EUROSÜSTEEM

Law of One Price in the euro area: an empirical investigation using Nielsen disaggregated price data

Dmitry Kulikov

Working Paper Series
10/2014

The Working Paper is available on the Eesti Pank web site at:
http://www.eestipank.ee/en/publications/series/working-papers

For information about subscription call: +372 668 0998; Fax: +372 668 0954
e-mail: publications@eestipank.ee

ISBN 978-9949-493-50-0
Eesti Pank. Working Paper Series, ISSN 1406-7161; 10

Law of One Price in the euro area: an empirical investigation using Nielsen disaggregated price data

Dmitry Kulikov*

Abstract

This paper examines the Law of One Price using Nielsen disaggregated price data covering 13 euro area countries and 45 different product categories over the time period 2008 to 2012. The empirical methodology is based on a non-structural log-linear regression with spatial effects in both the geographical and product-variety dimensions, estimated by the Bayesian methods. The models link the relative prices of homogenous products in the sample of euro area countries to four distinct groups of factors: product-specific consumption preferences, country-specific macroeconomic and regional characteristics, volatility of prices and volumes, and spatial effects. The estimated reduced-form Law of One Price models uncover a strong interdependence of relative prices both on the geographical scale and across "similar" product varieties, going beyond the included set of explanatory variables and warranting further empirical investigation.

JEL Code: C21, D40, E31

Keywords: disaggregated prices, spatial dependence, Bayesian estimation, Law of One Price

Author's e-mail address: dmitry.kulikov@eestipank.ee

*I wish to express my gratitude to the conference and research seminar participants at Eesti Pank, ECEE 2014 and Lietuvos Bankas for their useful comments and suggestions. As usually, all errors and omissions in the paper are my own.

Non-technical summary

The Law of One Price (LOP) asserts that the prices of identical goods and services in different locations (countries, regions, shops, etc.) are the same when expressed in a common denominator. Purchasing Power Parity (PPP) postulates that the LOP holds on average across baskets of similar goods and services in different locations. Both of these concepts are related to the cross-section of prices in different locations, but their time evolution is also important, especially in the context of global trade integration.

Many of the existing empirical studies of these concepts are data-limited: price data that are broad and wide in scope, but sufficiently homogeneous to permit meaningful comparisons across countries and products, and rich in terms of supplementary attributes are difficult to come by. In addition, volatile nominal exchange rates further complicate price comparisons between countries on the international scale. Set against this background, the euro area offers a unique opportunity to study the LOP and PPP in the framework of a common currency, a single market, increasing economic and financial integration, and a shared regulatory regime.

In spite of the inherently limited data, the considerable policy relevance of these issues stimulates ample academic research on understanding the price disparities of similar goods and services in the geographical dimension (countries, intra-country regions, supra-national economic and financial unions) using disaggregated price data. The research agenda in this field originates from the seminal "border effects" study of Engel and Rogers (1996), where significant differences between price pairs of similar products across several US and Canadian cities are documented and ascribed to the effect of border between the two countries. Since then, many empirical LOP studies have identified the following data regularity: while prices of similar goods and services tend to be quite homogeneous within a given country, a substantial and often difficult to explain price heterogeneity is frequently observed even between very geographically close and economically similar pairs of countries.

These issues are particularly relevant for the European Union (EU) in general and the euro area in particular. An overview of empirical price level convergence in Europe before and after the European Monetary Union (EMU) is given in Faber and Stokman (2009), where they document significant momentum towards price level harmonisation in the run up to the EMU, which appears to cease after 2002. They also show that the price level dispersion across the EMU is larger than a comparable statistic for the US. On the other hand, Crucini, Telmer and Zachariadis (2005) find, after controlling for the national income and VAT differences, that average consumer basket prices in the EU are quite similar, although large differences between countries still exist at

the level of prices for individual products.

This paper contributes to the empirical literature on the LOP deviations in the euro area using new data on disaggregated prices, which was recently made available by the ECB to several researchers at the euro area national central banks. The dataset, compiled by Nielsen Holdings N.V. from their proprietary commercial records of European retail prices, covers 13 euro area countries over the time period 2008 to 2012 and includes 45 homogenous product varieties, additionally split across the intra-country regions and store types. These rich data are particularly suitable for LOP studies, as they feature a large number of harmonised retail-level product varieties that permit price comparisons in the equivalent unit terms (expressed as euro prices per kilogram, litre and so on), and include a wealth of information on brand and selling unit-level price ranges and price volatilities within every product category.

The paper focuses on the empirical, i.e. "non-structural", LOP models that attribute the observed price heterogeneity across the sample of euro area countries to a number of factors linked to the product-specific consumption preferences, country-specific macroeconomic and regional characteristics, volatility of prices and sales volumes, and spatial effects. The empirical methodology is based on log-linear regressions with spatial dependence across the geographical and product-variety dimensions, estimated by the Bayesian methods. Therefore this paper offers a more nuanced view of the classical "border effect", which is made possible by the rich structure and depth of the Nielsen disaggregated price data.

In particular, the paper finds that relative variations in the income levels and economic growth rates across countries strongly and predictably affect relative prices. In addition, a number of significant price effects are linked to the economies of scale in the retail sector, and to consumer demand and preferences in different euro area economies. At the same time, the estimated reduced-form LOP regressions suggest a strong interdependence of relative prices in both the geographical and product-variety data dimensions, going beyond the included set of explanatory variables and warranting further empirical investigation.

The paper documents several empirical regularities in the observed distribution of relative euro area prices, which should help to improve the economic policy advice that depends on a detailed understating of price shock propagation across the euro area economies.

Contents

1. Introduction

The issue of pricing and price differentials of homogeneous goods and services in the euro area, fifteen years after the introduction of the common currency, remains high on the research agenda of central banks and academia across Europe. It is of great practical importance to understand and explain the price convergence and remaining price differentials in order to formulate effective economic and monetary policies in the common currency area. From the theoretical point of view, any systematic deviations from the Law of One Price (LOP) across the tightly integrated euro area countries and regions present a unique opportunity to advance our knowledge and understanding of the real-world pricing of goods and services, and improve our economic models and policy advice.

There is a large and growing literature on the price heterogeneity of similar goods and services in different countries and the closely related empirical LOP modelling using micro- and disaggregated prices. One of the seminal studies in the field, Engel and Rogers (1996), examines prices across nine Canadian and fourteen US cities, and ascribe the observed significant heterogeneity in price pairs of similar products across the US and Canada to the effect of the border between the two countries. Since the study of Engel and Rogers (1996), one of the main "stylised facts" in many empirical LOP models is that prices of similar products tend to be quite homogeneous within a country, but substantial and difficult to explain price differentials are often found even between closely neighbouring countries.

This issue is particularly relevant for the euro area, since one of the main advantages that common currency brings to markets is price transparency and the elimination of exchange rate risks. In addition, tight economic and financial integration in the euro area countries should enhance market competition and promote mobility of labour and capital, exerting further pressure on the remaining heterogeneity of prices of similar goods and services in Europe.

However, some significant LOP deviations in the euro area still remain and are well documented in many recent studies. An overview of the empirical price level convergence in Europe before and after the EMU is provided in Faber and Stokman (2009). They document a significant push towards price level convergence in the run up to EMU, but note that this process has essentially run out of momentum after 2002; cf. Allington, Kattuman and Waldemann (2005) and Kurkowiak (2012). Faber and Stokman (2009) show that the price level dispersion across the EMU is larger than a comparable statistic for the US. Similar conclusions are attained in other studies of the price convergence in the EMU using a variety of datasets and empirical methodologies, see Engel and Rogers (2004), Goldberg and Verboven (2005), Rogers (2007), ECB (2011), and Fischer (2012) among others. At the same time, the fact that price differentials within the EMU are significantly smaller relative to the countries outside the common currency area, coming as a consequence of the tight economic integration and elimination of the exchange rate volatility, is widely acknowledged, see Allington et al. (2005), Rogers (2007), Wolszczak-Derlacz (2008) and Imbs, Mumtaz, Ravn and Rey (2010) among many others.

However, many of the existing empirical studies of LOP deviations tend to suffer from available data limitations: broad, sufficiently homogeneous and comparable, and wide (in terms of products, locations and other dimensions) price data are inherently difficult to come by. In addition, many cross-country LOP studies are often complicated by volatile nominal exchange rates, sticky local currency prices and the need to use a common denominator; see Parsley and Wei (2001) and Cheung and Lai (2006). In this context, the euro area offers a unique opportunity for cross-border relative price comparison in the context of a single currency, the absence of trade restrictions, labour and capital mobility, and a shared regulatory framework.

This paper contributes to the empirical literature on the LOP deviations in the euro area using a new dataset of disaggregated prices, made available by the ECB to several researchers at the euro area national central banks. The dataset, compiled by Nielsen Holdings N.V. from their proprietary commercial records of European retail prices, covers 13 euro area countries over the time period 2008 to 2012 and includes 45 homogenous product varieties, with additional splits across the regional and store-type dimensions. This rich data enhance our understanding of LOP deviations in the euro area in a number of unique directions, not easily attainable with more traditional datasets. Firstly, the definitions of retail-level product categories are harmonised across countries in the sample and allow further narrowing to the level of specific brands and stock-keeping units (SKU). Secondly, the new data permit price comparisons in terms of equivalent units (kilograms, litres, rolls of paper, etc.) for each sample product variety, eliminating biases due to different consumer preferences for package sizes in different markets. Thirdly, the new data contain a wealth of information on market concentration, price ranges and volatilities within each product category, and many other micro-level attributes, which substantially contribute to the explanatory power of the estimated LOP models.

The paper focuses on the empirical i.e. "non-structural" LOP models that are simple statistics of price heterogeneity ascribed to various plausible explanatory factors. The empirical methodology is based on log-linear regressions with spatial effects in both the geographical and product-variety dimensions, estimated by the Bayesian methods. The paper shows that notable differences in relative prices of homogenous products in the euro area are linked to the following four groups of factors: (i) product-specific consumption preferences, (ii) country-specific macroeconomic and regional characteristics, (iii) volatility of prices and volumes, and (iv) spatial effects. The estimated reduced-form LOP models suggest a strong interdependence of the relative prices both on the geographical scale and across "similar" product varieties, which go beyond the included set of explanatory variables and warrant further empirical investigation.

