Manufacturers and Retailers in the Global Economy\(^1\)

**Horst Raff**  
Department of Economics  
and Kiel Institute for the World Economy  
Christian-Albrechts-Universität zu Kiel  
24098 Kiel, Germany  
Email: raff@econ-theory.uni-kiel.de

**Nicolas Schmitt**  
Department of Economics  
8888 University Drive  
Simon Fraser University  
Burnaby BC, V5A 1S6, Canada  
Email: schmitt@sfu.ca

March 2015

\(^1\)We would like to thank a co-editor, two anonymous referees, and participants at various workshops and conferences for helpful comments and the Social Sciences and Humanities Research Council of Canada for financial support.
Abstract

We develop a model with multi-product retailers acting as intermediaries between manufacturers and consumers. We show that the rise in retailer product assortment, the rise of up-front payments in many retail markets and the observed shift in employment from manufacturing to retailing may be the consequence of the global integration of product markets. We also identify a novel benefit from market integration consisting of efficiency gains in the vertical distribution chain.

JEL classification: F12, F15, L13

Keywords: international trade, market integration, retailing, multi-product firm, intermediation, up-front payment
1 Introduction

This paper develops a model with retailers acting as intermediaries between manufacturers and consumers. It has two main purposes. The first one is to investigate to what extent three important changes in the interaction between retailers and manufacturers that have taken place in the industrialized countries in the past decades can be explained by international market integration. These changes are (i) the shift in employment from manufacturing to retailing, (ii) the increase in retailer product assortment, and (iii) the emergence and subsequent increase in up-front payments by manufacturers to retailers. The second purpose is to examine the welfare impact of international market integration in a world in which retailers control access to consumers and thus determine the variety of goods available to consumers.

Consider the structural changes that have taken place in manufacturing and retailing over the last 40 years. First, there has been a fundamental increase in the importance of services in general and of wholesale and retail trade in particular. In the United States, for instance, this shift took place especially strongly from the end of the 1970s and it took place at the expense of manufacturing. Simply put, US employment fell in manufacturing between 1970 and 1990, but rose by 71% in wholesale and retail trade (see Blum, 2008). In 2008 retailing alone was the second largest industry in the US in terms of employment (11% of total employment, a higher share than in manufacturing; US Bureau of Labor Statistics, 2009) and accounted for $3.9 trillion in annual sales. Interestingly, this shift in employment can also be observed among the US wholesaling/retailing and manufacturing firms that engage in international trade.

Second, retailers have on average always carried a large variety of products, but their assortment has risen significantly over the last 30 years. Ac-
According to Quelch and Kenny (1994), the number of consumer-packaged-goods stock-keeping units (SKUs) grew 16% each year between 1985 and 1992. Grocery retailing is just one example in this respect, but a revealing one. In the United States this sector is dominated by supermarkets (i.e. stores with sales in excess of $2 million annually). In 2008, supermarkets sold on average 46,852 different items, up from an average of 35,000 items in 2002.

Third, up-front payments, which are lump-sum payments made by manufacturers to retailers, have become an important feature of retailing used in a variety of retail sectors ranging from grocery, apparel and footwear, to toys and games, automotive parts and e-commerce (see Sudhir and Rao 2006; Wilkie, Desrochers and Gundlach 2002; Bone, France and Riley 2006). Because of their proprietary nature, it is very difficult to have precise and consistent information about these payments whether across sectors or through time. However, the general view is that slotting allowances appeared first in 1984 in the US grocery sector (OECD, 2006, p13; Sullivan, 1997, p461) and that their use has quickly spread to other retail sectors. Today they are sufficiently widespread and common for the EU Commission (2009) to write that ‘slotting allowances are currently perceived by the majority of operators as a requirement of modern trade and as an intrinsic characteristic of the business model of modern retailers’ (Section 4.2.3).

---

4 In 2002, the sales of supermarkets represented 77% of all US grocery sales for a total sale value of $547.1 billion and they collectively employed 3.2 million workers; see www.fmi.org.

5 Upfront payments are fixed fees paid by manufacturers to retailers ostensibly to obtain access to shelf space, defray upfront costs, and support downstream promotional activities. The term is descriptive of when these payments are actually made, that is, at the time the contract is signed and/or at the beginning of each year if the length of the contract spans several years. Slotting allowances belong to this class of payments, as do so-called listing fees, pay-to-stay fees, and street money (Marx and Shaffer, 2007, p823).

6 Evidence also suggests that upfront payments have increased over time. Wilkie et al. (2002) for instance write that their survey results ‘clearly indicate an increase in slotting over the past decade in supermarkets.’ (p281). Citing a 2000 ACNielsen report, Wilkie et al. (2002) indicate that ‘85% of retailers reported charging slotting fees. Moreover, 42% of manufacturers reported that they were charged increased slotting allowances, whereas 28% of retailers reported having increased their slotting fees’. See also Hogsett (2006) about increases in up-front payments in the textile and apparel sector. Evidence also suggests that foreign firms make these payments, and that retailers in developing countries use them. On the first point, Yeung (2001) reports that ‘Vitasoy had paid to stores about HK$7 million as a "slotting allowance" to launch its natural soymilk in leading coastal
What has prompted these structural changes and especially the emergence of up-front payments in the 1980s and their subsequent widespread use across retailing? Up-front payments are generally explained by the fact that retailers became increasingly powerful gatekeepers during the 1980s because shelf space is scarce and the number of product varieties rose significantly around that time (see Sullivan, 1997). We suggest that trade liberalization has contributed to the explosion of product variety in a number of sectors and made the competition for shelf-space much more intense than it would otherwise have been. In fact, retailers have themselves contributed to this bottleneck whether it is through the introduction of private labels and/or foreign sourcing of goods (see Feenstra and Hamilton, 2006).

A simple way to illustrate this is to consider US import penetration of consumer products during the 1980s. From 1982 onward, import penetration of these products grew at more than 12% per year (12.4% in 1982-86 and 16% in 1987-91), whereas this rate was 8.6% per year in 1977-81. Even by using a very conservative method we estimate the number of new imported consumer good varieties to have grown on average over 2.5 times faster in 1982-86 as compared to 1977-81 (4.5% per year against 1.8%).

