Price Dispersion across U.S. Districts of Entry

Hakan Yilmazkuday*

March 3, 2014

Abstract

This paper investigates the price dispersion of U.S. imports at the good-category level across U.S. districts of entry. Although there is a large heterogeneity across goods, on average, the implied markups of a simple model explain about 31% of the price dispersion, while the implied marginal costs of production explain about 69%; the effects of trade costs, for which we have actual data, are almost none. The results are robust to the consideration of possible endogeneity problems, multiplicative versus additive trade costs, and measurement errors in prices.

JEL Classification: F12, F13, F14

Key Words: Price Dispersion; Variable Markups; Trade Costs.

*Department of Economics, Florida International University, Miami, FL 33199, USA; e-mail: hyilmazk@fiu.edu

1. Introduction

In international economics, typical components of prices are marginal costs of production (excluding trade costs), markups, and trade costs. Therefore, decomposing prices into their components is the key in understanding the price dispersion across locations and thus the deviations from the Law of One Price (LOP).[1] However, this is not an easy task, since data for such components are mostly not available; this has led researchers rather focus on the implications of economic models for estimating these components. For instance, in an influential study, Engel and Rogers (1996) have estimated the effects of trade barriers/costs on the price dispersion using variables such as distance and/or an international border and shown that such variables are highly significant in explaining the price dispersion across locations at the good-category level.

Using actual data on trade costs (i.e., cost, insurance, freight, and duties/tariffs), together with a simple model based on variable markups, this paper shows that marginal costs of production and markups are the main sources of variation in prices; the effects of trade costs are almost none. In particular, marginal costs of production explain about 69% and markups about 31% of the price dispersion of U.S. imports across U.S. districts of entry (i.e., the district in which merchandise clears customs) on average. The results are robust to the consideration of possible endogeneity problems, multiplicative versus additive trade costs (due to having actual data on trade costs), and measurement errors in prices. Therefore, studies that proxy the actual data on trade costs by distance/border effects may well be capturing any unmodeled part of preferences in utility functions, such as dyadic demand shifters, rather than actual trade costs. If preferences are the main source of trade barriers, policies aimed to increase welfare-improving trade would require more than just reducing duties/tariffs.

2. A Simple Model

We have a demand-side model where we distinguish between the utilities of importers located at different U.S. districts of entry. In particular, a typical importer located at district d of entry in the U.S. has the following utility U_d^g maximization out of consuming varieties of good g coming from different source countries, each denoted by s :

$$\max U_d^g = \sum_s \kappa_{ds}^g \left(1 - e^{-\alpha^g q_{ds}^g}\right) \tag{2.1}$$

where q_{ds}^g is the quantity traded, α^g is a good-specific parameter (to be connected to markups, below), and κ_{ds}^g represents preferences (i.e., demand shifters).[2] Maximization of this utility function

[1] Isard (1977) is one of the earliest studies showing such deviations from LOP.

[2] Behrens and Murata (2007) have shown that the type of this utility function, namely constant absolute risk aversion, implies variable markups. In the absence of actual data on trade costs, Yilmazkuday (2013) has used a similar utility function to investigate the deviations from LOP by including more structure on preferences; this paper deviates from Yilmazkuday (2013) by considering actual data on trade costs and source-specific marginal costs of production for the identification of markups versus marginal costs of production.

results in the following demand function:

$$q_{ds}^{g} = \frac{E_{d}^{g} - \frac{1}{\alpha^{g}} \sum\limits_{s'} \ln\left(\frac{p_{ds}^{g}\kappa_{ds'}^{g}}{p_{ds'}^{g}\kappa_{ds}^{g}}\right) p_{ds'}^{g}}{\sum\limits_{s'} p_{ds'}^{g}} \tag{2.2}$$

where p_{ds}^{g} represents the price per unit of q_{ds}^{g}. Taking the demand function into account, source country s follows a pricing-to-market strategy by maximizing its profits given by:

