VŠE / FACULTY OF INTERNATIONAL RELATIONS

Working Paper Series

Spreading the Good Apples Out: Market Entry Dynamics of Quality Differentiated Products

Working Paper 2024 - 03

Esteban Jaimovich

Boryana Madzharova

Vincenzo Merella

ISSN 2695-0820

Spreading the Good Apples Out: Market Entry Dynamics of Quality Differentiated Products

Esteban Jaimovich, Boryana Madzharova, Vincenzo Merella*

December 30, 2024

Abstract

This paper examines firms' market entry strategies for quality differentiated goods, focusing on both initial and subsequent product-level entry decisions. Using a theoretical framework incorporating nonhomothetic preferences, we show that premium products are more likely to enter wealthier markets earlier, where producers can capture higher mark-ups. Furthermore, the determinants of follow-up market selection differ by product quality: for premium products, income plays a dominant role in shaping expansion paths, whereas geographic proximity remains the primary driver for low-quality products. Empirically, we test these predictions using micro-level data from the refrigeration industry. Our results confirm a strong positive relationship between market order of entry and income, with this effect being particularly pronounced for high-quality products. Additionally, we observe that, as product quality increases, follow-up markets tend to be geographically more dispersed relative to earlier markets.

JEL Classification: F1; F14; F23; L68.

Keywords: Market entry; Gravity; Nonhomothetic preferences; Quality-differentiated products.

---

*Jaimovich: University of Turin, Corso Unione Sovietica 218 bis, 10134, Turin, Italy, esteban.jaimovich@unito.it; Madzharova: Central Bank of Ireland, New Wapping Street, North Wall Quay, Dublin 1, D01 F7X3, Ireland, Friedrich-Alexander-Universität Erlangen-Nürnberg and CESIfo, boryana.madzharova@centralbank.ie; Merella: University of Cagliari and Prague University of Economics and Business, W. Churchilla 1938/4, 130 67 Prague 3, Czech Republic, merella@unica.it. We thank conference and seminar participants at the Central Bank of Ireland, the Prague University of Economics and Business, and the Workshop in International Economics (University of Turin) for useful comments and suggestions. These are the authors' views, not those of the Central Bank of Ireland.

1 Introduction

Understanding the rationale behind firms' entry patterns across different markets is a fundamental question in international trade. While several contributions to the literature have shown that firms tend to expand sequentially into geographically and culturally proximate markets due to diminishing entry costs with accumulating export experience (e.g., Albornoz et al., 2012, 2023; Chaney, 2014; Defever, Heid, and Larch, 2015; Morales, Sheu, and Zahler, 2019; Egger et al., 2023), there is limited evidence of whether and how these dynamics vary across different levels of product quality.[1] This paper addresses this gap by investigating if firms employ distinct market-entry strategies for premium, high-quality products compared to more affordable, entry-level versions.

We establish that the impact of demand-side factors and geographic proximity in firms' market entry and expansion choices varies across the quality distribution. For premium products, firms prioritize entry into wealthier locations first, where they can charge higher mark-ups as consumers have a stronger preference for high quality and their demand for those products is less price-elastic. In contrast, mark-ups for lower-quality products are less sensitive to income variations, making geographic proximity to existing markets a stronger determinant for entry decisions. The underlying mechanism also drives the dynamic patterns of market expansion, where the gravity pull of existing markets weakens with product quality. Consequently, an "extended" version of the Alchian-Allen effect emerges, in which premium products (the "good apples") are gradually spread to more distant markets relative to earlier sales destinations.

Our paper thus intermingles path dependence in market entry frictions with demand-driven forces stemming from nonhomothetic preferences for vertically differentiated products. To this end, we develop a multi-country dynamic framework that incorporates elements of extended gravity and uses the Almost Ideal Demand System (AIDS) of Deaton and Muellbauer (1980) as a setting within which to introduce nonhomothetic consumer preferences over quality. This demand system accommodates flexible price elasticities tied to both consumers' incomes and product quality. Firms can thus leverage higher mark-ups on premium products in richer markets owing to the lower price sensitivity of wealthier consumers. These increased margins more than compensate for the higher trade costs associated with less proximate markets, resulting in the emergence of heterogeneous dynamic patterns of market expansion for goods of different quality.

Guided by the model's main predictions, we empirically investigate market entry and

[1] Following the terminology introduced in Morales et al. (2019), henceforth, we will refer to the cited collection of studies as the 'extended gravity' literature.

expansion patterns at the *product level* in the specific case of the refrigerator industry. For the purposes of our analysis, this industry has two main advantages: it is characterized by a substantial degree of vertical differentiation and covers goods that are sold in several markets over a number of years. Our empirical strategy relies on a panel dataset comprising product-level monthly sales and prices of refrigerators across 24 European countries in the period 2009-2017. The dataset allows us to track the life cycle of a product within a country and, importantly, virtually across the whole EU for most years. In addition, the dataset contains detailed information on several fridge attributes, which we utilize to generate product-specific quality indices using hedonic regressions.[2]

Based on the obtained quality indices, we investigate the impact of per-capita income on the choice of a first-ever market of entry and the entire order of entry into different markets for models of different quality. To this end, we first carry out a set of conditional logit regressions (to study choices of products' first market of entry) followed by rank-ordered logit regressions (to investigate products' entire order of market entry). The conditional logit analysis shows that first markets of entry tend to be invariably characterized by higher incomes when it comes to products of high quality. The rank-order logit estimates confirm that this pattern also applies to the order of entry and that the relation between choices of earlier market entry and high income weakens as we move down the quality dimension. We also show that these results remain robust when including several potential confounding country-level factors, such as population size and brand familiarity. More importantly, by restricting the analysis to a subsample of fridge models manufactured outside the 24 European countries included in our dataset, we establish that these dynamic patterns are unlikely to be explained by supply-side factors alone.

We next investigate how geographic distances from the first markets of entry influence further expansion choices. Focusing on the second market(s) of entry, we corroborate previous evidence by the extended gravity literature that geographic proximity is an important factor guiding entry: market expansions are more likely to occur in countries nearer to those where a fridge model was first introduced. However, our empirical results also uncover that the intensity of the "gravity pull" varies with quality; namely, it gets weaker at higher levels of quality. More specifically, we show that the proportion of market expansions taking place in markets that do *not* share a border with any of the first markets of entry is greater for products of higher quality. Similarly, the average

[2] As a preliminary step motivating our theoretical model, we first report evidence that the fridge industry engages in *nonhomothetic* pricing-to-market: higher quality fridge models command higher mark-ups in richer markets; conversely, mark-ups do not vary significantly with income for fridges of lower quality. These results confirm those previously obtained by Auer, Chaney, and Saure (2018) for the automobile industry.

distance between the first and second markets of entry is found to be larger for premium fridge models.

Recent literature has highlighted the profound distributional effects of international trade across income levels (e.g., Fajgelbaum, Grossman and Helpman, 2011; Fajgelbaum and Khandelwal, 2016). Our findings further emphasize the disparity in accessing higher-quality products driven by income heterogeneity across countries showing that wealthier markets tend to be the early recipients of improved products from established firms. These dynamic entry patterns across countries align with previous findings by Jaravel (2019) within the context of the U.S. retail sector, where higher-income households experienced a faster increase in product variety and quality. A disproportionately larger presence of premium products in high-income countries implies that such countries are the primary beneficiaries of product innovations. More generally, the unequal diffusion and availability of energy-efficient technologies, new drugs, and other innovations across countries would indicate unequal/delayed opportunities in tackling climate change, healthcare outcomes, and other central matters (e.g., Cockburn et al., 2016). Understanding the drivers behind quality-specialization patterns and their dynamics is thus an important precondition before policymakers can step in to address issues involving quality-specific market segmentation.

By integrating vertical differentiation and entry dynamics with geographic path dependence, our paper bridges two strands of international trade literature that have evolved in parallel. The first strand, the extended gravity literature, primarily explores the dynamic forces driving firms' decisions regarding market entry and expansion. These studies show that firms typically choose nearby markets first to test their export profitability before expanding afield (Albornoz et al., 2012), and seek to capitalize on reduced adaptation costs and improved market information (Defever et al., 2015). They also highlight that firms tend to exploit their existing export experience in similar markets (Morales et al., 2019) and their network of international contacts (Chaney, 2014) to explore more distant markets. We contribute to this literature by examining how differences in consumers' incomes may create deviations from entry choices guided exclusively by proximity considerations in the case of vertically differentiated products. Our focus also differs from the extended gravity tradition in that we examine the entry of new products offered by already established exporters within familiar markets, rather than the behavior of firms as newcomers to specific markets. Hence, our analysis shifts the attention to the entry dynamics at the product level, allowing us to investigate whether distinct patterns arise even in contexts where firms introduce new products of varying quality in markets where they already offered products in the past.

The second strand of literature concerns vertically differentiated industries and nonhomothetic demand schedules, which have typically been framed within static models of comparative advantage. These contributions have emphasized that income elasticities of demand for quality are a central driver of Linder-type (Linder, 1961) international specialization (Flam and Helpman, 1987; Murphy and Shleifer, 1997; Hallak, 2010), and that this could lead to home-market effects as a key determinant of trade flows (Fajgelbaum et al., 2011) as well as higher trade intensity at higher levels of quality (Hummels and Klenow, 2005; Jaimovich and Merella, 2015).[3] Our model and the ensuing empirical analysis are primarily concerned with the impact of nonhomotheticities on the dynamic considerations regarding entry and market expansion decisions along the quality dimension in the presence of geographic frictions.

Lastly, our empirical analysis, leveraging retail data of product-level sales to consumers across different final markets, contributes to the recent surge in the use of product-level data sourced from retail sales to study the impact of product availability and price variation at different levels of consumer income in the presence of nonhomothetic preferences (e.g., Jaravel, 2019; Handbury, 2021; Faber and Fally, 2022). These studies have been carried out for the case of the US based on the Nielsen home scanner data, which contains detailed information on household purchases of many nondurable goods. Our paper complements the work of this strand of the literature by examining multi-country data in a sector with a wide degree of vertical differentiation across durable goods.

The paper proceeds as follows. Section 2 discusses the main dataset and constructs product-specific quality indices from hedonic regressions. Section 3 provides empirical evidence of nonhomothetic pricing-to-market in the refrigeration industry. In Section 4 we document some stylized facts regarding product-market-entry sequences and their relation with income and product quality. Section 5 develops a theoretical multi-country framework embedding variable mark-ups and featuring products' market entry and expansion. Section 6 reports the results from choice models that support the theoretical predictions of faster entry of higher-quality products in richer markets. Section 7 further provides empirical evidence that geographic proximity matters less in the market expansion paths of premium goods relative to entry-level ones as gains from higher mark-ups outweigh distance-based costs. Section 8 concludes.

[3] These results consistently indicate that trade flows are closely linked to countries' income levels, with wealthier nations showing stronger demand for higher-quality products and less affluent markets gravitating towards lower-quality products, as documented by, e.g., Schott (2004), Hallak (2006), Verhoogen (2008), Khandelwal (2010), Hallak and Schott (2011), Manova and Zhang (2012).

2 Data and Construction of Quality Index

2.1 Scanner Data

The empirical analysis is primarily conducted with GfK GmbH's Retail Panel on Major Domestic Appliances, which is a product-level monthly-frequency database comprising the unit sales and VAT-inclusive scanner prices of different types of white goods.[4] As we are interested in products characterized by a high degree of vertical differentiation, we narrow our focus to refrigerators, which represent the type of white goods exhibiting the widest price distribution amongst those in our dataset. The panel covers 22 current EU members plus the UK and Serbia from January 2009 until September 2013 and extends to January 2017 for a subset of eight countries.[5] In addition to prices and monthly quantities, the data contains a number of product characteristics such as brand, energy label, and others summarized in Table A.2 in Appendix A.2.[6]

Forty-two different brands are present in the data. Nineteen of them account for 90% of all observations and, with few exceptions, the top brands are present in all 24 markets. Identical products in the dataset share the same unique identifier across countries. We can thus observe both the country-specific and (at least until 2013) nearly all EU-wide sales of a given product, as well as its contemporaneous prices across its various sales destinations. On average, the dataset records information on 5,421 unique refrigerator models annually, and a total of 11,529 products over the full duration of the panel. These products account for about 74.4% of the EU's aggregate expenditure on refrigerators between 2009-2013.[7]

[4] These include refrigerators, washing machines, dishwashers, and other household appliances. Unit sales are the total units sold of a product across all brick-and-mortar retailers in a given country on a given date (month-year), whereas scanner prices are the unit sales-weighted mean prices across these retailers over the same period (date). We do not observe geographically-differentiated or retailer-specific sales or prices within a country. Sales by online retailers are not part of the database. Throughout the analysis, we will refer to a unique product (e.g. Bosch KAG93AIEP) interchangeably as a 'product', 'model', or 'variety'.

[5] See Table A.1 in Appendix A.2 for detailed time-coverage by country. The UK was still part of the EU during the years covered by the dataset. The five EU markets not included in the data are Bulgaria, Cyprus, Ireland, Luxembourg, and Malta.

[6] The coverage of refrigerators' characteristics is also relatively more comprehensive than that of other types of products in the data (see Table A.2). These appliances are also more diverse in capacity and dimensions, energy efficiency, functionality and settings, and other important attributes.

[7] This share drops from 2014 onwards due to the fall in country coverage. For 2009-2013, the average share was estimated on the basis of yearly aggregate apparent consumption, defined as the value of production plus imports net of exports in the Prodcom database as reported in Table 7 in European Commission (2016). The relevant categories are 27511110 - Combined refrigerators-freezers, with separate external doors, 27511133 - Household-type refrigerators (incl. compression-type, electrical absorption-type) (excl. built-in) and 27511135 - Compression-type built-in refrigerators.

TABLE 1 – Descriptive Statistics

	All		By quality quantile							
			(1)		(2)		(3)		(4)	
	Mean	Mdn	Mean	Mdn	Mean	Mdn	Mean	Mdn	Mean	Mdn
Price	684.2	544.0	356.5	329.3	517.0	472.9	702.7	640.8	1,213.0	1060.0
	(493.7)		(128.3)		(194.5)		(291.1)		(676.7)	
N	912,951		238,651		236,309		223,187		214,804	
Units	35.7		49.0		37.3		31.3		24.0	
	(111.1)		(149.9)		(107.7)		(93.3)		(75.1)	
N	1,026,132		263,219		265,878		251,152		245,883	
Quality	0.148		-0.375		-0.020		0.283		0.777	
	(0.454)		(0.153)		(0.091)		(0.090)		(0.269)	
N	9,817		2,564		2,634		2,267		2,352	

Notes: The table shows descriptive statistics per product per date per country for a subsample that excludes the 1,250 products (about 15,000 observations) in the data sold in only one market throughout their life-cycle. The following basic cleaning has been performed: zero or negative prices are replaced with missing observations; negative unit sales are replaced with missing observations. 'Quality' is the time-invariant quality index constructed from the hedonic specification (2). Columns (1)-(4) report statistics for four quantiles of the quality index. 'N' denotes the number of observations. 'Mdn' abbreviates median value. All prices are in Euro. Table A.4 in the Appendix reports identical statistics for the full sample including single-sale-destination products.

2.2 Time-Invariant Quality Index

To segment the product space by quality, we construct a time-invariant product-specific quality index using a hedonic log-linear model relating the prices of goods to a set of observed essential attributes. In particular, we use the specification:

$$\ln \text{Price}_{jmd} = \sum_{a=1}^{4} b_a \kappa_{aj} + \lambda_{md} + u_{jmd}, \tag{1}$$

where $\ln \text{Price}_{jmd}$ is the natural logarithm of the price of product j in country/market m on date (month-year) d and κ_{aj} is attribute a, which is product-specific. We consider four separate attributes coded as categorical variables; namely, number of doors and freezer position, availability of no-frost function, energy label, and brand (see Table A.2).[8] Since there are multiple price observations for each product over countries and time, we explicitly control for country-by-date fixed effects λ_{md}. These indicators capture any country-specific time-varying confounders that may affect average prices of fridges in a given country on a given date, and nest as well country dummies and EU-wide time-

[8] In particular, high energy and cooling efficiency as captured by the energy label are closely associated with higher quality due to requirements of advanced compressor-technology.

varying confounders.

We compute product-specific quality indices using the estimated coefficients on each characteristic's marginal contribution to the product's market price, purposefully omitting country-date and idiosyncratic variation; namely:

$$\hat{q}_j = \sum_{a=1}^{4} \hat{b}_a \kappa_{aj}. \tag{2}$$

The point estimates from eq. (1) are reported in Table A.5 in Appendix A.3 and are in line with standard consumer valuation expectations. Thus, more energy-efficient fridges or those with multiple doors and a no-frost system are associated with higher prices than counterparts without these underlying attributes. Also, well-known high-end brands tend to exhibit a significant price premium.[9] Figure A.5 in Appendix A.3 displays a histogram with of the quality index, pointing to a substantial degree of vertical differentiation in the sector.[10]

As our analysis studies the pricing of the same product in multiple destinations and entry patterns in those destinations, we focus henceforth solely on multi-market appliances. These models constitute 90% of the sample. Descriptive statistics for all products and by quality quantiles are reported in Table 1 after removing single-sales-destination devices.[11] The typical refrigerator sells about 35 units per month per country at a mean price of 682 euro. Not surprisingly, segmenting the product space by quality translates into clear price and unit-sales separation: premium quality products (i.e. those in the top quality quartile) are close to four times more expensive and generate half the monthly sales of entry-level products (bottom quality quartile). Given the limited prevalence of single-country appliances, the sample in Table 1 accurately represents the characteristics of the full sample summarized in Table A.4 in Appendix A.2.

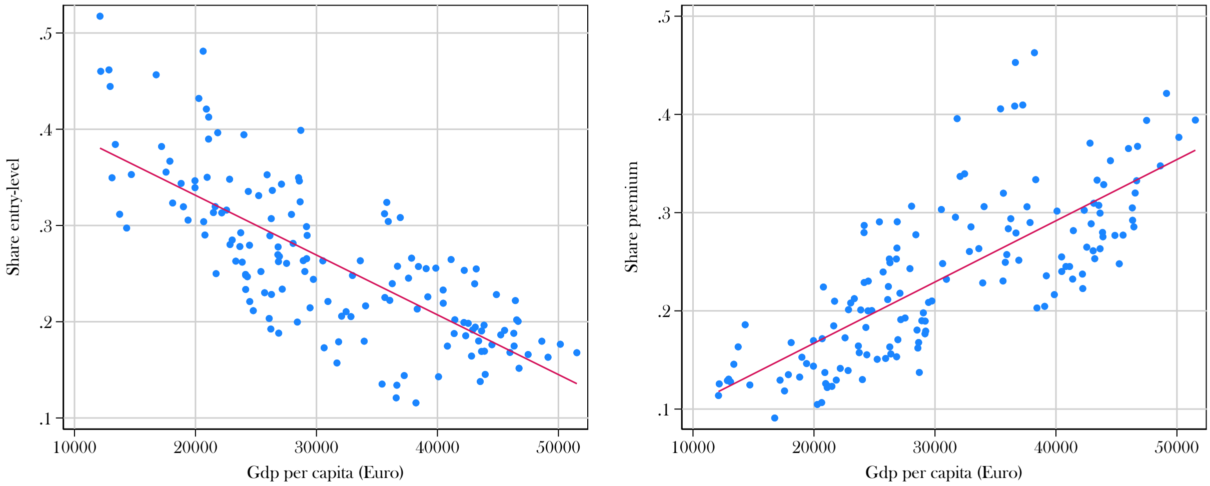
Figure 1 shows the relation between the shares of the highest quality and lowest quality products over the total number of products offered in a country and its income per head. The figure reveals a distinct pattern of quality differentiation in consumption linked to income levels. Specifically, panel (a) shows a negative correlation: as income increases, the share of entry-level products decreases. Conversely, panel (b) displays a

[9]Brand names are not explicitly reported in Table A.5 but are available from the authors upon request.

[10]The index ranges from -1.05 (a refrigerator model of brand PKM with an average price over its life-cycle of 254 Euro) to 2.31 (a refrigerator model of brand Gaggenau with an average price over its life-cycle of 6,421 Euro).

[11]There are 1,250 products sold in only one country throughout their life-cycle, 65% of which are solely marketed in Austria, while 37% are retailer-specific brands, classified as 'Tradebrands' in the data.

FIGURE 1 – Premium and Entry-level Product-shares by Income per Capita



(a) Entry-level

(b) Premium

Notes: The figure plots country-specific yearly shares of entry-level products (number of entry-level products over the total number of products per year) in (a) and of premium products (number of premium-quality products over the total number of products per year) in (b) vis-a-vis GDP per capita. Entry-level products are those in quartile one, and premium products in quartile four of the product-specific quality estimates obtained from eq. (2). The lines in both graphs are linear prediction plots. Figure A.6 in the Appendix shows that the same relationships hold with respect to the shares of unit sales from total sales for entry-level and premium products.

positive correlation, indicating that the proportion of top-quality products rises with higher income levels.

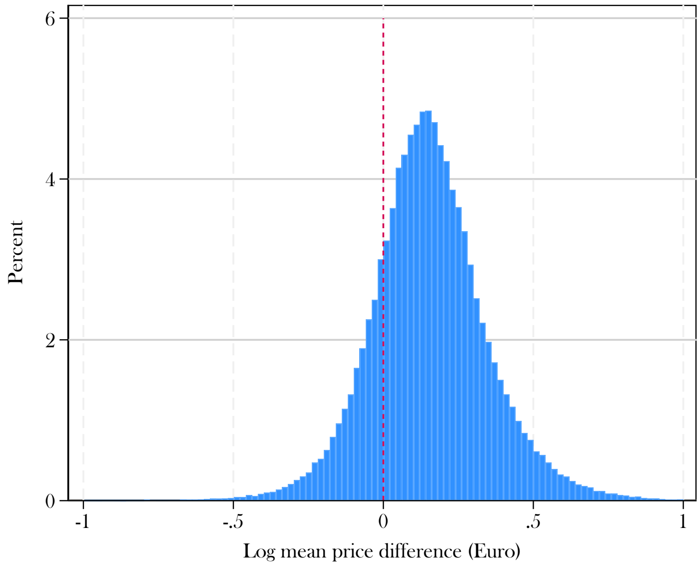
3 Pricing-to-Market along the Quality Dimension

To investigate whether the same products are priced differently in countries with different incomes, we begin with a simple exercise based on the raw data. Specifically, we focus on products sold concurrently in both below-median- and above-median-income countries, excluding the value-added tax (VAT) component from the prices. Figure 2 illustrates the *within-product within-date* log difference in mean prices between richer and poorer destination markets. The figure showcases that products are on average more expensive in higher-income countries. However, this pairwise comparison remains silent as to whether mark-ups vary with consumers' incomes, especially across layers of quality.[12]

The trade literature has long acknowledged that nonhomothetic preferences, especially in relation to quality, may imply that richer households display lower price elasticities for

[12] Simonovska (2015), for example, finds that for identical items supplied by a large and global clothing retailer, a substantial fraction of cross-country price variation is driven by pricing-to-market.

