Journal of International Economics 102 (2016) 291–309

Contents lists available at ScienceDirect

Journal of International Economics

journal homepage: www.elsevier.com/locate/jie

The price of development: The Penn–Balassa–Samuelson effect revisited☆

Fadi Hassan

Department of Economics, Trinity College Dublin, Ireland
Center for Economic Performance, LSE, United Kingdom

ARTICLE INFO

Article history:
Received 29 May 2015
Received in revised form 11 May 2016
Accepted 27 July 2016
Available online 12 August 2016

JEL classification:
F3
F4
O11

Keywords:
Penn effect
Balassa–Samuelson hypothesis
Developing countries
Real exchange rate
Structural transformation

ABSTRACT

The Penn–Balassa–Samuelson effect is the stylized fact about the positive correlation between cross-country price level and per-capita income. This paper provides evidence that the price–income relation is actually non-linear and turns negative among low income countries. The result is robust along both cross-section and panel dimensions. Additional robustness checks show that biases in PPP estimation and measurement error in low-income countries do not drive the result. Rather, the different stage of development between countries can explain this new finding. The paper shows that a model linking the price level to the process of structural transformation captures the non-monotonic pattern of the data. This provides additional understanding of real exchange rate determinants in developing countries.

1. Introduction

It is widely understood that market exchange rates do not give accurate measures of real income in different economies and that adjustment by purchasing power parity (PPP) factors is necessary for such measures. This understanding is based on an observed empirical regularity that richer countries have a higher price level than poorer countries.[1] The positive correlation between cross-country price level and per-capita income is generally regarded as a stylized fact. This result was documented for twelve developed countries in the seminal paper of Bela Balassa (1964), was confirmed for a large sample of countries as soon as data from the International Comparison Program (ICP) became available and is now renowned as the Penn–Balassa–Samuelson effect (Penn–BS).[2,3]

The paper makes an important qualification to this general understanding. Using non-parametric estimation, it provides evidence that the price–income relation is non-linear and turns negative in low-income countries, both along a cross-section and a panel dimension. Standard regression analysis in sub-samples of poor, middle-income and rich countries is consistent with this finding. The results of the paper are robust to possible sources of bias from PPP estimation and measurement error in low-income countries.

☆ I have greatly benefited from comments and discussions with Francesco Caselli, Douglas Gollin, Philip Lane, Guy Michaels, Branko Milanovic, Daniel Sturm, Silvana Tenreyro, Adrian Wood, and Alwyn Young. I am particularly indebted to Bernardo Guimaraes and Rachel Ngai for their precious guidance and support. All remaining errors are mine.

E-mail address: fhassan@tcd.ie (F. Hassan).

[1] Adjustment by PPPs is necessary as long as price levels vary across countries, even if the variation is not systematic with income.

[2] The Penn–BS effect was documented also by Barro (1991), Summers and Heston (1991), and Rogoff (1996). Samuelson (1994) stresses that the proper name for it would be *Ricardo–Viner–Harrod–Balassa–Samuelson–Penn–Bhagwati–et al. effect.*

[3] The Penn–BS effect should not be confused with the Balassa–Samuelson hypothesis. The latter provides the mainstream explanation for the former. The Balassa–Samuleson hypothesis argues that richer countries have a higher relative productivity in the tradable sector; under certain assumptions, this leads to a higher relative price of non-tradables, hence to a higher aggregate price level.

http://dx.doi.org/10.1016/j.jinteco.2016.07.009
0022-1996/

This paper argues that the non-monotonicity of the price–income relation is due to the different stages of development that characterize low- and high-income countries. We extend the standard Balassa–Samuelson model to a three sectors environment (agriculture, manufacturing and services) and trace the effects of agricultural productivity, sectoral expenditure and employment shares on the price level of low-income countries. This model captures the non-monotonic pattern of the data, in a way that the standard Balassa–Samuelson hypothesis, focused on productivity differences between tradables and non-tradables, does not. The intuition is that, when a poor country starts to develop, its productivity growth lies mainly in the agricultural sector. Since, at an early stage of development, agriculture is primarily non-traded and represents a big share of expenditure, this productivity growth reduces the relative price of agricultural goods, hence the overall price level. After a certain level of development, the role of agriculture becomes negligible and the overall price level is driven by the raise of the relative productivity of manufacturing respect to services, as in the classical Balassa–Samuelson hypothesis.

In economics, empirical regularities are rare and important. As Solow (1956) and Easterly and Levine (2001) point out, economists build models to match relevant empirical regularities and they use these models to understand economic events and give policy suggestions. The Penn–BS effect is the empirical regularity that the seminal models of Balassa (1964) and Samuelson (1964) try to reproduce. The mechanisms of these models are at the basis of our understanding of long-run real exchange rate movements, are incorporated into many new open-economy macroeconomic models and have been the initial point of reference for a vast literature on this subject.[4] The paper shows that the empirical regularity, which models in the literature are supposed to match, namely the Penn–BS effect, is not actually present in low income countries.[5]

The paper makes a significant empirical contribution by uncovering a twist to what has long been accepted as a well-established empirical regularity and offers a novel explanation of real exchange rate determinants in low income countries, based on the process of structural transformation. From a policy point of view, by showing that the price–income relation is negative in poor countries, the paper suggests that there is a “natural” depreciation of the real exchange rate along the development process. This is an important finding that central banks and governments of low-income countries should take into account as they formulate exchange rate policy. Moreover, the result of the paper suggests that current measures of real exchange rate undervaluation based on the Balassa–Samuelson hypothesis are biased for developing countries; for instance, once we account for the non-monotonic pattern of the price–income relationship, the Chinese Renminbi is 30% less undervalued than standard estimates suggest.[6] The new empirical regularity shown by the paper and its explanation can help us to better understand long-run real exchange rate movements in developing countries and lay the ground for further research on this subject.

The paper relates to the literature on PPPs and the Penn–BS effect as in Kravis et al. (1982), Heston and Summers (1992), and Feenstra et al. (2015). Our contribution is to identify the non-monotonic pattern of the price–income relation as a novel stylized fact and link this non-monotonicity to a plausible model of structural transformation.

The paper refers to the debate on PPPs and real exchange rate determinants in the long run, as in Balassa, 1964, Samuelson, 1964, Bhagwati (1984), De Gregorio et al. (1994), Rogoff (1996) and Taylor and Taylor (2004). Within this literature the papers close in spirit to our are Bergin et al. (2006) and Devereux (1999). The former shows that there is no Penn–BS effect before the 1970s; the latter presents a model of endogenous productivity growth in the distribution sector to explain real exchange rate depreciation in East Asian countries. Our paper provides a more generalized and systematic evidence of a counter Penn–BS effect and real exchange rate movements in developing countries.[7] Moreover, our explanation of this finding offers an original contribution of real exchange rate determinants in developing countries, based on structural transformation.

Finally, the paper is complementary to the literature on structural transformation and the role of agriculture as a driver of development as in Gollin et al. (2002, 2007) and Ngai and Pissarides (2007). We highlight the importance of structural transformation out of agriculture as a determinant of real exchange rates in developing countries.

The paper is structured as follows. Section 2 shows that the price–income relation is non-monotonic using both non-parametric and linear estimations. Section 3 establishes that the results are robust to measurement error, bias in the estimation of PPPs, and different databases. Section 4 argues that differences in economic structure can explain the results, derives a model that links the price level to the process of structural transformation, and analyzes the empirical prediction of the model, showing that it can capture the non-monotonicity of the data. Section 5 concludes by summarizing the main findings and discussing further research based on these results.

2. The price–income relation

In this section, we show that the price–income relation is non-monotonic. We provide evidence along a cross-section and panel dimension, through both linear and non-linear estimation. Following the literature on the Penn–BS effect, we measure income per capita in purchasing power parity (PPP) and define the price level as the ratio of PPP to the exchange rate with the US dollar.[8]

2.1. Cross-section dimension

In Fig. 1.1, we can see an example of the little attention that the literature has paid to the Penn–BS effect in developing countries. The figure illustrates the positive price–income relation reported in the review of the purchasing power parity puzzle by Rogoff (1996). Since observations with an income per capita lower than Syria are gathered in a cloud of points, it is difficult to properly disentangle the relation between price and income in poor countries.

Therefore, in Fig. 1.2, using the same data-set as in Rogoff (1996), we plot the log-values of income per capita.[9] We investigate the price–income relation using a non-parametric estimation technique

[4] The Balassa–Samuelson hypothesis hits more than 7000 entries on Google Scholar; see Rogoff (1996) and Taylor and Taylor (2004) for extended reviews and Bordo et al. (2014) and Berka et al. (2014) for the most recent applications at the time of writing.

[5] This can explain why there is not much evidence of the Balassa–Samuelson hypothesis in lower income countries as in Choudhri and Khan (2005) and Genius and Tzouvelekas (2008). Notice that they focus on the effect of relative productivity in the tradable sector on the real exchange rate (the Balassa–Samuelson hypothesis), whereas this paper focuses on the Penn effect which, to the best of our knowledge, is a novel contribution.

[6] Standard measures of undervaluation, as in Rodrik (2008), are the difference between the data and the fitted value of a linear regression of the price measure from Penn World Table on income.

[7] Notice that Feenstra et al. (2015) argue that the results of Bergin et al. (2006) are driven by interpolation issues of PPPs to past data; this critique does not apply to this paper because our main results are based on a cross-section dimension in benchmark years.

[8] We use income per capita at constant prices for the panel analysis and income at current prices for the cross-section analysis.

[9] This is Penn World Table 5.6 (prices' reference year 1985); he considers the year 1990.

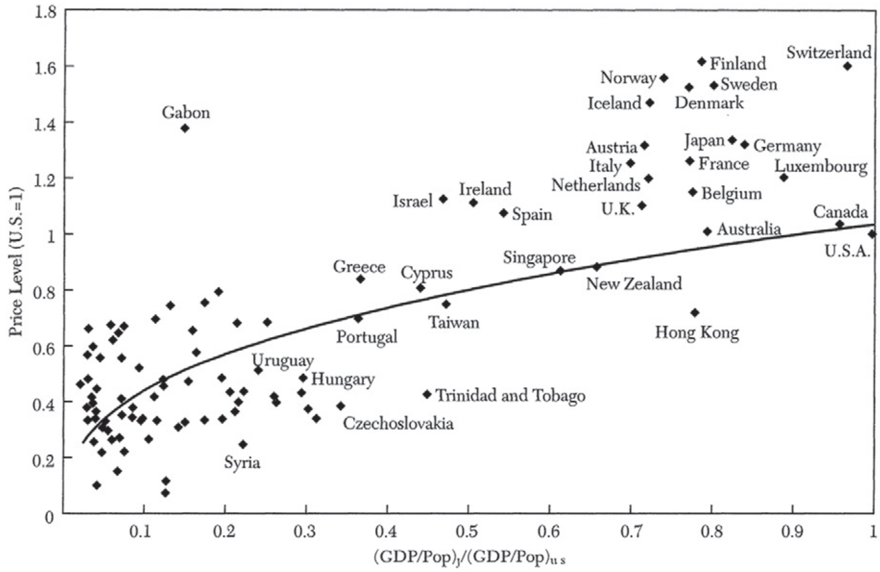

Fig. 1.1. Price level and income — Rogoff (1996).

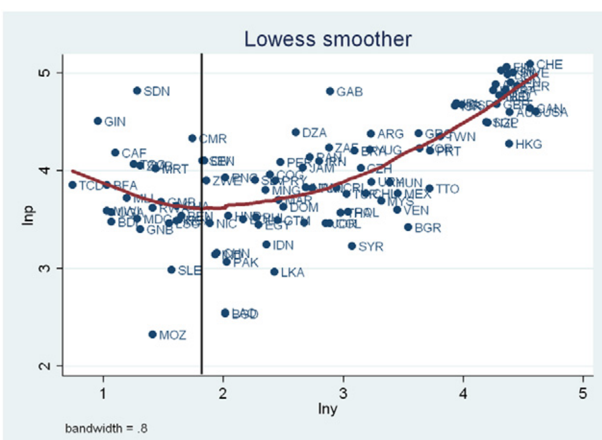

Fig. 1.2. Price level and income — Rogoff (1996); log-income & non-param. estimation.

known as LOWESS (locally weighted scatter smooth), which allows us to impose as little structure as possible on the functional form.[10] This estimation suggests that the Penn–BS effect does not hold in the poorest 25% of countries in the sample, where the relation is actually downward sloping. The minimum point of the curve corresponds to an income level of around 1350 PPP $ (1985 prices), which is equivalent to the income of Senegal in the year 1990.