$$\pi_{ds}^{g} = q_{ds}^{g}\left(p_{ds}^{g} - c_{ds}^{g}\right)$$

where c_{ds}^{g} represents marginal costs of exporting given by:

$$c_{ds}^{g} = w_{s}^{g}\tau_{ds}^{g}$$

where w_{s}^{g} represents source-specific marginal costs of production, and τ_{ds}^{g} represents trade costs. The profit maximization results in the following price expression:

$$p_{ds}^{g} = w_{s}^{g}\mu_{ds}^{g}\tau_{ds}^{g} \tag{2.3}$$

where $\mu_{ds}^{g} = \left(1 - \alpha^{g} q_{ds}^{g}\right)^{-1}$ represents gross variable markups (that change with quantity traded).

3. Data

The U.S. imports data are from the US. International Trade Commission (http://dataweb.usitc.gov/) covering imports from 232 source countries for 443 good categories[3] at the SITC 4-digit level measured at 41 U.S. districts of entry (i.e., the districts in which merchandise clears customs)[4] for the most recent year of 2012. The data set includes (i) customs value (quantity times price charged by exporters) measured at the dock of the source country, (ii) quantity traded, (iii) general import charges in values (i.e., the aggregate cost of all freight, insurance, and other charges incurred, excluding U.S. import duties), and (iv) calculated duties in values (i.e., the estimated import duties collected based on the applicable rates of duty as shown in the Harmonized Tariff Schedule).

Overall trade costs in multiplicative terms are calculated by dividing the sum of general import charges and calculated duties by the customs value; this calculation methodology effectively

[3] These are the good categories for which we have at least 120 observations for a robust estimation at the good level. The complete list of good categories is available upon request.

[4] The list of districts of entry is as follows: Anchorage, AK; Baltimore, MD; Boston, MA; Buffalo, NY; Charleston, SC; Charlotte, NC; Chicago, IL; Cleveland, OH; Port of Portland, OR Fort Worth, TX; Detroit, MI; Duluth, MN; El Paso, TX; Great Falls, MT; Honolulu, HI; Houston, TX; Laredo, TX; Los Angeles, CA; Miami, FL; Milwaukee, WI; Minneapolis, MN; Mobile, AL; New Orleans, LA; New York, NY; Nogales, AZ; Norfolk, VA; Ogdensburg, NY; Pembina, ND; Philadelphia, PA; Port Arthur, TX; Portland, ME; Providence, RI; San Diego, CA; San Francisco, CA; San Juan, Puerto Rico Savannah, GA; Seattle, WA; St. Albans, VT; St. Louis, MO; Tampa, FL; Washington, DC.

converts any type of trade costs (either additive or multiplicative) into multiplicative terms. For robustness, overall trade costs are decomposed into duties/tariffs and freight-related costs; duties/tariffs are calculated by dividing the calculated duties by the customs value, while freight-related costs are calculated by dividing the general import charges (excluding duties/tariffs) by the customs value.

We calculate unit destination prices by dividing the sum of customs value, general import charges and calculated duties by the quantity traded. Two typical examples are the prices of a kilogram of coffee (with an SITC code 711) exported by Argentina and Brazil to the U.S. where Chicago, IL and Miami, FL are the U.S. districts of entry, respectively; in this particular example, we are interested in understanding the sources of price dispersion between Chicago, IL and Miami, FL regarding coffee prices. Since these unit prices are subject to measurement errors, for robustness, while decomposing the destination prices into their components below, we will consider only the fitted value of prices obtained by our empirical methodology.