This is of course only suggestive as this does not provide a direct link between imports and up-front payments. But consider two examples: the textile and the athletic footwear markets. According to Hogsett (2001), the textile industry had to deal with negotiated slotting fees and other retail fees ‘for decades’ with suppliers complaining that ‘retailers ... deduct almost 10 percent of the price in the form of chargebacks and deductions’. It is also an industry where import penetration is generally high, reaching as much as 97.7% for men’s and boys’ sweaters in 1998.

This cities in the US.

7Based on own computations. Import penetration is measured by the ratio of imports to shipments at the 4-digit SIC level and growth in the number of new varieties is measured by the change in SIC-product/country-of-origin pairs (see Broda and Weinstein, 2006, for additional evidence).

8Import penetration is measured here as the ratio of imports to apparent consumption (production plus imports minus exports). Data from http://www.allcountries.org/uscensus/1246_selected_apparel_shipments_foreign_trade_and.html
footwear industry is even more extreme. In 1999, import penetration in the
US for footwear with rubber or plastic uppers and rubber or plastic soles was
98.6% (94.3% for footwear in general).\footnote{From http://www.allcountries.org/uscensus/1247_footwear_production_foreign_trade_and_apparent.html} Here, too, slotting fees have been
used. Cassidy (2001) reports in particular that the Venator Group, owner
of the Foot Locker and Champs Sports chains, had a 19% market share and
was a leader regarding the use of slotting allowances in this industry.

Thus, in several segments of the retail market, the proliferation of prod-
ucts and the use of up-front payments would hardly have been as extensive
without foreign sourcing. Multi-product retailers are then powerful simply
because they typically have little to lose by not selling a particular variety.
This has given incentives to retailers in several sectors to use up-front pay-
ments. But it is also true that they are not used by all retailers in a given
segment of the market, and they can vary across products in a given retailer.\footnote{In particular Wal-Mart and Costco, the first and third biggest retailers in the world
by revenues in 2012, claim not to use up-front payments, but Tesco (see Gale Group, 2001)
and Carrefour (the second and fourth in the world) do. FTC (2003) reports, for instance,
that up-front payments are more prevalent for ice cream and salad dressings than for bread
and hot dogs, even if they exist for these products as well. Sudhir and Rao (2006) find the
frequency of slotting allowances for one retailer particularly high for canned food products
(fruit, vegetable, juice and drinks), household supplies, health and beauty aids.}

This suggests that up-front payments are less the result of retailer character-
istics than of the retailer-manufacturer relationship. Our model sheds light
on the circumstances under which up-front payments arise in equilibrium and
on the factors determining their size.

In this paper we build a monopolistic competition model of retailers
and manufacturers to examine these stylized facts and to explore the con-
sequences of international trade for social welfare. Obviously our analysis
hangs on the fact that retailers play an important role in international trade.
They do, and this can be seen by the number of imported goods on their
shelves and by their direct involvement in importing. The Journal of Com-
merce (2014) ranks the top 100 US importers in 2013 using ocean containers.
The top 4 importers are retailers, and among the top 30 importers 16 are
retailers. Bernard, Jensen, Redding and Schott (2010) document the inter-
national trade activities of US retailers and wholesalers and find that 13%
of importing firms are pure retailers responsible for a small proportion of
overall US import value but 35% of the value of imports from China. Basker
and Van (2008) find that over the period 1997 to 2002 US imports from
China and other less-developed countries rose especially quickly in retail sec-
tors and that Wal-Mart alone accounts for around 15% of US consumer-good
imports from China (Basker and Van, 2010). This phenomenon is not lim-
ited to the United States and has taken place in many industries, including
electronics, computers, cameras, housewares, toys, games, clothing, footwear
and groceries.\textsuperscript{11} Interestingly, Blum, Claro and Horstmann (2010, 2011) find
that a considerable size difference exists between foreign exporters and the
importers they deal with. In particular, they find that large multi-product
retailers facilitate trade for small foreign exporters because they provide an
efficient way of reaching consumers who otherwise would be difficult to find.
Our model is precisely one that makes explicit the role of large multi-product
retailers dealing with smaller one-product manufacturers.

Our goal, however, is not to explain why manufacturers use retailers to
reach consumers. Rather we maintain that retailing is such an important
feature of consumer-goods markets that we take it for granted when exploring
the effects of international trade on retail markets as well as on resource
allocation between retailing and manufacturing and on social welfare.

The model has three main components. The first is a standard monop-
olistically competitive manufacturing sector where each manufacturer pro-
duces a single variety of a consumer good with an increasing-returns-to-scale
technology. Of course, this is a simplification as manufacturers are often
multi-product firms; however they typically produce a much smaller number
of varieties (see Eckel, Iacovone, Javorcik and Neary, 2015, for Mexico) than
sold by retailers. The second component is the retailing sector through which
all differentiated products are distributed. Retailers choose their product as-
sortment and retail prices. These two choices give them power although
limited by monopolistic competition. Moreover, each of them understands
that distributing more varieties within its own store leads to a cannibaliza-
tion effect in the sense that the demand for a new product ‘eats up’ some of
the demand for the other varieties sold in the store. We model this canni-
balization effect as in Feenstra and Ma (2008), who have developed this idea
for multi-product manufacturers.\textsuperscript{12}

\textsuperscript{11}For instance, in 2003, the share of imports in Canada was 55% for clothing, 82% for
clothing accessories, 86% for footwear, 100% for audio, video, small electrical appliances,
as well as for toys and games (Jacobson, 2006, Table 33).

\textsuperscript{12}See Dhingra (2010) for an alternative model of cannibalization and for showing that
intra-firm cannibalization is empirically relevant at the manufacturing level.
The third component is the wholesale market that connects manufacturers and retailers. Retailers charge an up-front payment for each product they stock. After determining which products to carry and receiving the up-front payment, each retailer bargains pair-wise with each of his manufacturers over the wholesale price. Even if this bargaining is efficient in the sense that the wholesale price maximizes the surplus of the retailer-manufacturer pair, the wholesale price nevertheless exceeds the marginal cost of production (and thus introduces a distortion). This is because the retailer-manufacturer pair takes into account the cannibalization effect that selling the respective variety generates for the retailer. The rent generated by this wholesale margin is transferred to the retailer through the up-front payment.

