CENTRO STUDI LUCA D'AGLIANO

WWW.DAGLIANO.UNIMI.IT

CENTRO STUDI LUCA D'AGLIANO
DEVELOPMENT STUDIES WORKING PAPERS

N. 322

October 2011

Verti-zontal Differentiation in Monopolistic Competition

*Francesco Di Comite**
*Jacques-François Thisse***
*Hylke Vandenbussche****

* Université Catholique de Louvain and European Commission
** Université Catholique de Louvain
*** National Bank of Belgium

Verti-zontal Differentiation in Monopolistic Competition*

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

October 1, 2011

The recent availability of trade data at a firm-product-country level calls for a new generation of models able to exploit the large variability detected across observations. By developing a model of monopolistic competition in which varieties enter preferences non-symmetrically, we show how consumer taste heterogeneity interacts with quality and cost heterogeneity to generate a new set of predictions. Applying our model to a unique micro-level dataset on Belgian exporters with product and destination market information, we find that heterogeneity in consumer tastes is the missing ingredient of existing monopolistic competition models necessary to account for observed data patterns.

JEL codes: D43 - F12 - F14 - L16

Keywords: Heterogeneous firms - Product Differentiation - Monopolistic Competition - Non-symmetric varieties.

---

*We thank P. Belleflamme, A. Bernard, D. Castellani, C. Eckel, M. Francesconi, C. Fuss, H. Famerée, K. Miyagiwa, G. Ottaviano, G. Peri, Ch. Piette, P. Picard, J. Smets, J. Vogel and X. Wauthy and seminar participants at Columbia University, Chicago Federal Reserve, Georgia Tech, Emory, Notre Dame, Dublin and Aarhus for helpful discussions and suggestions. We are grateful to Florian Mayneris and Mathieu Parenti for detailed comments on earlier drafts. 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.

†Corresponding author. Department of Economics (IRES), Université catholique de Louvain, and European Commission, Directorate-General Economic and Financial Affairs, Avenue de Beaulieu 1, 1150 Bruxelles, Belgium - Email: francesco.di-comite@ec.europa.eu - Tel: +32 (0)2 2988124. The usual disclaimer applies. The paper was completed before the author joined the European Commission. The opinions expressed in this paper do not necessarily reflect those of the institution to which the author is affiliated.

‡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, Belgium - Email: hylke.vandenbussche@uclouvain.be - Tel: + 32 (0)10 474137

1 Introduction

It is now widely recognized that firms are heterogeneous along different dimensions, even within the same sector and geographical market. Yet, an encompassing theoretical framework incorporating multi-dimensional heterogeneity is still lacking.

Early attempts to model heterogeneity focused on firms' differences in productivity within a demand system exhibiting constant elasticity of substitution (CES) which fitted the empirical evidence in an elegant and parsimonious way (Melitz, 2003). However, careful scrutiny of the properties of such models shows that, despite their relevance at the aggregate level, they fail to account for several empirical regularities at more disaggregate levels of analysis, which hinders our ability to deal with micro-level trade data. For example, heterogeneity in costs alone cannot account for exporters charging higher domestic prices than non-exporters. In addition, the observed price discrimination across markets cannot be explained by standard CES specifications.[1]

Because the availability of firm-product-level trade data is rapidly increasing, several papers aim to tackle these drawbacks by exploiting different demand specifications in order to have non-constant markups and richer interactions between firms and their competitive environments (Feenstra, 2003; Melitz and Ottaviano, 2008), or by introducing additional dimensions of heterogeneity, the most common being quality.[2] This paper not only merges these two approaches but nests multi-dimensional heterogeneity in the quadratic utility framework proposed by Ottaviano et al. (2002). Specifically, we will introduce vertical as well as variety-specific horizontal differentiation in a monopolistic competition model, which we refer to as "verti-zontal".

Only a few papers in the trade literature have tried to account for both vertical and horizontal differentiation within the Lancasterian definitional setting (Lancaster, 1979), which arguably provides the best analytical setting to study product differentiation because it allows for a precise definition of products' characteristics. To the best of our knowledge, all trade papers in the spirit of Lancaster are empirical contributions based on discrete choice models where variety-specific differentiation is mainly interpreted as a random demand shifter [3] A common feature in this strand of literature is that quality and marginal costs alone do not suffice to make full sense of how a product performs in a market. Heterogeneity in consumer tastes seems to play an important role too, as suggested by the common presence of a "home bias effect" in quantities in various contexts, ranging from the European car market (Goldberg and Verboven, 2001) to the wine sector (Friberg et al., 2010) through cultural industries (Hanson and Xiang, 2011).

Our paper aims to respond to these empirical challenges by combining different strands of literature. We build a model of monopolistic competition in which idiosyncratic elements of vertical and horizontal product differentiation determine prices, sales and market conditions reflecting the intensity of competition. In the spirit of Lancaster,

[1] See, for example, Fontagné et al. (2008), Gorg et al. (2010) and Schott (2004).

[2] Notable examples of quality-augmented CES models are Baldwin and Harrigan (2011); Fajgelbaum et al. (2009); Hallak and Sivadasan (2009) and Crozet et al. (2011).

[3] Recent examples include Katayama et al. (2009); Khandelwal (2010) and Verhoogen (2008).

we assume that vertical attributes are intrinsic to varieties and affect prices *similarly in all markets*. By contrast, horizontal attributes are allowed to be valued *differently across markets*. This new model corresponds better to empirical regularities described in the literature and to the patterns of prices and sales that we find using a very disaggregate firm-product-country dataset on Belgian exporters also used in Bernard et al. (2010). Models with either cost or quality as the only source of firm-product heterogeneity appear inadequate to predict sales patterns across destinations. Even models that combine cost and quality heterogeneity fail to generate predictions for prices and sales observed in export markets. Accounting for taste heterogeneity, in the way we do in our model, generates a set of predictions that correspond much better with what we observe in disaggregated firm-level data.

Defining a variety as a firm-product combination we find that heterogeneity in consumer tastes is the missing ingredient to account for the observed data patterns. To be precise, our empirical analysis relies on firm-product-destination data and is performed at five different levels of product aggregation. Because our model leads to clear-cut theoretical predictions about the equilibrium prices and quantities, our empirical strategy compares price and quantity of all the varieties exported from Belgium within and across a given set of destination countries. The analysis of rank correlations within and between markets, together with standard correlations and OLS regressions on dummies provide supportive evidence for the model developed in the paper.

The starting point of our model is a quasi-linear utility nesting a quadratic sub-utility in which we allow product characteristics to differ across varieties within the same sector. More precisely, the demand function for any particular variety is characterized by three elements: (i) a shifter affecting its intercept, which reflects the vertical dimension of differentiation; (ii) a parameter determining its slope, which captures consumers' taste for this particular variety; (iii) a substitutability parameter capturing the common degree of competitive pressure exerted by the other varieties. The choice of a quasi-linear utility model is driven by the ease with which we can introduce a rich parameterization on the demand side and keep the model tractable. Even though our quasi-linear model does not directly capture income effects, it is rich enough to allow for product prices to range from pure monopoly to marginal costs of production and to accommodate for the various competitive effects associated with different market conditions, which empirical evidence has consistently shown to matter (Gorg et al., 2010).

It is worth noting that our model encompasses important insights provided by models of industrial organization dealing with product differentiation. In this literature there has been a long tradition of clearly distinguishing vertical from horizontal differentiation because they generate very different results. The monopolistic competition model we present in this paper offers the same clear distinction between vertical and horizontal differentiation. In contrast to standard models of monopolistic competition in which parameters have no link with the Lancasterian framework of product differentiation, the parameters of the model we develop here can be given a precise definition. However, unlike industrial organization models which emphasize strategic interactions between firms, our approach focuses on "weak interactions" between firms, meaning that firms' behavior is influenced only by market aggregates which are themselves unaffected by the choices made by any single firm. In addition, an appealing property of our model is its ability to replicate several results obtained in differentiated oligopoly theory. This is achieved by using market aggregates of variables or parameters, weighted by variety-specific consumer tastes. For example, market prices are strongly (weakly) affected by the mass of varieties which have a good (bad) match with consumers' ideal varieties, very much as in Lancasterian models of production differentiation (Anderson et al., 1992). In this respect, we find it fair to say that our model provides a reconciliation of the two main approaches to competition on differentiated markets, which were then developed by Hotelling (1929) and Chamberlin (1933). Monopolistic competition models are known to be well equipped to deal with a large number of firms, which makes them empirically relevant to guide research in firm-level data sets. Our main finding is that *heterogeneity in consumer tastes is a necessary ingredient for models of monopolistic competition to account for observed micro-level data patterns.*

The remainder of the paper is organized as follows. The next section presents some first evidence to motivate the model's assumptions, section 3 presents the model and its properties, and section 4 investigates the empirical relevance of the model using a unique dataset on Belgian exporters with product and destination market information. Section 5 concludes.

2 Motivation

Before introducing our model, we first look at how micro-level evidence on prices and quantities typically looks like. For this purpose we turn to a free and publicly available dataset on the European car market used by Goldberg and Verboven (2001). The patterns arising from the car data are very similar to the ones we observe in the Belgian export dataset which will be presented in section 4. The reason why we prefer to use the car evidence to motivate our choice of assumptions is that these data can be easily verified by any reader, which is useful given that access to firm-level data is not always granted.[4]

In order to motivate our modeling strategy, we look at prices and quantities in the five countries reported in the dataset (France, Italy, Germany, UK and Belgium) in 1999, the year in which the highest number of identical car models, 72, were sold in all of them. We assign a price and a quantity rank to each car model in each market and, in Figure 1a, plot one against the other in all markets. Each dot in the figure represents a combination of a price and quantity rank in a particular geographical market for a particular car model.