FIGURE 2 – Within-product Comparison of Pricing in Below and Above-median GDP p.c. Sales Destinations



Notes: The sample includes only products sold simultaneously in both above- and below-median GDP per capita countries. The share (48%) of products only sold either in above-median, or below-median markets are excluded from this analysis. Prices are presented net of country-specific standard VAT rates. The plot shows the difference in a product's log mean prices in countries above- and below-median income within the same month-year.

higher-quality varieties. If this is indeed the case, mark-ups charged on higher-quality models would turn out to be relatively higher in richer markets.[13] We next provide evidence that quality-pricing-to-market is present in the cold-appliances sector. Specifically, we employ the following regression equation:

$$\ln \text{Price}_{jmd} \quad = \quad \alpha_{jd} + \delta_{jm} + \beta_1 \cdot \ln \text{Income}_{md(t)} + \beta_2 \cdot (\ln \text{Income}_{md(t)} \times \hat{q}_j) + \epsilon_{jmd}, \tag{3}$$

where $\ln \text{Income}_{md(t)}$ is the log of per capita income in country m in year t and $\hat{q}_j$ is the product-specific estimate of quality as per eq. (2). A positive β_2, the coefficient on the interaction term between per-capita income and the quality index would be indicative of higher-quality goods commanding relatively higher mark-ups in richer markets. The inclusion of product-date specific fixed effects α_{jd} ensures that price comparisons across countries occur within the same appliance j and date d, and thus that any time-varying product-specific shocks common to all markets (such as aging over the life-cycle or variation in the cost of model-specific inputs) are accounted for. Note that α_{jd} further subsume

[13] Auer et al. (2018) have found evidence of quality-pricing-to-market depending on income in the automobile industry.

brand indicators and control for all other permanent product attributes.

Variations in the prices of products across destinations at a given date will also reflect certain country-specific costs. In particular, as each products is usually manufactured in a single location and shipped to multiple destinations, part of the cross-country price variation for identical goods will stem from differences in transportation and handling costs. Additionally, because the data comprises sales in brick-and-mortar stores, price differences will also result from destination-specific rents, local wages, and distribution costs. To disentangle the effect of discretionary mark-ups from that of the above costs, we rely on further controls: eq. (3) includes product-by-country fixed effects δ_{jm}, which will capture the impact of factors such as shipping costs and different currencies, as well as absorb country dummies, and thus average differences in labor, rental, and other operational costs across countries. Furthermore, in our most demanding specification, we incorporate country-by-date fixed effects (θ_{md}), which control not only for the levels of any country-specific confounders but also for any changes in these over time.

Table 2 reports results from the estimation of eq. (3). Column (1) essentially replicates the raw exercise shown in Figure 2 yielding a statistically significant price elasticity with respect to income per head of 0.25%. The estimation exploits cross-country variation in prices within the same product within the same date. This estimate, however, ceases to be significant once product-by-country indicators enter the specification in Column (2). Assuming that a product is manufactured in only one location throughout its life-cycle, these fixed effects will tend to absorb the impact on final prices paid by consumers of product-specific across-destination variation in transport costs. They will also absorb any non-time-varying differences between countries such as geographical features, alongside long-standing disparities in consumer preferences, infrastructure, taxation, etc.

Specifications (3)-(7) include our main coefficient of interest: an interaction term between the quality estimate and log GDP per capita to capture heterogeneity in pricing-to-market. The sign and significance of the interaction term suggest that the impact of per-capita income on mark-ups is increasing in product quality. In terms of its quantitative impact, the estimate of β_2 in (3) indicates that relative to a product of median quality ($q = 0.118$), the effect of income on a product in the 90th percentile of quality is a 0.120 higher log-price ($(0.799 - 0.118) * 0.177$).[14]

In columns (4)-(7) we sequentially add a set of additional covariates. Specifically, based on the GfK data, we calculate brand-specific market shares (MS Brand) as the

[14] Table A.6 in Appendix A.3, which splits the sample into products of above- and below-median quality, reveals that mark-ups vary strongly with quality across income per capita only for higher quality appliances whereas this effect is not significant for below-median quality fridges.

Table 2 – Pricing-to-market: Income and Quality

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ln Income	0.251***	-0.021	-0.042	-0.061	-0.060	-0.059	
	(0.052)	(0.060)	(0.063)	(0.070)	(0.071)	(0.071)	
ln Income$\times\hat{q}_j$			0.177**	0.191***	0.191***	0.184**	0.192***
			(0.065)	(0.067)	(0.067)	(0.068)	(0.047)
ln Pop				-0.232	-0.230	-0.238	
				(0.377)	(0.372)	(0.370)	
ln MS Brand				-0.006***	-0.006***		
				(0.002)	(0.002)		
ln Retail				0.009	0.009	0.009	
				(0.020)	(0.020)	(0.020)	
ln Energy				0.060	0.060	0.061	
				(0.043)	(0.043)	(0.043)	
HHI					-0.018	-0.026	
					(0.128)	(0.118)	
HHI$\times\hat{q}_j$						0.189*	-0.005
						(0.109)	(0.070)
δ_{jm}		Yes	Yes	Yes	Yes	Yes	Yes
θ_{md}							Yes
N	785,232	778,341	778,341	778,341	778,341	778,341	778,341

Notes: The table shows results from the estimation of eq. (3). All specifications include product-by-date fixed effects, not reported. Specifications (2)-(7) control for product-by-country indicators, while (7) further incorporates country-by-date fixed effects. Ln Pop, ln Retail, ln MS Brand, and ln Energy are the logarithms of population, retail turnover index, brand market share, and bi-annual household energy prices. HHI is the Herfindahl-Hirschman index of market concentration. See Table A.3 for a detailed description of these variables and summary statistics. Standard errors are robust, and two-way clustered by product and by country throughout. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

ratio of a brand's unit sales in a given country on a given date to its total sales within the same country-date pair. We also use these shares to construct a Herfindahl-Hirschman index (HHI) of market concentration for each country-date. Table A.3 in the Appendix shows the mean of the HHI is 0.14 pointing to an overall relatively competitive industry, but with considerable country variation. We further supplement the GfK data with a set of country-specific variables at date, bi-annual, or yearly frequency, namely population, index of retail turnover, energy prices, and others. (These variables are defined and summarized in Table A.3.)

Column (4) adds (log) population to control for market size, country-date specific brand market shares, and retail index to capture any underlying brand and retail expenditure evolution over time. It also includes bi-annual household energy prices as in economies with relatively costly energy better quality refrigerators may be purchased due to their higher energy efficiency. Amongst these additional covariates, only the brand market share is found to be significantly associated with prices, specifically a 1% share increase (brand familiarity) is found to reduce prices by about 0.01%. The point estimate of the interaction term increases to 0.191, remaining also highly significant. In (5)-(6) we additionally control for market concentration through the HHI and its interaction with the quality estimates. The estimates and significance of β_2 in (5) and (6) remain virtually unaffected. Interestingly, the specification in (6) also suggests that higher quality products are also more expensive in less competitive markets.[15]

Lastly, in column (7) we add a a full set of country-date fixed effects. These fixed effects nest all previous controls except for the interaction terms. Remarkably, the point estimate of β_2 remains essentially intact and so does its level of statistical significance. That is, even when controlling for all time-varying factors within each country, fridge models of higher quality are found to command relatively higher prices when sold to consumers in richer markets. Given the inclusion of extensive sets of fixed effects, effectively absorbing different sources of price variability across destinations and time, we interpret the results in columns (3)-(7) as robust evidence of variable mark-ups along the quality dimension.

4 Market Entry and Quality

Section 3 shows supporting evidence for the presence of nonhomothetic preferences along the quality distribution enabling firms' engagement in pricing-to-market. In this section, we explore a series of qualitative dynamic patterns in terms of the market-entry order of

[15] The estimate of the coefficient associated with this interaction term loses significance, however, once we control for country-date fixed effects in specification (7).

products belonging to different quality layers.

4.1 Country-specific, EU-wide dates of entry and product-specific destination sequences

Within each month-year combination (date), each country and the EU market as a whole are composed of various cohorts of products launched on the current or earlier dates.[16] We differentiate between two types of entry dates per product: a country-specific (local) and an EU-wide (global) date. Let y_{jmd} be the unit sales of product j in country m on date d, and $y_{jm'd}$ be the sales of the same product in country m'. Then the country-specific dates of entry of j in m and in m' are given by the conditional minima: $\widetilde{d}_{jm} = \min\{d|y_{jmd} > 0\}$ and $\widetilde{d}_{jm'} = \min\{d|y_{jm'd} > 0\}$, respectively. In other words, we consider the country-specific date of entry in a given market to be the first date when the sales of j in this market are positive. The minimum of the set of all of j's country-specific entry dates, $\widetilde{d}_j = \min\{\widetilde{d}_{jm}, \widetilde{d}_{jm'}, ..., \widetilde{d}_{jz}\}$, then yields the EU-wide entry date of product j. The ordered sequence of country-specific entry dates $\{\widetilde{d}_{jm} \leq \widetilde{d}_{jm'} \leq, ..., \leq \widetilde{d}_{jz}\}$ maps directly into product j's destination entry-sequence, with country m being the first market of entry, country m' – the second (provided strict inequality applies), and so on. The weak inequality allows for market ties, i.e., products may be launched in different destinations on the same date.

Correctly identifying the timing and location of EU entry depends crucially on country coverage and necessitates that we trace a product from the very beginning of its life cycle. For this reason, even though throughout 2009-2017 we observe the sales of products introduced before 2009, it is not possible to fully recover their market sequences. Likewise, as in 2014 country coverage is substantially smaller, a similar identification problem arises concerning both the first market and the market sequence overall of new entrants in 2014-2017.[17] Consequently, in some of the follow-up analyses, we keep only products whose global first dates lie between 2009-2013 and observe their sales (if any) until 2017. This sample restriction does not preclude product-country sequences from being incomplete in later years, but it does guarantee that the EU-wide *first* market is correctly observed.[18] The distribution of the country-specific-first dates is depicted in

[16]Henceforth, by 'cohort' we will designate a group of products entering a given market on the same *date*, and by 'annual-cohort' all (maximum twelve) cohorts launched in the same *year*.

[17]See for example Plot (a) of Figure A.1 in the Appendix, which visualizes aggregate sales per annual-cohort over time. Sales' evolution is cut off for annual-cohort 2008 in the beginning, while for annual-cohorts 2014 onward (not depicted), developments in aggregate sales will not reflect all markets. In both instances, the first EU market of entry cannot be identified.

[18]Given an average EU-wide life-cycle of 4-5 years, country sequences will be close to exhaustive for

Figure A.4 in the Appendix.

4.2 Order of market entry vis-a-vis income across quality levels

In this subsection, we document differences in the entry dynamics and life-cycle properties of products sold in different income segments of the EU market using the product-specific market sequences defined in Section 4.1. Specifically, splitting the set of countries in the sample according to a year-specific median GDP per capita, we explore the characteristics of products sold only in higher income (above-median), only in lower income (below-median), and simultaneously in both types of countries throughout their entire life cycles. As discussed earlier, to maximize country coverage, we focus on annual cohorts 2009-2013. We also exclude products sold exclusively in one country throughout their life cycles, as these presumably serve local markets and are likely to be retailer-specific. Of the resulting sample, 29% of products are found to be marketed only in countries with above-median GDP per capita, 14% only in below-median-income countries, and the rest in a mixture of both.

Table 3 reports how product features differ based on the above sample split focusing on quality, prices, and unit sales in the first panel of the table. A clear pattern of quality (and price) segmentation is observed by sales-destination income: the mean quality estimate of products sold only in richer destinations is substantially higher than those exclusively offered in poorer economies, a finding already hinted at in Figure 1. The same pattern repeats within the group of products marketed in a mixture of higher and lower income locations: quality (and average prices) decrease with the share of below-median income destinations in which these appliances are present.

Market sequences are shortest for products sold exclusively in high-income countries (2.6 markets on average), and longest for products sold throughout the entire EU (8.4 markets on average). In terms of life-cycle duration, products sold exclusively in either richer or poorer markets are available for an average of 3.3 years, compared to 4.6 years for products present in both market types. The sequential entry lag, defined as the monthly difference between the first dates in sequential markets is 7.2 months on average for above-median-income-country products, whereas it takes 3.6 months on average for mixed-income-country products to enter new destinations. Most products in the sample reach more than half of their intended geographical coverage within their first year.[19]

We next turn to analyzing products' destination-entry patterns. Is the order of market

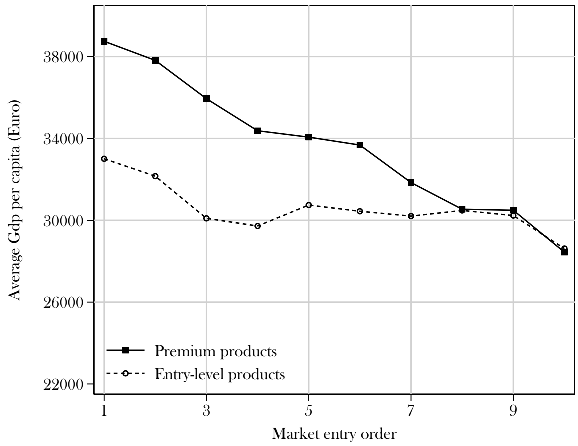
the 2009-2010 annual product cohorts, and least complete for the 2013 product cohort, whose second markets may be misidentified as the data's country coverage reduces to eight markets in 2014. Section A.2.1 in the Appendix discusses in some detail two methods that measure with reasonable accuracy the

Table 3 – Market Segmentation by Income

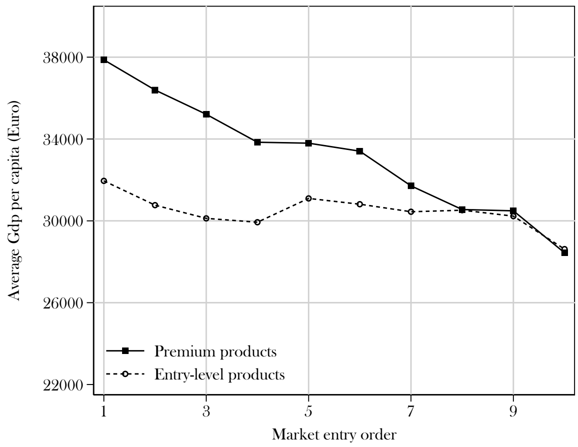
	Products sold throughout life-cycle in countries with incomes						
	A. Only above median	B. Only below median	C. Both				
				of which mean share below-median			
				0.1	0.4	0.6	0.9
Quality	0.144	-0.034	0.219	0.345	0.257	0.148	0.042
	(0.466)	(0.386)	(0.446)	(0.472)	(0.429)	(0.423)	(0.363)
Price (Euro)	634.0	438.5	659.2	821.2	701.6	659.4	526.2
	(454.2)	(274.7)	(517.4)	(608.2)	(441.6)	(579.7)	(375.2)
Units sold	42.4	32.1	36.7	34.5	36.2	40.6	38.6
	(103.0)	(97.4)	(112.1)	(119.7)	(111.8)	(124.7)	(102.6)
Destinations	2.6	3.4	8.4	7.1	8.8	9.7	8.9
	(1.5)	(2.1)	(4.3)	(3.1)	(4.9)	(5.1)	(3.7)
Destin. first year	1.8	2.3	5.3	4.2	5.5	6.1	6.1
	(1.1)	(1.7)	(3.7)	(2.7)	(4.0)	(4.4)	(3.7)
Seq. entry lag	7.2	6.2	3.6	3.6	3.0	2.7	2.8
	(8.7)	(8.0)	(5.3)	(4.4)	(3.9)	(3.5)	(3.6)
Life cycle (months)	40.2	41.9	55.2	54.7	55.4	55.7	55.5
	(21.0)	(17.5)	(16.4)	(16.5)	(17.3)	(16.5)	(15.3)
No. of products	1,880	832	2,500	872	600	465	563

Notes: The table provides descriptive statistics on the quality index, prices (in Euro), and units sold for products sold in: only above-median GDP per capita countries, only below-median GDP per capita countries, or a mixture of both. Median income is year-specific and includes all 24 countries in the estimation sample until 2013. Austria, Belgium, Denmark, Finland, France, Germany, Italy, the Netherlands, Sweden and the United Kingdom are always classified as above-median income countries; Croatia, the Czech Republic, Estonia, Greece, Hungary, Latvia, Lithuania, Poland, Portugal, Romania, Serbia, Slovakia, Slovenia are always classified as below-median income countries; Spain is above-median in some years, and classified as below-median in others. The sample excludes products sold in only one country and is based on annual cohorts 2009-2013 whose sales can be observed until 2017. For products sold in both below- and above-median-income countries, the last four columns report statistics by quantiles of the share of below-median-income countries in the total number of countries where these products are sold. Thus, quantile one has a 10% mean share of below-median-income countries, while the analogous for quantile four is 90% on average. Destinations are the average number of countries in which products are sold throughout their life-cycle while destinations in the first year are the average number of destinations a product enters in the first year it appears in the EU market anywhere. Sequential entry lag is the average delay measured in months between sequential market entry, i.e., between first and second market(s), second and third market(s), etc. The life cycle is the monthly difference between a product's EU-wide last date and its EU-wide first date.

Figure 3 – Order of Market Entry vis-a-vis Income



(a) all products



(b) products sold in above- and below -median income countries

Notes: Plot (a) depicts average income per capita by all products' order of market entry (first, second, third, etc. market) and quality (premium products (quantile four) shown as solid line, and entry-level products (quantile one) – as dashed line) for all products. Plot (b) visualizes the same relationship but only for products sold both in above-median and below-median income countries. The order of market entry is determined as per Table A.9. Quality quantiles are based on the full set of products.

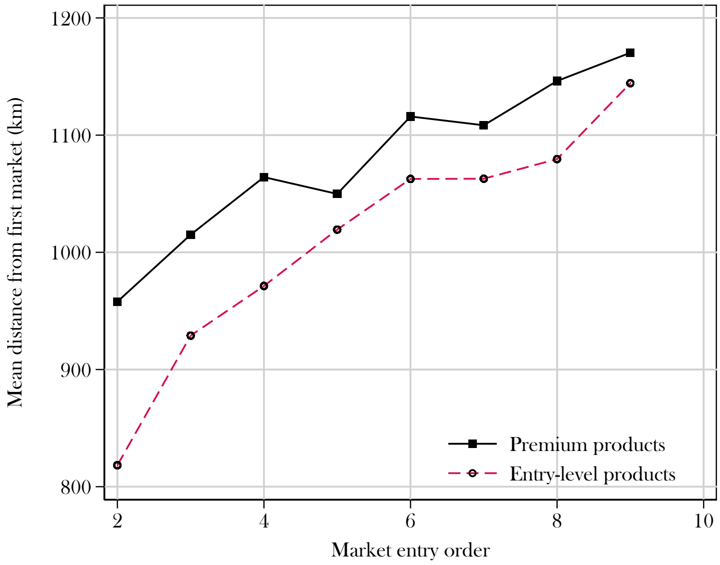
entry associated with income and does this pattern vary with quality? Figure 3 displays the average income per head by order of market entry for all products (Plot (a)) and for those in Panel C of Table 3 (Plot (b)). The plots show that market order is negatively correlated with income, with first (earlier) destinations generally associated with higher per-capita incomes than follow-up markets. Interestingly, this relationship is particularly pronounced for models in the top quartile (premium products), while it appears rather weak (or barely existent) for those in the bottom quartile (entry-level products). Figure A.7 in the Appendix replicates Figure 3 using only products sold in at least five destinations over their life cycle, confirming that the dynamic patterns showed in Figure 3 do not seem to be driven by selection with different product-composition per market-order.

If income is a strong predictor of (early) market entry choices for high-quality products, but much less so for low-quality products, as seemingly indicated by Figure 3, one could then expect that geographic proximity between sequential markets should matter proportionately less for the former than for the latter. Some preliminary supportive evidence of this hypothesis is displayed in Figure 4. This figure plots the average distance in km of products' follow-up markets relative to their first destinations of entry for premium

average life-cycle of products both within a single market and across the EU overall.

[19]Plot (c) in Figure A.1 in the Appendix which visualizes sales and sales destinations by entry cohort conveys the same finding.

FIGURE 4 – Distance of further markets relative to first destination of entry



Notes: The figure plots the average distance (in km) of products' sequential markets of entry (second, third, etc. markets) relative to their first destination market differentiating between premium and entry-level products. Bilateral distances between countries are retrieved from the CEPII data set, using intra-country agglomeration weighted measures.

products (solid line) and entry-level products (dashed line). Two points deserve a mention. First, premium products' second and further sales destinations are on average more distant from their first markets than entry-level products. Second, irrespective of quality, proximity still matters: markets nearer to the first destinations are selected earlier in the life cycle than more remote markets.

Regarding the first markets of entry, the stylized facts conveyed in Table 3, and in Figures 3 and 4, may arguably be capturing a home-market effect if high-quality goods are more likely to be produced in higher-income countries, which in turn also serve as their first markets of sale. Similarly, distance may matter less for them simply because such products are manufactured closer to a higher-income customer base. Using data on the countries of manufacture of refrigerator models taken from Jaimovich, Madzharova, and Merella (2023) and merging it with our sample, we identify the production locations of 854 products.[20] In this subsample, only 15.2% of products' first markets coincide with their manufacturing locations pointing to a weak link between the determinants of production and the choice of first-sales destinations.

[20] Jaimovich et al. (2023) use the same data to study how the home-market effect linked to nonhomothetic preferences affects patterns of quality specialization.

5 A Model with Variable Mark-ups and Gravity

This section introduces a tractable model with nonhomothetic preferences along the quality dimension building on the Almost Ideal Demand System proposed by Deaton and Muellbauer (1980). The framework generates demand functions with variable price elasticities linked to income elasticities, leading firms to optimally adjust mark-ups based on consumers' incomes. The first subsection analyzes this mechanism within a static, single-country setup. We then extend the framework to a dynamic multi-country model, where proximity to previously served markets facilitates market expansions.[21]

5.1 Static model in a single country

We consider an economy with a continuum of households with mass H. Each household has the same nominal income Y. There is a continuum of varieties of a differentiated good with mass N. Household preferences are summarized by the indirect utility function:

$$\left(\frac{Y}{a(\mathbf{p})}\right)^{\frac{1}{b(\mathbf{p})}},$$

where $a(\mathbf{p})$ and $b(\mathbf{p})$ are price aggregators given by:

$$a(\mathbf{p}) \equiv \exp\left(\int_{\mathcal{J}} \alpha_j \ln p_j \, dj + \frac{1}{2}\int_{\mathcal{J}} \left[\gamma_{jj} (\ln p_j)^2 + \int_{k\neq j} \gamma_{jk} \ln p_j \ln p_k \, dk\right] dj\right),$$
$$b(\mathbf{p}) \equiv \exp\left(\int_{\mathcal{J}} \beta_j \ln p_j \, dj\right),$$

where $j, k \in \mathcal{J}$ index varieties. As shown by Deaton and Muellbauer (1980), to abide by consumer rationality, $a(\mathbf{p})$ and $b(\mathbf{p})$ must satisfy the following restrictions:

$$i)\ \int_{\mathcal{J}} \alpha_j \, dj = 1,\ \alpha_j > 0; \quad ii)\ \int_{\mathcal{J}} \beta_j \, dj = 0; \quad iii)\ \gamma_{jj} + \int_{k\neq j} \gamma_{jk} \, dk = 0, \quad \gamma_{jk} = \gamma_{kj}. \quad (4)$$

We let henceforth $\Upsilon \equiv H \times Y$ denote the level of nominal GDP in the economy. Using Roy's identity, we can derive the market demand for variety j, which (provided it

[21] The AIDS structure has been recently used by Fajgelbaum and Khandelwal (2016) to model nonhomothetic demand and capture its implications on gains from trade across consumers with different incomes. Differently from our paper, these authors do not address quality differentiation. To the best of our knowledge, this is the first paper in the international trade literature to rely on the AIDS structure to account for nonhomotheticities along the vertical dimension.

is strictly positive) will be given by:

$$D_j = \left(\alpha_j + \gamma_{jj} \ln p_j + \int_{k \neq j} \gamma_{jk} \ln p_k \, dk + \beta_j \ln \left(\frac{Y}{a(\mathbf{p})} \right) \right) \frac{\Upsilon}{p_j}. \tag{5}$$

Following Deaton and Muellbauer (1980) $a(\mathbf{p})$ can be interpreted as a price index at the subsistence level.[22] We can thus define real income as:

$$y \equiv \frac{Y}{a(\mathbf{p})}.$$

We let $y \geq 1$ always hold, meaning that households' incomes lie above the subsistence level. In this context, β_j will govern the income elasticity of demand for variety j. Varieties whose $\beta_j > 0$ exhibit an income elasticity above one, and will thus account for larger expenditure shares in richer households. Conversely, the income elasticity of varieties with $\beta_j < 0$ will always remain below one, implying that households' expenditure shares in those varieties decline with y.

