although it can deal with income differences across countries. However, we do not see this as a major weakness of our approach because empirical researchers in trade typically have one observation for each product per destination country in their data and cannot distinguish between the consumers' types within a country. As a consequence, we find it reasonable to work with a representative consumer per country.

Our model also has some appealing features from the theoretical viewpoint. First, it is rooted in Lancaster (1979), which arguably provides the best analytical setting to study product differentiation, especially when dealing with asymmetric varieties through their precise positioning in the characteristics space. Ever since Hotelling (1929), two varieties of the same good are defined as horizontally differentiated when there is no common ranking of these varieties across consumers. By contrast, two varieties are vertically differentiated when all consumers agree on their ranking (Gabszewicz and Thisse, 1979). Combining these two types of differentiation, the encompassing model we present generates a set of predictions that can rationalize the micro trade patterns that we and others observe and that introduces a separate source of variation affecting sales but not prices within firm-products across destinations. This leads us to refer to it as a model of *verti-zontal* differentiation whose main purpose is to propose a richer parameterization on the demand side.

In addition, our model holds important insights from the industrial organization literature on product differentiation in which there has been a long tradition of distinguishing vertical from horizontal differentiation because they generate very different results. However, unlike industrial organization models which emphasize strategic interactions between firms, our approach focuses on *weak interactions*, meaning that firms' behavior is influenced only by taste-weighted price, cost and quality indices, which individual firms cannot affect. For example, the equilibrium outcomes are strongly (weakly) affected by the mass of varieties which have a good (bad) match with consumers' ideal varieties, as captured by the taste parameter. In doing so, we are able to determine how the degree of competition, whence the market outcome, is affected by the various differentiation parameters.

Without a clear separation between horizontal and vertical differentiation, researchers are likely to misinterpret high sales conditioning on prices and confound quality with consumer taste. Once we allow some varieties within a market to match local taste better, specific varieties can sell more than others even when offered at the same price and quality. But more importantly, once vertical differentiation is properly identified, it can be used to correct aggregate price indices. The verti-zontal model offers a convenient way to obtain a micro-level quality and taste measures at the firm-product level, which may then be used for macro-economic purposes to improve measurement of competitiveness indicators at higher levels of aggregation.

Empirically, we compare fob export prices and sales of more than 24,000 firm-product combinations exported by Belgian firms to different destination countries in a particular year. This cross-section allows us to compare prices and quantities of the same varieties across destination markets. The model we present provides a rational for the strong price correlation across markets which is a feature also present in models with cost or quality heterogeneity under CES (Melitz, 2003) or quadratic preferences (Foster, Haltiwanger and Syverson, 2008). But whereas existing models would also predict a similar strong correlation for quantity rankings across destination markets, the model presented here rationalizes quantity differences across countries as a result of different consumer taste for each firm-product in each destination market. Given that our main focus lies on the demand side, we disregard issues related to market participation and consider only products that are present in all the destination markets in our empirical analysis. In addition, aspects related to the multi-product nature of firms, such as cannibalization, while not included in the preferences could also be incorporated but are outside the scope of the current paper (see, e.g., Eckel and Neary, 2010, Dhingra, 2011, and Arkolakis and Muendler, 2010).

The next section presents the stylized facts obtained on a unique dataset of Belgian exporters with exports by product and country. Section 3 describes in detail our model and explains how it can rationalize the above stylized facts. Section 4 shows how to empirically identify its parameters. Section 5 tests the model's key assumptions and compares our findings to other models. Section 6 concludes.

2 Data and stylized facts

2.1 Data

We use a unique dataset on trade flows of Belgian exporters. The data is composed of fob (free on board) export prices and quantities by destination market at the firm-product level.[8] This allows us to compare prices and quantities of the same firm-products across destination markets as well as prices and quantities of different firm-products within the same destination market.

The Belgian export data are obtained from the National Bank of Belgium's Trade Database, which covers the entire population of recorded annualized trade flows by product and destination at the firm-level. Exactly which trade flows are recorded (i.e. whether firms are required to report their trade transactions) depends on their value and destination. For extra-EU trade, all transactions with a minimum value of 1,000 euros or weight of more than 1,000 kg have to be reported. For intra-EU trade, firms are only required to report their export flows if their total annual intra-EU export value is higher than 250,000 euros. The export data are recorded at the year-firm-product-country level, i.e. they provide information on firm-level

[8] Prices are unit values obtained by dividing values by quantities with the latter expressed in weight (kilograms) or units, depending on the product considered.

export flows by 8-digit Combined Nomenclature (CN8) product and by destination country.[9] Due to its hierarchical nature, all products expressed as CN8 are also classified as products at more aggregated levels. Incidentally, CN6 is identical to the HS 6 digit classification, which is the international product classification.[10] For firms with primary activity in manufacturing, the data includes over 5,000 exporters and over 7,000 different CN8 products, resulting in more than 60,000 firm-product varieties (firm-product combinations at the CN8 level) exported to 220 destination markets in a total of almost 250,000 observations in one year. We use cross-sectional export data for the year 2005 from manufacturing firms and for which both values and weights (or units shipped) are reported which allows us to compute prices

Given that our functional form is mainly suited for consumption goods, we only consider consumption goods as indicated by the BEC classification.[11]

2.2 The Chocolates Example

To illustrate our methodology we first turn to an example. A product that is frequently exported from Belgium and included in our data is chocolates. The three main destination markets for Belgian chocolates are Germany, France and the Netherlands. Considering the 6-digit classification (CN6=HS6) of chocolate, in 2005 there were 94 different chocolate varieties being exported from Belgium to all three markets. In Figure 1a each dot represents a combination of a price and quantity rank in a particular geographical market for a particular Belgian chocolate variety.[12]

[INSERT FIGURE 1a HERE.]

If one assumes, as most trade models implicitly do, that all chocolate varieties face the same demand in every market, and that the only difference between varieties is the cost at which they are produced, one would expect all observations to lie around the diagonal from top-left to bottom-right. Put differently, one would expect high-cost chocolates to rank high in the price ranking (close to the origin on the price axis) with few people buying them (top-left area of the figure). Low-cost chocolates, on the other hand, would sell a lot at a low price (bottom-right area of the figure). If instead one assumes that quality is the only source of heterogeneity and acts as a demand shifter, one would expect observations of different Belgian chocolate varieties to cluster around the diagonal running from bottom-left to top-right. One would expect high-quality chocolates to be highly priced and to sell a lot, while low-quality chocolates would be associated with low prices and would sell poorly in all markets.

Interestingly, Figure 1a shows that there is a very weak correlation pattern between price and quantity rankings.[13] This suggests that a particular chocolates variety with the same price

[9]The Combined Nomenclature is the European Union's product classification, with 8 digits being the most detailed level.

[10]The CN classification can be downloaded from the Eurostat Ramon server: http://ec.europa.eu/eurostat/ramon/.

[11]The BEC classification is an indicator of consumption goods at the 6 digit level. Thus, goods in sector CN8 and sector CN6 are easy to classify. However turning to more aggregate sectors like sector CN2, both consumption and other (capital, industrial) goods may occur. Our decision rule has been to include sectors CN2 and sectors CN4 when there was at least one CN6 consumption product.

[12]For example Neuhaus, Godiva, Leonidas, Guylian etc. are all examples of Belgian chocolate varieties.

[13]On average, the correlation between price and quantity rankings of chocolates within markets is around -39%, with rank correlations ranging from -48% in Germany to -28% in France, through -41% in Netherlands, all being significantly different from -1.

rank in the different markets, can sell relatively well in one market but badly in another. Such a pattern is inconsistent with a model where the only source of heterogeneity among firms is either productive efficiency or quality. Consequently, an important first observation arising from the chocolates market is that more than one source of heterogeneity appears to be present in the micro-level data.[14]

A second important observation arises from plotting price rankings in different countries, which we do in Figure 1b. Each dot in the figure now represents the ranking of a variety in a particular destination market, compared to the ranking of that chocolates variety in Germany (on the horizontal axis), in such a way that a perfect correlation between price ranks across markets would result in dots following the 45° line. Pairwise price rank correlations are surprisingly high, ranging from 79% between France and the Netherlands to 83% between the Netherlands and Germany. A strong and positive price correlation between markets corresponds to the prediction arising both from a pure cost and a pure intrinsic quality model.

[INSERT FIGURE 1b HERE.]

It is important to note that even though the price correlations between any pair of destination countries are high, prices in levels differ substantially. The average standard deviation of the 94 different chocolates varieties exported to France, Germany, Netherlands in the CN6 category is 2.41. Given that the average export price for Belgian chocolates in the sample is 8.31, the standard deviation is therefore roughly 30% of the average price. Considering that the three destinations are neighbouring countries with arguably similar shipping costs from Belgium, this also suggests that prices for the same variety differ across destinations.

The chocolate example is not an exception. Computing the ratio of the standard deviation over the average price of a variety across markets, we find that mean of this ratio across varieties is equal to 60% while the median is 40% when we consider unit expressed in weights.[15] Note also that firms charge the same fob price irrespective of the export destination in less than 1% of the cases, showing that virtually all firms price discriminate across markets. These findings are consistent with Gorg, Halpern and Murakozy (2010), who work with similar Hungarian export data.

A third observation arising from the exports of Belgian chocolates stems from Figure 1c. There we plot quantity rankings of chocolates varieties in the three countries in a similar way as we plotted price rankings. The pattern arising from quantity ranking is rather different from the price ranking. Pairwise quantity rank correlations for chocolates average 61%, ranging from 56% to 67%, which is much less than the corresponding price rank correlations, suggesting that price rankings of chocolate are much more stable than quantity rankings across markets.

[INSERT FIGURE 1c HERE.]

Existing trade models incorporating cost and/or quality heterogeneity suggest quantity rankings that are just as regular as price rankings. The above discussion tells us that the opposite

[14]Similar conclusions are reached through more formal analyses by authors such as Brooks (2006); Crozet, Head and Mayer (2012); Hummels and Klenow (2005).

[15]The corresponding numbers when the unit values are expressed in units are respectively 57% qnd 41%

holds. This suggests the presence of a third source of heterogeneity affecting quantities which is not only variety-specific but also market-specific (using US trade data, a similar point is made by Bernard et al., 2011). In other words, heterogeneity on the supply side needs to be supplemented with heterogeneity on the demand side. In this paper we rationalize this quantity variation through the existence of differences in consumer taste which we allow to differ for each variety in each destination market.

Building on this preliminary observation in the chocolates sector, we now discuss evidence based on a detailed micro-level dataset on Belgian exporters and show that the empirical regularities illustrated above turn out to be extremely robust and hold in virtually all markets and products considered.

2.3 Looking at prices and quantities within and between markets

In what follows, we explain the products and destination markets that have been included in our analysis. Their intersection determines the product-market samples on which price and quantity comparisons are conducted.