In commenting the result of Fig. 1.1, Rogoff (1996) stressed that "*The relation between income and prices is quite striking over the full data set (...); it is far less impressive when one looks either at the rich countries as a group, or at developing countries as group*". In this paper we take Rogoff's point further using a non-parametric estimation that shows that the relation is actually striking when looking at rich countries as a group and negative when looking at poor countries as a group. According to our knowledge, the non-monotonicity of the price–income relation has not been previously documented in the literature.

Next, we extend the analysis to the database of International Comparison Program (ICP) for the year 2011.[11] This dataset makes use of the latest available round of price collection coordinated by the ICP and the World Bank. As Deaton and Aten (2014) argue, this the most reliable round of data collection that has been implemented so far. Moreover, using only the benchmark countries and year minimizes the source of measurement error.[12]

In Fig. 2.1 we can confirm the strong positive relation predicted by the Penn–BS effect by running a standard linear estimation of price on income: the OLS coefficient is 0.24 with a robust t-statistic of 9.79.[13] However, once we allow for non-linearities, the Penn–BS effect breaks down for low income countries. Fig. 2.2 shows the results of running a LOWESS estimation between price and income,

[10] The LOWESS estimation works as follows: Consider an independent variable x_n and a dependent variable y_n. For each observation y_n the LOWESS estimation technique runs a regression of x_n using few data points around x_n. The regression is weighted so that the central point $(x_n; y_n)$ receives the highest weight and points further away get less weight. The fitted value of this regression evaluated at y_n represents the smoothed value y_n^S which is used to construct the non-parametric curve that links y and x. The procedure is repeated for each observation $(x_n; y_n)$. The number of regressions is equal to the number of observations, and the smoothed curve is the set of all $(x_n; y_n^S)$.

[11] We use only benchmark countries. We exclude countries with less than 1.3 million people in the year 2011 and oil countries, which tend to have very low prices given their income level; including these countries would reinforce our findings. The list of the countries included can be found in Appendix A.

[12] All the results presented in the paper hold also using the new generation of Penn World Tables introduced by Feenstra et al. (2015), as well as for older Penn World Table versions. In this part of the paper we prefer to use the ICP data because it relies on the latest round of price collection and not on the more controversial 2005 round. Using PWT 8.0 delivers the same results. PWT 8.0 is the latest available that includes data on prices from the expenditure side; the more updated PWT 8.1 provides data of prices only from the output side; our results are robust to this version too (we discuss this point more in details later).

[13] We run an OLS regression, with robust standard errors, of the log of the price level of GDP and the log of GDP per capita in PPPs at current prices. We use the expenditure-side of real GDP and price because of comparability with previous studies; these are the variables that the literature on the Penn–BS effect traditionally focused on. The results of the paper are robust to real GDP and prices measured from the output-side introduced by PWT 8.1. All the regression results of the paper are virtually unchanged, but the cross-section result of the LOWESS estimation are a bit weaker. Nevertheless, we believe that this is driven by a sample issue, as data of output-prices are not available for about 20% of countries in the low-income part of the distribution, which is the main object of our paper.

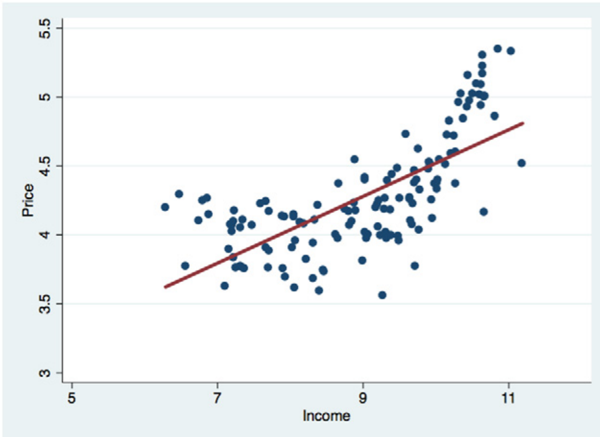

Fig. 2.1. Price level and income, ICP 2011: linear estimation.

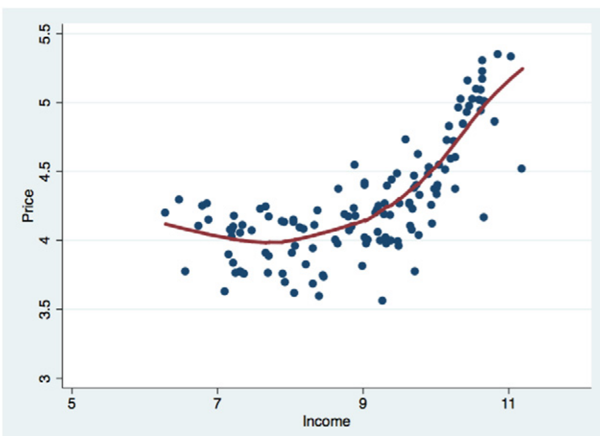

Fig. 2.2. Price level and income, ICP 2011: non-parametric estimation.

imposing little restriction on the functional form.[14] We can see that the expected upward sloping relation holds only for middle- and high-income countries. The relation is downward sloping for low-income countries; this involves almost 20% of the countries in the sample. The turning point is at 2130 PPP $ per-capita (2011 prices), which is equivalent to the income of Lesotho in the year 2011.

Fig. 3 reports 95% confidence bands of the LOWESS estimation derived from the standard errors of the smoothed values. The confidence interval confirms the non-monotonic pattern of the data. The pseudo-R^2 of the non-parametric estimation is 0.72, which is higher than the 0.50 R^2 of the linear model. The *F*-test comparing the non-parametric model to the linear one rejects the null hypothesis that the non-linear model does not provide a statistically significant better fit.

Standard cross-country OLS regression supports the finding of the non-parametric estimation. Table 1 shows that a quadratic specification of the price–income relations confirms the non-monotonic pattern. Both *Income* and *Income*^{2} are statistically significant. The coefficient associated to the linear term is negative and the quadratic one is positive, indicating a convex relation. The marginal effect of income on price turns positive around 2643 PPP $ per-capita (2011 prices), which is equivalent to the income of Cote d'Ivoire in the year 2011. The turning point from the quadratic specification is at a higher level of income than from the previous non-parametric estimation. The countries on the downward sloping path are listed in Table 3; we can notice that these are mainly African and some Asian (no Latin-American).[15]

Given the functional form $Price_i = \alpha + \beta\ Income_i + \gamma\ Income_i^2 + \epsilon_i$, Lind and Mehlum (2010) show that in order to test for the presence of a *U*-relation, it is necessary to formulate the following joint null hypothesis:

$$H_0 : \beta + 2\ \gamma\ \text{Income}_{\min} \geq 0\ \textit{and/or}\ \beta + 2\ \gamma\ \text{Income}_{\max} \leq 0 \tag{1}$$

against the alternative:

$$H_1 : \beta + 2\ \gamma\ \text{Income}_{\min} < 0\ \textit{and}\ \beta + 2\ \gamma\ \text{Income}_{\max} > 0 \tag{2}$$

Lind and Mehlum (2010) build a test for the joint hypotheses using Sasabuchi (1980) likelihood ratio approach. Table 2 shows that the marginal effect of income on price is negative and statistically significant at $\text{Income}_{\min}$ and positive and statistically significant at $\text{Income}_{\max}$. The last line of the table shows that the SLM test rejects H_0 in favor of the alternative and thus indicates that the result is consistent with the presence of a non-monotonic relation between price and income.

Finally, in Table 4 we divide the sample by income groups according to the standard World Bank classification. The price–income relation is negative, sizable, and significant for low-income countries; it is positive, but with a coefficient close to zero for the middle-income group; and it turns positive, sizable, and strongly significant for high-income countries. Also the results of these regressions are consistent with the non-monotonicity of the price–income relation.[16]

Therefore, looking at the results from the various approaches used to analyze the data, this section provides overall evidence of a non-monotonic price–income relation.

2.2. Panel dimension

In this section, we analyze the price–income relation in a panel dimension. We turn to the PWT 8.0 as the main database, because as Feenstra et al. (2015) argue, it provides a better methodology to compare real income and prices over time.[17] The ICP collects data prices over time in benchmark years. Then, the PWT used to estimate prices for other years by rescaling according to the inflation rate differential with the US. However, the new version of the Penn World

[14] LOWESS estimation requires that the bandwidth of observations included in the regression of each point be chosen. Specifying a large bandwidth provides a smoother estimation, but increases the risk of bias by including observations from other parts of the density. A small bandwidth can better identify genuine features of the underlying density, but increases the variance of the estimation. In the paper I use the default STATA bandwidth of 0.8, which is a conservative choice and provides a lower-bound of the non-monotonic pattern of the data. The pseudo-R^2 of the LOWESS estimation is maximized at a bandwidth of 0.4, which delivers a stronger non-monotonicity at the cost of higher variance. Using a Kernel estimation rather than a LOWESS conveys very similar results to the ones presented in the paper.

[15] This does not imply that our finding is related to African countries only, as we discuss in the panel analysis the result is a more general feature of the data.

[16] The observations per-income group are 28, 66, and 39 respectively. The World Bank threshold is 1005 US$ (2011) for low-income countries and 12,276 US$ (2011) for high-income countries.

[17] We use PWT 8.0 rather than 8.1 because it is the latest version to provide data also on prices from the expenditure side of GDP, which are, as we previously said, the traditional focus of the Penn–BS effect. Results hold if we use output-prices from PWT 8.1. Results also hold if we use World Bank data rather than PWTs. The World Bank data have the advantage to rely on the 2011 ICP round; however this is more useful for a cross-sectional dimension and it becomes less relevant if we want to extend the analysis over time. In this case the methodology provided by the new generation of PWTs, is a better reference, because it makes use of past benchmark years.

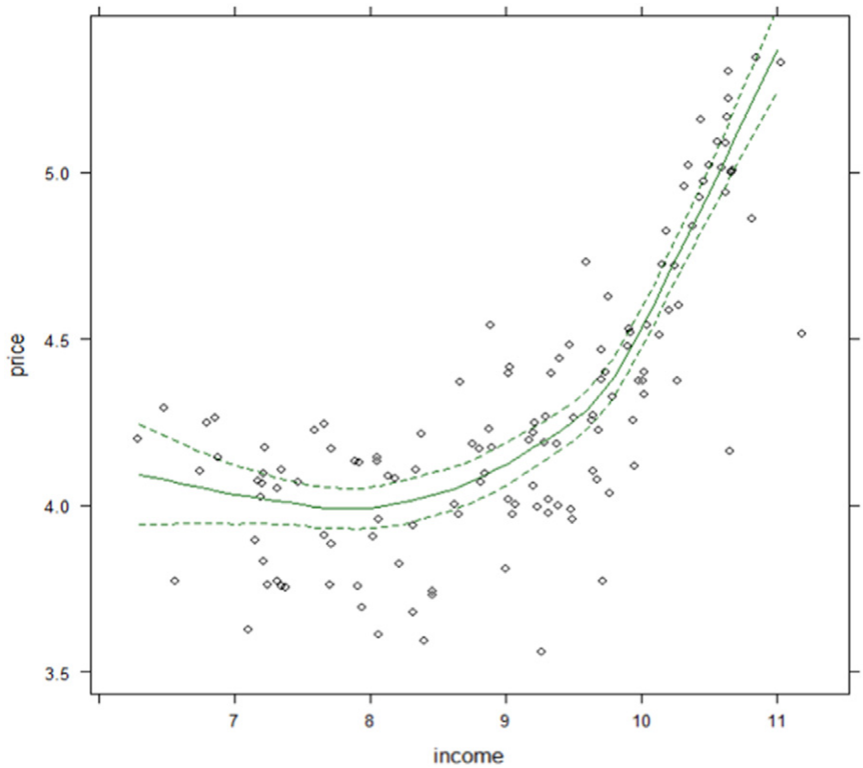

Fig. 3. Price and income, ICP 2011: non-parametric estimation, 95% confidence bands.

Table 1
Cross-country OLS regression: linear and quadratic specifications, year 2005.

Dependent var: ln *price*	(1)	(2)
ln*income*	0.24***	−2.69***
	(9.79)	(−6.50)
ln*income*^{2}		0.13***
		(7.09)
N. obs.	133	133
R^2	0.50	0.70
Turning point		2643 PPP $

*** Significant at the 1% level; robust t-statistics in parenthesis.

Table 2
Tests for a U-shape.