INSERT FIGURE 1a HERE.

If one assumes, as most trade models implicitly do, that all car models face the same demand in every market, and that the only difference between car models is the cost at which they are produced, one would expect all observations to lie around the diagonal

[4]The dataset can be found on http://www.econ.kuleuven.be/public/ndbad83/frank/cars.htm.

from top-left to bottom-right. Put differently, one would expect high-cost cars 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 cars, 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 car models to cluster around the diagonal running from bottom-left to top-right. Put differently, one would expect high-quality cars to be highly priced and to sell a lot, while low-quality cars would be associated with low prices and would sell poorly in all markets.

Interestingly, Figure 1a shows that there is no clear correlation pattern between price and quantity rankings.[5]. This suggests that a particular car model, displaying the same price ranking across 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 between models is productive efficiency or quality. Consequently, an important first observation arising from the car data is that more than one source of demand heterogeneity appears to be needed to fit micro-level data.[6]

A second important observation arises from plotting price rankings between countries, which we do in Figure 1b. Each dot in the figure now represents the ranking of a car model in a particular geographical market compared to the ranking of that car model in Belgium (horizontal axis), in such a way that a perfect correlation between price ranks across markets would result in dots following the 45° line. Looking at Figure 1b, we see that bilateral price rank correlations are in fact surprisingly high, ranging from 95.7% to 98.3%. A strong and positive price correlation between markets corresponds to the prediction arising both from a pure cost and a pure intrinsic quality model, but appears inconsistent with a model of differently perceived quality. A model that assumes quality to be perceived differently in every market would in fact result in a low price correlation between markets. When we introduce quality in the model, we will therefore assume it to be variety-specific but not market-variety-specific. This choice is also shown to be consistent with a rigorous interpretation of what is horizontal and what is vertical in product differentiation.

INSERT FIGURE 1b HERE.

A third observation arising from the car data stems from Figure 1c. There we plot quantity rankings of car models between countries in a similar way as we plotted price rankings. The pattern arising from quantity rankings is very different. Bilateral rank correlations of car models averages 66% and can be as low as 49.5%, which is much less than the corresponding price rank correlations. Hence, while price rankings of car models are quite stable across markets, quantity rankings are not. In section 4 we discuss evi-

[5]On average, the correlation between price and quantity rankings of car models within markets is around -11%, with rank correlations ranging from 10% in Germany to -30% in Italy, through -0.2% in Belgium

[6]Similar conclusions are reached through more formal analyses by authors such as Crozet et al. (2011); Hummels and Klenow (2005); Manova and Zhang (2011) .

dence based on a detailed micro-level dataset on Belgian exporters and show that these empirical regularities turn out to be extremely robust and hold in virtually all markets and products considered.[7].

INSERT FIGURE 1c HERE.

Based on existing trade models incorporating either cost or quality heterogeneity or both, we would expect quantity rankings to be just as regular as price rankings. What this observation is telling us, though, is that there appears to be a source of heterogeneity affecting quantities that is not just variety-specific but also market-specific. The introduction of an additional source of heterogeneity affecting quantities but not prices seems necessary to account for prices and quantities behaving so differently. Or put differently, *heterogeneity on the supply side needs to be supplemented by heterogeneity on the demand side*, and notably by idiosyncratic consumer taste.

We respond to these empirical challenges by extending a quasi-linear model of monopolistic competition with a quadratic sub-utility for the differentiated good in a way such that each variety may be viewed as a different bundle of horizontal and vertical attributes. In the spirit of Lancaster, we assume that vertical attributes are intrinsic to varieties, affecting prices *similarly in all markets*. By contrast, horizontal attributes are allowed to be valued *differently across markets*. The vertical attributes are captured by a demand-shifting parameter; the horizontal attributes will be interpreted as measuring taste mismatch between varieties' characteristics and consumers' ideals as it has been developed in industrial organization (Anderson et al., 1992). In line with the overwhelming majority of trade models and empirical evidence, we also allow for cost heterogeneity.

By choosing the quadratic utility model, we further acknowledge that competition effects are important and that they can differ in geographical markets. Empirical evidence has shown indeed that absolute price levels can be very different between countries. Exploiting product-level Hungarian custom data, Gorg et al. (2010) show that even the same firm-product may be sold at very different prices in different markets. This suggests the existence of important local market effects, which operate like a market-specific demand shifter (but which does not affect price rankings). In other words, markets appear to be segmented, with the intensity of local competition playing a role as important as individual product characteristics in affecting prices and quantities.

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

In this section, we present a model that builds directly on the above-mentioned stylized facts, embedding them in a rigorous model of product differentiation inspired by the industrial organization literature.

[7]This finding is consistent with the observation of a puzzlingly weak relationship between firms' productivity and size found by Brooks (2006) and Hallak and Sivadasan (2009) and with the evidence of a bias towards the consumption of domestic varieties (Ferreira and Waldfogel, 2010).

There are several definitions of vertical and horizontal differentiation, which are (more or less) equivalent. Ever since Hotelling (1929) and Lancaster (1979), 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 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 (Gabszewicz and Thisse, 1979; Shaked and Sutton, 1983). 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.[8] 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.

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:

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

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

[8]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 Kugler and Verhoogen (2008) and Eckel et al. (2011).

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 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 is shifted 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 used by (Gabszewicz and Thisse, 1979; Shaked and Sutton, 1983), 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, it follows that α_s can be interpreted as a measure 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. Note, finally, that α may reflect effects other than quality. We will return to this issue in section 3.3.

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 quadratic utility models is that a priori their main parameters cannot be interpreted within a characteristics space.[9] 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 consumers are assumed to make mutually exclusive purchases, in the verti-zontal model we develop they 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 2, 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 2 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

[9] Anderson et al. (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.

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 $[1-\beta_{\max}+\gamma/2, \beta_{\max}-\gamma/2]$.

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 (2) encapsulate both vertical (α_s) and horizontal (β_s) differentiation features. This specification is also flexible enough to retain the tractability of the standard quadratic utility model.

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

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

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 one typically decreases with the consumer's income, we may rank consumers by increasing 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 $alpha_s$ 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.

3.4 The multi-variety case

For notational simplicity, we return to the case of one market whose demand side is represented by a consumer and consider the standard setting of monopolistic competition in which the differentiated good is available as a continuum $S \equiv [0, N]$ of varieties, 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 substitutability between variety s and any other variety $r \in S$. 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

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

Note that the density over S is equal to 1 because each variety is supplied by a single firm.

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 variety having a small (large) β_r has a strong (weak) impact on the demand for variety s because the representative consumer is (not) willing to buy much of it.[10] In contrast, a variety with a small β_r has a strong impact on the consumption of variety s because the representative consumer highly values its horizontal characteristics. This explains why β_r appears in the denominator of the three aggregates.

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 consumer's 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

[10] Formally, we should consider an open interval of varieties containing r because the impact of a single variety upon another is zero.

vary significantly from one market to another, even when the same range of varieties is supplied in 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.

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:[11]

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

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 harder the market penetration of its variety. Thus, it is fair to say that, 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 aggregated indices:

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

where the cost index is defined as

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

[11] The supply side of the model is kept as simple as possible. Our main purpose is to explore a richer demand framework which can then be incorporated in a full-fledged trade model.

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:

$$p_s^* - c_s = \frac{\alpha_s - c_s}{2} - \mathcal{T}\left(\frac{\mathbb{A} - \mathbb{C}}{2\mathbb{N}}\right) \tag{9}$$

Note that the first term 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 a coefficient that accounts for the *toughness of competition*, i.e.

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

which depends on 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.[12]

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. This is because the price elasticity is given by

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

This expression ranges from 0, when $p_s = 0$, to ∞, when prices equal the intercept of the inverse demand function, $\alpha_s - \gamma Q$. This implies that β_s does not affect ϵ_s and,

[12] 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.

therefore, has no impact on p_s. However, the whole distribution β_r matters because it influences Q.

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

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

These various properties show that our model retains the flexibility displayed by the standard quadratic utility model, while enabling to capture several new effects.

3.6 From theory to empirics

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 i populated by M_i consumers. Consumers living in the same country share the same preferences. The theory then tells us what to expect as price and per capita 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. On the other side, 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:

$$p_{s,i}^* = \frac{\alpha_s + c_s}{2} - \mathcal{T}_i\left(\frac{\mathbb{A}_i - \mathbb{C}_i}{2\mathbb{N}_i}\right) \tag{11}$$

$$q_{s,i}^* = \frac{M_i}{\beta_{s,i}}\left[\frac{\alpha_s - c_s}{2} - \mathcal{T}_i\left(\frac{\mathbb{A}_i - \mathbb{C}_i}{2\mathbb{N}_i}\right)\right] \tag{12}$$

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. This market effect acts like a shifter for all prices in a particular market. Thus, although the general level of prices can differ across markets, 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 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.[13].

It follows from the above analysis that firm-product quantities across destination markets should display more variability than prices. We verify in the next section if this is what we observe in the data.

4 Empirical evidence

The aim of this section is to confront the above model with micro-level data. To this end, we use a unique dataset on Belgian exporters similar to the one used by Bernard et al. (2010). The data is composed of fob (free on board) export prices and quantities by destination market at the firm-product level.[14] 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.

4.1 Data

The Belgian export data used in this paper 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 export flows by 8-digit Combined Nomenclature (CN8) product and by destination country.[15] For firms with primary activity in manufacturing, the data includes over 5,000 exporters and over 7,000

[13]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.