4. Empirical Methodology

We are interested in decomposing the destination prices p_{ds}^g into source-specific marginal costs of production w_s^g, markups μ_{ds}^g, and trade costs τ_{ds}^g. Accordingly, we consider the stochastic version of Equation 2.2 to estimate the key parameter α^g at the good level (that we need to obtain implied markups):

$$\underbrace{q_{ds}^g}_{\text{Quantity Traded}} = \underbrace{\left(\frac{E_d^g + \frac{1}{\alpha^g} \sum_{s'} \ln\left(p_{ds'}^g\right) p_{ds'}^g}{\sum_{s'} p_{ds'}^g} \right)}_{\text{Destination-and-Good Fixed Effects}} - \underbrace{\frac{\ln p_{ds}^g}{\alpha^g}}_{\text{Prices}} + \underbrace{\frac{\ln \kappa_{ds}^g}{\alpha^g}}_{\text{Residuals}}$$

where we employ preferences as residuals (as in Yilmazkuday, 2012). However, since prices p_{ds}^g also depend on quantity traded q_{ds}^g according to Equation 2.3 (due to markups), there is a potential endogeneity/simultaneity problem. Accordingly, we use two stage least squares (TSLS) as an estimation methodology, and estimate the reduced form of log destination prices in the first stage of TSLS estimation approximated by the following stochastic version of Equation 2.3:

$$\underbrace{\ln p_{ds}^g - \frac{\ln \tau_{ds}^g}{2}}_{\text{Data on Prices and Trade Costs}} \approx \underbrace{\frac{\ln w_s^g}{2}}_{\text{Source-and-Good Fixed Effects}} + \underbrace{\left(\frac{\alpha^g E_d^g + \sum_{s'} \ln\left(p_{ds'}^g\right) p_{ds'}^g}{2 \sum_{s'} p_{ds'}^g} \right)}_{\text{Destination-and-Good Fixed Effects}} + \underbrace{\frac{\ln \kappa_{ds}^g}{2}}_{\text{Residuals}} \quad (4.1)$$

where we have used $\ln \mu_{ds}^g \approx \alpha^g q_{ds}^g$ (for simplicity) in order to obtain a linear relationship between q_{ds}^g and $\ln p_{ds}^g$.[5] It is important to emphasize that preferences κ_{ds}^g's enter into the price expression as residuals; if κ_{ds}^g's depend on any source- or destination-specific measures, such as quality, these

[5] The fixed effects on the right hand side correspond to the instruments of TSLS.

would be captured by source and destination fixed effects as well. However, any unmodeled dyadic demand shifter (at the good level), including any distance/border effects that are independent of measured trade costs, would be reflected as residuals, because we already have data on trade costs. Afterwards, we calculate the fitted values for log destination prices according to:

$$\underbrace{\widehat{\ln p_{ds}^{g}}}_{\text{Fitted Prices}} = \underbrace{\left(\frac{\widehat{\ln w_{s}^{g}}}{2}\right)}_{\text{Fitted Source-and-Good Fixed Effects}} + \underbrace{\left(\frac{\alpha^{g} E_{d}^{g} + \widehat{\sum_{s'} \ln}\left(p_{ds'}^{g}\right) p_{ds'}^{g}}{2 \sum_{s'} p_{ds'}^{g}}\right)}_{\text{Fitted Destination-and-Good Fixed Effects}} + \underbrace{\frac{\ln \tau_{ds}^{g}}{2}}_{\text{Data on Trade Costs}}$$

to be further used in the second stage of TSLS estimation to estimate α^{g}'s.

Once α^{g}'s are estimated at the good level, which we achieve by good-level TSLS regressions, we further use them, together with the fitted values of quantity traded $\widehat{q_{ds}^{g}}$, to obtain markups according to the approximation given by $\mu_{ds}^{g} \approx \exp\left(\widehat{\alpha^{g}}\widehat{q_{ds}^{g}}\right)$. We obtain estimates for source-specific marginal costs of production by the fitted value of source-and-good fixed effects in the first stage of TSLS, above (i.e., by using $(\ln \widehat{w_{s}^{g}}/2)$). This identification strategy results in the following decomposition of the fitted values for log destination prices:

$$\underbrace{\widehat{\ln p_{ds}^{g}}}_{\text{Log Fitted Prices}} = \underbrace{\widehat{\ln w_{s}^{g}}}_{\text{Log Estimated Marginal Costs}} + \underbrace{\widehat{\alpha^{g}}\widehat{q_{ds}^{g}}}_{\text{Log Estimated Markups}} + \underbrace{\ln \tau_{ds}^{g}}_{\text{Log Data on Trade Costs}}$$

