LIS
Working Page Series

No. 752

How do regional price levels affect income inequality? Household-level evidence from 21 countries

Petr Janský and Marek Šedivý

October 2018

Luxembourg Income Study (LIS), asbl

How do regional price levels affect income inequality? Household-level evidence from 21 countries

Petr Janský and Marek Šedivý[1]

[1] Petr Janský (jansky.peta@gmail.com), Marek Šedivý (73090835@fsv.cuni.cz), Institute of Economic Studies, Faculty of Social Sciences, Charles University (Opletalova 21, 170 00 Prague, Czech Republic). We are grateful to prof. Zhong Zhao for his advice about the Chinese labour market data, prof. Steff van Buuren for the insight he provided into the methodology of multiple imputation and Bettina Aten for the valuable comments she provided on an early draft of this paper. We also want to express gratitude to Bartlomiej Rokicki and Luigi Biggeri for comments on the final version of this paper. Our thanks also go to the Luxembourg Income Study for providing access to their database and to all the national statistical offices that helped us to construct our data set, namely – Central Statistical Office (Ireland), Statistics Denmark, National Statistics Office of Georgia, Turkish Statistical Institute, Philippine Statistics Authority and General Statistics Office of Vietnam.

Abstract

Regional differences in prices levels are substantial in many countries, but little is known about how important they are for income inequality and relative poverty. To bridge this gap, we provide new evidence on the basis of the best available data and a novel two-step approach. First, we collect the largest cross-country dataset of regional price level estimates from 12 countries and use it to predict regional price levels in other countries. We then combine all these regional prices levels with household-level data from the Luxembourg Income Study, which gives us results for a final sample of 21 countries. We find that for some countries Gini coefficients and headcount poverty ratios are statistically significantly different when adjusted for regional price levels. For example, we show that adjusting for regional price levels would lower the Gini coefficients by 2% for Italy, 3% for Columbia and by 4% for Georgia, while it would increase the headcount poverty ratio by 6% for France and by 7% for Ireland. We conclude that regional price levels affect income inequality to a varying extent and should be taken into account by policy makers and in future research.

Keywords: income inequality; relative poverty; regional price levels; regional purchasing power parities; Luxembourg Income Study

JEL classification: D63; O18; R11; R58

1 Introduction

Price levels have a substantial impact on differences in standards of living among countries and regions and knowledge of them is therefore essential for all assessments of income distribution. The Purchasing Power Parities (PPPs) that are applied to convert monetary data into comparable units are usually available only at the national level, despite substantial evidence of intra-national disparities in prices. Regional price levels, or regional purchasing power parities, provide information about the relative price level in the considered regions to the national level. To the best of our knowledge estimates of these exist for just fifteen countries and only two statistical offices in the world - those of the United States and Turkey, regularly provide measures of regional price levels. However, failure to adjust for intra-country spatial price differences when constructing income-based economic indicators could lead to inaccurate inference about the distribution of income. One possible consequence of failure to adjust for regional price levels is, for example, a bias in estimates of income inequality since spatial price differences may lead to overvaluation of income in regions with high price levels and undervaluation of income in regions with low regional price levels.

The influence of regional price levels on income-based economic indicators has been assessed by a number of studies focused always focused on a single country. For example, Deaton (2003) provides an assessment for India and Brandt and Holz (2006) for China. More generally, Aten (2017) argues that we need accurate measures of regional price levels in order to evaluate within-country income inequality. This importance has accentuated by recent shifts in the geography of poverty. During the past two decades a shift of the global poor from low to middle income countries has been observed. Consequently the world of the “bottom billion” outlined by Collier (2007) has been replaced by the “new bottom billion” of Sumner (2010). About three quarters of the global poor live in middle income countries. Considering that, for example, both the huge and regionally diverse countries of India and China belong to this group, failure to account for differences in regional price levels could have severe consequences.

So far, to the best of our knowledge, no systematic assessment of the impact of intra-country spatial price differences on indicators of income inequality and relative poverty has yet been conducted for a group of countries. We aim to bridge this gap by estimating regional price levels for a sample of 21 high and middle income countries and by estimating the impact of spatial price disparities on indicators of income inequality and relative poverty. For that purpose we use existing estimates of regional price levels in combination with important assumptions to estimate regional price levels for other countries. We combine the resulting regional price estimates with household-level income data provided by the Luxembourg Income Study (LIS) database in order to construct national level Gini coefficients and poverty headcount ratios, both unadjusted and also adjusted for intra-country spatial price differences. We provide an assessment of the scale and statistical significance of the changes in the considered indicators induced by adjustment for regional price levels.

In our empirical results, we find that for some countries Gini coefficients and headcount poverty ratios are statistically significantly different when adjusted for regional price levels. For example, we show that adjusting for regional price levels would lower the Gini coefficients by 2% for Italy, 3% for Columbia and by 4% for Georgia, while it would increase the headcount poverty ratio by 6% for France and by 7% for Ireland. Our results are primarily of indicative value, since assumptions had to be made to obtain them, as we explain below. Still, we conclude that regional price levels affect income inequality to a varying extent and should be taken into account by policy makers and in future research. For example, we point out below some possible applications of the growing literature in the field of intra-country spatial price differences and its possible use in studies focused on groups of countries rather than a single country.

The remainder of this paper is organized as follows. Section 2 contains a brief literature review focused on the description of various methodologies used to construct regional price levels. Section 3 describes the data on regional price levels and household incomes. Section 4 describes the methodological approaches both to construct an econometric model that estimates regional price levels and to adjust income inequality indicators for regional price levels. Section 5 presents the results of these two steps: first, we estimate the econometric model, test its out-of-sample predictive powers and arrive at the regional price levels for 21 countries. These are then used in combination with household-level data to test the impact of regional price levels on the national level Gini coefficient and poverty headcount ratios based on different poverty lines. The final section concludes.

2 Related literature

Despite the significant attention devoted to the study of temporal differences in price levels the literature dedicated to the analysis of spatial price differences remains more constrained. Even though this paper looks at estimation of intra-country spatial price differences, literature studying the determination of price levels on the national level can provide valuable insight. Ahec-Šonje and Nestić (2002) study the determinants of PPPs based on the price levels provided by the European Comparison Program for 39 countries. Gelb and Diofasi (2015) carry out a similar exercise while benefiting from a wider data set of 168 national price levels constructed by the International Comparison Program (ICP) of the World Bank. Each of these studies proposes a set of control variables, such as income, institutional quality and the openness of labour markets. However, intra-country spatial price differences also exist. One possible theoretical explanation for this phenomenon is provided by Suedekum (2006), who adds a home-goods sector to the model outlined in the seminal work by Krugman (1991) and shows that a core-periphery structure with higher price levels in the core can appear. A variety of approaches have been adopted by researchers and statistical offices in order to estimate these differences.

The majority of estimates of regional price levels are based on data collected for the construction of temporal price indices. These are usually combined with a superlative price index such as the Geary-Khamis (GK) or Gini-Èltetö-Köves-Szulc (GEKS) formula, or a Country-Product-Dummy (CPD) methodology as proposed by Summers (1973). In some instances the two approaches are combined. A comprehensive overview of the CPD methodology is provided by Biggeri et al. (2017), while Deaton and Heston (2010) provide a description of the superlative price indexes. Aten and Menezes (2002) use data collected for the Consumer Price Index (CPI) to estimate spatial price indices for 11 Brazilian cities. Data collected for creation of CPI are also used in a joint project of the Bureau of Economic Analysis (BEA) and the Bureau of Labour Statistics (BLS), where they are combined with indicators of housing prices in order to produce Regional Purchasing Parities (RPP) for the states of the United States (US). Aten (2017) provides a detailed description of the applied methodology, which consists of applying the CPD methodology and the GK formula. The impacts of the RPP on the distribution of income are discussed for example by Aten et al. (2011) and Aten et al. (2012).

To the best of our knowledge, in addition to the US government the only other statistical office which regularly provides spatial price indices is Turkstat, which uses the data collected to construct a CPI and applies the GEKS formula. The Office for National Statistics (2011) in the United Kingdom also combined data collected for the construction of its CPI with a survey on regional prices and applied the GEKS formula to construct Relative Regional Consumer Price Levels (RCPLs) for the UKs NUTS 1 regions. The GEKS formula is also applied by Biggeri et al. (2017) to construct Chinese regional price levels. Kramulová et al. (2016) adopt an approach inspired by the methodology of the Organisation for Economic Co-operation and Development (OECD) to estimate regional price levels for the Czech Republic on the NUTS 3 level, based on data collected for the construction of the CPI. A similar methodology is also used by Rokicki (2015) who provides regional price levels for Poland. Statistics Bureau of the Japanese Ministry of Internal Affairs and Communications produced Regional Difference Index of Prices for the year 2007. The Asian Development Bank estimated regional price levels in the Philippines during a one-off exercise. Its estimation was based on data collected for the CPI and a combination of the CPD methodology and geometric Laspeyeres index. The resulting regional purchasing power parities for the years 2005 to 2010 are reported in Dikhanov et al. (2011). Biggeri et al. (2017) proposes a set of CPD models and estimates comparable price levels for 19 Italian regional capitals.