Next we consider the comparative static properties of the model, concentrating on the effects of international market integration. The model allows us to distinguish between two different types of integration. One is product-market integration, i.e., allowing manufacturers to export their products to more countries and allowing retailers to source differentiated products from different countries. The other is retail-market integration, i.e., allowing retailers to access overseas consumers by opening stores abroad. We find that the rise in retailer product assortment, the increase in up-front payments, as well as the shift in employment from manufacturing to retailing are consistent with international product-market integration. We also show that retail-market integration, but not product-market integration, reduces up-front payments and thus the distortion in the wholesale market.

Essentially, product-market integration helps manufacturers, not retailers, by spreading their fixed cost of production across multiple markets. Welfare rises because consumers benefit from the presence of more product varieties. But it is precisely the presence of more varieties that also creates bottlenecks at the retail level and higher up-front payments. Retail-market integration on the other hand helps retailers, not manufacturers, since retailers can spread headquarter fixed costs over multiple markets. This implies more competition among retailers, less bottlenecks at the retail level and thus lower up-front payments. Welfare increases, too, but now because the vertical distribution chain is more efficient.

Our paper is most closely related to Eckel (2009) and Raff and Schmitt (2012) as both papers also use a monopolistic competition approach to determine how trade liberalization affects retailing and social welfare when manufacturers have to go through retailers to reach consumers. What differentiates this paper is the vertical link between manufacturers and retailers
and its implications for up-front payments, the product assortment of retailers, the allocation of labor between manufacturing and retailing, and the welfare implications of product and retail market integration.

Other papers specifically examining the interaction between trade and retailing include Richardson (2004) who studies market access to retail distribution, Raff and Schmitt (2005, 2006, 2009) who examine the effects of trade liberalization on markets where either manufacturers or retailers have power over the other group of firms, and Francois and Wooton (2010) who show that market structure in distribution becomes increasingly important for trade as tariffs fall. They also include Basker and Van (2008) who investigate the effects of trade liberalization on competition between a chain retailer and small single-market retailers, concluding that trade liberalization raises the number of stores of the chain retailer, and that the growth of the chain gives an additional boost to imports. But again these papers do not consider the three stylized facts we are after.

There is a large industrial organization literature that examines the causes and consequences of up-front payments. A seminal contribution is Shaffer (1991) who shows that these payments may be used by retailers to soften price competition and shift rents from manufacturers to retailers. Others, such as Sullivan (1997) and Klein and Wright (2006), view up-front payments, such as slotting allowances, as a price for scarce shelf space. But the emphasis of these partial-equilibrium papers is on competition issues rather than, as in our paper, on explaining industry- or economy-wide phenomena and trying to link those phenomena to global market integration.13

The paper continues as follows. In Section 2, we present a simple model with manufacturers and retailers. Section 3 characterizes the equilibrium of the closed economy, and Section 4 examines the effects of product market integration. Section 5 introduces retail market integration and deals with the welfare effects of both product and retail market integration. Section 6 concludes, and the Appendix contains proofs.

13There is also a growing literature on the role of intermediaries in international trade and their welfare impact. It includes Antras and Costinot (2011), Bhum, Claro and Horstmann (2011), and Bardhan, Mookherjee and Tsumagari (2013). But these papers are mainly concerned with the role of wholesalers not retailers.
2 The Model

In this section, we develop a simple model of manufacturing and retailing. We first develop a model of a closed economy and later turn to a world economy consisting of $C$ identical countries with integrated product and/or retail markets.

2.1 Households

Consider an economy with $L$ consumers/workers, each endowed with one unit of labor. Individual preferences are given by the utility function

$$U = y_0 + \rho \ln(Y_d), \quad \rho < 1,$$

where $y_0$ denotes the consumption of an outside good, taken as the numeraire, and $Y_d$ is the aggregate individual consumption of a differentiated product. Letting $y_d(i)$ denote the quantity consumed of variety $i$, we assume that $Y_d$ takes the following CES form:

$$Y_d = \left( \int_{i \in \tilde{M}} y_d(i) \frac{n-1}{n} \, di \right)^{\frac{n}{n-1}},$$

where $\eta > 1$ is the elasticity of substitution between varieties and $\tilde{M}$ is the endogenous set of varieties.

Labor, the only factor of production, is inelastically supplied and perfectly mobile between the production and the retailing sectors. The numeraire good, $y_0$, is produced by a competitive industry under constant returns to scale and a unit labor requirement of one. The price of labor is hence also equal to one. Maximizing utility subject to the consumer’s budget constraint and aggregating individual demands over the $L$ consumers yields the following total demand for variety $i$:

$$y_d(i) = \frac{\rho L}{P^{1-\eta}} p(i)^{-\eta},$$

where $p(i)$ is the retail price of variety $i$, and $P$ is the CES price index.

2.2 Firms

There are two kinds of firms: manufacturers and retailers. Retailers engage in monopolistic competition, and they are large relative to manufacturers in the
sense that each manufacturer produces a single variety and sells that variety exclusively through one retailer, whereas retailers carry many varieties. Each retailer decides what mass of varieties to carry and sets the retail price of each variety. Since, in addition, the number of retailers is endogenously determined by free entry, the total mass of varieties is also endogenous.

Our modelling of retailers as multi-product firms follows Feenstra and Ma (2008) who develop this approach to study producers. There are $R$ retailers and the mass of varieties handled by retailer $r$ is $M_r > 0$. Given our assumption of exclusive dealing, each retailer carries a different set of varieties. Without loss of generality we choose the ordering of the products such that retailer 1 carries the first $M_1$ varieties, retailer 2 the following $M_2$ varieties, and so on. Hence the total mass of varieties consumed is $\bar{M} \equiv \sum_{r=1}^{R} M_r$, and the aggregate consumption of varieties is

$$Y_d = \left( \int_{0}^{M_1} y_d(i)^{\eta} di + \int_{M_1}^{M_1+M_2} y_d(i)^{\eta} di + \cdots + \int_{M-M_R}^{M} y_d(i)^{\eta} di \right)^{\frac{1}{\eta}}.$$  

(4)

Similarly, the CES price index is given by

$$P = \left( \int_{0}^{M_1} p(i)^{1-\eta} di + \int_{M_1}^{M_1+M_2} p(i)^{1-\eta} di + \cdots + \int_{M-M_R}^{M} p(i)^{1-\eta} di \right)^{\frac{1}{1-\eta}}.$$  