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

[15]The Combined Nomenclature is the European Union's product classification, with 8 digits being the most detailed level. Due to its hierarchical nature, all products expressed as CN8 are also classified as products at more aggregated level such as CN6, CN4 and CN2. Incidentally, CN6 is identical to the HS 6 digit classification, which is the international product classification. The CN classification can be downloaded from the Eurostat Ramon server: http://ec.europa.eu/eurostat/ramon/.

different CN8 products, resulting in more than 60,000 firm-product varieties exported to 220 destination markets in a total of almost 250,000 observations. 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 the theory is about consumption goods, we only consider consumption goods as indicated by the BEC classification.[16]

Because CN8 is the most detailed product-level classification available, we define a *variety* s as a firm-CN8 combination. While our definition of a variety does not change throughout the analysis, the definition of a product and the size of the product-market S_i is allowed to change with the level of product aggregation. This means that, when we use the CN8 as our product definition, each variety is associated with a specific firm. At higher levels of aggregation, a firm can supply several varieties.

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 do not retain a single level of aggregation but repeat our analysis for four levels of aggregation, the CN8, CN6, CN4 and CN2. In a more aggregated product classification, a *product* will then be defined as a collection of varieties (firm-CN8) sharing the same CN code. More broadly defined product-markets will have a higher number of varieties, but the varieties included will be poorer substitutes and, therefore, the assumption of symmetry in substitutability becomes more stringent.[17]

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

Product selection. For each level of aggregation (CN8, CN6, CN4, CN2), we have to choose which products to include in the analysis. Within each product, we analyze differences across markets and varieties. Therefore, we must focus on products which are sold in a sufficiently large number of varieties and markets. In order to ensure that there are enough varieties in enough markets when comparing firm-product prices and quantities, we retain the five products which are supplied under the highest number of varieties at each level of aggregation. The products yielding the highest number of varieties are listed in Table 10 with corresponding CN codes and descriptions.

Market selection. Since our analysis focuses on price and quantity variations 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

[16]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.

[17]Our model assumes that product-markets are characterized by the same pattern of substitutability, γ. Note that constant patterns of substitutability between varieties within a product category, or even the entire economy, is the standard assumption virtually all trade models, be they based on CES or linear quadratic utility functions.

number of markets to compare. However, we also need a sufficient number of varieties to be simultaneously sold in all the markets. The trade-off arises because the number of varieties simultaneously present in all markets drops significantly with each additional destination market. Since 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 *a maximum number of observations to base our analysis on.* We start by considering only those destination markets that are important outlets 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 1. 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 1, 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.[18] Thus, at the bottom of the second column, we obtain the number of varieties present in all markets which is close to 400.

INSERT TABLE 1 HERE.

Product-market samples. The intersection of all the best N-market combinations with the 20 products (i.e., five products for each of the four levels of aggregation) leads to 220 potential data samples. Since some samples are very small, having just 2 or 3 varieties, we further restrict ourselves to samples with more than 10 varieties in order to permit a meaningful correlation analysis between markets. This results in 171 samples. Across these samples, Table 2 provides the effective number of varieties used in our analysis for each level of aggregation (rows) and each best N-country combination (columns).

INSERT TABLE 2 HERE.

4.2 Looking at prices and quantities: rank correlations

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 imposing a less strict interpretation of the theory. This will be relaxed later where we show results also to hold for actual prices and quantities.

Price-quantity ranking correlations within markets. Similarly to what we have shown on the car data example in section 2, we investigate whether, within each

[18] As it turns out, the best N-market combinations happen to be always a sub-group of the (N+1)-market combinations, which allows us to display them in this order.

market, rankings of prices and quantities are significantly correlated. 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 productive differences. Both a Spearman's and a Kendall's rank correlation is applied on the samples resulting from our market and product selection.[19] Results are given in Table 3a where we report them by product-market 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.

Interestingly, results vary a lot depending on the level of aggregation and N-market combination considered. Overall, for the entire sample, the data reject a significant correlation of prices and quantities within markets in about 1/3 of the times. Evaluated in the narrowest product definition, the CN8 level, the rejection rate of a significant correlation is much higher and lies between 76% and 78% of the cases, depending on the statistic used. These results seem to confirm the notion 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.

We now turn to statistics for quantity rank correlations and price rank correlations across markets. Results are reported in Table 3b in a similar format as in table 3a. 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 equal to zero in about 60% of the cases. In table 3c, we show the corresponding results for price rankings between markets. It is striking how much lower the rejection rates are for prices as compared to quantities. The Spearman rank statistic, considers prices to be significantly correlated in 98% of cases, while the corresponding value for the Kendall Tau statistic is about 97%. Put differently, both measures of rank correlations estimate price correlations not to be correlated in only 2 to 3% of the cases.

INSERT TABLES 3b AND 3c HERE.

Between-market price and quantity rank correlations. The between-market predictions are what truly delineate the verti-zontal model from a model with only cost and quality heterogeneity. These two sources of heterogeneity cannot explain a systematically different rank correlation across markets for prices as compared to quantities, which is what we observe. Only the introduction of a third source involving idiosyncratic taste can do it.

An illustrative example: chocolate products. To make our analysis more concrete and to illustrate the discrepancy between price correlations and quantity corre-

[19] The difference between these two approaches to rank correlation is that, whereas the Spearman rank correlation transforms actual values into their relative rank and then compute a standard correlation, the Kendall tau rank correlation measures the frequency of concordant pairs, i.e. observations whose rank coincides.

lations between markets, we focus on one particular product. A product frequently exported from Belgium and included in our data is Belgian chocolates. At the CN8 level, Belgian chocolates fall under "Chocolate products not containing alcohol". For the sake of illustration, we show results limiting ourselves to the best 3-destination market combination, which involves Germany, France and the Netherlands, for which we identify 34 different varieties exported to each of the three destination markets. The values of the pairwise ranking correlations are provided in the top panel of Table 4. We note that price rank correlations ($corr(pp)$) are systematically higher than quantity rank correlations ($corr(qq)$) which suggests that the relative price ranking across the three destination markets is more regular than the quantity ranking. This is true not only for the average correlations across country pairs, but for any country pair correlation, even when CN6 and CN4 definitions of chocolate products are used, which are reported in the middle and bottom panel of Table 4 respectively.

INSERT TABLE 4 HERE.

The general case. While the chocolate example reported the correlation coefficients for 3 chocolate-related CN samples, the same analysis can be repeated for the remaining 168 samples in our data and, for robustness, for the whole manufacturing. In order to give the reader a sense of the pattern that emerges from all the pairwise correlations considered, we report averages.[20] So for reporting purposes 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.[21]

Tables 5a and A.1 report average Spearman and Kendall correlation coefficients and show that that *average price rank correlations between markets are systematically higher than average quantity rank correlations.* This holds irrespective of the number of varieties included (column dimension) and the number of markets considered (row dimension). The difference lies around 15 percentage points, which is relatively similar across the samples.

INSERT TABLES 5a AND A.1 HERE.

As a robustness check, the same rank correlation analysis can be repeated considering the entire manufacturing sector and, thus, there is only one correlation coefficient. Table 5b shows that when doing so previous results are even stronger, i.e. high price correlation but low quantity correlation between markets

INSERT TABLE 5b HERE.

[20] When 3 markets are considered, for example, 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.

[21] All the coefficients associated to each individual sample can be provided upon request.

As noted in section 2, these results do not appear consistent with any combination of two sources of heterogeneity, be they both variety- or market-variety specific. Price correlations between markets are high, suggesting that quality and/or productive efficiency are intrinsic and not market-specific. Yet, quantity correlations are lower, indicating that an additional source of heterogeneity must be present at a market-variety level.

Graphically this can easily be visualized. The coefficients reported in Table 5a are averaged by best N-market combinations and level of disaggregation and plotted in Figure 3a. The simple average by product is instead shown in Figure 3b. The square dots show average price rank correlations for the considered samples, while triangle dots show quantity rank correlations. In the two graphs, these averages are additionally averaged by level of product disaggregation (CN2, CN4, CN6 and CN8), which is represented through the solid line for prices and the dashed line for quantities. It can be observed that price correlations consistently lie well above quantity correlations, especially at narrowest levels of product definitions.

INSERT FIGURES 3a and 3b HERE.

These results support the idea that a third source of heterogeneity needs to be taken into account when dealing with micro-level trade data. Ideally, this third source should affect only quantities sold in different markets, or at least should affect quantities to a larger extent than prices.

4.3 Taking the verti-zontal model to the data

A general feature of quadratic utility functions is that they generate extremely tractable demand functions. Whereas this represents a clear advantage in terms of theoretical developments, it may pose some problems when confronted with real data, as it imposes a linear demand on the data. A legitimate concern may then arise on how restrictive this linearity assumption is. We explore this issue in two ways.

First, we repeat the previous correlation analysis looking at the actual values instead of rankings. If we find correlations on absolute values of prices and quantities to be similar to rank correlations, this suggests that the assumption of linear demand is not very restrictive. To see this, consider the case where demand is non-linear. If the rankings of prices show a strong positive correlation, this may just imply that prices are monotonic (not necessary linear) in quality, marginal costs of production and local market characteristics. But when the absolute value of prices shows a similar positive correlation, it must be the case that a linear structure is a good approximation and that local market effects are shifting the demand for all the varieties in a parallel way.

Second, we run an OLS regression on market and variety dummies and consider the variability explained. This will tell us how well a linear regression line fits the cloud of observed prices. The goodness-of-fit of such a regression will tell us something about the validity of our linearity assumption.