Dependent var: ln *price*	
Slope at $Income_{min}$	−0.42***
	(−4.85)
Slope at $Income_{max}$	0.86***
	(8.87)
SLM test for U-shape	4.85
p-Value	0.00

*** Significant at the 1% level; robust t-statistics in parenthesis.

Tables makes use of historical ICP benchmarks to extrapolate the time series of prices and real incomes.[18] Moreover, as Feenstra et al. (2015) stress the method of aggregation of goods' prices that the PWT use to compute PPPs allow to extrapolate or interpolate prices outside the benchmark years using price indexes of each country from national accounts.[19] For these reasons, in the panel analysis we rely on the PWT as the main database.

Despite the methodological innovations of the last version of the PWTs, there is a higher degree of uncertainty about PPP outside benchmark years. Nevertheless, panel regressions of price on income are commonly used to build measures of real exchange rate over/undervaluation. Thus, it is relevant to assess if the non-monotonicity of the price–income relation holds also along a panel dimension.

If we extend the analysis to a panel of countries between 1950–2009, standard linear estimation of price on income confirms the positive relation predicted by the Penn–BS effect: the OLS coefficient is 0.15 with a t-statistic of 32.7 (Fig. 4.1).[20] However, non-parametric estimation shows that the price–income relation is also non-monotonic along a panel dimension. The Penn–BS effect holds for middle- and high-income countries, but in low-income countries the relation is negative (Fig. 4.2).

Fig. 4.3 reports the fitted value of the LOWESS estimation. The turning point is at 1421 PPP $ per-capita (2005 prices), which corresponds to the income of Senegal in the year 2000. The downward sloping arm of the curve includes 27% of the total observations, and 45% of the countries in the sample. The countries on the downward sloping arm and their frequencies are reported in Table 5. We can see

[18] Nevertheless, it is important to keep in mind that many countries, especially developing ones like China or India, did not participate in all the benchmark collections; this makes the computation of prices and real incomes in non-benchmark years more uncertain.

[19] We provide more details on this in the next section.

[20] This is for a sample of 126 benchmark countries from 1950 to 2009 using PWT 8.0. Countries with less than 1.3 million people in the year 2010 were dropped. We also drop Zimbabwe and Tajikistan which are clear outliers, adding them would reinforce our results. We run an OLS regression of the log of the price level of GDP (variable $pl_g dpe$) and the log of GDP per capita in PPPs at constant chained prices (*rgdpe*/*pop*).

Table 3
Countries before the minimum, cross-section dimension.

Benin	Madagascar
Burkina Faso	Malawi
Central African Republic	Mali
Chad	Mozambique
Congo, Dem Rep.	Nepal
Ethiopia	Niger
Gambia	Rwanda
Guinea	Senegal
Guinea-Bissau	Sierra Leone
Haiti	Tajikistan
Kenya	Tanzania
Lesotho	Togo
Lesotho	Uganda
Liberia	Zimbabwe

Table 4
Cross-country OLS regression by income groups, year 2011.

Dependent var: ln*price*	ln *income*
Low income	−0.29***
	(−4.49)
Middle income	0.08**
	(2.28)
High income	0.52***
	(3.21)
Full sample	0.24***
	(9.79)

** Significant at the 5% level; robust t-statistics in parenthesis.
*** Significant at the 1% level; robust t-statistics in parenthesis.

that some of the countries are persistently on the downward-sloping arm (i.e. Ethiopia and Tanzania); others moved along the curve (i.e. China and Vietnam). In this case we can see that the non-monotonic pattern of the data does not involve only African countries, but it is more extended.

Standard panel-data analysis, Table 6, confirms the result of the non-parametric estimation. I take 5-year averages of price and income between 1950–2009. I show that for developing countries the relation between price and income is negative and significant with and without country fixed-effects. I do this by running a regression for the full sample, and then for developing countries only.[21] This result comes despite a broad definition of developing countries and a linear restriction on the price–income relation.[22]

Tables 7 and 8 extend the panel analysis and show that a quadratic regression supports a J/U-shaped price–income relation across all specifications both for the full sample and for developing countries only. It is important to stress that the non-monotonic result holds also for the full sample using country fixed-effects. This implies that even when we exploit only the within-country variation in the complete sample, the price–income relation turns to be non-monotonic. This reinforces the point that a non-monotonic Penn–BS effect is a general pattern in the data; as it holds in a cross-section, it holds in a panel, and it holds also using within-country variation only.

[21] I define developing countries those below the World Bank's threshold of high-income countries; a stricter definition of developing countries reinforces my result. Notice that in the full sample with country fixed effects the coefficient is not significantly different from zero.

[22] As an additional robustness, we run the panel specifications excluding sub-Saharan countries and the results hold.

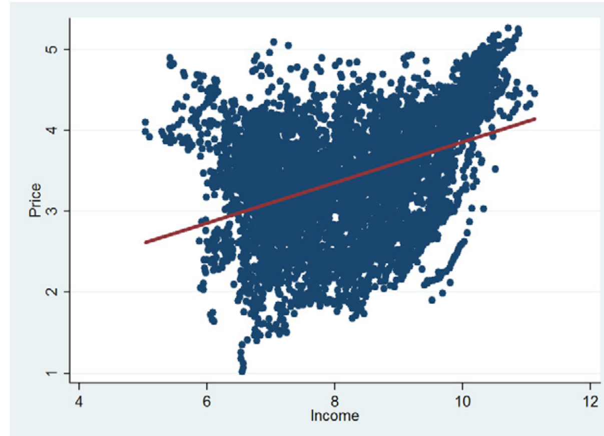

Fig. 4.1. Prices and income 1950–2011: OLS estimation.

3. Robustness checks

The data used to estimate the price–income relationship are PPPs, exchange rates, and GDP per-capita.[23] Most of the robustness analysis focuses on PPPs by looking at measurement error in prices and at bias in the construction of PPPs, which are arguably the main source of concern. Moreover, given that in developing countries official exchange rates can be different from black market rates, we control for this possible source of bias. Finally, we show that results are robust to different versions of the Penn World Tables.[24]

3.1. Classical measurement error

Chen et al. (2007) analyze the bias of the OLS estimation of price on income when there is measurement error in prices. In this case the independent variable becomes correlated with the error term, so that the standard assumptions for a consistent and unbiased least square estimator break down.[25] Chen et al. (2007) conclude that the OLS estimate will be biased downwards and can become negative if the variance of the measurement error is sufficiently high. In fact, they show that[26]

$$\text{plim} \quad \hat{\beta} = \frac{\beta^{true} - \frac{\sigma_\eta^2}{\sigma_{y*}^2}}{1 + \frac{\sigma_\eta^2}{\sigma_{y*}^2}} \tag{3}$$

where σ_η^2 is the variance of measurement error and σ_{y*}^2 is the variance of the true real income per-capita. From this expression we can see that as the variance of the measurement error σ_η^2 increases, the estimated $\hat{\beta}$ can turn negative.

[23] We remind the reader that $p = \frac{PPP}{XRAT}$ and $y = \frac{GDP}{PPP}$.

[24] In general we may also have measurement error in GDP data; however, these are of lower concern. Gollin et al. (2014) analyze the definitions and measurement approaches used in the construction of national accounts data in poor countries. They conclude that these aggregate data are robust to problems associated with informality or household production and that there is no reason to believe that they are intrinsically flawed. Therefore, we do not focus the robustness discussion on estimates of GDP per-capita.

[25] The econometric specification of the price–income relation is such that $p_i^* = \alpha + \beta y_i^* + \epsilon_i$, where variables are expressed in logs and p_i^* is the true price level without measurement error and $y_i^* = Y_i - p_i^*$ is the true real income per-capita. Consider the case where the measured price level p_i contains an error such that $p_i = p_i^* + \eta_i$, where η_i has mean zero and is normally distributed; then the regressor and the error term become correlated.

[26] Assuming that the measurement error is uncorrelated with the true dependent and independent variables as well as with the equation error, Eq. (3) follows.

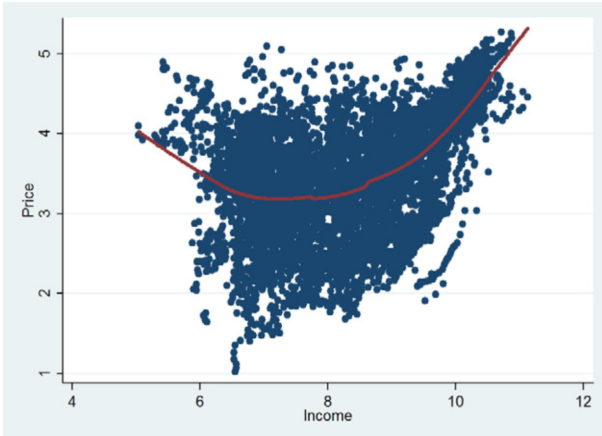

Fig. 4.2. Prices and income 1950–2011: non-parametric estimation.

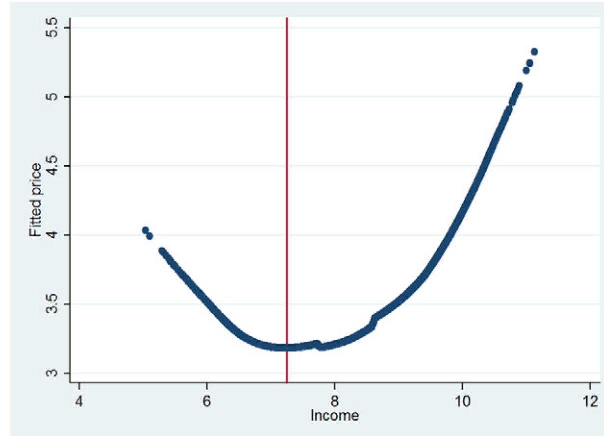

Fig. 4.3. Prices and income 1950–2011: non-parametric estimation, fitted values.

If we look at the group of low-income countries in Table 4, the OLS estimate of price on income is −0.29 (Table 4). What is the level of measurement error's variance needed to drive this result? Assuming that measurement error is correlated to the level of income but not to the level of price, we can rewrite Eq. (3) as[27]

$$\text{plim}\,\hat{\beta} = \frac{\beta^{true} - \frac{\sigma_\eta^2}{\sigma_{y^*}^2}}{1 + \frac{\sigma_\eta^2}{\sigma_{y^*}^2}} = \frac{\beta^{true} - \frac{\sigma_\eta^2}{\sigma_Y^2+\sigma_p^2+\sigma_\eta^2-2\sigma_{Yp}}}{1 + \frac{\sigma_\eta^2}{\sigma_Y^2+\sigma_p^2+\sigma_\eta^2-2\sigma_{Yp}}} \tag{4}$$

In the sub-sample of countries where the price–income relation is negative, we have $\sigma_Y^2 = 1.48$, $\sigma_p^2 = 0.16$, $\sigma_{Yp} = 0.35$ (remember that all the variables are expressed in logs).

The variance of measurement error that would lead to the negative estimation of −0.29 depends on the value of β^{true}. Let's suppose that β^{true} is equal to the OLS estimation over the full sample (0.24). In this case, in order to get $\hat{\beta} = -0.24$, we would need $\sigma_\eta^2 = 0.68$: the measurement error on prices should have a variance more than four times higher than the variance of observed prices over the full sample. If instead, we assume that in low income countries β^{true} is zero, we would need $\sigma_\eta^2 = 0.34$: hence in this case the variance of the measurement error on prices in this sub-sample of countries should be more than double than the variance of the observed prices.

Therefore, even if measurement error could potentially drive the results of the paper, an improbable high variance of the measurement error itself is required to obtain the negative price–income relation presented in the paper.

3.2. Purchasing power parities bias

The process of computing PPPs is subject to intrinsic fragilities, making the comparison of real income and prices across countries a difficult exercise. The underlying source of data is collected by the International Comparison Program (ICP). The ICP coordinates the collection of prices of about 1500–2000 items across a large set of countries. In the latest round the ICP selected a list of 618 global core items which were representative enough to be priced in each country. The regional offices then provided the final list of items to be priced in each region trying to incorporate as much product as possible from the global core list.[28] The ICP then aggregates items' prices into 155 goods called basic headings. A basic heading is the most disaggregated level at which expenditure data are available from national accounts. The ICP collects quotes for different items within each basic heading and then computes a unique price through a country-product dummy weighted regression (CPD-W), where each item is assigned a 3:1 weight according to its representativity.[29] Once the prices of all 155 basic headings are obtained for each country, these are used to compute PPPs and to compare real income across countries.