(5)

The assumption that a retailer carries a whole mass of varieties implies that any adjustment he makes to his assortment or to prices across his assortment has an effect on the price index. It is both realistic and useful to assume that retailers take this into account when choosing their assortment and setting prices. Consider first the implications for the retailer’s pricing decision. Note that manufacturers in our model are symmetric so that a retailer faces identical wholesale prices across all the varieties he sells. The retailer thus sets the same retail price for all varieties in his assortment and responds to a change in these wholesale prices by adjusting his retail prices across the

---

14Exclusive dealing is common in many industries. Although it does not prove exclusivity, there is very little overlap between the products sold by different retailers when one considers barcode data. Broda and Romalis (2009) find that only around 2% of the 61,119 food Universal Product Code categories sold by either Wal-Mart or Wholefoods are sold by both. The choice between exclusive and non-exclusive dealing contracts has been studied in a trade context by Raff and Schmitt (2006, 2009). We have nothing new to add to the analysis of this choice and therefore do not model it here.
board. Denoting the price retailer $r$ charges for each of the varieties he sells by $p_r$, the CES price index (5) simplifies to:

$$P = \left( \sum_{r=1}^{R} M_r p_r^{1-\eta} \right)^{\frac{1}{1-\eta}}. \quad (6)$$

Taking into account that the price index $P$ is increasing in $p_r$ for $M_r > 0$ means that demand for each variety reacts less to price changes than in the usual CES framework. More precisely, the price elasticity of demand is not constant but rather decreasing in $r$’s market share, $s_r$:

$$-\frac{\partial y_r}{\partial p_r} \frac{p_r}{y_r} = \eta - (\eta - 1)s_r, \quad (7)$$

with

$$s_r = \frac{M_r p_r y_r}{\sum_{r=1}^{R} M_r p_r y_r} = \frac{M_r p_r^{1-\eta}}{\sum_{r=1}^{R} M_r p_r^{1-\eta}} \quad (8)$$

and $y_r$ denoting the quantity of an individual variety sold by retailer $r$.

Next consider how a multi-product retailer’s effect on the price index affects the assortment choice. From (6) we see that an increase in $M_r$ reduces $P$. Using this in (3), we observe that demand for each variety falls. In other words, the retailer acknowledges that adding a product to his assortment lowers the demand for the other products he carries. As we will see below, this ‘cannibalization’ effect becomes bigger as the retailer adds products, thus putting a limit on his product assortment.

Retailing traditionally involves a spatial dimension, with consumers visiting local retailers to purchase different bundles of differentiated products.\(^\text{15}\)

We deal with the spatial dimension in a very simple way. In particular, we assume that each retailer sells its goods to consumers in its home market through one (or a fixed number of) retail store(s). The fixed cost of operating a store (or stores) is denoted by $k_s$. Retailing also involves a fixed cost

---

\(^{15}\)The CES utility function (2) may be given a microfoundation in terms of a spatial model of product differentiation. Anderson, de Palma and Thisse (1992, ch. 5) demonstrate that it can be reinterpreted as arising from an address model in which each consumer buys only one variety (in our case from one retailer), namely the one whose location in characteristic space provides the greatest conditional indirect utility. Note, however, that the models may not have the same welfare implications, as shown by Tito (2014) for the case of asymmetric trade liberalization.
for headquarter services, $k_h$, such as payments for information technology
that play a crucial role in retailing. Define the total fixed cost of retailing as
$k_0 \equiv k_h + k_s$.

The fixed cost per variety carried, for instance the cost of shelf space,
is $k_1$. These costs turn out to be important for the analysis. Retailers
are homogeneous in that they all use the same technology; we may therefore
normalize the marginal cost of selling a unit of a given variety to zero, and we
drop retailer subscripts whenever this can be done without causing confusion.

Hence the labor requirement of a retailer carrying a mass $M_r$ of varieties
in its local store(s) is given by

$$l_r = k_0 + k_1 M_r.$$  \hfill (9)

Manufacturers are single-product firms. There is free entry into manufactur-
ing so that retailers are free to choose which manufacturers to do business
with. We make the standard assumption that production requires a fixed la-
or input, $\alpha$, and a variable labor input, $\beta$, both identical costs across firms.
Hence the total labor input required to produce $y_r$ units of a given variety is
given by

$$l_m = \alpha + \beta y_r, \quad \alpha, \beta > 0.$$ \hfill (10)

2.3 The Wholesale Market

The manufacturing and the retailing side of the economy are linked through
the wholesale market. Each retailer first chooses which manufacturers to do
business with and collects an up-front payment, $T$, for each product he carries
(where $T$ can be negative). He then bargains simultaneously and bilaterally
with each of his manufacturers over a wholesale price $w$.\footnote{Note that it
would be difficult, even illegal, for a retailer to get together with all
his suppliers to jointly fix wholesale prices. The assumption of simultaneous
bilateral bargaining between multi-product retailers and individual manufacturers (or of manufacturers
dealing with more than one retailer) is standard in the industrial organization literature on buyer power (see, for instance, Dobson and Waterson, 1997).} We assume for
simplicity that the wholesale price chosen in this bargaining maximizes the
joint surplus of each retailer-manufacturer pair. This allows us to shift the
focus away from possible bargaining inefficiencies to the market inefficiency
that arises naturally when a multi-product retailer chooses his assortment
but negotiates the wholesale price individually with each manufacturer. As
will be shown below, due to the cannibalization effect, the wholesale price exceeds the marginal production cost. The rent generated by this markup is what may lead to a positive up-front payment in equilibrium.

3 Equilibrium of the Closed Economy

In this section we characterize the equilibrium of the closed economy. A retailer chooses the retail price \( p \) and the mass of varieties \( M \) to maximize:

\[
\Pi_r = M (p - w) y - M (k_1 - T) - k_0. \tag{11}
\]

Substituting for \( y \) from (3), the corresponding first-order condition with respect to the retail price reads:

\[
p = \left( 1 + \frac{1}{(\eta - 1)(1 - s)} \right) w. \tag{12}
\]

We observe that the higher is a retailer’s market share, \( s \), the higher is his mark-up. To derive the first-order condition with respect to \( M \), recall that we have to take into account that \( y \) is a function of \( M \) through the effect each retailer has on the price index. The first-order condition then reads:

\[
(p - w) y - s (p - w) y = k_1 - T. \tag{13}
\]

The left-hand side of (13) gives the marginal benefit of adding a variety. It has two elements: the first term is the additional operating profit generated by this variety. The second term represents the cannibalization effect, that is, the reduction in the demand for the other varieties sold by the retailer times the mark-up on these other varieties. The higher is the retailer’s market share, the bigger is this cannibalization effect. On the right-hand side of (13) we have the marginal cost of adding a variety, which consists of the direct cost, \( k_1 \), minus any up-front payment received from the manufacturer producing the additional variety.