The paper is organised as follows. Section 2 gives a detailed overview of the new Nielsen disaggregated price data and offers model-free statistical evidence of the relative price differentials in the sample euro area countries. Section 3 provides a conceptual overview of the "non-structural" LOP and Purchasing Power Parity (PPP) models, related data treatments and the econometric methodology. Section 4 documents the main empirical findings and discusses the results. Concluding remarks and potential future research directions are outlined in the last part of the paper.

2. Nielsen disaggregated price data

The "Eurosystem Project on Retail Price Analysis" was organised by the ECB in early 2012 with a mandate to address research questions pertaining to the structural price disparities of goods and services in the euro area. This project is a continuation of the past ECB efforts summarised in the 2011 Structural Issues Report on the distributive trade sector, see ECB (2011). The report, entitled "Structural features of distributive trades and their impacts on prices in the euro area", identifies a considerable heterogeneity of consumer prices in the euro area and points to the role of the structural features of the distributive trades sector in this regard. One of the conclusions of this report highlighted the lack of an in-depth analysis of the LOP deviations and related topics using a detailed disaggregated data on prices of goods and services across the euro area. Such data would cover multiple geographical sites and product varieties and be sufficiently homogeneous to allow meaningful comparisons across the whole range of product categories and locations, providing a rich supplement of additional data attributes and dimensions.

Such a dataset was compiled by Nielsen Holdings N.V. from their proprietary commercial price records of European retailers and was made available by the ECB to the ESCB researches in the beginning of 2013.[1] The dataset is very extensive, containing around 4.5 million unique records at all disaggrega-

[1]The roadmap of the ECB "Eurosystem Project on Retail Price Analysis" calls for wider availability of this research dataset for academic use by the end of 2014.

tion levels, and requires careful processing and handling prior to its use for research purposes; see Meyler (2013). This section provides a brief overview of the data and its various transformations, and presents some model-free statistics of price deviations across the sample product varieties and countries in the euro area.

The following three major dimensions of the Nielsen disaggregated prices provide a useful bird's-eye overview of the entire dataset:[2]

- *Geographical dimension*: 13 euro area countries (AT, BE, DE, EE, ES, FR, GR, IE, IT, NL, PT, SI, SK), and four to eight intra-country regions;[3]
- *Product dimension*: 45 distinct product categories, all of them groceries available from a typical European retail outlet, further split into two pan-European brands, two local brands and one private label, which in turn are sub-dividied into up to three separate SKU-s;[4]
- *Temporal dimension*: from late 2008 to early 2012, depending on the particular country and product category, with heterogeneous sampling frequencies ranging from four weeks to two months for different country-product pairs.

Each of the three aforementioned main data dimensions in turn consists of the following detailed records:

- *Sales*: expressed in euros and available at the aggregated country level, including all recorded sales for the specific product category in the proprietary Nielsen Holdings N.V. commercial data, and separately for up to thirteen selected brands and specific SKU-s;
- *Volumes*: expressed in selling units (SKU-s), and separately in equivalent units, which can be kilograms, litres, rolls of paper etc. depending on the particular product category;
- *Prices*: ratio of sales to volumes, available separately as the selling unit prices and the equivalent unit prices.

[2]The following ISO-3166-1 alpha-2 two-letter designations, shown in parentheses, are used for the euro area countres: Austria (AT), Belgium (BE), Germany (DE), Estonia (EE), Spain (ES), France (FR), Greece (GR), Ireland (IE), Italy (IT), the Netherlands (NL), Portugal (PT), Slovenia (SI), Slovakia (SK). All empirical results in this paper are ordered alphabetically by the corresponding two-letter designations.

[3]In addition to the aforementioned geographical dimension, the dataset also features the store type dimension, which can be substituted for the intra-country geographical dimension. The store types in the data range from department stores, to drug stores, and to gas stations, but are currently not sufficiently harmonised across countries and are not used in this study.

[4]The full list of available product categories can be found in the first column of Table A1.

This extensive research price data is a snapshot of the wealth of information available in the Nielsen Holdings N.V. proprietary commercial databank that contains detailed SKU-level scanner information from the participating European retailers.[5] However, it is not generic scanner-level price data: the sales and volumes along each of the three main data dimensions are time-aggregates of the underlying scanner-level data over the specific sampling frequency periods. Hence, prices in this dataset are simple ratios of sales to volumes over a specific time period, which can be as short as four weeks and as long as two calendar month depending on the particular country-product pair.

From the perspective of this study, the available price data have several advantages and a few shortcomings. The greatest advantage of the Nielsen disaggregated price dataset for studying LOP deviations in the euro area is undoubtedly its breadth and width in terms of available country-product pairs. For example, with regard to the country coverage and the number of available price observations at the SKU level, the Nielsen disaggregated price data is comparable to the Eurostat data sample in Crucini et al. (2005), widely deemed as one of the benchmark studies in the field. On the other hand, unlike Crucini et al. (2005), our data are lacking in terms of the variety of available homogenenous product categories (only groceries, covering approximately 90% of the first COICOP category, plus a few other household consumption items) and the length of the encompassing time period. However, the limited menu of available product categories is balanced by the depth of SKU-level information within each variety, which is the lowest possible denomination for a direct retail price comparison across different locations.[6]

The most serious limitation of the available data from the perspective of this paper is their short and heterogenous time dimension, making it difficult to study the dynamics of the price convergence, or lack of it, in the sample. The relative fractions of the first and last sampling dates for all available country-product pairs in the Nielsen disaggregated price dataset are depicted in Figure 1. The average time span for a typical country-product pair is between December 2008 and September 2011, corresponding to two full years and ten

[5]Nielsen Holdings N.V. is a well-established provider of commercial retail-level data, and therefore enjoys broad country coverage in terms of the participating retailers, ranging from 60% to 95%, depending on the particular country-product pair; see Meyler (2013). This ensures that the available research data are representative of the actual prices along all three main data dimensions.

[6]The limited number of product varieties and lack of services in the Nielsen disaggregated price data arguably leads to stricter product category definitions, which are easier to compare across countries and locations. While a pack of refrigerated milk may come in a number of shapes and flavours, all bunched together in a single product category in our data, a male haircut in different countries and locations ultimately lends itself to a much wider range of possible accompanying services and environments, and is therefore harder to harmonise across various data collection sites.

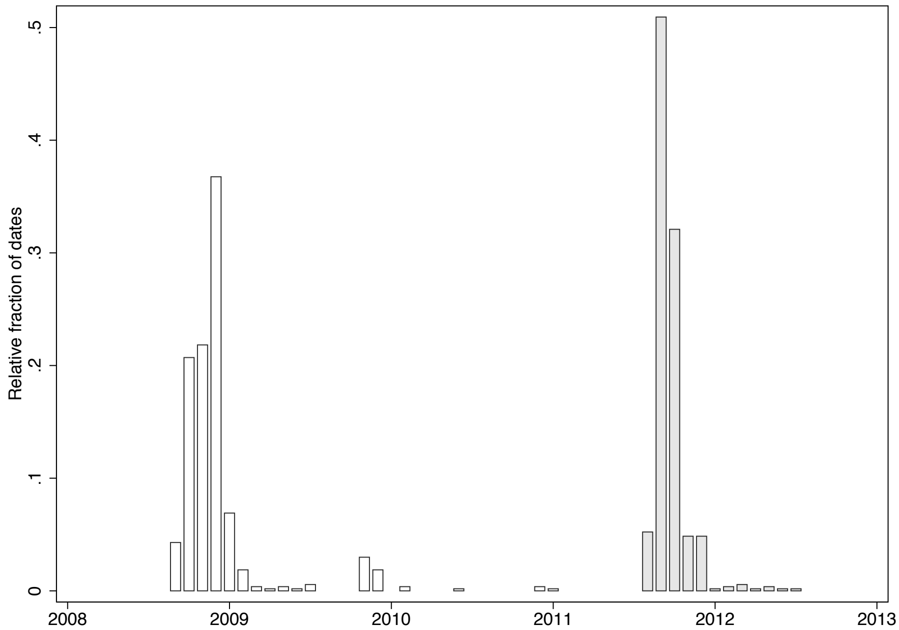

Figure 1: Relative fractions of first (white bars) and last (grey bars) sampling dates; bars correspond to calendar months

months of available data. In addition to these heterogeneous time spans, the sampling frequency of observations on sales, volumes and prices differs across countries and products, making it necessary to re-sample the observations at regular intervals prior to subsequent analyses.[7]

Some of the practical issues in handling the time dimension of the Nielsen disaggregated price dataset are illustrated in Figures 2 and 3. For example, monthly ground coffee prices in the sample euro area countries, depicted in Figure 2 , appear to converge by the end of 2011. However, an in-depth examination reveals the effect of sharply increased global coffee prices, which have virtually doubled between 2009 and 2011, pulling together the retail coffee prices in the low and high-price countries, where the latter have enjoyed a virtually unchanged price for ground coffee throughout the whole sample period. Nevertheless, there is no *a priori* assurance that the relatively low dispersion of the retail coffee prices across the euro area countries observed in our sample in the end of 2011 would not revert back to it's pre-2009 level should global

[7]For the purpose of this study, the raw data on sales and volumes were initially up-sampled to weekly frequency using the closest-neighbour linear interpolation technique, and then down-sampled to monthly frequency by summing weekly sales and volumes, accounting for the number of full and fractional weeks in each calendar month. Monthly prices were then re-calculated as the ratios of monthly sales to monthly volumes. Monthly and bi-monthly raw data, comprising up to 15% of the original sample, were left unchanged; repeated observations were used to covert bi-monthly data to the monthly frequency.