Since CN8 is the most detailed product-level classification available, we define a *variety* s as a firm-CN8 combination. Our definition of a variety does not change throughout the analysis, but the size of the product-market S_i is allowed to change. When defining a relevant product-market, the level of product aggregation must be traded off against the number of varieties, which falls dramatically as the product-market narrows. For this reason, we repeat our analysis for four levels of product-market aggregation. Next to the CN8, we also verify results for the CN6, CN4 and CN2 level. In a more aggregated product classification than CN8, a product-market will then be defined as a collection of varieties (firm-CN8) sharing the same CN code and sold in the same market.[16]

To ensure that there are enough varieties in enough product-markets, we retain the five products which are associated with the highest number of varieties exported. These products are listed in Table 1 with corresponding CN codes and descriptions.

[INSERT TABLE 1 HERE.]

Since our analysis focuses on price and quantity variations not just within but also across destination markets, another trade-off involves the number of countries to consider. Since we are interested in price and quantity differences across markets, we need a sufficient number of markets to compare. However, we also need a sufficient number of varieties to be simultaneously sold in all the markets considered. The trade-off arises because the number of varieties simultaneously present in all markets drops significantly with each additional destination market. As there is no clear-cut rule to settle this issue, we follow a data driven approach, the aim of which is to retain a set of countries and products that allow for *the maximum number of observations to base our analysis on.* We start by considering only those destination markets that are important outlets

[16] An example can clarify this. Belgian chocolates at the CN8 are given by the code 18069019. In our analysis, the product-market for Belgian chocolates at the CN8 level then consists of all Belgian exporters exporting 18069019 which results in 34 varieties (firm-products) being exported from Belgium. By defining the product-market at a higher level of aggregation, say the CN6 we then include all firm-CN8 in that fall within the CN6 180690. This results in 94 varieties (firm-product) being exported from Belgium.

for Belgian exporters in terms of the number of firm-products. This leads us to include only those destination markets that import at least 5,000 varieties. This results in 12 destination markets, which are listed in Table 2. Next, we explore all possible market combinations to find how many varieties are exported simultaneously to $N = 2, 3, ..., 12$ countries and, for each value of N we identify a *best N-market combination*. In the first column of Table 2, we report the number of varieties shipped to each of these 12 markets. The second column gives the total number of varieties sold simultaneously in each best N-market combination, which is obtained by adding the corresponding country to all the countries listed in the previous rows.

[INSERT TABLE 2 HERE.]

The combination of markets and products described above leads to 220 potential data samples (a combination of 5 products, 4 levels of aggregation, 11 markets) maximizing the number of observations for the analysis. However, even following this procedure, some samples are still too small to be used, having just 2 or 3 varieties. We then further restrict ourselves to samples with more than 10 varieties, in order to permit a meaningful correlation analysis between markets. This results in 171 effective samples as shown in Table 3 which is what we will base our analysis on.

[INSERT TABLE 3 HERE.]

We start by considering rank correlations of prices and quantities within and between markets. The use of rank correlations allows us to capture general features of the data, even in the context of non-linear or non-additive demand functions. Put differently, by considering rank correlations we are not imposing any particular set of demand preferences, but we are just testing whether prices and quantities of a certain group of varieties keep their relative ranking structure in different markets. This will be relaxed in Section 5 where we show by means of regression that our findings also hold for actual prices and quantities.

Within-markets rank correlations. First of all, similarly to what we have shown on the chocolate example above, we investigate whether the rankings of prices and quantities of different varieties are significantly correlated within each market. In a model where only quality or only cost efficiency matters, they should be. If at least both elements are at play, then the relationship should be generally weak or insignificant, with the exception of sectors in which there is not much scope for quality or productivity differences. We thus investigate whether price and quantity rankings are significantly correlated within each of the samples resulting from our market and product selection. Results are given in Table 4, where we report them first for all the samples, and then disaggregated by product-market level of aggregation and number of countries included in the analysis. In particular we report how many times the within market price-quantity correlation is not significantly different from 0 at a 5% level of confidence out of the total number of samples considered.

[INSERT TABLE 4.]

Interestingly, results vary a lot depending on the level of aggregation and the number of destination markets considered. Evaluated in the narrowest product definition, the CN8 level, the rejection rate of a significant correlation of prices and quantities within markets is 76.3%. Over all the sample samples considered, at different level of CN aggregation, the data reject a significant correlation of prices and quantities within markets in about 35.3% of the times. These results seem to confirm the idea that any theory should at least involve two sources of heterogeneity to explain the pattern of prices and quantities observed in the data. This is most evident in narrowly defined product-markets.

Between-markets rank correlations. We now turn to statistics for quantity and price rank correlations between markets. It can be noted that quantity rank correlations between markets are often not significantly different from 0, at a 5% level. At the narrowest product-level, which is the CN8, the quantity correlations are not significantly different from zero in about 60.5% of the cases, which is almost as much as price-quantity correlations within markets. On the other hand, it is striking how much lower the rejection rates are for price ranks correlations between markets, as shown again in Table 4. The Spearman rank statistic points at prices being significantly correlated in 98% of cases. Put differently, the Spearman rank correlation rejects a correlation of price rankings across markets in only 2.9% of the cases, against the 60.5% of quantity rankings.

After having shown that price rank correlations are significant much more often than quantity rankings, we also look at the value of these price and quantity correlations between markets. For reporting purposes, in order to give the reader a sense of the pattern that emerges from all the pairwise correlations considered, we average the pairwise coefficients arising from comparing rankings in any two destination markets at the sample level and then average these sample coefficients by level of product-aggregation and market-combination.[17]

In Figure 2 we illustrate the results graphically. We see that *average price correlations are systematically higher than average quantity correlations*, for all the market combinations considered and at every level of product aggregation. The difference between them lies around 15 percentage points, which is relatively similar across the samples.[18]

[INSERT FIGURE 2 HERE.]

These results are not consistent with a combination of cost and quality heterogeneity alone. Price correlations between markets are high, suggesting that quality and/or productive efficiency are intrinsic to a variety rather than market-specific. Yet, quantity correlations are systematically lower, indicating that an additional source of heterogeneity is present at the market-variety level. In other words, a third source of heterogeneity needs to be accounted for when dealing with

[17] For example, when 3 markets are considered as in chocolates, 3 pairwise market correlations for prices and 3 for quantities are obtained; when 4 markets are considered, the coefficients are 6, and so on up to 12 markets, at which point 66 bilateral correlations are obtained. All the coefficients associated to each individual sample can be provided upon request.

[18] In addition, as a robustness check, we repeat the same correlation analysis considering the entire manufacturing sector. Results are similarly strong, i.e. high price correlation but low quantity correlation between markets.

micro-level trade data. In the model below we rationalize this by allowing consumer taste to differ for each firm-product in each destination market.[19]

2.4 A robustness check: does geography matter?

A legitimate concern is whether our results are not driven by the fact that most destination countries included in our analysis are European (see Table 2), which may have a dampening effect on price differences. If the high price correlation is the result of arbitrage or lack of border controls, this would drive the results and would not provide much information, although the large differences in export price levels across countries should rule out this possibility. However, as a consistency check, we investigate whether a different country selection could affect our results. We do so by considering a range of heterogeneous and most remote countries (Brazil, South Africa, Australia, Turkey, China, India, Japan, US, and Canada) together with the three main trading partners of Belgium (France, Netherlands and Germany). Out of the whole manufacturing sector, this selection of destination countries results in 87 varieties exported in 2005 to these 12 countries. The rank correlation pairs for these 87 varieties are plotted in Figure 3 for prices and quantities, sorting them by decreasing quantity rank correlation. The results are in line with earlier findings. Price rank correlations range between 84% and 97% for all the country pairs, while quantity rank correlations can be as low as 50%, averaging 71%. This result is reassuring since it confirms that prices are surprisingly similar across markets, even when including countries outside the European Union.

In fact, if anything it appears that the original samples containing mostly European countries may generate results against our modelling choices. This can also be seen from Figure 3. Of all the countries included in this new sample, the ones displaying the highest pairwise quantity rank correlations are the three European countries, with an average price rank correlations also above average. This means that, if anything, our original samples containing mostly European countries may have overestimated the regularity of quantities sold across markets and underestimated the real difference between price and quantity coefficients.

3 Re-thinking product differentiation in monopolistic competition: Chamberlin and Hotelling unified

In this section, we present a model that can rationalize the above-mentioned stylized facts. We do so by constructing an encompassing model where consumer preferences are quadratic and vertical and horizontal differentiation enter asymmetrically and in an unprecedented way. We feel it is important to offer a clear interpretation for the parameters that arise from the model and do so by pointing to their micro-foundations. This is useful for future researchers interested in applying the identification procedure for firm-product quality and taste laid out in Section 4.

Recall that two varieties of the same good are said to be horizontally differentiated when there is no common ranking of these varieties across consumers. In other words, horizontal

[19] We label this source of variation taste, but we do not exclude other interpretations proposed in the literature such as distribution channels (Arkolakis, 2010) or demand accumulation (Foster, Haltiwanger and Syverson, 2012), but consumer taste seems a plausible explanation and ultimately results in a convenient way to separate it from quality.

differentiation reflects consumers' idiosyncratic tastes. By contrast, two varieties are vertically differentiated when all consumers agree on their rankings. Vertical differentiation thus refers to the idea of quality being intrinsic to these varieties.[20] Such definitions of horizontal and vertical differentiation have hitherto been proposed for indivisible varieties with consumers making mutually exclusive choices. In what follows, we first formulate our model within the Lancasterian definitional setting and then generalize it to allow (i) consumers to buy more than one variety and (ii) the differentiated good to be divisible.[21] Defining horizontal differentiation when consumers have a love for variety is straightforward because such a preference relies on horizontally differentiated varieties. By contrast, defining vertical differentiation is more problematic because the ranking of varieties may change with consumption levels, an issue that we address below.

3.1 The one-variety case

Imagine an economy with one consumer whose income is y. There are two goods: the first one is differentiated while the second one is a Hicksian composite good which is used as the numéraire. Consider one variety s of the differentiated good. The utility from consuming the quantity $q_s > 0$ of this variety and the quantity $q_0 > 0$ of the numéraire is given by

$$u_s = \alpha_s q_s - \frac{\beta_s}{2} q_s^2 + q_0$$

where α_s and β_s are positive constants, which both reflect different aspects of the desirability of variety s with respect to the numéraire. The budget constraint is

$$p_s q_s + q_0 = y$$

where p_s is the price of variety s. Plugging the budget constraint in u_s and differentiating with respect to q_s yields the inverse demand for variety s:

$$p_s = \max\left\{\alpha_s - \beta_s q_s, 0\right\}. \tag{1}$$

In this expression, p_s is the highest price the consumer is willing to pay to acquire the quantity q_s of variety s, i.e. her willingness-to-pay (WTP). When the good is indivisible, the WTP depends only on α and β. Here, instead, it declines with consumption, following the decrease in its marginal utility. As long as the WTP for one additional unit of variety s is positive, a consumer chooses to acquire more of this variety. In contrast, she chooses to consume more of the numéraire when the WTP is negative. The equilibrium consumption is obtained when the WTP is equal to zero.