which effectively eliminates any measurement errors in the price data due to ignoring residuals (that represent preferences if we literally consider the implications of the model). Once we have this expression, we further have a variance decomposition analysis (in order to understand the sources of price dispersion) according to two different methodologies, for robustness. For the first variance decomposition methodology, we use the following expression:

$$\underbrace{\frac{var\left(\widehat{\ln p_{ds}^{g}}\right)}{\left(avg\left(\widehat{\ln p_{ds}^{g}}\right)\right)^{2}}}_{\text{Price Dispersion}} = \underbrace{\frac{cov\left(\widehat{\ln w_{s}^{g}}, \widehat{\ln p_{ds}^{g}}\right)}{\left(avg\left(\widehat{\ln p_{ds}^{g}}\right)\right)^{2}}}_{\text{Contribution of Marginal Costs}} + \underbrace{\frac{cov\left(\widehat{\alpha^{g}}\widehat{q_{ds}^{g}}, \widehat{\ln p_{ds}^{g}}\right)}{\left(avg\left(\widehat{\ln p_{ds}^{g}}\right)\right)^{2}}}_{\text{Contribution of Markups}} + \underbrace{\frac{cov\left(\ln \tau_{ds}^{g}, \widehat{\ln p_{ds}^{g}}\right)}{\left(avg\left(\widehat{\ln p_{ds}^{g}}\right)\right)^{2}}}_{\text{Contribution of Trade Costs}}$$

where avg, var and cov are operators of average, variance and covariance, respectively; this expression holds with equality due to the properties of covariance operator. Dividing both sides by the square of the corresponding average price is just to control for scale effects at the good level so that the obtained numbers are comparable across goods. Since this first methodology has an implicit assumption that the right hand side variables are independent from each other, we also consider a second variance decomposition methodology according to the following approximation:

$$\underbrace{\frac{var\left(\widehat{\ln p_{ds}^{g}}\right)}{\left(avg\left(\widehat{\ln p_{ds}^{g}}\right)\right)^{2}}}_{\text{Price Dispersion}} \approx \underbrace{\frac{var\left(\widehat{\ln w_{s}^{g}}\right)}{\left(avg\left(\widehat{\ln p_{ds}^{g}}\right)\right)^{2}}}_{\text{Contribution of Marginal Costs}} + \underbrace{\frac{var\left(\widehat{\alpha^{g}}\widehat{q_{ds}^{g}}\right)}{\left(avg\left(\widehat{\ln p_{ds}^{g}}\right)\right)^{2}}}_{\text{Contribution of Markups}} + \underbrace{\frac{var\left(\ln \tau_{ds}^{g}\right)}{\left(avg\left(\widehat{\ln p_{ds}^{g}}\right)\right)^{2}}}_{\text{Contribution of Trade Costs}}$$

where we have ignored the covariance terms to focus on the pure effects of right hand side variables on price dispersion.[6]

5. Empirical Results

The summary of implications for price dispersion are given in Table 1, where, on average (across goods), the price dispersion of 1.89 is due to markups by 27% (35%) and due to source-specific marginal costs of production by 73% (65%) when the first (second) variance decomposition methodology is used; the effects of trade costs, which we decompose into the effects due to freight-related costs and duties/tariffs, are virtually none on average.[7] The decomposition of price dispersion for all goods is given in Figure 1 in levels and in Figure 2 in percentage terms when the first methodology is used for variance decomposition.[8] As is evident, the contribution of markups in the price dispersion are up to 85%, while the contribution of marginal costs are up to 100% across goods; the contribution of trade costs are again almost none in most cases (although there are few exceptions).

When we consider the categorization of these goods according to Rauch (1999), we see that the contribution of markups are higher for homogenous products, while it is lower for differentiated products, independent of the variance decomposition methodology used. Therefore, as implied by our model (i.e., markups are positively related to quantities traded), exporters charge higher markups for homogenous products as they sell more to the U.S., while differentiated goods have already-distinguished marginal costs of production measured at the source.