Each variety j is characterized by a level of quality q_j. For simplicity, we assume that there are only two levels of quality, $q = l, h$, with $l < h$, and let N_q be the mass of the set $\mathcal{J}_q$ of varieties of quality q, with $N_l + N_h = N$. The trade literature has consistently shown empirically that higher-quality varieties are associated with greater income elasticity of demand (e.g., Hallak, 2006; Verhoogen, 2008; Khandelwal, 2010). Accordingly, we let the parameter governing the income elasticity of variety j (β_j) be positively tied to its level of quality relative to the average level of quality in the market.

Assumption 1 (nonhomotheticities)**.** *Let* $\Gamma \equiv (N_l l + N_h h)/N$*. Then:*

$$\beta_j = q_j - \Gamma. \tag{6}$$

Assumption 1 implies that varieties whose quality is greater than the average quality in the market ($q_j > \Gamma$) have income elasticity above one (hence, they are luxuries). Varieties with $q_j < \Gamma$ conversely display income elasticity below one.[23] Without any loss of generality, we will henceforth let $l > 0$ and $h - l = 1$.

We next impose some additional structure on the demand functions in (5) in terms of the patterns of demand cross-substitution across varieties.

[22] $a(\mathbf{p})$ becomes in fact the exact price index at any income level when $\beta_j = 0$ for all j.

[23] Note that Assumption 1 ensures that the condition $\int_{\mathcal{J}} \beta_j \, dj = 0$ in (4) will always be satisfied in our model.

Assumption 2 (cross-substitution)**.** *Let $j \in \mathcal{J}_q$ and $k \in \mathcal{J}_{q'}$, with $q, q' = l, h$. Then:*

$$i)\ \gamma_{jk} = \frac{1}{N_q}, \quad \textit{for all } k \neq j \textit{ such that } q' = q;$$

$$ii)\ \gamma_{jk} = 0, \quad \textit{for all } k \neq j \textit{ such that } q' \neq q.$$

Assumption 2.i entails that the degree of cross-substitution across all pairs of varieties within the same quality level is strictly positive and inversely related to the total mass of these varieties present in the market. Assumption 2.ii precludes cross-substitution across varieties belonging to different layers of quality.[24]

To comply with the parametric restrictions (i) and (iii) in (4), we lastly impose specific structure on α_j and on γ_{jj}. For the former, we will henceforth assume that $\alpha_j = 1/N$ for all $j \in \mathcal{J}$. With regards to γ_{jj}, Assumption 2 entails that $\gamma_{jj} = -\int_{k \in \mathcal{J}_q, k \neq j} N_q^{-1} dk$ must hold for any variety $j \in \mathcal{J}_q$. This, in turn, boils down to $\gamma_{jj} = -1$ for every variety j regardless of its quality level.

Assumption 2 coupled with the parametric restrictions in (4) allows us to rewrite the demand for variety $j \in \mathcal{J}_q$ as:

$$D_j = \left[-\ln p_j + (q_j - \Gamma) \ln y + \frac{1}{N} + \frac{1}{N_q} \int_{k \in \mathcal{J}_q, k \neq j} \ln p_k \, dk \right] \frac{\Upsilon}{p_j},$$

from which we can obtain the price elasticity of D_j:

$$\varepsilon_j \equiv -\frac{\partial \ln D_j}{\partial \ln p_j} = 1 + \frac{1}{-\ln p_j + (q_j - \Gamma) \ln y + \frac{1}{N} + \frac{1}{N_q} \int_{k \neq j} \ln p_k \, dk}. \tag{7}$$

Note that $D_j > 0$ implicitly entails $\varepsilon_j > 1$. In the optimum, the producer of variety j will thus set $p_j = [\varepsilon_j / (\varepsilon_j - 1)] c_j$, where c_j denotes the marginal cost j.[25]

We let henceforth $\mu_j \equiv \varepsilon_j / (\varepsilon_j - 1)$ denote the mark-up charged by firm j. Thus, firms optimally set $p_j = \mu_j c_j$. We assume that the marginal cost of production is constant and identical for all varieties in the same quality layer, that is, $c_j = c_q$ for all $j \in \mathcal{J}_q$. From

[24] In conjunction with Assumption 2.i, fixing cross-substitution between varieties belonging to different quality layers at zero through Assumption 2.ii is an extreme way to model the notion that consumers are more likely to substitute between similar rather than markedly different levels of quality. The assumption can be relaxed at the cost of more complicated algebra.

[25] Recall that $p_j = [\varepsilon_j / (\varepsilon_j - 1)] c_j$ is the standard mark-up over the marginal cost that obtains from solving the optimization problem: $\max_{p_j} \Pi_j = (p_j - c_j) \times D_j (p_j)$, where $\varepsilon_j \equiv -(\partial D_j / \partial p_j) \times (D_j / p_j)$.

(7), it then follows that:

$$\mu_j + \ln \mu_j = 1 + (q_j - \Gamma) \ln y + \frac{1}{N} + \frac{1}{N_q} \int_{k \neq j} \ln \mu_{k_q} \, dk. \tag{8}$$

Equilibrium mark-ups

The previous analysis has focused on the optimal behavior of a generic firm j in isolation, taking the behavior of other firms in the market as given. In equilibrium, each firm will behave optimally given the behavior of their competitors. We consider the strategy of each firm j as consisting of choosing the optimal mark-up μ_j^* given the mark-ups chosen by all other firms $\{\mu_k\}_{k \neq j}$.

Equation (8) pins down the optimal mark-up charged by firm j producing a variety of quality q given the mark-ups charged by all other firms producing varieties of the same quality level. Those firms will also choose their mark-ups optimally following an expression analogous to that in (8). The Nash equilibrium will therefore be characterized by a full set of conditions like (8) holding simultaneously for all firms. Lemma 2 in Appendix A.1 shows that, in a Nash equilibrium, μ_j^* will necessarily be symmetric for all $j \in \mathcal{J}_q$. Furthermore, the lemma shows that the Nash equilibrium solution is always unique and that all mark-ups will be strictly larger than one ($\mu_j^* > 1$ for all j).[26]

Denote henceforth by μ_q^* the equilibrium mark-ups charged on varieties of quality $q = l, h$. The following proposition demonstrates that mark-ups behave heterogeneously across different quality and income levels:

Proposition 1. *Given the mass N_q of varieties of quality $q = l, h$, where $N_l + N_h = N$, the equilibrium mark-ups are given by:*

$$\mu_l^* = 1 + \frac{1}{N} - \frac{N_h}{N} \ln y, \tag{9}$$

$$\mu_h^* = 1 + \frac{1}{N} + \frac{N_l}{N} \ln y. \tag{10}$$

The main implication of Proposition 1 is that mark-ups vary heterogeneously with the level of income y: mark-ups charged on higher-quality varieties are increasing in y, whereas the opposite holds for low-quality varieties.[27] This result stems from the interplay

[26] The underlying reason for the uniqueness of the equilibrium is that, although the model features strategic complementarities across firms, their implicit positive spillovers are never strong enough to lead to multiple fixed points in (8).

[27] To be perfectly rigorous, one can observe from (9) and (10) that μ_q^* varies with y provided that $0 < N_q < N$. This is indeed the case, as according to Assumption 1 the income elasticity of demand depends on the quality of a variety relative to the average quality in the market. As a result, when

between Assumption 1 and the fact that, within the AIDS structure, higher income elasticities translate into weaker price elasticity of demand via the term $(q - \Gamma)\ln y$ in (7). Assumption 1 ties income elasticities to the levels of quality of each variety. In turn, as clear from (7), the presence of the term $(q - \Gamma)\ln y$ means that income-elastic varieties tend to display lower price elasticity as real incomes grow. Consequently, higher-quality varieties will end up exhibiting lower price elasticity in richer markets, thereby allowing firms to charge higher mark-ups in those markets.

We can lastly compute the equilibrium profit (π_j) earned by the producer of variety j. Since $\pi_j = (p_j - c_j)\, D_j$, using (5) and (7) together with $\mu_j = \varepsilon_j/(\varepsilon_j - 1)$, we can obtain:

$$\pi_j = \frac{(\mu_j - 1)^2}{\mu_j}\Upsilon. \tag{11}$$

Notice from (11) that $\mu_j > 1$ implies that $\partial\pi_j/\partial\mu_j > 0$. Based on (9) and (10), it follows that holding the value of Υ constant, producers of high-quality varieties (resp. low-quality varieties) will tend to earn higher profits in richer markets (resp. poorer markets).

5.2 A multiple-country dynamic framework: Market entry and expansion

This subsection extends the single-market static framework to a dynamic setup with multiple markets. The goal is to allow entry of newly designed varieties and study their dynamics in terms of first market entry and subsequent expansions to additional markets. Our main interest lies in exploring whether entry dynamics vary with the quality of incoming varieties.

We assume there is an even number $M > 4$ of countries/markets in the world economy $(\mathcal{W})$, indexed by $m \in \mathcal{W}$. Half of the countries host low-income households and the other half host high-income households. We assume that the level of household income in the poorer markets is $y_p = 1$, whereas that of richer households is $y_r > 1$. Since the main intention of the model is to focus on the impact of real income differences on mark-ups and entry dynamics, we deliberately shut down any variation of aggregate market size across countries. For that reason, we further assume that nominal GDP is identical across markets; that is, we let $\Upsilon_m = \Upsilon$.[28] In addition, we normalise $\Upsilon = 1$.

Life evolves along a discrete-time infinite horizon framework $t = 1, 2, .., \infty$. In each period t, a mass $\widetilde{\rho}_q > 0$ of newly designed varieties of quality $q = l, h$ becomes available.

$N_q = N$ for some q, all varieties offered on the market are of equal quality, and demand elasticity is one for all of them.

[28] This essentially amounts to assuming that while countries may differ in their per-capita nominal income (Y_m) and population (H_m), the product $Y_m \times H_m$ is equal to Υ for all m.

For simplicity, we assume henceforth that $\widetilde{\rho}_q = \widetilde{\rho}$ for both $q = l, h$. We also assume that delaying the introduction of a newly designed model is not feasible: A new model manufactured in period t may only be marketed for the first time in t, or else it will never be introduced in any market. At the beginning of each period t, a share $\delta_q = \delta \in (0, 1)$ of the varieties of quality q already present in $t-1$ in a given market m are randomly removed forever from that market. In other words, a variety present in a market in $t-1$ faces a probability δ of exiting that market in t.

To keep the dynamic analysis relatively simple, we assume that *only one* market may be entered as the first market in any period t. Conversely, we allow producers of a variety first introduced in period t to expand their market coverage (if desired) to multiple additional countries as 'second markets' in period $t+1$. We will restrict, however, any market expansion to a single round, i.e. for a variety first introduced in period t, subsequent market entry takes place only in the following period $t+1$.[29]

5.2.1 World economy geography and entry costs

We assume that $\mathcal{W}$ exhibits a "fractal" geographic structure. Specifically, regardless of its income level (y_p or y_r), each country $m \in \mathcal{W}$ is surrounded by two *neighboring* countries, one inhabited by low-income households and the other by high-income households. In addition to its *neighboring* countries, each m also has two *nearby* (albeit non-neighboring) countries, which are again characterized by different levels of household income. We denote by m_{dp} (resp. m_{dr}) the low-income (resp. high-income) market located at a distance d from m. We normalize the distance between m and each of its neighbors to be equal to zero ($d = 0$). The two nearby markets are equidistantly located at $d = 1$.

We will refer to all remaining markets in the set $\mathcal{W}$ as *faraway* markets assumed to be positioned at a distance greater than one ($d > 1$) relative to m. Lastly, when considering any pair of countries $m, m' \in \mathcal{W}$, with $m \neq m'$, we assume that the sets $\{m_{0p}, m_{0r}\}$ and $\left\{m'_{0p}, m'_{0r}\right\}$ will have at most one element in common. Analogously the sets $\{m_{1p}, m_{1r}\}$ and $\left\{m'_{1p}, m'_{1r}\right\}$ also share at most one element.[30]

[29] Allowing multiple (simultaneous) first markets of entry or more than one round of market expansions would rapidly increase the dimensionality of the choice set faced by firms rendering the theoretical model essentially intractable. In particular, relaxing any of these two assumptions would mean that a firm's optimal plan must take into account all possible period-by-period deviations in terms of market entry sequences. A similar problem is dealt with empirically by Morales et al. (2019) who rely on moment inequalities derived using Euler's perturbation method in discrete time.

[30] At first glance, such a geographic structure appears ad-hoc, but the purpose is to build a framework in which all markets are (ex-ante) identical from the viewpoint of a generic producer of a new variety except for their incomes and exact realizations of the entry cost (see Assumptions 3 and 4 below). In particular, we wish to ensure that within the framework the producer of a new model j designed in the period t will optimally choose j's first market in t without any regard for strategic considerations about

If m is chosen as the first market of entry for a new variety j, its producer must incur an entry cost of $\phi_{jm} > 0$. We impose the following structure governing entry costs in a first market of entry:

Assumption 3 (entry cost: first market of entry)**.** *When a newly designed variety j enters market m as its first market, the producer incurs an entry cost ϕ_{jm} independently drawn from a probability distribution with cdf $F(\phi)$, which satisfies: i) $F\left(\underline{\phi}\right) = 0$, ii) $F\left(\overline{\phi}\right) = 1$, and iii) $F'(\phi) > 0$ for all $\phi \in \left[\underline{\phi}, \overline{\phi}\right)$, where $0 < \underline{\phi} < \overline{\phi}$.*

A variety j first introduced in market m in period t and not withdrawn from this market in $t+1$ may enter new destinations in $t+1$. To do so, j's producer must incur additional entry costs. The costs associated with expansion to country $m' \neq m$ will be influenced by the distance between m' and m. Specifically, we impose the following structure for the entry costs to additional markets:

Assumption 4 (entry costs: further market expansions)**.** *Consider a variety j, first introduced in market $m \in \mathcal{W}$ in period t. If variety j enters market $m' \in \mathcal{W}$, with $m' \neq m$, in period $t+1$, its producer must incur an entry cost $\varphi_{jm'}$, where:*

i) Neighboring markets — If $m' \in \{m_{0p}, m_{0r}\}$, then $\varphi_{jm'}$ is independently drawn from a probability distribution $G_0(\varphi) = G(\varphi)$ such that $G(0) = 0$, $G(\infty) = 1$, and $G'(\varphi) > 0$ for all $\varphi \geq 0$;

ii) Nearby (non-neighboring) markets — If $m' \in \{m_{1p}, m_{1r}\}$, then $\varphi_{jm'}$ is independently drawn from a probability distribution $G_1(\varphi) = (G(\varphi))^{\lambda}$, where $\lambda > 1$;

iii) Faraway markets — If $m' \notin \{m_{0p}, m_{0r}\}$ and $m' \notin \{m_{1p}, m_{1r}\}$, then $\varphi_{jm'} = \infty$.

Assumption 4 implies that expanding market coverage of variety j beyond m involves additional entry costs. As $G_0(\varphi)$ first-order stochastically dominates $G_1(\varphi)$, these costs are likely smaller in neighboring rather than in nearby markets. Note also that Assumption 4.iii means that it will never be profitable to enter a faraway market.[31]

For the sake of tractability, in what follows we let the "exit rate" δ be small enough to ensure that entry in some market m will always be profitable for the producer of a newly designed variety.[32] As a result, in each period t the mass $\widetilde{\rho}$ of newly designed varieties of each quality level will always enter some market within the set $\mathcal{W}$ for the first time. To

future market expansions in $t+1$.

[31]Assumption 4.iii is posed to simplify the model's dynamic behavior regarding market expansions.

[32]This assumption is posed essentially to ensure that the present value of the flow of expected profit upon entry of a newly designed variety in period t, given probability of survival until period $\tau \geqslant t$, namely $(1-\delta)^{\tau-t}$, will be large enough to justify entry in some market $m \in \mathcal{W}$ even when the entry costs is equal to $\overline{\phi}$. None of the main results in this section depend crucially on this assumption.

ease notation, we henceforth let $\rho \equiv \widetilde{\rho}/M$, which equals the mass of new models (of each of the two quality levels) per country in the world at each date.

5.2.2 Market entry dynamics

We focus on the equilibrium dynamics along a *steady state* characterized by a constant mass of varieties of each quality level in each market m. We further restrict the analysis to *symmetric* steady states, where the mass of varieties of a given quality level is identical across markets with the same income. We let henceforth N_{qr}^* (resp. N_{qp}^*) denote the steady-state mass of varieties of quality $q = l, h$ in a market with income y_r (resp. y_p).

Let also Λ_{qy} denote the (intertemporal) expected profit flow generated by a variety of quality q in a market with consumer income y. Given the exit rate δ, the expression in (11), and the normalization $\Upsilon = 1$, in the steady state:

$$\Lambda_{qy}^* = \frac{1}{\delta} \frac{\left(\mu_{qy}^* - 1\right)^2}{\mu_{qy}^*}, \tag{12}$$

where μ_{qy}^* denotes now the steady-state level of mark-ups charged on varieties of quality q in a market with income level $y = y_p, y_r$.

A share $1 - \delta$ amongst the newly designed varieties first introduced in period t will remain present in their first market of entry in period $t + 1$. Those surviving varieties may undergo market expansions in $t+1$. Assumption 4 indicates that any variety $j \in \mathcal{J}_q$ first introduced in m at time t may branch out into all possible combinations between the elements of $\{m_{0p}, m_{0r}, m_{1p}, m_{1r}\}$. Additionally, it may be the case that entry cost realizations are too high such that no market expansion proves profitable. Bearing in mind (12), it follows that the producer of variety j will expand its market coverage to any market $m' \neq m$ for which $\Lambda_{qy}^* > \varphi_{jm'}$.

A steady state must feature equality between the total entry of varieties of quality q (either as the first market of entry of newly designed varieties or as a second market of entry during market expansions) and the exit of existing varieties of the same quality, in every market in the world economy in any period t. Let $\mathcal{P}_{qy}$ denote the probability that a country with income y is chosen as the first market of entry for a variety of quality q, and let $\Gamma\left(\mu_{qy}^*\right) \equiv G\left(\Lambda\left(\mu_{qy}^*\right)\right) + \left(G\left(\Lambda\left(\mu_{qy}^*\right)\right)\right)^\lambda$ equal the probability of varieties of quality q expanding towards a market with income y from its neighboring and nearby countries, respectively. Then, in the steady state, the following equality must hold:

$$\mathcal{P}_{qy}\rho M + (1-\delta)\,\mathcal{P}_{qy}\rho M\,\Gamma\left(\mu_{qy}^*\right) + (1-\delta)\,\mathcal{P}_{qy'}\rho M\,\Gamma\left(\mu_{qy}^*\right) = \delta N_{qy}^*, \tag{13}$$

where $y' \neq y$. To interpret (13), note that the first term on the left-hand side equals the total mass of varieties of quality q that enter a market m with income y as first market in t, while the sum of the remaining two terms is the total mass of varieties of quality q entering market m as second market of entry in t. In particular, the second term is the mass of varieties expanding towards m from either its neighboring or nearby market having the same income levels as m (i.e. either $m_{0,y}$ or $m_{1,y}$), while the third term constitutes the mass of varieties expanding towards m from either its neighboring or nearby market having a different income level than m (i.e. either $m_{0,y'}$ or $m_{1,y'}$).

Lemma 3 in Appendix A.1 shows that, letting $\Delta_{qy} \equiv \Lambda_{qy} - \Lambda_{qy'}$ for a country with income y and $y' \neq y$, where Λ_{qy} and $\Lambda_{qy'}$ are as in (12), we get:

$$\mathcal{P}_{qy} = \Phi\left(\Delta_{qy}\right), \tag{14}$$

where $\Phi(\Delta_{qy})$ is an increasing function of Δ_{qy} and $\Phi(0) = 1/M$. That is, the probability that a market m with income level y becomes the first market of entry of a variety $j \in \mathcal{J}_q$ is increasing in the difference between j's intertemporal expected profit throughout its life cycle in m and the one corresponding to any other market with income level $y' \neq y$. Note that, should $\Lambda_{q,y} = \Lambda_{q,y'}$, then all markets (regardless of their income level) would face an identical probability $1/M$ of becoming j's first market of entry.

Restricting the analysis to the class of symmetric steady states entails that a steady state will be determined by four dynamic equilibrium conditions stemming from the country-level dynamic conditions (13). Each of the four conditions applies to a specific quality-income combination — see equations (A.31)-(A.34) in the Appendix. From these, we derive the following result.

Proposition 2. *There exists a unique steady state (amongst the class of symmetric steady states). In the steady state, the mark-ups charged on high-quality varieties in rich markets are higher than those charged in poorer markets, while the opposite is true for low-quality varieties. That is, in the steady state,* $\mu^*_{hr} > \mu^*_{hp}$ *and* $\mu^*_{lp} > \mu^*_{lr}$. *Furthermore, the steady-state features identical mark-ups for high- and low-quality varieties in poorer markets, which in turn implies that* $\mu^*_{hr} > \mu^*_{hp} = \mu^*_{lp} > \mu^*_{lr}$.

In the steady state, mark-ups will vary across quality levels depending on the income of the market where varieties are sold. Furthermore, the variation of mark-ups across quality-income combinations is non-monotonic: Along the steady state, high quality varieties command greater mark-ups in richer markets relative to those charged on varieties of low quality, but the opposite qualitative pattern is observed for varieties of low quality.

The gap in mark-ups across high- and low-quality varieties does, however, depend on income and it is greater in richer than in poorer markets; that is, $\mu^*_{hr} - \mu^*_{lr} > \mu^*_{hp} - \mu^*_{lp}$.

Bearing in mind the expression in (14) and the definition of Λ^*_{qy} in (12), the probability of becoming the first market of entry of a variety of quality q is directly tied to the gap in expected profits in a given market relative to other markets. In addition, from (13) in conjunction with (14) it follows that within each of the two quality layers, the left-hand side of (13) is found to be greater for markets commanding higher mark-ups along the steady state. These two observations lead to the following corollary:

Corollary 1. *The fact that the steady state features $\mu^*_{hr} > \mu^*_{hp}$ and $\mu^*_{lp} > \mu^*_{lr}$ implies that:*

1. *The proportion of newly designed high-quality (resp. low-quality) varieties entering a richer market as their first market is greater (resp. smaller) than the proportion entering a poorer market as their first market.*

2. *The mass of varieties of high quality (resp. low quality) offered in a richer market is greater (resp. smaller) than the mass of those varieties offered in a poorer market. That is, along the steady state, $N^*_{hr} > N^*_{hp}$ and $N^*_{lr} < N^*_{lp}$.*

5.2.3 Geographic patterns of market expansion

When market expansions are allowed, the specifics of the geographic structure of the model will play a role in subsequent market choices. In general, one implication of Assumption 4 is that expansions to neighboring markets tend to be more profitable than non-neighboring ones. Yet, the frictions brought about by entry costs will be heterogeneous across quality layers and, in particular, will crucially depend on how mark-ups for different qualities respond to income.

To characterize the geographic patterns of market expansions, let $S_q \in [0, 1]$ denote the ratio of the number of market expansions to non-neighboring markets over the total number of market expansions of varieties of quality $q = l, h$. Bearing in mind Assumption 4 coupled with (12), the following result obtains.