As an alternative approach to the estimation of regional price levels, Coondoo et al. (2004) propose to use the analysis of the Engel's curve and estimation of a Quadratic Almost Ideal Demand System (QUAIDS), for which consumer expenditure data are required. This methodology is applied for example by Coondoo et al. (2004), Majumder et al. (2011), Mishra and Ray (2014) or Majumder and Ray (2017) to estimate regional price levels in Vietnam, Indonesia, Australia and India. The regional price levels in India are estimated based on data provided by the Indian National Sample Survey (NSS). A similar data source is used by Deaton (2003) to construct Laspeyers, Paasche, Fischer Ideal and Törnqvist price indexes for the rural and urban parts of the 17 Indian regions. Brandt and Holz (2006) use data published in the Price Statistical Yearbook of the Chinese National Bureau of Statistics (NBS) to estimate price levels for the rural and urban parts of the 16 Chinese regions as well as one combined price level for each region. Li, Zhang, and Du (2005) also provide regional price levels for the regions of China constructed by the Feixuan formula. However these are based on a larger basket of goods. Radvansky and Fuchs (2012) provide regional price levels for Slovak NUTS 3 regions based on household budget survey.

Another approach is proposed by Roos (2006), who uses the data set provided by Ströhl (1994) to construct an OLS model with price levels as the controlled variable. By performing out-of-sample prediction and shifting the estimated indices by the application of regional CPI he estimates regional price levels for German NUTS 1 regions. A similar approach is adopted by Janský and Kolcunová (2017) for the NUTS 2 regions of EU countries. The authors construct an OLS model based on existing estimates of regional price levels and use it in a second step for one of the largest estimations of regional price levels yet conducted. However Blien et al. (2009) criticize this approach asserting that the construction of the econometric model based on estimated values for which standard errors are unknown leads to a bias in the estimated standard errors. As a solution, they propose the application of the multiple imputation framework and use it to test whether the agglomeration wage differential was not at least partly caused by unobserved differences in regional price levels.

A number of existing estimates of regional price levels were used to provide an assessment of their impacts on income-based economic indicators and their consequences for measurement of poverty. Pittau et al. (2011) find statistically significant distortions in Italian Gini, Theil and Decile ratios resulting from failure to adjust for differences in inter-regional price levels. Mogstad et al. (2007) provide an indicative assessment of the impact of region-specific poverty lines on both national level indicators as well as intra-country geography of poverty for Norway. They conclude that application of region-specific poverty lines affects geographical as well as demographical distribution of poverty. However this may be a consequence of the fact that their regional price levels are based solely on prices of household rents. Brandt and Holz (2006) show an illustrative example of 30% decrease of the 1990 Chinese Gini coefficient. Probably the most thorough assessment is provided by the literature based on the Regional purchasing power parities produced by the joint project of the BEA and BLS, e.g. Aten et al. (2012) or Aten et al. (2011). Bajgar and Janský (2014) study the effects of regional prices estimated by Čadil et al. (2014) on wages and pensions in Czech Republic. They find that after controlling for spatial price differences and composition of the labour force a worker in the capital city of Prague earns the same real average wage as workers in other regions despite her wage being 43% higher in nominal terms. Rokicki and Hewings (2016) study the impact of regional price levels on regional differences in income concluding that they should be of major interest in the framework of the EU Cohesion policy. Janský and Kolcunová (2017) study the implications of regional price levels on the assignment of funds to the EU regions in the framework of the EU cohesion policy. They find that nearly 7% of the regions are currently misclassified. Their work is perhaps the only one assessing the impacts of regional price levels on income-based economic indicators for a group of countries and therefore we aim to provide develop this literature by focusing on economic inequality and using microeconomic household-level data for the first time.

3 Data

We first describe the data we use for regional price levels and subsequently we briefly introduce also the household-level data with which we combine them. Since our objective is to assess the impact of regional price levels on income-based economic indicators, the variable of interest is the indicator of the price level in a given region relative to the national price level. A value lower than 100 indicates that the price level in the given region is lower than the national price level, whereas a value above 100 indicates the opposite. To the best of our knowledge, indicators of regional price levels are available only from four statistical offices in the world, namely the joint project of the BEA and BLS for US, ONS for UK, Statistics Japan for Japan and Turkstat for Turkey. In addition to these four countries, we are aware that the construction of regional price levels was considered by Statistics New Zealand, but no data from this exercise are available, as confirmed by Biggeri et al. (2017). We therefore needed to supplement this data with sources from academic literature. Given the existence of competing estimates for some countries (e.g. the UK, China and the US) we had to choose between the available sources. In general, we preferred indices based upon more representative baskets of goods. Preference was also given to more contemporary estimates i.e. estimates based upon more recent data (the bulk of empirical analysis was concluded in 2017 and therefore estimates of regional purchasing parities published after the date are not included in the data). In total, we use data on regional price levels for fourteen countries from the statistical offices and academic literature.

However, we were unable to include all the available regional price levels in our data set. The unavailability of control variables at the appropriate regional level prevented us from using the regional price levels provided by Majumder et al. (2011) for Vietnam (the General Statistics Office of Vietnam confirmed that they do not have data on the given regional level) and indices provided for India such as Majumder and Ray (2017) or Chakrabarty et al. (2015). In the case of India, the the Indian Ministry of Statistics and Program Implementation is only able to provide labour market indicators on the appropriate level and, although Dikhanov (2010) provides regional price levels for 10 regions of India, we were unable to include these in our data set due to the unavailability of control variables. In the case of China, we had to choose between estimates by Brandt and Holz (2006), Li et al. (2005) and Biggeri et al. (2017); we selected the regional price levels provided by Brandt and Holz (2006) even though both Li et al. (2005) and Biggeri et al. (2017) provide regional price levels based on larger and more up-to-date baskets of goods. The main reason for this was that the latter two studies do not provide regional price levels in the desired form i.e. the price level of the given region relative to the whole country. Furthermore the approach used by Li et al. (2005) was criticised by Biggeri et al. (2017). The selection of potentially less representative price indices is reflected during the selection of control variables for the considered model.

Also for Germany there are two different sets of estimates , specifically Roos (2006) and Deckers et al. (2016). Despite the fact that the latter is based on a more recent data set, we use the former., because Deckers et al. (2016) do not provide the regional price levels in the desired form. We are also aware that the Italian National Institute of Statistics provided indices for the capital cities of Italian regions of which Giusti et al. (2017) and Marchetti and Secondi (2017) provide examples of application. However we have preferred the ones used by Pittau et al. (2011) as these are provided in the desired form i.e. regional price level relative to the whole country. Japanese regional price levels could not be used because of unavailability of control variables at the appropriate regional level.

Even though different methodologies were employed to construct the regional price levels, all the estimates we use can be considered sufficiently representative of differences in regional price levels. Consequently the model we construct in the following section is based on a data set consisting of 227 regional price levels in 12 countries. We acknowledge that pooling estimates constructed using different methodologies and based on data of varied coverage involves making a strong assumption that these measure an identical phenomenon and are comparable. However, we consider the regional indices to be sufficiently robust estimates for our application. A complete list of the sources of the regional price levels we use, with short descriptions is provided in Table 1.

(Table 1 here)

The selection of explanatory variables was inspired by the models constructed in Roos (2006), Janský and Kolcunová (2017) and Blien et al. (2009). Valuable insights into the selection of control variables are also provided by Gelb and Diofasi (2015) and Ahec-Šonje and Nestić (2002). During the construction of the model we adopt assumptions identical to those in Roos (2006) and Janský and Kolcunová (2017). Under the following three assumptions, regional price levels are determined only by differences in regional supply and demand: spatial segmentation of regional markets which makes any strategic price setting or spatial arbitrage impossible; short term immobility of consumers and firms; trading of intermediate inputs between the regions at no transportation cost and at the same price in each of the regions.

We have selected the following variables for the outlined reasons. Regional disposable income can be considered a major determinant of the strength of regional demand, and the number of consumers living within a given region is also highly correlated with the strength of regional demand. Population density may be of influence for example through underlying quality differences, associated mainly with the service sector. Area and GDP per capita are also tested as explanatory variables. We also considered a set of variables describing the labour market, which included the employment, unemployment and participation rate, as well as a set of dummy variables describing the characteristics of the region. Dummy variables indicating the presence of the capital city within the region or a city with more than 1% or 2% of the total populations were tested. Access to the sea and the presence of an international airport in the region could translate into a bigger touristic attractiveness, which can be considered an important source of demand. A dummy variable for Chinese regions was also included to control for the above described imprecisions in the dependent variable as well as imprecisions in the control variables that are described below. A full list of control variables is outlined in Table 2.

(Table 2 here)

The majority of the explanatory variables needed to construct the data set were available from the respective statistical offices or supranational institutions such as Eurostat or OECD. However, some were more difficult to obtain. For the Philippines, the required data are available from the Philippine Statistical Yearbook, published annually by the Philippine Statistical Authority, but some of the time series are only provided in three-year intervals thus making it impossible to construct the data set for one base year. Although the majority of data were available for the year 2010, these did not include disposable income, which was released for the years 2009 and 2012. Since there were only minimal changes in the regional differences in the distribution of disposable income between the years 2009 and 2012, we decided to use the data for the year 2012 as we believe these sufficiently represent the differences in disposable income between regions.

Labour market statistics for Chinese regions required us to make further adjustments: the Chinese National Bureau of Statistics (NBS) only provides data on the regional unemployment rate. However, sufficient data are provided in the Statistical yearbook published by the NBS to enable us to construct both participation and employment rates. Nevertheless, an important caveat remains. The labour market statistics and the unemployment rate in particular have been judged highly unrepresentative by a number of academic studies such as Liu (2012) and Giles et al. (2005). Cai et al. (2013) even consider them “almost useless” as they are likely to underestimate the true values. Even though national unemployment rates for China have been estimated, we found no source of regional unemployment rates for China in the academic literature. We therefore used the labour market indicators based on the Statistical Yearbook published by the NBS. As we noted above, the imprecisions in the Chinese data are controlled for in our analysis by a dummy variable for Chinese regions.

