BGPE Discussion Paper

No. 102

Are Wages Equal Across Sectors of Production? A Panel Data Analysis for Tradable and Non-Tradable Goods

Achim Schmillen

June 2011

ISSN 1863-5733

Editor: Prof. Regina T. Riphahn, Ph.D.
Friedrich-Alexander-University Erlangen-Nuremberg

BGPE Discussion Paper Series

— *June 26, 2011*—

ARE WAGES EQUAL ACROSS SECTORS OF PRODUCTION? — A PANEL DATA ANALYSIS FOR TRADABLE AND NON-TRADABLE GOODS

Achim Schmillen[1]

The assumption that national labor markets are homogenous across tradable and non-tradable goods is common in multisector (open-economy) macro models and crucial for the prominent Balassa-Samuelson hypothesis. This study tests it with a novel method to distinguish the tradable and non-tradable sectors grounded in economic theory, modern empirical methods and a large and detailed macro data set. It finds that both the internal relationship between productivity and wages in the tradable and non-tradable sectors postulated by the Balassa-Samuelson hypothesis and its external transmission mechanism are rejected.

Keywords: Balassa-Samuelson, Wage equalization, Tradability
JEL-Classification: F31, F41.

1. INTRODUCTION

Multisector macro models commonly assume wage equalization across sectors of production. At the same time, the labor economics literature has been heavily contesting this assumption with micro data evidence since at least the late 1980s [cf. Dickens and Katz (1987) and Krueger and Summers (1987)]. Starting with Bernard and Jensen (1995) the international trade literature has similarly used micro data to show that exporters tend to pay higher wages than non-exporters.

This study shifts the attention from such micro case studies to a cross-country perspective in order to test whether labor markets are homogenous or not in a macro context. More specifically, it is concerned with one of the most prominent open-economy macro frameworks — the Balassa-Samuelson hypothesis.

The Balassa-Samuelson (BS) hypothesis [formulated by Balassa (1964) and Samuelson (1964) and sometimes also referred to as Harrod-Balassa-Samuelson hypothesis in reference to a much earlier contribution by Harrod (1933)] is the dominant approach for explaining long-run real exchange rate developments. Both in its original presentation and in more recent formulations by Ghironi and Melitz (2005), Bergin et al. (2006) or Herrendorf and Valentinyi (2006) it crucially relies on the assumption that labor markets are homogenous across tradable and non-tradable goods. The main contribution of this study is to show that this assumption cannot be confirmed empirically. Thus, conventional models of real exchange rate determination miss an important factor and should be

[1]Osteuropa-Institut Regensburg, Institute for Employment Research and University of Regensburg. Correspondence to: Osteuropa-Institut Regensburg, Landshuter Strasse 4, D-93047 Regensburg, Germany; e-mail: schmillen@osteuropa-institut.de; phone: +49 (941) 943-5422; fax: +49 (941) 943-5427.

amended to account for heterogeneous labor markets across sectors of production. More generally, results from all multisector (open-economy) macro models that assume homogenous labor markets may have to be regarded with skepticism.

Even though the assumption of wage equalization across sectors of production is central in the BS context it has so far been tested by barely a handful of cross-country panel studies. These studies — most prominently Strauss and Ferris (1996), Strauss (1997), Strauss (1998) and Lee (2005) — almost unanimously reject the assumption of wage equalization in levels across sectors of production. However, all relevant studies suffer from a number of drawbacks: First, they use rather small samples of industrial countries that never include those economies for which the BS hypothesis is often held to be politically most relevant (that is especially those from Central or Eastern Europe). Second, they *ad hoc* decide which sectors to label "tradable" and which "non-tradable", a crucial distinction in the BS context. Third, they tend to rely on econometric methods that might not by totally unproblematic.

In contrast, this study tests the empirical validity of the assumption of homogenous labor markets across sectors of production with a bigger, more detailed and more up-to-date country sample that includes a number of Central or Eastern European economies. It also introduces a new method to distinguish tradable and non-tradable sectors to the BS literature that is grounded in economic theory. Finally, it utilizes modern econometric methods that allow relevant variables to be analyzed in levels and are robust to possible non-stationarities.

The remainder of this paper is structured as follows: Section 2 deals with the theoretical basis of this study. Section 3 introduces the data and discusses the crucial question of how to define the tradable sector and other methodological issues. Section 4 contains the main results and Section 5 concludes.

2. THEORETICAL CONSIDERATIONS

The simplest version of the BS hypothesis is stated within a framework of two small open economies, two homogenous goods (one tradable, one non-tradable), and one factor of production (labor) building on the following assumptions.

Assumption 1. Markets are competitive with real wages equal to labor productivity:

$$P_i^j = \frac{W_i^j}{A_i^j} \quad \forall i, j, \tag{2.1}$$

with P denoting prices, W nominal wages and $A = \frac{Y}{L}$ labor productivity. Superscript $j = \{H, F\}$ gives the country (***H****ome* or ***F****oreign*) and subscript $i = \{T, NT\}$ the sector (***T****radables* or ***N****on-****T****radables*).