Lemma 1. *Along the steady state, the ratio of market expansions to non-neighboring markets relative to total market expansions in the world economy $\mathcal{W}$ is given by:*

$$S^*_q = \frac{\big(G(\Lambda^*_{qr})\big)^{\lambda} + \big(G(\Lambda^*_{qp})\big)^{\lambda}}{G(\Lambda^*_{qr}) + G(\Lambda^*_{qp}) + \big(G(\Lambda^*_{qr})\big)^{\lambda} + \big(G(\Lambda^*_{qp})\big)^{\lambda}}, \quad \textit{for } q = l, h. \tag{15}$$

From (15) and since $\lambda > 1$, one may observe that $S^*_q < 1 - S^*_q$ for both $q = l$ and $q = h$. Put another way, irrespective of the level of quality, market expansions to non-neighboring markets are always less likely than to neighboring markets. Yet, in the model, the dampening effect of geographic distance on the likelihood of market expansions is stronger for lower-quality varieties than for higher-quality ones.

Proposition 3. *The share of market expansions that take place in non-neighboring markets is strictly larger for higher-quality varieties than for lower-quality ones. That is, bearing in mind (15):* $S_h^* > S_l^*$.

Proposition 3 formally characterizes the heterogeneous impact of geographic distance on market expansion choices along the quality dimension. In our model, two main factors guide these choices for an existing variety. One is geographic proximity, which tends to lower the entry cost faced by producers. The other factor is the impact of households' incomes on relative demand and equilibrium mark-ups for different qualities through the AIDS nonhomothetic structure. The result in Proposition 3 may be interpreted as stating that the impact of geographic distance on market expansion choices becomes relatively weaker for higher-quality varieties. The intuition behind the result lies in the fact that mark-ups on higher-quality varieties increase with income and are larger than those on lower-quality varieties. As a result, geographic proximity tends to have a smaller relative impact on market expansions for higher-quality varieties, especially when producers plan to expand into other wealthy markets.

6 Empirical Analysis I: First Market of Entry

One key result of the dynamic model presented in Section 5.2 is that richer markets are more likely to become the first market of entry for higher-quality varieties (Corollary 1). While the theoretical model has imposed several simplifying assumptions to describe the entry dynamics tractably, the result linking consumers' incomes and first-market-of-entry choices for different qualities essentially stems from the presence of variable mark-ups. More precisely, the nonhomothetic framework yields demand functions where the price elasticity of higher-quality varieties is relatively more sensitive to variations in income, and hence such varieties can command greater mark-ups in richer markets. Within our dynamic framework, this in turn leads to faster entry of new varieties of high quality in richer markets.

In this section, we aim to test whether this dynamic prediction is supported in the data. To that end, we exploit the fact that we follow the entire life cycles of the same fridge models across 24 European markets for annual cohorts 2009-2013.

6.1 Market entry: Empirical framework

Collapsing the longitudinal data into a cross-section of products alongside the first dates in which they enter each of their respective markets yields product-specific sequences of countries in which products are sold throughout their life cycles. These sequences can be interpreted as 'rankings' of firms' preferred locations of entry for their products. The rank order is determined by a differential *timing-of-entry* in each market (including the possibility of ties), with an *earlier* presence in a particular location indicating a more desirable choice out of a set of available alternatives.

Does the presence of nonhomothetic preferences allowing premium products to charge higher mark-ups in high-income countries translate into earlier market entry of those products in richer destinations as previewed in Figure 3? To assess the dynamics of entry patterns empirically, we let each product j face a choice set Ω_j of size C_j consisting of countries (alternatives) where it can be launched. Given the dataset's coverage, we consider that producers must decide the timing of entry into (potentially) all 24 European markets present in the dataset, such that $\Omega_j = \Omega$ and $C_j = C = 24$ for all j. Further, each country $m \in \Omega$ is associated with a number of characteristics such as income per capita, population, and level of infrastructure, amongst others. Since products' life cycles start and end at different times, the markets' characteristics will be time-dependent.[33] These can be summarized by a vector $\mathbf{X_{mt}}$. One key element of $\mathbf{X_{mt}}$ in our specific context is the level of income per capita of country m in period t, namely y_{mt}. We also consider that each product is characterized by a set of product-specific attributes, where the most relevant in our context is its level of quality q_j.

The producer of model j to be first introduced at time t will rank countries based on the revenue stream potential of product j in country m, denoted by R_{jmt}. We let R_{jmt} comprise a deterministic component (V_{jmt}) and a random component (ϵ_{jmt}):

$$R_{jmt} = V_{jmt} + \epsilon_{jmt}. \tag{16}$$

The deterministic component V_{jmt} is assumed in turn to be additive and linear in the vector $\mathbf{X_{mt}}$, in the set of product-specific attributes, and also in a set of interaction terms between $\mathbf{X_{mt}}$ and q_j. We henceforth subsume the set of product-specific attributes (herein the quality index q_j) within a set of product fixed effects denoted by ζ_j. More

[33] This means that two different products with the same country sequences will face different values of markets' characteristics if their life cycles do not coincide in time.

specifically, we let V_{jmt} be given by:

$$V_{jmt} = \Theta \cdot \mathbf{X_{mt}} + \Gamma \cdot (\mathbf{X_{mt}} * q_j) + \zeta_j + \gamma_m + \tau_t, \tag{17}$$

where in addition to the previously described terms in (17), V_{jmt} also includes a set of country dummies γ_m which absorb all country-specific constant characteristics that may influence the overall intensity of entry of new products in country m, and also a set of year indicators τ_t dated at the year of first entry of model j in a respective market.[34]

Let $r_{jmt} = 1$ denote the case in which a producer chooses m as the first market of entry for product j in period t. That means that $R_{jmt} \geq \max\{R_{jm't}, ..., R_{jCt}\}$. Assuming further that all random terms ϵ_{jmt} are independent, the probability that country m is the most preferred first market for product j would be given by:

$$\begin{aligned} Pr(r_{jmt} = 1) &= Pr(R_{jmt} \geq \max\{R_{jm't}, ..., R_{jCt}\}) \\ &= \prod_{m' \neq m}^{C} Pr(\epsilon_{jm't} < (V_{jmt} - V_{jm't}) + \epsilon_{jmt}) \end{aligned} \tag{18}$$

To operationalize the extreme value probability expression in (18) we further assume that all random terms follow an extreme value type-1 distribution.[35] Therefore, (18) boils down to the standard logit choice probability; namely:

$$Pr(r_{jmt} = 1) = \frac{e^{V_{jmt}}}{\sum_{k=1}^{C} e^{V_{jkt}}}, \tag{19}$$

where the values of each V_{jkt} in (19) are given by (17) with $k = m$. The first set of regressions we rely on to test the predictions of the model (Corollary 1) will follow this empirical framework. That is, we will use a conditional logit model (CML) to initially assess whether products of higher quality will exhibit a higher likelihood of entering richer

[34] The country dummies γ_m will, for instance, capture the fact that the specific geographic location or tariff structure of certain countries may make them more likely to become earlier markets for new models. They would also control for the fact that certain countries may host the production of larger sets of brands. The time dummies τ_t will in turn take into account the possibility of time trends (or cycles) in the overall arrival of new products.

[35] The match between ϵ_{jmt} and the random component in the theoretical model for entry costs posed by Assumption 3 is not perfect as we have assumed that entry costs are independently drawn from a bounded probability distribution with support over a subset of $\mathbb{R}_+$, while the extreme value type-1 distribution is an unbounded distribution defined over the entire set $\mathbb{R}$. Naturally, we let ϵ_{jmt} follow such a distribution so that we may obtain a known closed-form solution for the empirical expression (18) that we may then be able to bring to the data. One could, nevertheless interpret the random component ϵ_{jmt} as absorbing not only the impact of entry costs but also other sources of possible product-destination-time specific randomness in demand factors.

markets first than products of lower quality.

The above empirical strategy follows from the implicit logic of the theoretical model which yields predictions regarding the first market of entry while abstracting from the geographic structure of the world economy. Nevertheless resorting to a CLM means disregarding potentially useful information contained within the entire sequence of market entry over products' life cycles. This sequence could be rationalized as informative of a complete rank of preferences over entry choices. With ranked preference data, an econometric model for estimating the influence of specific variables associated with products/countries on the locations-of-entry process is the rank-ordered logit (ROL). While the ROL is typically applied in hypothetical situations based on survey data, in this paper we have the advantage of observing actual geographic choices by firms at the product level.[36]

The ROL cross-sectional setting implicitly assumes that decision-making regarding a product's geographical coverage is made all at once: a product's entire market sequence is chosen once and for all rather than following dynamic consecutive choices made at different points in time. This assumption does not seem unreasonable in our setup given the maturity of the sector under consideration: the top 19 brands in our dataset, which collectively account for 90% of the sample, are well-established and generally active in all 24 EU markets included in the analysis. Firms with extensive exporting experience such as those we study are likely to have good knowledge of their profitability potential in various destinations, allowing for an ex-ante profit-maximizing selection of market sequences before observing actual performance in individual locations. Importantly, the assumption of pre-determined market sequences is not inconsistent with actual staggered market entry over time, which may occur due to logistical constraints such as transportation, distribution, and other costs associated with launching a product in new markets.

Eq. (19) disregards the ranking of countries by order of preference within a product's choice set and focuses only on the selection of the most preferred alternative (first market of entry). Relying on the ROL framework increases the precision of the estimation of the parameters of the choice model, by allowing a model specification that yields the probability of the complete ranking instead of that of the most preferred alternative alone (e.g Beggs et al., 1981; Punj and Staelin, 1978).

Let now r_{jmt} denote the rank assigned to a given market m for product j to be first introduced in period t. The full ranking of countries/alternatives for product j is given

[36]In other words, the observed sequence is not merely a stated preference ordering but an actual choice. Thus, unlike survey data where respondents may be less attentive to their choices below their most preferred alternative, the scanner data reflects actual rather than hypothetical market situations.

by $r_j = (r_{jm_1t}, r_{jm_2t}, ..., r_{jm_Ct})$ if all alternatives are ranked fully in subsequent order. Consequently, the probability of such ranking is:

$$L_j = Pr(R_{jm_1} > R_{jm_2} > ... > R_{jm_C}) = \prod_{j=1}^{C} \left[\frac{e^{V_{jmt}}}{\sum_{k=1}^{C} \rho_{mk} e^{V_{jkt}}} \right], \tag{20}$$

where $\rho_{mk} = 1$ if $r_{jm} > r_{jk}$, and zero otherwise.[37]

Implementing the ROL model requires setting out a complete sequence of entries for each product in the dataset. Table A.9 in the Appendix outlines how country-sequences (ranks) are generated for each product based on the first-ever date of entry in a given market, $\tilde{d}_j$, and first dates of entry in any subsequent markets. The full entry order is associated with a value of one for the first market of entry up to n for the n-th market. The market entry order is the same for countries in which a product is introduced on the same date. The Rank variable is then constructed by assigning a value of 24 to the first market(s), 23 to the second market(s), and so on until the last market of entry, which is product-specific (i.e. market-sequences will have different depths across products). The ranking information may be incomplete at the bottom, as some countries may never end up being actually entered by a given product. The cases of countries in which product j never enters are unranked and coded as zeros. In this way, we follow the traditional set-up of the ROL by assigning higher values to more preferred alternatives.

6.2 Choice of first (earlier) market(s)

First market selection. We start with specifications focusing solely on the first market of entry. Specifically, we implement the conditional logit model (CML).[38] Note that the CLM can flexibly account for products with two or more simultaneous first markets of entry in the estimation.

Results from the CLM specifications are reported in Panel A of Table 4. These include the log of income as well as its interaction with the quality estimate as explanatory variables. The interaction term captures the presence of a differential impact of income per head on the choice of first markets at different levels of quality. The estimation implicitly controls for product fixed effects by specifying products as an identifier variable. In (2),

[37]For N number of products, the log-likelihood is given by: Ł$=\sum_{j=1}^{N} logL_j = \sum_{j=1}^{N} \sum_{m=1}^{C} e^{V_{jmt}} - \sum_{j=1}^{N} \sum_{m=1}^{C} log \left[\sum_{k=1}^{C} \rho_{mk} e^{V_{jkt}} \right]$.

[38]This framework is typically used in the location choice literature analyzing, for example, determinants of firms' decisions on where to locate plants (e.g., Head and Mayer, 2004; Devereux, Griffith, and Simpson, 2007). In the current context, the CLM can be applied similarly, but the unit of analysis is products and how they choose a first market of entry out of 24 possible destinations.

we further include country fixed effects, which capture all non-time-varying country-level confounders such as geographic proximity to main manufacturing hubs, systematic differences in tax legislation, etc. Since except for Serbia all countries in the period under consideration are part of the EU single market, there are virtually no differential technical, legal, or other barriers to trade at the country level or over time that may influence the choice of first market. Both specifications yield a positive and statistically significant estimate of the interaction term between log income and quality, the main coefficient of interest, which is in line with the predictions of the theoretical model. That is, higher-quality products are found to be significantly more likely to enter higher-income markets first relative to entry-level products. The coefficient of log income is also positive and statistically significant in both (1) and (2) but as we will show later on the effect of income alone is neither qualitatively nor quantitatively robust across specifications once we include additional controls.

The theory developed in Section 5.1 indicates that nonhomothetic *demand* preferences, allowing firms to charge higher mark-ups for their higher-quality goods in richer countries, are the primary driver behind the launch of these goods earlier in these more profitable markets. The theoretical predictions are corroborated empirically both by our pricing-to-market results in Table 2 and the CLM estimates above. One threat to our interpretation of the observed mechanism, however, is the possibility that we also capture *supply-side* factors. Specifically, if the first market of sale frequently coincides with the market in which a product is manufactured, then we cannot cleanly disentangle supply and demand determinants of the (first) market choice. To tackle this issue, we would ideally have information on the country of production for each product in the sample. As a robustness check, we can then restrict the estimation only to first markets and market sequences which are export destinations, but not manufacturing locations. Arguably, for products manufactured outside those countries present in our sample, the selection of the first market of entry (within those present in the sample) should not be driven by supply-side factors in their respective countries of manufacturing.

As noted in Section 4.2, for a fraction of the estimation sample we do observe refrigerators' country of manufacture borrowed from Jaimovich et al. (2023). This allows us to conduct the above exercise on a reduced sample that excludes the possibility that the country of manufacturing and first market of entry coincide. More precisely, we keep only products made outside the EU or in Bulgaria (a country not in our sample). We further include brands which, to the best of our knowledge, do not have production facilities in

Table 4 – Role of Income and Quality in First (Earlier) Market Entry

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	A. Conditional logit			B. Rank-ordered logit			
ln Income	1.509***	1.458**	1.043	0.617**	0.675*	-1.112	-0.450
	(0.466)	(0.646)	(1.440)	(0.289)	(0.353)	(0.881)	(0.552)
ln Income×$\hat{q}$	2.008***	1.814***	0.931**	1.393***	1.557***	0.770***	1.259***
	(0.548)	(0.564)	(0.418)	(0.354)	(0.399)	(0.264)	(0.376)
γ_m	No	Yes	Yes	No	Yes	Yes	Yes
τ_t	No	No	No	No	Yes	Yes	Yes
ζ_j	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Products	5,212	5,212	280	5,212	5,212	280	1,710
Brands	41	41	19	41	41	19	41
N	125,088	125,088	6,720	125,088	125,088	6,720	41,040

Notes: The method of estimation is conditional logit in (1)-(3) and rank-ordered logit in (4)-(7). The dependent variable in (1)-(3) equals one for the first market(s) of entry and zero for all the remaining countries in the dataset, and it is the market sequence or Rank in (4)-(7). The identification of first and subsequent markets at the product level is discussed in Section 4.1 and summarized in Table A.9 in the Appendix. The method of handling ties in the rank-ordered logit is Efron's. Specifications (2)-(3) and (5)-(7) include country fixed effects, γ_m. Specifications (5)-(7) further include country-specific (first)-year fixed effects, τ_t. Product fixed effects, ζ_j, are implicitly taken into account in the estimation by specifying the identifier variable in the conditional logit estimation, and by `cmset` in the ordered-logit model, which declares the data to be cross-sectional choice model data. In specifications (3) and (6) the sample is restricted to products manufactured in China, Russia, South Korea, Thailand, Turkey, or Bulgaria, in addition to the following brands: Atlant, Hisense, Shivaki, Nord, Vestel and Sharp, which do not have manufacturing facilities in Europe. In specification (7), the sample is reduced to annual cohorts 2009 and 2010. $\hat{q}$ is the product-specific quality index from eq. (1). Standard errors are clustered by brand except in (3) and (7) where clustering is by product. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

the EU.[39] The resulting sample consists of 280 products and the CLM estimates based on it are reported in column (3) of Table 4. The effect of log income is not statistically different from zero in this specification, but the coefficient on the interaction term with quality remains positive and significant at the 5% level, albeit its magnitude is slightly smaller than the benchmark estimate in (2). Importantly, the results using export destinations alone confirm that the likely mechanism at play in market selection is demand-driven, and cannot be explained by supply-side considerations.[40]

Rank-ordered market sequences. –The CLM estimates presented in Panel A of Table 4, while shedding light on the importance of income and quality in the selection of the

[39] See the note to Table 4 for details on countries of origin and brands.

[40] The smaller estimated effect of the interaction term in (3) is not surprising given that products in this sub-sample are of slightly lower quality compared to the 2009-2013 product cohorts overall. The average quality index is 0.170 (s.e. 449) compared to 0.190 (s.e.0.442) in the full estimation sample. As expected, the quality difference is also reflected in prices as products in (3) are 74 euro cheaper, on average.

most preferred alternative (country) do not take into account the additional information contained in the full product-specific temporal market sequences. To do so requires an alternative estimation approach that also considers the choice of subsequent markets. As discussed above, we interpret the temporal sequence of countries as a ranking of their profitability potential for a given product with an earlier entry indicative of a more preferred alternative (market).

Based on the above interpretation and our construction of the dependent variable 'Rank' (see Table A.9), Panel B of Table 4 employs the ranked-ordered logit model to study the determinants of *earlier* market entry. The most highly valued market is the first market of entry and has the highest rank. A positive association between a market's attribute and the dependent variable, for instance, would imply that a higher value of the attribute increases the chance of an earlier entry in that market, or equivalently the chance of a higher rank. Column (4), similarly to (1), incorporates only product fixed effects, while (5) additionally includes market and first-year-of-entry indicators, τ_t. Note that, since the global date of (EU) entry does not vary within a product, τ_t are subsumed by ζ_j in the conditional logit estimates in Panel A. In Panel B, however, as product j can be launched across countries in different years, τ_t can be identified from within-product variation. Regardless of the fixed-effects structure, specifications (4) and (5) both unambiguously highlight the importance of quality for the geographic positioning of products: goods of high quality are consistently found to enter high-income markets earlier than other destinations.

Two types of robustness checks are performed in (6)-(7) in Panel B. In (6), we repeat the specification restricting the sample to a group of products whose market sequences are certain to consist solely of export destinations. Similarly to (3), the coefficient on log income turns insignificant and the effect of the interaction term is weaker but economically meaningful, pointing to underlying demand-side fundamentals in market choice. Last but not least, in (7), we address the possibility of errors in the product-specific market sequences, which only affect second and further markets and can be especially pronounced for product cohorts 2012-2013. As discussed earlier, because country coverage decreases sharply in 2014, later market entries other than the first EU market may be misidentified. To check how this possibility plays out in the estimation, in (8) we report results based only on product-cohorts 2009-2010, i.e., the cohorts with the most accurate and complete market sequences in the data as, for them, data for all 24 markets is available the longest. We continue to find a statistically significant earlier entry of premium goods in wealthier markets.

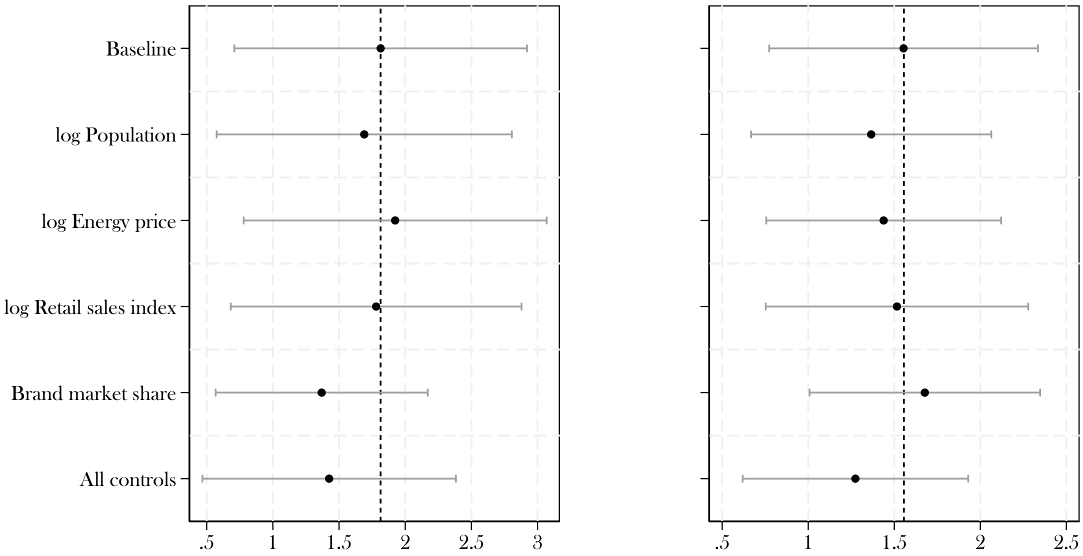
6.3 Robustness to confounding factors

Despite the extensive fixed-effects structure of the specifications in Table 4, thus far we empirically model the potential profit of product j in market m at time t with a single time-varying market attribute – income per capita, alongside its interaction with quality. Can other market-specific or product-market-specific covariates such as market size and brand familiarity explain away the effect of income on the market selection of high(er) quality goods? We explore the presence of confounding factors by introducing a series of covariates (and their interaction with quality) and examining the impact of their inclusion on the estimate of our main coefficient of interest. For convenience, the results are presented graphically in Figure 5, while Tables A.7 and A.8 in the Appendix report a full set of coefficients for the CLM and the ROL, respectively.

We first check how our results perform once we control for market size through the inclusion of log population. Next, we consider residential energy prices to possibly gauge preferences for energy performance. We expect that these differ across consumers facing varying energy costs and may influence demand for higher quality appliances, which typically are also more energy efficient. To capture consumer sentiment and economic conditions to some degree, we also include the log of the index of turnover for retail trade. None of these variables (and their interaction with quality) influence the finding of an earlier entry in richer markets for products positioned at the higher end of the quality distribution: lines 2, 3, and 4 on Figure 5 indicate that the estimate of the income-quality interaction remains very close to the baseline without additional covariates (line 1).

We next examine whether differential brand presence in a market can be an important confounding factor in market selection. Are new high-quality models first introduced in high-income markets because they happen to be the markets where high-quality brands have their largest market shares, and thus largest consumer familiarity with their products? Market shares proxy for brand loyalty (e.g., Villas-Boas, 2004) and thereby indirectly for the degree to which public perceptions of a product belonging to a given brand are well-established (for example, perceptions of quality, environmental impact, design appeal, etc.). Brand loyalty can also go hand-in-hand with the presence of so-called switching costs such as learning costs involved in transitioning from familiar to unfamiliar brands, which presumably consumers want to avoid (e.g., Beggs and Klemperer, 1992). Differential market shares may also be indicative of the maturity of distribution networks and other related infrastructure pertaining to brand-country-specific supply chains. In this respect, they may also be informative of costs as brands with large market shares are likely to already have long-established logistics facilities and local market knowledge,

FIGURE 5 – Effect of ln Income x Quality: Alternative Specifications



(a) Conditional logit: Choice of first market

(b) Rank-ordered logit: Choice of market sequence

Notes: The figure plots conditional-logit estimates (a) and rank-ordered logit estimates in (b) of the interaction term of the log of income with the product-specific quality estimate for models that include log population. the log of household energy prices, the log of retail sales index (100=2009), the brand market share, and all of the above controls. In all specifications, the interaction term(s) of the respective covariate(s) with quality is also included. Coefficient estimates are reported in Tables A.7 and A.8 in the Appendix. The dashed line denotes our baseline estimate from column (2) in Table 4 in (a) and column (6) from the same table in (b). 95% confidence intervals are also depicted.

and thus potentially lower costs when placing a new product.