To make the data comparable, we transform the variables according to the formula, similarly to Janský and Kolcunová (2017). The transformation is necessary as some of the regions within the data set are of similar size as some countries. Given that we estimate a relative indicator, i.e. regional price level relative to the price level of the country, we adjust the variables by dividing the regional value by average national value and by multiplying it by 100.

Once we have estimated the regional price levels, we then combine them with reliable household-level microdata to construct robust Gini coefficients and poverty headcount ratios. We use household income survey data gathered and harmonized by the Luxembourg Income Study (LIS). Even though the LIS database provides data sets for 49 countries, the variable indicating the data's regional origin was not included in all the data sets and control variables were not available for all the considered countries. We were therefore able to obtain regional price levels and indicators of poverty and inequality for a total of 21 countries. The full list includes two lower-middle-income countries (India and Georgia), two upper-middle-income countries (Colombia and Mexico) and most of the countries are classified as high income by the World Bank (Australia, Austria, Canada, Czech Republic, Denmark, France, Germany, Greece, Hungary, Ireland, Italy, Poland, Slovakia, Spain, Switzerland, UK and US).

4 Methodology

In this section we first describe how we estimate the regional price levels and then how we adjust the income inequality indicators for these regional prices.

4.1 Estimating regional price levels

We adopt the methodology used by Roos (2006) and Janský and Kolcunová (2017), after carefully considering other feasible frameworks for the estimation of regional price levels. We thus construct an OLS model and use it to make an out-of-sample prediction of regional price levels for countries for which there are no existing estimates.

We prefer this regression framework over the CPD methodology and over constructing spatial indices based on data collected for temporal ones because of the high demand the latter two approaches impose on the input data that we would not be able to meet for a number of countries. We deem the chosen methodology suitable for our objectives, but we believe it is important to provide a discussion of its weaknesses. As we noted in the literature review, Blien et al. (2009) pointed out that the OLS regression framework involves estimating new indices based on estimates for which the standard errors are unknown, which can lead to bias in the standard errors formulas and erroneous inference. As a remedy, we considered adopting the multiple imputation methodology. However, the multiple imputation framework adopted by Blien et al. (2009) was superior to the OLS regression framework because the authors have used the completed data set as a basis for further regression. The proposed approach suffers, as detailed, for example, by Van Buuren (2012), from weaknesses identical to the OLS regression framework if the estimated regional price levels are used for adjustment of income for intra-country differences of price levels.

There are some other alternatives to the cross-sectional regression for estimation of regional price levels, which could be replaced by other regression frameworks. For example, application of a general equilibrium model might be more appropriate as some of the explanatory variables may be determined simultaneously. Instrumentation for the endogenous variables could be a possible solution too. However, the available data prevented us from using these methods. Another possible approach towards the construction of the model would be to create a single time series model for each country. As we noted in the previous section, the underlying data set consists of 227 regional price levels observed in 12 countries. Although we consider this data set to be sufficiently representative of the mechanisms underlying the determination of regional price levels, panel data are only available for the US and Turkey. We constructed a linear model using an OLS regression under the assumption that the effect of the explanatory variables is the same across all regions of the considered countries. We acknowledge the strength of the assumption required for the application of this framework, but, in combination with the best available data outlined in the previous section, this seems the most suitable approach to shedding more light on regional income inequality.

Although our chosen model is sufficient for predicting regional price levels, there are opportunities for improvement and further research. Perhaps its main weakness is that it does not include any variable indicating housing prices within the given region, even though these have been found to be significant determinants of regional price levels, for example by Aten (2017). Due to the unavailability of any unified measure for all the countries in our data set, these could not be included in the model. The unavailability of a unified indicator of the volume of tourists visiting each region also restricts the model's predictive powers because tourists serve as an important source of demand. Price levels could also be influenced by institutional quality, as Gelb and Diofasi (2015) suggest, but no usable measure of intra-country differences in this aspect could be found either.

4.2 Adjusting income inequality and relative poverty indicators for regional price levels

To assess the possible bias caused by regional price levels in the measures of income inequality and poverty, we chose - as good representatives of a range of measures - the Gini coefficient and the poverty headcount ratio (defined as the ratio of a number of people living below the poverty threshold to total population). The Gini coefficient was selected as one of the most widely applied indicators in academic research. We have constructed it upon three types of income indicators – household income, per capita income and equalised income, our preferred type. We obtain equalised income by dividing household income by the square root of the number of members in the house. For a poverty indicator, after careful consideration of the selected sample of countries, we chose a relative poverty line rather than an absolute threshold, which makes this poverty measure an alternative income inequality indicator. We set the relative poverty line at the official OECD poverty threshold which is 50% of median income. We also consider additional poverty lines at 40% and 60% of median income.

To adjust the Gini coefficient and poverty headcount ratio for regional price levels, we proceed in a very straightforward way that can be summarized by the following equation:

$$Y_i^{adj} = \frac{Y_i^{unadj}}{regional\ price\ level_i}$$

To get income Y_i^{adj} adjusted for regional price levels, we divide household-level incomes in region *i*, Y_i^{unadj}, by the relevant regional price level. With this adjusted income Y_i^{adj} we re-estimate the Gini coefficient and poverty headcount ratio. Also, we use a bootstrap procedure to test the statistical significance of difference between the unadjusted and adjusted income inequality estimates. A random sample of length equal to the number of observations in the income survey for the given country is drawn with replacement from the underlying data set. A Gini coefficient is then constructed based on this sample and this procedure was repeated 20 thousand times. Under the assumption of a normal distribution the standard deviation of the resulting distribution of the obtained indicators is used to test for statistical significance.

5 Results

In this section we first discuss our estimated regional price levels for countries for which there have been none available so far. Subsequently, we discuss how income inequality and relative poverty indicators change when we adjust them for regional price levels.

5.1 Estimating regional price levels

Our objective is to estimate a regression model explaining the determination of regional price levels. After running a series of regressions, we found that a level-level model provides the best fit for the data and we adopted the following algorithm of model construction - all variables outlined in Table 2 are regressed on the dependent variable and the most statistically significant one is kept in the model. The remaining variables are then added successively to the chosen variable so as to identify the second most significant one, which is then added to the model. These steps are repeated until none of the remaining variables are found to be of statistical significance. In order to be considered as significant the control variable has to be statistically significant at least at the 5% level.

The best fit is provided by model 1 presented in Table 3. We find four variables to be significant at the desired level – disposable income, population density, area of the region and the dummy variable for Chinese regions. Nearly 40% of the variance in the explanatory variable is explained by this model. Even though this may seem low compared to previous applications of this methodology, the significant heterogeneity of the considered sample should be borne in mind. The fit of the resulting model cannot be as high as in the case of a more homogenous sample of countries such as member countries of the EU or regions within a single country. However, to assess the quality of the model we teste the assumptions necessary for its construction. We conduct a Ramsey regression specification error test (RESET) to test for possible functional form misspecification, but we find no evidence of this (even at the 20% significance level). We apply the White's test for heteroskedasticity and we reject the null hypothesis of homoskedasticity at 1% significance level, indicating that there is enough evidence of heteroskedasticity in the data set. We thus use White's standard errors to account for its presence and we report also these results in Table 3. Consequently, the population density and area variables were no longer significant even at the 10% level.

(Table 3 here)

Furthermore, we identify a large number of outliers and high leverage points in the data set and we therefore construct the model based on a restricted data set, so as to evaluate their impact on the estimation. We identify outliers by application of Cook's distance and all points with Cook's distance higher than the conventional cut-off point 4/n, where n is the number of observations in the data set. Using this procedure, we identify seventeen regions.[2] The model based on the restricted data set is reported as Model 2 in Table 3. We test for heteroskedasticity and functional form misspecification by the same means as for the model based on the full data set and we did not find evidence of either at 10% significance level. We conduct a regression on the unrestricted data set using the Huber weighting function and we report the resulting model in Table 3. As can be seen from the reported estimates the outlying observations have a significant effect on the estimated coefficients and it is therefore important to take their influence into consideration when selecting the model to use to predict regional price levels. After careful consideration, we select model 2 since it does not suffer from the influence of outliers nor it is subject to heteroskedasticity. We now turn to testing this preferred model.

[2] Prague (Czechia), Banska Bystrica (Slovakia), Scotland (United Kingdom), District of Columbia (US), Beijing (China), Neimenggu (China), Shanghai (China), Zhejiang (China), Hainan (China), Chongqing (China), Yunnan (China), Xizang (China), Qinghai (China), Manila (Philippines), Queensland (Australia), Northern Territory (Australia) and Western Australia (Australia).

We use the econometric model for of out-of-sample predictions of regional price levels, so it is necessary to evaluate its predictive powers. To this end, we perform a set of cross validations, consisting of leave one out and 3-fold cross validations and we record the resulting root mean squared errors. While the root mean square error of the selected model is 5.334, after performing the leave one out and 3-fold cross validations we found that the root mean square error grew to 5.5 and 5.574 respectively. We consider this a modest increase, and we judge the model as suitable for the prediction of regional price levels.