Assumption 2. PPP holds for tradables:

$$\frac{P_T^F}{EP_T^H} = 1, \tag{2.2}$$

with E as the bilateral exchange rate between H and F. An increase in E corresponds to a depreciation of H's currency. If one further assumes that PPP holds for tradables not only between H and F but also between these two counties and the rest of the world then the assumption that H and F are small open economies implies that they are price-takers with respect to tradable goods.

Assumption 3. If in addition to that national labor markets are homogenous across sectors of production,

$$W_T^j = W_{NT}^j \quad \forall j, \tag{2.3}$$

then one arrives at what Égert et al. (2003) call the BS effect's *internal transmission mechanism*: A productivity increase in the tradables sector leads to a wage increase in this sector. This in turn implies higher wages for non-tradables and ultimately higher prices for non-tradable goods. Formally, taking logarithms (denoted by small type) and rearranging yields

$$w_T^j = p_T^j + a_T^j \quad \forall j, \tag{2.4}$$

$$w_{NT}^j = p_T^j + a_T^j \quad \forall j, \tag{2.5}$$

$$p_T^j = p_T^j + a_T^j - a_{NT}^j \quad \forall j. \tag{2.6}$$

Assumption 4. Moreover, assume that the national price level is given by the geometric mean of sectoral price levels and equal preferences across countries are described by constant and equal consumption expenditure shares for tradables and non-tradables, θ and $1-\theta$, respectively (with $0 < \theta < 1$):

$$P^j = \left(P_T^j\right)^\theta \left(P_{NT}^j\right)^{1-\theta} \quad \forall j. \tag{2.7}$$

A combination of Assumptions 1 to 4 with the definition of the bilateral real exchange rate ($Q = \frac{P^F}{EP^H}$) leads to an expression [dubbed the BS effect's *external transmission mechanism* by Égert et al. (2003)],

$$Q = \left(\frac{A_T^F A_{NT}^H}{A_T^H A_{NT}^F}\right)^{1-\theta}, \tag{2.8}$$

where real exchange rates depend on sectoral productivities but in no way on sectoral wages.

To sum up, based on the assumption of homogenous labor markets the BS hypothesis predicts (1) that "productivity growth in the traded sector should be closely related to wage growth in both the traded and nontraded sector" and (2) that "real exchange rates depend solely on productivity differentials between the tradable and non-tradable sectors" (Strauss, 1997, p. 393).[1]

[1] During the last decades much more elaborated formulations of the BS hypothesis have been developed than the one presented here. While these relinquish many of the assumptions of the most basic version even recent formulations [like Bergin et al. (2006) or Herrendorf and Valentinyi (2006)] crucially rely on the assumption of homogenous labor markets.

3. DATA AND METHODOLOGY

3.1. *Data*

The two main data sources of this study are the March 2008 release of the EU KLEMS database and version 6.3 of the Penn World Tables.[2] All comparable studies mentioned in Section 1 rely on data sets that are much older, allow the distinction of only a handful of sectors and cover about a dozen countries at best. What is more, none of these studies includes any Central or Eastern European economies while both the EU KLEMS database and the Penn World Tables contain data on the Czech Republic, Estonia, Hungary, Latvia, Lithuania, Poland, Slovenia and Slovakia as well as a on a number of OECD countries.[3]

From the EU KLEMS database sectoral data on gross value added, price indices of value added, compensation of employees and total hours worked by persons engaged are extracted for 49 sectors in 29 countries from 1995 to 2005. From the raw data real labor productivity (A) and real compensation per employee (W) are calculated, both in U.S. dollars and relative to the United States.

Additionally, data on *comparative prices* (also relative to the United States) are retrieved from the Penn World Tables. Comparative prices are a measure of a country's weighted real exchange rate. As Frensch (2006) shows their yearly changes are highly correlated with those of trade-weighted real effective exchange rate indices. Compared to real effective exchange rates comparative prices have the enormous advantages of being more widely available and also of being internationally comparable in level terms. That is why this study uses them as its real exchange rate measure (Q).

3.2. *Tradability*

In the BS context the distinction between goods that are tradable and those that are non-tradable is crucial. Mostly due to widespread data limitations the overwhelming majority of relevant empirical studies decides *ad hoc* which sectors to label tradable and which non-tradable. Together with the very crude sectoral breakdown commonly employed such an approach has the danger to seriously distort empirical results. In particular, almost all relevant studies count the whole service sector (a very diverse sector that accounts for more than two thirds of GDP in most advanced economies) as non-tradable. Thus they completely ignore the importance of trade in services [noted for instance by Eichengreen and Gupta (2009)].