All this process generates various potential sources of bias in the estimation of PPPs. The main ones are as follows: the method of aggregation of basic headings' prices into PPPs index; quality matching; and items' representativity (Deaton and Heston, 2010; ICP, 2007). The direction of the PPP bias can have a key influence on our results. Let's suppose that the true price–income relationship is flat. Fig. 5 shows that if in low-income countries PPPs tend to be overestimated a negative price–income relationship would arise because of that bias; however, if PPPs are underestimated, a Penn–BS effect would emerge.[30]

The literature has so far established that PPPs in low-income countries tend to be underestimated (Almas, 2012; Deaton and Heston, 2010; Hill, 2004; Neary, 2004; Nuxoll, 1994). This implies that the negative price–income relationship in poor countries shown in the paper is likely to be a lower bound of the true one.

The method of aggregation of basic headings' prices into the PPP index differs between the ICP/World Bank and the PWTs. The ICP focuses on a regional approach. They firstly compute within-region PPPs and then they obtain between-region PPPs using the global core products prices while maintaining the within-region ranking of

[27] From the specification of Chen et al. (2007), we have that $y_i^* = Y_i - p_i + \eta_i$; keeping the same independence assumptions of their paper, such that $Cov(Y, \eta) = 0$, which is plausible for the subsample of countries we are looking at, Eq. (4) follows.

[28] The list and description of items for a particular cluster of products are elaborated at a regional level with the collaboration of national statistical offices. The list provides a standardized product description (SPD). An example of SPD is as follows: "Men's shirt, well known brands, 100% cotton, light material, classic styling, uniform color, short sleeves, classic collar, buttons fastener" (ICP, 2007). The ICP regions are Africa, Asia-Pacific, CIS, South America, OECD-Eurostat, and Western Asia.

[29] For instance, for the basic heading *rice*, the ICP collects quotes for six different kinds of rice, including long-grained, short-grained, and brown rice. The country-product-dummy weighted regression is then used to obtain a price of the basic heading rice. See ICP (2015) for an explanation of the aggregation procedure.

[30] The underlying assumption of Fig. 5 is that PPP bias affects mostly poorer countries.

Table 5
Countries on the downward sloping arm of the LOWESS estimation, panel dimension.

Country	Frequency	Country	Frequency	Country	Frequency
Bangladesh	38	Guinea	24	Nigeria	14
Benin	53	Guinea-Bissau	52	Pakistan	20
Bolivia	7	India	45	Paraguay	5
Bosnia Herzegovina	4	Indonesia	15	Philippines	3
Botswana	16	Iraq	1	Romania	2
Brazil	2	Kenya	24	Rwanda	41
Burkina Faso	53	Korea	14	Senegal	4
Cambodia	35	Laos	24	Sierra Leone	48
Cameroon	15	Lesotho	51	Sudan	33
Central African Rep.	52	Liberia	33	Syria	16
Chad	44	Madagascar	52	Taiwan	2
China	30	Malawi	58	Tanzania	50
Congo, Dem. Rep.	62	Mali	48	Thailand	17
Congo, Republic of	20	Mauritania	26	Togo	52
Cote d'Ivoire	2	Mongolia	13	Tunisia	1
Egypt	34	Morocco	11	Uganda	46
Ethiopia	62	Mozambique	52	Vietnam	11
Gambia	52	Nepal	52	Yemen	15
Ghana	13	Niger	52	Zambia	20

countries.[31] Whereas, the PWT firstly aggregates goods in three different categories (consumption, investment, and government) using a GEKS methodology and then aggregate prices of these categories into the final PPP index for GDP using a Geary–Khamis method (GK); this is done across all countries in a single step.

There are different advantages and disadvantages between these methods. As far as our discussion on PPP bias is concerned, the main issue is that the GK method tend to understate PPPs in poor countries; so, without this bias, our results would be reinforced. GK compares domestic prices with world prices. The world price of a good is defined as a weighted average of its price in all countries and the weights are given by a country's share in the global consumption of that good. Hence, countries with a larger physical volume of consumption get a greater weight in the construction of world prices. This implies that the vector of international prices used as a reference is closer to the price of rich rather than poor countries.[32] This generates a Gershenkron effect for low income countries according to which PPP is lower the more the price of a country differs from the price of reference (Gershenkron, 1947; Nuxoll, 1994). This effect stems from the substitution bias that characterizes indexes with a single reference price vector as in the GK method. This arises because these type of indexes do not account for utility maximizing agents switching towards cheaper goods as relative prices change (Hill, 2000).[33]

The method of aggregation is not the only source of bias of PPPs. Quality matching is also a problem because the estimation of PPPs makes use of a set of homogeneous goods. As Deaton and Heston (2010) stress, one of the most criticized issues of ICP rounds is that lower quality goods and services in poor countries are often matched to higher quality items in rich countries. Quality mismatch leads to an underestimation of the price level in poor countries; hence also this source of bias reinforces the results of the paper.

Finally, the representativity of the items whose prices are collected is also a potential source of bias. This relates both to the aggregation of items into a basic heading and to the urban bias in collecting prices. If an item within the basic heading is representative in some countries but not in others, PPPs may be estimated incorrectly.[34] This is a common problem in ICP rounds. However, in order to mitigate this issue, in the 2011 round items were weighted according to their representativity in the basic headings' aggregation process.[35]

There is much debate about the impact of representativity on the 2005 round. Diewert (2008) argues that if non-representative prices are well-distributed across all countries in a region, they may not cause serious distortions. Moreover, Deaton (2010) computes a Tornqvist index to measure how much different goods moves the overall PPP-index in Africa and Asia.[36] He concludes that there is no evidence to support the idea that prices in Africa or in the Asia-Pacific region are systematically overstated by representativity. Nevertheless, once comparing the 2011 and 2005 round Deaton and Aten (2014) and Inklaar and Rao (2014) show that representativity issues, especially related to the so called ring-approach for linking regions, overstated PPPs in the 2005 round. Following, the findings of Inklaar and Rao (2014) the PWT 8.1 adjusts PPPs estimates accounting for potential representativity bias and our results are robust to this.[37] Therefore, given that our findings hold for the 2011 round and the adjusted 2005 round, we conclude that representativity bias is unlikely to drive our results.

Feenstra et al. (2013) show that in the 2005 round the price level in China was overstated because of a urban bias in the data collection.[38] In order to account for this bias the PWT introduces

[31] The ICP computes within-region PPPs using the Gini–Elteto–Koves–Szulc index (GEKS), which basically takes a geometric mean over all the possible Fisher indexes of all countries. Then, it computes a set of five regional prices for the global core products provided by all countries. These prices are used to compute between-region basic headings' PPPs linking each region to a base region. Finally, the within-region basic heading PPPs is multiplied by the between-region PPPs so that it is converted into a global PPP, where the relative ranking between economies in the same region remains. See ICP (2015) for further details.

[32] Nuxoll (1994) shows that international prices are closest to that of a moderately prosperous country like Hungary.

[33] Neary (2004) shows that the GK method of aggregation is exact if preferences are Leontief; in this case goods are perfect complements and the substitution bias does not arise.

[34] See for example the wheat versus teff example in Deaton and Heston (2010).

[35] Something similar was tried also in the 2005 round, but it did not work systematically across all regions. The Latin American region tried to overcome this issue in the 2005 round by using an extended CPD method, adding a representativity dummy. The OECD/Eurostat and CIS regions used an EKS method based on Javon indexes of representative products between countries; see Rao (2004) and ICP (2007) for a brief description of this method.

[36] He estimates a pairwise Tornqvist index for the ring African countries vs. the UK and at regional level for Africa and Asia-Pacific vs. OECD/Eurostat.

[37] See Section 2 for a brief discussion of our results with PWT 8.1.

[38] However, there is no clear evidence of price overestimation for other countries due to the urban bias. Actually, Atkin and Donaldson (2012) show that the price of detailed products in Ethiopia and Nigeria are on average 5–12% higher in rural areas.

Table 6
Panel evidence of price and real income, 1950–2009 (5-year average).

Dependent var: ln *price*	Full sample		Developing countries	
	(1)	(2)	(3)	(4)
ln *income*	0.08***	0.002	−0.11***	−0.18***
	(2.38)	(0.04)	(−2.51)	(−2.79)
Country, fe	No	Yes	No	Yes
Time dummies	Yes	Yes	Yes	Yes
No. of countries	126	126	94	94
Avg obs per country	9.7	9.7	9.4	9.4

*** Significant at the 1% level; robust t- and z-statistics in parenthesis.

Table 7
Panel evidence of non-linear price and real income relation, 1950–2009 (5-year average).

Dependent var: ln *price*	Full sample		Developing countries	
	(1)	(2)	(3)	(4)
ln *income*	−2.04***	−2.03***	−1.4***	−1.28**
	(−7.25)	(−6.42)	(−2.83)	(−2.41)
ln $income^2$	0.13***	0.12***	0.08***	0.07**
	(7.83)	(6.79)	(2.70)	(2.15)
Country, fe	No	Yes	No	Yes
Time dummies	Yes	Yes	Yes	Yes
No. of countries	126	126	94	94
Avg obs per country	9.7	9.7	9.4	9.4
Turning point, PPP $ (2005)	2749	3608	4208	6799

** Significant at the 5% level; robust t- and z-statistics in parenthesis.
*** Significant at the 1% level; robust t- and z-statistics in parenthesis.

a uniform reduction of 20% to the ICP prices. Our results account for this downward revision. Urban bias is less of a concern for the 2011 round, as it ensures adequate coverage of rural outlets in large countries. Therefore, urban bias is not be leading the results of the paper.

To summarize, the method of aggregation and quality matching tend to bias downwards the estimation of PPPs in low-income countries compared the “true” values. Moreover, the latest ICP round and PWT 8.1 account for products' representativity bias. Therefore, we conclude that the non-monotonicity shown in Section 2 is unlikely to be driven by measurement issues and it is more likely to be a lower-bound.

3.3. Previous versions of the Penn World Tables and black market exchange rates

The analysis of the paper makes use of data from ICP and from the new generation of Penn World Tables. The former relies on the 2011 ICP round and the latter on the 2005 round. It is comforting that results hold across different rounds and different adjustments. Moreover, they hold also for previous versions of the PWTs.

In Fig. 6 we run a series of cross-section LOWESS estimations of the price–income relation for benchmark years and benchmark countries of subsequent versions of the PWT.[39] The non-monotonicity of the price–income relation is confirmed also for these older versions of the PWT.[40] Notice that Fig. 6 provides only a LOWESS estimation with a conservative bandwidth; regression analysis with a non-linear term and by group of income as in Section 2 confirms the non-monotonic results. Finally, it is interesting to observe that the relative income of the turning point

Table 8
Tests for a U-shape.

Dependent var: ln *price*	Full sample		Developing countries	
	(1)	(2)	(3)	(4)
Slope at $Income_{min}$	−0.65***	−0.70***	−0.49***	−0.50***
	(−6.16)	(−5.57)	(−3.04)	(−2.85)
Slope at $Income_{max}$	0.78***	0.69***	0.43**	0.31*
	(9.26)	(6.96)	(2.27)	(1.40)
SLM test for U-shape	6.16	5.57	2.28	1.40
p-Value	0.00	0.00	0.01	0.08

* Significant at the 10% level; robust t-statistics in parenthesis.
** Significant at the 5% level; robust t-statistics in parenthesis.
*** Significant at the 1% level; robust t-statistics in parenthesis.

of the relation decreases over time, so we observe an increasing Penn–Balassa–Samuelson effect as stressed by Bergin et al. (2006).

Another potential issue to account for is that the PWTs use official exchange rates to compute the price level, but in developing countries the official rates can greatly differ from the one actually used in daily transactions, above all in the early years of our sample. Nevertheless, this issue does not undermine the finding of the paper. As Reinhart and Rogoff (2004) argue, multiple exchange rate arrangements decreased greatly over time and apply mainly until the 1980s, while the non-monotonicity of the price–income relation shown in the paper takes the year 2005 as a benchmark. However, I have run a non-parametric estimation of price on income using black market exchange rates for the year 1996 and the non-monotonicity of the relation is confirmed also in this case.[41]

This section has shown that the results of the paper are robust to classical measurement error, bias in PPPs estimation, that they hold for different versions of the PWTs and are not affected by using black market exchange rates. All this provides evidence that the non-monotonicity of the price–income relation is not a spurious result, but a hitherto-undocumented economic fact.