We predict regional price levels for the 21 countries using the model described by the equation below, while we report it in detail in Table 4 in the Appendix.

$$Price\ level_i = 0.197118 * income_i + 0.264190 * area_i + 0.009841 * population\ density_i + 10.045246 * China_i + 77.115475$$

We also profited from existing official estimates of regional price levels and used these to assess the quality of our predictions. More specifically, we compared our estimated regional price levels with the official indices constructed by the ONS for the UK and the BEA for the US. Even though other official sources of regional price level exist, these are the only countries for which they are available for the considered year. In the case of the UK, our model has an overall tendency to overestimate regional price levels. The average bias is 0,03% with the West Midlands, North East England and Northern Ireland suffering from the biggest underestimations, while Scotland and London suffer from the biggest overestimations. The average bias in the case of the US is -0,44%. The biggest overestimations of regional price levels occur for the District of Columbia and South and North Dakota. This can be explained by the fact that no variable indicating differences in rents was included in the model; according to Aten (2017) high price levels in the District of Columbia and Hawaii are due to high rents. The states of Hawaii, California and New York suffered from the biggest underestimation. For the UK and US, both our predicted and the official published values will be used for further calculations and we will primarily discuss indicators based on the latter, since the official indices should be most informative of the true differences in regional price levels.

5.2 Adjusting income inequality and relative poverty indicators for regional price levels

We first discuss our results for the Gini coefficient and then similarly for the poverty headcount ratio; we include complete results in Table 5 in Appendix. We calculate Gini coefficients for all considered countries both adjusted and unadjusted for regional price levels. In the majority of cases the adjustment for regional price levels led to downward adjustment of the respective indicators. In the case of Gini coefficients based on household income, there are 15 cases of downward adjustments and 6 cases in which the adjustment for differences in regional price levels lead to no change in the considered indicator. Income inequality is overestimated for 19 out of the 21 countries by Gini coefficients based on per capita income, while it is underestimated in the remaining two cases. Considering the Gini coefficient based on equalised income is associated with overestimation of income inequality in 16 cases, underestimation in 2 cases and no change in 3 cases.

(Graph 1 here)

The resulting changes in the Gini coefficients based on equalised income are presented in Graph 1. The Gini coefficient based on data adjusted for regional price levels is plotted on the horizontal axes and the absolute value of the difference from the Gini based on income data unadjusted for regional price levels is plotted on the vertical axis. We find the largest distortions in absolute terms for middle income countries, namely Georgia and Columbia. Among the high income countrie,s the Gini coefficients of Italy and Hungary are subject to the biggest changes. The changes in Table 5 may seem significant in nominal terms; we proceed to assess their statistical significance using the following bootstrap procedure. We use red points to mark countries with statistically significant - at the 5% level - adjustments and blue points countries with non-significant adjustments.

The adjustment for differences in regional price levels results in statistically non-significant adjustments for the majority of considered countries. Georgia is the only country for which the induced adjustments are statistically significant for Gini coefficients based on all types of income. In the case of the Gini coefficient based on the per capita income statistically significant adjustments are also induced for the

Czech Republic and Hungary. Statistically significant differences between adjusted and unadjusted Gini coefficients based on equalised income are also discovered for Colombia, Czech Republic and Italy. The difference induced by the adjustment for regional price levels was found for the Gini coefficient of Italy based on total household income.

We calculate poverty headcount ratios, as with the Gini coefficients, both adjusted and unadjusted for regional price levels. We provide the estimated values in Table 6 in the Appendix using equalised disposable income is used to construct the poverty headcount ratios. In the case of the poverty headcount ratio based on the official OECD poverty line of 50% of median income, negative adjustment occurred in 12 cases, while 11 and 13 negative adjustments (respectively) were observed for the poverty ratios based on the additional poverty lines of 40% and 60% of median income. We find upward adjustments for nine countries for the poverty headcount ratio based on the middle and low poverty lines and for seven countries we find upwards adjustments for the indicator based on the highest poverty line. In one case there is no adjustment to the indicators based on the low and high poverty lines.

(Graph 2 here)

Graph 2 provides a representation of the results similar to that shown in Graph 1. The poverty thresholds and absolute values of the adjustments are based on the 60% of median income poverty threshold. Unlike in the case of the Gini coefficients, the biggest adjustments are encountered for two high income countries: Italy and Ireland. Large adjustments are also encountered in the group of middle income countries. We assess the statistical significance of the adjustments using a bootstrap procedure similar to the one used for the Gini coefficient. The only country for which the differences between the unadjusted and adjusted poverty headcount ratios are statistically significant for all considered poverty lines is France. For the poverty headcount ratio based on the lowest poverty line no other statistically significant differences between the adjusted and unadjusted indicators are identified. Statistically significant differences are found for Italy and Mexico in the case of the indicator based on the middle poverty line and for Colombia, Georgia, Ireland and Italy in the case of the indicator based on the highest poverty line.

6 Conclusion

In this paper we have assessed the influence of regional price levels on indicators of income inequality and relative poverty. Since, people, not countries, are poor and where the poorest people live can change as we have seen with the shift from the world of the "bottom billion" by Collier (2007) to the "new bottom billion" of Sumner (2011). We argue that our understanding of the geography of inequality and poverty might further change if we knew more about the characteristics of the regions in which people live and what implications these might have for poverty. One such characteristic is regional price level, which has been shown to substantially differ across regions within specific countries. However, less is known about how important regional price levels are for income inequality and relative poverty.

To fill in this gap, we have provided new cross-country evidence on the basis of the best available data and a two-step approach. First, we have estimated regional price levels for a selected group of 21 countries using an econometric model based on existing regional price level data for 12 countries provided by previous academic research and national statistical offices. Second, we have combined our estimated regional prices levels with LIS household-level data to construct Gini coefficients and poverty headcount ratios, both adjusted and unadjusted for differences in regional price levels. Consequently, we have found that in the majority of the considered cases failure to adjust for intra-country spatial price differences results in an overestimation of income inequality and poverty, but only for a minority of those these differences are statistically significant.

We find that for some countries, the Gini coefficients and poverty headcount ratios are statistically significantly different when adjusted for the regional price levels. For example, adjusting for regional price levels lowers the Gini coefficients by 2% for Italy, 3% for Columbia and 4% for Georgia, while it increases the poverty headcount ratio by 6% for France and 7% for Ireland. We conclude that regional price levels affect income inequality and should be taken into account by policy makers and future research, in particular where middle income countries are concerned.

References

Ahec-Šonje, Amina, a Danijel Nestić. 2002. „International Differences in Price Levels: An Empirical Analysis". *Croatian Economic Survey*, č. 4: 85–120.

Aten, Betina, a Tatiane Menezes. 2002. „Poverty price levels: an application to Brazilian metropolitan areas". In *World Bank ICP Conference, Washington, DC*, 11–15. http://www.rug.nl/ggdc/productivity/pwt/related-research-papers/ba-tatwb.pdf.

Aten, Bettina H. 2017. „Regional price parities and real regional income for the United States". *Social Indicators Research* 131 (1): 123–143.

Aten, Bettina H., Eric B. Figueroa, a Troy M. Martin. 2012. „Regional price parities for states and metropolitan areas, 2006-2010". *Survey of Current Business* 92 (8): 229–242.

Aten, Bettina H., Eric B. Figueroa, a B. M. Vengelen. 2011. „Real personal income and regional price parities for states and metropolitan areas, 2008–2012". *change* 2012: 2012.

Biggeri, Luigi, Guido Ferrari, a Yanyun Zhao. 2017. „Estimating cross province and municipal city price level differences in China: Some experiments and results". *Social Indicators Research* 131 (1): 169–187.

Biggeri, Luigi, Tiziana Laureti, a Federico Polidoro. 2017. „Computing sub-national PPPs with CPI data: an empirical analysis on Italian data using country product dummy models". *Social Indicators Research* 131 (1): 93–121.

Blien, Uwe, Hermann Gartner, Heiko Stüber, a Katja Wolf. 2009. „Regional price levels and the agglomeration wage differential in western Germany". *The Annals of Regional Science* 43 (1): 71–88.

Brandt, Loren, a Carsten A. Holz. 2006. „Spatial price differences in China: Estimates and implications". *Economic Development and Cultural Change* 55 (1): 43–86.

Collier, Paul. 2007. „The Bottom Billion: Why the Poorest Countries Are Failing and What Can Be Done About It."

Coondoo, Dipankor, Amita Majumder, a Ranjan Ray. 2004. „A method of calculating regional consumer price differentials with illustrative evidence from India". *Review of Income and Wealth* 50 (1): 51–68.

Deaton, Angus. 2003. „Prices and poverty in India, 1987-2000". *Economic and Political Weekly*, 362–368.

Deaton, Angus, a Alan Heston. 2010. „Understanding PPPs and PPP-based national accounts". *American Economic Journal: Macroeconomics* 2 (4): 1–35.

Deckers, Thomas, Armin Falk, a Hannah Schildberg-Hörisch. 2016. „Nominal or real? The impact of regional price levels on satisfaction with life". *The BE Journal of Economic Analysis & Policy* 16 (3): 1337–1358.

Dikhanov, Yuri. 2010. „Income Effect and Urban-Rural Price Differentials from the Household Survey Perspective". *ICP Global Office, February*. http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2014/09/18/000470435_20140918134506/Rendered/PDF/904660WP004.060l0Box0385317B0PUBLIC.pdf.