Actual correlations of prices and quantities across markets. In Table 6 we show the correlations of actual prices and quantities across markets, which can usefully be compared to the results in Table 5a where Spearman rank correlations have been displayed.

INSERT TABLE 6 HERE.

The average difference between price and quantity correlations across destination markets when using actual values (column 1) is surprisingly similar to the rank correlations, ranging from 15% to more than 25% depending on the sample considered. Correlations lose however, some of their strength due to the possible presence of outliers, different transport costs across markets and any other possible measurement error whose importance was reduced through the use of rankings. This is shown in Figure A.1 and A.2, which are the counterparts of Figures 3a and 3b when actual values are considered instead of rankings. Again it can be noted that average price correlations (square dots) are much higher than average quantity correlations (triangle dots) independent of the product aggregation and independent of the number of destination markets that are included in the sample.

INSERT FIGURES A.1 and A.2 HERE.

These results suggest that prices across markets depend on some variety-specific characteristics which have a similar impact across markets, while quantities sold appear to be affected by "something else". In our model, this "something else" is captured by market-variety specific differences in the liking by consumers of a set of product characteristics. It is also worth noting that if destination market-specific factors, such as institutions or market size, affected Belgian exports in a similar fashion, this would not affect correlation coefficients within a product category.[22] We build on this point in the next step of our exploratory analysis, where we show that variety- and market-dummies capture the variability of prices across markets much better than for quantities.

OLS regression and goodness of fit. Once we accept that at least three sources of heterogeneity seem to be present in micro-level trade data, we go one step further and see if the way in which they are combined in the verti-zontal model is consistent with the prices and quantities observed. Turning to equation (11), we observe that profit-maximizing prices depend on a variety-specific component, indexed by s, and a market specific component, indexed by i. Thus, the first term differs across varieties but not across markets, whereas the last term varies across destination markets, capturing relevant dimensions of local competitive pressure. As shown by equation (12), profit-

[22]Note that bigger markets could be expected to buy more products of a particular type. But this does not necessarily mean that each variety will sell more, as a bigger market is typically served by more varieties (Baldwin and Harrigan, 2011). Hence the effect of market size on the actual sales of a particular variety is not clear a priori.

maximizing quantities also depend on a market-variety specific taste component(β).

These implications of the model can be empirically tested by regressing individual firm-product prices and per capita quantities ($y_{s,i}$) on variety-specific dummies and destination market-specific dummies as in (13):[23]

$$y_{s,i} = \delta_0 + \delta_1 Variety_s + \delta_2 Market_i + \epsilon_{s,i} \tag{13}$$

We run the specification in (13) on the 171 data samples identified. Note that the unit of observation is always an individual variety, defined by the combination of a firm and a CN8 product code, 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 market.

In terms of the verti-zontal model, the first dummy on the RHS in (13) is meant to capture all the variety-specific characteristics, i.e. marginal cost of production and idiosyncratic quality while the second dummy is expected to capture destination market-wide differences. A high R^2 for prices then suggests that each variety has some intrinsic characteristics determining pricing decisions. Based on the equilibrium quantity expression, we would expect a systematically lower R^2 for quantities, due to the presence of market-variety characteristics which vary both per variety - s - and destination market - i - thus reducing the amount of sample variability explained by the two dummies. As a benchmark, the reader can bear in mind the implications of alternative models other than the verti-zontal model. In a pure cost or quality model we would expect the independent variables in (13) to explain an equal amount of variability of both prices and quantities, which is not what we find in the data. Also, the predictions of the verti-zontal model can be contrasted with a model of market-specific demand shifters (capturing, say, differently perceived quality), rather than a variety-specific demand shifter in the verti-zontal model. Based on such a model we would expect only a negligible amount of variability to be explained by our two sets of dummies for both quantities and prices, while results suggest the opposite. The average (R^2) for regression (13) are summarized in Table 7.

INSERT TABLE 7 HERE.

The price regressions have an R^2 of between 60 to 70% depending on the sample that is used, which is systematically higher than the one associated with quantity regressions that ranges between 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% in the captured variability between price and quantity regressions. Browsing Table 7, we see that this difference is systematically present, no matter which product-market definition or market combination is used. This consistently higher goodness-of-fit for price as opposed to

[23]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 qualitatively the same as those obtained here.

quantity regressions can be interpreted as the effect on quantities of different tastes in different markets.

The differences in goodness-of-fit are displayed in Figure B.1 and B.2, 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. It can be noted that the OLS fit is systematically better in the price regressions that in the quantity regressions

INSERT FIGURES B.1 and B.2 HERE.

Omitted variable tests. In order to complement our analysis of the variability explained by the regressions, we run a test especially designed to verify the functional forms used in the theory and to test for omitted variable bias, which is the Ramsey's RESET. We know that a low R^2 may be caused by omitted variables or non-linear functional forms involving variety- and market-specific effects. In what follows, we use the RESET to assess their respective role. This test is performed for each of the actual samples on which regressions are run. Table 8 shows how many times the RESET test is passed. The results are strikingly different for price and quantity regressions. The top row 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-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 8 disaggregates this by levels of product aggregation and best-N market combinations. The difference between price regressions and quantity regressions is again striking, especially at the narrowest levels of product aggregation and for an intermediate number of destination markets. 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.

INSERT TABLE 8 HERE.

Overall variability explained. Up to this point, our analysis has always been restricted to the 171 product-markets identified earlier. As a robustness check, the same regression analysis can be repeated but now considering the entire manufacturing sector. Implicitly this amounts to attributing the same pattern of substitutability to all the varieties produced, which is a convenient assumption also present in CES models used to study economy-wide issues. Put differently, we now consider the entire economy as a product-market and introduce market and variety dummies as before. In addition to market-dummies, for robustness we also verify results when substituting market dummies by market-product dummies with products defined at a 2-digit CN level. In this way, we can spot differences in local competitive pressure across products, which may affect prices and quantities differently. To this end, the empirical specification in (13) may be rewritten as follows:

$$y_{s,i} = \delta_0 + \delta_1 Variety_s + \delta_2 ProductMarket_i + \epsilon_{s,i} \tag{14}$$

An important caveat is that the unit of measurement in which per capita quantities are expressed in the data can differ when dealing with the whole manufacturing sector. While in the large majority of cases quantities are expressed in kilograms, for some products another unit of measurement is used (liters, pairs, square meters and so on). This did not constitute a problem as long as our analysis was restricted to specific product definitions, which are always measured in the same way, but it becomes more of an issue in (14), as different units of measure now co-exist in the sample. To account for this, we consider the results for varieties whose quantity is expressed in terms of weight (kilograms) separately from those varieties whose quantities are expressed in units.

The results are listed in Table 9. By and large, we see that the main determinants of the model still explain a substantial part of the variation, even when including the entire set of varieties in the manufacturing sector. This is true both for varieties expressed in units (columns 1,2) and for varieties whose quantities are expressed in weight (columns 3, 4).

INSERT TABLE 9 HERE.

Consistently with our previous results, the amount of variation captured by these two simple sets of dummies is impressive, and so is their difference. It is also interesting to note how the R^2 of price regressions on dummies remains virtually identical as we move from the inclusion of pure market dummies (columns 1, 3) to product-market dummies (columns 2, 4), suggesting that regulation or any other product-level source of variability within a geographical market does not add much information in the determination of variety profit-maximizing prices. Surprisingly, this is again not true when looking at quantity regressions, whose R^2 is indeed sensitive to the kind of product-market dummy considered. In other words, price differences across markets are the same for all product categories, whereas quantity differences are not. For example, shoes and beers exported to France can be more expensive than shoes and beers exported to Poland, but the French may want to buy more shoes than the Polish whereas the Polish may prefer to buy Belgian beers rather than shoes. In our model, this quantity effect is captured by the parameter β. That market characteristics, such as population size, wealth and institutions, are less relevant for quantities than for prices is evidence that there exist a source of variability that affects quantities and not prices.

Does geography matter? Finally, we ask ourselves whether our results may be driven by the fact that most destination countries included in our analysis are European (see Table 1). Indeed, European integration may have a dampening effect on price differences as a result of arbitrage, proximity or lack of border controls, which could explain the high price correlation observed in the data. Even if we find it hard to see how this could explain the low correlation in terms of quantities sold, we consider this a legitimate concern. For this reason, we check whether a different country selection could have affected our results. We do so by considering a range of heterogeneous and 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 choice 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 4 for prices and quantities, sorting them by decreasing quantity rank correlation. The results are again surprising but in line with earlier results. 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, whereas quantities sold are far less similar.

In fact, if anything it appears that the original sample selection containing mostly European countries may generate results against our modelling choices. This can be seen again from Figure 4. Of all the countries included in this new sample, the ones displaying the highest pairwise quantity rank correlations are the 3 European countries, with an average price rank correlations also above average. In our setting, this would be associated with countries sharing similar tastes or, more precisely, countries with similar taste mismatch between their ideal variety characteristics and the actual characteristics of the 87 varieties considered. This means that our original sample selection containing mostly European countries may have overestimated the regularity of quantities sold across markets and underestimated the real distance between price and quantity coefficients in correlation and regression analyses.

5 Conclusions

Existing trade models are unable to explain the richness of new firm-product-country level trade data, thus calling for a new generation of models. This paper proposes a generalization of the quadratic utility model to respond to this challenge. By enriching the demand side to account for non-symmetric varieties through a precise interpretation of horizontal and vertical differentiation, we have developed a framework in which taste heterogeneity interacts with quality and cost heterogeneity to shape the market outcome. In particular, the vertical attributes of each variety interact with the local market perception of its horizontal characteristics in such a way that even the same mass of varieties can generate a different level of competitive pressure in different markets.