4. Theoretical explanation

The most accepted explanation of the Penn–BS effect is the Balassa–Samuelson (BS) hypothesis. This explanation focuses on productivity differentials between the tradable and the non-tradable sector. Assuming free labor mobility across sectors and that the law of one price holds for tradables, the BS hypothesis shows that countries with higher relative productivity in the tradable sector have a higher price level. Since richer countries tend to have higher relative productivity in the tradable sector, the price level should then raise with per-capita income.[42]

This classical explanation cannot capture the non-monotonicity of the price–income relation highlighted in the paper. We argue that we need a modified BS framework that accounts for the relevance of the agricultural sector in poor countries and for the fact that low-income and high-income countries are at different stages of their process of structural transformation, defined as the reallocation of economic activity across agriculture, manufacturing and services.

[39] I use PWT 5.6 for 1985, PWT 6.1 for 1996, and PWT 7 for 2005.
[40] The non-monotonicity holds also for the panel dimension; results available upon request

[41] I choose the year 1996 because this is the oldest benchmark year for which both black market rates and raw PPPs are available. Results available upon request. Data on black market rates are taken from Reinhart and Rogoff (2004). Prices are computed dividing PPPs from PWT 6.1 by the black market exchange rates.
[42] Devereux (1999) shows that a counter Penn–BS effect can arise if there is higher productivity growth in the non-tradable sector, due to, for instance, improvements in the distribution of the service sector. Higher productivity in the non-tradable sector and a reclassification of the non-tradable sector are keys in this paper.

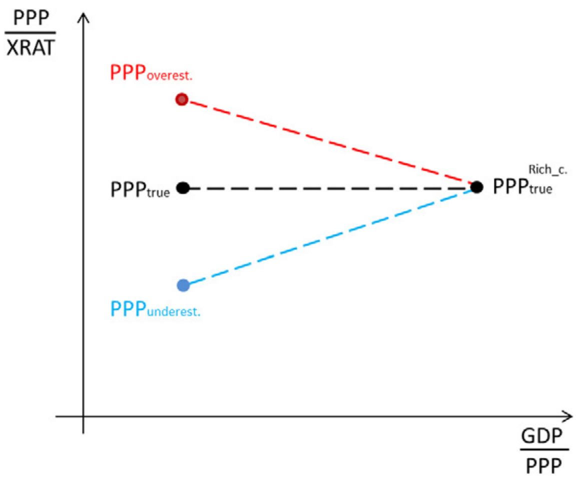

Fig. 5. The effect of PPP bias.

4.1. Beyond the Balassa–Samuelson hypothesis

Using disaggregated data kindly provided by the International Comparison Program at the World Bank, we can compute sectoral PPPs and price levels.[43] Table 9 shows that there is a different structure of relative prices by level of development.[44] Perhaps contrary to conventional wisdom, the relative price of agriculture in terms of both services and manufacturing turns to be higher in low-income than in rich countries.[45] Moreover, the average price level of services and manufacturing increases by income group, but the price level of agriculture decreases between the bottom and the intermediate group. Non-parametric estimations of sectoral prices on income confirm this pattern: Figs. 7.1–7.3 show that the price dynamics of the agricultural sector accounts for most of the non-monotonicity of the overall price–income relation.[46]

Moreover, in Table 9 we can see that countries in the bottom income group have a remarkably different structure in terms of sectoral valued added, expenditure, and employment shares. The most significant differences refer to the agricultural sector: the first group of countries, where the price–income relation is negative, have a 10 times higher valued-added share in agriculture, a five times higher expenditure share and a nine times higher employment share than the countries in the top group of income. This clearly reflects the stage of development that characterizes these countries, and it is consistent with the facts of structural transformation, as summarized by Herrendorf et al. (2014).

Finally, Fig. 8 provides a preliminary check about the role of structural transformation on the non-monotonicity of the price–income relation. It shows that there is a non-monotonic pattern between the price level and both expenditure and employment shares in agriculture, which are two key proxies for the stage of development at which countries are.

All this hints to the fact that structural transformation and agriculture play a key role to explain the non-monotonic pattern of the price–income relation. The mechanism that can link structural transformation to the non-monotonic price income relation is the following: As a poor country develops, productivity in agriculture

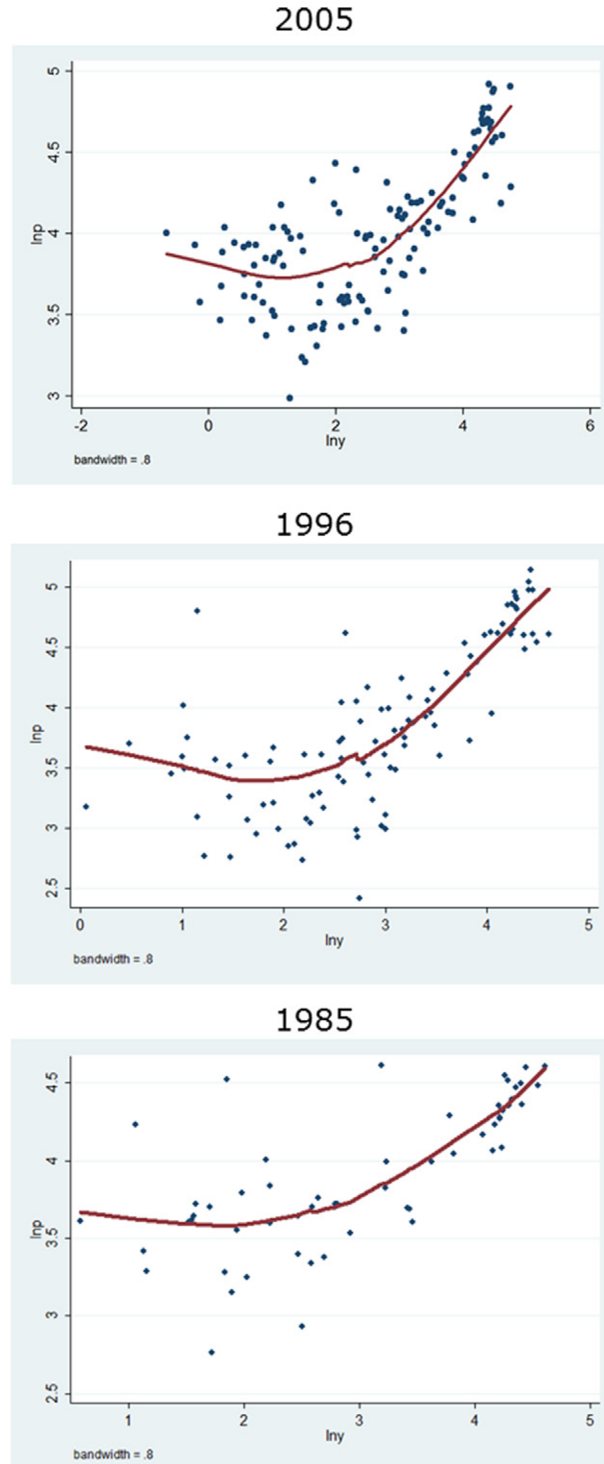

Fig. 6. Price and income: benchmark years and countries.

[43] The price level of sector i is given by $p_i = PPP_i/XRAT$ with $p_i^{US} = 1$. In order to preserve aggregation at the GDP level, I use the Geary–Khamis method to compute sectoral PPPs. See Appendix A.4 for a detailed description of sectoral classification of goods; as suggested by Herrendorf and Valentinyi (2012), we map the agricultural sector with the food sector.

[44] In Table 9 we consider the benchmark countries of PWT 8.0 for the year 2005. We rank countries by their level of income and divide the sample by income group as defined by the World Bank. Then, following the tradition of the development of macroeconomics literature, we focus on a sectoral division of the economy between agriculture, manufacturing, and services.

[45] Caselli (2005) hints at this possibility in a footnote. Lagakos and Waugh (2013) have a similar finding.

[46] Notice, that the ICP-based price of agriculture is based on consumer prices of food products. Those consumer prices are not only driven by prices of agricultural products, but also by transportation and distribution activities. As a robustness check, we use also farm-gate PPPs kindly provided by the FAO statistical office for the 2011 round. The non-monotonicity of the price–income relation in agriculture is confirmed also in this case; results are available upon request.

Table 9
Price–income relation and the stage of development.

Price–income relation		1st tercile Negative	2nd tercile Flat	3rd tercile Positive
Price level				
	Agriculture	0.67	0.63	1.06
	Manufacturing	0.56	0.63	1.03
	Services	0.19	0.27	0.77
Value-added share of GDP				
	Agriculture	30.46	11.09	2.84
	Manufacturing	26.42	37.00	31.95
	Services	43.12	51.92	65.21
Employment share				
	Agriculture	60.61	28.02	6.65
	Manufacturing	10.50	22.10	26.01
	Services	28.33	49.13	66.97
Expenditure share				
	Agriculture	35.08	20.45	8.47
	Manufacturing	41.71	43.86	41.42
	Services	20.28	25.15	29.91

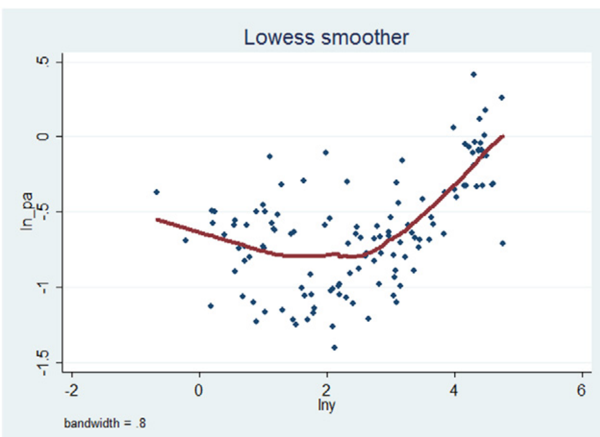

Fig. 7.1. Price of agriculture and income: non-narametric estimation.

increases and this drives down the price of agricultural goods. Given that for low-income countries agriculture represents a large share of the consumption basket, the overall price level goes down. This is at the basis of the downward sloping part of the price–income relation. After a certain level of development, the share of the agricultural sector in the economy becomes negligible; so, even if productivity growth in agriculture is still high and the price of agriculture declines, it will have little effect on the overall price level. At this point what becomes important is the increase of productivity in the manufacturing sector, which drives the overall price level up as in the standard Balassa–Samuelson hypothesis.

The non-monotonicity of the price–income relation arises naturally as a consequence of the process of structural transformation. However, this explanation relies on two underlying assumptions. One is related to the non-tradability of agricultural goods and the other to the ranking of sectoral productivity growth, which needs to be faster in agriculture. We now discuss these more in details.

An important element of our explanation is that agricultural goods are non-tradable, so that there is no price equalization of agricultural products and agricultural prices are relatively higher in poor countries because of lower productivity. More precisely, we do not assume that agricultural goods are intrinsically non-tradable, but that in practice are not traded, at least from the perspective of low-income countries. This assumption is consistent with the empirical findings of Tombe (2015) and Gollin et al. (2007). Tombe (2015) shows that trade costs lead to minimal food imports in poor countries despite the low productivity in agriculture. Moreover, Gollin et al. (2007) argue that *"it is reasonable to view most [poor] economies as closed from the perspective of trade in food"*. They show that in the year 2000 about 70% of arable land in 159 developing countries was devoted to staple food

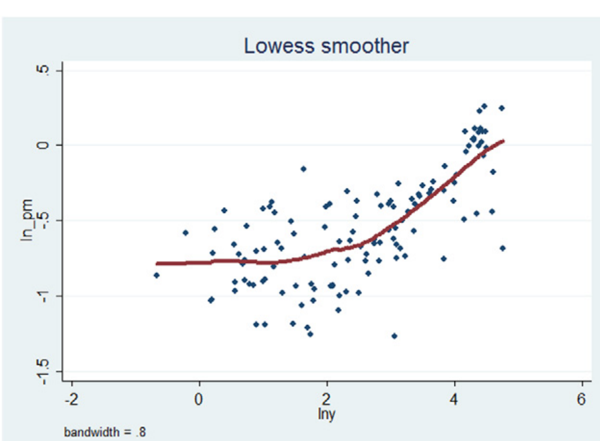

Fig. 7.2. Price of manufacturing and income: non-parametric estimation.