A manufacturer’s profit, \( \Pi^m \), is given by

\[
\Pi^m = (w - \beta) y - T - \alpha. \tag{14}
\]

The surplus that is generated when a retailer adds the manufacturer’s product to his assortment is equal to the sum of \( \Pi^m \) and the incremental profit
of the retailer, which we have already encountered in (13). We thus have as joint surplus of a manufacturer-retailer pair:

\[(w - \beta)y + (1 - s)(p - w)y - k_1 - \alpha.\]  

(15)

As explained above, we assume that the manufacturer-retailer pair settles on the wholesale price that maximizes this joint surplus. Taking the derivative with respect to \(w\) and then substituting for \((p - w)\) from (12), we can write the corresponding first-order condition as:

\[
\frac{\eta y}{\eta - 1} + \left(\frac{\eta w}{\eta - 1} - \beta\right) \frac{dy}{dp} \frac{dp}{dw} = 0.
\]

(16)

Recognizing that \(dy/dp\) follows from (7) and \(dp/dw\) from (12), we can solve (16) for the equilibrium wholesale price:

\[
w = \beta + \frac{s \beta}{\eta(1 - s)}.
\]

(17)

The wholesale price thus exceeds the manufacturer’s marginal cost by a margin that is increasing in the retailer’s market share, \(s\). This distortion is due to the cannibalization effect: the retailer-manufacturer pair takes into account that additional sales of one variety reduce demand for other varieties sold by the retailer. This effect is stronger the greater the retailer’s market share, which implies that the wholesale price is increasing in the market share.

Given that there is free entry into manufacturing, retailers can extract any rents earned by manufacturers through the up-front payment, which implies that \(\Pi^m = 0\). As can be seen from (14), the up-front payment from the manufacturer to the retailer hence equals the rent earned by the manufacturer, \((w - \beta)y\), net of the fixed cost of production, \(\alpha\):

\[
T = (w - \beta)y - \alpha.
\]

(18)

Naturally, if a manufacturer did not earn any rent, he would be unable to pay a retailer for adding his products to the assortment.

Using (17) and (18) in (13), we can then solve for the output of each variety

\[
y = (1 - s)\frac{(k_1 + \alpha)(\eta - 1)}{\beta}.
\]

(19)
This output is decreasing in \( s \) due to the fact that the wholesale price is increasing in \( s \) as explained above.

To close the model we impose zero-profit conditions on retailers and a labor-market clearing condition. The retailer zero-profit condition is obtained by setting the profit in (11) equal to zero. This yields an expression for the mass of varieties carried by each retailer as a function of the number of retailers:

\[
M = \frac{k_0}{(k_1 + T)} (R - 1).
\]  

(20)

A second equation linking \( M \) and \( R \) is the labor-market clearing condition. Since in equilibrium a fraction \( \rho \) of the labor force is employed in the differentiated-good industry (i.e., manufacturing and retailing), this condition can be written as:

\[
Rk_0 + RM (k_1 + \alpha) + RMy\beta = \rho L.
\]  

(21)

We can now easily solve for the equilibrium number of retailers,

\[
\hat{R} = \frac{1}{\eta} \left( \frac{\eta - 1}{2} + \sqrt{\frac{(\eta - 1)^2}{4} + \frac{\eta \rho L}{k_0}} \right),
\]  

(22)

and the mass of varieties carried by each retailer:

\[
\hat{M} = \left( \frac{\eta}{\eta(1 - \hat{s}) + \hat{s}} \right) k_0 (1 - \hat{s}) \left( k_1 + \alpha \right) \hat{s},
\]  

(23)

where \( \hat{s} = 1/\hat{R} \). We observe from (22) that an increase in \( k_0 \) reduces the equilibrium number of retailers (\( d\hat{R}/dk_0 < 0 \)), leading to greater retail market concentration. The effect on the retailer product assortment is non-trivial, since \( k_0 \) affects the equilibrium assortment directly and indirectly through the effect on the number of retailers. It can be shown that \( d\hat{M}/dk_0 > 0 \). That is, an increase in \( k_0 \), whether it is through the fixed cost of operating a store or the headquarter fixed cost, requires retailers to carry a larger product assortment in order to avoid making losses.

Using (12) and (17) we observe that the equilibrium retail price exceeds the marginal production cost due to both the retailer mark-up and the wholesale mark-up:

\[
\hat{p} = \left( 1 + \frac{1}{(\eta - 1)(1 - \hat{s})} \right) \left( 1 + \frac{\hat{s}}{\eta(1 - \hat{s})} \right) \beta.
\]  

(24)
The equilibrium value of output per variety can be obtained by using $\hat{s}$ in (19):

$$\hat{y} = (1 - \hat{s}) \frac{(k_1 + \alpha) (\eta - 1)}{\beta},$$

and the equilibrium transfer from a manufacturer to a retailer is

$$\hat{T} = \hat{s} \frac{\eta - 1}{\eta} (k_1 + \alpha) - \alpha.$$  

An up-front payment by the manufacturer occurs in equilibrium when $\hat{T} > 0$. It has the following properties:

**Proposition 1** The equilibrium up-front payment by a manufacturer is increasing in the retailer fixed cost ($k_0$), the cost of adding a variety ($k_1$), the elasticity of substitution ($\eta$), and decreasing in the manufacturer’s fixed cost ($\alpha$) and the fraction of income spent on differentiated goods ($\rho L$).