In this way, our model can address the concerns raised by a growing number of empirical studies that fail to find evidence in support of existing models when confronted with micro-level data. To further illustrate this point, we have used a unique dataset on Belgian exporters, with information on products and destinations, and find that one of the weakest points of existing theories lies in assuming prices and quantities to be determined in equilibrium by the same set of parameters in segmented markets. By looking at prices and quantities across markets, we show that this assumption is unfounded. We tackle this issue by accounting for taste differences through a new way of dealing with horizontal differentiation, which makes our model a valuable building block to be integrated in models where the supply side is more developed.

To keep the model as general as possible, we did not assume any particular link between cost, quality and taste distributions. While other papers require quality and marginal cost to be positively correlated, the model presented here does not impose any restrictions on how quality is brought about, be it through higher marginal costs, 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. It is a matter of further empirical work to determine whether high quality goods sell relatively more in richer countries. Indeed, since model we develop remains largely agnostic about the supply side of the economy, the improvements proposed on the modelling of the demand side be directly used as a module that can be incorporated into any future trade model.

It is worth noting that accounting for taste differences may also have implications for the way in which quality is currently measured. Once we allow for markets to be characterized by different tastes, specific varieties can sell more than others at the same price and quality because they match local tastes better. This suggests being careful when trying to infer quality by looking only at prices and quantities sold in one market.

A final word: our model provides a reconciliation between localized competition à la Hotelling and non-localized competition à la Chamberlin. Indeed, in our setting global competition is affected by the proximity/remoteness among varieties through simple and intuitive market aggregates.

Figure 1a: Scatterplot of price against quantity rankings for car models sold in Belgium, France, Germany, Italy and UK within each market.

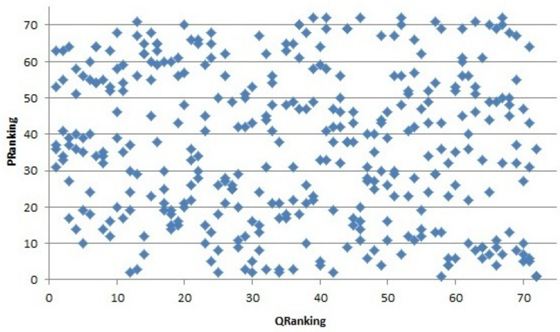


Figure 1b: Price ranks in France, Germany, Italy and UK against Belgium.

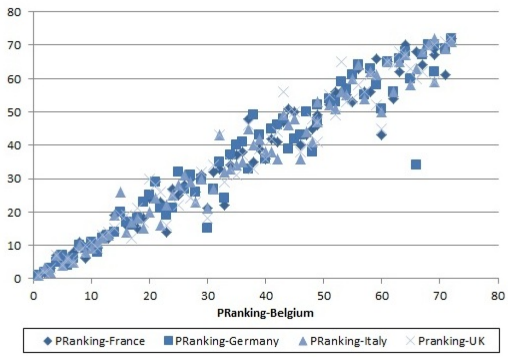


Figure 1c: Quantity ranks in France, Germany, Italy and UK against Belgium.

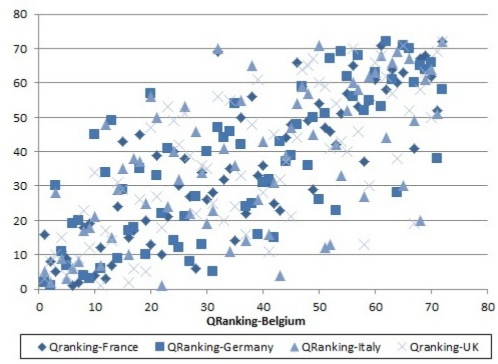


Figure 2: Graphical intuition of the spatial problem

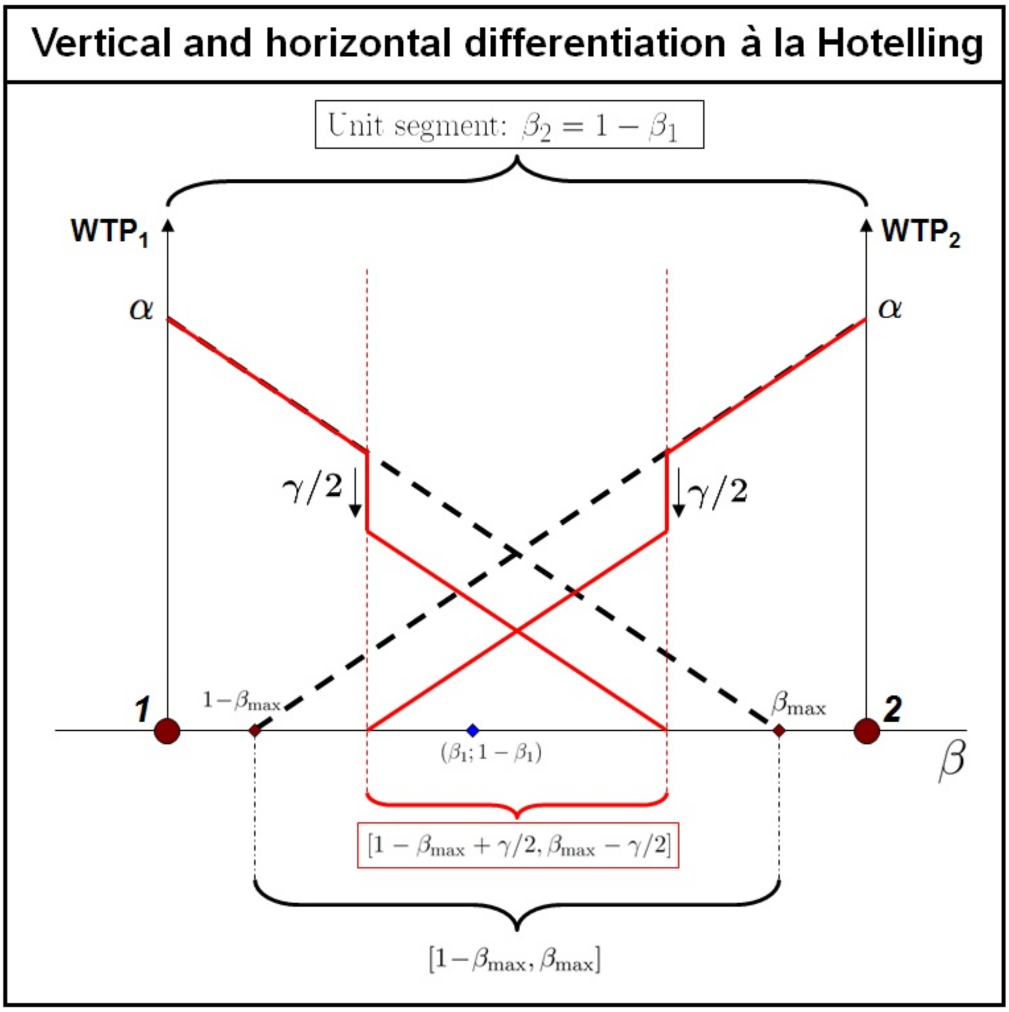
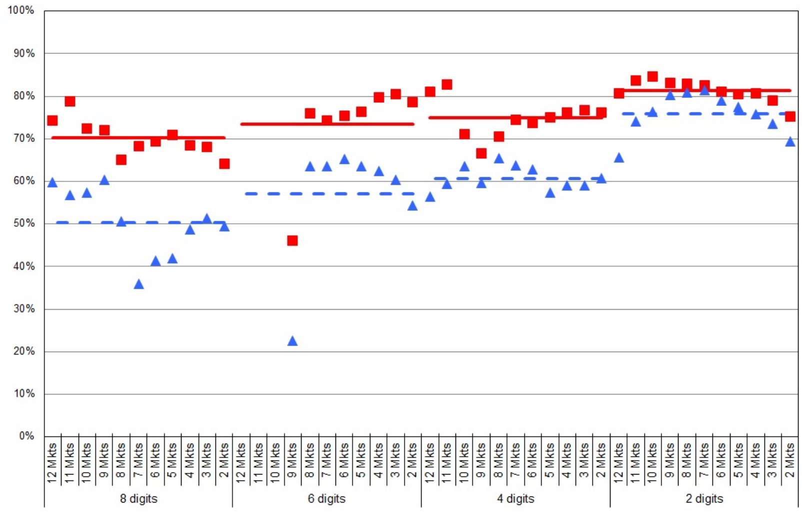
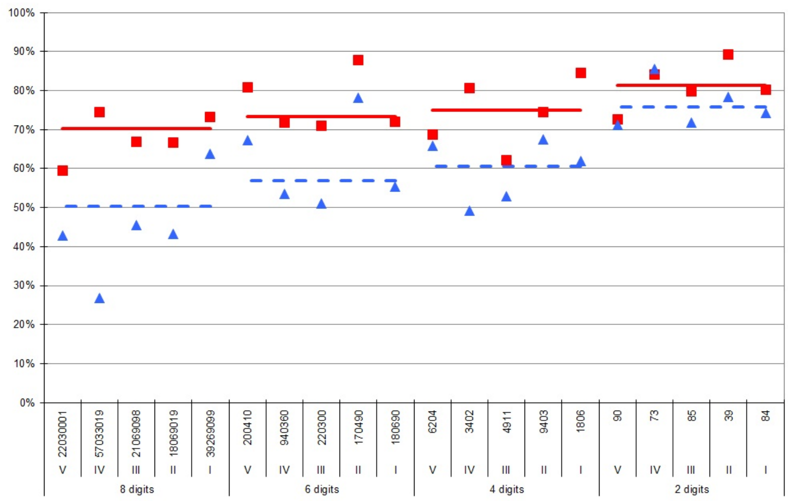


Figure 3a: Visual representation of the results reported in Table 5a.