The utility u_s being quasi-linear, the above expressions do not involve any income effect. However, we will see below how our model can capture the impact of income differences across

[20]Our approach is consistent with assuming that quality is intrinsic to a firm (brand), as in Aw, Batra and Roberts, 2001 and Eckel and Neary, 2010. Indeed, the latter would imply that the quality parameter is the same across products supplied by the same firm which is a less strict assumption than the one used here where we allow quality to vary by firm-product.

[21]Note that our approach, like most models of monopolistic competition, abstracts from the way product characteristics are chosen by firms. This issue has been tackled in a handful of theoretical papers (Hallak and Sivadasan, 2009) and analyzed empirically by Eckel, Iacovone, Javorcik and Neary (2011).

markets.

3.2 The two-variety case: a spatial interpretation

Consider now the case of two varieties, whose degree of substitutability is captured by a parameter $\gamma > 0$. That γ is positive and finite implies that varieties are imperfect substitutes entering symmetrically into preferences. The utility of variety $s = 1, 2$ is now given by

$$u_s = \alpha_s q_s - \frac{\beta_s}{2} q_s^2 - \frac{\gamma}{2} q_s q_r + q_0 \tag{2}$$

where q_r is the amount consumed of the other variety.

In this case, $\alpha_s - \gamma q_r/2$ is the marginal utility derived from consuming an arbitrarily small amount of variety s when q_r units of variety r are consumed. This marginal utility varies inversely with the total consumption of the other variety because the consumer values less variety s when her consumption of its substitute r is larger. Note that the intercept is positive provided that the desirability of variety s (α_s) dominates the negative impact of the consumption of the other variety, q_r, weighted by the degree of substitutability between the two varieties (γ). As q_s increases, the WTP of this variety decreases and variety s is consumed as long as its WTP is positive.

Repeating the procedure to obtain the inverse demand as in (1), the WTP of variety s becomes

$$p_s = \alpha_s - \frac{\gamma}{2} q_r - \beta_s q_s. \tag{3}$$

Compared to (1), the WTP for variety s shifts downward to account for the fact that the two varieties are substitutes; the value of the shifter increases with the total consumption of the other variety and the degree of substitutability.

Following the literature, we define two varieties as vertically differentiated when consumers view the vertical characteristics of variety 1 as dominating those of variety 2. Therefore, in line with the definition of vertical differentiation, we say that varieties 1 and 2 are vertically differentiated when all consumers' WTP for the first marginal unit of variety 1 exceeds that of variety 2, i.e. $\alpha_1 > \alpha_2$. Because a higher α_s implies that the WTP increases regardless of the quantity consumed, α_s can be interpreted as an index of the *quality* of variety s. Since the WTP for a variety decreases with its level of consumption, an alternative definition would be to say that varieties 1 and 2 are vertically differentiated when $\alpha_1 - \beta_1 q > \alpha_2 - \beta_2 q$ for all $q > 0$. However, this definition overlaps with the very definition of the WTP that captures more features than vertical attributes. Furthermore, we will see that the equilibrium price of variety s always increases with α_s, which we find sufficient to express the idea that a higher quality variety is expected to be priced at a higher level.

We now come to the interpretation of parameter β_s. It is well known that the best approach to the theory of differentiated markets is the one developed by Hotelling (1929) and Lancaster (1979) in which products are defined as bundles of characteristics in a multi-dimensional space. In this respect, one of the major drawbacks encountered in using aggregate preferences such as the CES and existing quadratic utility models is that a priori their main parameters cannot be interpreted within a characteristics space.[22] This is why we find it critical to provide an unambiguous interpretation of β_s within the Lancasterian framework, such that each parameter of the model we develop here is given a precise and specific definition. In addition, the differentiated good being divisible in monopolistic competition, the interpretation of these parameters must be independent of the unit in which the good is measured.

Our spatial metaphor involves a continuum of heterogeneous consumers. Whereas in Hotelling's model they are assumed to make mutually exclusive purchases, in the verti-zontal model, consumers are allowed to visit several shops. In the spirit of spatial models of product differentiation, we first assume here that consumers buy one unit of the good in each shop they visit, an assumption that will be later relaxed.

In Figure 4, we depict a spatial setting in which two varieties/shops, indexed $s = 1$ and $r = 2$ respectively are located at the endpoints of a unit segment, where $\alpha_1 = \alpha_2 = \alpha$ and $\beta_2 = 1 - \beta_1 > 0$. Using (3), the WTP for, say, variety 1 has an intercept equal to $\alpha - \gamma/2$, while β_1 is the distance between shop 1 and consumers, the transport rate being normalized to 1. The consumer's WTP for variety 1 equals zero at

$$\beta_{\max} = \alpha - \gamma/2.$$

[INSERT FIGURE 4 HERE.]

Treading in Hotelling's footsteps, we say that a consumer located at $\beta_1 \in [0, \beta_{\max}]$ is willing to buy variety 1 when her WTP for one unit of the good from shop 1 is positive, that is, when the distance to this shop is smaller than $\beta_{\max}$. Therefore, a high (low) value of β_1 amounts to saying that the consumer is far from (close to) shop 1. As a result, we may view β_s in (2) as a parameter expressing the *idiosyncratic mismatch* between the horizontal characteristics of variety s and the consumer's ideal. This interpretation of β_s is nicely related to the concavity of u_s. As the mismatch between variety s and the consumer's ideal horizontal characteristics β_s increases, it is natural to expect the consumer to reach faster the level of satiation. In other words, if our consumer prefers vanilla to chocolate as an ice-cream flavor, the utility of an additional chocolate scoop will decrease faster than that of a vanilla scoop.

We now proceed by exploring the links between the above spatial setting and our model of monopolistic competition. When $\beta_1 < \beta_{\max}$, the consumer visits at least shop 1. However, as long as $\alpha - \gamma/2 - \beta$ is positive at $1/2$, then there is another segment $[1 - \beta_{\max}, \beta_{\max}]$ in which both $\alpha - \gamma/2 - \beta_1$ and $\alpha - \gamma/2 - (1 - \beta_1)$ are positive. Indeed, since consumers have a love for variety, a consumer located in the vicinity of $1/2$ may want to visit both shops. For this to happen, we must account that the consumer has already acquired one unit of the good so that the two WTP-lines shift downward by $\gamma/2$. Therefore, the segment over which both shops are actually visited is narrower than $[1 - \beta_{\max}, \beta_{\max}]$ and given by $[1 - \beta_{\max} + \gamma/2, \beta_{\max} - \gamma/2]$. Consequently, when the consumer is located at $\beta_1 < 1 - \beta_{\max} + \gamma/2$ she visits shop 1 only, whereas she visits both shops when her location belongs to $[1 - \beta_{\max} + \gamma/2, \beta_{\max} - \gamma/2]$.

[22] Anderson, de Palma and Thisse (1992) have pinned down the Lancasterian foundations of the CES utility. To be precise, they show that there exists a one-to-one relationship between the elasticity of substitution across varieties and the distance between these varieties in the characteristics space: the larger the distance between varieties, the smaller the elasticity of substitution.

The foregoing argument shows how our spatial model can cope with consumers buying one or two varieties of the differentiated good. In particular, regardless of her location β_1, any consumer acquires the two varieties when the interval $[1-\beta_{\max}+\gamma/2, \beta_{\max}-\gamma/2]$ is wide enough. This will be so if and only if

$$\alpha - \gamma > 1.$$

This condition holds when the desirability of the differentiated good is high, the substitutability between the two varieties is low, or both.

Conversely, it is readily verified that, regardless of her location, our consumer acquires a single variety if and only if

$$1 > 2(\alpha - \gamma) \Leftrightarrow \gamma > \alpha - \frac{1}{2}.$$

In other words, when varieties are very good substitutes, consumers choose to behave like in the Hotelling model: despite their love for variety, they patronize a single shop because the utility derived from buying from the second shop is overcome by the cost of patronizing this shop. In particular, consumers located near the ends of the segment buy only one variety and consumers located in the central area buy both if and only if

$$\alpha - \gamma < 1 < 2(\alpha - \gamma).$$

Note that, when α is sufficiently small, a consumer located in the central area does not shop at all because both her desirability of the differentiated good is low and her taste mismatch is high. In the standard Hotelling framework, this corresponds to the case in which the price of the good plus the transport cost borne by the consumer exceeds her reservation price.

Summing up, we find it fair to say that the preferences in (2) encapsulate both vertical (α_s) and horizontal (β_s) differentiation features. Indeed, we drew a parallel between the taste parameter β_s and the distance a consumer has to travel to the shop. From this parallel it is clear that a large value of β_s corresponds to a bad match because of the longer distance one has to travel. In other words, when β_s is large, the consumer's ideal variety is far from the actual variety. This interpretation of beta allows us to refer to this parameter as an inverse indicator of taste.[23]

3.3 A digression: how income matters

In the foregoing, income had no impact on the demand for the differentiated good. Yet, it is reasonable to expect consumers with different incomes to have different WTP. When the product under consideration accounts for a small share of their total consumption and the numéraire is interpreted as capturing a bundle of consumption of all the other products, we may capture this effect by slightly modifying the utility function $u_{s,i}$ of consumer $i = 1, .., n$. Specifically, consumer i's utility of variety s is now given by

$$u_{s,i} = \alpha_s q_s - \frac{\beta_{s,i}}{2} q_s^2 + q_{0,i}$$

[23] Note that the degree of acquaintance of a consumer with a particular product can be included among the horizontal characteristics captured by β_s, generating the trade patterns observed by Arkolakis (2010).

where $q_{0,i} = \delta_i q_0$ and $\beta_{s,i}$ is consumer's taste mismatch, which may be interpreted as in the foregoing. In this reformulation, $\delta_i > 0$ measures the consumer's marginal utility of income. Because this typically decreases with the consumer's income, we may rank consumers by decreasing order of income, and thus $\delta_1 < \delta_2 < ... < \delta_n$ where $\delta_1 = 1$ and $q_{0,1} = q_0$ by normalization.