The market share of a brand indeed emerges as an important determinant in the choice of products' first and earlier markets as demonstrated in both Tables A.7 and A.8 – the better established a brand in a given market, the higher the chance of new varieties being introduced in that market first. This effect does not appear to vary heterogeneously with quality though. Despite the large estimated effect of brand market share, line 5 of Figure 5 in both plots (a) and (b) reveals that the effect of the interaction between income and quality remains still close to our benchmark estimate and highly economically relevant for market choice.

Last but not least, line 6 of the figure shows the change in the coefficient of ln Income x Quality when all of the above covariates themselves alongside their interactions with quality enter the estimation. Once again, our coefficient of interest remains virtually unchanged. We further tested the sensitivity of our results to the inclusion of one additional covariate reported in columns (6) of Tables A.7 and A.8. Specifically, we study the effect of entry costs in the form of transportation costs measured as the distance (in km) from the country of manufacture to a country of sale based on the sample of products whose country of origin is identified. Transportation costs are corroborated to play an important role in first(earlier)-location choices. As expected, farther markets relative to production locations are found to be less attractive for new product launches. We do not find a differential effect along the quality distribution for this result. Reassuringly, the income-quality effect remains an important factor despite being larger in magnitude than our baseline estimate in this sub-sample.

7 Empirical Analysis II: Geographic Distance and Market Expansions by Quality

Our model predicts a heterogeneous impact of geographic distance on market expansion decisions for different quality layers. More precisely, Proposition 3 states that while geographic frictions render market expansions more likely in locations contiguous to the first market of entry, the impact of those frictions becomes relatively less pressing for higher-quality varieties than for lower-quality ones. Figure 4 provided a quick snapshot suggesting that the average distance of market expansions is wider at higher levels of quality, especially in the cases of the second market entry. This subsection will offer a more thorough econometric analysis of the impact of geography on market expansion choices guided by the results in Section 5.2.

In Section 5.2, we imposed some simplifying assumptions for the sake of analytical tractability, which require further discussion before making contact with the data. Firstly, we assumed that only one market may be selected as the first market of entry. Secondly, we imposed a ("fractal") geographic structure in which all markets are "geographically identical" in terms of location relative to other markets with a given income level. Neither assumption is in fact crucial to the prediction that geographic distance matters relatively less for expansion decisions at higher levels of quality; the underlying reason for this result rests actually on the variable equilibrium mark-ups stemming from our nonhomothetic demand structure. Nevertheless, to assess empirically the validity of the main result in Section 5.2, we need now to account for the following two facts: i) the dataset includes instances with multiple first markets of entry; ii) the geographic distribution of the 24 countries in the dataset is far from being "fractal" (in particular, neither every single country has the same number of neighboring countries, nor the distribution of income of the respective neighbors is identical for all).

To account for the possibility of multiple first markets of entry on a given date, we let $\Phi_j \subseteq \Omega$ denote the subset of markets within Ω (the set of 24 European countries in the panel) where variety j was first introduced. There are a total of 5,212 different varieties. The set Φ_j comprises one single element (i.e., there is one single first market of entry) in 76.8% of the sample, two elements in 16.1%, three in 4.2%, four in 1.9%, and five, six and seven elements for 0.7%, 0.2% and 0.1% of the sample, respectively. Notice that there are in total as many different compositions of the sets Φ_j as combinations of first market(s) of entry present in the data.[41] In the benchmark regressions, we will treat all varieties equally, regardless of the size of the set Φ_j. We will, however, control for the set of first markets of entry by including a full set of fixed effects for all the combinations of Φ_j present in the data. In addition, we will also show results restricting the analysis to the subsample of models for which Φ_j contains one single element.

Denote by $\Phi'_j \subset \{(\Omega - \Phi_j) \cup \emptyset\}$ the subset of market(s) —possibly empty— where variety j is subsequently introduced as second market(s) of entry. There are 4,213 (80.8%) models for which $\Phi'_j \neq \emptyset$. These models represent our relevant benchmark sample for the regressions in this section.[42] Restricting the attention to the cases in which $\Phi'_j \neq \emptyset$, we let $\Theta_j = \Phi_j \times \Phi'_j$ denote the set of all possible combinations of country pairs between $m \in \Phi_j$ and $m' \in \Phi'_j$ for model j, and next define the index function $\mathbb{I}_j(m, m') : \Theta_j \to \{0, 1\}$, such that $\mathbb{I}_j(m, m') = 1$ if and only if m' and m are neighbouring countries.

Henceforth, we will consider a market expansion to m' to be a 'non-neighboring market expansion' if and if only this market does not share any border with any of model j's first markets of entry; that is, if and only if $\mathbb{I}_j(m, m') = 0$ for all $m \in \Phi_j$. As a result, the share $S_j \in [0, 1]$ of market expansions that take place in non-neighboring countries relative to the total number of market expansions for model j is given by:

$$S_j = \frac{\displaystyle\sum_{m' \in \Omega'_j} \left(1 - \max_{m \in \Phi_j} \{\mathbb{I}(m, m')\}\right)}{n(\Phi'_j)}, \tag{21}$$

where $n(\Phi'_j)$ is the number of elements of Φ'_j – we naturally restrict the analysis to $n(\Phi'_j) \geq 1$.

[41] There are 165 different combinations of 'first market(s) of entry' present in our dataset. Note that if the set of 'first market(s) of entry' of two different varieties j and k are identical, then $\Phi_j = \Phi_k$.

[42] Amongst the 4,213 models that exhibit a market expansion during the sample years 2009-13, 74% has one single element in Φ'_j, 15.8% has two elements, 5.8% three elements, 2.4% four elements, and 1%, 0.5%, 0.2%, and 0.1% have five, six, seven and eight elements, respectively. The remainder are a few odd cases in which Φ'_j comprises 9, 10 and 11 markets.

TABLE 5 – Share of Non-neighboring Market Expansions and Quality

	(1)	(2)	(3)	(4)	(5)	(6)
Quality	0.061***	0.055***	0.059***	0.060***	0.056***	0.060***
	(0.015)	(0.014)	(0.021)	(0.016)	(0.015)	(0.022)
Constant	0.404***	0.385***	0.384***	0.413***	0.387***	0.386***
	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)	(0.007)
First Market(s) FE	Yes	Yes	Yes	Yes	Yes	Yes
Second Market(s) FE		Yes	Yes		Yes	Yes
Brand FE			Yes			Yes
Sample (Models)	All	All	All	Single First	Single First	Single First
Observations	4,213	4,213	4,213	3,155	3,155	3,155

Notes: The dependent variable is the share of non-contiguous market expansions to a second market of entry. Columns (1)-(3) are performed on all models that experience market expansions regardless of the number of first markets of entry. Columns (4)-(6) are performed on the sub-sample of models that exhibit only one single first market of entry. Robust standard errors are reported in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Our benchmark regression in Table 5, column (1), shows the results of the following OLS regression:

$$S_j = constant \ + \beta \cdot q_j + \psi_{\Phi_j} + \varepsilon_j, \tag{22}$$

where each S_j stems from (21), q_j is the level of quality characterizing variety j, and ψ_{Φ_j} denotes a full set of fixed effects for each combination of 'first market(s) of entry' present in the dataset. Note that by including ψ_{Φ_j}, we are controlling for the fact that each country has different sets of neighboring countries with varying characteristics (and in particular, with different incomes per capita). Since those factors may expectably be correlated with q_j, we wish to explicitly control for the exact set of countries in which each model was first introduced and exploit only variation in quality within each of them. Notice that ψ_{Φ_j} will also take into account that a set Φ_j of larger size will tend to (mechanically) have more neighboring markets, albeit at the same time less scope for total expansion (as a smaller number of markets are possibly left for expansion within the original set of 24 countries), which in turn will affect both the numerator and denominator in (21).

The result of the OLS regression (22) reported in column (1) yields a positive and highly significant estimate of β. This entails that, abstracting from variations in the set of first market of entry Φ_j, the share of expansions that take place in markets that are non-neighboring to any of the first markets of entry tends to be greater for models of higher quality. In terms of its quantitative impact, denoting by q_X the value of the quality measure at the X^{th} percentile, and computing the value of $\hat{S}(q_X)$ based on the estimates in column (1) for the set Φ_j for which $\hat{\psi}_{\Phi_j} = 0$, we obtain that $\hat{S}(q_5) = 0.372$ and $\hat{S}(q_{95}) = 0.462$.[43] Within our sample of 24 European markets, the average number of non-neighboring markets is 21.5. The results for $\hat{S}(q_5)$ and $\hat{S}(q_{95})$ then suggest that market expansions are in general much more likely to involve contiguous markets, but also that the quantitative difference in the probabilities of non-contiguous market expansions across top and bottom qualities is nonetheless quite wide. That is, our results confirm that geographic distance matters when assessing where to further expand the market penetration of new products. Yet, they also show that the influence of such "gravity pull" declines at higher levels of quality.

In column (2) we add to the benchmark regression specification (22) a full set of fixed effects for all the combinations of 'second market(s) of entry' present in the data. These fixed effects will control for the fact that certain countries may tend to be more regularly those where market expansions take place after the initial entry of a new model. In addition, these fixed effects will account for the specificities of the geography of each country, which will differ in terms of the number of neighboring countries (and their levels of income).[44] Next, in column (3) we also include a set of brand fixed effects. The reason for this would be to control for the possibility that certain brands may in general be present in certain markets, and absent in others. Since different brands tend to vary as well in terms of the range of quality of models they offer to consumers, an uneven geographic distribution of brand presence could represent a source of bias to the results. As we can observe from columns (2) and (3), both the magnitude and statistical significance of the estimate for the coefficient of interest remain quite stable after including these additional controls.

As a further robustness check, columns (4)-(6) repeat in turn the regressions carried out in columns (1)-(3), but restrict the analysis to the subsample of models with one single market of entry. In these cases the set of fixed effects ψ_{Φ_j} becomes a set of 'first country of entry' fixed effects. Restricting the sample to models whose ψ_{Φ_j} identifies a singleton seems an important robustness check for two separate reasons. Firstly, it represents a more natural empirical setup against which to compare the predictions of the dynamic version of the model in Section 5.2. Secondly, in cases of models with multiple first markets of entry, further market expansions could exhibit a different degree of sensitivity

[43] The value of those quality indices stemming from our Hedonic price regressions are q_5=-0.519 and q_{95}=0.938

[44] Note that the set of fixed effects in column (1), ψ_{Φ_j}, had controlled for the geographic specificities of the 'first market(s) of entry' but not for those of the 'second market(s) of entry'.

to variations in quality. Reassuringly, the results in columns (4)-(6) remain quantitatively very similar to their counterparts in columns (1)-(3).

Table 5 exploits a binary distinction between neighboring vs. non-neighboring pairs of countries. While this approach has the appeal of highlighting the fluidity of sharing a border, it overlooks important aspects of the geography of countries in our dataset. One is that the dependent variable in Table 5 treats all pairs of non-contiguous countries identically regardless of how distant they are from each other.[45] Analogously, all contiguous countries are also treated identically, as if they all share an equally long border and have their entire populations sitting side by side across their common border.

To account for these nuances, in Table 6, we replace the dependent variable S_j in (21) by one that reflects average distances between markets. For each pair of countries m and m', we use their bilateral weighted geographic distance $dist\,(m, m') = dist\,(m', m)$, where weights are based on city-level population distribution in each of the two countries.[46] Based on the bilateral distances for each pair of countries, we next define:

$$Mean\_Dist_j = \frac{1}{n(\Phi'_j)} \cdot \sum_{m' \in \Phi'_j} \min_{m \in \Phi_j} \left\{ dist\,(m, m') \right\}, \tag{23}$$

to be used as the dependent variable throughout all the specifications in Table 6.[47]

All the columns in Table 6 follow the same pattern of controls and samples as those in Table 5. As it may be readily observed, all the results based on (23) as the dependent variable remain qualitatively in line with the previous ones in Table 5. In terms of its quantitative interpretation, the estimate for β in column (1) entails that the average distance between the first market(s) of entry and those to where the first wave of market expansion takes place tends to be approximately 17% greater for the variety in 95^{th}-percentile of quality than for the variety in 5^{th}-percentile of quality.

[45] That means we are treating an expansion from Germany to Italy as equal to one from Germany to Portugal, even though the shortest distance between the former pair is approx. 70 km while between the latter is almost 2000 km.

[46] Bilateral distance data are sourced from CEPII database – see Conte, Cotterlaz and Mayer (2022).

[47] The definition of (23) computes the mean distance of market expansion for model j using the bilateral distances between a second market of entry $m' \in \Phi'_j$ and its nearest first market of entry for that model. All the results in Table 6 are robust to using the alternate dependent variable:

$$Mean\_Dist_j = \frac{1}{n\left(\Phi'_j\right)} \cdot \sum_{m' \in \Phi'_j} \left(\frac{1}{n\left(\Phi'_j\right)} \cdot \sum_{m \in \Phi(j)} dist\,(m, m') \right).$$

Table 6 – Market Expansions Distance and Quality

	(1)	(2)	(3)	(4)	(5)	(6)
Quality	79.43***	47.49***	39.87***	79.55***	51.97***	42.77***
	(14.69)	(11.23)	(18.73)	(15.19)	(12.14)	(20.12)
Constant	733.86***	736.49***	737.83***	773.67***	754.55***	756.09***
	(5.78)	(5.36)	(5.86)	(6.65)	(5.96)	(6.39)
First Market(s) FE	Yes	Yes	Yes	Yes	Yes	Yes
Second Market(s) FE		Yes	Yes		Yes	Yes
Brand FE			Yes			Yes
Sample (Models)	All	All	All	Single First	Single First	Single First
Observations	4,213	4,213	4,213	3,155	3,155	3,155

Notes: The dependent variable is the average distance between second market(s) of entry and its nearest first market of entry. Columns (1)-(3) are performed on all models that experience market expansions regardless of the number of first markets of entry. Columns (4)-(6) are performed on the sub-sample of models that exhibit only one single first market of entry. Robust standard errors are reported in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Spreading the Good Apples Out

The results in Tables 5 and 6 demonstrate that market expansions for higher-quality products tend to occur farther away from the initial entry markets. If the markets where products are first launched coincide with the country where they are manufactured, one plausible explanation for this pattern could be the Alchian-Allen effect (Alchian and Allen, 1964), often referred to as "shipping the good apples out." This hypothesis states that, when faced with per-unit transaction or shipping costs, producers find it more profitable to sell higher-quality versions of their products in more distant markets [see, e.g., Hummels and Skiba (2004), Baldwin and Harrigan (2011), and Crozet et al. (2012)].

From one perspective, our findings on the geography of market expansions can be seen as a dynamic extension of the static Alchian-Allen effect. However, our model goes further, offering additional predictions beyond the traditional "shipping the good apples out" framework. According to our model, the observed greater distances between initial entry markets and subsequent expansion markets for higher-quality varieties result from nonhomothetic demand factors pulling more strongly against gravity forces for premium products than entry-level ones. This interplay –or "tug-of-war"– between variable markups and gravity effects occurs regardless of the manufacturing country's location. As a result, our model predicts similar patterns in the average distance of market expansions whether or not the first entry market coincides with the manufacturing country. We term

Table 7 – Market Expansions and Quality: Subsample with known market of manufacturing

	(1)	(2)	(3)	(4)	(5)	(6)
	Share Non-neighboring			Average Distance		
Quality	0.101***	0.101**	0.092*	133.68***	100.19**	149.10**
	(0.037)	(0.042)	(0.053)	(37.83)	(47.96)	(60.73)
Constant	0.522***	0.580***	0.582***	820.04***	891.14***	879.36***
	(0.018)	(0.020)	(0.022)	(18.60)	(24.11)	(25.47)
First Market(s) FE	Yes	Yes	Yes	Yes	Yes	Yes
Country of Origin FE	No	No	Yes	No	No	Yes
Sample (Models)	all with known origin	restricted	restricted	all with origin origin	restricted	restricted
Observations	625	400	400	625	400	400

Notes: The dependent variable in columns (1)-(3) is the share of non-contiguous market expansions to a second market of entry, while in columns (4)-(6) it is the average distance between second market(s) of entry and its nearest first market of entry. Regressions in columns (1) and (4) are performed on the entire sample of models with a known country of manufacturing, while those in (2), (3), (5), and (6) are performed on the subsample of models for which the country of manufacturing is neither among the first market(s) of entry nor among the second market(s) of entry. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

this broader pattern as "spreading the good apples out."

To assess whether the data on market expansions is also consistent with the idea of "spreading the good apples out", in Table 7 we report a series of regressions conducted on the sample of products for which we have located their country of manufacture. This leaves us with 625 fridge models that will eventually undergo a market expansion – that is, 14.8% of the sample used in columns (1)-(3) in Tables 5 and 6.

In columns (1) and (4) of Table 7, we replicate the regressions from column (1) of Tables 5 and 6, respectively, focusing on products with known country of manufacture. In columns (2) and (5), we refine the analysis by restricting the sample to products where the country of manufacture is neither the first nor the second market of entry. These results demonstrate that even when analyzing products initially introduced in markets different from their country of manufacture, and whose first market expansions also occur in other countries, geographic proximity still plays a relatively smaller role in market expansion decisions for higher-quality products. Furthermore, as shown in columns (3) and (6), these results remain robust when controlling for country of manufacture fixed effects.

8 Conclusion

This paper has studied the dynamics of entry of vertically differentiated products into a world economy where market expansions face geographic frictions and consumers exhibit nonhomothetic preferences for quality. Our findings reveal that geographic proximity is a key determinant in the market expansion strategies of lower-quality products. However, its importance diminishes for higher-quality products, as high-end producers prioritize catering to wealthier consumers' preference for quality over spatial considerations. This distinction in strategy underscores the critical role of income distribution in shaping market expansion paths within vertically differentiated industries.

The empirical analysis of the refrigerator industry across 24 different European countries validates our model's main predictions, showing that high-quality products are indeed introduced first in wealthier countries and subsequently expanded to more distant, high-demand markets. Meanwhile, lower-quality products adhere more strictly to geographic proximity in their market entry sequences.

These results highlight how income-related preferences impact the broader structure of global trade networks, with significant implications for export strategy formulation. By acknowledging that high- and low-quality goods tend to follow distinct geographic and income-based pathways, policymakers may tailor more refined export-promoting platforms that account for quality differences. For example, logistical support and policies aimed at overcoming trade frictions may have a much stronger impact on lower-quality versions of goods than on premium ones. On the other hand, promoting quality upgrading would call for an enhanced effort in the establishment of trade agreements with richer economies without excessive concern in prioritizing geographic considerations.

While our empirical exercises focus on a specific sector, the framework could be adapted to a broader range of contexts. In our setting, a 'producer' represents a brand offering a range of refrigerator models, with each model treated as a unique 'variety'. However, this structure could be readily adapted to other scenarios. For instance, a 'producer' might represent a country, and a 'variety' a given product category within customs data – a common approach in the international trade literature. Conceptually, this model thus provides a versatile foundation for examining income-driven trade dynamics both within and across diverse sectors. Consequently, our model opens avenues for further research on nonhomothetic preferences and vertical differentiation, potentially aiding both theoretical advancements and policy-making in the field of international trade.

References

Albornoz, F., Pardo, H.F.C., Corcos, G., and Ornelas, E. (2012). Sequential exporting. *Journal of International Economics*, 88, 17-31.

Albornoz, F., Pardo, H.F.C., Corcos, G., and Ornelas, E. (2023). Sequentially exporting products across countries. *Journal of International Economics*, 142, 103735.

Alchian, Armen A., and William R. Allen. 1964. *University Economics*. Belmont,Calif.: Wadsworth

Auer, R., Chaney, T., and Saure, P. (2018). Quality pricing-to-market. *Journal of International Economics*, 110, 87-102.

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

Beggs, S. and Cardell, S. (1981). Assessing the potential demand for electric cars. *Journal of Econometrics*, 16, 1-19.

Beggs, A. and Klemperer, P. (1992). Multi-period competition with switching costs. *Econometrica*, 60, 651-666.

Chaney, T. (2014). The network structure of international trade. *American Economic Review*, 104, 3600-3634.

Conte, M., Cotterlaz, P., and Mayer, T. (2022). The CEPII gravity database. Working Papers 2022-05, CEPII research center.

Cockburn, I. M., Lanjouw, J. O., and Schankerman, M. (2016). Patents and the global diffusion of new drugs. *American Economic Review*, 106(01), 136-164.

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

Deaton, A. and Muellbauer, J. (1980). An almost ideal demand system. *American Economic Review*, 70, 312-326.

Defever, F., Heid, B., and Larch, M. (2015). Spatial exporters. *Journal of International Economics*, 95, 145-156.

Devereux, M.P., Griffith, R., and Simpson, H. (2007). Firm location decisions, regional grants and agglomeration externalities. *Journal of Public Economics*, 91, 413-435.

Egger, P., Foellmi, R., Schetter, U., and Torun, D. (2023). Gravity with History: On Incumbency Effects in International Trade (No. 153a). Center for International Development at Harvard University.

European Commission (2016). Preparatory/review study. Commission Regulation (EC) No. 643/2009 with regard to ecodesign requirements for household refrigeration appliances and Commission Delegated Regulation (EU) No. 1060/2010 with regard to energy labelling of household refrigeration appliances. Final Report.

Faber, B. and Fally, T. (2022). Firm heterogeneity in consumption baskets: Evidence from home and store scanner data. *Review of Economic Studies*, 89, 1420-1459.

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

Fajgelbaum, P. and Khandelwal, A. (2016). Measuring the unequal gains from trade. *Quarterly Journal of Economics*, 131, 1113–1180.

Flam, H. and Helpman, E. (1987). Vertical product differentiation and north-south trade. *American Economic Review*, 77, 810-822.

Fok, D., Paap, R., and Van Dijk, B. (2012). A rank-ordered logit model with unobserved heterogeneity in ranking capabilities. *Journal of Applied Econometrics*, 27, 831-846.

Hallak, J.C. (2006). Product quality and the direction of trade. *Journal of International Economics*, 68, 238-265.

Hallak, J.C. (2010). A product-quality view of the Linder hypothesis. *Review of Economics and Statistics*, 92, 453-466.

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

Handbury, J. (2021). Are poor cities cheap for everyone? Non-homotheticity and the cost of living across U.S. cities. *Econometrica*, 89, 2679-2715.

Head, K., and Mayer, T. (2004). Market potential and the location of Japanese investment in the European Union. *Review of Economics and Statistics*, 86, 959-972.

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

Hummels, D. and Skiba, A. (2004). Shipping the Good Apples Out? An Empirical Confirmation of the Alchian-Allen Conjecture. *Journal of Political Economy*, 112, 1384-1402.

Jaimovich, E. and Merella, V. (2015). Love for quality, comparative advantage, and trade. *Journal of International Economics*, 97, 376-391.

Jaimovich, E., Madzharova, B., and Merella, V. (2023). Inside the white box: Unpacking the determinants of quality and vertical specialization. *European Economic Review*, 152, 104374.

Jaravel, X. (2019). The unequal gains from product innovations: Evidence from the US retail sector. *Quarterly Journal of Economics*, 134, 715-783.

Khandelwal, A. (2010). The long and short (of) quality ladders. *Review of Economic*

Studies, 77, 1450-1476.