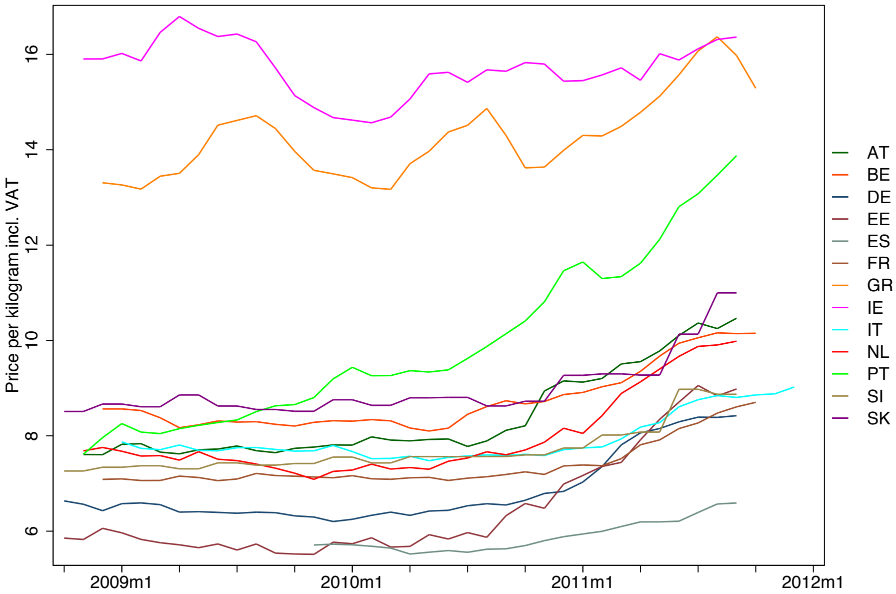

Figure 2: Monthly ground coffee prices in 13 euro area countries over the period 2008 to 2011

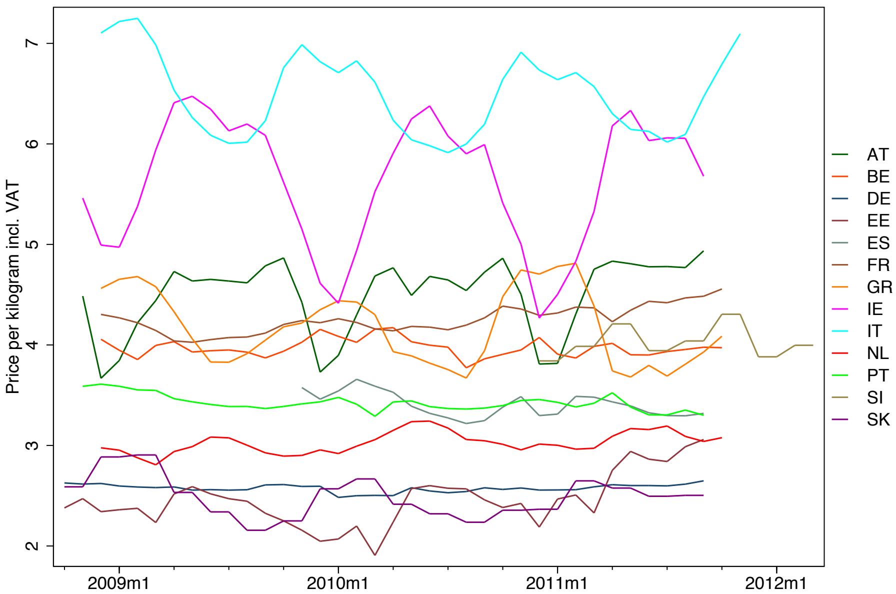

Figure 3: Monthly ice cream prices in 13 euro area countries over the period 2008 to 2012

coffee prices decrease again, since the economic fundamentals, such as diverse income levels, consumer preferences and competitive forces in the retail sector of different countries, remain virtually unchanged over the short time span of our data. Ideally, in the context of empirical LOP models, one needs to take into consideration global commodity prices to be able to distinguish different price convergence or divergence scenarios.

Figure 3 provides another glimpse of disaggregated prices along the temporal dimension. In this case, the *relative* retail prices of ice cream in different euro area countries in our sample remain virtually unchanged throughout the sample period, but there is a strong and distinct seasonal pattern in the ice cream prices in several relatively expensive countries. While deeper economic forces might be behind the observed seasonality of disaggregated prices, being oblivious to these data issues in the empirical LOP models may lead to potentially confounded findings.[8]

Considering the aforementioned short time span limitation of the Nielsen disaggregated prices and potential model specification difficulties when simultaneously operating with all three available data dimensions, we choose to average out the time span of prices in our sample and focus solely on the remaining two data dimensions.[9] The resulting product variety vis-à-vis geographical location data arrangement, akin to the layout of Table A1, simplifies the discussion of both model-free and model-based statistical evidence of relative price variations in our sample by making reference to the vertical data dimension (product-by-product over all countries) as "the LOP data window" and to the horizontal data dimension (country-by-country over all products) as "the PPP data window". All our subsequent empirical analyses in this paper will be framed according to these two complementary data windows. A similar approach was recently adopted in Crucini and Yilmazkuday (2013) in a related study of long-run factors behind the observed price dispersions in a panel of global countries.

Following conventions in the empirical LOP literature, in order to unify the scale for all diverse product prices in our sample and simplifying their comparison and the estimation of the empirical LOP models, we consider the fol-

[8]On the link between the fundamental economic forces and seasonal demand patterns, MacDonald (2000) and Chevalier, Kashyap and Rossi (2003) conject that high retail margins may facilitate steep seasonal discounts on imperfectly competitive markets and lead to lower prices during seasonal demand peaks. Curiously, just two out of four seasonal ice cream price patterns in Figure 3 appear to support this hypothesis. The empirical LOP models in Section 4 control for the retail market concentration across different country-product pairs in our sample.

[9]However, we make extensive use of the available time dimension of sales, volumes and prices in several of our data-derived explanatory variables in the empirical LOP models; see Table 1.

lowing two alternative transformations of prices per equivalent unit observed in the data into the same-scale relative prices:[10]

$$
\begin{array}{ll}
\textit{Equally-weighted}\text{ relative prices:} & p_{ij}^{E} = \log P_{ij} - \log \sum_{j=1}^{M} \frac{1}{M} P_{ij} \\
\textit{Market size-weighted}\text{ relative prices:} & p_{ij}^{M} = \log P_{ij} - \log \sum_{j=1}^{M} s_{ij} P_{ij}
\end{array}
\tag{1}
$$

where P_{ij} denotes price per equivalent unit of product $i \in \{1, \ldots, N\}$ in country $j \in \{1, \ldots, M\}$, computed from the original data as ratios of cumulated sales over cumulated volumes over the available time span, and s_{ij} is the market share of country j by sales of product i among the sample countries. The market size-weighted relative prices, while not commonly used in the empirical LOP literature, have the advantage of measuring individual country-product prices relative to the "common euro area price" for a specific product category, in the same way the regional prices would usually be measured within the boundaries of a single country.[11] Since all the countries in our sample are members of the common European market and have the same currency in circulation, we make use of the market size-weighted relative prices in our empirical analyses.[12]

Figures 4 and 5 depict "the PPP data window" into the Nielsen disaggregated price data for the equally and market–size weighted relative prices. The relative ordering of countries in the sample is almost identical across the two alternative weighting schemes, with the exception of Belgium and Slovenia, although the means of relative price distributions with respect to the "common euro area price" have different locations in the two figures. Nevertheless, both data views place Germany on the cheapest side in terms of the relative basket prices containing 45 available product varieties, while Ireland is ranked

[10] Similar product price transformations, with slight variations depending on the available data and empirical focus, are used in Crucini et al. (2005), Crucini and Shintani (2008), Lee (2010) and others.

[11] The reason it is not commonly used in the the empirical LOP literature is twofold. Firstly, most of the widely cited studies in the field compare prices across wide geographical swathes, where individual markets often have different currencies and share no common border. In this context a market size-weighted relative price lacks an obvious economic interpretation. Secondly, an admittedly extreme interpretation of the LOP posits the same price for the same products regardless of any features of the market or geography; in this context it makes sense to compute prices relative to the equally-weighted benchmark.

[12] Estonia and Slovakia are the two late entrants to the common currency area in our sample. Prior to becoming a full euro area member in January 2011, Estonia adhered to an orthodox currency board arrangement vis-à-vis the euro, and therefore had no significant exchange rate risks. After more than three years of participation in the European Exchange Rate Mechanism II, Slovakia finally joined the euro area in January 2009, with only a few observations in the Nielsen disaggregated price data prior to this date.

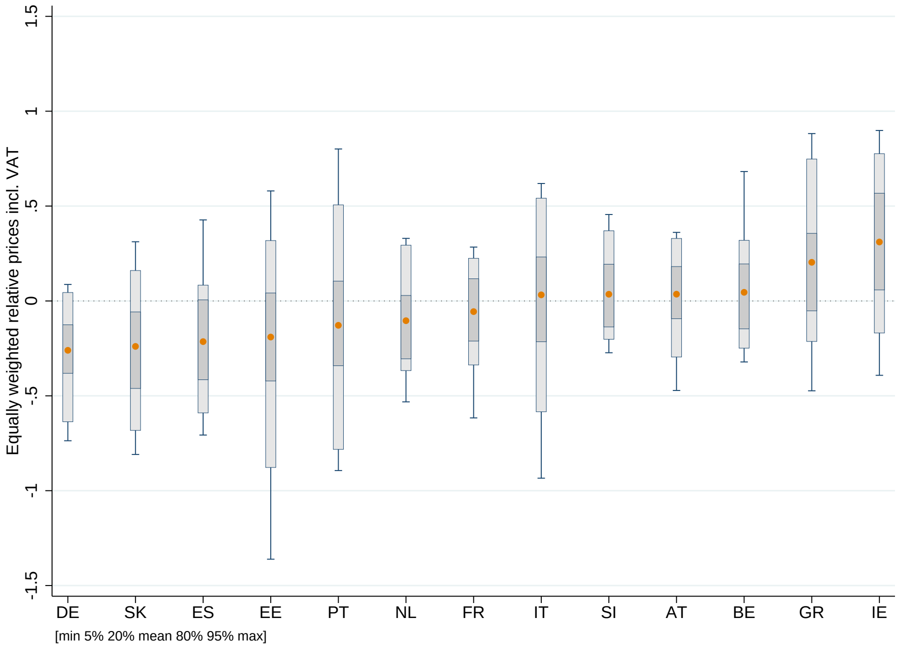

Figure 4: PPP data window: empirical quantiles of p_{ij}^{E} in 13 euro area countries