Apart from an approach by Frensch and Schmillen (forthcoming) that circumvents an explicit distinction of tradables and non-tradables by using trade-based measures of product variety which are, by definition, only available for tradables and have additionally been identified as proxies for productivity, the only serious

[2]For detailed descriptions of the EU KLEMS database and the Penn World Tables see Timmer et al. (2007) and Summers and Heston (1996), respectively.

[3]For a complete list of countries and sectors covered see Appendix A.

attempt to relinquish an *ad hoc* distinction between tradables and non-tradables in the context of the empirical BS literature was made by De Gregorio et al. (1994). De Gregorio et al. (1994) equate tradability of a good with actual trade in that good. More specifically, they classify a sector as tradable if more than 10 percent of its total production is exported. Unfortunately, this approach is not without drawbacks. Apart from the issue that tradability and actual trade might not always be the same thing data issues somewhat limit its usefulness. In particular, data on the export share in total production is not available for a detailed breakdown of the service sector.

This study introduces an alternative method to distinguish the tradable and non-tradable sectors to the BS literature. The basic idea — based on a model by Helpman and Krugman (1985) — is that tradables should tend to be geographically concentrated in order to take advantage of economies of scale. In contrast, such geographic concentrations would not be possible for non-tradables which should more or less be distributed uniformly with population and income.[4]

As an illustration Krugman (1991, p. 65) uses the example of the service sector to note that "(i)n the late twentieth century the great bulk of our labor force makes services rather than goods. Many of these services are nontradable and simply follow the geographical distribution of the goods-producing population – fast-food outlets, day-care providers, divorce lawyers surely have locational Ginis pretty close to zero. Some services, however, especially in the financial sector, can be traded. Hartford is an insurance city; Chicago the center of futures trading; Los Angeles the entertainment capital; and so on."

Jensen and Kletzer (2006) empirically implement the idea of Helpman and Krugman (1985) and Krugman (1991) with very detailed regional and sectoral data for the United States in order to understand the scope and impact of services offshoring. They use locational Gini coefficients to measure the geographical concentration of different sectors and classify sectors with a Gini coefficient below 0.1 as non-tradable and all others as tradable.

This study applies the idea of distinguishing the tradable and non-tradable sectors with the help of measures of geographic concentration to the BS context. More specifically, it directly relies on the distinction of tradable and non-tradable sectors made by Jensen and Kletzer (2006).[5] While this approach largely confirms the basic division of tradable and non-tradable sectors by De Gregorio et al. (1994) it allows a much more detailed breakdown of the service sector. A large part of this sector is classified as non-tradable but an important proportion turns out to be among the tradable part of the economy. In particular, this is the case

[4] The model by Helpman and Krugman (1985) combines economies of scale with monopolistic competition. While in Section 2 the BS hypothesis was formulated under perfect competition it can easily be incorporated into a monopolistic competition model [cf. Frensch (2006)].

[5] Table VI in Appendix A.2 displays the sectoral breakdown used by this study and whether a sector is classified as tradable or non-tradable. Appendix A.2 also covers issues of mapping Jensen and Kletzer (2006)'s division of tradable and non-tradable sectors to the data sets used here.

for a number of financial or business services.

3.3. *Pooled Mean Group Estimator*

The standard approach to empirically test the assumption of wage equalization between tradables and non-tradables would be to estimate a fixed or random effects model. However, these estimators have a number of important drawbacks: Most severely, the use of fixed or random effects models only leads to consistent estimates if slope parameters are homogenous across groups, an assumption that is very often inappropriate [cf. Pesaran and Smith (1995) or Phillips and Moon (2000)]. In contrast, the Pooled Mean Group Estimator introduced by Pesaran et al. (1999) and used by this study is consistent in the presence of slope heterogeneity as well as dynamic effects.

The Pooled Mean Group Estimator relies on an Autoregressive Distributed Lag Model [ARDL$(p, q, q, ..., q)$ model],

$$y_{i,t} = \sum_{j=1}^{p} \lambda_{i,j} y_{i,t-j} + \sum_{j=0}^{q} \boldsymbol{\delta'_{i,j}} \boldsymbol{x_{i,t-j}} + \mu_i + \epsilon_{i,t}. \tag{3.1}$$

Here $t = 1, 2, ..., T$ identifies time periods and $i = 1, 2, ..., N$ groups; y is the dependent and $\boldsymbol{x_{i,t}}(k \times 1)$ the vector of independent variables for group i; the coefficients of the lagged dependent variable, λ_{ij}, are scalars while those of the independent variables, $\boldsymbol{\delta_{i,j}}$, are $k \times 1$ vectors; μ_i is the group-specific effect and $\epsilon_{i,t}$ an i.i.d. error term.