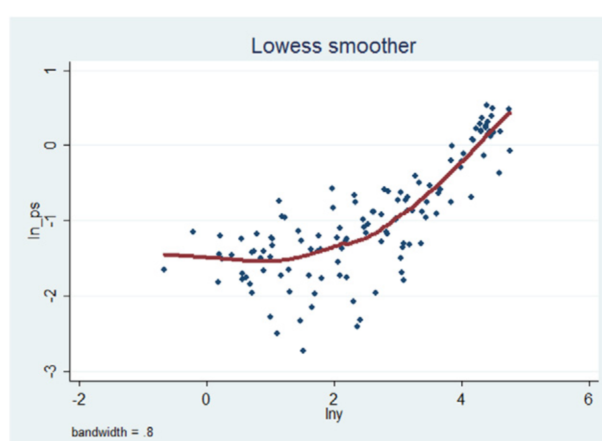

Fig. 7.3. Price of services and income: non-parametric estimation.

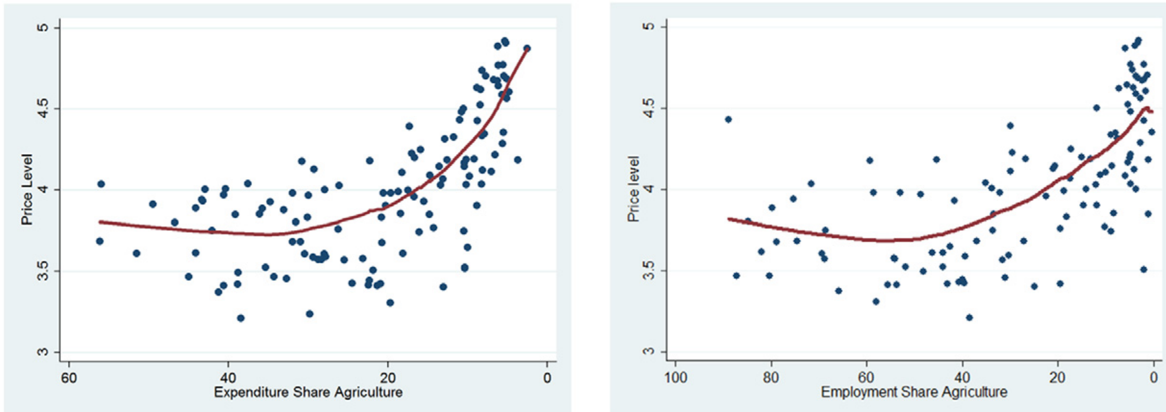

Fig. 8. Price level, expenditure and employment share of agriculture (reversed scale): non-parametric estimation.

crops. With the exception of few developing countries, almost all of the resulting production was for domestic consumption. Using FAOSTAT data for 2005, we find that the share of cereal exports relative to overall production is respectively 3%, 12%, and 37% for the countries where the price–income relation is negative, flat, and positive. Also on the import side there is evidence of low tradability. In fact food imports and food aid are not a major source of food for poor countries and they supply around 5% of total calories consumed.

The other underlying assumption of the mechanism described above is that productivity growth in agriculture is faster than in the other sectors. This is consistent with the Duarte and Restuccia (2010), who show for a panel of 29 countries between 1956–2004 that productivity growth was 4% in agriculture, 3% in manufacturing and 1.3% in services. This finding goes back to the seminal paper of Nurkse (1953) and it is a central aspect in the literature on structural transformation as in Gollin et al. (2002, 2007) and Ngai and Pissarides (2007). Our assumption is also consistent with the "labor push" hypothesis of structural transformation, as in Alvarez-Cuadrado and Poschke (2011). They show that productivity growth in agriculture was the main driver of structural transformation after World War II, when TFP growth in agriculture turned higher than in manufacturing thanks to key innovations in cultivation processes and mechanization.[47]

4.2. Structural transformation and the price level

In order to illustrate the role of structural transformation in the price–income relation, we extend the Balassa–Samuelson model connecting the price level to the process of structural change. We do so by staying as close as possible to the framework and assumptions of Balassa–Samuelson, so we can preserve simplicity and comparability with the standard model.

The consumption-based price index derived in the classical version of the Balassa–Samuelson hypothesis is

$$\log P_z^{BS} = \gamma_{zNT} \left(\log A_{zT} - \log A_{zNT}\right) \tag{5}$$

where γ_{zNT} is the expenditure share on non-tradables in country z, A_{zT} is TFP in the tradable sector, and A_{zNT} is TFP in the non-tradable sector. We can observe that as richer countries have a higher relative productivity in the tradable sector, they will have a higher price level for any given expenditure share on non-tradables.

We extend the standard model to a three-sector environment with agriculture, manufacturing, and services. Moreover, as argued in the previous section, we assume that agriculture is non-traded. Finally, taking the model of Ngai and Pissarides (2007) as a reference, we link the sectoral weights of the price level to the employment shares, so that we can account for the stage of development at which countries are.[48] The price level equation implied by our extended Balassa–Samuelson model is[49]

$$\log P_z^{BS+} = (\gamma_{zA} + \gamma_{zS}) \left[\log A_{zM} - \left(\frac{l_{zA}}{l_{zA} + l_{zs}} \log A_{zA} + \frac{l_{zS}}{l_{zA} + l_{zS}} \log A_{zS}\right)\right] \tag{6}$$

where A_{zi} is TFP of country z in sector i ($i = A, M, S$; agriculture, manufacturing and services); l_{zi} and γ_{zi} are employment shares and expenditures shares of country z in sector i. We label this price equation "Balassa–Samuelson+" because Eqs. (5) and (6) are very similar. The differences are *i)* that in the Balassa–Samuelson+ there is a better focus on the agricultural sector and *ii)* the sectoral TFPs of agriculture and services are weighted by the relative employment shares, so that the price index reflects the stage of structural transformation. If we shut down the focus on the agricultural sector

[47] For periods before World-War II, Alvarez-Cuadrado and Poschke (2011) show that "labor pull" — higher productivity growth in the manufacturing sector — was the main driver of the process of structural transformation.

[48] We choose Ngai and Pissarides (2007) as the main reference between the models of structural transformation along a generalized balanced growth path, because it can generate both a decline in the employment share of agriculture and a change in sectoral relative prices, which is consistent with what we observe in the data. Alternative models like Kongsamut et al. (2001) can generate a decreasing employment share of agriculture, but they imply constant relative prices which is at odds with the empirical evidence of this paper. See Herrendorf et al. (2014) for a detailed discussion of alternative models of structural transformation.

[49] We present the derivation of the model in Appendix A. As in Balassa–Samuleson, we use Cobb–Douglas preferences and take manufacturing as the numeraire. Cobb–Douglas preferences imply constant labor shares, which can sound at odds with the structural transformation explanation of the paper. However, given the cross-sectional nature of the data we are trying to match, we solve the model as a series of cross-sections, so that employment shares and expenditure shares are going to differ across countries and can capture the point of structural change at which countries are. This approach allows us to provide the simplest extension of the Balassa–Samuelson model to account for the role of structural transformation.

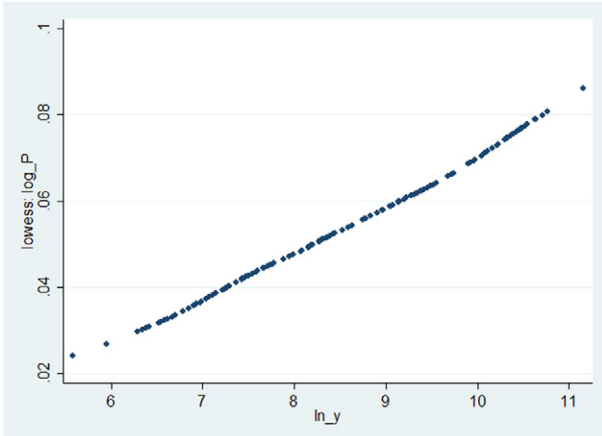

Fig. 9.1. The price level in the Balassa–Samuelson hypothesis: non-parametric estimation of the price–income relation, fitted values.

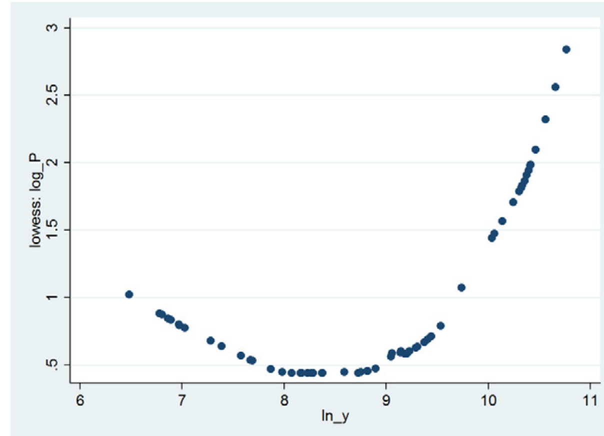

Fig. 9.2. The price level in the Balassa–Samuelson+ hypothesis: non-parametric estimation of the price–income relation, fitted values.

by setting γ_{zA} and l_{zA} equal to zero, as if they were absorbed by the manufacturing sector, we are back to the standard Balassa–Samuelson hypothesis.

Looking at Eq. (6), we can see that as TFP in agriculture increases across countries at an early stage of development, the overall price level can decrease given the high relative employment share of agriculture in these countries. However, in countries at a further stage of structural transformation the employment share in agriculture becomes negligible; hence, the relative TFP in manufacturing turns to be the main driver of the price difference across countries as in the standard Balassa–Samuelson model.

Next, we feed Eqs. (5) and (6) with data on sectoral TFP, expenditure shares, and employment shares. We obtain sectoral estimates of TFP across countries following the methodology of Herrendorf and Valentinyi (2012).[50] Employment shares are taken by the WDI database and by national sources. The consumption share in agriculture and service are given by the expenditure shares from the ICP database.[51]

Finally we run a non-parametric estimation of the price levels implied by the two models and income per-capita. We then compare the two estimates with the one obtained using prices from the PWT.[52] Fig. 9.1 shows the fitted values of the non-parametric estimation of the price–income relation, where prices are given by Eq. (5): I am able to confirm the strictly positive relation predicted by the Balassa–Samuelson hypothesis.

However, Fig. 9.2 shows that the price implied by the "BS+" hypothesis allows for more flexibility in the price–income relation and can generate a negative pattern at low levels of development. Therefore, by taking into account that countries are at a different stage of their process of structural transformation, I am able to match better the actual pattern of the data reported in Fig. 9.3.

Table 10 analyzes the quantitative fit: under the BS+ hypothesis 26% of countries in the sample are on the downward sloping path of the price–income relation; in the standard BS hypothesis this is 0% and in the actual data it is 20% of the sample. The variance of prices generated by the BS+ hypothesis is two and half times higher than in the data (1.02 vs 0.41). Finally, the turning point of the BS+ model is around 3000 PPP $, but in the data it is around 1440 PPP $.

The quantitative result of the "Balassa–Samuelson+" hypothesis clearly outperforms that of the Balassa–Samuelson hypothesis. The model derived in this paper is relatively simple, a richer approach that accounts for other factors like the tradability of agriculture in rich countries or the reduction of trade costs as a country develops might deliver a better quantitative fit. However, adding simple extensions to the Balassa–Samuelson model is sufficient to generate a non-monotonic price–income relation; this is encouraging and lay the ground for further theoretical and empirical research on the relation between structural transformation and the real exchange rate.

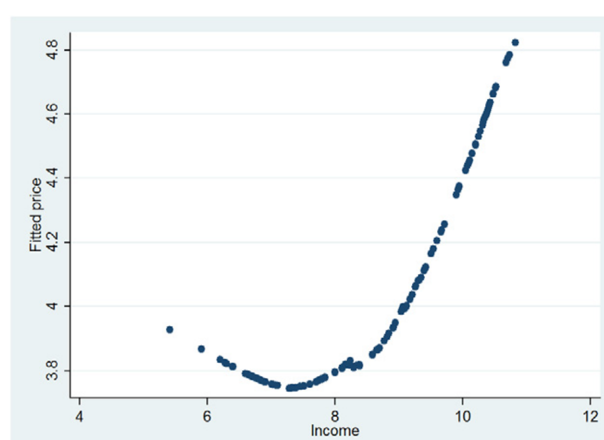

Fig. 9.3. Penn World Table 8.0 (2005): non-parametric estimation of the price–income relation, fitted values.

[50] They elaborate a development accounting framework to compute sectoral productivities using the Penn World Tables; see Appendix A for a detailed description.

[51] We are able to compute the price levels for 60 countries out of 127 because of the lack of sectoral employment data in many poor countries and lack of investment data necessary for computing TFP in middle-income and former USSR countries; following Caselli (2005) I exclude countries with data on investment starting only after the '70s.