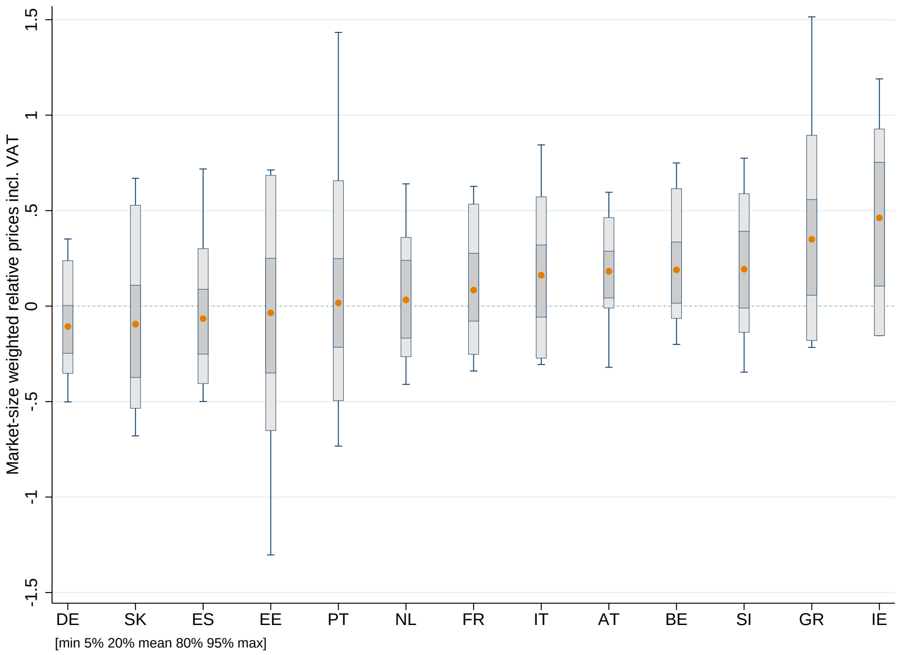

Figure 5: PPP data window: empirical quantiles of p_{ij}^{M} in 13 euro area countries

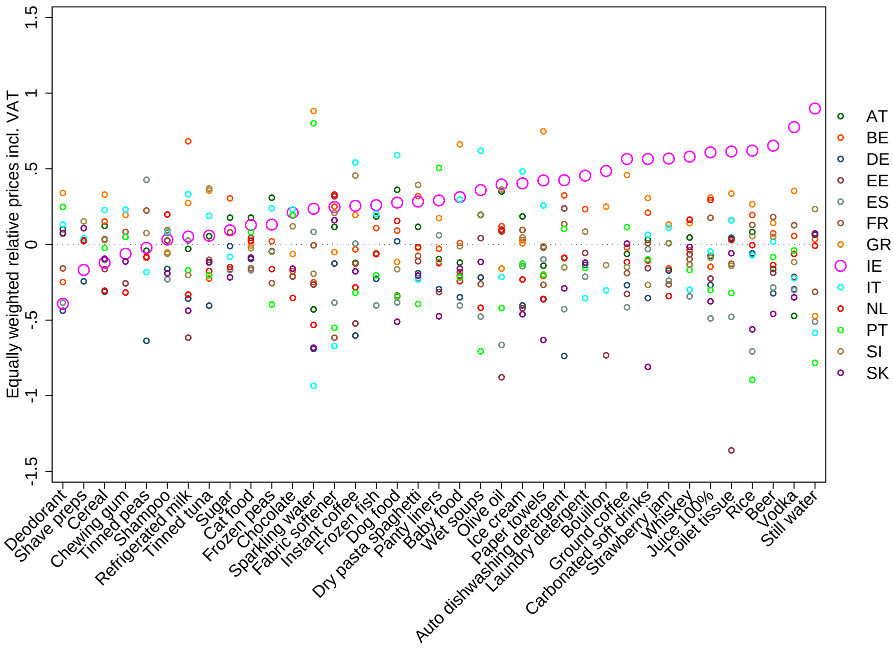

Figure 6: LOP data window: distribution of p_{ij}^E, ordered by 37 products available for Ireland

the most expensive country according to our PPP data view.[13] In addition, the smaller countries in our sample, such as Estonia, Greece, Ireland and Portugal, tend to have wider distributions of the relative basket prices. This observation appears to be consistent with Crucini et al. (2005), who obtain similar distributions of relative price deviations using a different dataset, but do not draw specific attention to this empirical feature.

Figures 6 and 7 illustrate "the LOP data window" into the Nielsen disaggregated price data for two countries in our sample, Ireland and Portugal, and two alternative price weighting schemes. Ireland, the most expensive country on the PPP scale, has only five product categories with the relative prices below or equal to the equally weighted sample average, as depicted in Figure 6. In terms of the market-size weighted relative prices, Portugal's basket of 39 available product categories is very close to the "common euro area price", as seen through the PPP data window in Figure 5 and further elaborated by the LOP view in Figure 7. However, Portugal's relative prices are quite disperse, with

[13]Our findings imply a different relative ordering of the euro area countries from that reported in Kurkowiak (2012), where the Eurostat data on price levels are used. In particular, the German consumption basket should have been more expensive than that in Estonia, Greece, Portugal, Slovenia, Slovakia and Spain in 2011 according to the Eurostat results. The statistical methodology and dataset used in Kurkowiak (2012) are markedly different from the present study. In particular, all items in our PPP baskets in Figures 4 and 5 have equal weights.

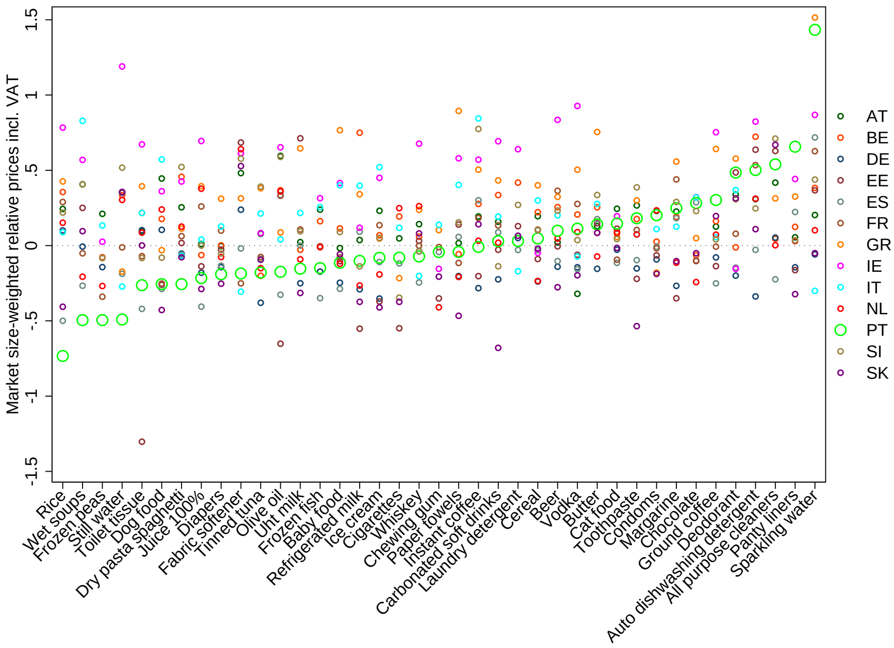

Figure 7: LOP data window: distribution of p^M_{ij}, ordered by 39 products available for Portugal

about nine product varieties outside the $\pm 50\%$ relative price interval around the market-size weighted sample average.[14]

In summary, the model-free statistical evidence of LOP price deviations in the euro area, based on the Nielsen disaggregated price data and documented in this section, yields empirical results that are broadly similar to the previous findings in Crucini et al. (2005) and others. A reduced-form model-based approach to explain the observed variation of relative prices is pursued in the remaining part of this paper.

3. Law of One Price models: empirical methodology

The Law of One Price states that identical goods and services in different locations (countries, regions, shops, etc.) should have identical prices ex-

[14]Our prices, by construction, are relative to the weighted or unweighted sample averages for each product category. However, in terms of the empirical distributions shown in Figures 4 to 7, each product category has an equal weight in the consumption basket. An alternative approach, not pursued in this paper, is to re-weight the product varieties according to their observed consumption basket weight.

pressed in the common currency due to forces of competition and arbitrage. Purchasing Power Parity is a related notion, whereby the LOP should hold on average: similar baskets of products should cost the same in different locations. Both of these concepts feature static and dynamic components: at each point in time they characterise cross-sections of prices of goods and services and act as a "gravity force", pulling together prices of similar goods and services over time across different locations; see Rogoff (1996).

The microeconometric LOP and PPP studies offer a variety of alternative models and explanations of the observed price heterogeneity of similar products across different countries and locations. The predominant conceptual framework is due to Engel and Rogers (1996), Engel and Rogers (2004) and Crucini et al. (2005), where the final retail price is determined by a combination of tradable and non-tradable inputs, where the tradable input prices are largely homogeneous, whereas the non-tradable input prices exhibit significant heterogeneity across countries:

$$P_{ij} = a_{ij}\, W_j^{\gamma_{ij}}\, T_i^{1-\gamma_{ij}}\ ,$$

where a_{ij} is the retail mark-up, W_j is a non-tradable country-specific retail input price, T_i is a tradable product-specific retail input price, γ_{ij} is the retail technology parameter, and the country and product indices defined in (1). Although conceptually simple, this retail technology model is not trivial to apply to the real-world data due to the large number of free parameters that are hard to estimate or find a good proxy for.

One of the major features of the Nielsen disaggregated prices, setting them apart from many previous studies in the empirical LOP literature, is the manifestly retail nature of all product varieties in the sample together with the complete absence of a consumer services dimension in the available data. One of the predominant explanatory factors behind the LOP deviations in the seminal studies by Engel and Rogers (1996) and Crucini et al. (2005) is the variation of the retail technology parameter γ_{ij} across different country-product pairs. The popular empirical methodology for estimating γ_{ij} in these models is to use the input-output tables for a cross-section of countries and some broadly defined categories of products and services. However, given the common European market framework and rather homogeneous retail technology across all the countries in our sample, we are unlikely to find sufficient variation in γ_{ij} among our retail grocery product varieties to explain the cross-product and cross-country LOP deviations. We therefore pursue a "non-structural" approach to empirical LOP regressions in this paper. In particular, we group different explanatory variables in our empirical models according to their presumed similarity to W_j, which is largely a "country-specific" factor, or to T_i, which can be viewed as a "product-specific" feature, or to a_{ij}, which combines both.