Reparametrization and stacking of time-series observations for each group yields the error correction formulation of equation (3.1),

$$\Delta \boldsymbol{y_i} = \phi_i(\boldsymbol{y_{i,-1}} - \boldsymbol{X_i}\boldsymbol{\theta_i}) + \sum_{j=1}^{p-1} \lambda^*_{i,j} \Delta \boldsymbol{y_{i,-j}} + \sum_{j=0}^{q-1} \Delta \boldsymbol{X_{i,-j}} \boldsymbol{\delta^*_{i,j}} + \mu_i \boldsymbol{\iota} + \boldsymbol{\epsilon_i}, \quad i = 1, 2, ..., N, \tag{3.2}$$

where the error-correction coefficient $\phi_i = -(1 - \sum_{j=1}^{p} \lambda_{i,j})$ gives the speed of adjustment to the long-run equilibrium (negative if a long-run relationship between $y_{i,t}$ and $\boldsymbol{x_{i,t}}$ exists). $\boldsymbol{\theta_i} = -(\sum_{j=0}^{q} \boldsymbol{\delta_{i,j}})/\phi_i$ is the vector of long-run coefficients on $\boldsymbol{X_i}$. Furthermore, $\boldsymbol{\iota} = (1, ..., 1)'$ is a $T \times 1$ vector of ones, $\Delta \boldsymbol{y_i} = \boldsymbol{y_i} - \boldsymbol{y_{i,-1}}$, $\Delta \boldsymbol{X_i} = \boldsymbol{X_i} - \boldsymbol{X_{i,-1}}$, $\lambda^*_{i,j} = -\sum_{m=j+1}^{p} \lambda_{i,m}$ and $\boldsymbol{\delta^*_{i,j}} = -\sum_{m=j+1}^{q} \boldsymbol{\delta_{i,m}}$.

The peculiarity of the Pooled Mean Group Estimator is that it constrains long-run coefficients to be identical across groups,

$$\boldsymbol{\theta_i} = \boldsymbol{\theta}, \quad i = 1, 2, ..., N. \tag{3.3}$$

At the same time, the Pooled Mean Group Estimator allows intercepts and short-run marginal effects to differ freely. Pesaran et al. (1999) show that contrary to fixed or random effects models it is consistent in the presence of dynamic effects and slope heterogeneity irrespective of whether relevant variables have a unit root (under the assumption that a cointegration relationship exists) or not.[6]

An alternative strategy might have been to use the so-called Mean Group Estimator proposed by Pesaran and Smith (1995). This estimator calculates separate equations across groups and examines the distribution of the estimated coefficients across groups. It does not rely on the constraint of equation (3.3) but is also less efficient if long-run coefficients are indeed identical across groups. Pesaran et al. (1999, p. 621) note that "(t)here are often good reasons to expect the long-run equilibrium relationships between variables to be similar across groups. (...) The reasons for assuming that short-run dynamics (...) should be the same tend to be less compelling." In the context of this study it could for instance be the case that arbitrage costs prevent wage equalization across sectors of production in the short but not in the long run. In any case, in the next section outputs of Hausman tests for the null hypothesis of no difference between Pooled Mean Group and Mean Group estimates will always be reported.

4. RESULTS

4.1. *Internal Transmission Mechanism*

Following Strauss (1997) this study proceeds in two steps. First, it estimates the relationship between productivity and wages in the tradable and non-tradable sectors inside the sample countries. Then it turns to the international relations between productivity, wages and comparative prices.

The choice of the exact estimation equations for the internal relationship between productivity and wages in the tradable and non-tradable sectors is motivated by Equations 2.4 and 2.5 of Section 2. These define w_T and w_{NT} in terms of p_T and a_T which are considered exogenous in the the context of the theoretical considerations if Section 2. Because of the assumed exogeneity of p_T and a_T regressions of wages in the tradable sector on productivity in this sector and of wages in the non-tradable sector on productivity in the tradable sector should lead to consistent estimates of the coefficients associated with a_T even if p_T is not included as a regressor. If the theoretical considerations outlined in Section 2 were correct one would expect these coefficients to be positive and statistically significant.[7] At the same time, a regression of wages in the non-tradable sector on productivity in this sector should not lead to a statistically significant coefficient.

Table I summarizes Pooled Mean Group estimation results for the relations between (1) wages in the tradable sector and productivity in this sector, (2)

[6]The Pooled Mean Group estimations were carried out with a Gauss program available on Hashem Pesaran's website at http://www.econ.cam.ac.uk/faculty/pesaran/jasa.exe.

[7]If one takes the simplest version of the BS hypothesis outlined in Section 2 literally they should even be equal to one.