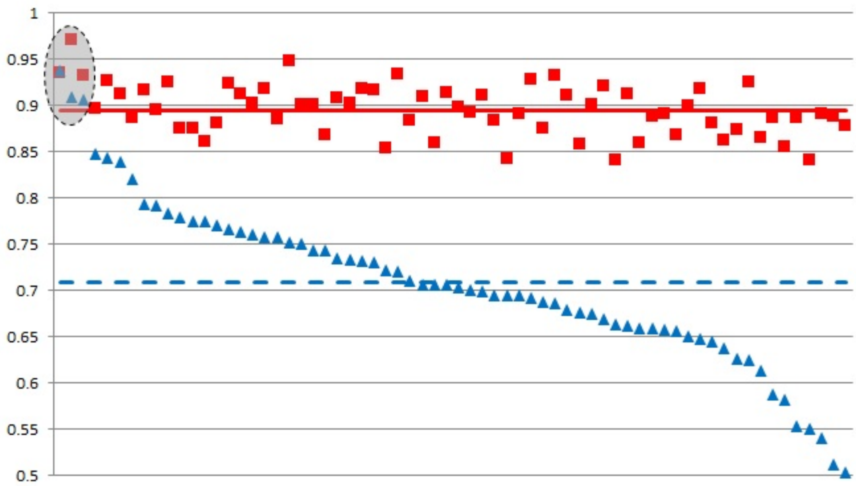
Notes: Square dots indicate average price rank correlations by best N-market combination across product codes, triangle dots indicate the same for quantity rank correlations. The horizontal line segments refer to average rank correlations across best N-market combinations by level of product disaggregation: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

Figure 3b: Visual representation of the results reported in Table 5a.



Notes: Square dots indicate average price rank correlations by product code across best N-market combinations, triangle dots indicate the same for quantity rank correlations. The horizontal line segments refer to average rank correlations across product codes by level of product disaggregation: the solid one refers to prices, the dashed one to quantities.

Figure 4: 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.

Table 1: 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

Note: 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 2: Varieties considered in each intersection of best N-market combination and level of product disaggregation.

Number of **varieties** considered	**Best N-country combinations**										
	Countries N=2	Countries N=3	Countries N=4	Countries N=5	Countries N=6	Countries N=7	Countries N=8	Countries N=9	Countries N=10	Countries N=11	Countries 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

Note: 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 1 become some varieties report both weight and units and therefore are counted only once in Table 1.

Table 3a: Rejection rates for within-market rank correlations.

Spearman and ***Kendall Tau*** rank correlation: Rejection of significance for ***price-quantity correlations within markets***

All the samples	**Spearman**	**35.3%**
	Kendall	**37.6%**
	Samples	(171)

		CN8	CN6	CN4	CN2
By level of product aggregation:	**Spearman**	**76.3%**	**25.7%**	**48.9%**	**1.8%**
	Kendall	**78.9%**	**34.3%**	**48.9%**	**1.8%**
	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:	**Spearman**	**50.0%**	**37.5%**	**44.4%**	**54.5%**	**47.1%**	**55.0%**	**40.0%**	**35.0%**	**25.0%**	**15.0%**	**15.0%**
	Kendall	**50.0%**	**37.5%**	**44.4%**	**54.5%**	**52.9%**	**60.0%**	**40.0%**	**35.0%**	**30.0%**	**20.0%**	**15.0%**
	Samples	(6)	(8)	(9)	(11)	(17)	(20)	(20)	(20)	(20)	(20)	(20)

Note: Percentages of samples not significantly correlated at a 5% level are reported by product aggregation and market combination. The number of samples considered is reported in brackets. For example, looking at Spearman rank correlations at a CN8 level of product aggregation, 76.3% of the 38 samples considered are not significantly different from 0.

Table 3b: Rejection rates for between-market quantity rank correlations.

Spearman and ***Kendall Tau*** rank correlation: Rejection of significance for ***quantity correlations between markets***												
All the samples	**Spearman**	**19.1%**										
	Kendall	**19.7%**										
	Samples	(171)										
By level of product aggregation:		CN8		CN6		CN4		CN2				
	Spearman	**60.5%**		**8.6%**		**15.6%**		**0.0%**				
	Kendall	**60.5%**		**11.4%**		**15.6%**		**0.0%**				
	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
	Spearman	**50.0%**	**37.5%**	**22.2%**	**18.2%**	**29.4%**	**25.0%**	**20.0%**	**15.0%**	**15.0%**	**10.0%**	**5.0%**
	Kendall	**50.0%**	**37.5%**	**22.2%**	**18.2%**	**29.4%**	**30.0%**	**20.0%**	**15.0%**	**15.0%**	**10.0%**	**5.0%**
	Samples	(6)	(8)	(9)	(11)	(17)	(20)	(20)	(20)	(20)	(20)	(20)

Note: Percentages of samples not significantly correlated at a 5% level are reported by product aggregation and market combination. The number of samples considered is reported in brackets. For example, looking at both Spearman and Kendall rank correlations at a CN8 level of product aggregation, 60.5% of the 38 samples considered are not significantly different from 0.

Table 3c: Rejection rates for between-market price rank correlations.

Spearman and ***Kendall Tau*** rank correlation: Rejection of significance for ***price correlations between markets***												
All the samples	**Spearman**	**2.9%**										
	Kendall	**3.5%**										
	Samples	(171)										
By level of product aggregation:		CN8		CN6		CN4		CN2				
	Spearman	**5.3%**		**2.9%**		**4.4%**		**0.0%**				
	Kendall	**10.5%**		**2.9%**		**2.2%**		**0.0%**				
	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
	Spearman	**0.0%**	**0.0%**	**11.1%**	**18.2%**	**11.8%**	**0.0%**	**0.0%**	**0.0%**	**0.0%**	**0.0%**	**0.0%**
	Kendall	**0.0%**	**0.0%**	**0.0%**	**18.2%**	**11.8%**	**5.0%**	**0.0%**	**0.0%**	**0.0%**	**5.0%**	**0.0%**
	Samples	(6)	(8)	(9)	(11)	(17)	(20)	(20)	(20)	(20)	(20)	(20)

Note: Percentages of samples not significantly correlated at a 5% level are reported by product aggregation and market combination. The number of samples considered is reported in brackets. For example, looking at Spearman rank correlations at a CN8 level of product aggregation, 5.3% of the 38 samples considered are not significantly different from 0.

Table 4: Spearman rank correlations for chocolate products.

CN8 - 18069019 (Best 3 markets)	**Chocolate products not contanining alcohol**			
	(1)	(2)	(3)	(4)
Market pairs	**Average**	FR-NL	FR-DE	NL-DE
Rank Corr(pp)	**71.35%**	69.38%	66.81%	77.85%
Rank Corr(qq)	**56.01%**	44.17%	56.17%	67.70%
Varieties	**34**	34	34	34
CN6 - 180690 (Best 3 markets)	**Chocolate products**			
	(1)	(2)	(3)	(4)
Market pairs	**Average**	FR-NL	FR-DE	NL-DE
Rank Corr(pp)	**80.99%**	78.79%	80.85%	83.32%
Rank Corr(qq)	**60.67%**	56.25%	59.09%	66.67%
Varieties	**94**	94	94	94
CN4 - 1806 (Best 3 markets)	**Chocolate and other food preparations containing cocoa**			
	(1)	(2)	(3)	(4)
Market pairs	**Average**	FR-NL	FR-DE	NL-DE
Rank Corr(pp)	**83.84%**	82.56%	84.77%	84.18%
Rank Corr(qq)	**65.52%**	64.47%	61.95%	70.15%
Varieties	**150**	150	150	150

Note: Spearman rank correlations for prices and quantities between markets are reported for the product codes involving chocolate present in our "top 5" product list, considering the "best 3 destination markets" .

Table 5a: Between-market Spearman price and quantity rank correlations.

Spearman Rank correlations		Average of averages	CN8 (Average of Top 5)	CN6 (Average of Top 5)	CN4 (Average of Top 5)	CN2 (Average of Top 5)
		(1)	(2)	(3)	(4)	(5)
Average of Averages	p	**75.29%**	**70.17%**	**73.36%**	**74.90%**	**81.25%**
	q	**61.38%**	**50.36%**	**56.99%**	**60.66%**	**75.79%**
2-market combination	p	**73.55%**	64.16%	78.67%	76.14%	75.24%
	q	**58.51%**	49.38%	54.43%	60.83%	69.42%
3-market combination	p	**76.05%**	68.00%	80.48%	76.69%	79.00%
	q	**61.13%**	51.43%	60.37%	59.13%	73.59%
4-market combination	p	**76.25%**	68.52%	79.68%	76.10%	80.70%
	q	**61.49%**	48.71%	62.40%	59.09%	75.77%
5-market combination	p	**75.71%**	70.96%	76.41%	75.09%	80.40%
	q	**60.09%**	41.99%	63.64%	57.32%	77.41%
6-market combination	p	**74.93%**	69.45%	75.50%	73.71%	81.05%
	q	**62.12%**	41.45%	65.28%	62.72%	79.02%
7-market combination	p	**74.85%**	68.25%	74.22%	74.46%	82.46%
	q	**61.19%**	36.00%	63.62%	63.83%	81.33%
8-market combination	p	**73.56%**	65.02%	75.91%	70.46%	82.86%
	q	**65.13%**	50.61%	63.64%	65.43%	80.83%
9-market combination	p	**66.92%**	72.10%	45.99%	66.50%	83.09%
	q	**55.69%**	60.42%	22.53%	59.58%	80.22%
10-market combination	p	**76.01%**	72.36%		71.03%	84.64%
	q	**65.76%**	57.43%		63.56%	76.28%
11-market combination	p	**81.71%**	78.73%		82.76%	83.63%
	q	**63.45%**	56.81%		59.38%	74.17%
12-market combination	p	**78.68%**	74.37%		80.96%	80.69%
	q	**60.59%**	59.72%		56.43%	65.61%

Note: Between-market Spearman price and quantity rank correlations are reported for the varieties present in Table 2. Coefficients are averaged across the number of samples present per intersection of best N-market combination and product disaggregation.