The type of a "non-structural" LOP regression for relative prices, considered in this paper, is given by:

$$p_{ij} = \tilde{\boldsymbol{x}}_{ij}^{\mathrm{T}} \boldsymbol{\beta} + \varepsilon_{ij} \,, \tag{2}$$

where $p_{ij} \in \{p_{ij}^E, p_{ij}^M\}$ are defined in (1) and the regressors $\tilde{\boldsymbol{x}}_{ij}$ are organised around the previously outlined conceptual retail technology framework, falling into the product-specific, country-specific and country-product specific factors, but without any explicitly attached "structural" interpretations. Hence, our empirical approach to the Nielsen disaggregated prices combines both the "LOP data window" and the "PPP data windows" in terms of the interpretation of the reduced-form parameters $\boldsymbol{\beta}$.

In order to proceed from the concept of "non-structural" LOP regression in (2) to its actual application to the real-word data, we consider a possible dependence structure of the innovations ε_{ij}, which is crucial for the reliability of the statistical inference on the parameters $\boldsymbol{\beta}$. As detailed in Section 2, we treat our sample as a two-dimentional table of product varieties vis-à-vis geographical locations, after integrating out the time span of raw disaggregated prices. It is therefore essential to consider interactions across equations given in (2) in both of these two dimensions simultaneously: a higher price of a given product variety in a given country (a larger ε_{ij} innovation) may spill over to the "neighbouring" countries within the same product category (larger ε_{ij}-s along j for the same i) *and* to "similar" products within the same country (larger ε_{ij}-s along i for the same j). From the statistical perspective, the two-dimensional array of innovations $\{\varepsilon_{ij} : 1 \leq i \leq N, 1 \leq j \leq M\}$ in our "non-structural" LOP regression (2) may not be an array of independent disturbances in either dimension, and this needs to be considered in the empirical applications of the model; see Whittle (1954) and Ord (1975).[15]

This type of dependence is known to be spatial in nature and is described by the class of simultaneous autoregressions (SAR):

$$y_k = \boldsymbol{x}_k^{\mathrm{T}} \boldsymbol{\beta} + \phi \cdot \sum_{m=1}^{K} w_{km} (y_m - \boldsymbol{x}_m^{\mathrm{T}} \boldsymbol{\beta}) + \epsilon_k \,, \tag{3}$$

[15]In this study we do not offer any "structural" justifications of the cross-border and cross-product interactions of relative prices, beyond the aforementioned statistical considerations. A more structural approach to relative price spillovers across borders and product varieties, along the lines of Behrens et al. (2012), may potentially be of interest. For example, in a more structural framework, the cross-border spillovers may be linked to competition and transportation costs, while the cross-product interactions may depend on the underlying elasticity of substitution parameters between the product varieties.

where the observations are indexed by $k = 1, \ldots, K$, and K denotes the sample size.[16] Deterministic spatial weights in (3) are denoted by w_{km}, where $w_{kk} := 0$ for all $1 \leq k \leq K$, representing the generalised notion of "proximity" and dependence between different cross-sectional units in the sample. The unknown parameter ϕ determines the overall degree of spatial interaction in the data, where the case of $\phi = 0$ corresponds to the classical linear regression model with i.i.d. innovations. This model can be written in the matrix notation as follows, revealing the dependence structure of innovations between the cross-sectional units:

$$\boldsymbol{y} = \mathbf{X}\boldsymbol{\beta} + (\mathbf{I}_K - \phi\mathbf{W})^{-1}\boldsymbol{\epsilon}\,,$$

where $\mathbf{W}$ is a $K \times K$ non-stochastic matrix of spatial weights, and $\mathbf{I}_K$ denotes an $K \times K$ identity matrix.

The appropriate parameterisation of spatial weights in the SAR model (3) is crucial for obtaining robust statistical inference on the unknown parameters $\boldsymbol{\beta}$ and ϕ. Motivated by the previously outlined interactions between relative prices in a two-dimensional table of product varieties vis-à-vis geographical locations, we propose a two-layered structure of the spatial weights in our empirical LOP model. To be specific, we superimpose the generalised spatial dependence across "similar" product varieties in our data on the conventional geographical proximity-based spillovers between relative prices across the sample euro area countries. In particular, the estimated "non-structural" LOP regressions in Section 4 make use of the following two spatial weighting matrices: (i) $\mathbf{W}_M$ is a 13×13 "neighbourhood" matrix with the elements linked to geographical proximity between the national capitals in our sample; (ii) $\mathbf{W}_N$ is a 45×45 "similarity" matrix with the elements related to the generalised proximity of the sample product varieties in terms of their COICOP codes.

The conventional cross-border "neighbourhood" matrix $\mathbf{W}_M$ is based on the widely used inverse distance weights, where the geographical distances are measured in kilometres between the national capitals of the 13 countries in our sample; see Cliff and Ord (1981). In addition, $\mathbf{W}_M$ is normalised by dividing all elements with the maximum spatial weight between the two closest countries in the sample (Austria and Slovakia).

The cross-product "similarity" matrix $\mathbf{W}_N$ is based on the generalised notion of proximity, whereby weighted absolute differences of the COICOP

[16] The spatial dependence is traditionally associated with cross-correlations throughout the physical space, e.g. crop yields in a field divided into smaller plots or crime rates across the US states; see Cliff and Ord (1981). However, the mathematical definition of proximity spaces is inherently abstract, extending beyond the obvious geography and time dimensions. Recent econometric applications of spatial methods based on a generalised notion of "similarity" between cross-sectional units include Conley and Topa (2002) and Behrens et al. (2012).

codes, shown in Table A1 for each product variety in the sample, are used to define the spatial weights in the following way:

$$S_{ip} := 1000 \cdot |C_{1i} - C_{1p}| + 100 \cdot |C_{2i} - C_{2p}| + 10 \cdot |C_{3i} - C_{3p}| + 1 \cdot |C_{4i} - C_{4p}| \,,$$

where the individual digits of the four-digit COICOP codes are $C_{1i}C_{2i}C_{3i}C_{4i}$, and $1 \leq i, p \leq N$ index over the sample products. The elements of $\mathbf{W}_N$ matrix are given by $\frac{1}{1+S_{ip}}$, where the product varieties belonging to exact same category in terms of their COICOP codes receive the maximum "similarity" weight of one.

Finally, the combined symmetric $K \times K$ spatial weighting matrix $\mathbf{W}$ is defined as:

$$\mathbf{W} := \frac{\mathbf{W}_N \otimes \mathbf{W}_M}{\lambda_{\max}} \,, \tag{4}$$

where the normalisation factor $\lambda_{\max}$ makes the largest positive eigenvalue of $\mathbf{W}$ equal to one, and all elements of $\mathbf{W}$ along the main diagonal are set to zero.[17]

Given the structure of spatial weights shown in (4), the country-product pairs of relative prices and other explanatory variables need to be flattened into the vector of dependent variables $\boldsymbol{y}$ and the matrix of regressors $\mathbf{X}$ by sorting in product-major order. In terms of the "non-structural" LOP regression (2), $\boldsymbol{y}$ consists of relative prices p_{ij} and the rows of $\mathbf{X}$ are made up of $\tilde{\boldsymbol{x}}_{ij}^{\mathrm{T}}$, where the indices are sorted first by i and then by j in the order they appear respectively in $\mathbf{W}_N$ and $\mathbf{W}_M$. The vertical dimension of the resulting data matrices is $K = N \times M$, before adjusting for the missing observations. In the empirical LOP regressions in Section 4 the missing data are tackled by selectively deleting the corresponding rows and columns of $\mathbf{W}$ according to Table A1.

Under the assumption of independent normally distributed innovations in the SAR model (3), $\epsilon_k \sim$ i.i.d. $\mathcal{N}(0, \sigma^2)$, the data distribution is normal with the following parameters:

$$\boldsymbol{y} \sim \mathcal{N}(\mathbf{X}\boldsymbol{\beta}, \sigma^2(\mathbf{I}_K - \phi\mathbf{W})^{-2}) \,,$$

and the likelihood function has the standard form. Given the "non-structural" nature of our empirical LOP analyses in Section 4, we impose a minimum of *a priori* information on the model parameters, following the recommendations in Oliveira and Song (2008). In particular, we impose a flat non-informative prior on $\boldsymbol{\beta}$, the usual flat Jeffreys prior on σ^{-2} and a uniform prior on the spatial parameter $\phi \in (\lambda_{\min}^{-1}, 1)$, where $\lambda_{\min}$ denotes the smallest eigenvalue of the spatial weighting matrix $\mathbf{W}$.

[17]The normalisation of $\mathbf{W}$ by $\lambda_{\max}$ is optional, ensuring that $\phi \leq 1$; see Oliveira and Song (2008).

The posterior sampling algorithm for the SAR model (3) with the aforementioned priors is outlined in Oliveira and Song (2008). It is based on a combination of the usual Gibbs sampler for $\boldsymbol{\beta}$ and σ^2 with a random walk Metropolis-Hastings (RWMH) step for the spatial parameter ϕ. In all empirical analyses in Section 4 we run the Gibbs-RWMH sampler for 4000 iterations, starting from a suitable prior distribution draw, and use the tail 50% of simulations for calculating the posterior statistics.[18]

In the context of empirical LOP regressions, the spatial approach similar to the one proposed in this paper is used by Keller and Shiue (2007) in a historical study of rice prices across 121 prefectures in 18th century China. However, since they examine a single product, their spatial weighting matrix is based on the usual notion of geographical proximity. This paper proposes an extension of the spatial LOP modelling approach by considering spillovers across different product varieties in addition to the conventional geographical dimension. Section 4 presents empirical results of the LOP regressions estimated using the Nielsen disaggregated price data.

4. Empirical LOP regressions for the euro area

This section provides a discussion of the main empirical findings from the LOP models estimated using the Nielsen disaggregated price data. The main feature of the models, as outlined in Section 3, is spatial dependence across both the geographical and product variety dimensions of the data.