Linder, S. (1961). An essay on trade and transformation. Uppsala: Almqvist and Wiksell.

Punj, G.N. and Staelin, R. (1978). The choice process for graduate business schools. *Journal of Marketing Research*, 15, 588-598.

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

Morales, E., Sheu, G., and Zhaler, A. (2019). Extended gravity. *Review of Economic Studies*, 86, 2668-2712.

Murphy, K. and Shleifer, A. (1997). Quality and trade. *Journal of Development Economics*, 53, 1-15.

Schott, P. (2004). Across-product versus within-product specialization in international trade. *Quarterly Journal of Economics*, 119, 647-678.

Simonovska, I. (2015). Income differences and prices of tradables: Insights from an online retailer. *Review of Economic Studies*, 82, 1612-1656.

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

Villas-Boas, J.M. (2004). Consumer learning, brand loyalty, and competition. *Marketing Science*, 23, 134-145.

Appendix

A.1 Omitted Proofs

Definition 1. *Let* $M_q \equiv \left(\int_{\mathcal{J}_q} \ln \mu_j \, dj\right)/N_q$*, for* $q = l, h$*.*

Lemma 2. *A Nash equilibrium features symmetry in the mark-ups of varieties within quality layers; that is,* $\mu_j^* = \mu_q^*$ *for all* $j \in \mathcal{J}_q$ *and* $q = l, h$*. Furthermore, it is always unique and features* $\mu_j^* > 1$ *for all* j*.*

Proof. Using Definition 1, we may rewrite (8) as:

$$\mu_j + \ln \mu_j = 1 + (q - \Gamma) \ln y + \frac{1}{N} + M_q, \tag{A.24}$$

where we have exploited $\left(\int_{k \neq j} \ln \mu_k \, dk\right)/N_q = M_q - \left(\int_{k=j} \ln \mu_k \, dk\right)/N_q$ and the fact that an integral taken over a degenerate interval equals zero. The symmetry of mark-ups then immediately follows from noticing that the sum $\mu_j^* + \ln \mu_j^*$ in the LHS of (A.24), written for the optimal mark-up μ_j^*, is strictly increasing in μ_j^*, while the RHS of (A.24), $1 + (q - \Gamma) \ln y + 1/N + M_q^*$, is constant for a given level of quality $q = l, h$.

Consider now the definition $\mu_j \equiv \varepsilon_j / (\varepsilon_j - 1)$ jointly with (**??**) and (7). It follows that $\mu_j > 1$ whenever $D_j > 0$ (i.e., for all firms in the market). As a consequence, there cannot be a Nash equilibrium with $M_q^* = 0$, as this would immediately imply $\mu_j^* = 1$ for all j. Next, notice from (A.24) that since the best-response functions $\mu_j(M_q)$ are such that $\partial \mu_j / \partial M_q > 0$, a sufficient condition for the existence of a unique equilibrium where all $1 < \mu_j^* < \infty$ for all $j \in \mathcal{J}_q$ will be:

$$\frac{\partial \ln \mu_j}{\partial M_q} < 1 \quad \text{and} \quad \frac{\partial^2 \ln \mu_j}{(\partial M_q)^2} < 0 \text{ for all } M_q > 0. \tag{A.25}$$

Differentiating the best-response functions $\mu_j(M_q)$ in (8) with respect to M_q and rearranging yields:

$$\frac{\partial \ln \mu_j}{\partial M_q} = \frac{1}{1 + \mu_j}, \tag{A.26}$$

and the expression in (A.26) straightforwardly implies that both conditions in (A.25) hold true. □

Proof of Proposition 1. Recall Definition 1, and notice that from Lemma 2 it immediately follows that $M_q^* \equiv \left(\int_{\mathcal{J}_q} \ln \mu_q^* \, dk_q\right)/N_q = \ln \mu_q^*$, with $q = l, h$. We can thus write down the conditions for the Nash equilibrium stemming from (A.24) as follows:

$$\mu_l^* = 1 + 1/N + (l - \Gamma)\ln y \quad \text{and} \quad \mu_h^* = 1 + 1/N + (h - \Gamma)\ln y,$$

from where (9) and (10) obtain after plugging $\Gamma = (N_l l + N_h h)/N$ in the above expressions, and letting $h - l = 1$. □

Lemma 3. *Let $\Delta_{qy} \equiv \Lambda_{qy} - \Lambda_{qy'}$, with $y' \neq y$, where $\Lambda_{qy} = \delta^{-1}(\mu_{qy} - 1)^2/\mu_{qy}$. Then $\mathcal{P}_{qy} = \Phi(\Delta_{qy})$, where $\Phi(\Delta_{qy})$ is an increasing function of Δ_{qy} and $\Phi(0) = 1/M$.*

Proof. Bearing in mind (11), we can observe that for a generic country m with real per-capita income y and a generic newly designed variety $j \in \mathcal{J}_q$, we have that the expected value of the intertemporal stream of profit net of the entry cost in market m is given by:

$$\Pi_{jy} = \Lambda_{qy} - \phi_{jm}. \tag{A.27}$$

Consider a pair of generic countries m' and m'' with income y and $y' \neq y$, respectively. Using (A.27), it follows that variety j is introduced first in m if the following two conditions hold simultaneously:

$$\phi_{jm} < \phi_{jm'}, \text{ for all } m' \neq m, \tag{A.28}$$

$$\phi_{jm} < \phi_{jm''} + \Delta_{qy}, \text{ for all } m''. \tag{A.29}$$

Note that, if $\Delta_{qy} < 0$, there exist a subset of values of ϕ_{jm} for which variety j is not introduced first in m, while if $\Delta_{qy} > 0$, there exist a subset of values of ϕ_{jm} for which variety j is not introduced first in any m''. Hence, letting $\tilde{\phi}_b \equiv \max\left\{\underline{\phi}, \underline{\phi} + \Delta_{qy}\right\}$ and $\tilde{\phi}_a \equiv \min\left\{\overline{\phi} + \Delta_{qy}, \overline{\phi}\right\}$, and provided that $\underline{\phi} - \overline{\phi} < \Delta_{qy} < \overline{\phi} - \underline{\phi}$, from the set of conditions (A.28) and (A.29), the probability that the newly designed variety of quality q will be introduced first in a market with income y is given by:

$$\mathcal{P}_{qy} = \int_{\underline{\phi}}^{\tilde{\phi}_b} (1 - F(\phi_j))^{\frac{M}{2}-1} f(\phi_j)\, d\phi_j + \int_{\tilde{\phi}_b}^{\tilde{\phi}_a} (1 - F(\phi_j))^{\frac{M}{2}-1} (1 - F(\phi_j - \Delta_{qy}))^{\frac{M}{2}} f(\phi_j)\, d\phi_j, \tag{A.30}$$

where $f(\cdot)$ denotes the pdf function associated to $F(\cdot)$. From (A.30) we can observe that this probability is a function of Δ_{qy}. We can thus write:

$$\mathcal{P}_{qy} = \Phi(\Delta_{qy}).$$

Notice that $\Phi(0) = 1/M$. In addition, when $\Delta_{qy} > 0$ and $\Delta_{qy} < \overline{\phi} - \underline{\phi}$ we have $\partial \mathcal{P}_{qy}/\partial \Delta_{qy} > 0$, while when $\Delta_{qy} < 0$ and $\Delta_{qy} > \underline{\phi} - \overline{\phi}$ we also have $\partial \mathcal{P}_{qy}/\partial \Delta_{qy} > 0$.[48] $\square$

Proof of Proposition 2. Bearing in mind (12) and (14), and recalling that $\Delta_{qy} \equiv \Lambda_{qy} - \Lambda_{qy'}$ and $\mathcal{P}_{qy'} = 2/M - \mathcal{P}_{qy}$, the steady-state equilibrium will be characterised by the (simultaneous) solution of the following four equations stemming from (13):

$$LP\text{-locus}: \quad \tilde{\Phi}\left(\mu^*_{lp}, \mu^*_{lr}\right) + \frac{2(1-\delta)}{M}\Gamma\left(\mu^*_{lp}\right) = \frac{\delta}{\rho M} N^*_{lp}, \tag{A.31}$$

$$LR\text{-locus}: \quad \frac{2}{M} - \tilde{\Phi}\left(\mu^*_{lp}, \mu^*_{lr}\right) + \frac{2(1-\delta)}{M}\Gamma\left(\mu^*_{lr}\right) = \frac{\delta}{\rho M} N^*_{lr}, \tag{A.32}$$

$$HP\text{-locus}: \quad \frac{2}{M} - \tilde{\Phi}\left(\mu^*_{hr}, \mu^*_{hp}\right) + \frac{2(1-\delta)}{M}\Gamma\left(\mu^*_{lr}\right) = \frac{\delta}{\rho M} N^*_{hp}, \tag{A.33}$$

$$HR\text{-locus}: \quad \tilde{\Phi}\left(\mu^*_{hr}, \mu^*_{hp}\right) + \frac{2(1-\delta)}{M}\Gamma\left(\mu^*_{hr}\right) = \frac{\delta}{\rho M} N^*_{hr}, \tag{A.34}$$

where $\tilde{\Phi}\left(\mu^*_{qy}, \mu^*_{qy'}\right) \equiv \Phi\left(\Lambda\left(\mu^*_{qy}\right) - \Lambda\left(\mu^*_{qy'}\right)\right) \in [0, 2/M]$, with $\partial\tilde{\Phi}\left(\mu^*_{qy}, \mu^*_{qy'}\right)/\partial\mu^*_{qy} > 0$ and $\partial\tilde{\Phi}\left(\mu^*_{qy}, \mu^*_{qy'}\right)/\partial\mu^*_{qy'} < 0$, and $\Gamma\left(\mu^*_{qy}\right) \geq 0$, with $\Gamma'\left(\mu^*_{qy}\right) \equiv \partial\Gamma\left(\mu^*_{qy}\right)/\partial\mu^*_{qy} > 0$. Each of these four equations represent the locus along which the mass of varieties of a given quality in markets of a given income (namely, the one depicted on the RHS of the relevant equation) remains constant over time.

Notice that (9) and (10) in Proposition 1 imply that, in a steady state, $\mu^*_{hp} = \mu^*_{lp} = \mu^*_p$

[48] In particular, when $\Delta_{qy} > 0$ and $\Delta_{qy} < \overline{\phi} - \underline{\phi}$

$$\frac{\partial \mathcal{P}_{qy}}{\partial \Delta_{qy}} = \int_{\underline{\phi}+\Delta_{qy}}^{\overline{\phi}} \frac{M}{2}\left(1 - F(\phi_j)\right)^{\frac{M}{2}-1}\left(1 - F(\phi_j - \Delta_{qy})\right)^{\frac{M}{2}-1} f(\phi_j - \Delta_{qy}) f(\phi_j)\, d\phi_j,$$

whereas when $\Delta_{qy} < 0$ and $\Delta_{qy} > \underline{\phi} - \overline{\phi}$

$$\frac{\partial \mathcal{P}_{qy}}{\partial \Delta_{qy}} = \int_{\underline{\phi}}^{\overline{\phi}+\Delta_{qy}} \frac{M}{2}\left(1 - F(\phi_j)\right)^{\frac{M}{2}-1}\left(1 - F(\phi_j - \Delta_{qy})\right)^{\frac{M}{2}-1} f(\phi_j - \Delta_{qy}) f(\phi_j)\, d\phi_j.$$

and $\mu_{lr}^* = \mu_{hr}^* - \ln y$. Furthermore, the following two relations must be verified:

$$N_{hp}^* = \frac{1}{\mu_p^* - 1} - N_{lp}^*, \tag{A.35}$$

$$N_{hr}^* = \frac{1}{\mu_{hr}^* - 1} - \frac{\mu_{hr}^* - 1 - \ln y}{\mu_{hr}^* - 1} N_{lr}^*. \tag{A.36}$$

Using (A.35) to equalise (A.31) and (A.33), (A.36) to equalise (A.32) and (A.34), and the equations $\mu_{hp}^* = \mu_{lp}^* = \mu_p^*$ and $\mu_{lr}^* = \mu_{hr}^* - \ln y$, we may reduce (A.31)-(A.34) to the system of two equations in two unknowns (namely, μ_p^* and μ_{hr}^*):

$$\begin{aligned} P\text{-locus}: \quad & \tilde{\Phi}\left(\mu_p^*, \mu_{hr}^* - \ln y\right) + \frac{2(1-\delta)}{M}\Gamma\left(\mu_p^*\right) + \frac{2}{M} - \tilde{\Phi}\left(\mu_{hr}^*, \mu_p^*\right) \\ & \qquad + \frac{2(1-\delta)}{M}\Gamma\left(\mu_{hr}^* - \ln y\right) - \frac{\delta}{\rho M}\frac{1}{\mu_p^* - 1} = 0, \end{aligned} \tag{A.37}$$

$$\begin{aligned} R\text{-locus}: \quad & \frac{2}{M} - \tilde{\Phi}\left(\mu_p^*, \mu_{hr}^* - \ln y\right) + \frac{2(1-\delta)}{M}\Gamma\left(\mu_{hr}^* - \ln y\right) \\ & + \chi\left(\mu_{hr}^*\right)\left[\tilde{\Phi}\left(\mu_{hr}^*, \mu_p^*\right) + \frac{2(1-\delta)}{M}\Gamma\left(\mu_{hr}^*\right) - \frac{\delta}{\rho M}\frac{1}{\mu_{hr}^* - 1}\right] = 0, \end{aligned} \tag{A.38}$$

where $\chi\left(\mu_{hr}^*\right) \equiv \left(\mu_{hr}^* - 1\right) / \left(\mu_{hr}^* - 1 - \ln y\right) > 1$, with:

$$\chi_{hr}' \equiv \frac{\partial \chi\left(\mu_{hr}^*\right)}{\partial \mu_{hr}^*} = \frac{1 - \chi_{hr}}{\mu_{hr}^* - 1 - \ln y}. \tag{A.39}$$

The proof proceeds in three steps. Specifically, with reference to a $\left(\mu_{hr}^*, \mu_p^*\right)$ Cartesian representation, we show that: (i) if the loci ever cross each other, at that point the R-locus is steeper than the P-locus; (ii) the R-locus lies below the P-locus for sufficiently low mark-up levels; (iii) the R-locus lies above the P-locus when it crosses the 45° line.

Part (i). Note that the system (A.37)-(A.38) only includes $\tilde{\Phi}\left(\mu_p^*, \mu_{hr}^* - \ln y\right)$ and $\tilde{\Phi}\left(\mu_{hr}^*, \mu_p^*\right)$ as explicit first-entry probabilities. With a slight abuse of notation, we let:

$$\tilde{\Phi}_{lp}' \equiv \frac{\partial \tilde{\Phi}\left(\mu_p^*, \mu_{hr}^* - \ln y\right)}{\partial \mu_p^*} > 0, \quad \tilde{\Phi}_{lr}' \equiv \frac{\partial \tilde{\Phi}\left(\mu_p^*, \mu_{hr}^* - \ln y\right)}{\partial \mu_{hr}^*} < 0,$$
$$\tilde{\Phi}_{hp}' \equiv \frac{\partial \tilde{\Phi}\left(\mu_{hr}^*, \mu_p^*\right)}{\partial \mu_p^*} < 0, \quad \tilde{\Phi}_{hr}' \equiv \frac{\partial \tilde{\Phi}\left(\mu_{hr}^*, \mu_p^*\right)}{\partial \mu_{hr}^*} > 0,$$

Furthermore, we also let:

$$\Gamma'_p \equiv \frac{\partial \Gamma\left(\mu_p^*\right)}{\partial \mu_p^*} > 0, \quad \Gamma'_{lr} \equiv \frac{\partial \Gamma\left(\mu_{hr}^* - \ln y\right)}{\partial \mu_{hr}^*} > 0, \quad \Gamma'_{hr} \equiv \frac{\partial \Gamma\left(\mu_{hr}^*\right)}{\partial \mu_{hr}^*} > 0.$$

Differentiating (A.37) and (A.38) yields:

$$\left(\tilde{\Phi}'_{hr} - \tilde{\Phi}'_{lr}\right) d\mu_{hr}^* = \left(\tilde{\Phi}'_{lp} - \tilde{\Phi}'_{hp} + A\right) d\mu_p^*, \tag{A.40}$$

$$\left(\chi_{hr}\left(\mu_{hr}^*\right) \tilde{\Phi}'_{hr} - \tilde{\Phi}'_{lr} + B\right) d\mu_{hr}^* = \left(\tilde{\Phi}'_{lp} - \chi_{hr}\left(\mu_{hr}^*\right) \tilde{\Phi}'_{hp}\right) d\mu_p^*, \tag{A.41}$$

where:

$$\begin{aligned}
A &\equiv \frac{2\left(1-\delta\right)\left(\Gamma'_p + \Gamma'_{lr}\right)}{M} + \frac{\delta}{\rho M \left(\mu_p^* - 1\right)^2} > 0, \\
B &\equiv \frac{2\left(1-\delta\right)\left(\Gamma'_{lr} + \chi_{hr}\left(\mu_{hr}^*\right)\Gamma'_{hr}\right)}{M} \\
&\qquad + \frac{2 - M\tilde{\Phi}\left(\mu_p^*, \mu_{hr}^* - \ln y\right) + 2\left(1-\delta\right)\Gamma\left(\mu_{hr}^* - \ln y\right)}{M\left(\mu_{hr}^* - 1 - \ln y\right)} > 0,
\end{aligned}$$

and in (A.41) we used (A.39) in conjunction with (A.38) to replace:

$$- \frac{\chi\left(\mu_{hr}^*\right)\left[\frac{2(1-\delta)\Gamma\left(\mu_{hr}^*\right)}{M} + \tilde{\Phi}\left(\mu_{hr}^*, \mu_p^*\right) - \frac{\delta}{\rho M}\frac{1}{\mu_{hr}^* - 1}\right]}{\mu_{hr}^* - 1 - \ln y} = \frac{\frac{2}{M} - \tilde{\Phi}\left(\mu_p^*, \mu_{hr}^* - \ln y\right) + \frac{2(1-\delta)\Gamma\left(\mu_{hr}^* - \ln y\right)}{M}}{\mu_{hr}^* - 1 - \ln y}.$$

Rearranging, we have:

$$\left.\frac{d\mu_p^*}{d\mu_{hr}^*}\right|_P = \frac{\tilde{\Phi}'_{hr} - \tilde{\Phi}'_{lr}}{\tilde{\Phi}'_{lp} - \tilde{\Phi}'_{hp} + A} > 0, \tag{A.42}$$

$$\left.\frac{d\mu_p^*}{d\mu_{hr}^*}\right|_R = \frac{\chi\left(\mu_{hr}^*\right)\tilde{\Phi}'_{hr} - \tilde{\Phi}'_{lr} + B}{\tilde{\Phi}'_{lp} - \chi\left(\mu_{hr}^*\right)\tilde{\Phi}'_{hp}} > 0. \tag{A.43}$$

From (A.42) and (A.43), we observe:

$$\left.\frac{d\mu_p^*}{d\mu_{hr}^*}\right|_P < \left.\frac{d\mu_p^*}{d\mu_{hr}^*}\right|_R. \tag{A.44}$$

Suppose $d\mu_p^*/d\mu_{hr}^*\big|_R \leq d\mu_p^*/d\mu_{hr}^*\big|_P$. This inequality requires:

$$\left(\tilde{\Phi}'_{lp} - \tilde{\Phi}'_{hp} + A\right)\left(\chi\left(\mu_{hr}^*\right)\tilde{\Phi}'_{hr} - \tilde{\Phi}'_{lr} + B\right) \leq \left(\tilde{\Phi}'_{lp} - \chi\left(\mu_{hr}^*\right)\tilde{\Phi}'_{hp}\right)\left(\tilde{\Phi}'_{hr} - \tilde{\Phi}'_{lr}\right). \quad \text{(A.45)}$$

Since $A, B > 0$, (A.45) also implies:

$$\left(\tilde{\Phi}'_{lp} - \tilde{\Phi}'_{hp}\right)\left(\chi\left(\mu_{hr}^*\right)\tilde{\Phi}'_{hr} - \tilde{\Phi}'_{lr}\right) < \left(\tilde{\Phi}'_{lp} - \chi\left(\mu_{hr}^*\right)\tilde{\Phi}'_{hp}\right)\left(\tilde{\Phi}'_{hr} - \tilde{\Phi}'_{lr}\right),$$

which in turn boils down to:

$$\tilde{\Phi}'_{hr}\tilde{\Phi}'_{lp} < \tilde{\Phi}'_{hp}\tilde{\Phi}'_{lr}. \quad \text{(A.46)}$$

Recall that:

$$\tilde{\Phi}\left(\mu_{hr}^*, \mu_p^*\right) \equiv \Phi\left(\Lambda\left(\mu_{hr}^*\right) - \Lambda\left(\mu_{hp}^*\right)\right), \quad \text{(A.47)}$$

$$\tilde{\Phi}\left(\mu_p^*, \mu_{hr}^* - \ln y\right) \equiv \Phi\left(\Lambda\left(\mu_{lp}^*\right) - \Lambda\left(\mu_{lr}^*\right)\right). \quad \text{(A.48)}$$

Differentiating (A.47) and (A.48) yield, respectively:

$$\tilde{\Phi}'_{hr} = \Phi'\left(\cdot\right)\Lambda'\left(\mu_{hr}^*\right) \quad \text{and} \quad \tilde{\Phi}'_{hp} = -\Phi'\left(\cdot\right)\Lambda'\left(\mu_{hp}^*\right), \quad \text{(A.49)}$$

$$\tilde{\Phi}'_{lp} = \Phi'\left(\cdot\right)\Lambda'\left(\mu_{lp}^*\right) \quad \text{and} \quad \tilde{\Phi}'_{lr} = -\Phi'\left(\cdot\right)\Lambda'\left(\mu_{hr}^* - \ln y\right). \quad \text{(A.50)}$$

Since $\mu_{hp}^* = \mu_{lp}^* = \mu_p^*$, plugging the expressions in (A.49) and (A.50) into (A.46) leads to the following sufficient condition for (A.44) to hold:

$$\Lambda'\left(\mu_{hr}^*\right) < \Lambda'\left(\mu_{hr}^* - \ln y\right).$$

Recall that, whenever $\mu > 1$, $\Lambda'\left(\mu\right) = \delta^{-1}\left(1 - \mu^{-2}\right) > 0$ and $\Lambda''\left(\mu\right) = 2\delta^{-1}\mu^{-3} > 0$. Hence, it follows that $\Lambda'\left(\mu_{hr}^*\right) \geq \Lambda'\left(\mu_{hr}^* - \ln y\right)$, leading to a contradiction. Therefore, (A.44) must hold true.

Part (ii). Recall that $\tilde{\Phi}\left(\cdot\right)$ is bounded above and below and $\Gamma\left(1\right) = 0$. Let $\mu_p^* = 1$ and denote by $\breve{\mu}_R^*$ the level of μ_{hr}^* that makes (A.38) hold true when $\mu_p^* = 1$. The resulting expression reads:

$$\frac{2}{M} - \tilde{\Phi}\left(1, \breve{\mu}_R^* - \ln y\right) + \frac{2\left(1-\delta\right)}{M}\Gamma\left(\breve{\mu}_R^* - \ln y\right) \\ + \chi\left(\breve{\mu}_R^*\right)\left[\tilde{\Phi}\left(\breve{\mu}_R^*, 1\right) + \frac{2\left(1-\delta\right)}{M}\Gamma\left(\breve{\mu}_R^*\right) - \frac{\delta}{\rho M}\frac{1}{\breve{\mu}_R^* - 1}\right] = 0,$$

thus it must be the case that $\breve{\mu}_R^* > 1$. Let now $\breve{\mu}_P^*$ denote the value of μ_p^* satisfying (A.37) when $\mu_{hr}^* = \breve{\mu}_R^*$, and suppose that $\breve{\mu}_P^* = 1$. From (A.37), it follows that:

$$\tilde{\Phi}\left(1, \breve{\mu}_R^* - \ln y\right) + \frac{2}{M} - \tilde{\Phi}\left(\breve{\mu}_R^*, 1\right) + \frac{2\left(1-\delta\right)}{M}\Gamma\left(\breve{\mu}_R^* - \ln y\right) - \frac{\delta}{\rho M}\lim_{\breve{\mu}_P^* \to 1}\left(\frac{1}{\breve{\mu}_P^* - 1}\right) = 0,$$

where the LHS goes to $-\infty$ as $\breve{\mu}_P^* \to 1$. Thus, it must be that $\breve{\mu}_P^* > 1$ for (A.37) to hold, which implies the P-locus lies above the R-locus for sufficiently low mark-ups.