Dikhanov, Yuri, Chellam Palanyandy, a Eileen Capilit. 2011. „Subnational Purchasing Power Parities toward Integration of International Comparison Program and Consumer Price Index: The Case of the Philippines". https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1997802.

Gelb, Alan, a Anna Diofasi. 2015. „What Determines Purchasing Power Parity Exchange Rates?" https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2671574.

Giusti, Caterina, Lucio Masserini, a Monica Pratesi. 2017. „Local comparisons of small area estimates of poverty: an application within the tuscany region in italy". *Social Indicators Research* 131 (1): 235–254.

Chakrabarty, Manisha, Amita Majumder, a Ranjan Ray. 2015. „Preferences, spatial prices and inequality". *The Journal of Development Studies* 51 (11): 1488–1501.

Janskỳ, Petr, a Dominika Kolcunová. 2017. „Regional differences in price levels across the European Union and their implications for its regional policy". *The Annals of Regional Science* 58 (3): 641–660.

Krugman, Paul. 1991. „Increasing returns and economic geography". *Journal of political economy* 99 (3): 483–499.

Li, Xiumin, Lili Zhang, a Yashu Du. 2005. „Study on the method of regional purchasing power parity in China". *China-US Business Review* 4 (7): 41–48.

Luxembourg Income Study (LIS) Database, http://www.lisdatacenter.org (multiple countries; February 2017 - September 2017). Luxembourg: LIS.

Majumder, Amita, a Ranjan Ray. 2017. „Estimates of Spatial Prices in India and Their Sensitivity to Alternative Estimation Methods and Choice of Commodities". *Social Indicators Research* 131 (1): 145–167.

Majumder, Amita, Ranjan Ray, Kompal Sinha, a others. 2011. „Estimating intra country and cross country purchasing power parities from household expenditure data using single equation and complete demand systems approach: India and Vietnam". *Economics discussion paper* 34 (11). https://www.researchgate.net/profile/Ranjan_Ray/publication/254429660_Estimating_Intra_Country_and_Cross_Country_Purchasing_Power_Parities_from_Household_Expenditure_Data_Using_Single_Equation_and_Complete_Demand_Systems_Approach_India_and_Vietnam/links/0f31753b1f2e264270000000.pdf.

Marchetti, Stefano, a Luca Secondi. 2017. „Estimates of Household Consumption Expenditure at Provincial Level in Italy by Using Small Area Estimation Methods:“Real” Comparisons Using Purchasing Power Parities". *Social Indicators Research* 131 (1): 215–234.

Mishra, Ankita, a Ranjan Ray. 2014. „Spatial variation in prices and expenditure inequalities in Australia". *Economic Record* 90 (289): 137–159.

Mogstad, Magne, Audun Langørgen, a Rolf Aaberge. 2007. „Region-specific versus country-specific poverty lines in analysis of poverty". *The Journal of Economic Inequality* 5 (1): 115–122.

Pittau, Maria Grazia, Roberto Zelli, a Riccardo Massari. 2011. „Do spatial price indices reshuffle the Italian income distribution?" *Modern Economy* 2 (03): 259.

Radvansky, Marek, Ladislav Fuchs, a others. 2012. „Computing real income at NUTS 3 regions". EcoMod.

https://ecomod.net/system/files/Computing%20real%20income%20at%20NUTS%203%20regions%20Radvansky%20Fuchs.pdf.

„Reale Kaukraft 2008: Einkommen unter Berücksichtigung Des Regionalen Preisniveaus". 2009. Österrechische Gesellshaft für Marketing.

Rokicki, Bart\lomiej. 2015. „Regionalne deflatory PPP w Polsce oraz ich wykorzystanie". *Ekonomista*, č. 1: 67–80.

Roos, Michael WM. 2006. „Regional price levels in Germany". *Applied Economics* 38 (13): 1553–1566.

Suedekum, Jens. 2006. „Agglomeration and regional costs of living". *Journal of Regional Science* 46 (3): 529–543.

Summers, Robert. 1973. „International price comparisons based upon incomplete data". *Review of Income and Wealth* 19 (1): 1–16.

Sumner, Andy. 2010. „Global Poverty and the New Bottom Billion: What if Three-Quarters of the World's Poor Live in Middle-income Countries?" *IDS Working Papers* 2010 (349): 01–43.

Van Buuren, Stef. 2012. *Flexible imputation of missing data*. CRC press. https://www.google.com/books?hl=cs&lr=&id=M89TDSml-FoC&oi=fnd&pg=PP1&dq=van+buuren+flexible+imputation&ots=BcpOnqWyke&sig=O0hluyTrsXNphzsh19zDI_PKRCM.

Table 1 – Sources of regional price levels

Country	Year	Source
Australia	*2009*	Estimation done by the the Exact Affine Stone Index demand system, based on the data from the Household Expenditure Survey. *Source: MISHRA, Ankita; RAY, Ranjan. Spatial variation in prices and expenditure inequalities in Australia. Economic Record, 2014, 90.289: 137-159.*
Austria	*2008*	Joint paper of the Austrian Statistical Office and Österreichische Gessellschaft für Marketing. *Source: "Reale Kaukraft 2008: Einkommen Unter Berücksichtigung Des Regionalen Preisniveaus" (2009)*
China	*2014*	Regional deflators based on a 1990 basket published by the NBS. Shifted by CPI. *Source: BRANDT, Loren; HOLZ, Carsten A. Spatial price differences in China: Estimates and implications. Economic development and cultural change, 2006, 55.1: 43-86.*
Czech Republic	*2012*	Estimation based on adjusted Eurostat methodology and consumption data. *Source: KRAMULOVÁ, Jana; MUSIL, Petr; ZEMAN, Jan; MICHLOVÁ, Radka. 2016. "Regional Price Levels in the Czech Republic - Past and Current Perspectives."*
Germany	*2010*	Regional deflators based on a prediction model constructed on data for 50 German cities. 1993 indices shifted to 2010 by regional CPI. *Source: ROOS, Michael WM. Regional price levels in Germany. Applied Economics, 2006, 38.13: 1553-1566.*

Italy	*2006*	Estimates produced by the National Bank of Italy in a joint project with Italian Office of Statistics. *Source: PITTAU, Maria Grazia; ZELLI, Roberto; MASSARI, Riccardo. Do spatial price indices reshuffle the Italian income distribution?. Modern Economy, 2011, 2.03: 259.*
Philippines	*2010*	Data set constructed by a project of the ADB, based on data collected for the construction of CPI *Source: DIKHANOV, Yuri; PALANYANDY, Chellam; CAPILIT, Eileen. Subnational Purchasing Power Parities toward Integration of International Comparison Program and Consumer Price Index: The Case of the Philippines. 2011.*
Turkey	*2014*	Data constructed by Turkstat as a by-product of the construction of the Spatial Adjustment Factor, based on a 2012 survey and shifted to 2014 by regional CPI. *Source: "Purchasing Power Parity (PPP)." 2017. Turkish Statistical Institute. http://www.turkstat.gov.tr/PreTablo.do?alt_id=1065.*
United Kingdom	*2010*	Data produced by the Office for National Statistics during the construction of the UK Spatial Adjustment Factors for Eurostat. *Source: "UK Relative Regional Consumer Price Levels for Goods and Services for 2010." 2011. Office for National Statistics. https://data.gov.uk/dataset/regional_consumer_price_levels .*
United States	*2014*	Price levels constructed by the BEA and BLS, based on consumption data, following the methodology outlined in Aten (2017).

	Source: "Regional Data." 2017. *Bureau of Economic Analysis*. February 20. https://www.bea.gov/iTable/iTable.cfm?reqid=70&step=1&isuri=1&acrdn=8#reqid=70&step=30&isuri=1&7022=101&7023=8&7024=non-industry&7033=-1&7025=0&7026=xx&7027=2014&7001=8101&7028=-1&7031=0&7040=-1&7083=levels&7029=101&7090=70.

Source: Authors.

Table 2 – Considered Explanatory variables

Variable	Description	Source
Income	Disposable household income. (relative value)	1, 2, 3, 6, 7
GDP	GDP per capita. (relative value)	1, 2, 6, 7
Employment	Employment rate - people aged 15 and over relative to total working population aged 15 and over. (relative value)	1, 2, 6, 7
Unemployment	Unemployment rate - people aged 15 and over relative to total working population aged 15 and over. (relative value)	1, 2, 3, 4, 5, 6, 7
Participation Rate	Participation rate - people aged 15 and over relative to total working population aged 15 and over. (relative value)	1,2, 6, 7
Population density	Regional population density. Inhabitants per kilometre squared. (relative value)	1, 2, 6, 7
Area	Total area covered by the respective region. In kilometre squared. (percentage of country total)	1, 2, 6, 7
Population	Total population living in the given region. (percentage of country total)	1, 2, 6, 7
Capital	Indicates the presence of capital in the region. (Dummy)	

City1	Indicates the presence of a city with more than 1% of total country population. (Dummy)
City2	Indicates the presence of a city with more than 2% of total country population. (Dummy)
Sea	Indicates that the region has access to sea. (Dummy)
Airport	Indicates presence of an international airport in the region. (Dummy)

Notes: 1: OECD 2: Eurostat 3: Turkstat 4: Czech statistical office 5: Slovak statistical office 6: Philippine Statistical Yearbook 7: National Bureau of Statistics (China)

Source: Authors.