In order to provide the reader a better idea about the results, we also consider a selected sample of goods (ranked with respect to the contribution of markups in price dispersion) in Table 1 where there is evidence for heterogeneity across goods regarding the percentage contribution of markups versus marginal costs of production; e.g., for *coffee* (with SITC code of 711), markups contribute about 71% (69%), while for *piano* (with SITC code of 8981), markups contribute only about 4% (18%) when the first (second) variance decomposition methodology is used.[9] One of the highest contribution of trade costs is for *plate of iron* (with SITC code of 6741) where the contribution of freight-related costs is more than 10%, mostly due to the heavy structure of the product.

[6] In this second methodology, we decompose the summation of the right hand side of the expression into its components.

[7] All estimated α^g's are significant at the 10% level. The average R-squared values are about 0.54 and 0.16 for the first and the second stage of TSLS, respectively, where 0.54 is an indicator of strong instruments. It is important to emphasize that low R-squared values of the second stage are not comparable to the high values in the gravity literature that have been obtained by log-linear equations where the left hand side is in logs (rather than levels as in this paper). All of the good-level results are available upon request.

[8] The figures are virtually the same when the second methodology is used for variance decomposition; such figures are available upon request.

[9] The results for apparel are in line with Simonovska (2010) who shows that roughly a third of the observed variation in prices of apparel are due to variable markups.

6. Conclusion

We decomposed the price dispersion of U.S. imports at the good-category level across U.S. districts of entry. On average (across goods and decomposition methodologies), while marginal costs contribute about 69% to the price dispersion, markups contribute about 31%. The surprising part is that trade costs, for which we have actual data, have almost no effects on the price dispersion on average; this is against many studies in the literature which mostly rely on unobserved measures of trade costs (e.g., distance-related effects or border effects as in Engel and Rogers, 1996). One possible explanation is that, in such studies, distance-related or border effects may be capturing the effects due to preferences when one literally considers the implications of economic models (e.g., dyadic demand shifters or time-to-trade as a part of κ_{ds}^{g}'s in this paper) rather than freight-related costs or duties/tariffs; if preferences are the main source of trade barriers, policies aimed to increase welfare-improving trade would require more than just reducing duties/tariffs. Understanding such linkages requires a richer model with more structure on preferences (i.e., κ_{ds}^{g}'s in this paper) together with a richer data set (including actual data on cost, insurance, freight, and duties/tariffs, which are not available, to our knowledge beyond the data set of this paper on the U.S. trade patterns) that covers more than one (preferably many) destination countries so that possible dyadic demand shifters can be identified.

References

[1] Behrens, K. and Murata, Y. (2007), "General equilibrium models of monopolistic competition: A new approach" Journal of Economic Theory. 13: 776 – 787.

[2] Engel, C. and Rogers, J., (1996) "How Wide is the Border?" American Economic Review, 86(5): 1112-1125.

[3] Isard, P. (1977) "How Far Can We Push the Law of One Price?" American Economic Review, 67(5): 942-48.

[4] Rauch, J.E. (1999), "Networks versus markets in international trade" Journal of International Economics 48: 7–35.

[5] Simonovska, I. (2010): "Income Differences and Prices of Tradables," NBER Working Paper No. 16233.

[6] Yilmazkuday, H. (2012) "Understanding Interstate Trade Patterns" Journal of International Economics. 86:158–166.