**Proof:** see Appendix.

The up-front payment reflects the costs and the benefits for a retailer to take on an additional variety. Thus the higher is the retailing cost per variety $k_1$ and the greater is the elasticity of substitution $\eta$, the less attractive it is for a retailer to take on an additional variety forcing the manufacturer’s up-front payment to be higher to induce the retailer to do so. The same is true for a higher fixed cost of retailing $k_0$ and a lower $\rho L$ which both reduce the number of retailers, thus making them bigger and strengthening the cannibalization effect.

In the next section, we investigate the impact of international market integration on the equilibrium and in particular whether product market integration can explain the stylized facts about retailing discussed in the introduction, including the rise in up-front payments.

## 4 Product-Market Integration

By product-market integration, we mean a scenario in which goods become tradable across countries but retail services remain non-tradable. Manufactu-
turers are thus able to reach more consumers by exporting goods to foreign markets. From the point of view of retailers, however, the number of households served does not change when product trade is liberalized, simply because there is no cross-border shopping.

We may examine the effects of product-market integration by considering a world consisting of identical countries indexed by \( c = 1, \ldots, C \) and studying how free trade in goods between them affects the equilibrium. From the point of view of a manufacturer, free trade means that his market has expanded as he is now able to sell his products to \( C \) retailers, one each in the \( C \) countries comprising the integrated world economy. Another way of saying this is that the manufacturer is able to spread his fixed cost over \( C \) markets. In effect, the fixed cost of manufacturing per country becomes \( \frac{\alpha}{C} \).

Since product-market integration only amounts to a reduction of the fixed cost of production per market, it neither affects the determination of the wholesale and retail prices, nor does it change the number of retailers. What changes is output and the number of varieties. To show this formally, we have to make a few straightforward modifications to our notation. Let the assortment that each retailer carries now be given by \( M = CM_c \), where \( M_c \) is the number of varieties produced in country \( c \). Let \( y_c \) denote the quantity sold in country \( c \) and \( T_c \) denote the transfer received by a retailer in that country.

With this notation, we can examine how the labor market equilibrium in a given country is affected by free trade. In particular, only a mass \( RM_c \) of varieties sold by retailers in a given country are locally produced varieties, but each local producer now has an output equal to \( C y_c \). Hence \( RM y_c \beta \) units of labor are needed to cover the variable labor requirement in production. The fixed labor requirement in production absorbs \( RM \alpha = RM \alpha / C \) units of labor, and the remaining labor is allocated to retailing. The new labor market clearing condition in a country is then

\[
Rk_0 + RM \left( k_1 + \frac{\alpha}{C} \right) + RM y_c \beta = \rho L. \tag{27}
\]

Noting that the number of retailers in each country and hence retailer market share remains unchanged at \( \hat{s} \), we can compute the mass of varieties (local and imported) carried by a retailer and local consumption of each variety by replacing \( \alpha \) with \( \alpha / C \) in (23) and (25):
Product-market integration thus leads to a market equilibrium in which there is a larger mass of product varieties carried by each retailer \( (d\tilde{M}/dC > 0) \), a larger total mass of varieties available to consumers (since the number of retailers remains unaffected), and a decrease in the consumption of each variety \( (d\tilde{y}_c/dC < 0) \).\(^{18}\)

By replacing \( \alpha \) with \( \alpha/C \) in (26), we can compute the up-front payment that a manufacturer has to pay each of the \( C \) retailers carrying his product:

\[
\tilde{T}_c = \hat{s}(\eta - 1) \frac{\eta - 1}{\eta} (k_1 + \alpha/C) - \alpha/C.
\]  

Product-market integration obviously erodes the quasi-rent earned by the manufacturer, the first term in (30). But the fixed cost falls by even more so that, on balance, the up-front payment paid by a manufacturer to each retailer rises as \( C \) goes up. In addition, even if there were no up-front payment in autarky \( (\tilde{T} \leq 0 \) in (26)), it has to be the case that in free trade \( \tilde{T}_c > 0 \) if \( C \) is sufficiently big.

Another interesting result is the impact of product-market integration on the allocation of labor between manufacturing and retailing. Since resources are being saved in manufacturing, product-market integration implies a shift in resources from manufacturing into the retail sector. This can be seen from (27) where the amount of labor allocated to retailing \( (\tilde{R}\tilde{M}\hat{c}k_1) \) rises, the fixed labor requirement in manufacturing \( (\tilde{R}\tilde{M}\alpha/C) \) declines, while the variable labor input in manufacturing \( (\tilde{R}\tilde{M}\tilde{y}_c\beta) \) remains unchanged. What makes this reallocation of labor possible is the fact that while the mass of varieties available to consumers rises with market integration, the mass of varieties produced in each country falls so that less labor is required in manufacturing.

We therefore conclude that the three stylized facts listed in the introduction are consistent with product market integration:

---

\(^{18}\)The change in consumption is non-standard in a model with CES preferences, but is due to the fact that in our model the price elasticity of demand is not constant.
Proposition 2 Product-market integration (i) raises the product assortment carried by each retailer and the total mass of varieties available to consumers, while reducing the quantity consumed of each variety; (ii) raises up-front payments by manufacturers; and (iii) leads to a reallocation of labor from manufacturing to retailing.

5 The Welfare Effects of Market Integration

In this section we examine the welfare effects of both product and retail market integration. These effects are non-trivial, since there is a distortion in the relationship between independent multi-product retailers and manufacturers. In particular, the market equilibrium is inefficient, since it involves double marginalization in the vertical distribution chain. In order to understand the normative aspects of market integration, it is useful to first compare the market allocation in autarky with a second-best allocation. Using this second-best allocation as the benchmark, we then analyze how market integration affects welfare.

The second-best allocation that we consider as the benchmark is the one in which double marginalization is ruled out by imposing a wholesale price equal to the marginal production cost, i.e., $w^B = \beta$, where the superscript $B$ denotes the second-best allocation. In this second best, the transfer that guarantees manufacturers zero profit is $T^B = -\alpha$. The second best thus amounts to maximizing the sum of a retailer’s profit and the profits of all the manufacturers he deals with.