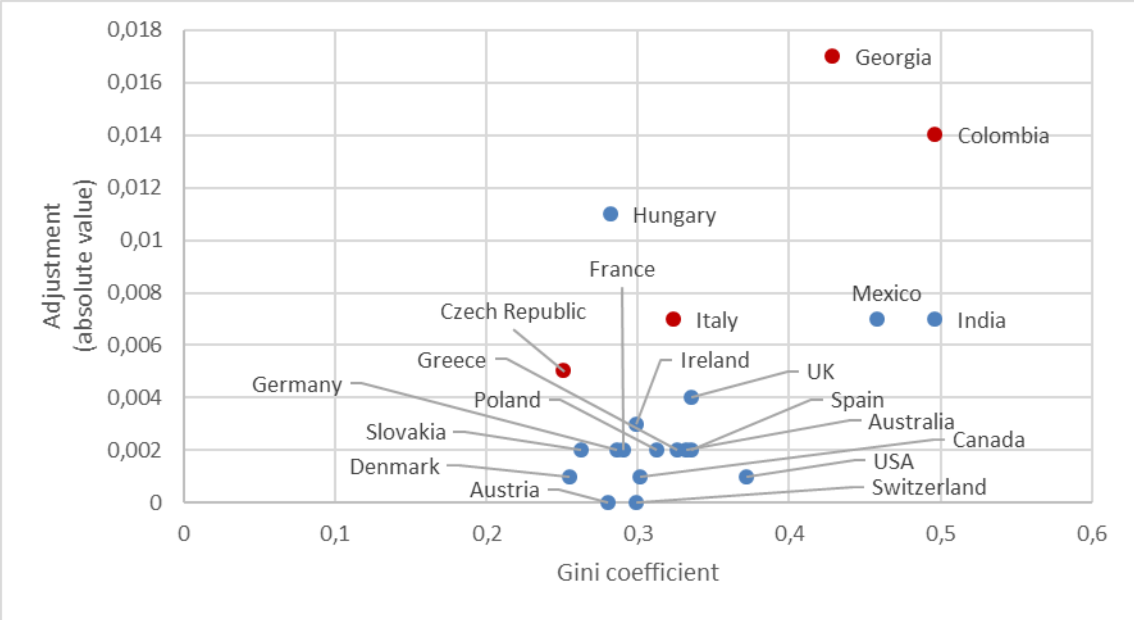
Table 3 – Estimation results

Variable	Model 1	Model 1 (heterosked. robust)	Model 2	Model 3 (Huber)
Intercept	75.546362 *** (2.898164)	75.546362 *** (4.3770363)	77.115475 *** (2.303701)	76.1884 *** (1.9643)
Income	0.197891 *** (0.028557)	0.197891 *** (0.0333076)	0.197118 *** (0.023013)	0.2091 *** (0.0194)
Population Density	0.006718 * (0.002751)	0.006718 (0.0042125)	0.009841 ** (0.003031)	0.0050 ** (0.0019)
Area	0.620339 *** (0.114824)	0.620339 (0.4348886)	0.264190 ** (0.094214)	0.2408 ** (0.0778)
China	12.507829 *** (1.681590)	12.507829 *** (2.1475756)	10.045246 *** (1.248757)	11.4376 *** (1.1398)
Adjusted R-squared	39.15	39.15	43.94	
F-statistics	37.35 on 4 and 222 DF		40.17 on 4 and 205 DF	

Notes: “***” p < 0,001; “**” p < 0,01; “*” p < 0,05; “.” p < 0,1

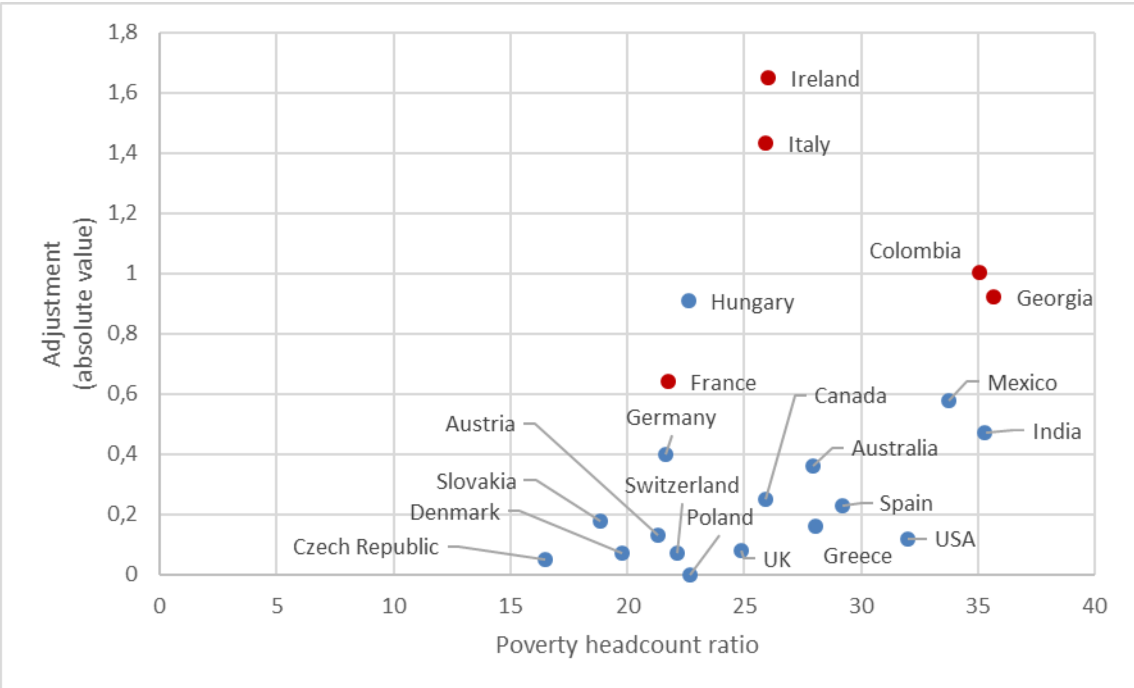
Source: Authors.

Graph 1



Source: Authors.

Graph 2



Source: Authors.

Appendix

Table 4 – Estimated regional price levels

Country	Regions	Regional Price Level
Australia	AU1: New South Wales	97.68452
	AU2: Victoria	95.48913
	AU3: Queensland	99.89687
	AU4: South Australia	97.03205
Year: 2010	AU5: Western Australia	104.37522
	AU6: Tasmania	93.18972
	AU7: Northern Territory	101.90838
	AU8: Australian Capital Territory	118.92521
Source: OECD		
Austria	AT11: Burgenland (AT)	97.93571
Year: 2010	AT12: Lower Austria	103.86642
	AT13: Vienna	104.92771
	AT21: Carinthia	99.21879
	AT22: Styria	101.75205
	AT31: Upper Austria	100.80905
	AT32: Salzburg	99.53404
	AT33: Tyrol	100.35769
	AT34: Vorarlberg	98.31994
Source: OECD, Eurostat		
Canada	CA10: Newfoundland and Labrador	96.30647
Year: 2010	CA11: Prince Edward Island	94.03379
	CA12: Nova Scotia	94.50578
	CA13: New Brunswick	94.52463

	CA24: Quebec	98.03537
	CA35: Ontario	99.74375
	CA46: Manitoba	95.98719
	CA47: Saskatchewan	98.30711
	CA48: Alberta	102.77842
	CA59: British Columbia	99.52089
Source: OECD		
Czech republic	CZ010: Prague	113.34114
Year: 2010	CZ020: Central Bohemia	103.42646
	CZ031: South Bohemia	99.79608
	CZ032: Plzen	99.98591
	CZ041: Karlovy Vary	96.61041
	CZ042: Ústí nad Labem	97.46901
	CZ051: Liberec	97.75348
	CZ052: Hradec Králové	98.61385
	CZ053: Pardubice	97.63502
	CZ063: Vysocina	98.75145
	CZ064: South Moravia	99.79084
	CZ071: Olomouc	97.31707
	CZ072: Zlín	97.33448
	CZ080: Moravia-Silesia	97.95307
Source: OECD, Eurostat		
Germany	DE1: Baden-Württemberg	102.36874
Year: 2010	DE2: Bavaria	105.17737
	DE3: Berlin	101.35535
	DE4: Brandenburg	97.36341
	DE5: Bremen	99.47814

	DE6: Hamburg	103.62341
	DE7: Hesse	100.42821
	DE8: Mecklenburg-Vorpommern	95.82989
	DE9: Lower Saxony	100.3701
	DEA: North Rhine-Westphalia	100.89659
	DEB: Rhineland-Palatinate	99.90994
	DEC: Saarland	97.18598
	DED: Saxony	96.59458
	DEE: Saxony-Anhalt	95.8451
	DEF: Schleswig-Holstein	99.19458
	DEG: Thuringia	95.77964
Source: OECD, Eurostat		
Greece	EL30: Attica	101.59483
Year: 2010	EL41: North Aegean	98.87398
	EL42: South Aegean	99.94105
	EL43: Crete	96.33402
	EL51: Eastern Macedonia, Thrace	97.76114
	EL52: Central Macedonia	100.2758
	EL53: Western Macedonia	99.27692
	EL54: Epirus	98.96155
	EL61: Thessaly	99.52924
	EL62: Ionian Islands	97.72783
	EL63: Western Greece	97.77542
	EL64: Central Greece	98.45352
	EL65: Peloponnese	99.65282
Source: OECD		
Italy	ITC1: Piedmont	102.90487