Part (iii) Recall that $\tilde{\Phi}\left(\mu, \mu\right) = 1/M$, and note that $\chi\left(\infty\right) = 1$. Furthermore, let $\hat{\Gamma}$ denote the upper-bound of $\Gamma\left(\mu\right)$. Firstly, let $\hat{\mu}_R^*$ denote the solution of (A.38) when $\mu_p^* = \mu_{hr}^* = \hat{\mu}_R^*$, and suppose that $\hat{\mu}_R^* \to \infty$. It follows that:

$$\frac{2}{M} + \frac{4\left(1-\delta\right)}{M}\hat{\Gamma} - \frac{\delta}{\rho M}\lim_{\hat{\mu}_R^* \to \infty}\left(\frac{1}{\hat{\mu}_R^* - 1}\right) = 0,$$

where the LHS is strictly positive in the limit. Therefore, it must be that $\hat{\mu}_R^* < \infty$, implying that the R-locus crosses the 45° line. Secondly, let $\hat{\mu}_P^*$ denote the solution of (A.37) when $\mu_p^* = \mu_{hr}^* = \hat{\mu}_P^*$, and suppose that $\hat{\mu}_P^* = \hat{\mu}_R^*$. From (A.37), we have:

$$\tilde{\Phi}\left(\hat{\mu}_R^*, \hat{\mu}_R^* - \ln y\right) + \frac{2\left(1-\delta\right)}{M}\Gamma\left(\hat{\mu}_R^* - \ln y\right) \\ + \frac{1}{M} + \frac{2\left(1-\delta\right)}{M}\Gamma\left(\hat{\mu}_R^*\right) - \frac{\delta}{\rho M}\frac{1}{\hat{\mu}_R^* - 1} = 0. \quad \text{(A.51)}$$

Note from (A.38) that, in the case $\mu_p^* = \mu_{hr}^* = \hat{\mu}_R^*$, we can write:

$$-\frac{\frac{2}{M} - \tilde{\Phi}\left(\hat{\mu}_R^*, \hat{\mu}_R^* - \ln y\right) + \frac{2(1-\delta)}{M}\Gamma\left(\hat{\mu}_R^* - \ln y\right)}{\chi\left(\hat{\mu}_R^*\right)} = \frac{1}{M} + \frac{2\left(1-\delta\right)}{M}\Gamma\left(\hat{\mu}_R^*\right) - \frac{\delta}{\rho M}\frac{1}{\hat{\mu}_R^* - 1}.$$

Plugging this expression into (A.51) yields:

$$\chi\left(\hat{\mu}_R^*\right)\tilde{\Phi}\left(\hat{\mu}_R^*, \hat{\mu}_R^* - \ln y\right) - \left(\frac{2}{M} - \tilde{\Phi}\left(\hat{\mu}_R^*, \hat{\mu}_R^* - \ln y\right)\right)$$
$$+ \left(\chi\left(\hat{\mu}_R^*\right) - 1\right)\frac{2\left(1-\delta\right)}{M}\Gamma\left(\hat{\mu}_R^* - \ln y\right) > 0,$$

Therefore, it must be that $\hat{\mu}_P^* < \hat{\mu}_R^*$, implying that the L-locus lies below the R-locus when the latter crosses the 45° line.

Bearing in mind all the previous three steps, it follows that there must exist one single combination $\left(\mu_{hr}^*, \mu_p^*\right) \in (1, \infty) \times (1, \infty)$, where also $\mu_{hr}^* > \mu_p^*$, satisfying (A.37) and (A.38).

Finally, it must be that $\mu_{lp}^* > \mu_{lr}^*$. To see this, note preliminarily that $\mu_{hr}^* > \mu_{hp}^*$ implies $\tilde{\Phi}\left(\mu_{hr}^*, \mu_{hp}^*\right) > 1/M$, hence from (A.33) and (A.34) it follows that $N_{hr}^* > N_{hp}^*$. Now, suppose that $\mu_{lp}^* \leq \mu_{lr}^*$. We have $\tilde{\Phi}\left(\mu_{lp}^*, \mu_{lr}^*\right) \leq 1/M$, which in conjuction with (A.31) and (A.32) yields $N_{lp}^* \leq N_{lr}^*$. However, (9) requires that $1/\left(N_{lr}^* + N_{hr}^*\right) \leq \left(1 - N_{hr}^* \ln y\right)\left(N_{lp}^* + N_{hp}^*\right)$, contradicting $N_{lp}^* \leq N_{lr}^*$. □

Proof of Lemma 1. Consider a generic country $m \in \mathcal{W}$ with income y. The sets of country m's neighbouring and nearby markets are $\{m_{0p}, m_{0r}\}$ and $\{m_{1p}, m_{1r}\}$, repsectively. Bearing in mind Assumption 4 coupled with (12), it follows that, along the steady state, m will be receiving as secondary market expansions a mass of varieties of quality q equal to $\left[G\left(\Lambda_{yq}^*\right) + \left(G\left(\Lambda_{yq}^*\right)\right)^\lambda\right]\left(\mathcal{P}_{qy} + \mathcal{P}_{qy'}\right)\rho M$, with $y' \neq y$. Hence, since there are $M/2$ richer and $M/2$ poorer markets in $\mathcal{W}$, it follows that the total number of market expansions of varieties of quality q along the steady state in $\mathcal{W}$ is equal to:

$$\left\{\left[G\left(\Lambda_{qr}^*\right) + \left(G\left(\Lambda_{qr}^*\right)\right)^\lambda\right] + \left[G\left(\Lambda_{qp}^*\right) + \left(G\left(\Lambda_{qp}^*\right)\right)^\lambda\right]\right\}\left(\mathcal{P}_{qr} + \mathcal{P}_{qp}\right)\rho M \times \frac{M}{2}. \tag{A.52}$$

Similarly, it follows that the total number of expansions (considering the whole world economy) to nearby (non-neighbouring) markets along the steady state is equal to:

$$\left[\left(G\left(\Lambda_{qr}^*\right)\right)^\lambda + \left(G\left(\Lambda_{qp}^*\right)\right)^\lambda\right]\left(\mathcal{P}_{qr} + \mathcal{P}_{qp}\right)\rho M \times \frac{M}{2}. \tag{A.53}$$

Therefore, dividing (A.53) by (A.52) the result in (15) obtains. □

Proof of Proposition 3. Notice that, since $y_p = 1$ entails $\Lambda^*_{hp} = \Lambda^*_{lp} = \Lambda^*_p$, we may write:

$$S^*_h = \frac{\left(G\left(\Lambda^*_{hr}\right)\right)^\lambda + \left(G\left(\Lambda^*_p\right)\right)^\lambda}{G\left(\Lambda^*_{hr}\right) + G\left(\Lambda^*_p\right) + \left(G\left(\Lambda^*_{hr}\right)\right)^\lambda + \left(G\left(\Lambda^*_p\right)\right)^\lambda}, \tag{A.54}$$

$$S^*_l = \frac{\left(G\left(\Lambda^*_{lr}\right)\right)^\lambda + \left(G\left(\Lambda^*_p\right)\right)^\lambda}{G\left(\Lambda^*_{lr}\right) + G\left(\Lambda^*_p\right) + \left(G\left(\Lambda^*_{lr}\right)\right)^\lambda + \left(G\left(\Lambda^*_p\right)\right)^\lambda}. \tag{A.55}$$

Recall that $G\left(\Lambda^*_{hr}\right) > G\left(\Lambda^*_p\right) > G\left(\Lambda^*_{lr}\right)$ since $\Lambda^*_{hr} > \Lambda^*_p > \Lambda^*_{lr}$.

Note that, letting:

$$\Sigma\left(x\right) \equiv \frac{\left(G\left(x\right)\right)^\lambda + \left(G\left(\Lambda^*_p\right)\right)^\lambda}{G\left(x\right) + G\left(\Lambda^*_p\right)}, \tag{A.56}$$

we may re-express (A.54)-(A.55) as:

$$S^*_h = \left[1 + \left(1/\Sigma\left(\Lambda^*_{hr}\right)\right)\right]^{-1},$$
$$S^*_l = \left[1 + \left(1/\Sigma\left(\Lambda^*_{lr}\right)\right)\right]^{-1}.$$

Therefore, proving $S^*_h > S^*_l$ amounts to proving that $\Sigma\left(\Lambda^*_{hr}\right) > \Sigma\left(\Lambda^*_{lr}\right)$. Let thus $\theta_q \equiv G\left(\Lambda^*_{qr}\right)/G\left(\Lambda^*_p\right)$, note that $\theta_h > 1 > \theta_l$. Noting that $G\left(\Lambda^*_{qr}\right) = \theta_q G\left(\Lambda^*_p\right)$, and plugging the relevant expressions into (A.56) the following condition obtains:

$$\Sigma\left(\Lambda^*_{hr}\right) > \Sigma\left(\Lambda^*_{lr}\right) \quad \Leftrightarrow \quad \frac{\left(1 + \theta^\lambda_h\right)}{\left(1 + \theta_h\right)} > \frac{\left(1 + \theta^\lambda_l\right)}{\left(1 + \theta_l\right)},$$

which indeed holds true because $\lambda > 1$. □

A.2 Additional Data Descriptives

Table A.1 – Data Coverage

Span	Country
01.2009-09.2013	Belgium, Denmark, Estonia, Finland, France, Greece, Italy, Latvia, Lithuania, Netherlands, Portugal, Slovakia, Spain, Sweden, United Kingdom
01.2009-01.2017	Austria, Croatia, Czech Republic, Germany, Hungary, Poland, Serbia, Slovenia

Notes: The table reports data coverage in month-years per country. On average 86% of all products on the EU market for the period 2009-2013 are present in the 8-country sample whose coverage extends to 2017. The eight countries on average account for 49.6% of total expenditure in the data for 2009-2013.

Table A.2 – Description of Refrigerators' Characteristics in Main Data

Characteristics	Description or Values
Annual energy use	Annual energy consumption measured in kilowatt hours per year. Mean 272 kWh; s.e. 95.7 kWh
Energy label	A+++ (most efficient) (4.1%); A++ (23.8%); A+ (50.2%); A (21.0%); B (0.86%); C (0.02%)
No-frost system	Yes (35.9%); No (64.1%)
Brand	Liebherr (11.5%); Bosch (9.6%); Whirlpool (8.1%), Gorenje (7.5%); Siemens (7.4%); Electrolux (6.4%), Samsung (5.5%), others (44%)
Number of doors	One door >90cm (20.6%); 2 doors, freezer bottom (56.6%); 2 doors, freezer top (13.1%); 3+ doors (2.2%); side-by-side (7.5%)
Installation	Built-in/built-under (23.2%); freestanding (76.8%)

Notes: The data lists the available refrigerator characteristics in the data. For categorical variables, the percent of each value from total observations is reported in parentheses based on the sample reported in Panel A of Table 1.

Table A.3 – Description of Variables in Supplemental Data

Variable	Description and Statistics
Energy	Energy prices per kWh (Euro) for household consumers with consumption from 2500-4999 kWh, all taxes and levies included. Frequency: bi-annual. Coverage: h12009-h12017. Variation: Country-by-half-year. Source: Eurostat. Mean: 0.168 s.dev.: 0.054. Min: 0.056 Max: 0.305. N: 408.
Income	GDP per capita (Euro). Frequency: yearly. Coverage 2009-2017. Variation: Country-by-year. Source: Penn World Tables. Mean: 32,316 s.dev.: 9,792. Min: 12,108 Max: 51,524 N: 216.
Pop	Population (in millions). Frequency: yearly. Coverage 2009-2017. Variation: Country-by-year. Source: Penn World Tables. Mean: 20.9 s.dev.: 23.8. Min: 1.31 Max: 82.1 N: 216.
Retail	Index of turnover for retail trade, except for motor vehicles and motorcycles (2010=100). Frequency: Monthly. Coverage: January 2009-January 2017. Variation: Country-by-date(month-year). Source: Eurostat. Mean: 105.1 s.dev.: 15.21. Min: 60.6 Max: 174.8 N: 2,328.
VAT rate	Standard value-added tax rate. Frequency: monthly. Coverage 2009-2017. Variation: Country-by-date. Source: European Commission. Mean: 0.21 s.dev.: 0.026. Min: 0.15 Max: 0.27 N: 2,328.
MS Brand	Ratio of total unit sales of a brand in a given country on a given date to total sales within this country-date. Frequency: monthly. Coverage: January 2009-January 2017. Variation: Brand-by-country-by-date. Source: GfK GmbH. Mean: 0.037 s.dev.: 0.059. Min: 0.00 Max: 0.57 N: 42,963.
Herfindahl-Hirschman Index (HHI)	Mean of sum of MS Brand squared across brands within a country-date. Frequency: monthly. Coverage: January 2009-January 2017. Variation: Country-by-date. Source: GfK GmbH. Mean: 0.14 s.dev.: 0.054. Min: 0.065 Max: 0.39 N: 1,688.

Notes: The table describes the additional variables added to the main GfK data, their frequency, coverage, variation, sources, and basic descriptive statistics, including the number of unique observations (N).

A.2.1 Market presence by annual cohorts and product's average life-cycles

Figure A.2 reveals the annual-cohort-specific market composition as separate annual-cohorts' shares (based on EU-wide entry years, which we observe even for already existing products in 2009 when the data starts) from the total number of products on offer in a given year. Focusing on 2009, the figure reveals that appliances in that year are a mixture of annual cohorts 2004-2009, with the share of newly introduced products amounting to 19.8%. The top sub-bar on each bar indicates that throughout 2009-2016, the annual share of new market entrants remains stable at close to 20%. To understand the cohort-specific extent of market misidentification in the generation of market sequences, we need a reasonably accurate measure of average life cycles. While it is difficult to provide descriptive statistics on the average product life-span given the more limited country coverage from 2014 onwards and the end of the data set in January 2017, based on annual-cohort 2009, which has the longest presence in multiple markets, products remain

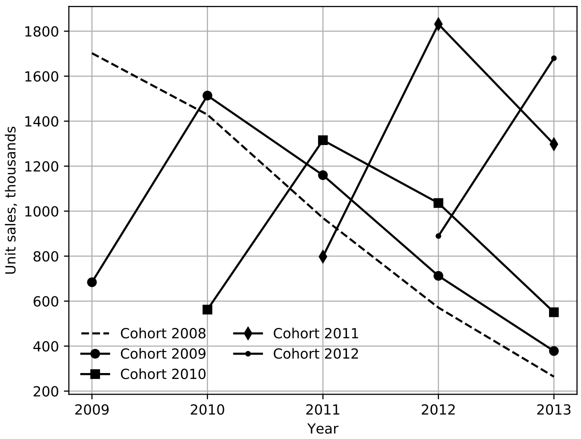
Table A.4 – Descriptive Statistics: Full Sample

	All		By quality quartile							
			(1)		(2)		(3)		(4)	
	Mean	Mdn	Mean	Mdn	Mean	Mdn	Mean	Mdn	Mean	Mdn
Price	682.5	542.3	356.7	329.5	517.1	473.2	702.0	640.4	1,213.3	1060.0
	(492.8)		(128.6)		(194.2)		(290.6)		(677.1)	
N	926,183		244,258		240,470		224,912		216,543	
Units	35.3		48.09		36.8		31.1		23.9	
	(110.3)		(148.3)		(106.8)		(92.9)		(74.8)	
N	1,041,854		269,724		270,857		253,164		248,109	
Quality	0.128		-0.379		-0.023		0.282		0.779	
	(0.453)		(0.155)		(0.091)		(0.091)		(0.267)	
N	11,547		3,167		3,233		2,524		2,623	

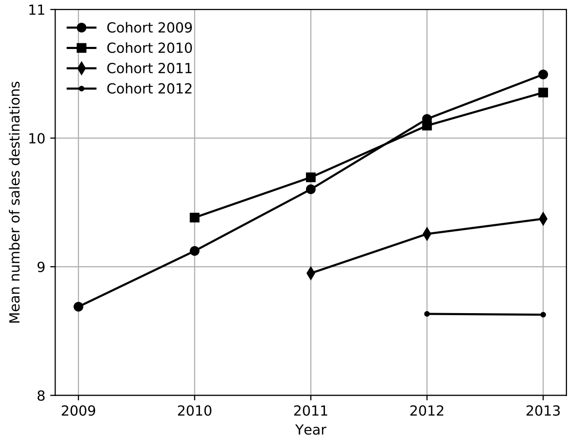
Notes: The table shows descriptive statistics per product per date per country for the full sample spanning January 2009 to January 2017. The following basic cleaning of the data has been performed: zero or negative prices are replaced with missing observations; negative unit sales are replaced with missing observations. 'Quality' is the time-invariant quality index constructed from the hedonic specification (1). Columns (1)-(4) report statistics for four quantiles of the quality index. 'N' denotes the number of observations. 'Mdn' abbreviates median value. All prices are in Euro.

within a country for on average 3.75 years (s.e. 1.25 years), and EU-wide – for 5.5 years. Conducting the same exercise with the sub-sample of eight countries spanning 2009-2017, the country-specific life-cycle is 3.48 years (s.e 2.15), and the sub-sample-wide – 5.57 (2.29 years), both values very close to the ones obtained with the 2009-annual-cohort alone. Despite the presence of multiple annual cohorts per year as indicated in Figure A.2, Figure A.3 shows that in terms of sales, invariably the four newest annual cohorts account for more than 80% of the market share per year, such that with respect to meaningful market shares, EU life-spans appear to not exceed 4 years. Given that country coverage in the data is close to complete for the EU in the period 2009-2013, a global life-cycle of about 4 years would indicate that market sequences are most accurately recovered for annual cohorts 2009-2010 and least accurately for cohort 2013. For annual cohort 2013 only the first market of entry is ensured to be correctly identified as the data's country coverage reduces to 8 countries in 2014. Thus, the second market of entry in the data for this cohort will not necessarily match the actual second market of entry, which may not be in the data.

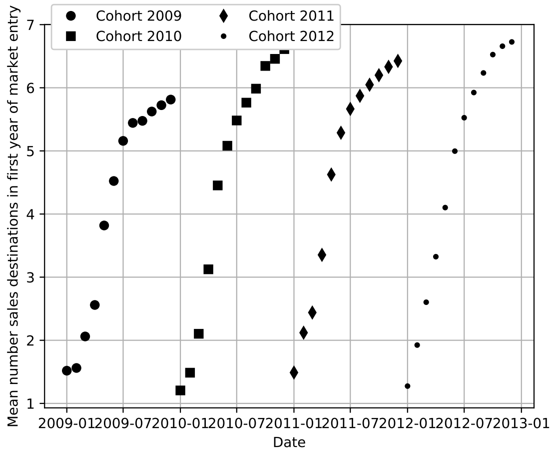
Figure A.1 – Cohort-specific Sales and Sales Destinations



(a) aggregate unit sales over life-cycle



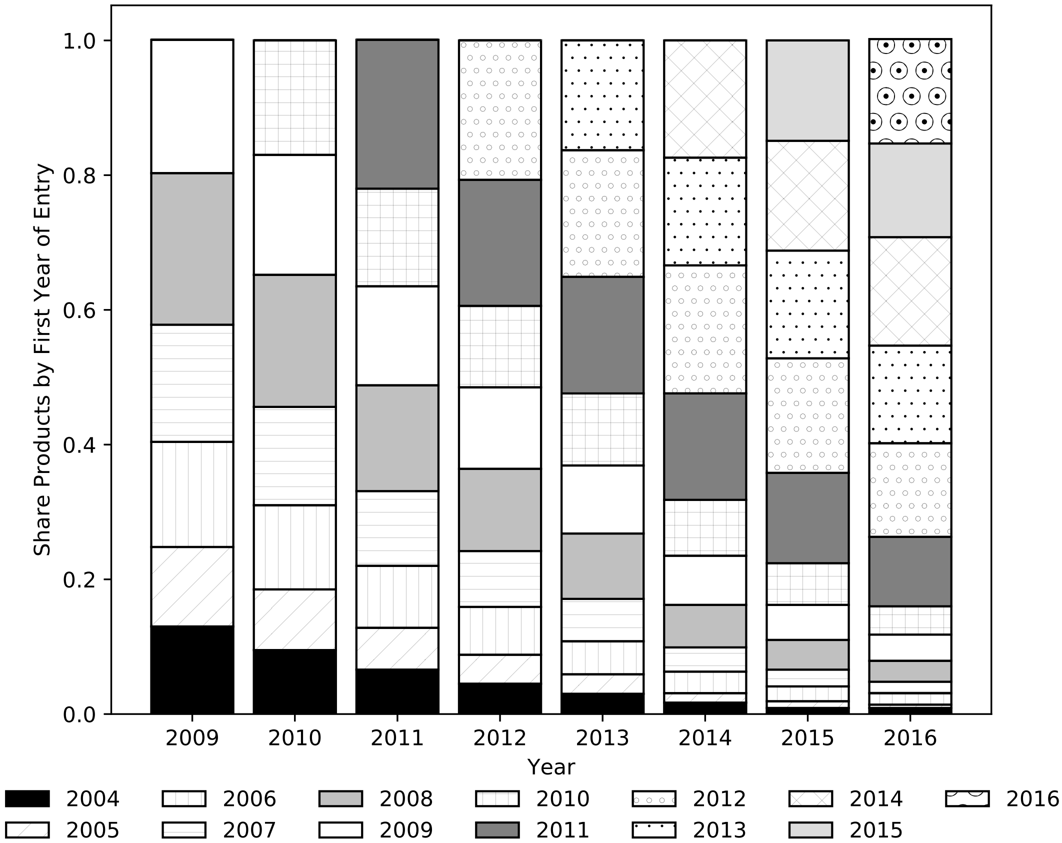
(b) sales destinations over life-cycle



(c) sales destinations by date within first year

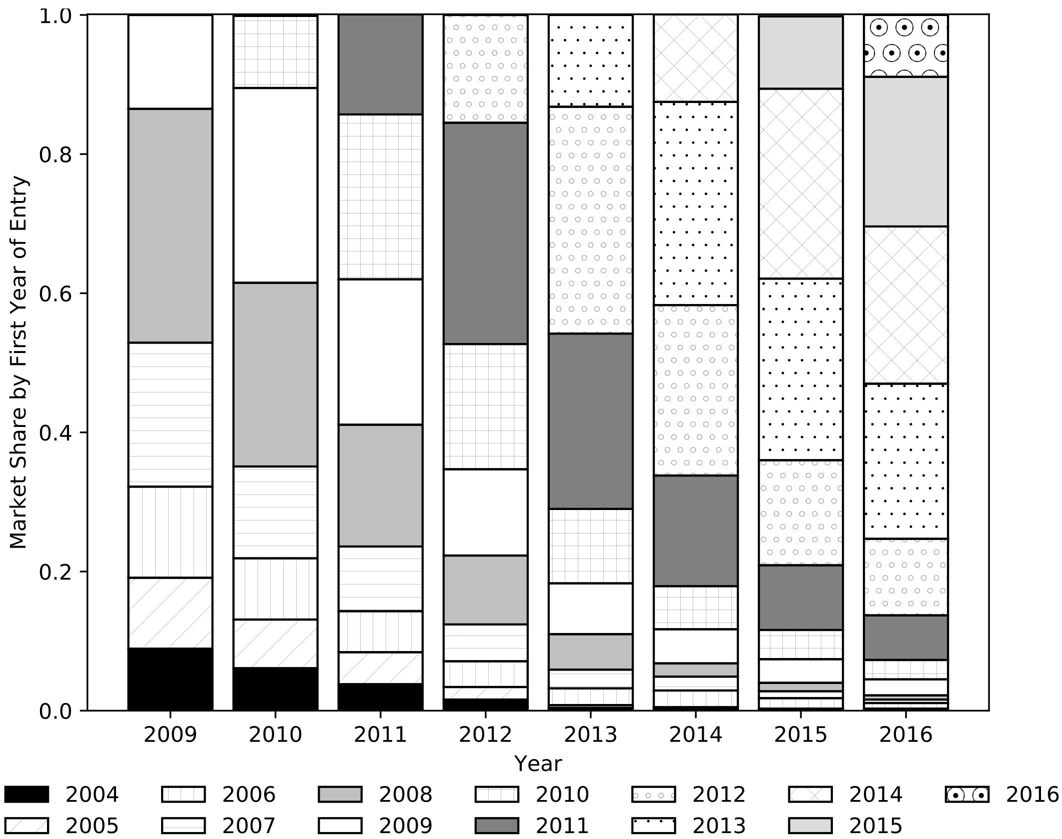
Notes: Plot (a) shows yearly sales aggregated across 24 countries (sales destinations) by cohort of products entering in 2009, 2010, 2011, or 2012 over the period 2009-2013 due to different country coverage from 2014 onward (see Table A.1). Plot (b) shows the cohort-specific average number of sales destinations by year for the same period. Plot (c) focuses solely on sale-destination-entry by cohort by date within the *first* year of market entry.