A detailed breakdown of the available sample of relative prices by country-product pairs is given in Table A1. There are $K = 517$ unique country-product pairs of prices in the dataset, with a fairly uniform coverage in terms of product varieties by country and vice versa. Corresponding to each country-product relative price observation, a set of fourteen explanatory variables was constructed according to the "non-structural" LOP regression concept discussed in Section 3. Detailed definitions of the explanatory variables used in the empirical LOP models are given in Table 1. Most of the available explanatory variables are country and product specific; many of them are derived using the time, brand and SKU-level dimensions of the Nielsen disaggregated price dataset, while the dependent variable is based on the time-averaged aggregated country-level data, as detailed in Section 2. We also make use of certain macroeconomic observables external to the Nielsen disaggregated price dataset, e.g. income per capita and average economic growth as proxies for

[18] All computations in this paper are carried out in Ox matrix programming language; see Doornik (2007). An in-depth exposition of the Markov chain Monte Carlo sampling methods can be found in Robert and Casella (2004).

the country-specific factors. Finally, our empirical LOP models also include a full set of product dummies for characterising the pan-European product market features, such as degree of market concentration, pricing power and product-specific market regulations for each available product variety.

Estimation results for the empirical LOP models based on equally and market size-weighted relative prices are shown in Tables 2, 3 and 4. The first set of results in Table 2 comprise the most comprehensive empirical LOP model that includes the spatial dependence across both the geographical and product market dimensions of the data. The estimated spatial parameter ϕ indicates a strong dependence between the country-product pairs of relative prices in the dataset, even after conditioning on the set of included explanatory factors. This may be an indication of omitted explanatory variables in our reduced-form LOP regressions or a sign of strong intrinsic links between country and product pairs of the relative prices in the sample. In either case, the statistical implications of this finding are important.

In order to disentangle the spatial dependence in the sample and shed some additional light on its nature, we re-estimate our empirical LOP regressions including one of the two spatial weight matrices at a time. The results are shown in Tables 3 and 4. It can be observed from the estimated spatial parameters in the tables that the strongest link between relative prices in our sample is found across product varieties, while the degree of conventional geographical dependence among the relative prices in the sample is relatively muted at this aggregation level.[19] Interestingly, the combined effect of the cross-product and cross-country spatial dependence appears to be higher than the sum of the individual effects, as witnessed by the relative magnitudes of the estimated spatial parameters in respectively Table 2 and Tables 3 and 4.

Although we do not give a complete structural interpretation of the observed spatial dependence of relative prices, the likely interpretation of the results in Tables 2 to 4 is related to the common production technology and demand characteristics of different product varieties in our sample. To be specific, the production technology of "similar" retail goods in all countries may depend on a certain single global commodity price, giving rise to a global price shock propagation channel throughout a subset of "similar" product varieties in all sample countries simultaneously. At the same time, different degrees of market concentration, pricing power and country-by-country market demand variations for a given product category are likely to lead to geographically correlated price shock propagation mechanisms in our data.

[19]Recall that we define country-product pairs of relative prices at the highest data aggregation level, i.e. whole country aggregates for each product variety; see Section 2. It likely that the degree of geographical dependence between relative prices at the lower aggregation level of intra-country regions will be substantially more pronounced; see Keller and Shiue (2007).

Table 1: Definitions of explanatory variables in the empirical LOP models

Explanatory variable	*Definition*
Log max/min equivalent unit price over time[(a)]	Range of prices over time, calculated as the log ratio of maximum equivalent unit price to its minimum over the monthly series from 2008 to 2012 at the aggregated country level
Log max/min equivalent unit price in cross-section[(a)]	Range of prices in cross-section, calculated as the log ratio of maximum equivalent unit price to its minimum over the time-aggregated brand and SKU-level data
Log max/min equivalent unit volume in cross-section[(a)]	Range of volumes in cross-section, calculated as the log ratio of maximum equivalent unit volume to its minimum over the time-aggregated brand and SKU-level data
Log country-product share by selling units[(a)]	Log of a given country share in terms of its SKU volumes in the aggregated SKU volume of all sample countries
Log average selling unit size[(a)]	Log of an average SKU size, computed as the SKU volume over the corresponding equivalent unit volume
Log equivalent units volume per capita[(a)¶]	Log of per capita product consumption, calculated as the equivalent unit volume over the country population 15 to 99 years of age
Price elasticity[(a)]	Price elasticity coefficient, estimated using both the time- and SKU-level information in the dataset
Log selling units share of the cheapest brand[(a)]	Log of SKU volume share of the cheapest brand in terms of its price per one selling unit
Log real GDP per capita[(b)†]	Log of country real GDP per capita, averaged over the sample time period from 2008 to 2012
Annual real GDP growth[(b)†]	Annual GDP growth rate, averaged over the sample time period from 2008 to 2012
Log dispersion of regional GDP[(b)‡]	Log of dispersion of regional GDP at the NUTS 3 disaggregation level
Annual inflation by COICOP[(a)§]	Annual inflation rate derived from national HICP sub-indcies, where the product correspondence is by its COICOP category, averaged over the sample time period from 2008 to 2012
Log product weight in HICP by COICOP[(a)§]	Log of product weight in the national HICP of a given country, where the product correspondence is by its COICOP category
Log VAT rate[(a)#]	Log of product and country-specific VAT rate, averaged over the sample time period from 2008 to 2012
Product-specific dummies[(c)]	Are included in the empirical LOP models to describe product-specific features of prices, such as a degree of pan-European market concentration and regulations for the specific product variety

Notes: Prices of cigarettes, beer, vodka and whiskey are adjusted for country and time-specific excise duties. The explanatory variables are classified as (a) country and product specific, (b) country specific, (c) product specific. The external data sources used in the explanatory variable definitions are:

¶ Eurostat population data
† Eurostat macroeconomic data
‡ Eurostat regional data
§ Eurostat price data
# ECB internal data

Table 2: Empirical LOP models with spatial dependence across products and countries

Model	*[2.30%*	*16.0%*	*50.0%*	*84.0%*	*97.7%]*
Linear regression of p_{ij}^E on:					
Spatial effects (ϕ)	0.487	0.630	0.753	0.859	0.933
Log max/min equivalent unit price over time	–0.299	–0.275	–0.255	–0.237	–0.219
Log max/min equivalent unit price in cross-section	–0.026	–0.024	–0.021	–0.019	–0.017
Log max/min equivalent unit volume in cross-section	–0.015	–0.015	–0.014	–0.014	–0.013
Log country-product share by selling units	–0.034	–0.032	–0.031	–0.029	–0.027
Log average selling unit size	–0.346	–0.341	–0.337	–0.333	–0.328
Log equivalent units volume per capita	–0.094	–0.092	–0.090	–0.088	–0.086
Price elasticity	–0.110	–0.106	–0.101	–0.097	–0.093
Log selling units share of the cheapest brand	–0.002	–0.001	–0.000	0.001	0.002
Log real GDP per capita	0.362	0.365	0.369	0.372	0.376
Annual real GDP growth	–0.058	–0.057	–0.056	–0.054	–0.053
Log dispersion of regional GDP	0.238	0.246	0.253	0.261	0.269
Annual inflation by COICOP	–0.000	0.000	0.001	0.002	0.002
Log product weight in HICP by COICOP	0.046	0.049	0.051	0.054	0.056
Log VAT rate	0.022	0.024	0.027	0.029	0.031
σ^2 $(\cdot 10^{-3})$	0.361	0.385	0.413	0.441	0.477
Linear regression of p_{ij}^M on:					
Spatial effects (ϕ)	0.481	0.616	0.758	0.853	0.927
Log max/min equivalent unit price over time	–0.298	–0.275	–0.253	–0.234	–0.218
Log max/min equivalent unit price in cross-section	–0.027	–0.025	–0.023	–0.020	–0.018
Log max/min equivalent unit volume in cross-section	–0.016	–0.016	–0.015	–0.015	–0.014
Log country-product share by selling units	–0.037	–0.035	–0.034	–0.033	–0.031
Log average selling unit size	–0.356	–0.352	–0.348	–0.342	–0.337
Log equivalent units volume per capita	–0.087	–0.085	–0.083	–0.081	–0.079
Price elasticity	–0.122	–0.118	–0.114	–0.109	–0.105
Log selling units share of the cheapest brand	0.001	0.002	0.003	0.004	0.004
Log real GDP per capita	0.360	0.364	0.367	0.371	0.374
Annual real GDP growth	–0.058	–0.057	–0.055	–0.054	–0.052
Log dispersion of regional GDP	0.226	0.234	0.242	0.250	0.258
Annual inflation by COICOP	–0.000	0.000	0.001	0.001	0.002
Log product weight in HICP by COICOP	0.047	0.050	0.053	0.055	0.058
Log VAT rate	0.026	0.028	0.030	0.032	0.034
σ^2 $(\cdot 10^{-3})$	0.368	0.394	0.420	0.448	0.483

Notes: SAR model with spatial dependence between products and countries based on a combination of inverse COICOP proximity and distance weights. Posterior 96% and 68% confidence sets are displayed in the columns. The full set of product dummies is included in each estimated model, but not shown in the table. The number of Gibbs-RWMH draws is 4000, where the last 2000 draws are used for the posterior inference. Sample size $K = 517$.