[52] Prices in the PWT are derived from prices of a set of goods across countries collected in local currency units. In order to make this local prices comparable, they need to be converted and aggregated using an appropriate methodology (i.e. a PPP conversion or simple conversion in USD). In the case of the PWTs this is done with a PPP conversion using the Geary–Khamis method. The theoretical prices computed by the models are the result of TFP levels, expenditure shares, and employment shares, which are directly comparable across countries, so there is no need to apply a Geary–Khamis method to these prices.

Table 10
Data and models.

	Data	BS+ model	BS model
Countries on the downward sloping path	20%	26%	0%
Price, std. deviation	0.41	1.02	0.02
Turning point	1464 PPP $	3070 PPP	–

5. Conclusions

We show that the relation between price and income is non-monotonic. To our knowledge, this is an original finding, and it is a hitherto undocumented empirical regularity. This result contradicts the conventional wisdom of a positive price–income relation, which draws upon a linear estimation. If we apply a non-parametric estimation or allow for non-linearities in standard regressions, the price–income relation turns out to be significantly negative in poor countries. This finding is robust along both cross-section and panel dimensions. The new evidence presented in this paper raises general questions about the relation between the process of economic development and the price level, as well as about the long-run determinants of real exchange rates in poor countries.

The paper shows that a model linking the price level to the process of structural transformation that characterizes developing countries can generate a non-monotonic pattern of the price–income relation. This result suggests that structural change and, more generally, inter-sectoral dynamics can be important determinants of real exchange rates movements.

Nevertheless, a richer theoretical approach could improve the quantitative fit. For instance, the model does not account for the role of trade costs. Trade costs are much higher than is generally recognized, even for traded goods: Anderson and Van Wincoop (2004) estimate that, for developed countries, trade costs average 170% of production costs, of which roughly half is international trade costs and half internal trade costs. For developing countries, they claim that this ratio is often higher, and many studies do indeed show strikingly high transport costs for individual developing countries or groups thereof (Limao and Venables, 2001).

Trade costs and the ratio of trade costs to production costs may well vary systematically with the level of development. For example as a low-income country starts developing, its infrastructure improves reducing both internal and external trade costs as well as the ratio of trade costs to production. This might turn to be a key element in explaining the initial negative pattern of the price–income relation and deserves further investigation. This is consistent with Du et al. (2013) who show that transport infrastructure is an important determinant of exchange rate especially in developing countries.

The tradability of agriculture in more developed countries is another feature that a richer model should account for. In the current model, agriculture is completely non-tradable and this could partly explain the high variance of prices and the turning point's high level of income that the model predicts.

Finally, a possible empirical extension of the paper could focus on regional variation within countries like India or China, where there are regions at very different stages of development. This kind of regional variation would be ideal to verify if the process of structural transformation is at the basis of the non-monotonic price–income relation.

This paper lays the ground for further theoretical and empirical research on the relation between economic development and the price level. The results presented, although surprising, should not be disturbing. It is probable that Samuelson himself would not have been startled. In his 1994 article for the thirty-year anniversary of the Balassa–Samuelson model, he wrote that "*The Penn–Balassa–Samuelson effect is an important phenomenon of actual history but not an inevitable fact of life. It can quantitatively vary and,* ***in different times and places, trace to quite different processes***".

Appendix A

A.1. Countries in the cross-section analysis of section

Albania	Cote d'Ivoire	Israel	Mexico	Slovenia
Algeria	Croatia	Italy	Moldova	South Africa
Angola	Czech Republic	Jamaica	Mongolia	Spain
Armenia	Denmark	Japan	Morocco	Sri Lanka
Australia	Dominican Rep.	Jordan	Mozambique	Sudan
Austria	Ecuador	Kazakhstan	Myanmar	Sweden
Azerbaijan	Egypt	Kenya	Namibia	Switzerland
Bangladesh	Estonia	Korea	Nepal	Tajikistan
Belarus	Ethiopia	Kyrgyzstan	Netherlands	Tanzania
Belgium	Finland	Laos	New Zealand	Thailand
Benin	France	Jamaica	Nicaragua	Togo
Bolivia	Gabon	Japan	Niger	Trinidad & Tobago
Bosnia and Herz.	Gambia, The	Jordan	Nigeria	Tunisia
Botswana	Georgia	Kazakhstan	Norway	Turkey
Brazil	Germany	Kenya	Oman	Uganda
Bulgaria	Ghana	Korea	Pakistan	Ukraine
Burkina Faso	Greece	Kyrgyzstan	Panama	United Kingdom
Burundi	Guatemala	Laos	Paraguay	United States
Cambodia	Guinea	Latvia	Peru	Uruguay
Cameroon	Guinea-Bissau	Lesotho	Philippines	Venezuela
Canada	Haiti	Liberia	Poland	Vietnam
Central Afr. Rep.	Honduras	Lithuania	Romania	Yemen
Chad	Hong Kong	Macedonia	Russia	Zambia
Chile	Hungary	Madagascar	Rwanda	Zimbabwe
China	India	Malawi	Senegal	
Colombia	Indonesia	Malaysia	Serbia	
Congo, Dem. Rep.	Iran	Mali	Sierra Leone	
Congo, Rep. of	Iraq	Mauritania	Singapore	
Costa Rica	Ireland	Mauritius	Slovak Rep.	

A.2. Derivation of the Balassa–Samuelson + price equation

A.2.1. Model setup

A representative consumer in country z maximizes the following utility function across three aggregate goods in agriculture, manufacturing, and services:[53]

$$U(C_a, C_m, C_s) = C_a^{\gamma_a} C_m^{\gamma_m} C_s^{\gamma_s} \tag{7}$$

Firms in each sector maximize a Cobb–Douglas production function technology with capital and labor such that

$$F_i\,(k_i, l_i) = A_i k_i^{\alpha} n_i^{1-\alpha}; \quad i = a, m, s \tag{8}$$

Market clearing must then satisfy

$$\sum_{i=1}^{m} l_i = 1; \quad \sum_{i=1}^{m} k_i = k; \tag{9}$$

Finally, we assume $F_i = c_i$ for $i = a, s$ and that manufacturing produces both a final consumption good and the economy's capital stock so that $\dot{k} = F^m - c_m - (\delta + n)\,k$. This means that manufacturing is the only tradable good and that trade is balanced period by period.[54] This assumption implies that the effect of trade is to equalize the price of manufacturing across countries and that there is financial autarky across countries, which is a reasonable assumption for low-income countries.

The consumption-based price index measures the least expenditure that buys a unit of the consumption index on which period utility depends. Given our utility specification this is going to be

$$\log P = \gamma_a \log P_a + \gamma_m \log P_m + \gamma_s \log P_s \tag{10}$$

Accounting for the cross-country equalization of the price of manufacturing through trade and normalizing it to one, the consumption-based price index can be written as

$$\log P = \gamma_a \log p_a + \gamma_s \log p_s \tag{11}$$

Solving the problem for the Cobb–Douglas case can sound at odds with the explanation of structural transformation provided in the paper. This is because under Cobb–Douglas preferences expenditure and employment shares are constant for a country in a time series dimension. However, given the empirical data that our model is trying to match, we are solving the problem as a series of cross-sections so that employment shares and expenditure shares are going to differ across countries and to capture the point of structural transformation at which each country is. This approach allows us to keep the model easily comparable with the standard Balassa–Samuelson hypothesis and it is consistent with the fact that we match a cross-sectional empirical result.

A.2.2. Production maximization, relative prices, and employment shares

From the supply-side, static efficiency condition requires equal marginal rate of technical substitution across sectors, so that $k_i = k$; while free movement of capital and labor leads to equal remuneration of the factors of production. Therefore, firms' profit maximization implies

$$\frac{P_a}{P_m} = \frac{A_m}{A_a} \tag{12}$$

$$\frac{P_s}{P_m} = \frac{A_s}{A_a} \tag{13}$$

From consumer's optimization problem we can write the relative expenditure of agriculture and services respect to manufacturing as

$$\frac{P_a c_a}{P_m c_m} = \frac{\gamma_a}{\gamma_m}\left(\frac{P_a}{P_m}\right)^{1-\theta} \equiv x_a \tag{14}$$

$$\frac{P_s c_s}{P_m c_m} = \frac{\gamma_s}{\gamma_m}\left(\frac{P_s}{P_m}\right)^{1-\theta} \equiv x_s \tag{15}$$

We then define $X = x_a + x_s + x_m$, where clearly $x_m = 1$. We also define

$$c \equiv \sum_{i=1}^{m} P_i c_i; \quad y \equiv \sum_{i=1}^{m} P_i F^i \tag{16}$$

[53] To save on notation we dismiss the country subscript z for the rest of Appendix A.

[54] This is similar to the one in the standard Balassa–Samuelson model and it helps to keep our model as close and as comparable as possible to the standard one.

Using Eqs. (14) and (15) and the efficiency conditions, we can rewrite Eq. (16) as

$$c = P_m c_m X; \quad y = P_m A_m k^{\alpha} \tag{17}$$

Notice that the technology parameter for output is TFP in manufacturing not an average of all sectors.

As in Ngai and Pissarides (2007) we can link relative expenditure with the employment shares. If we substitute we substitute $F^i = c_i$ for $i = a, s$ in Eqs. (14) and (15), using the market clearing conditions in Eq. (9), we can show that it results

$$l_a = \frac{c}{y}\frac{x_a}{X} \tag{18}$$

$$l_s = \frac{c}{y}\frac{x_s}{X} \tag{19}$$

The employment share in the manufacturing sector is derived by firstly observing that $l_m = 1 - l_a - l_s$, so that we have

$$l_m = \frac{c}{y}\frac{x_m}{X} + \left(1 - \frac{c}{y}\right) \tag{20}$$

Let's consider the case where $\theta = 1$ and manufacturing is the numeraire. In this case the price index is given by $\log P = \gamma_a logp_a + \gamma_s logp_s$. By using firm's optimality conditions (12) and (13) as well as (18) and (19) we can write the price level as

$$\log P = (\gamma_a + \gamma_s)\left[\log A_m - \left(\frac{l_a}{l_a + l_s}\log A_a + \frac{l_s}{l_a + l_s}\log A_s\right)\right] \tag{21}$$

A.3. Sectoral TFPs methodology

In order to compute sectoral TFPs, I use the methodology of Herrendorf and Valentinyi (2012) who elaborate a sectoral development accounting framework that allows to compute sectoral TFPs using PWT. The key assumptions of their methodology are as follows: competitive markets; factor's mobility across sectors; Cobb–Douglas production function with factor shares common to all countries.

The production function for sector i in country z is given by

$$y_i^z = A_i^z \left(k_i^z\right)^{\theta_i} \left(l_i^z\right)^{\phi_i} \left(h_i^z\right)^{1-\theta_i-\phi_i} \tag{22}$$

where k is capital, l is land, and h is human capital.

Under the assumption stated above, Herrendorf and Valentinyi (2012) show that the sectoral factors of production are

$$k_i^z = \frac{\theta_i p_i^z y_i^z}{\sum_j \theta_j p_j^z y_j^z} \sum_i k_i^z \tag{23}$$

$$l_i^z = \frac{\phi_i p_i^z y_i^z}{\sum_j \phi_j p_j^z y_j^z} \sum_i l_i^z \tag{24}$$

$$h_i^z = \frac{(1-\theta_i-\phi_i)\, p_i^z y_i^z}{\sum_j (1-\theta_j-\phi_j)\, p_j^z y_j^z} \sum_i h_i^z \tag{25}$$

In order to compute sectoral TFPs, I take the sectoral factor shares from Herrendorf and Valentinyi (2012), who calculate them from the US input–output tables. Then, following their methodology, I compute the capital stock in the economy k^z with the perpetual inventory method as in Caselli (2005). Land l^z is arable land for agriculture and urban land for manufacturing and services. I take data on arable land from FAOSTAT and following World Bank (2006) estimates, I set urban land equal to 24% of physical capital. Finally, I compute human capital h^z as in Caselli (2005) and it is an increasing function of average years of schooling per worker.