TABLE I

Internal transmission mechanism [model: ARDL(1,1)]

	(1)	(2)	(3)
regressand	w_T	w_{NT}	w_{NT}
long-run coefficients			
a_T	1.191***		0.476***
	(0.028)		(0.031)
a_{NT}		0.604***	
		(0.021)	
error correction coefficients	-0.324***	-0.230***	-0.335***
	(0.080)	(0.084)	(0.066)
short-run coefficients			
a_T	0.386***		0.160***
	(0.080)		(0.032)
La_T	0.363***		0.446***
	(0.086)		(0.085)
a_{NT}		0.139***	
		(0.051)	
La_{NT}		0.680***	
		(0.078)	
intercept	0.009	0.017	0.040
	(0.020)	(0.038)	(0.046)
Hausman test statistic	1.35	0.15	2.65
	[0.25]	[0.70]	[0.42]
sample size	290	290	290
cross-sections/time	29/1995–2005	29/1995–2005	29/1995–2005

Notes: Asymptotic standard errors in parentheses, p-values in brackets. * , (**), (***) indicates significance at the 10, (5), (1) per cent level, L the lag parameter. For a detailed description of variables used see Section 3.1.

wages in the non-tradable sector and productivity in this sector and (3) wages in the non-tradable sector and productivity in the tradable sector. For these and all further regressions logarithms were taken of all variables and cross-section demeaned data was used as suggested by Pesaran et al. (1999). For the Pooled Mean Group estimations the initial estimates of the long-run parameters were obtained from the Mean Group estimates, the Newton-Raphson method (which uses both the first and second derivatives of the log-likelihood function) was chosen for the numerical algorithms and both p and q were set to one.

For all three estimations summarized in Table I error correction coefficients are significant and have negative signs. This is evidence in favor of the dynamic stability of the empirical models. Moreover, Hausman tests with the null hypothesis of no difference between the Pooled Mean Group and Mean Group Estimator are rejected in all cases. Thus the imposition of long-run homogeneity does not appear to be a problem and in general the empirical models appear to be well-specified.

Concerning the individual regressions, Columns (1) and (3) of Table I show significant and strong positive relationships between wages and productivity inside the tradable sector and between wages in the non-tradable sector and productivity in the tradable sector. These results are in accordance with the BS hypothesis. This is not the case when it comes to the output reported in Column (2) which makes it clear that the relationship between wages and productivity inside the non-tradable sector seems to be close as well, with a long-run elasticity of about one half and a high level of statistical significance. All in all, the empirical model cannot fully confirm the relationships between productivity and wages in the tradable and non-tradable sectors inside the sample countries postulated by the BS hypothesis.

4.2. *External Transmission Mechanism*

Now the focus is shifted to the international relations between productivity, wages and comparative prices described in Section 2. More specifically, Pooled Mean Group estimates are computed for a variant of equation (2.8), where q is the regressand and productivity components are combined into one variable to save degrees of freedom:

$$\mathcal{A} \equiv \ln\left(\frac{A_T^F A_{NT}^H}{A_T^H A_{NT}^F}\right) \tag{4.1}$$

In addition to $\mathcal{A}$, a second regressor is included which is modeled after $\mathcal{A}$ and consists of the corresponding wage components:

$$\mathcal{W} \equiv \ln\left(\frac{W_T^F W_{NT}^H}{W_T^H W_{NT}^F}\right) \tag{4.2}$$

One issue here is the choice of a benchmark. All relevant studies mentioned in Section 1 more or less *ad hoc* choose one country as their benchmark (mostly the United States). In contrast Betts and Kehoe (2008) and Frensch and Schmillen

TABLE II

External transmission mechanism [model: ARDL(1,1,1)]

regressand	q
long-run coefficients	
$\mathcal{A}$	-1.187***
	(0.042)
$\mathcal{W}$	0.408***
	(0.042)
error correction coefficients	-0.911***
	(0.017)
short-run coefficients	
$\mathcal{A}$	-1.081***
	(0.020)
L$\mathcal{A}$	-0.050**
	(0.022)
$\mathcal{W}$	0.372***
	(0.007)
L$\mathcal{W}$	0.203
	(0.018)
intercept	-0.017
	(0.015)
joint Hausman test statistic	4.01
	[0.11]
sample size	4060
cross-sections/time	406/1995–2005

Notes: Asymptotic standard errors in parentheses, p-values in brackets. * , (**), (***) indicates significance at the 10, (5), (1) per cent level, L the lag parameter. For a detailed description of variables used see Section 3.1.

(forthcoming) do not use one benchmark economy but instead evaluate all bilateral country pairs in their respective samples. This study also relies on such a bilateral approach and is the first to use it for an evaluation of whether labor markets are homogenous across tradables and non-tradables, dramatically expanding the scope of the empirical investigation.

Table II shows Pooled Mean Group estimates for a regression of comparative prices on $\mathcal{A}$ and $\mathcal{W}$ for 406 country pairs. If labor markets were homogenous real exchange rates should depend solely on productivity differentials between the tradable and non-tradable sectors and in a regression of comparative prices on $\mathcal{A}$ and $\mathcal{W}$ the latter should not be associated with a statistically significant coefficient.