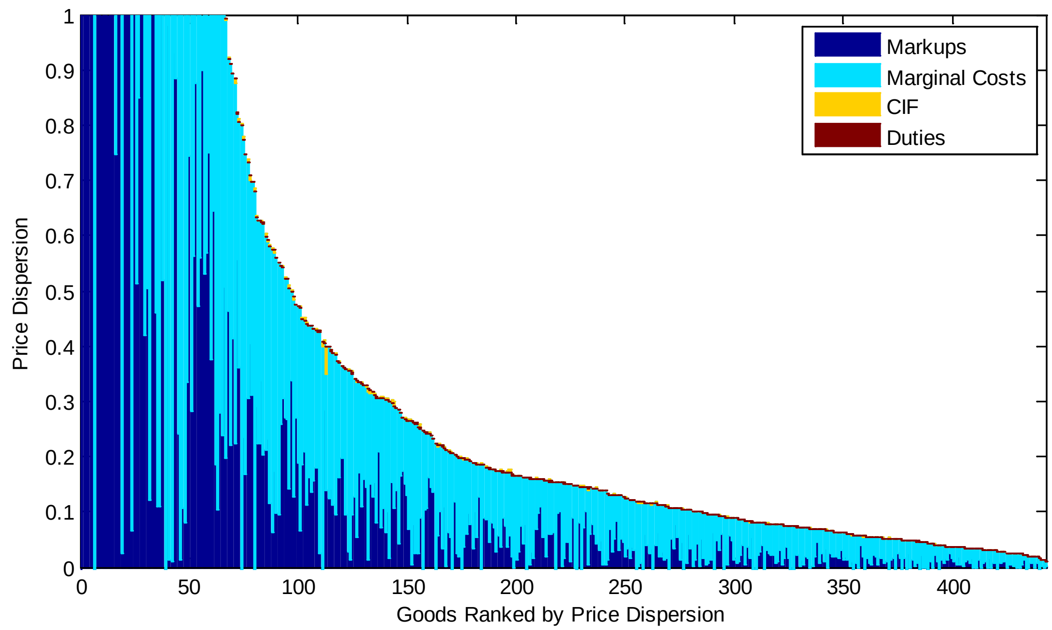
[7] Yilmazkuday, H. (2013) "Constant versus Variable Markups: Implications for the Law of One Price" Available at SSRN: http://ssrn.com/abstract=2144037 or http://dx.doi.org/10.2139/ssrn.2144037.

Table 1 - Summary of Price Dispersion

Goods	Price Dispersion	% Contribution of: Variable Markups	Marginal Costs of Production	Trade Costs (Freight)	Trade Costs (Duties)
Average across All Goods	1.89	27.02 (35.03)	73.16 (64.60)	-0.21 (0.34)	0.02 (0.03)
Average across Homogenous Goods	0.79	37.88 (41.51)	62.43 (58.12)	-0.30 (0.31)	-0.01 (0.06)
Average across Reference-Priced Goods	2.91	34.44 (40.75)	65.56 (58.74)	-0.03 (0.49)	0.04 (0.01)
Average across Differentiated Goods	1.88	23.73 (32.75)	76.52 (66.90)	-0.29 (0.32)	0.02 (0.04)
Selected Sample of Goods					
Coffee [711]	0.11	71.15 (69.03)	28.36 (30.59)	0.48 (0.38)	0.00 (0.00)
Natural Honey [616]	0.31	68.42 (68.19)	31.61 (31.75)	0.03 (0.07)	-0.06 (0.00)
Caviar [371]	0.82	68.08 (68.59)	32.22 (31.26)	-0.12 (0.09)	-0.18 (0.06)
Wine [1121]	2.13	65.99 (68.08)	34.35 (31.85)	-0.26 (0.06)	-0.09 (0.01)
Carpet [6592]	0.03	48.34 (48.17)	51.56 (50.82)	0.09 (1.00)	0.01 (0.01)
Women's Suits [8432]	0.17	43.92 (44.95)	56.11 (54.37)	-0.58 (0.48)	0.52 (0.21)
Men's Coats [8421]	0.05	40.81 (39.19)	59.33 (60.04)	-0.71 (0.62)	0.55 (0.15)
Trousers [8423]	0.08	38.41 (36.94)	62.49 (62.29)	-1.25 (0.60)	0.24 (0.17)
Men's Suits [8422]	0.29	36.63 (38.07)	65.49 (61.24)	-2.16 (0.57)	-0.22 (0.11)
Motorcycle [7851]	0.07	36.00 (38.42)	64.43 (61.48)	-0.51 (0.10)	0.08 (0.00)
Plate of Iron [6741]	0.40	34.74 (37.74)	52.71 (54.99)	12.55 (10.26)	0.00 (0.00)
Frozen Fish [342]	0.30	33.07 (36.55)	67.90 (63.30)	-0.97 (0.15)	-0.01 (0.00)
Beer [1123]	1.08	25.78 (33.03)	77.95 (65.55)	-3.73 (1.42)	0.00 (0.00)
Tea [741]	0.04	17.42 (30.33)	82.33 (69.43)	0.25 (0.25)	0.00 (0.01)
Refined Sugar [612]	3.46	14.86 (28.95)	85.00 (70.12)	-0.36 (0.15)	0.53 (0.78)
Refrigerator [7414]	0.05	5.61 (26.82)	94.61 (73.11)	-0.21 (0.08)	0.00 (0.00)
Piano [8981]	0.07	4.11 (17.92)	95.96 (82.02)	-0.24 (0.05)	0.15 (0.01)