A.4. ICP 2005, classification of goods

Category	Basic heading	BS–SC framework: sector allocation	BS framework: tradability
Food	Rice	A	T
	Other cereals and flour	A	T
	Bread	A	T
	Other bakery products	A	T
	Pasta products	A	T
	Beef and veal	A	T
	Pork	A	T
	Lamb, mutton and goat	A	T
	Poultry	A	T
	Other meats and preparations	A	T
	Fresh or frozen fish and seafood	A	T
	Preserved fish and seafood	A	T
	Fresh milk	A	T
	Preserved milk and milk products	A	T
	Cheese	A	T
	Eggs and egg-based products	A	T
	Butter and margarine	A	T
	Other edible oils and fats	A	T
	Fresh or chilled fruit	A	T
	Frozen, preserved or processed fruits	A	T
	Fresh or chilled vegetables	A	T
	Fresh or chilled potatoes	A	T
	Frozen or preserved vegetables	A	T
	Sugar	A	T
	Jams, marmalades and honey	A	T
	Confectionery, chocolate and ice cream	A	T
	Food products n.e.c.	A	T
Beverages and tobacco	Coffee, tea and cocoa	M	T
	Mineral waters, soft drinks, fruit and veg juices	M	T
	Spirits	M	T
	Wine	M	T
	Beer	M	T
	Tobacco	M	T
Clothing and footwear	Clothing materials and accessories	M	T
	Garments	M	T
	Cleaning and repair of clothing	S	NT
	Footwear	M	T
	Repair and hire of footwear	S	NT
Housing, water, electricity and gas	Actual and imputed rentals for housing	S	NT
	Maintenance and repair of the dwelling	S	NT
	Water supply and miscellaneous services relating to the dwelling	S	NT
	Miscellaneous services relating to the dwelling	S	NT
	Electricity	M	T
	Gas	M	T
	Other fuels	M	T
Furniture, household equipment and maintenance	Furniture and furnishings	M	T
	Carpets and other floor coverings	M	T
	Repair of furniture, furnishings and floor coverings	S	NT
	Household textiles	M	T
	Major household appliances whether electric or not	M	T
	Small electric household appliances	M	T
	Repair of household appliances	S	NT
	Glassware, tableware and household utensils	M	T
	Major tools and equipment	M	T
	Small tools and miscellaneous accessories	M	T
	Non-durable household goods	M	T
	Domestic services	S	NT
	Household services	S	NT
Health	Pharmaceutical products	M	T
	Other medical products	M	T
	Therapeutical appliances and equipment	M	T
	Medical services	S	NT
	Dental services	S	NT
	Paramedical services	S	NT
	Hospital services	S	NT
Transport	Motor cars	M	T
	Motor cycles	M	T
	Bicycles	M	T
	Fuels and lubricants for personal transport equipment	M	T
	Maintenance and repair of personal transport equipment	S	NT
	Other services in respect of personal transport equipment	S	NT
	Passenger transport by railway	S	NT

(continued on next page)

(*continued*)

Category	Basic heading	BS–SC framework: sector allocation	BS framework: tradability
Transport	Passenger transport by road	S	NT
	Passenger transport by air	S	NT
	Passenger transport by sea and inland waterway	S	NT
	Combined passenger transport	S	NT
	Other purchased transport services	S	NT
Communication	Postal services	S	NT
	Telephone and telefax equipment	M	T
	Telephone and telefax services	S	NT
Recreation and culture	Audio–visual, photographic and information processing equipment	M	T
	Recording media	M	T
	Repair of audio–visual, photographic and information processing equipment	S	NT
	Major durables for outdoor and indoor recreation	M	T
	Other recreational items and equipment	M	T
	Gardens and pets	S	NT
	Veterinary and other services for pets	S	NT
	Recreational and sporting services	S	NT
	Cultural services	S	NT
	Games of chance	S	NT
	Newspapers, books and stationery	S	NT
	Package holidays	S	NT
Education	Education	S	NT
Restaurant and hotels	Catering services	S	NT
	Accommodation services	S	NT
Miscellaneous goods and services	Hairdressing salons and personal grooming establishments	S	NT
	Appliances, articles and products for personal care	S	NT
	Prostitution	S	NT
	Jewelry, clocks and watches	M	T
	Other personal effects	M	T
	Social protection	S	NT
	Insurance	S	NT
	FISIM	S	NT
	Other financial services n.e.c	S	NT
	Other services n.e.c.	S	NT
Government expenditure	Government compensation of employees	S	NT
	Government intermediate consumption	M	T
	Government gross operating surplus	S	NT
	Government net taxes on production	S	NT
	Government receipts from sales	S	NT
Capital formation	Metal products and equipment	M	T
	Transport equipment	M	T
	Residential buildings	M	T
	Non-residential buildings	M	T
	Civil engineering works	M	T
	Other products	M	T
Inventories	Changes in inventories and acquisitions	M	T

A=agriculture; M=manufacturing; S=services; T=tradable; NT=non-tradable.

The sectoral allocation and the tradability allocation apply respectively to the estimation of the Balassa–Samuelson-Structural-Change and the Balassa–Samuelson framework in Section 4.

References

Almas, I., 2012. International income inequality: measuring PPP bias by estimating Engel curves for food. Am. Econ. Rev. 102, 1093–1117.

Alvarez-Cuadrado, F., Poschke, M., 2011. Structural change out of agriculture: labor push versus labor pull. Am. Econ. J. Macroecon. 3, 127–158.

Anderson, J.E., Van Wincoop, E., 2004. Trade costs. J. Econ. Lit. 42, 691–751.

Atkin, D., Donaldson, D., 2012. Who's getting globalized? Size and nature of intranational trade costs, mimeo, MIT.

Balassa, B., 1964. The purchasing-power parity doctrine: a reappraisal. J. Polit. Econ. 72, 584–596.

Barro, R.J., 1991. Economic growth in a cross section of countries. Q. J. Econ. 106, 407–443.

Bergin, P.R., Glick, R., Taylor, A.M., 2006. Productivity, tradability, and the long-run price puzzle. J. Monet. Econ. 53, 2041–2066.

Berka, M., Devereux, M.B., Engel, C., 2014. Real exchange rates and sectoral productivity in the Eurozone. NBER Working Paper. no. 20510.

Bhagwati, J.N., 1984. Why are services cheaper in poor countries? Econ. J. 94, 279–286.

Bordo, M.D., Choudri, E.U., Fazio, G., MacDonald, R., 2014. The real exchange rate in the long run: Balassa–Samuelson effects reconsidered. NBER Working Paper. no. 20228.

Caselli, F., 2005. Accounting for cross-country income differences. In: Aghion, P., Durlauf, S. (Eds.), Handbook of Economic Growth. North-Holland, Amsterdam.

Chen, L.L., Choi, S., Devereux, J., 2007. Searching for Balassa Samuelson in post-war data. CRIF Working Paper Series No. 2.

Choudhri, E.U., Khan, M.S., 2005. Real exchange rates in developing countries: are Balassa–Samuelson effects present? IMF Staff. Pap. 52, 387–409.

Deaton, A.S., 2010. Price indexes, inequality, and the measurement of world poverty. Am. Econ. Rev. 100, 5–34.

Deaton, A.S., Aten, B., 2014. Trying to understand the PPPs in ICP2011: why are the results so different? NBER Working Paper. No. 20224.

Deaton, A.S., Heston, A.A., 2010. Understanding PPPs and PPP-based national accounts. Am. Econ. J. Macroecon. 2, 1–35.

De Gregorio, J., Giovannini, A., Wolf, H.C., 1994. International evidence on tradables and non-tradables inflation. Eur. Econ. Rev. 38, 1225–1244.

Devereux, M.B., 1999. Real exchange rate trends and growth: a model of East Asia. Rev. Int. Econ. 7, 509–521.

Diewert, E., 2008. New methodology for linking the regions. University of British Columbia Working Paper.

Du, Q., Wei, S.-J., Peichu, X., 2013. Roads and the real exchange rate. NBER Working Paper No. 19291.

Duarte, M., Restuccia, D., 2010. The role of the structural transformation in aggregate productivity. Q. J. Econ. 125, 129–173.

Easterly, W., Levine, R., 2001. What have we learned from a decade of empirical research on growth? It's not factor accumulation: stylized facts and growth models. World Bank Econ. Rev. 15, 177–219.

Feenstra, R.C., Ma, H., Neary, P.J., Rao, P.D.S., 2013. Who shrunk China? Puzzles in the measurement of real GDP. Econ. J. 123, 1100–1129.

Feenstra, R.C., Inklaar, R., Timmer, M., 2015. The next generation of the Penn World Table. Am. Econ. Rev. 105, 3150–3182.
Genius, M., Tzouvelekas, V., 2008. The Balassa–Samuelson productivity bias hypothesis: further evidence using panel data. Agric. Econ. Rev. 9, 31–40.
Gershenkron, A., 1947. The Soviet indices of industrial production. Rev. Econ. Stud. 29, 217–226.
Gollin, D., Lagakos, D., Waugh, M.E., 2014. The agricultural productivity gap in developing countries. Q. J. Econ. 129, 939–993.
Gollin, D., Parente, S.L., Rogerson, R., 2002. The role of agriculture in development. Am. Econ. Rev. 92, 160–164.
Gollin, D., Parente, S.L., Rogerson, R., 2007. The food problem and the evolution of international income levels. J. Monet. Econ. 54, 1230–1255.
Herrendorf, B., Valentinyi, A., 2012. Which sectors make poor countries so unproductive? J. Eur. Econ. Assoc. 10, 323–341.
Herrendorf, B., Rogerson, R., Valentinyi, A., Aghion, P., Durlauf, S., 2014. Growth and structural transformation. Handbook of Economic Growth, North-Holland, Oxford.
Hill, R.J., 2000. Measuring substitution bias in international comparison based on additive purchasing power methods. Eur. Econ. Rev. 44, 145–162.
Hill, R.J., 2004. Constructing price indexes across space and time: the case of the European Union. Am. Econ. Rev. 94, 1379–1410.
ICP, 2007. ICP Methodological Handbook. The World Bank, Washington D.C.
ICP, 2015. Purchasing Power Parities and the Real Size of World Economies. The World Bank, Washington D.C.
Inklaar, R., Rao, D.S.P., 2014. Cross-country income levels over time: did the developing world suddenly become much richer? GGDC Research Memorandum no. 151.
Kongsamut, P., Rebelo, S., Xie, D., 2001. Beyond balanced growth. Rev. Econ. Stud. 68, 869–882.
Kravis, I.B., Summers, R., Heston, A., 1982. World Product and Income. John Hopkins University Press.
Lagakos, D., Waugh, M.E., 2013. Selection, agriculture, and cross-country productivity differences. Am. Econ. Rev. 103, 948–980.
Limao, N., Venables, A.J., 2001. Infrastructure, geographical disadvantage, transport costs and trade. World Bank Econ. Rev. 15, 451–479.
Lind, J.T., Mehlum, H., 2010. With or without U? The appropriate test for a U-shaped relationship. Oxf. Bull. Econ. Stat. 72, 109–118.
Neary, J., 2004. Peter rationalizing the Penn World Tables. Am. Econ. Rev. 94, 1411–1428.
Ngai, R., Pissarides, C., 2007. Structural change in a multisector model of growth. Am. Econ. Rev. 97, 429–448.
Nurkse, R., 1953. Problems of Capital Formation in Underdeveloped Countries. Oxford University Press, New York.
Nuxoll, D.A., 1994. Differences in relative prices and international differences in growth rates. Am. Econ. Rev. 84, 1423–1436.
Rao, P.D.S., 2004. The country-product-dummy method: a stochastic approach to the computation of purchasing power parities in the ICP. CEPA Working Papers Series No. 32004.
Reinhart, C.M., Rogoff, K.S., 2004. The modern history of exchange rate arrangements: a reinterpretation. Q. J. Econ. 119, 1–48.
Rodrik, D., 2008. The real exchange rate and economic growth. Brook. Pap. Econ. Act. 39, 365–439.
Rogoff, K.S., 1996. The purchasing-power parity puzzle. J. Econ. Lit. 34, 647–668.
Sasabuchi, S., 1980. A test of a multivariate normal mean with composite hypotheses determined by linear inequalities. Biometrika 67, 429–439.
Samuelson, P.A., 1964. Theoretical notes on trade problems. Rev. Econ. Stat. 46, 145–154.
Samuelson, P.A., 1994. Facets of Balassa–Samuelson thirty years later. Rev. Int. Econ. 2, 201–226.
Solow, R., 1956. A contribution to the theory of economic growth. Q. J. Econ. 70, 65–94.
Summers, R., Heston, A., 1991. The Penn-World Table (mark 5): an expanded set of international comparisons, 1950-1988. Q. J. Econ. 106, 327–368.
Taylor, A.M., Taylor, M.P., 2004. The purchasing-power-parity debate. J. Econ. Perspect. 18, 135–158.
Tombe, T., 2015. The missing food problem: trade, agriculture, and international productivity differences. Am. Econ. J. Macroecon. 7, 1–33.