Figure A.2 – Product composition by annual-cohorts



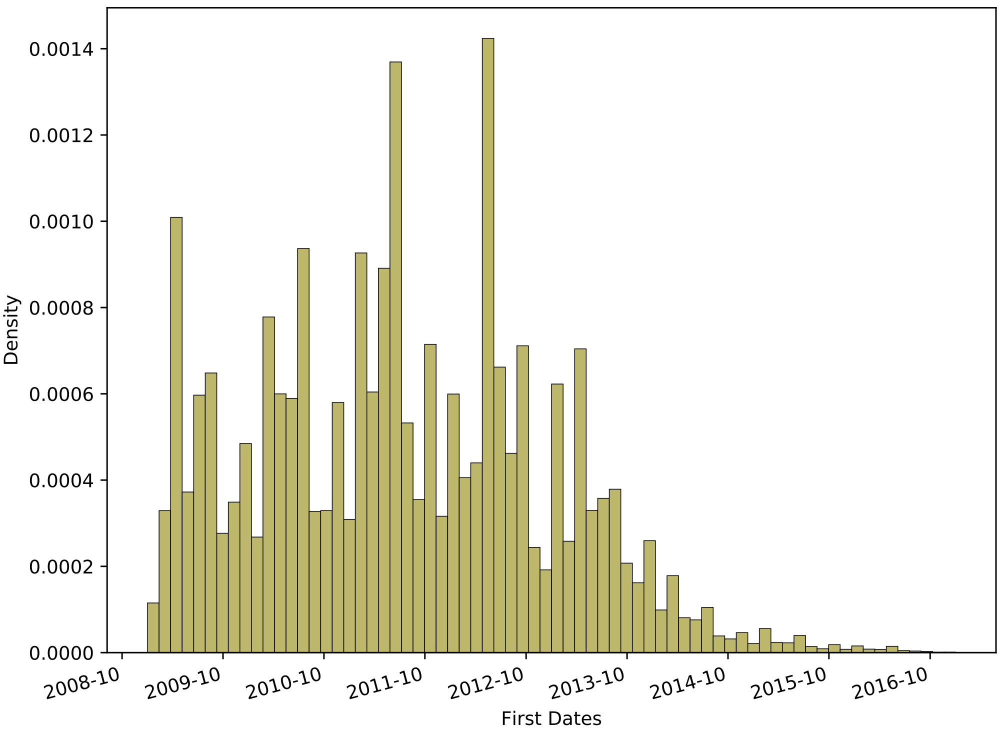
Notes: The bar chart shows the share of products by cohorts of first-year of market entry within a year, spanning 2009-2016. Thus, in 2009, the share of products that entered the EU market in 2005 and are still sold in 2009 is 0.118, those that entered in 2006 – 0.156, 2007 – 0.174, 2008 – 0.225, and newly introduced products in 2009 account for 0.198 of all products on the market in that year. Until 2013, the total number of products and cohort-specific shares are based on 24 countries, and from 2014 onwards – on 8 EU countries. First-year of market entry is the first year in which a product appears on the EU market anywhere. The first year is truncated at 2004, such that the shares of products over time that enter the market in 2004 are likely also capturing products introduced before 2004. Within each bar, sub-bars are stacked in such a way that the share of new market entrants in year y is always on top, followed immediately by cohort $y-1$, $y-2$, etc.

FIGURE A.3 – Market share composition by annual-cohorts



Notes: The bar chart shows the market share by cohorts of first-year of market entry within a year, spanning 2009-2016. Thus, in 2009, the market share of products that entered the EU market in 2005 and are still sold in 2009 is 0.10, those that entered in 2006 – 0.13, 2007 – 0.21, 2008 – 0.34, and newly introduced products in 2009 account for 0.14 of all unit sales on the market in that year. Until 2013, unit sales and cohort-specific market shares are based on 24 EU countries, and from 2014 onwards – on 8 EU countries. First-year of market entry is the first year in which a product appears on the EU market anywhere. The first year is truncated at 2004, such that the market share of products over time that enter the market in 2004 is likely also capturing that of products introduced prior to 2004. Within each bar, sub-bars are stacked in such a way that the market share of new market entrants in year y is always on top, followed immediately by the market share of cohort $y-1$, $y-2$, etc.

FIGURE A.4 – Country-specific Entry Dates



Notes: The figure plots the country-specific first dates (month-year combination) of a product's entry in a given market. These are determined by finding a) the first year, in which a product appears in a given country, and b) the first month within the first year, in which units sold are not missing and not zero. The plot pertains to all products with 'global' first dates between 2009-2013. First dates plotted for 2014-2016 therefore capture such products' first dates in the second, third, etc. markets.

A.3 Additional Empirical Results and Robustness Checks

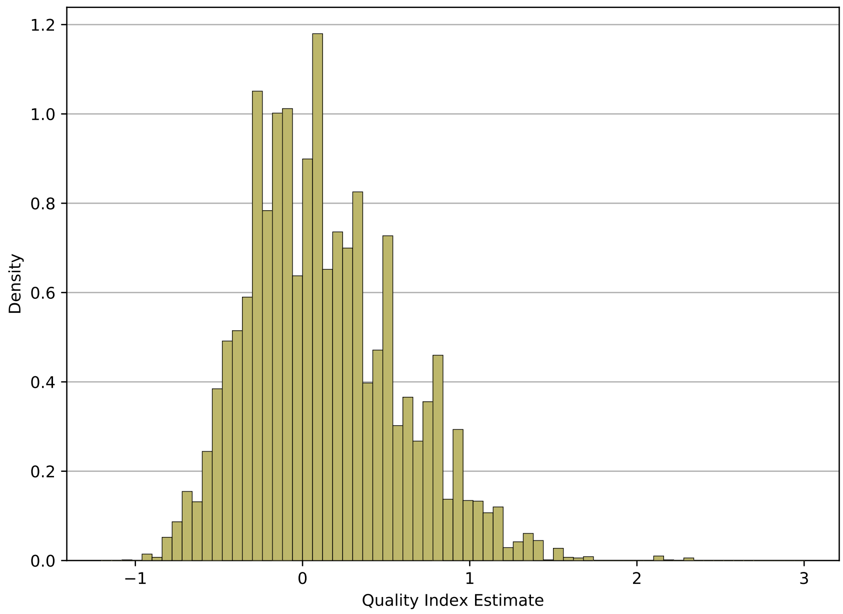
Table A.5 – Quality Index Generation

	Coef. (1)	S.e. (2)
noforst		
yes	0.318***	0.011
doors		
2 doors frz btm	-0.031**	0.012
2 doors frz top	-0.200***	0.016
3+ doors	0.717***	0.049
side by side	0.690***	0.019
energy label		
A+	0.148***	0.010
A++	0.305***	0.013
A+++	0.511***	0.016
B	-0.036	0.029
C	0.183	0.193
brand		
2.b	-0.650***	0.027
3.b	-0.342***	0.039
4.b	-0.204***	0.023
5.b	-0.502***	0.021
6.b	-0.236***	0.039
7.b	-0.741***	0.030
8.b	-0.036*	0.021
9.b	-0.468***	0.023
10.b	-0.160***	0.040
11.b	-0.468***	0.031
12.b	-0.266***	0.022
13.b	-0.113***	0.031
14.b	-0.710***	0.026
15.b	-0.258***	0.035
16.b	1.126***	0.110
17.b	-0.218***	0.023
18.b	-0.709***	0.053
19.b	-0.377***	0.026
20.b	-0.318***	0.028
21.b	-0.467***	0.021
22.b	0.079*	0.046
23.b	-0.599***	0.042
24.b	0.352***	0.045

	Coef. (1)	S.e. (2)
25.b	-0.245***	0.025
26.b	0.220***	0.026
27.b	0.512***	0.036
28.b	0.191***	0.029
29.b	-0.125***	0.034
30.b	-0.852***	0.033
31.b	-0.331***	0.031
32.b	-0.591***	0.045
33.b	-0.318***	0.025
34.b	-0.347***	0.055
35.b	-0.524***	0.030
36.b	0.004	0.053
37.b	0.041*	0.024
38.b	0.495***	0.055
39.b	-0.503***	0.025
40.b	-0.396***	0.031
41.b	-0.266***	0.021
42.b	-0.419***	0.025
N	926,183	

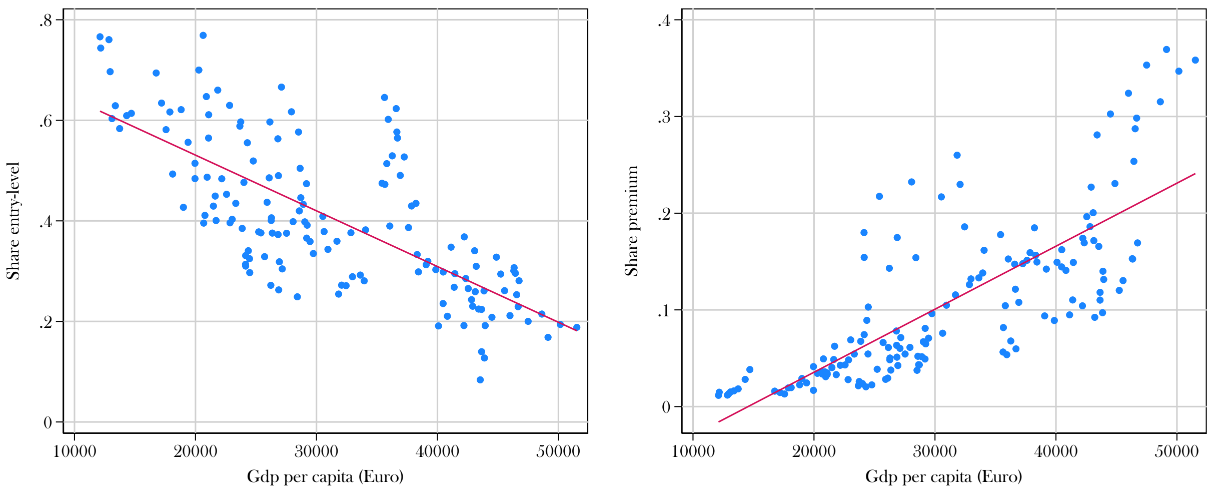
Notes: The table shows the estimated coefficients from eq. (1) used in the calculation of the product-specific quality index. The excluded type of refrigerator belongs to brand AEG, has an energy class A, no 'nofrost' function, and one door greater than 90 cm. See a detailed description of these variables in Table A.2. Further information on the 41 brands shown in the table is available upon request from the authors. Standard errors are robust and clustered by product. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure A.5 – Quality Index Distribution



Notes: Histogram of the estimated quality index for each product in the data based on eq. (2).

Figure A.6 – Premium and Entry-level Unit-sale-shares by Income per Capita



(a) Entry-level

(b) Premium

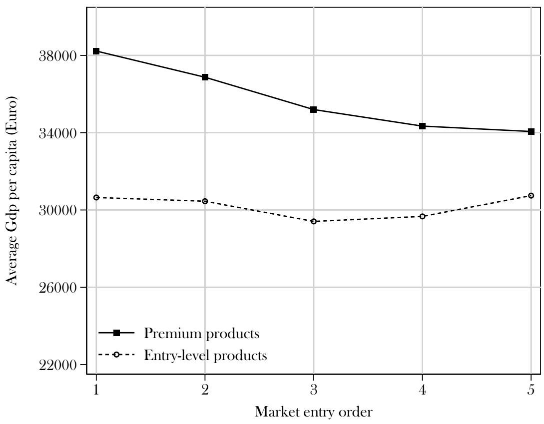
Notes: The figure plots country-specific yearly shares of the unit sales of entry-level products from total unit sales in (a) and of premium products (unit sales of premium-quality products over total unit sales per year) in (b) vis-a-vis GDP per capita. Entry-level products are those in quantile one, and premium products – in quantile four of the product-specific quality estimates obtained from eq. (2). The lines in both graphs are linear prediction plots.

Table A.6 – Pricing-to-market: Heterogeneity across Quality

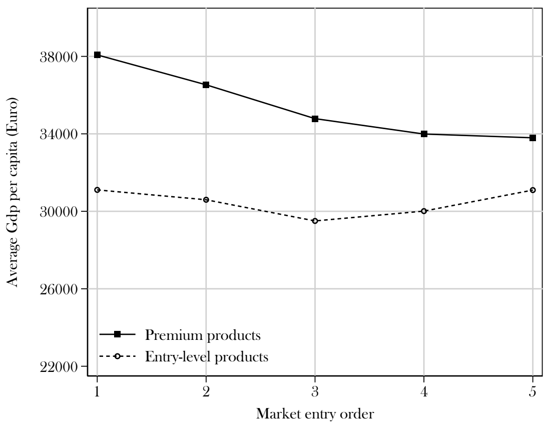
	(1)	(2)	(3)	(4)	(5)	(6)
	Products of above-median quality			Products of below-median quality		
ln Income	-0.192**	-0.188**		-0.099	-0.097	
	(0.080)	(0.080)		(0.099)	(0.099)	
ln Income×$\hat{q}_i$	0.421***	0.413***	0.327***	0.035	0.036	0.059
	(0.119)	(0.120)	(0.108)	(0.113)	(0.112)	(0.103)
ln Pop	0.093	0.088		-0.597	-0.597	
	(0.388)	(0.385)		(0.416)	(0.414)	
ln MS Brand	-0.006**			-0.006***		
	(0.003)			(0.002)		
ln Retail	0.006	0.007		0.011	0.011	
	(0.019)	(0.019)		(0.023)	(0.023)	
ln Energy	0.022	0.022		0.097**	0.099**	
	(0.044)	(0.044)		(0.040)	(0.041)	
HHI	0.078	-0.040		-0.101	-0.145	
	(0.153)	(0.136)		(0.107)	(0.113)	
HHI×$\hat{q}_i$		0.269**	0.078		-0.208	-0.269
		(0.124)	(0.086)		(0.200)	(0.173)
δ_{ic}	Yes	Yes	Yes	Yes	Yes	Yes
θ_{cd}			Yes			Yes
N	388,459	388,459	388,459	389,774	389,774	389,774

Notes: The table shows results from the estimation of eq. (3). All specifications include product-by-date fixed effects, not reported. Specifications (1)-(2) and (4)-(5) control for product-by-country indicators, while (3) and (6) further incorporate country-by-date effects. The sample pertains to products of above-median quality in columns (1)-(3), and to products of below-median quality in (4)-(6). Ln Pop, ln Retail, ln MS Brand, and ln Energy are the natural logarithms of population, a retail turnover index, brand market share, and bi-annual household energy prices. HHI is the Herfindahl-Hirschman index of market concentration. See Table A.3 for a detailed description of these variables and summary statistics. Standard errors are robust, and two-way clustered by product and by country throughout. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure A.7 – Order of Market Entry vis-a-vis Income for Products with at Least 5 Destinations



(a) all products sold in at least five markets



(b) all products sold in at least five above- and below-median income countries

Notes: The figure replicates Plots (a) and (b) in Figure 3 in the main text, but reduces the sample only to products with at least 5 sales destinations (1,998 devices). In this figure, therefore, the composition of products per market-entry order is the same. Plot (a) depicts average income per capita by products' order of market entry (first, second, third, etc. market) and quality (premium products (quantile four) shown as solid line, and entry-level products (quantile one) – as dashed line) for all products. Plot (b) depicts the same relationship but pertains to products sold both in above-median and below-median income countries. The order of market entry is determined as per Table A.9. Quality quantiles are based on the full set of products. For comparability, the range of the y-axis is held constant across plots.

Table A.7 – Determinants of First Market Entry (Most Preferred Alternative): Additional Covariates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log Income	1.458**	1.565**	1.296**	0.254	1.163	2.378**	0.373
	(0.646)	(0.763)	(0.640)	(0.752)	(0.815)	(0.945)	(0.978)
ln Income$\times\hat{q}$	1.814***	1.691***	1.924***	1.780***	1.369***	3.184***	1.425***
	(0.564)	(0.569)	(0.584)	(0.560)	(0.409)	(0.923)	(0.489)
ln Pop		3.492					5.074
		(7.859)					(6.674)
ln Pop$\times\hat{q}$		0.097					0.095
		(0.090)					(0.062)
ln Energy			-0.859				-0.528
			(0.601)				(0.562)
ln Energy$\times\hat{q}$			-0.146				-0.282
			(0.452)				(0.337)
ln Retail				1.721***			1.486***
				(0.516)			(0.535)
ln Retail$\times\hat{q}$				-2.149**			-2.692***
				(0.983)			(0.935)
MS brand					9.655***		9.635***
					(0.994)		(0.978)
MS brand$\times\hat{q}$					-0.120		-0.113
					(1.473)		(1.409)
ln Distance						-0.524**	
						(0.208)	
ln Distance$\times\hat{q}$						0.223	
						(0.283)	
N	125,088	125,088	125,088	125,088	99,376	20,496	99,376

Notes: The method of estimation is conditional logit. The dependent variable equals one for the first market(s) of entry and is set to zero for the remaining countries in a set of 24 possible European destinations. The first market(s) are determined by the earliest date a product appears in any location. All specifications include country fixed effects, not reported. For convenience, the baseline specification without covariates is reported in column (1). ln Income, ln Pop, ln Retail and ln Energy are the natural logarithms of annual GDP per capita (in PPP, Penn World Tables), annual population (Penn World Tables), monthly index of turnover for retail trade (except for motor vehicles and motorcycles) sourced from Eurostat, and half-yearly electricity prices for household consumers, inclusive of all taxes and fees for consumption band from 2 500 kWh to 4 999 kWh, sourced from Eurostat. "MS brand" is a country-date-brand-specific market share. $\hat{q}$ is the product-specific quality index from eq. (1). Standard errors are clustered by brand throughout. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

TABLE A.8 – Determinants of Earlier Market Entry: Additional Covariates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
ln Income	0.777**	0.340	0.738**	0.330	-0.040	0.084	-0.797**
	(0.335)	(0.390)	(0.299)	(0.335)	(0.280)	(0.674)	(0.316)
ln Income×$\hat{q}$	1.563***	1.373***	1.447***	1.530***	1.703***	2.454***	1.380***
	(0.400)	(0.357)	(0.355)	(0.391)	(0.323)	(0.648)	(0.334)
ln Pop		-5.837***					-6.151***
		(2.223)					(1.273)
ln Pop×$\hat{q}$		0.197***					0.236***
		(0.067)					(0.081)
ln Energy			-0.737***				-0.057
			(0.214)				(0.248)
ln Energy×$\hat{q}$			0.158				0.102
			(0.193)				(0.246)
ln Retail				0.673***			0.440*
				(0.203)			(0.246)
ln Retail×$\hat{q}$				-0.901**			-0.123
				(0.384)			(0.546)
MS brand					6.401***		6.242***
					(1.033)		(0.991)
MS brand×$\hat{q}$					-0.168		0.245
					(1.329)		(1.294)
ln Distance						-0.296***	
						(0.094)	
ln Distance×$\hat{q}$						0.120	
						(0.110)	
N	125,088	125,088	125,088	125,088	50,160	20,496	50,160

Notes: The method of estimation is a ranked-orderd logit. The dependent variable is Rank, which gives the highest value to the first market of entry of product j up to a value of zero for any of the 24 EU markets in the sample, in which product j never enters. See also Table A.9 for further discussion on the construction of the dependent variable. All specifications include country and year fixed effects, not reported. For convenience, the baseline specification without additional covariates is shown in column (1). Log Income, log Pop, log Retail and log Energy are the natural logarithms of annual GDP per capita (in PPP, Penn World Tables), annual population (Penn World Tables), monthly index of turnover for retail trade (except for motor vehicles and motorcycles) sourced from Eurostat, and half-yearly electricity prices for household consumers, inclusive of all taxes and fees for consumption band from 2 500 kWh to 4 999 kWh, sourced from Eurostat. "Market share brand" is a country-date-brand-specific market share. "Quality" is the product-specific quality measure estimated in Section 2.2. Standard errors are robust and clustered by brand throughout. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.9 – Market Entry and Ranking

Market	First date	Market entry order	Rank	Market entry ties	Rank with with ties
1st Market$_j$	$\tilde{d}_j$	1	24	1	24
2nd Market$_j$	$\tilde{d}_j$+entry_lag$_{12}$	2	23	1	24
3rd Market$_j$	$\tilde{d}_j$ +entry_lag$_{13}$	3	22	2	23
...	...	...	...	...	...
...	...	...	...	...	...
...	...	...	...	...	...
nth Market$_j$	$\tilde{d}_j$ +entry_lag$_{1n}$	n_j	$24 - n_j + 1$	$n_j - \tilde{n}_j$	24- $(n_j - \tilde{n}_j + 1)$
Other markets$_j$	–	–	0	–	0

Notes: The table shows how the market entry sequence of individual products over their life cycle is determined by identifying country-specific first dates. First-date is the first month-year in which product j has non-zero sales in country m. An 'EU-wide' first-date, $\tilde{d}_j$, is defined as the first time product j is introduced anywhere in the EU. entry_lag is the monthly difference between the first and any sequential markets. Thus entry_lag$_{13}$ is the monthly difference of the first dates in the first and third markets of entry. The series of country-specific first dates directly translates into market-entry order. We construct the Rank variable by assigning a value of 24 to the first market(s) reflecting the total number of countries in the data, 23 to the second, and so on until the last market of entry, which is given a rank of $24 - n_j + 1$. In many instances, products enter several markets simultaneously. In the example shown in the fifth column, the 1st and 2nd markets are contemporaneous such that entry_lag$_{12}$ = 0. In this case, the market entry order is the same, and so are the assigned ranks. The total number of markets n_j for products with tied entry are therefore adjusted down by the number of markets with tied entry $\tilde{n}_j$. For example, if four countries are entered as third markets, then this tie has three duplicates. The assumption is that firms value ties equally. Other markets$_j$ are those markets out of the 24, in which product j is never sold. All of these markets are assigned a rank of zero.