Consumer i's WTP for variety s becomes

$$p_{s,i} = \max\left\{\frac{\alpha_s - \beta_{s,i}q_s}{\delta_i}, 0\right\}$$

where $p_{s,i}$ is expressed in terms of the numéraire of the richest consumer: the lower δ, the higher the WTP for the differentiated good. Thus, we indirectly capture the impact of income on demand. Therefore, though we find it convenient to refer to α_s as the quality of variety s, we acknowledge that this parameter interacts with some other variables, such as income. It is readily verified that such variables generate market effects akin to what we call quality. Hence what we show is that a quasi-linear model like ours can deal with income differences across countries but cannot deal with income inequality within countries. Still, to work with a representative consumer model seems a reasonable assumption in view of the current data availability. Most available trade data, like ours, only have one observation per firm-product-country which does not allow analysis of consumer differences within countries.

3.4 The multi-variety case

We now move to the case of a consumer having access to a large number (formally, a continuum) of varieties, $S \equiv [0, N]$, where N is the mass of varieties.

$$u_s = \alpha_s q_s - \frac{\beta_s}{2}q_s^2 - \frac{\gamma}{2}q_s\left[\int_S q_r dr\right] + q_0$$

$$= \alpha_s q_s - \frac{\beta_s}{2}q_s^2 - \frac{\gamma}{2}q_s Q + q_0 \tag{4}$$

where $\gamma > 0$ and Q is the consumer's total consumption of the differentiated good. In this expression, γ measures the direct substitutability between variety s and any other variety $r \in S$. This parameter is assumed to be the same between any two varieties of the same product because $\beta_s >$ already captures asymmetries in preferences. Allowing γ to vary across varieties would make the algebra more cumbersome. While analytically feasible, it would be difficult to measure empirically the substitution patterns between any pair of varieties. We thus follow an assumption common to virtually all models of trade and consider substitutability as a product characteristic common to all the varieties of a particular good.

Consequently, the two-variety WTP now generalizes into

$$p_s = \alpha_s - \frac{\gamma}{2}Q - \beta_s q_s. \tag{5}$$

Compared to (1), the WTP for variety s is shifted downward to account for the fact that all varieties are substitutes; the value of the shifter increases with the total consumption of the differentiated good and the substitutability across varieties.

Integrating (4) over the set S of varieties consumed, yields the utility function

$$U = \int_S \alpha_s q_s ds - \frac{1}{2}\int_S \beta_s q_s^2 ds - \frac{\gamma}{2}\left[\int_S q_s ds\right]^2 + q_0$$

where α_s and β_s are two positive and continuous functions defined on S, the former measuring the intrinsic quality of variety s and the latter capturing the distance between the consumer's ideal and variety s. The above expression is to be contrasted to the standard quadratic utility in which α and β are identical across varieties, which means that all varieties have the same quality and taste mismatch.

The budget constraint is

$$\int_S q_s p_s ds + q_0 = y.$$

Using (5), we readily see that the demand for variety s is given by

$$q_s = \frac{\alpha_s - p_s}{\beta_s} - \frac{\gamma(\mathbb{A} - \mathbb{P})}{\beta_s(1 + \gamma\mathbb{N})} \tag{6}$$

where

$$\mathbb{N} \equiv \int_S \frac{dr}{\beta_r} \qquad \mathbb{A} \equiv \int_S \frac{\alpha_r}{\beta_r} dr \qquad \mathbb{P} \equiv \int_S \frac{p_r}{\beta_r} dr.$$

Like in most models of monopolistic competition, the demand for a variety depends on a few market aggregates, here three (Vives, 2001), which are market-specific. Using the interpretation of β_r given above, it is straightforward to see $1/\beta_r$ as a measure of the proximity of variety r to the representative consumer's ideal set of characteristics. Consequently, a group of neighborhing varieties characterized by small (large) values of β_r have a strong (weak) impact on the demand for variety s because the representative consumer is (not) willing to buy much of them, as she (dis)likes its horizontal characteristics. This explains why β_r appears in the denominator of the three aggregates, $\mathbb{N}$, $\mathbb{A}$ and $\mathbb{P}$.

Having this in mind, it should be clear why each variety is *weighted by the inverse of its taste mismatch* to determine the *effective* mass of varieties, given by $\mathbb{N}$. It is $\mathbb{N}$ and not the unweighted mass of varieties, N, that affects the consumers' demand for a given variety. Indeed, adding or deleting varieties with bad taste matches, for example, does not affect much the demand for the others, whereas the opposite holds when the match is good. Note that $\mathbb{N}$ may be larger or smaller than N according to the distribution of taste mismatches. Similarly, the quality and price of a variety are weighted by the inverse of its taste mismatch to determine the effective quality index $\mathbb{A}$ and the effective price index $\mathbb{P}$. In particular, varieties displaying the same quality (or price) may have very different impacts on the demand for other varieties according to their taste mismatches. These three aggregates show that taste heterogeneity affects demand and, therefore, the market outcome. In addition, two different markets are typically associated with two different β-distributions. Consequently, the nature and intensity of competition may vary significantly from one market to another, even when the same range of varieties is supplied to both.

The above discussion shows that it is possible to introduce heterogeneity across varieties on the consumer side in order to generate a large array of new features in consumer demand. In what follows, we call *verti-zontal differentiation* this new interaction of vertical and horizontal characteristics.

It should be clear from (6) that the demand for a variety depends on its own horizontal and vertical attributes, but also on the aggregate $\mathbb{A}$, $\mathbb{P}$ and $\mathbb{N}$, which together determine how a particular firm *meets the competition.*

Finally, also note that Q, as in (6), is given by

$$Q = \frac{\mathbb{A} - \mathbb{P}}{1 + \gamma\mathbb{N}}$$

which shows once more how the utility of a variety depends on the distribution of the taste parameter β_s.

3.5 Monopolistic competition under verti-zontal differentiation

When each variety s is associated with a marginal production cost $c_s > 0$, operating profits earned from variety s are as follows:

$$\Pi_s = (p_s - c_s)q_s$$

where q_s is given by (6). Differentiating this expression with respect to p_s yields

$$p_s^*(\mathbb{P}) = \frac{\alpha_s + c_s}{2} - \frac{\gamma(\mathbb{A} - \mathbb{P})}{2(1 + \gamma\mathbb{N})}. \tag{7}$$

The natural interpretation of this expression is that it represents firm s' best-reply to the market conditions. These conditions are defined by the aggregate behavior of all producers, which is summarized here by the price index $\mathbb{P}$. The best-reply function is upward sloping because varieties are substitutable: a rise in the effective price index $\mathbb{P}$ relaxes price competition and enables each firm to sell its variety at a higher price. Even though the price index is endogenous, $\mathbb{P}$ is accurately treated parametrically because each variety is negligible to the market. In contrast, $\mathbb{A}$ and $\mathbb{N}$ are exogenously determined by the distributions of quality and tastes over S. In particular, by shifting the best reply downward, a larger effective mass $\mathbb{N}$ of firms makes competition tougher and reduces prices. Similarly, when the quality index $\mathbb{A}$ rises, each firm faces varieties having in the aggregate a higher quality, thus making the market penetration of its variety harder. Thus, through market aggregates determined by the asymmetric distribution of varieties, our model of monopolistic competition manages to reconcile weak interactions, typical of Chamberlin-like models, with several of the main features of Hotelling-like models of product differentiation.

Integrating (7) over S shows that the equilibrium price index can be expressed in terms of three market indices:

$$\mathbb{P}^* = \mathbb{C} + \frac{\mathbb{A} - \mathbb{C}}{2 + \gamma\mathbb{N}} \tag{8}$$

where the cost index is defined as

$$\mathbb{C} \equiv \int_S \frac{c_r}{\beta_r} dr.$$

In this expression, varieties' costs are weighted as in the above indices for the same reasons as in the foregoing. Hence, efficiently produced varieties may have a low impact on the cost index when they have a bad match with the consumer's ideal. Note also that $\mathbb{A}$ affects prices positively, even though it affects each individual variety's price negatively.

Plugging $\mathbb{P}^*$ into (7), we obtain the (absolute) markup of variety s:

$$p_s^* - c_s = \frac{\alpha_s - c_s}{2} - \mathcal{T}\left(\frac{\tilde{a} - \tilde{c}}{2}\right) \tag{9}$$

where taste-weighted average quality and cost indices are obtained by dividing cost and quality indices by the effective number of varieties in the market:

$$\tilde{a} \equiv \mathbb{A}/\mathbb{N} \qquad \tilde{c} \equiv \mathbb{C}/\mathbb{N}$$

and where

$$\mathcal{T} \equiv \frac{\gamma\mathbb{N}}{2 + \gamma\mathbb{N}} \in [0; 1].$$

Note that the first term of (9) is variety-specific, but the second term is not. Since it affects identically all the varieties in a market, we refer to it as a market effect (ME). In words, a variety markup is equal to half of its social value minus half of the average social value of all varieties, the second term being weighted by $\mathcal{T}$ that accounts for the *toughness of competition* through the effective mass of firms and the degree of substitutability across varieties. In particular, only the varieties with the highest social value will survive, very much as in oligopolistic models of product differentiation (Shaked and Sutton, 1983). When $\gamma\mathbb{N}$ is arbitrarily small, each variety is supplied at its monopoly price since $\mathcal{T} \to 0$. On the other hand, when $\mathcal{T} \to 1$, the market outcome converges toward perfect competition. The benefits of assuming that γ is the same across varieties are reaped by capturing the degree of competition on a particular market through $\mathcal{T}$. In addition, the toughness of competition may vary from one market to another because $\mathcal{T}$ depends on the effective mass of varieties.[24]

Last, suppose that the average effective quality $\mathbb{A}/\mathbb{N}$ increases by $\Delta > 0$. Then, if the quality upgrade Δ_s of variety s is such that

$$\Delta_s > \mathcal{T}\Delta$$

then its markup and price will increase, even though the quality upgrade Δ_s may be lower than Δ. In contrast, if the quality upgrade of variety s is smaller than $\mathcal{T}\Delta$, then its markup and price will decrease, even though the quality upgrade Δ_s is positive. In other words, quality differences are exacerbated by the toughness of competition in the determination of markups.

Note that the equilibrium price of variety s is independent of β_s because the price elasticity

[24]This parameter can be nicely related to the existence of different price ranges across sectors observed by Khandelwal (2010). Noting that each variety is characterized by an idiosyncratic quality and cost parameter, we can show that, paraphrasing Khandelwal, it is *the length of the markup ladder* that varies across sectors in our model: the tougher the competition, the shorter the ladder.

itself is independent of β_s and given by

$$\epsilon_s = -\frac{p_s}{\alpha_s - \gamma Q - p_s}.$$

This expression ranges from 0, when $p_s = 0$, to $-\infty$, when prices equal the intercept of the inverse demand function, $\alpha_s - \gamma Q$. Note that β_s does not affect ϵ_s and, therefore, has no impact on p_s. However, the whole distribution β_r matters because it influences the equilibrium value of Q, which is the aggregate sales of all other firms selling a similar product to consumers.