Instead, in both the short and the long run $\mathcal{W}$ is *ceteris paribus* significantly associated with bilateral comparative prices. So this study's estimation of the BS effect's external transmission mechanism rejects the assumption of wage equal-

TABLE III

INTERNAL TRANSMISSION MECHANISM

	(1)	(2)	(3)	(4)
regressand		w_T		
long-run coefficients				
a_T	1.191***	0.698***	1.267***	1.211***
	(0.028)	(0.211)	(0.022)	(0.023)
regressand		w_{NT}		
long-run coefficients				
a_{NT}	0.604***	0.738***	0.312**	0.974***
	(0.021)	(0.173)	(0.151)	(0.012)
regressand		w_{NT}		
long-run coefficients				
a_T	0.476***	0.654***	0.520***	1.062***
	(0.031)	(0.161)	(0.092)	(0.031)
	equivalent to column (3) of table I	*dynamic fixed effects estimator*	*1970 – 2005*	model: SIC

Notes: Asymptotic standard errors in parentheses [in Column (2) corrected for possible heteroscedasticity]. Sample size: 290 [in Column (3) 630]. * , (**), (***) indicates significance at the 10, (5), (1) per cent level. For a detailed description of variables used see Section 3.1.

ization across sectors of production.

Two other points should also be noted. First, it turns out that the long-run coefficient associated with $\mathcal{A}$ is negative and both statistically and economically extremely significant. This result is perfectly in line both with the stylized fact that at the going exchange rate aggregate price levels are generally higher in richer than in poorer economies and with the basic BS hypothesis. Second, for the estimation reported in Table II the error correction coefficient is significant and has the expected negative sign. As for all regressions discussed in the last subsection, a Hausman test cannot reject the null hypothesis of no difference between the Pooled Mean Group and the Mean Group estimates.

4.3. *Alternative Specifications*

This section reports the outcomes of a number of checks that evaluate whether the results presented above are robust to variations of the empirical setup. Tables III and IV summarize long-run coefficients for the baseline approach from the last subsection as well as for a number of alternative specifications. The first columns of Tables III and IV repeat Table I and Table II, respectively, while in Columns (2) to (4) results are reported for the following alternative specifications:

First, a dynamic fixed effects estimator is used instead of the Pooled Mean

TABLE IV

External transmission mechanism

	(1)	(2)	(3)	(4)
regressand		q		
long-run coefficients				
$\mathcal{A}$	-1.187***	-1.126***	-1.049***	-1.264***
	(0.042)	(0.048)	(0.036)	(0.025)
$\mathcal{W}$	0.408***	0.450***	0.132***	0.596***
	(0.042)	(0.058)	(0.024)	(0.025)
	equivalent to table II	*dynamic fixed effects estimator*	*1970 – 2005*	model: SIC

Notes: Asymptotic standard errors in parentheses [in Column (2) corrected for possible heteroscedasticity]. Sample size: 4060 [in Column (3) 5355]. * , (**), (***) indicates significance at the 10, (5), (1) per cent level. For a detailed description of variables used see Section 3.1.

Group Estimator. While Section 3.3 argued that using the Pooled Mean Group Estimator offers several advantages this might serve as a robustness check as to whether the results concerning the internal and external transmission mechanisms of labor market homogeneity reported in the last subsections are sensitive to the choice of the estimator.[8]

Second, estimations are repeated for an alternative sample consisting of 18 countries for the time span 1970 to 2005.[9] This entails the exclusion of all Central and Eastern European economies but provides much longer time series.

Finally, instead of setting p and q equal to one the lag structure is determined by minimizing the value of the Schwarz Information Criterion [SIC, cf. Schwarz (1978)] for each country, subject to a maximum lag of one. This exercise is meant to test whether results are robust to the order of the ARDL model.

As Table III shows, results for the BS effect's internal transmission mechanism are qualitatively very robust to the alternative specifications presented here. Most importantly, the internal relationship between productivity in the non-tradable sector and wages in this sector stays statistically positive and significant for all alternative specifications. The same is also true for the relationship between wages and productivity inside the tradable sector and for wages in the non-tradable sector and productivity in the tradable sector.

[8]One issue with the Pooled Mean Group Estimator is that its favorable asymptotic properties require $N \to \infty$ as well as $T \to \infty$. In contrast, the popular "Difference" and "System" GMM estimators have good asymptotic properties for $N \to \infty$ without requiring $T \to \infty$. At the same time, they suffer from a number of drawbacks: First, they are generally inconsistent in the presence of nonstationary time series. Second, they necessitate the validity of moment conditions that are often questionable. Third, results are often very sensitive to the choice of the exact specification (like the number of instruments). Still, I experimented with "Difference" and "System" GMM estimations. As could have been expected, results were unstable.

[9]For a list of countries for which consistent data are available from 1970 to 2005 see Appendix A.1.