Table 3: Empirical LOP models with spatial dependence across product varieties

Model	*[2.30%*	*16.0%*	*50.0%*	*84.0%*	*97.7%]*
Linear regression of p_{ij}^{E} on:					
Spatial effects (ϕ)	0.331	0.438	0.519	0.611	0.691
Log max/min equivalent unit price over time	–0.270	–0.246	–0.222	–0.198	–0.175
Log max/min equivalent unit price in cross-section	–0.027	–0.025	–0.023	–0.020	–0.018
Log max/min equivalent unit volume in cross-section	–0.016	–0.015	–0.015	–0.014	–0.014
Log country-product share by selling units	–0.037	–0.035	–0.034	–0.032	–0.031
Log average selling unit size	–0.358	–0.353	–0.347	–0.341	–0.335
Log equivalent units volume per capita	–0.090	–0.087	–0.085	–0.082	–0.080
Price elasticity	–0.113	–0.108	–0.104	–0.100	–0.095
Log selling units share of the cheapest brand	–0.004	–0.003	–0.002	–0.001	0.000
Log real GDP per capita	0.368	0.373	0.377	0.382	0.387
Annual real GDP growth	–0.056	–0.054	–0.053	–0.052	–0.050
Log dispersion of regional GDP	0.198	0.207	0.216	0.224	0.232
Annual inflation by COICOP	–0.001	0.000	0.001	0.001	0.002
Log product weight in HICP by COICOP	0.037	0.040	0.044	0.047	0.050
Log VAT rate	0.024	0.026	0.028	0.030	0.033
σ^2 $(\cdot 10^{-3})$	0.369	0.395	0.421	0.450	0.484
Linear regression of p_{ij}^{M} on:					
Spatial effects (ϕ)	0.356	0.444	0.528	0.604	0.693
Log max/min equivalent unit price over time	–0.269	–0.244	–0.220	–0.198	–0.176
Log max/min equivalent unit price in cross-section	–0.028	–0.026	–0.024	–0.022	–0.020
Log max/min equivalent unit volume in cross-section	–0.017	–0.016	–0.015	–0.015	–0.014
Log country-product share by selling units	–0.041	–0.039	–0.037	–0.036	–0.034
Log average selling unit size	–0.372	–0.365	–0.358	–0.352	–0.345
Log equivalent units volume per capita	–0.083	–0.080	–0.077	–0.074	–0.071
Price elasticity	–0.126	–0.122	–0.117	–0.112	–0.108
Log selling units share of the cheapest brand	–0.001	0.000	0.001	0.002	0.003
Log real GDP per capita	0.367	0.371	0.375	0.380	0.385
Annual real GDP growth	–0.056	–0.055	–0.053	–0.052	–0.050
Log dispersion of regional GDP	0.186	0.195	0.204	0.213	0.221
Annual inflation by COICOP	–0.001	0.000	0.001	0.001	0.002
Log product weight in HICP by COICOP	0.039	0.042	0.045	0.049	0.052
Log VAT rate	0.027	0.029	0.031	0.034	0.036
σ^2 $(\cdot 10^{-3})$	0.373	0.401	0.428	0.458	0.489

Notes: SAR model with spatial dependence between product varieties based on inverse COICOP proximity metric. Posterior 96% and 68% confidence sets are displayed in the columns. The full set of product dummies is included in each estimated model, but not shown in the table. The number of Gibbs-RWMH draws is 4000, where the last 2000 draws are used for the posterior inference. Sample size $K = 517$.

Table 4: Empirical LOP models with spatial dependence across countries

Model	*[2.30%*	*16.0%*	*50.0%*	*84.0%*	*97.7%]*
Linear regression of p_{ij}^{E} on:					
Spatial effects (ϕ)	–0.074	0.039	0.151	0.272	0.377
Log max/min equivalent unit price over time	–0.391	–0.378	–0.366	–0.355	–0.343
Log max/min equivalent unit price in cross-section	–0.030	–0.028	–0.026	–0.024	–0.022
Log max/min equivalent unit volume in cross-section	–0.016	–0.016	–0.015	–0.014	–0.014
Log country-product share by selling units	–0.035	–0.034	–0.033	–0.032	–0.031
Log average selling unit size	–0.322	–0.319	–0.315	–0.312	–0.308
Log equivalent units volume per capita	–0.099	–0.097	–0.095	–0.092	–0.090
Price elasticity	–0.110	–0.106	–0.102	–0.097	–0.093
Log selling units share of the cheapest brand	0.001	0.002	0.003	0.004	0.004
Log real GDP per capita	0.362	0.365	0.369	0.372	0.376
Annual real GDP growth	–0.058	–0.057	–0.056	–0.055	–0.054
Log dispersion of regional GDP	0.235	0.242	0.249	0.255	0.262
Annual inflation by COICOP	0.001	0.001	0.002	0.002	0.003
Log product weight in HICP by COICOP	0.052	0.054	0.057	0.059	0.062
Log VAT rate	0.022	0.024	0.026	0.028	0.030
$\sigma^2\ (\cdot 10^{-3})$	0.341	0.365	0.389	0.415	0.444
Linear regression of p_{ij}^{M} on:					
Spatial effects (ϕ)	–0.054	0.060	0.168	0.278	0.374
Log max/min equivalent unit price over time	–0.389	–0.377	–0.365	–0.352	–0.341
Log max/min equivalent unit price in cross-section	–0.030	–0.028	–0.026	–0.024	–0.022
Log max/min equivalent unit volume in cross-section	–0.017	–0.016	–0.016	–0.015	–0.014
Log country-product share by selling units	–0.037	–0.036	–0.035	–0.034	–0.033
Log average selling unit size	–0.329	–0.326	–0.323	–0.319	–0.316
Log equivalent units volume per capita	–0.093	–0.091	–0.089	–0.087	–0.085
Price elasticity	–0.119	–0.115	–0.110	–0.106	–0.102
Log selling units share of the cheapest brand	0.004	0.005	0.006	0.006	0.007
Log real GDP per capita	0.360	0.363	0.367	0.370	0.374
Annual real GDP growth	–0.058	–0.057	–0.056	–0.055	–0.054
Log dispersion of regional GDP	0.227	0.233	0.240	0.246	0.253
Annual inflation by COICOP	0.001	0.001	0.002	0.002	0.002
Log product weight in HICP by COICOP	0.054	0.057	0.059	0.062	0.064
Log VAT rate	0.025	0.027	0.029	0.031	0.033
$\sigma^2\ (\cdot 10^{-3})$	0.350	0.370	0.396	0.422	0.451

Notes: SAR model with spatial dependence between euro area countries based on inverse distance weights. Posterior 96% and 68% confidence sets are displayed in the columns. The full set of product dummies is included in each estimated model, but not shown in the table. The number of Gibbs-RWMH draws is 4000, where the last 2000 draws are used for the posterior inference. Sample size $K = 517$.

Turning now to the estimated effects of other explanatory variables in our empirical LOP models, we report a number of reduced-form results that appear to be in line with the conceptual retail technology framework outlined in Section 3. In particular, there is a strong effect of the non-tradable input cost component in the relative prices, as verified by the large and positive effects of the real per capita GDP and the dispersion of the intra-country regional GDP explanatory factors. At the same time, a higher level of annual real GDP growth appears to have a moderating impact on the relative prices, which can be explained by the pro-cyclical competition effects in the retail and product markets.

There also is a compelling set of results linked to economies of scale and market demand effects on the relative prices across products and countries in our sample. Country and product-specific shares and volumes per capita (as a product demand proxy) show a strong negative effect on the relative prices, as does the average selling unit size explanatory variable. These factors are linked to the retail markup part of the conceptual retail technology framework outlined in Section 3, but are likely to reflect both the retail-side scale effects and the demand-side consumption preferences in different countries. Another robust result from our empirical LOP regressions is linked to the estimated country and product-specific price elasticity factor: more price-sensitive product varieties exhibit lower relative prices in line with the prevailing economic intuition. At the same time, we find that the volume effect of the cheapest brand, which most of the time tends to be a supermarket private label, on the relative prices is close to zero.

A set of explanatory factors linked to the volatility and width of price distribution across products and countries is also reported to be a statistically important explanatory factor in our empirical LOP regressions in Tables 2, 3 and 4. A particularly pronounced negative effect on the relative prices is found for the time volatility of prices. A plausible interpretation of this reduced-form explanatory factor may again be related to the degree of retail competition in a specific market: a prevalence of seasonal sales and discounts, which increases the observed time volatility of prices, may indicate stronger retail sector competition and lower relative prices.

Finally, the effect of product market regulations is reflected in the estimated positive contribution of the average VAT rate across products and countries in our sample.

In summary, the estimated empirical LOP regressions using the Nielsen disaggregated price dataset reveal a nuanced picture of the relative prices across a sample of 13 euro area countries and 45 product varieties. In particular, we find that the relative income levels and economic growth strongly affect relative prices in our sample. In addition, a number of significant price effects are linked to the economies of scale, market demand and consumer preferences. At the same time, our empirical LOP regressions suggest a strong interdependence of relative prices across both the geographical and product-variety data dimensions, which go beyond the included set of explanatory variables and warrant further empirical investigation.

5. Conclusion

This paper documents the results of the empirical LOP regressions estimated using the newly available ECB Nielsen disaggregated price dataset, which covers 13 euro area countries and 45 homogeneous product varieties over the time period from 2008 to 2012. The empirical methodology is based on non-structural log-linear regressions with spatial effects in both the geographical and product-variety dimensions, estimated by the Bayesian methods. The models link the relative prices of homogenous products to four distinct groups of factors: product-specific consumption preferences, country-specific macroeconomic and regional characteristics, volatility of prices and volumes, and spatial effects.

The estimated empirical LOP regressions reveal a nuanced picture of relative price variations across the sample of euro area countries and different product varieties. In particular, we find that the relative income levels and economic growth strongly affect relative prices in our sample. In addition, a number of significant price effects are linked to economies of scale, market demand and consumer preferences. At the same time, our empirical LOP regressions suggest a strong interdependence of relative prices across both the geographical and product-variety data dimensions, which go beyond the included set of explanatory variables and warrant further empirical investigation.

There are several issues which go beyond the scope of the present paper, but need to be addressed in future research. They include structural explanations of the observed price disparities, possibly utilising the more homogeneous product varieties available in the Nielsen disaggregated price dataset, such as brand and SKU-level relative prices. Additional breakdowns across the available data dimensions, e.g. intra-country regions, store types, etc., may also be considered. Finally, the impact of additional explanatory factors, such as country-level net exports, competition indicators and product-level brand concentrations, need to be taken into account in order to further advance our understanding of the observed price disparities in the euro area and give appropriate and informed policy advice.

References

ALLINGTON, N. F. B., P. A. KATTUMAN, AND F. A. WALDEMANN (2005): One market, one money, one price? *International Journal of Central Banking*, 1 (3), pp. 73–115.

BEHRENS, K., C. ERTUR, AND W. KOCH (2012): "Dual" gravity: using spatial econometrics to control for multilateral resistance. *Journal of Applied Econometrics*, 27 (5), pp. 773–794.