Using the properties of linear demand functions, we readily verify that the equilibrium output of each variety is given by

$$q_s^* = \frac{1}{\beta_s}(p_s^* - c_s) \tag{10}$$

while the corresponding equilibrium operating profits are

$$\pi_s = \frac{1}{\beta_s}(p_s^* - c_s)^2.$$

It is also interesting to notice how cost, quality and taste interact in determining the market share of a certain variety, in terms of quantities sold:

$$\frac{q_s}{Q} = \frac{\gamma}{2\beta_s}\left(\frac{\alpha_s - c_s}{\tilde{a} - \tilde{c}}\frac{1}{\mathcal{T}} - 1\right).$$

The relative sales of a particular variety in a market is here shown to depend not only on the quality and cost of a particular variety with respect to the rest of the market, but also on variety-specific taste mismatch, β_s, market specific toughness of competition, $\mathcal{T}$, and product substitutability, γ. Specifically, varieties with higher quality, α_s, or lower costs, c_s, or a better match with local taste, β_s, will have higher market shares. The more so, the lower is the toughness of competition in the market or the higher is product substitutability.

3.6 Trade under verti-zontal differentiation

While the model has been solved for one consumer, from this point forward we interpret the model in a trade context where the world consists of different countries, where each country i is considered as a single consumer whose preferences have been described above. The theory then tells us what to expect as price and quantity determinants in each destination market. Variety-specific determinants of prices and per capita quantities (captured by subscript s), such as cost and quality, do not vary by destination market and influence prices and quantities in a similar way in all countries. The idiosyncratic taste parameter, β, varies by variety and country, so it is indexed by i and s. Since we follow the literature in assuming that markets are segmented, market aggregates such as the price index $\mathbb{P}$, the mass of competing varieties $\mathbb{N}$ and the quality index $\mathbb{A}$ are also considered as country-specific variables having an effect on local prices and per capita quantities. The relevant product-market in which varieties are competing, S, is composed by all the varieties s of a certain good in a specific market i.

Equilibrium prices and quantities can then be written as follows:

$$p^*_{s,i} = \frac{\alpha_s + c_s}{2} - \mathcal{T}_i\left(\frac{\tilde{a}_i - \tilde{c}_i}{2}\right) \tag{11}$$

$$q^*_{s,i} = \frac{1}{\beta_{s,i}}\left[\frac{\alpha_s - c_s}{2} - \mathcal{T}_i\left(\frac{\tilde{a}_i - \tilde{c}_i}{2}\right)\right] \tag{12}$$

Note that the second terms on the RHS of (11) and (12) shows that absolute prices and quantities of varieties can differ across geographical markets due to a common market effect (composed of all the terms indexed by i) which can be thought of as local competitive conditions and substitution patterns. This market effect implies that the level of prices can differ across markets through the β-weighted market aggregates.

The model would predict that if a variety is sold at a relatively high price in a market, it will remain relatively expensive in another market because its cost and quality parameters have a same effect on prices anywhere. Furthermore, the same variety may be sold in different markets at different prices and in different quantities, even when the differences in costs are negligible. Prices and markups depend on the vertical attributes of each variety and on the market-specific degree of competitiveness, which can be fully captured by taste-weighted price, quality and cost indices as well as by the effective mass of competitors. Quantities additionally also depend on market variety-specific mismatch.

In what follows, we assume transport costs to be product-specific and identical for all products going from the same origin country (Belgium in our case) to the same destination market, thus they will not affect price ranks of varieties across markets. Transport costs will consequently cancel out and will not need to be modelled explicitly.[25]

4 Parameter Identification

An interesting feature of our model is that it can be used to address many specific empirical questions because its parameters are easily identifiable. However, for this to be possible, not just information on quantities and prices is required, but also on costs (or makups), which is not often readily available. Its particular functional form allows researchers to directly measure quality differences between varieties (firm-products) and taste mismatch parameter at the market-variety level. In addition, it can be used to estimate the absolute level of quality and the degree of substitutability between varieties.

The parameter capturing market-variety-specific horizontal differentiation is the most easily identifiable. In particular, $\beta_{s,i}$ can be interpreted as the distance in the horizontal characteristics space between the consumer i's ideal combination of attributes and the actual attributes of a particular variety s, as noted in the previous section. To determine its value, we can simply rewrite equation (10) and show it is the ratio between each variety's markups and its quantities sold, at any given point in time, t:

$$\beta_{s,i,t} = \frac{p^*_{s,i,t} - c_{s,t}}{q^*_{s,i,t}} \tag{13}$$

[25]Note that our approach would be consistent with the assumption of both linear or iceberg transport costs, as long as they are product-specific and do not vary by variety.

As for the parameter capturing variety-specific vertical differentiation, it can be either measured in relative terms or estimated in absolute terms, together with the substitutability parameter γ_s. Notice that α_s is the value in terms of numéraire attributed by all consumers to the vertical characteristics of a particular variety s. In relative terms, quality differentials between any couple of varieties (say, α_s and α_r) can be readily obtained by exploiting the property that relative prices of different varieties in a market depend only on differences in costs and quality (as shown in equation (11)):

$$\Delta\alpha_{sr,t} = \alpha_{s,t} - \alpha_{r,t} = 2[(p^*_{s,i,t} - p^*_{r,i,t}) - (c_{s,t} - c_{r,t})] \tag{14}$$

Measuring the distance between the quality of all the varieties, α_s, t, and the quality of the worst variety, $\alpha_{0,t}$, we can identify the relative quality distribution of the varieties present in a market at time t and eventually normalize it to have an idea of the relative distribution of varieties' quality in a market. This means that no additional information or indirect estimation methodology is needed to capture the relative quality of each variety in a market.

However, if we want to use this framework to improve macroeconomic indicators, we need to go beyond relative quality and identify absolute level of quality, defined independently of other varieties' quality.[26] As can be observed from equation (8), aggregate quality and cost levels positively affects the price index. But we can see from equation (7) that a quality-driven increase in prices will have very different implications as compared to a cost-driven price increase in macroeconomic terms. Specifically, a market where prices increase because a range of firms upgrade the quality of their products becomes more competitive, thus reducing the markup of the firms that keep the quality of their products unchanged. The opposite would happen in the case of a cost-driven price rise. Thus, price increases in a market may not always be associated with a loss of competitiveness. To correctly assess its macroeconomic implications and policy action, it is important to disentangle quality and cost, once they are accurately weighted by differences in tastes as they are in the indices $\mathbb{A}$ and $\mathbb{C}$ as in (8).

In order to go beyond a relative definition of quality, we can follow a procedure that exploits the time dimension of the data, such as variety-level prices and costs as well as total sales of a certain product in a market (i.e., Q_i).[27] To this end, equation (7) can be rewritten as:

$$p^*_{s,t}(Q_{i,t}) = \frac{\alpha_{s,t} + c_{s,t} - \gamma Q_{i,t}}{2}. \tag{15}$$

where t denotes the time dimension. This can be further seen as:

$$2p^*_{s,t} - c_{s,t} = \alpha_s - \gamma Q_{i,t}. \tag{16}$$

Note that the left-hand side of the equation consists of observables, i.e. time-varying variety-

[26] It is worth stressing that we provide a precise interpretation to the parameter capturing the absolute level of quality in our model. Our parameter α_s represents the price any consumer would be willing to pay, in terms of its numéraire, for the first marginal unit of that particular variety in the absence of substitutes, which amounts to considering either a substitution parameter equal to 0 or that no other varieties of the same good are consumed.

[27] To this end, information on the total quantities consumed of a certain good, Q_i, is needed. For example, this information can be obtained by merging trade datasets, such as UN COMTRADE or Eurostat COMEXT, which capture total imports of a good in a market, with production datasets such as Eurostat PRODCOM, which reports total consumption in a market.

specific prices and costs. On the right-hand side, however, we do not directly observe neither α_s nor γ. But holding the substitution parameter fixed over time and maintaining that the worst-quality variety, $s = 0$, is worth always the same amount of the numéraire, we can estimate $\hat{\gamma}$ through a simple regression analysis. By regressing the left-hand side of (16) for this worst-quality variety on Q_i, we obtain a value for the substitutability parameter, γ, as the coefficient associated with Q_i and α_0 as the constant of the regression.[28] Notice that the identity of the worst variety can be determined from equation (14).

Plugging then back $\hat{\gamma}$ into equation (15), since all the other variables are now observed or identified, we can determine the absolute quality levels for all the varieties of a particular product at a particular time and have all the parameters of the model empirically identified.

As long as individual prices, quantities and costs (or markups) can be observed or estimated, all the idiosyncratic and market variables can be fully identified and be used to address empirical questions on quality upgrading or competitiveness dynamics. However, this is not the focus of the current paper. Instead, we want to investigate whether the underlying assumptions of the model are confirmed or rejected by the export data, which is what we do in the next section. This will allow us to verify whether initial results in rankings are robust to an analysis in levels.

5 Are the assumptions rejected by the data?

In the verti-zontal model we have assumed vertical differentiation to be intrinsic to a variety (α_s), rather than allowing it to vary by destination market. Also, marginal cost has been assumed to be variety specific (c_s). This resulted in a price equation (11) where in equilibrium both firm-product quality and cost affect prices in a similar and linear way. So, even without identifying quality and without disentangling quality and cost, a simple OLS regression of export prices on firm-product dummies is expected to capture this variation and to explain an important part of the price data. According to our model, the other determinants of export prices are all country-level effects indexed by subscript i in (11) and also enter the price equation in a linear way.

Based on the theory, the joint inclusion of firm-product and country dummies is expected to yield a good fit in a regression on individual firm-product prices. But the same set of variables is expected to perform less well in explaining variation of per capita quantities across markets. This becomes apparent from equation (12) in the theory. Quantities are not just a function of firm-product cost and quality and country-level competition effects, but are also determined by idiosyncratic taste ($\beta_{s,i}$) which renders the quantity equation in (12) a non-linear one.

Therefore, based on the model we would expect a substantially lower *goodness-of-fit* when regressing variety-specific dummies and destination market-specific dummies on individual firm-product prices then on per capita quantities ($y_{s,i}$) as in (17):[29]

[28] The interpretation of such a parameter as constant over time amounts to considering the numéraire as a Hicksian composite good representing a bundle of all the other goods purchased by the consumer and allowing that the evolution of the worst variety's quality of a particular good follows the evolution of quality of the economy at large, as captured by the Hicksian composite good and implicitly normalized to 1 in the numéraire (or linked to income levels, as stressed in the digression section on how income matters).