Results for the external transmission mechanism of labor market homogeneity are as robust as those for the internal transmission mechanism. $\mathcal{W}$ always exhibits a significantly positive long-run coefficient whose order of magnitude also stays remarkably similar throughout all specifications [except for the one reported in Column (3) of Table III]. What is more, statistical significance and size of short-run and error correction coefficients (not reported here) remain largely unchanged as does the relationship between $\mathcal{A}$ and bilateral comparative prices which is negative and highly significant throughout all specifications.

4.4. *Unit Root Tests*

One final concern is that all results presented so far might be the outcome of "spurious regressions" [cf. Granger and Newbold (1974)]. While the Pooled Mean Group Estimator is consistent in the presence of unit roots (under the assumption that a cointegration relationship exists) it has so far not been tested if relevant variables are non-stationarity and — if yes — whether they are cointegrated.

Unfortunately, testing for unit roots seem unfeasible for this study's benchmark sample because of its relatively short length. Instead, Table V reports the results of panel unit root tests for the alternative sample already mentioned in the last subsection which consists of 18 countries for the time span 1970 to 2005. For all relevant variables four different tests are employed: One by Levin et al. (2002) that relies on the assumption of a common unit root process across cross-sections and three [one developed by Im et al. (2003) and two Fisher-type tests based on the popular Augmented Dickey-Fuller and Phillips-Perron unit root tests, respectively] that allow for individual unit root processes.[10]

For all tests the null hypothesis is the existence of a unit root. As Table V shows this null hypothesis can always be overwhelmingly rejected. Thus all relevant variables appear to be stationary and there seems to be little need to worry about whether results from the last subsections are caused by "spurious regressions".

5. CONCLUSION

This study introduced a novel, theory-based method to distinguish the tradable and non-tradable sectors to the BS literature. It used this distinction, modern empirical methods and a large and detailed macro data set to assess an assumption that is crucial in the BS context and common in many multisector (open-economy) macro models, namely the homogeneity of labor markets across sectors of production. It found that this assumption cannot be confirmed empirically across tradable and non-tradable goods.

[10]For the Levin-Lin-Chu, Im-Pesaran-Shin and Augmented Dickey-Fuller panel unit root tests one lag difference term is always included, for the Levin-Lin-Chu and Phillips-Perron tests a Bartlett kernel type and the Newey-West method for bandwidth selection are used. Besides, individual effects are always included as exogenous variables. Results are qualitatively very robust to changes in the exact test specifications, for instance to the inclusion of individual trends.

TABLE V
Unit root tests

variable	Levin-Lin-Chu	Im-Pesaran-Shin	Augmented Dickey-Fuller	Phillips-Perron
w_T	-4.633***	-5.397***	91.923***	48.768**
w_{NT}	-4.265***	-5.831***	94.622***	57.533***
a_T	-4.183***	-5.253***	89.064***	47.992*
a_{NT}	-4.231***	-5.718***	92.145***	54.446**
q	-32.783***	-48.715***	2614.405***	3298.491***
$\mathcal{W}$	-37.833***	-45.925***	2404.337***	3234.448***
$\mathcal{A}$	-37.512***	-52.678***	2880.324***	3852.922***

Notes: Null hypotheses: unit root. * , (**), (***) indicates significance at the 10, (5), (1) per cent level. Probabilities for Fisher-type tests are computed using an asymptotic Chi-square distribution. All other tests assume asymptotic normality. Individual effects are included as exogenous variables. For a detailed description of variables used see Section 3.1.

This result implies that multisector macro models should relinquish the assumption of homogenous labor markets. Instead, one might model segmented labor markets along the lines of the corresponding labor economics literature in order to paint a more realistic picture of wage interactions across sectors of production.

Concerning the BS hypothesis, this study's results have even more direct consequences. The hypothesis is very prominent and of high political relevance — for instance in connection with any conclusion as to whether a (former) transition economy is ready to join the eurozone. While this study fails to confirm wage equalization across tradable and non-tradable goods it does not reject an important role or even the existence of the BS effect. This calls for further theoretical investigations into the BS hypothesis and the exact role of the assumptions underlying it.

ACKNOWLEDGEMENTS

I thank Jennifer Abel-Koch, Jürgen Jerger and Michael Stops as well as conference and seminar participants in Kiel, Luxembourg and Regensburg for helpful comments and suggestions. Special thanks are due to Joachim Möller and Richard Frensch for their constant guidance and support throughout this project. Financial support from the German Research Foundation is gratefully acknowledged. The usual disclaimer applies.

REFERENCES

Balassa, Béla (1964). The Purchasing Power Parity Doctrine: A Reappraisal. *Journal of Political Economy* **72** 584–596.

Bergin, Paul; Glick, Reuven and Taylor, Alan (2006). Productivity, Tradability, and the Long-Run Price Puzzle. *Journal of Monetary Economics* **53** 2041–2066.

Bernard, Andrew and Jensen, Bradford (1995). Exporters, Jobs, and Wages in U.S. Manufacturing: 1976-1987. *Brookings Papers on Economic Activity: Microeconomics* **1995** 67-119.