Given these values of $w^B$ and $T^B$, it is straightforward to establish that $R^B = \hat{R}$ and

\begin{align}
M^B &= \frac{\eta(1 - \hat{s}) + \hat{s}}{\eta} \hat{M} < \hat{M}, 
\end{align}

\begin{align}
y^B &= \frac{\hat{y}}{1 - \hat{s}} > \hat{y}
\end{align}

\begin{align}
p^B &= \left(1 + \frac{1}{(\eta - 1)(1 - \hat{s})}\right) \beta < \hat{p}.
\end{align}

Hence not only is the retail price in the market equilibrium too high compared with the second best, but each retailer carries too many varieties and sells too little of each variety.
Our benchmark can also be used to evaluate the allocation of resources between manufacturing and retailing. In particular, it is immediate from (31) and (23) that in equilibrium too much labor is devoted to distributing the mass of varieties compared to the second best, $\hat{R}(k_0+k_1 M) > \hat{R}(k_0+k_1 M^B)$. With a fixed amount of labor devoted to the differentiated good industry ($\rho L$), this implies that too little labor is left over for the production of each variety.

This has implications for welfare. Despite the fact that the mass of varieties is bigger in equilibrium than in the second best, equilibrium welfare is lower precisely because there is too little consumption of each variety. These results can be summarized as follows:

**Proposition 3** In comparison to the second best, the equilibrium in autarky exhibits: (i) a larger product assortment of each retailer and smaller sales per variety; (ii) a larger total mass of differentiated products in the economy; (iii) a greater allocation of labor to retailing, and (iv) a lower social welfare.

**Proof:** see Appendix.

Now consider the effects of market integration. It is straightforward to see that product-market integration leaves the distortion unchanged, because retailer market share remains unchanged at $\hat{s}$. This implies that, even with product-market integration, product variety remains too large, output per variety too small and too much labor is allocated to retailing when compared with the second best. Of course social welfare increases with product-market integration but this gain is entirely due to gains from variety. Thus,

**Proposition 4** Product-market integration, while raising social welfare, does not move the equilibrium allocation closer to the second best.

If product-market integration is unable to close the gap with the second best, is there any form of market integration that would do it in our model? We show that retail-market integration has this property. In particular, we prove:

**Proposition 5** Retail-market integration raises social welfare and moves the equilibrium allocation closer to the second best.
Proof: see Appendix.

By retail-market integration we mean that retailers gain access to foreign customers by opening stores in each country, but there is no trade. In terms of fixed costs, a retailer operating a store in each of $C$ countries faces fixed costs of $k_h + C k_s < C k_0$. That is, the benefit to the retailer is that he can spread the headquarter fixed cost across $C$ countries. This implies the following labor market clearing condition for each individual country:

$$R \left( \frac{k_h}{C} + k_s \right) k_0 + RM (k_1 + \alpha) + RMy_c\beta = \rho L.$$  \hfill (34)

Using $k_h + C k_s$ instead of $C k_0$ in (22), it is easy to see that retail market integration raises the total number of retailers and thus lowers the market share of each retailer, $\hat{s}$. A lower retail market share reduces the distortion in the wholesale price, moving it closer to marginal cost $\beta$, as can be seen from (17). A lower wholesale mark-up is equivalent to a lower up-front payment. Another way to see this is to note that a smaller $\hat{s}$ reduces the cannibalization effect and hence the payment manufacturers have to offer retailers to obtain distribution for their products. The retail price declines due to the reduced wholesale price and because a retailer with a lower market share perceives a higher price elasticity of demand and thus charges a smaller retail mark-up. Output of each variety obviously has to increase when retail prices fall. Retail-market integration also impacts the allocation of labor between manufacturing and retailing by making it less concentrated in retailing.

To understand the effect of retail-market integration on retailer product assortment, it is useful to rewrite (23) as

$$\hat{M} = \left( \frac{\eta}{\eta(1 - \hat{s}) + \hat{s}} \right) M^B,$$ \hfill (35)

where the first term comes from the market distortion. The reduction in the cannibalization effect associated with a smaller $\hat{s}$ increases directly $M^B$. However, the distortion also becomes smaller which decreases the first term. As shown in the Appendix, the effect on $M^B$ dominates so that retailer product assortment rises. Social welfare must unambiguously rise, since retail prices fall and overall product variety in the economy increases. Finally, as the distortion in the wholesale market shrinks, equilibrium welfare approaches the second-best level.
6 Conclusions

This paper proposed a simple monopolistic competition model to examine whether product-market integration may explain several important stylized facts regarding retailing and manufacturing. We showed that, in the context of the model, product-market integration contributes to the rise of up-front payments by manufacturers, the larger product assortment carried by retailers as well as the shift in employment from manufacturing to retailing. To our knowledge, no one has established a direct link between the use of up-front payments and retailers’ imports, but as noted in the introduction there is at least circumstantial evidence that the emergence of up-front payments in the early 1980s and their subsequent spread across retailing coincide with higher import penetration of consumer products in the US, and the use of these payments in industries like apparel and footwear.

Up-front payments by manufacturers arise in our model, because manufacturers and retailers understand that a retailer selling one more variety reduces the demand for the other varieties he sells. This cannibalization effect implies that the wholesale price is set in excess of the marginal cost and is combined with a fixed payment which, when it is paid by a manufacturer to a retailer, represents an up-front payment. This payment and the inefficiency it is associated with are thus directly linked to the fact that retailers are multi-product firms, but they do not depend on our simple modeling of manufacturers producing a single good. The same inefficiency would persist with multi-product manufacturers as long as one manufacturer is not the only provider of the products sold by a retailer and thus as long as each manufacturer produces a smaller mass of varieties than sold by a retailer.

We also showed that retail-market integration leads to welfare gains that are distinct from those achieved through product-market integration. Retail-market integration, which we defined as an expansion of retailers to other countries, reduces the gap between the market equilibrium wholesale price and the marginal cost of production. The recent internationalization of retailing through foreign direct investment makes this scenario increasingly likely with the implication that up-front payments may not be as important in the future as they have been over the last 30 years.

The observed shift in employment from manufacturing to retailing may, of course, also be a consequence of offshoring of production and differences in technologies between services (such as retailing) and manufacturing. Interestingly, these forces do not need to be present for this shift in resources to
occur. Indeed, given the vertical links between retailers and manufacturers, this paper shows that this shift toward retailing is a natural consequence of market integration.

In this paper, manufacturers go through independent retailers in order to have their product picked up by consumers. Vertical integration between manufacturers and retailers could easily be examined in our model as well. In fact, to the extent that vertical integration eliminates the inefficiency between each retailer and the manufacturers it deals with, the market outcome would be identical to the second best derived in Section 3. This shows one more time that a central point of this paper is linked to the inefficiency that manufacturers and multi-product retailers generate when they are independent.