[29] Since countries have different sizes M_i, the quantities used in our analysis are the total quantities divided by the population size of each destination country, $q_{s,i}/M_i$. Using instead total quantities yields results that are

$$y_{s,i} = \delta_0 + \delta_1 Variety_s + \delta_2 Market_i + \epsilon_{s,i} \quad (17)$$

In the regressions we use price and quantity level data as dependent variables and run the specification in (17) on the 171 data samples identified. Note that the unit of observation is always an individual variety defined by a firm-CN8 in a particular destination market. Each variety will then be associated with a specific dummy in all the markets where it is sold. Similarly, all the varieties present in the same destination market will be assigned a dummy equal to one when observed in that specific destination.

The average (R^2) for regression (17) are displayed in Figure 5, where the square dots should now be read as average R^2 resulting from the price regressions and the triangle dots are the R^2 from the quantity regressions. The horizontal line segments indicate the average R^2 by level of product aggregation, while the individual dots show the averages by number of markets considered for each level of product aggregation. The solid line shows average prices while the dashed line shows average quantities in different samples.

[INSERT FIGURE 5 HERE.]

The price regressions have an R^2 that varies between 60 and 70% depending on the sample that is used, which is systematically higher than the one associated with quantity regressions that ranges from 40 to 50%. Looking at the top row, column (1), we can see that the average of the averages across all samples displays a difference of 20 percentage points in the captured variability between price and quantity regressions. This difference is systematically present, no matter which product-market definition or market combination is used. A consistently higher goodness-of-fit for price as opposed to quantity regressions is consistent with the model's prediction that taste variation is important, which is not captured by the dummy model in (1). By no means do we present nor interpret our OLS regressions as a test of the model. All we are saying is that the difference in goodness-of-fit between the price versus quantity regression of the same set of dummies is what we would expect on the basis of the model. Indeed, the low R^2 in the quantity regression is suggestive of an important omitted variable.

We run a test especially designed to test for omitted variable bias, which is the Ramsey's RESET. The top row in Table 5 shows that the price regression passes the Ramsey test in 71.9% of the samples, while the comparable number of the quantity regression is 9.4%. A natural way to interpret this is that the high R^2 for the price regression suggests that the linear functional form is reasonable and no important variables are omitted. The opposite holds true for quantity regressions, which supports the idea that a market-variety-specific taste parameter is missing in the regression and structural parameters affecting equilibrium quantities do so in a non-linear way.

The rest of the Table 5 disaggregates this by product aggregation and by best N- market combinations. The difference between price regressions and quantity regressions remains striking. For example, when 7 markets are considered, only 1 quantity regression out of 20 passes the RESET test, whereas 16 out of 20 do so for the price regressions on dummies.

qualitatively the same as those obtained here.

[INSERT TABLE 5 HERE.]

A pertinent question is how our predictions are different from Melitz and Ottaviano (2008). Since our firm-product dummies in (17) do not distinguish between quality and cost, they could also just be picking up cost heterogeneity. Thus it has to be pointed out that the dummy approach in the price regression in (17) is not really what distinguishes our model from others, but the quantity regression does.

In the Melitz and Ottaviano (2008) model where the taste parameter beta is assumed constant across countries, the quantity regression is predicted to capture the same variability as the price regression. Hence, we would expect the goodness-of-fit in the quantity regression to be equally good and we would not expect to see a difference in the R^2 of price and quantity regressions, since Melitz and Ottaviano (2008) assume both quantity and price to be determined by the same variables. The verti-zontal model in contrast predicts the quantity regression to have a non-linear omitted variable, which is why the fit in the price regression is expected to be higher than the fit in the quantity regression. The empirical results on (17) seem to favor a model with varying demand and pure quantity shifters.

Another assumption we make in the model is that quality is intrinsic to a variety. This does not seem to be refuted by the data. Since if quality (or perceived quality by all the consumers) of the same firm-product would substantially differ per destination country, we would expect a much lower R^2 in the price regression. If firms would systematically offer different qualities of the same firm-product across markets, price correlations across markets within firm-products would be low and only a negligible amount of variability would be explained by our two sets of dummies while results suggest the opposite.

6 Conclusions

Existing trade models explain an important number of stylized facts arising from firm-product-country level trade data, but not all. This paper presents an encompassing model that includes the main features from other models in the literature as well as some new ones to respond to this challenge. By enriching the demand side to account for non-symmetric varieties, we develop a tractable framework in which taste heterogeneity interacts with cost heterogeneity and vertical differentiation. We call it a *verti-zontal* model to stress its vertical and horizontal attributes based on Lancasterian definitions. This model offers a tractable and fully identifiable alternative to existing models of monopolistic competition. One of its distinctive features is the definition of market indices that better reflect the competitiveness of firms at the macro-economic level. The construction of such indices lies beyond the scope of this paper but is clearly a promising future line of research. In this respect, section 4 can be viewed as a starting point. In addition, as in most models of monopolistic competition, we treated cost, quality and taste as exogenous. A valuable extension would be to allow firms to endogenize these parameters, but this is left for future research.

To keep the model as general as possible, we do not assume any particular link between cost, quality and taste distributions. Several other papers require quality and marginal cost to be correlated, or assume a link between marginal cost of serving a market and the size of the market. We do not impose any restrictions on whether quality is associated with marginal costs or with fixed investments in research and development, or advertising. The same is true for the relationship between quality and taste. Yet one could think of cases where high quality products are mainly sold in rich countries reflecting a different taste for quality. Put differently, the predictions of the model seem to capture the stylized facts arising from the data without imposing any correlation between the sources of heterogeneity in the model.

The model we present here has the advantage of being identifiable, consistent with the empirical evidence, and at the same time endowed with a clear and univocal interpretation of its parameters. By contrast, it remains largely agnostic about the supply side of the economy. For example, neither firms' entry and exit nor the multi- or single-product nature of firms are explicitly treated. However, the improvements proposed on the modeling of the demand side of the economy can be directly used as a *module* that can be incorporated into trade models where the supply side has additional features.

Figure 1a: Price and quantity ranks of Belgian chocolate exports to France, Germany and the Netherlands.

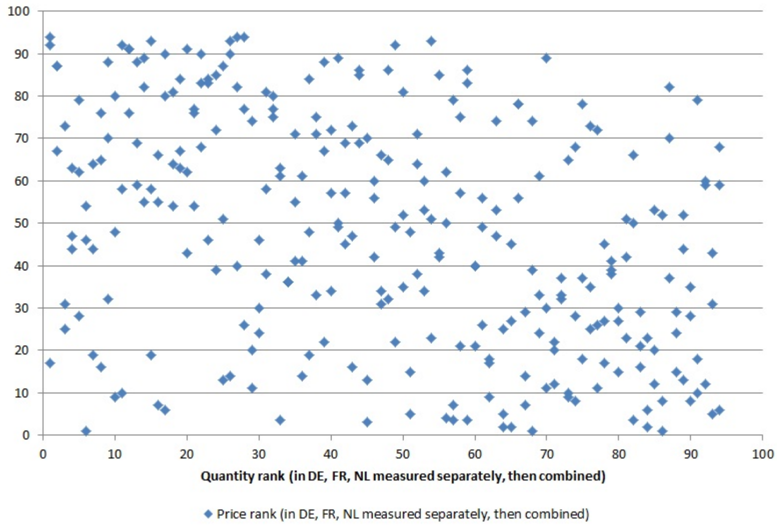


Figure 1b: Price ranks of Belgian chocolate exports to France, Germany and the Netherlands.

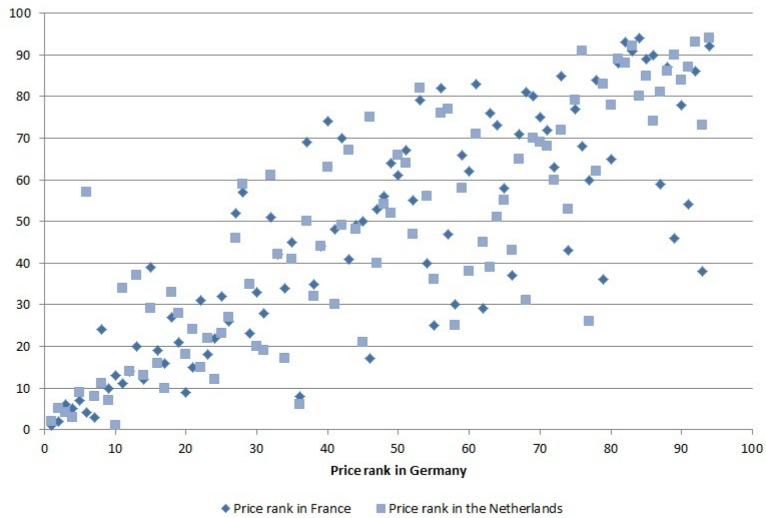


Figure 1c: Quantity ranks of Belgian chocolate exports to France, Germany and the Netherlands.

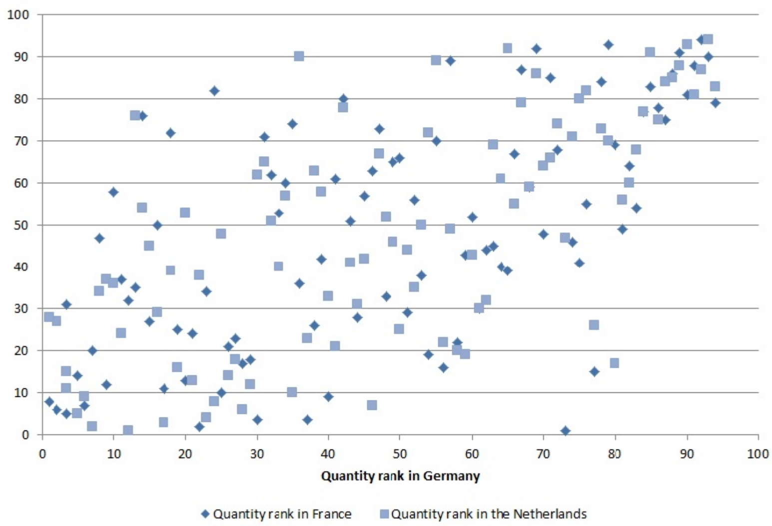
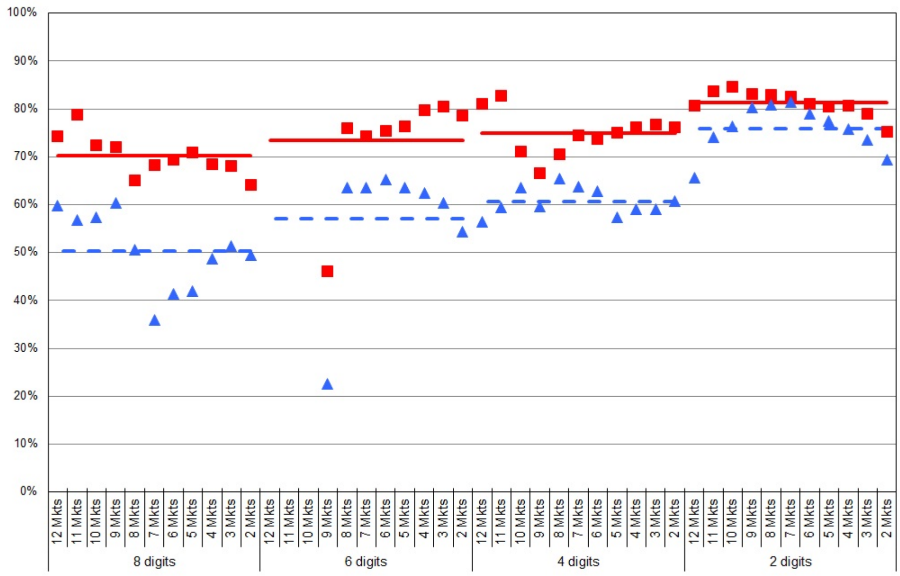
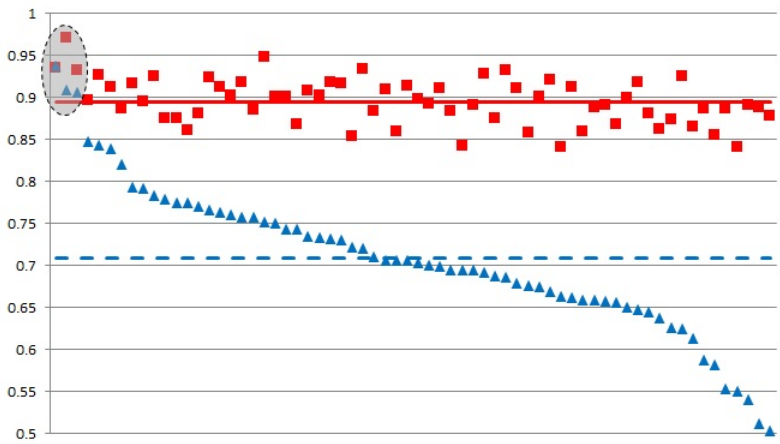