Table 5b: Between-market price and quantity rank correlations for the whole manufacturing.

Rank correlations		Volumes expressed in Units		Volumes expressed in Weight	
		Spearman	***Kendall***	***Spearman***	***Kendall***
		(1)	(2)	(3)	(4)
Average of Averages	Price	**95.65%**	**84.85%**	**92.93%**	**79.56%**
	Quantity	**77.66%**	**59.29%**	**79.50%**	**60.91%**
2-market combination	Price	95.10%	83.98%	91.46%	77.43%
	Quantity	75.64%	56.71%	74.62%	55.73%
3-market combination	Price	95.49%	84.83%	92.40%	78.93%
	Quantity	76.78%	58.09%	77.29%	58.53%
4-market combination	Price	95.64%	85.19%	93.17%	79.81%
	Quantity	78.11%	59.46%	78.38%	59.68%
5-market combination	Price	95.86%	85.66%	93.45%	80.26%
	Quantity	79.41%	60.75%	80.05%	61.37%
6-market combination	Price	96.14%	85.89%	93.46%	80.43%
	Quantity	78.86%	60.58%	81.39%	62.86%
7-market combination	Price	96.12%	85.85%	93.32%	80.21%
	Quantity	78.62%	60.33%	82.07%	63.63%
8-market combination	Price	95.91%	85.22%	93.16%	80.03%
	Quantity	77.91%	59.59%	82.41%	63.95%
9-market combination	Price	95.87%	84.97%	93.21%	80.04%
	Quantity	76.38%	58.05%	82.33%	63.73%
10-market combination	Price	95.67%	84.73%	93.95%	81.13%
	Quantity	76.80%	58.85%	80.65%	62.01%
11-market combination	Price	95.60%	84.17%	92.34%	78.71%
	Quantity	77.42%	59.28%	77.97%	59.51%
12-market combination	Price	94.71%	82.91%	92.32%	78.17%
	Quantity	78.32%	60.50%	77.38%	58.97%

Note: Between-market Kendall and Spearman price and quantity rank correlations are reported for the whole manufacturing in each best N-market combination. Correlations are computed separately for varieties whose quantities are reported in weigh and varieties whose quantities are reported in units.

Table 6: Between-market price and quantity simple correlations.

Between-market correlations		Average of averages (1)	CN8 (Average of Top 5) (2)	CN6 (Average of Top 5) (3)	CN4 (Average of Top 5) (4)	CN2 (Average of Top 5) (5)
Average of Averages	p	**71.55%**	**74.31%**	**74.36%**	**70.98%**	**66.20%**
	q	**56.35%**	**52.39%**	**50.43%**	**57.32%**	**61.87%**
2-market combination	p	**57.56%**	51.58%	68.90%	59.22%	50.55%
	q	**46.94%**	40.78%	51.49%	34.97%	60.51%
3-market combination	p	**70.88%**	70.31%	81.28%	70.48%	61.47%
	q	**49.19%**	37.76%	50.82%	41.36%	66.83%
4-market combination	p	**73.34%**	73.49%	80.13%	72.15%	67.59%
	q	**50.22%**	41.21%	49.78%	43.06%	66.81%
5-market combination	p	**73.51%**	78.03%	78.59%	68.53%	68.89%
	q	**51.80%**	40.85%	53.94%	50.52%	61.91%
6-market combination	p	**72.26%**	74.60%	74.88%	67.80%	71.75%
	q	**54.61%**	43.06%	54.19%	57.77%	63.40%
7-market combination	p	**73.72%**	75.32%	78.78%	70.81%	69.97%
	q	**55.52%**	42.97%	51.12%	62.62%	65.36%
8-market combination	p	**73.63%**	74.15%	82.19%	67.09%	71.09%
	q	**61.52%**	62.62%	51.13%	66.64%	65.68%
9-market combination	p	**65.06%**	78.74%	50.15%	62.47%	68.89%
	q	**61.42%**	78.32%	41.01%	64.56%	61.78%
10-market combination	p	**75.52%**	83.09%		74.30%	69.18%
	q	**68.08%**	72.10%		71.45%	60.70%
11-market combination	p	**75.90%**	78.52%		83.61%	65.57%
	q	**62.34%**	58.32%		68.40%	60.30%
12-market combination	p	**75.71%**	79.57%		84.38%	63.20%
	q	**58.26%**	58.34%		69.12%	47.33%

Note: Between-market price and quantity correlations are reported for the varieties present in Table 2. Coefficients are averaged across the number of samples present per intersection of best N-market combination and product disaggregation.

Table 7: R^2 associated with prices and quantities regressed on dummies.

R-squared in regressions on dummies		Average of averages	CN8 (Average of Top 5)	CN6 (Average of Top 5)	CN4 (Average of Top 5)	CN2 (Average of Top 5)
		(1)	(2)	(3)	(4)	(5)
Average of Averages	Price	**72.13%**	**74.35%**	**79.16%**	**69.22%**	**65.78%**
	Quantity	**49.89%**	**49.70%**	**50.88%**	**46.28%**	**52.71%**
2-market combination	Price	**77.05%**	75.46%	81.12%	77.15%	74.48%
	Quantity	**67.48%**	64.35%	68.48%	60.58%	76.51%
3-market combination	Price	**78.15%**	77.26%	86.03%	77.95%	71.37%
	Quantity	**58.32%**	52.78%	58.75%	50.91%	70.82%
4-market combination	Price	**76.74%**	78.20%	83.34%	76.90%	68.51%
	Quantity	**51.93%**	46.41%	51.20%	45.98%	64.13%
5-market combination	Price	**75.11%**	80.84%	81.58%	69.17%	68.85%
	Quantity	**47.19%**	43.25%	47.68%	41.03%	56.81%
6-market combination	Price	**71.92%**	72.77%	72.76%	68.58%	73.59%
	Quantity	**45.70%**	40.06%	45.50%	41.49%	55.75%
7-market combination	Price	**73.81%**	76.40%	79.14%	68.87%	70.85%
	Quantity	**42.07%**	37.83%	41.48%	39.91%	49.07%
8-market combination	Price	**73.72%**	77.84%	79.24%	66.23%	71.55%
	Quantity	**41.60%**	37.98%	41.44%	38.68%	48.32%
9-market combination	Price	**67.08%**	77.59%	70.03%	62.46%	58.25%
	Quantity	**46.18%**	46.91%	52.53%	40.40%	44.88%
10-market combination	Price	**62.49%**	69.59%		58.95%	58.93%
	Quantity	**46.61%**	51.56%		43.88%	44.38%
11-market combination	Price	**60.98%**	62.53%		62.97%	57.43%
	Quantity	**53.36%**	58.93%		59.74%	41.40%
12-market combination	Price	**63.80%**	69.35%		72.23%	49.83%
	Quantity	**46.97%**	66.62%		46.50%	27.78%

Note: This table reports R^2 associated with OLS regressions of prices and per capita quantities on dummies for the varieties present in Table 2. Coefficients are averaged across the number of samples present per intersection of best N-market combination and product disaggregation.

Table 8: Success rates in tests for omitted variables in the regressions on dummies run for Table 7.

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)

Note: 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.

Table 9: R^2 associated with prices and quantities regressed on dummies for the entire manufacturing.

R-squared in regressions on dummies		Volumes expressed in Units		Volumes expressed in Weight	
		Variety and market dummies	Variety and market-product dummies	Variety and market dummies	Variety and market-product dummies
		(1)	(2)	(3)	(4)
Average of Averages	Price	**88.90%**	**89.77%**	**64.03%**	**65.36%**
	Quantity	**36.32%**	**52.19%**	**33.34%**	**44.91%**
2-market combination	Price	92.77%	92.86%	84.38%	85.29%
	Quantity	54.00%	56.15%	55.74%	57.45%
3-market combination	Price	91.95%	92.06%	87.48%	87.64%
	Quantity	37.40%	40.35%	38.54%	42.30%
4-market combination	Price	89.03%	89.32%	88.59%	88.92%
	Quantity	45.38%	54.26%	33.79%	41.76%
5-market combination	Price	80.56%	80.89%	70.76%	71.57%
	Quantity	39.00%	51.74%	27.18%	38.84%
6-market combination	Price	79.44%	79.67%	66.26%	67.61%
	Quantity	42.49%	50.04%	23.28%	34.64%
7-market combination	Price	95.33%	95.47%	58.17%	59.52%
	Quantity	38.63%	45.25%	22.82%	40.80%
8-market combination	Price	95.78%	95.95%	53.43%	55.28%
	Quantity	35.57%	44.68%	23.93%	37.15%
9-market combination	Price	95.92%	96.10%	36.86%	37.42%
	Quantity	32.80%	48.89%	43.16%	49.58%
10-market combination	Price	95.98%	96.30%	58.02%	59.24%
	Quantity	31.85%	42.05%	43.77%	56.46%
11-market combination	Price	95.56%	95.83%	54.83%	56.98%
	Quantity	23.80%	44.56%	44.31%	55.90%
12-market combination	Price	65.61%	73.04%	45.60%	49.45%
	Quantity	18.63%	96.15%	10.26%	39.13%

Note: R^2 associated with prices and per capita quantities regressed on dummies for the entire manufacturing, i.e. for all the varieties present in each best N-market combination. Regressions are run separately for varieties whose quantities are reported in weight and varieties whose quantities are reported in units.