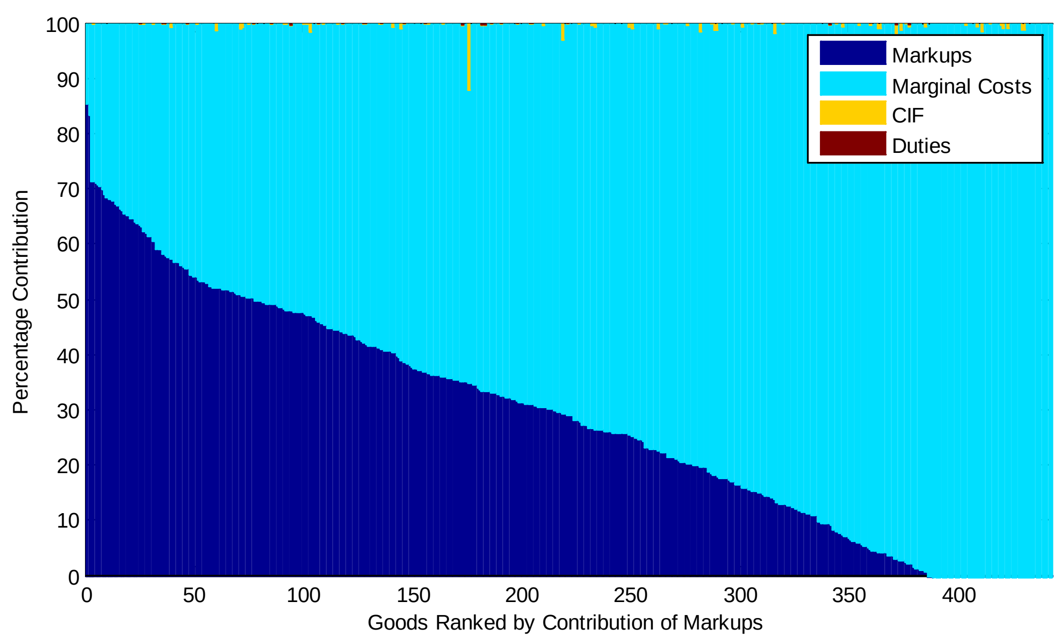
Notes: The percentage contribution values represent the variance decomposition of prices calculated according to the first methodology described in the text, while the values in parenthesis represent the variance decomposition according to the second methodology. Goods have been categorized according to Rauch (1999). For selected goods, which have been ranked according to the percentage contribution of markups, the corresponding SITC codes are given in the brackets.

Figure 1 - Decomposition of Price Dispersion in Levels



Notes: The decomposition has been achieved by using the first methodology described in the text. We limited the maximum of the vertical axis to one for presentational purposes.

Figure 2 - Decomposition of Price Dispersion in Percentage Terms



Notes: The decomposition has been achieved by using the first methodology described in the text.