Figure 2: Average Spearman rank correlations for prices and quantities between markets



Notes: Square dots indicate average Spearman rank correlation for regressions of prices on dummies by best N-market combination across product codes, triangle dots indicate the same for regressions of quantities on dummies. The horizontal line segments refer to average Spearman rank correlation across best N-market combinations by level of product disaggregation: the solid one refers to prices, the dashed one to quantities.

Figure 3: Pairwise rank correlations for a sample of the 12 relevant export markets selected from across the globe



Notes: The countries considered are: France, Netherlands, Germany, US, Canada, Brasil, South Africa, Australia, Turkey, China, India, Japan. The square dots indicate price rank correlations for all the 66 country pair combinations, triangle dots indicate pairwise quantity rank correlations. The horizontal line segments refer to the averages: the solid one refers to prices, the dashed one to quantities. Note that for illustrative purposes country pairs have been sorted in decreasing quantity rank correlation order. The shaded area covers the three most correlated country pairs in terms of quantity ranks: France-Netherlands; Germany-France; Germany-Netherlands.

Figure 4: Graphical intuition of the spatial problem

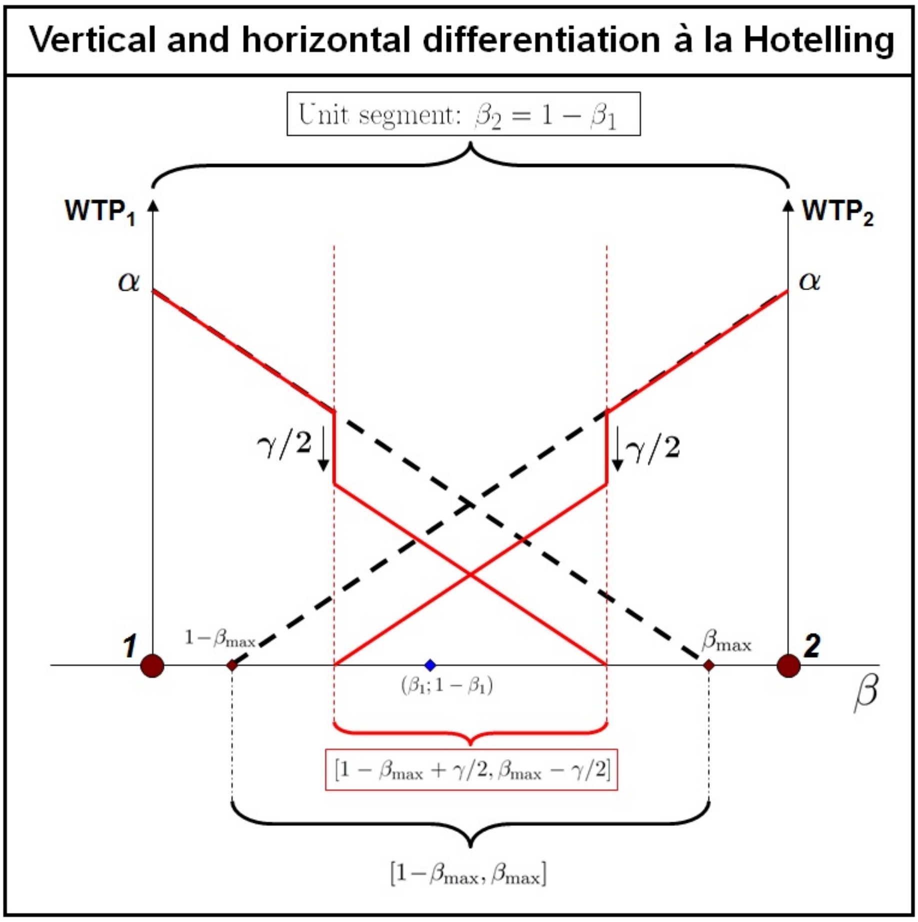
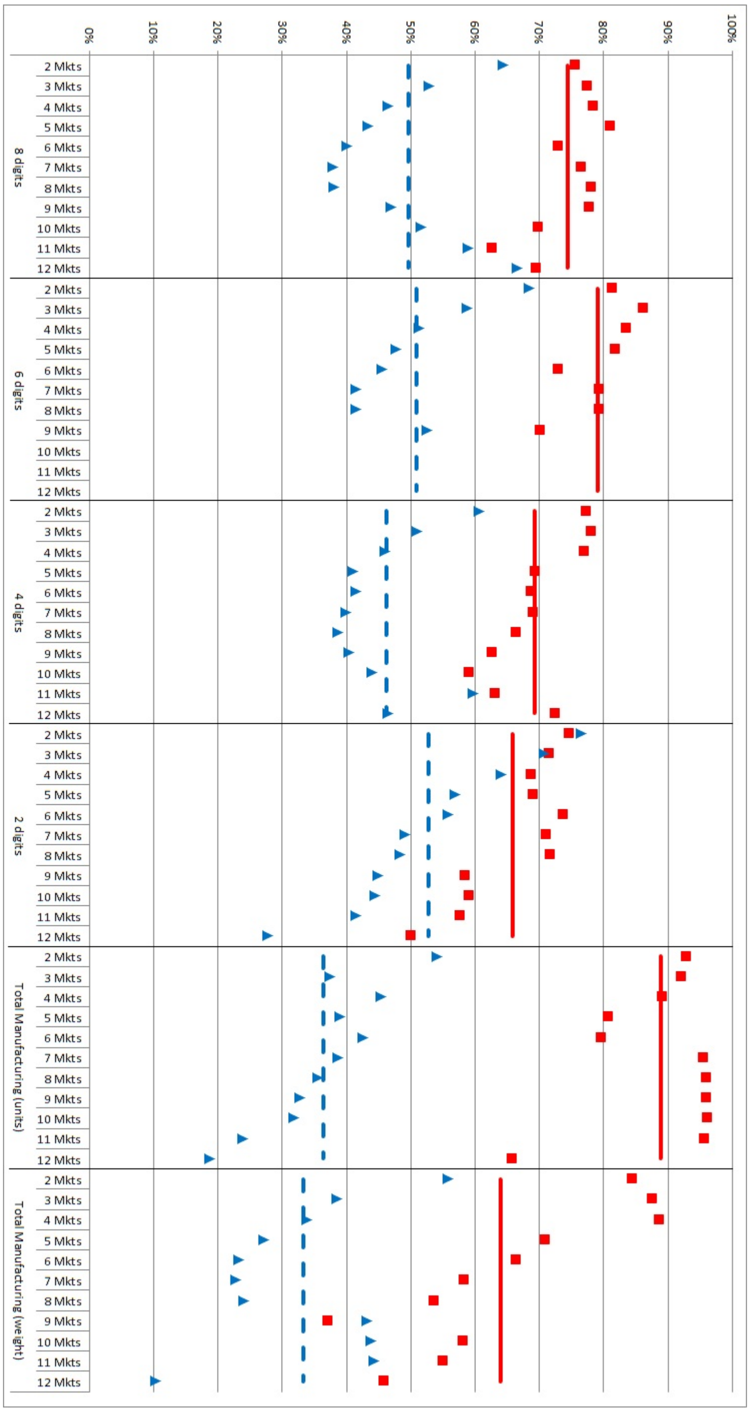


Figure 5: Average R^2 associated with price and quantity dummy regressions.



Notes: Square dots indicate average R^2 for regressions of prices on dummies by best N-market combination across product codes, triangle dots indicate the same for regressions of quantities on dummies. The horizontal line segments refer to average R^2 across best N-market combinations by level of product disaggregation: the solid one refers to prices, the dashed one to quantities.

Table 1: Product codes considered for each level of product disaggregation.

"Top 5" Combined Nomenclature product codes							
CN2	**Short description**	*CN4*	**Short description**	*CN6*	**Short description**	*CN8*	**Short description**
84	Machinery and mechanical appliances	1806	Chocolate and food preparations with cocoa	180690	Chocolate products	39269099	Other articles of plastics
39	Plastics and articles thereof	3926	Articles of plastics	170490	Sugar confectionery not containing cocoa	18069019	Chocolate products not contanining alcohol
85	Electrical machinery and equipment	0710	All frozen vegetables	220300	Beer made from malt	21069098	Food preparations
73	Articles of iron or steel	9403	Furniture and parts thereof	210690	Food preparations	57033019	Polypropylene carpets and floor coverings
90	Optical, measuring, precision, medical, or surgical instruments	4911	Printed matter, including printed pictures and photographs	071080	Frozen vegetables	22030001	Bottled beer made from malt

Notes: The BEC classification is an indicator of consumption goods at the 6 digit level. Thus, goods in sector CN8 and sector CN6 are easy to classify. However turning to more aggregate sectors like sector CN2, both consumption and other (capital, industrial) goods may occur. Our decision rule has been to include sectors CN2 and sectors CN4 when there was at least one CN6 consumption product.

Table 2: Varieties by destination marketsand destination-market combinations.