More broadly, this paper suggests that, in order to understand economic integration or the policies associated with it in today’s world, our attention should not be restricted exclusively to freer trade in goods or services separately. Indeed when services such as retailing are closely associated with the products themselves because there is a complementarity between production and distribution, then not only may market integration take different forms but its impact may differ as well depending on the specific form of the integration. It is then crucial to understand how the production and distribution of goods interact. This is what this paper has started to do by proposing a model giving multi-product retailers a central role in an environment consistent with three stylized facts that have unfolded over the last few decades.

7 Appendix

7.1 Proof of Proposition 1

Given (22) and (26), the changes in $\hat{T}$ caused by changes in $k_0$, $k_1$, $\alpha$ and $\rho L$ are straightforward. To determine the comparative statics with respect to $\eta$ rewrite $\hat{T}$ as

$$\hat{T} = \frac{(\eta - 1) (k_1 + \alpha)}{\left(\frac{\eta - 1}{2} + \sqrt{(\frac{\eta - 1}{4})^2 + \frac{\eta \rho L}{k_0}}\right)} = \frac{(\eta - 1) (k_1 + \alpha)}{D}.$$
Thus

\[ \frac{\partial \hat{T}}{\partial \eta} = \frac{1}{D^2} \left[ (k_1 + \alpha) D - (\eta - 1) (k_1 + \alpha) \left\{ \frac{1}{2} + \frac{1}{2} \sqrt{\frac{(\eta - 1)^2 + \rho L}{k_0}} \right\} \right]. \]

\[ \text{sign} \frac{\partial \hat{T}}{\partial \eta} = \text{sign} \left[ \frac{D - \eta - 1}{2} \left\{ 1 + \frac{(\eta - 1)\rho L}{k_0} \right\} \right] \]

\[ = \text{sign} \left[ \frac{(\eta - 1)^2}{4} + \frac{\eta \rho L}{k_0} - \frac{(\eta - 1)\rho L}{k_0} \right] \]

\[ = \text{sign} \frac{\rho L}{k_0} \left( \frac{\eta + 1}{2} \right) > 0. \]

### 7.2 Proof of Proposition 3

We prove (iv) only. Since consumers spend a fixed share of their income on differentiated goods, social welfare (i.e., indirect utility) is strictly decreasing in the price index for differentiated goods. The price indices in equilibrium and in the second best are given respectively by

\[ \hat{P} = \hat{p} \left( \hat{R} \hat{M} \right)^{\frac{1}{1-\eta}} \text{ and } P^B = p^B \left( \hat{R} M^B \right)^{\frac{1}{1-\eta}}. \]

Given that the number of retailers is the same in equilibrium and in the second best, the respective price indices can be written as

\[ \hat{P} = \hat{p} \left( \hat{R} \hat{M} \right)^{\frac{1}{1-\eta}} \text{ and } P^B = p^B \left( \hat{R} M^B \right)^{\frac{1}{1-\eta}}. \quad (36) \]

We hence have

\[ \hat{P} - P^B = \hat{R}^{\frac{1}{1-\eta}} \left[ \hat{p} \hat{M}^{\frac{1}{1-\eta}} - p^B M^B \right] \]

\[ = p^B \left( \hat{R} M^B \right)^{\frac{1}{1-\eta}} \left[ \frac{1}{(1 - \hat{s}) \left( \frac{\eta}{\eta(1 - \hat{s}) + \hat{s}} \right)} \right] - 1 \]. \quad (37)

\[ \hat{P} - P^B > 0 \text{ provided that the expression in brackets is positive. This is the case if} \]

\[ f(\hat{s}, \eta) \equiv \hat{s} - \eta (1 - \hat{s}) \left[ (1 - \hat{s})^{-\frac{1}{\eta}} - 1 \right] > 0 \]

(38)
for $\eta > 1$ and $\hat{s} \in (0, 1)$. Note that $f(0, \eta) = 0$. The proof proceeds by showing that $f(\hat{s}, \eta)$ reaches a minimum in $\hat{s}$ at $\hat{s} = 0$, which is guaranteed by

$$
\frac{\partial f(\hat{s}, \eta)}{\partial \hat{s}} = 1 + \eta \left[ (1 - \hat{s})^{-\frac{1}{\eta}} - 1 \right] - (1 - \hat{s})^{-\frac{1}{\eta}} = 0 \quad \text{at } \hat{s} = 0,
$$

and

$$
\frac{\partial^2 f(\hat{s}, \eta)}{\partial \hat{s}^2} = \left( 1 - \frac{1}{\eta} \right) (1 - \hat{s})^{-\frac{1+\eta}{\eta}} > 0 \quad \forall \hat{s} \in [0, 1) \text{ and } \eta > 1.
$$

### 7.3 Proof of Proposition 5

Given an increase in $C$, the increase in $R$ (and decrease in $\hat{s}$) follows immediately from (22). The decrease in $\hat{s}$ reduces $\hat{p}$ and $\hat{w}$, as can be seen in (24) and (17), respectively. The effect on $\hat{M}$ is given by:

$$
\frac{d\hat{M}}{d\hat{s}} = -\frac{k_0(1 - \hat{s})}{(k_1 + \alpha) \hat{s} (\eta(1 - \hat{s}) + \hat{s})^2} \frac{\eta}{(k_1 + \alpha) \hat{s}^2 \eta(1 - \hat{s}) + \hat{s}} - \frac{k_0}{(k_1 + \alpha) \hat{s}^2 (\eta(1 - \hat{s}) + \hat{s})^2} [1 - s(\eta(1 - s) + s)+s] < 0.
$$

Overall product variety, $\hat{R}\hat{M}$, rises, since both components increase.

The rise in social welfare follows directly from the fall in the price index due to the decrease in retail prices and the increase in $\hat{R}\hat{M}$. To see why equilibrium welfare approaches the second best, note from the proof of Proposition 3 that $\hat{P} - P^B$ is proportional to $f(\hat{s}, \eta)$, as defined in (39); but $f(\hat{s}, \eta)$ approaches zero as $\hat{s}$ falls.

### References


[40] Shanghai Daily (2011) "New rules on fees planned," February 18