Year: 2010	ITC2: Aosta Valley	100.14943
	ITC3: Liguria	101.69777
	ITC4: Lombardy	105.96176
	ITF1: Abruzzo	96.59713
	ITF2: Molise	94.13801
	ITF3: Campania	94.96977
	ITF4: Apulia	94.52658
	ITF5: Basilicata	92.72396
	ITF6: Calabria	93.11513
	ITG1: Sicily	95.00983
	ITG2: Sardinia	95.65279
	ITH2: Province of Trento	101.53935
	ITH3: Veneto	101.53806
	ITH4: Friuli-Venezia Giulia	100.8577
	ITH5: Emilia-Romagna	103.99997
	ITI1: Tuscany	101.56546
	ITI2: Umbria	98.92282
	ITI3: Marche	98.96507
	ITI4: Lazio	102.28112
Source: OECD		
Mexico	ME01: Aguascalientes	100.1665
Year: 2010	ME02: Baja California Norte	106.4809
	ME03: Baja California Sur	105.54329
	ME04: Campeche	97.61403
	ME05: Coahuila	102.75654
	ME06: Colima	100.10858
	ME07: Chiapas	87.80632

	ME08: Chihuahua	100.93639
	ME09: Federal District (MX)	110.13456
	ME10: Durango	94.40735
	ME11: Guanajuato	94.0195
	ME12: Guerrero	90.53681
	ME13: Hidalgo	92.19274
	ME14: Jalisco	100.31621
	ME15: Mexico	97.80863
	ME16: Michoacan	93.57045
	ME17: Morelos	96.30574
	ME18: Nayarit	98.27394
	ME19: Nuevo Leon	110.41675
	ME20: Oaxaca	90.88205
	ME21: Puebla	92.42439
	ME22: Queretaro	98.77855
	ME23: Quintana Roo	104.08923
	ME24: San Luis Potosi	94.22833
	ME25: Sinaloa	98.86402
	ME26: Sonora	102.22831
	ME27: Tabasco	92.97001
	ME28: Tamaulipas	99.07517
	ME29: Tlaxcala	90.54269
	ME30: Veracruz	93.99582
	ME31: Yucatan	95.60933
	ME32: Zacatecas	92.79387
Source: OECD		
Poland	PL11: Lodzkie	100.71886

Year: 2010	PL12: Mazovia	104.98454
	PL21: Lesser Poland	99.31566
	PL22: Silesia	104.03775
	PL31: Lublin Province	97.64077
	PL32: Podkarpacia	95.76661
	PL33: Swietokrzyskie	97.10206
	PL34: Podlasie	96.92752
	PL41: Greater Poland	101.77201
	PL42: West Pomerania	100.14248
	PL43: Lubusz	98.14985
	PL51: Lower Silesia	100.96771
	PL52: Opole region	97.89255
	PL61: Kuyavian-Pomerania	98.33448
	PL62: Warmian-Masuria	97.70195
	PL63: Pomerania	99.94627
Source: OECD		
Slovakia	SK010: Bratislava Region	110.10203
Year: 2010	SK021: Trnava Region	100.39153
	SK022: Trencín Region	99.60711
	SK023: Nitra Region	100.92458
	SK031: Žilina Region	100.63328
	SK032: Banská Bystrica Region	100.39939
	SK041: Prešov Region	98.66668
	SK042: Košice Region	98.18542
Source: OECD		
Spain	ES11: Galicia	96.99961
Year: 2010	ES12: Asturias	98.17674

	ES13: Cantabria	97.03078
	ES21: Basque Country	104.20253
	ES22: Navarra	102.96119
	ES23: La Rioja	97.95787
	ES24: Aragon	101.32729
	ES30: Madrid	102.5979
	ES41: Castile and León	101.96578
	ES42: Castile-La Mancha	98.20629
	ES43: Extremadura	94.52463
	ES51: Catalonia	102.00895
	ES52: Valencia	96.17299
	ES53: Balearic Islands	97.36986
	ES61: Andalusia	97.78843
	ES62: Murcia	94.17045
	ES63: Ceuta	95.36365
	ES64: Melilla	93.76291
	ES70: Canary Islands	94.53415
Source: OECD		
Switzerland	CH01: Lake Geneva Region	101.73309
Year: 2010	CH02: Espace Mittelland	101.87372
	CH03: Northwestern Switzerland	99.08786
	CH04: Zürich	102.02712
	CH05: Eastern Switzerland	103.33436
	CH06: Central Switzerland	100.39696
	CH07: Ticino	96.74117
Source: OECD		
UK	UKC: North East England	95.77799

Year: 2010	UKD: North West England	97.55211
	UKE: Yorkshire and The Humber	97.07756
	UKF: East Midlands	97.68279
	UKG: West Midlands	97.16223
	UKH: East of England	101.21784
	UKI: Greater London	111.06019
	UKJ: South East England	103.31145
	UKK: South West England	100.82673
	UKL: Wales	97.1938
	UKM: Scotland	105.34398
	UKN: Northern Ireland	95.94888
Source: OECD		
US	US01: Alabama	94.32168
Year: 2010	US02: Alaska	102.92904
	US04: Arizona	94.32584
	US05: Arkansas	93.2444
	US06: California	99.08507
	US08: Colorado	96.88259
	US09: Connecticut	108.60466
	US10: Delaware	98.45376
	US11: District of Columbia	131.87896
	US12: Florida	97.48133
	US13: Georgia	94.79952
	US15: Hawaii	98.48113
	US16: Idaho	93.24702
	US17: Illinois	98.20394
	US18: Indiana	95.04015

US19: Iowa	96.20468
US20: Kansas	96.73821
US21: Kentucky	93.85364
US22: Louisiana	96.14932
US23: Maine	95.93634
US24: Maryland	102.71348
US25: Massachusetts	104.60288
US26: Michigan	95.14412
US27: Minnesota	97.86499
US28: Mississippi	92.76801
US29: Missouri	95.7206
US30: Montana (US)	95.11728
US31: Nebraska	97.46582
US32: Nevada	95.80573
US33: New Hampshire	101.09297
US34: New Jersey	104.92027
US35: New Mexico	94.10157
US36: New York	100.8382
US37: North Carolina	95.20302
US38: North Dakota	98.99385
US39: Ohio	95.80482
US40: Oklahoma	95.36345
US41: Oregon	94.86601
US42: Pennsylvania	98.53974
US44: Rhode Island	100.80122
US45: South Carolina	93.66462
US46: South Dakota	98.1785

	US47: Tennessee	95.73642
	US48: Texas	96.94689
	US49: Utah	93.11095
	US50: Vermont	97.69791
	US51: Virginia	99.74422
	US53: Washington	98.62971
	US54: West Virginia	93.33564
	US55: Wisconsin	96.4743
	US56: Wyoming	99.85102
Source: OECD		
India	IN01: Jammu and Kashmir	100.60004
Year: 2011	IN02: Himachal Pradesh	102.23309
	IN03: National Capital Territory of Delhi	114.48398
	IN04: Rajasthan	94.43811
	IN05: Uttar Pradesh	89.78485
	IN06: Sikkim	103.72771
	IN07: Arunachal Pradesh	118.48507
	IN08: Nagaland	105.92145
	IN09: Meghalaya	96.98235
	IN10: Assam	94.01247
	IN11: West Bengal	93.75037
	IN12: Gujarat	95.06701
	IN13: Dadra & Nagar Haveli	90.92459
	IN14: Maharashtra	94.79852
	IN15: Daman & Diu	99.54625
	IN17: Kerala	100.72192
	IN18: Punjab	100.81648

	IN19: Chandigarh	143.36527
	IN20: Haryana	97.14339
	IN21: Uttaranchal	93.69676
	IN22: Bihar	87.9142
	IN23: Jharkhand	89.04439
	IN24: Manipur	98.6274
	IN25: Mizoram	108.77647
	IN26: Tripura	90.06643
	IN27: Orissa	88.92586
	IN28: Madhya Pradesh	89.7943
	IN29: Chhattisgarh	89.12806
	IN30: Andhra Pradesh	91.2471
	IN31: Karnataka	94.00923
	IN32: Goa	95.42965
	IN33: Tamil Nadu	96.24076
	IN34: Puducherry	104.6927
Source: OECD, LIS		
Ireland	IE011: Border	100.30334
Year: 2010	IE012: Midlands	98.12135
	IE013: West	101.9853
	IE021: Dublin	105.57426
	IE022: Mid-East	100.03557
	IE023: Mid-West	100.63211
	IE024: South-East (IE)	100.32102
	IE025: South-West (IE)	101.9371
Source: OECD, Eurostat, Central Statistics Office		
Hungary	HU101: Budapest	120.63548

Year: 2012	HU102: Pest	98.0334
	HU211: Fejér	99.28554
	HU212: Komárom-Esztergom	101.70226
	HU213: Veszprém	97.59883
	HU221: Gyor-Moson-Sopron	99.98002
	HU222: Vas	96.0251
	HU223: Zala	97.22529
	HU231: Baranya	94.84158
	HU232: Somogy	97.91448
	HU233: Tolna	95.05931
	HU311: Borsod-Abaúj-Zemplén	100.38288
	HU312: Heves	102.47845
	HU313: Nógrád	100.01699
	HU321: Hajdú-Bihar	99.66641
	HU322: Jász-Nagykun-Szolnok	92.95895
	HU323: Szabolcs-Szatmár-Bereg	95.78271
	HU331: Bács-Kiskun	97.80714
	HU332: Békés	96.51163
	HU333: Csongrád	98.74007
Source: OECD, LIS		
Colombia	CO05: Antioquia	103.47823
Year: 2010	CO08: Atlántico	93.9945
	CO11: Bogotá Capital District	122.5191
	CO13: Bolívar	92.53499
	CO15: Boyacá	100.27239
	CO17: Caldas	110.48294
	CO18: Caquetá	97.1859