Markets	**Varieties exported to this particular destination market**	**Varieties shipped to this market and all the previous**
France	24,612	24,612
Netherlands	24,183	13,608
Germany	17,911	9,347
UK	11,956	6,367
Spain	8,799	4,419
Italy	8,869	3,572
Denmark	5,540	2,519
Sweden	5,530	2,047
Poland	6,227	1,498
Switzerland	5,732	966
U.S.	6,592	649
Luxembourg	10,317	393

Notes: In the first column is reported, for each destination market, the number of exported varieties for which units or Kilograms shipped are available. In the second column only varieties that are present simultaneously also in all the destination markets listed in the previous rows are counted.

Table 3: Varieties considered in each intersection of best N-market combination and level of product disaggregation.

Number of **varieties** considered	**Best N-country combinations**										
	N=2	N=3	N=4	N=5	N=6	N=7	N=8	N=9	N=10	N=11	N=12
top 5 CN8	275	221	174	139	117	93	66	24	15	12	11
top 5 CN6	333	263	215	174	130	100	72	10	0	0	0
top 5 CN4	818	604	464	339	250	174	134	83	41	24	22
top 5 CN2	3674	2591	1835	1352	1123	811	698	535	358	259	135
Whole Manufaturing (weight)	12981	8908	6040	4166	3361	2362	1908	1407	893	599	355
Whole Manufacturing (units)	2831	1913	1306	879	701	502	412	311	212	146	81

Notes: Each intersection is composed of 5 samples at most, but there could be less, as samples are considered valid for our analysis when they are composed of at least 10 varieties. On the last two rows, all the varieties are reported for which we observe quantities shipped in Kilograms (weight) or other units of measure (units). The sum of the last two rows is higher than the second column of Table 2 become some varieties report both weight and units and therefore are counted only once in Table 2.

Table 4: Rejection rates for rank correlations.

Spearman's rank correlations: *Rejection of significance for within markets price-quantity correlations (**Price-Quantity**) and between-markets price and quantity correlations (respectively, **Quantity-Quantity** and **Price-Price**)*

All the samples	**Price-Quantity**	**35.3%**
	Quantity-Quantity	**19.1%**
	Price-Price	**2.9%**
	Samples	*(171)*

		CN8	CN6	CN4	CN2
By level of product aggregation:	**Price-Quantity**	**76.3%**	**25.7%**	**48.9%**	**1.8%**
	Quantity-Quantity	**60.5%**	**8.6%**	**15.6%**	**0.0%**
	Price-Price	**5.3%**	**2.9%**	**4.4%**	**0.0%**
	Samples	*(38)*	*(35)*	*(45)*	*(53)*

		12 Mkts	11 Mkts	10 Mkts	9 Mkts	8 Mkts	7 Mkts	6 Mkts	5 Mkts	4 Mkts	3 Mkts	2 Mkts
By best N-market combinations:	**Price-Quantity**	**50.0%**	**37.5%**	**44.4%**	**54.5%**	**47.1%**	**55.0%**	**40.0%**	**35.0%**	**25.0%**	**15.0%**	**15.0%**
	Quantity-Quantity	**50.0%**	**37.5%**	**22.2%**	**18.2%**	**29.4%**	**25.0%**	**20.0%**	**15.0%**	**15.0%**	**10.0%**	**5.0%**
	Price-Price	**0.0%**	**0.0%**	**11.1%**	**18.2%**	**11.8%**	**0.0%**	**0.0%**	**0.0%**	**0.0%**	**0.0%**	**0.0%**
	Samples	*(6)*	*(8)*	*(9)*	*(11)*	*(17)*	*(20)*	*(20)*	*(20)*	*(20)*	*(20)*	*(20)*

Notes: Percentages of samples not significantly correlated at a 5% level are reported by product aggregation and market combination. For each sample we consider the average level of significance of all the pairwise correlations involved. The number of samples considered is reported in brackets. For example, looking at Spearman between-markets price rank correlations at a CN8 level of product aggregation, 2 of the 38 samples considered are not significantly different from 0, which yields the reported 5.3%.

Table 5: Success rates in tests for omitted variables in the dummy regressions in Figure 5.

Share of samples passing the regression specification error test (***RESET***) for omitted variables.												
All the samples	**Price**	**71.93%**										
	Quantity	**9.36%**										
	Samples	(171)										
By level of product disaggregation:		CN8	CN6	CN4	CN2							
	Price	**76.32%**	**88.57%**	**62.22%**	**66.04%**							
	Quantity	**10.53%**	**5.71%**	**6.67%**	**13.21%**							
	Samples	(38)	(35)	(45)	(53)							
By best N-market combinations:		12 Mkts	11 Mkts	10 Mkts	9 Mkts	8 Mkts	7 Mkts	6 Mkts	5 Mkts	4 Mkts	3 Mkts	2 Mkts
	Price	**50.0%**	**50.0%**	**44.4%**	**54.6%**	**70.6%**	**80.0%**	**70.0%**	**70.0%**	**85.0%**	**80.0%**	**85.0%**
	Quantity	**0.0%**	**0.0%**	**0.0%**	**0.0%**	**0.0%**	**5.0%**	**5.0%**	**5.0%**	**10.0%**	**25.0%**	**30.0%**
	Samples	(6)	(8)	(9)	(11)	(17)	(20)	(20)	(20)	(20)	(20)	(20)

Notes: Percentages of samples passing the RESET test for omitted variables are reported by product disaggregation and market combination. The number of samples considered is reported in brackets. For example, at a CN8 level product disaggregation, 63.7% of the 38 samples considered passed the test when prices were regressed on dummies, but only 21.1% passed the test when quantities regressions were considered.

References

Anderson, S., de Palma, A. and Thisse, J. (1992). *Discrete Choice Theory of Product Differentiation*. Cambridge, MA: MIT Press.

Arkolakis, C. (2010). Market penetration costs and the new consumers margin in international trade. *Journal of Political Economy* 118: 1151 – 1199.

Arkolakis, C. and Muendler, M.-A. (2010). The Extensive Margin of Exporting Products: A Firm-level Analysis. NBER Working Papers 16641, National Bureau of Economic Research, Inc.

Aw, B. Y., Batra, G. and Roberts, M. J. (2001). Firm heterogeneity and export-domestic price differentials: A study of taiwanese electronics products. *Journal of International Economics* 54: 149–169.

Aw, B. Y., Chung, S. and Roberts, M. J. (2000). Productivity and turnover in the export market: micro-level evidence from the republic of Korea and Taiwan (China). *World Bank Economic Review* 14: 65–90.

Baldwin, R. and Harrigan, J. (2011). Zeros, quality, and space: trade theory and trade evidence. *American Economic Journal: Microeconomics* 3: 60–88.

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

Bernard, A. B., Redding, S. J. and Schott, P. K. (2010). Multiple-product firms and product switching. *American Economic Review* 100: 70–97.

Bernard, A. B., Redding, S. J. and Schott, P. K. (2011). Multi-product firms and trade liberalization. *Quarterly Journal of Economics* 126: 1271–1318.

Brooks, E. L. (2006). Why don't firms export more? Product quality and colombian plants. *Journal of Development Economics* 80: 160–178.

Crozet, M., Head, K. and Mayer, T. (2012). Quality sorting and trade: firm-level evidence for french wine. *Review of Economic Studies* 79: 609–644.

Dhingra, S. (2011). Trading Away Wide Brands for Cheap Brands. CEP Discussion Papers dp1103, Centre for Economic Performance, LSE.

Eaton, J., Kortum, S. and Kramarz, F. (2011). An anatomy of international trade: evidence from french firms. *Econometrica* 79: 1453–1498.

Eckel, C., Iacovone, L., Javorcik, B. and Neary, J. P. (2011). Multi-product firms at home and away: cost- versus quality-based competence. CEPR Discussion Papers 8186.

Eckel, C. and Neary, J. P. (2010). Multi-product firms and flexible manufacturing in the global economy. *Review of Economic Studies* 77: 188–217.

Fajgelbaum, P., Grossman, G. M. and Helpman, E. (2011). Income distribution, product quality, and international trade. *Journal of Political Economy* 119: 721 – 765.

Feenstra, R. C. and Romalis, J. (2012). International prices and endogenous quality. mimeo.

Foster, L., Haltiwanger, J. and Syverson, C. (2008). Reallocation, firm turnover, and efficiency: selection on productivity or profitability? *American Economic Review* 98: 394–425.

Foster, L., Haltiwanger, J. C. and Syverson, C. (2012). The slow growth of new plants: learning about demand? NBER Working Papers 17853, National Bureau of Economic Research, Inc.

Gabszewicz, J. J. and Thisse, J. (1979). Price competition, quality and income disparities. *Journal of Economic Theory* 20: 340–359.

Gorg, H., Halpern, L. and Murakozy, B. (2010). Why do within firm-product export prices differ across markets? CEPR Discussion Papers 7708.

Hallak, J. C. and Schott, P. K. (2011). Estimating cross-country differences in product quality. *Quarterly Journal of Economics* 126: 417–474.

Hallak, J. C. and Sivadasan, J. (2009). Firms' exporting behavior under quality constraints. Working Papers 09-13, Center for Economic Studies, U.S. Census Bureau.

Hotelling, H. (1929). Stability in competition. *The Economic Journal* 39: 41–57.

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

Katayama, H., Lu, S. and Tybout, J. R. (2009). Firm-level productivity studies: illusions and a solution. *International Journal of Industrial Organization* 27: 403–413.

Kee, H. L. and Krishna, K. (2008). Firm-level heterogeneous productivity and demand shocks: Evidence from Bangladesh. *American Economic Review* 98: 457–62.

Khandelwal, A. (2010). The long and short (of) quality ladders. *Review of Economic Studies* 77: 1450–1476.

Kugler, M. and Verhoogen, E. (2008). The quality-complementarity hypothesis: theory and evidence from Colombia. NBER Working Papers 14418.

Lancaster, K. J. (1979). *Variety, Equity and Efficiency*. Oxford: Basil Blackwell.

Manova, K. and Zhang, Z. (2012). Export prices across firms and destinations. *Quarterly Journal of Economics* 127: 379–436.

Martin, J. (2012). Markups, quality, and transport costs. *European Economic Review* 56: 777 – 791.

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

Melitz, M. J. and Ottaviano, G. I. P. (2008). Market size, trade, and productivity. *Review of Economic Studies* 75: 295–316.

Shaked, A. and Sutton, J. (1983). Natural oligopolies. *Econometrica* 51: 1469–1483.

Verhoogen, E. A. (2008). Trade, quality upgrading, and wage inequality in the mexican manufacturing sector. *Quarterly Journal of Economics* 123: 489–530.

Vives, X. (2001). *Oligopoly Pricing: Old Ideas and New Tools*. Cambridge, MA: MIT Press.