CHEUNG, Y.-W., AND K. S. LAI (2006): A reappraisal of the border effect on the relative price volatility. *International Economic Journal*, 20 (4), pp. 495–513.

CHEVALIER, J. A., A. K. KASHYAP, AND P. E. ROSSI (2003): Why dont't prices rise during periods of peak demand? Evidence from scanner data. *American Economic Review*, 93 (1), pp. 15–37.

CLIFF, A. D., AND J. K. ORD (1981): *Spatial Processes, Models and Applications*. Pion, London.

CONLEY, T. G., AND G. TOPA (2002): Socio-economic distance and spatial patterns in unemployment. *Journal of Applied Econometrics*, 17 (4), pp. 303–327.

CRUCINI, M. J., AND M. SHINTANI (2008): Persistence in law of one price deviations: Evidence from micro-data. *Journal of Monetary Economics*, 55, pp. 629–644.

CRUCINI, M. J., C. I. TELMER, AND M. ZACHARIADIS (2005): Understanding European real exchange rates. *American Economic Review*, 95 (3), pp. 724–738.

CRUCINI, M. J., AND H. YILMAZKUDAY (2013): Understanding long-run price dispersion. *NBER Working Paper*, 18811, National Bureau of Economic Research.

DOORNIK, J. A. (2007): *Object-Oriented Matrix Programming Using Ox*. 3rd edn. Timerlake Consultants Press.

ECB (2011): Structural features of distributive trades and their impact on prices in the euro area. *Structural Issues Report September 2011*. European Central Bank.

ENGEL, C., AND J. H. ROGERS (1996): How wide is the border? *American Economic Review*, 86 (5), pp. 1112–1125.

ENGEL, C., AND J. H. ROGERS (2004): European product market integration after the euro. *Economic Policy*, pp. 347–384.

FABER, R. P., AND A. C. J. STOKMAN (2009): A short history of price level convergence in Europe. *Journal of Money, Credit and Banking*, 41 (2–3), pp. 461–477.

FISCHER, C. (2012): Price convergence in the EMU? Evidence from micro data. *European Economic Review*, 56, pp. 757–776.

GOLDBERG, P. K., AND F. VERBOVEN (2005): Market integration and convergence to the Law of One Price: evidence from the European car market. *Journal of International Economics*, 65, pp. 49–73.

IMBS, J. M., H. MUMTAZ, M. O. RAVN, AND H. REY (2010): One TV, one price? *Scandinavian Journal of Economics*, 112 (4), pp. 753–781.

KELLER, W., AND C. H. SHIUE (2007): The origin of spatial interaction. *Journal of Econometrics*, 140, pp. 304–332.

KURKOWIAK, B. (2012): Major dispersion in consumer prices across Europe: Comparative price levels in 37 European countries for 2011. *Statistics in Focus*, Eurostat.

LEE, I. (2010): Geographical price dispersion in retail markets: Evidence from micro-data. *Journal of Macroeconomics*, 32, pp. 1169–1177.

MACDONALD, J. M. (2000): Demand, information, and competition: why do food prices fall at seasonal demand peaks? *Journal of Industrial Economics*, XLVIII (1), pp. 27–45.

MEYLER, A. (2013): *An overview of the Nielsen disaggregated price dataset provided to the Eurosystem.* Internal Document, European Central Bank.

OLIVEIRA, V. D., AND J. J. SONG (2008): Bayesian analysis of simultaneous autoregressive models. *Sankhya: The Indian Journal of Statistics*, 70-B (2), pp. 323–350.

ORD, K. (1975): Estimation methods for models of spatial interaction. *Journal of the American Statistical Association*, 70 (349), pp. 120–126.

PARSLEY, D. C., AND S.-J. WEI (2001): Explaining the border effect: the role of exchange rate variability, shipping costs, and geography. *Journal of International Economics*, 55, pp. 87–105.

ROBERT, C. P., AND G. CASELLA (2004): *Monte Carlo Statistical Methods*. 2nd edn, Springer.

ROGERS, J. H. (2007): Monetary union, price level convergence, and inflation: How close is Europe to the USA? *Journal of Monetary Economics*, 54, pp. 785–796.

ROGOFF, K. (1996): The purchasing power parity puzzle. *Journal of Economic Literature*, 34 (2), pp. 647–668.

WHITTLE, P. (1954): On stationary processes in the plane. *Biometrika*, 41 (3/4), pp. 434–449.

WOLSZCZAK-DERLACZ, J. (2008): Price convergence in the EU — an aggregate and disaggregate approach. *International Economics and Economic Policy*, 5, pp. 25–47.

Table A1: Sample coverage table by country and product category

Product	*COICOP*	*AT*	*BE*	*DE*	*EE*	*ES*	*FR*	*GR*	*IE*	*IT*	*NL*	*PT*	*SI*	*SK*	*Total*
All purpose cleaners	0561	•	•	•	•	•		•			•	•	•	•	10
Auto dishwashing detergent	0561	•	•	•	•	•	•		•		•	•	•	•	11
Baby food	0119	•	•	•	•	•	•	•	•	•	•	•	•	•	13
Beer	0213	•	•	•	•	•	•	•	•	•	•	•	•	•	13
Bouillon	0119				•			•	•	•			•		5
Butter	0115	•	•	•	•	•	•	•		•	•	•	•	•	12
Carbonated soft drinks	0122	•	•	•	•	•	•	•	•	•	•	•	•	•	13
Cat food	0934	•	•	•	•	•	•	•	•		•	•	•	•	12
Cereal	0111	•	•	•	•	•	•	•	•	•	•	•	•	•	13
Chewing gum	0118				•	•	•	•	•	•	•	•		•	9
Chocolate	0118		•		•	•	•	•	•	•	•	•	•	•	11
Cigarettes	0220	•	•		•	•		•		•	•	•	•	•	10
Condoms	0612	•	•	•	•	•	•	•		•	•	•		•	11
Deodorant	1213	•	•	•	•	•	•	•	•	•		•	•	•	12
Diapers	1213	•	•	•	•	•	•	•		•	•	•	•	•	12
Dog food	0934	•	•	•	•	•	•	•	•	•	•	•	•	•	13
Dry pasta spaghetti	0111	•	•	•	•	•	•	•	•	•	•	•	•	•	13
Fabric softener	0561	•	•	•	•	•	•	•	•	•	•	•	•	•	13
Frozen fish	0113	•	•	•		•	•		•	•	•	•			9
Frozen peas	0117	•	•	•			•		•	•	•	•	•		9
Ground coffee	0121	•	•	•	•	•	•	•	•	•	•	•	•	•	13
Ice cream	0118	•	•	•	•	•	•	•	•	•	•	•	•	•	13
Instant coffee	0121	•	•	•	•	•	•	•	•	•	•	•	•	•	13
Juice 100%	0122	•	•	•	•	•	•	•	•	•	•	•	•	•	13
Laundry detergent	0561		•	•	•	•			•	•		•	•	•	9
Margarine	0115	•	•	•	•	•	•	•		•	•	•	•	•	12
Olive oil	0115	•	•	•	•	•	•	•	•	•	•	•	•		12
Panty liners	1213	•	•	•	•	•		•	•		•	•	•	•	11
Paper towels	0561	•	•	•	•	•	•	•	•	•	•	•	•	•	13
Refrigerated milk	0114	•	•	•	•	•		•	•	•	•	•	•	•	12
Rice	0111	•	•	•		•	•	•	•	•	•	•	•	•	12
Shampoo	1213	•	•	•	•	•	•	•	•	•	•		•	•	12
Shave preps	1213			•	•				•	•	•		•	•	7
Sparkling water	0122	•	•	•	•	•	•	•	•	•	•	•	•	•	13
Still water	0122	•	•	•	•	•	•	•	•	•	•	•	•	•	13
Strawberry jam	0118	•	•	•		•	•	•	•	•	•		•		10
Sugar	0118	•	•	•	•	•	•		•	•	•			•	10
Tinned peas	0117	•	•	•		•	•		•	•	•		•		9
Tinned tuna	0113	•	•	•		•	•	•	•	•	•	•	•	•	12
Toilet tissue	1213	•	•	•	•	•	•	•	•	•	•	•	•	•	13
Toothpaste	1213	•	•	•	•	•	•	•		•	•	•	•	•	12
UHT milk	0114	•	•	•	•	•	•	•		•	•	•	•	•	12
Vodka	0211	•	•	•	•	•	•	•	•	•	•	•	•	•	13
Wet soups	0119	•		•	•	•	•		•	•	•	•	•	•	11
Whiskey	0211	•	•	•	•	•	•	•	•	•	•	•	•	•	13
Total		40	41	41	39	42	38	37	37	41	42	39	41	39	517

Notes: Bullets signify presence of a specific country-product pair in the sample used for empirical LOP analyses in Section 4. The four-digit COICOP codes are used to compute the spatial weights as detailed in Section 3.

Working Papers of Eesti Pank 2014

No 1
Jaanika Meriküll, Tairi Rõõm. Are Foreign-Owned Firms Different? Comparison of Employment Volatility and Elasticity of Labour Demand

No 2
Merike Kukk. Distinguishing the Components of Household Financial Wealth: the Impact of Liabilities on Assets in Euro Area Countries

No 3
Juan Carlos Cuestas, Luis A. Gil-Alana, Paulo José Regis. On the Changes in the Sustainability of European External Debt: What Have We Learned

No 4
Juan Carlos Cuestas, Karsten Staehr. The Great (De)leveraging in the GIIPS Countries. Domestic Credit and Net Foreign Liabilities 1998–2013

No 5
Fabio Filipozzi, Kersti Harkmann. Currency Hedge – Walking on the Edge?

No 6
Tairi Rõõm, Katri Urke. The Euro Changeover in Estonia: Implications for Inflation

No 7
Jaanika Meriküll, Tairi Rõõm. One Currency, One Price? Euro Changeover-Related Inflation in Estonia

No 8
Liina Malk. Determinants of Reservation Wages: Empirical Evidence for Estonia

No 9
Jaanika Meriküll, Pille Mõtsmees. Do you get what you ask? The gender gap in desired and realised wages