	CO19: Cauca	93.40026
	CO20: Cesar	94.00617
	CO23: Córdoba (CO)	96.59216
	CO25: Cundinamarca	95.07848
	CO27: Chocó	94.15329
	CO41: Huila	100.95977
	CO44: La Guajira	100.14996
	CO47: Magdalena	94.33076
	CO50: Meta	103.38765
	CO52: Nariño	92.86112
	CO54: Norte de Santander	89.91734
	CO63: Quindio	97.12577
	CO66: Risaralda	105.81102
	CO68: Santander	103.60819
	CO70: Sucre	94.32307
	CO73: Tolima	101.62101
	CO76: Valle del Cauca	96.09793
Source: OECD, LIS		
France	Île de France	138.09277
Year: 2010	Bassin Parisien	124.49122
	Nord - Pas-de-Calais	92.6996
	Est (FR)	99.53522
	Ouest (FR)	113.37341
	Sud-Ouest (FR)	108.22125
	Centre-Est (FR)	110.586
	Méditerranée	110.3167
	Guadeloupe	79.57267

	Martinique	79.56737
	Guyane	81.05839
	La Réunion	80.89873
	Mayotte	79.55359
Source: Eurostat		
Denmark	DK011: City of Copenhagen	103.48806
Year: 2010	DK012: Copenhagen suburbs	100.35473
	DK013: North Zealand	99.75029
	DK014: Bornholm	96.79779
	DK021: East Zealand	98.12853
	DK022: West and South Zealand	100.47078
	DK031: Fyn	98.50168
	DK032: South Jutland	101.96899
	DK041: West Jutland	101.12064
	DK042: East Jutland	100.29786
	DK050: North Jutland	101.4648
Source: OECD, Statistics Denmark		
Georgia	Kakheti	94.33005
Year: 2010	Tbilisi	137.2332
	Shida Kartli	89.22772
	Kvemo Kartli	92.39516
	Adjara (Automomous Region)	90.6118
	Samegrelo-Zemo Svaneti	96.45067
	Imereti, Racha-Lechkhumi and Kvemo Svaneti	110.12043
	Other regions	98.54101
Source: National Statistics Office of Georgia		

Source: Authors.

Table 5 – Gini coefficients and impacts of regional price levels

	Household income		Per capita income		Equalised income	
Country	**Adjusted**	**Unadjusted**	**Adjusted**	**Unadjusted**	**Adjusted**	**Unadjusted**
Australia	0.389 (0.39)	0.39	0.347 (0.387)	0.348	0.332 (0.274)	0.334
Austria	0.353 (no change)	0.353	0.303 (0.404)	0.304	0.28 (no change)	0.28
Based on: ***Reale Kaukraft 2008***	0.354 (0.394)		0.303 (0.404)		0.28 (no change)	
Canada	0.363 (0.358)	0.364	0.321 (0.386)	0.322	0.301 (0.364)	0.302
Colombia	0.51 (0.07)	0.523	0.53 (0.066)	0.543	0.496 (0.042)	0.511

Czech Republic	0.319 (0.114)	0.323	0.259 (0.028)	0.266	0.251 (0.071)	0.256
Based on: Kramulová et al. (2016)	0.319 (0.114)		0.258 (0.015)		0.25 (0.039)	
Denmark	0.347 (no change)	0.347	0.262 (0.332)	0.263	0.255 (0.323)	0.254
France	0.342 (no change)	0.342	0.322 (0.408)	0.323	0.292 (no change)	0.292
Georgia	0.466 (0.0165)	0.479	0.436 (0.0005)	0.457	0.427 (0.0008)	0.446
Germany	0.353 (0.291)	0.355	0.303 (0.297)	0.305	0.286 (0.293)	0.288

Greece	0.367 (no change)	0.367	0.337 (0.431)	0.338	0.327 (0.427)	0.328
Hungary	0.343 (0.183)	0.35	0.301 (0.0347)	0.317	0.281 (0.0787)	0.293
India	0.52 (0.131)	0.524	0.511 (0.148)	0.515	0.499 (0.141)	0.503
Ireland	0.351 (no change)	0.351	0.323 (0.296)	0.32	0.298 (0.345)	0.296
Italy	0.357 (0.138)	0.362	0.348 (0.086)	0.355	0.323 (0.044)	0.331
Based on: ***Pittau et al. (2011)***	0.354 (0.041)		0.348 (0.086)		0.323 (0.044)	
Mexico	0.474	0.479	0.493	0.499	0.459	0.465

	(0.079)		(0.097)		(0.061)	
Poland	0.36 (0.204)	0.362	0.344 (0.329)	0.346	0.312 (0.276)	0.314
Based on: Rokicki and Hewings (2011)	0.36 (0.204)		0.344 (0.329)		0.312 (0.276)	
Slovakia	0.343 (0.426)	0.344	0.27 (0.372)	0.274	0.262 (0.403)	0.264
Based on: Radvansky and Fuchs (2012)	0.343 (0.426)		0.27 (0.372)		0.26 (0.312)	
Spain	0.37 (0.235)	0.372	0.349 (0.159)	0.352	0.335 (0.234)	0.337
Switzerland	0.348	0.348	0.33	0.329	0.299	0.299

	(no change)		(0.425)		(no change)	
UK	0.384 (0.144)	0.387	0.353 (0.168)	0.356	0.336 (0.155)	0.339
based on ONS RCPL	0.384 (0.144)		0.353 (0.168)		0.336 (0.155)	
US	0.414 (0.261)	0.415	0.403 (0.312)	0.404	0.372 (0.286)	0.373
based on BEA RPP	0.413 (0.101)		0,403 (0.312)		0.372 (0.286)	

Notes: p-values are provided in parenthesis.

Source: Authors' calculation based on LIS.

Table 5 – Poverty headcount ratios and impacts of regional price levels

Poverty threshold	40% of median		50% of median		60% of median	
Country	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted
Australia	3.92 (0.357)	3.85	13.79 (0.223)	14.06	27.96 (0.15)	28.32
Austria	3.34 (0.455)	3.37	9.23 (0.283)	9.02	21.2 (no change)	21.2
Based on: ***Reale Kaukraft 2008***	3.35 (0.470)		9.14 (0.365)		21.69 (0.148)	
Canada	4.75 (0.449)	4.72	12.39 (0.361)	12.5	26.02 (0.336)	26.18
Colombia	10.48 (0.406)	10.6	19.35 (0.111)	20	34.95 (0.016)	36.1

Czech Republic	2.16 (0.384)	2.09	6.1 (0.232)	6.33	16.5 (0.438)	16.44
Based on: Kramulová et al. (2016)	2.21 (0.313)		6.05 (0.177)		16.47 (0.467)	
Denmark	2.54 (0.426)	2.53	6.4 (0.118)	6.32	19.78 (0.279)	19.72
France	3.82 (0.012)	3.39	9.99 (0.0014)	9.18	22.47 (0.000036)	21.13
Georgia	8.98 (0.347)	8.81	19.14 (0.214)	19.56	35.11 (0.0042)	36.64
Germany	2.79 (no change)	2.79	9.5 (0.347)	9.39	21.63 (0.127)	22.03

Greece	5.82 (0.471)	5.79	13.81 (0.334)	14.02	28.1 (0.402)	28.24
Hungary	4.32 (0.298)	4.61	10.14 (0.271)	10.58	22.42 (0.088)	23.56
India	9.39 (0.443)	9.42	19.41 (0.098)	19.72	35.41 (0.0526)	35.81
Ireland	4.15 (0.459)	4.2	10.28 (0.059)	9.41	26.12 (0.0051)	24.39
Italy	5.43 (0.25)	5.71	11.93 (0.049)	12.74	25.81 (0.00395)	27.38
Based on: ***Pittau et al. (2011)***	5.37 (0.201)		10.88 (0.00008)		24.92 (0.00003)	
Mexico	10.33	10.53	19.6	20.18	33.89	34.32

	(0.21)		(0.0197)		(0.084)	
Poland	3.51 (0.324)	3.56	9.65 (0.418)	9.62	22.74 (0.391)	22.69
Based on: ***Rokicki and Hewings (2011)***	3.55 (0.464)		9.65 (0.423)		22.71 (0.457)	
Slovakia	3.64 (0.331)	3.79	8.15 (0.381)	8.03	18.7 (0.475)	18.67
Based on: ***Radvansky and Fuchs (2012)***	3.73 (0.431)		8.11 (0.421)		18.57 (0.42)	
Spain	7.57 (0.398)	7.65	15.31 (0.242)	15.57	29.21 (0.301)	29.41
Switzerland	3.35	3.3	9.18	9.15	22.28	22.07

	(0.429)		(0.465)		(0.337)	
UK	3.89 (0.269)	3.79	9.81 (0.458)	9.79	24.84 (0.349)	24.93
based on ONS RCPL	3.86 (0.333)		9.76 (0.438)		24.9 (0.449)	
US	8 (0.433)	8.02	17.28 (0.442)	17.3	31.98 (0.205)	32.1
based on BEA RPP	8 (0.433)		17.37 (0.306)		31.93 (0.121)	

Notes: p-values are provided in parenthesis.

Source: Authors' calculation based on LIS.