Table 10: 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	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

References

Anderson, S., Palma, A. and Thisse, J. (1992). *Discrete choice theory of product differentiation*. The MIT Press.

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., Beveren, I. V. and Vandenbussche, H. (2010). Multi-product exporters, carry-along trade and the margins of trade. Research series 201010-203, National Bank of Belgium.

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

Chamberlin, E. (1933). *The theory of monopolistic competition*. Cambridge (Mass.): Harvard University Press.

Crozet, M., Head, K. and Mayer, T. (2011). Quality sorting and trade: firm-level evidence for french wine. *Review of Economic Studies* .

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.

Fajgelbaum, P. D., Grossman, G. M. and Helpman, E. (2009). Income distribution, product quality, and international trade. NBER Working Papers 15329.

Feenstra, R. C. (2003). A homothetic utility function for monopolistic competition models, without constant price elasticity. *Economics Letters* 78: 79–86.

Ferreira, F. and Waldfogel, J. (2010). Pop internationalism: has a half century of world music trade displaced local culture? NBER Working Papers 15964.

Fontagné, L., Gaulier, G. and Zignago, S. (2008). Specialization across varieties and North-South competition. *Economic Policy* 23: 51–91.

Friberg, R., Paterson, R. W. and Richardson, A. D. (2010). Why is there a home bias? A case study of wine. CEPR Discussion Papers 7885.

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

Goldberg, P. K. and Verboven, F. (2001). The evolution of price dispersion in the european car market. *Review of Economic Studies* 68: 811–48.

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

Hanson, G. and Xiang, C. (2011). Trade barriers and trade flows with product heterogeneity: an application to US motion picture exports. *Journal of International Economics* 83: 14 – 26.

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.

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. (2011). Export prices across firms and destinations. *The Quarterly Journal of Economics* (forthcoming), NBER Working Paper 15342.

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.

Ottaviano, G. I. P., Tabuchi, T. and Thisse, J. (2002). Agglomeration and trade revisited. *International Economic Review* 43: 409–436.

Schott, P. K. (2004). Across-product versus within-product specialization in international trade. *The Quarterly Journal of Economics* 119: 646–677.

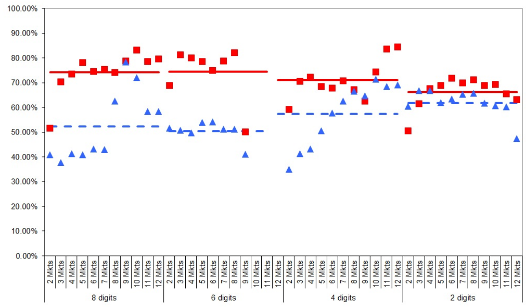
Shaked, A. and Sutton, J. (1983). Natural oligopolies. *Econometrica* 51: 1469–83.

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

Vives, X. (2001). *Oligopoly pricing: old ideas and new tools*. No. 026272040x in MIT Press Books. The MIT Press.

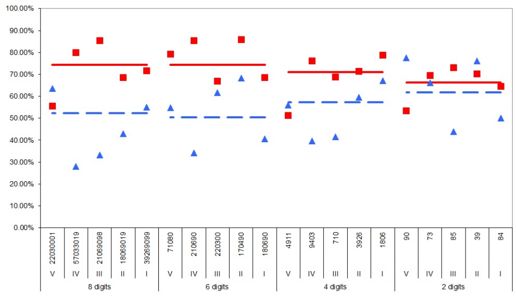
Appendix: Additional Figures and Tables

Figure A.1: Visual representation of the results reported in Table 6.



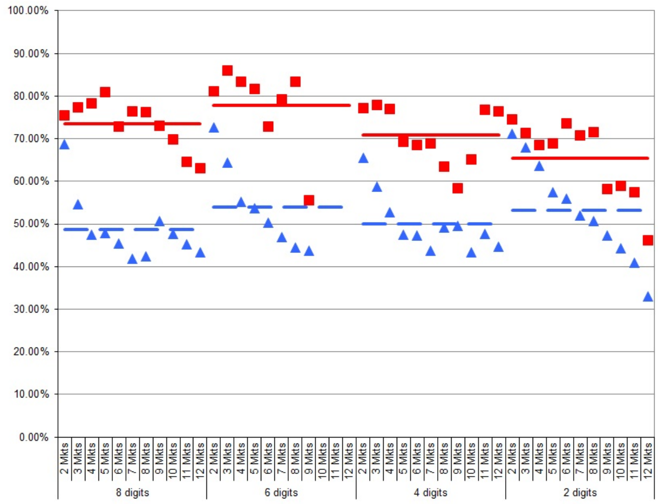
Notes: Square dots indicate average actual price correlations by best N-market combination across product codes, triangle dots indicate the same for quantity correlations. The horizontal line segments refer to average correlations across best N-market combinations by level of product disaggregation: the solid one refers to prices, the dashed one to quantities.

Figure A.2: Visual representation of the results reported in Table 6.



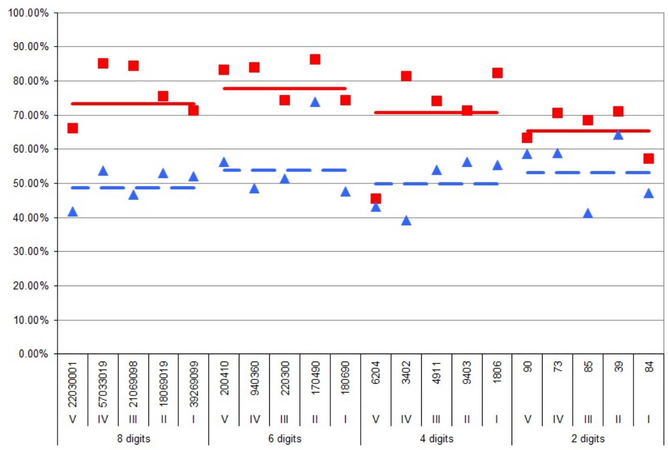
Notes: Square dots indicate average actual price correlations by product code across best N-market combinations, triangle dots indicate the same for quantity correlations. The horizontal line segments refer to average correlations across product codes by level of product disaggregation: the solid one refers to prices, the dashed one to quantities.

Figure B.1: Visual representation of the results reported in Table 7.



Notes: Square dots indicate average R^2 for regressions of prices on dummiesby 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.

Figure B.2: Visual representation of the results reported in Table 7.



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

Table A.1: Between-market Kendall price and quantity rank correlations.

Kendall Rank correlations		Average of averages	CN8 (Average of Top 5)	CN6 (Average of Top 5)	CN4 (Average of Top 5)	CN2 (Average of Top 5)
		(1)	(2)	(3)	(4)	(5)
Average of Averages	p	**59.38%**	**54.51%**	**58.09%**	**59.30%**	**65.60%**
	q	**46.05%**	**37.92%**	**42.40%**	**45.45%**	**58.44%**
2-market combination	p	**57.27%**	49.65%	61.65%	59.09%	58.68%
	q	**42.51%**	35.39%	39.25%	43.68%	51.72%
3-market combination	p	**59.48%**	51.83%	63.42%	60.04%	62.61%
	q	**44.91%**	36.80%	44.22%	43.07%	55.56%
4-market combination	p	**59.97%**	52.57%	63.30%	59.58%	64.42%
	q	**45.57%**	35.45%	46.17%	42.80%	57.85%
5-market combination	p	**59.74%**	55.15%	60.45%	59.01%	64.37%
	q	**44.85%**	30.93%	47.20%	41.96%	59.32%
6-market combination	p	**59.23%**	54.01%	59.88%	58.10%	64.92%
	q	**46.92%**	31.16%	49.19%	46.38%	60.96%
7-market combination	p	**59.40%**	53.03%	59.15%	59.05%	66.38%
	q	**46.86%**	27.32%	48.14%	48.24%	63.74%
8-market combination	p	**58.51%**	50.86%	60.82%	55.15%	67.20%
	q	**50.09%**	38.85%	48.75%	49.35%	63.40%
9-market combination	p	**52.88%**	55.35%	36.05%	52.25%	67.85%
	q	**42.33%**	45.27%	16.30%	44.88%	62.88%
10-market combination	p	**60.60%**	55.64%		56.14%	70.01%
	q	**51.30%**	44.44%		49.80%	59.66%
11-market combination	p	**66.40%**	62.31%		67.48%	69.40%
	q	**49.50%**	44.46%		46.00%	58.04%
12-market combination	p	**63.80%**	59.17%		66.45%	65.78%
	q	**46.87%**	47.05%		43.83%	49.72%

Note: Between-market Kendall price and quantity rank correlations are reported for the varieties present in Table 2. Coefficients are averaged across the number of samples present per intersection of best N-market combination and product disaggregation.