Betts, Caroline and Kehoe, Timothy (2008). Real Exchange Rate Movements and the Relative Price of Non-traded Goods. *NBER Working Paper No. 14437*.

De Gregorio, José; Giovannini, Alberto and Wolf, Holger (1994). International Evidence on Tradables and Nontradables Inflation. *European Economic Review* **38** 1225–1244.

Dickens, William and Katz, Lawrence (1987). Inter-Industry Wage Differences and Theories of Wage Determination. *NBER Working Paper No. 2271*.

Égert, Balázs, Drine, Imed, Lommatzsch, Kirsten and Rault, Christophe (2003). The Balassa-Samuelson Effect in Central and Eastern Europe: Myth or Reality? *Journal of Comparative Economics* **31** 552–572.

Eichengreen, Barry and Gupta, Poonam (2009). The Two Waves of Service Sector Growth. *NBER Working Paper No. 14968*.

Frensch, Richard (2006). Balassa-Samuelson, Product Differentiation and Transition. *Osteuropa-Institut Regensburg Working Paper No. 266*.

Frensch, Richard and Schmillen, Achim (forthcoming). Can We Identify Balassa-Samuelson Effects with Measures of Product Variety? *Economic Systems*.

Ghironi, Fabio and Melitz, Marc (2005). International Trade and Macroeconomic Dynamics with Heterogeneous Firms. *The Quarterly Journal of Economics* **120** 865–915.

Granger, Clive and Newbold, Paul (1974). Spurious Regressions in Econometrics. *Journal of Econometrics* **2** 111–120.

Harrod, Roy (1933). *International Economics*, 2nd revised ed., Nisbet and Cambridge University Press, London and Cambridge.

Helpman, Elhanan and Krugman, Paul (1985). *Market Structure and Foreign Trade: Increasing Returns, Imperfect Competition, and the International Economy*, MIT Press, Cambridge.

Herrendorf, Berthold and Valentinyi, Akos (2006). The Balassa-Samuelson Hypothesis Forty Years Later. *Arizona State University, Department of Economics Working Paper No. 2180043*.

Im, Kyung So, Pesaran, Hashem and Shin, Yongcheol (2003). Testing for Unit Roots in Heterogeneous Panels. *Journal of Econometrics* **115** 53–74.

Jensen, Bradford and Kletzer, Lori (2006). Tradable Services: Understanding the Scope and Impact of Services Offshoring. In Brainard, L. and Collins, S. (Eds.). *Offshoring White-Collar Work – The Issues and the Implications*, Brookings Institution, Washington, DC.

Krueger, Alan and Summers, Lawrence (1987). Efficiency Wages and the Inter-industry Wage Structure. *Econometrica* **56** 259–293.

Krugman, Paul (1991). *Geography and Trade*, MIT Press, Cambridge.

Lee, Jeong-Joon (2005). Persistent Wage Differential and its Implications on the Balassa-Samuelson Hypothesis. *Applied Economics Letters* **15** 643–648.

Levin, Andrew, Lin, Chien-Fu and Chu, Chia-Shang (2002). Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties. *Journal of Econometrics* **108** 1–24.

Pesaran, Hashem and Smith, Ron (1995). Estimating Long-Run Relationships from Dynamic Heterogeneous Panels. *Journal of Econometrics* **68** 79–11.

Pesaran, Hashem, Shin, Yongcheol and Smith, Ron (1999). Pooled Mean Group Estimation of Dynamic Heterogeneous Panels. *Journal of the American Statistical Association* **94** 621–634.

Phillips, Peter and Moon, Hyungsik (2000). Nonstationary Panel Data Analysis: An Overview of Some Recent Developments. *Econometrics Reviews* **19** 263–286.

Samuelson, Paul (1964). Theoretical Notes on Trade Problems. *Review of Economics and Statistics* **46** 145–154.

Schwarz, Gideon (1978). Estimating the Dimension of a Model. *Annuals of Statistics* **6** 461–464.

Strauss, Jack (1997). The Influence of Traded and Nontraded Wages on Relative Prices and

Real Exchange Rates. *Economics Letters* **55** 391–395.
Strauss, Jack (1998). Relative Price Determination in the Medium Run: The Influence of Wages, Productivity and International Prices. *Southern Economic Journal* **65** 223–244.
Strauss, Jack and Ferris, Mark (1996). The Role of Nontraded and Traded Wages in the Productivity Differential Model. *Southern Economic Journal* **63** 327–338.
Summers, Robert and Heston, Alan (1996). The Penn World Table (Mark 5): An Expanded Set of International Comparisons. *The Quarterly Journal of Economics* **106** 327–368.
Tica, Josip and Družić, Ivo (2006). The Harrod-Balassa-Samuelson Effect: A Survey of Empirical Evidence. *University of Zagreb, Faculty of Economics and Business, Working Paper No. 06-07*.
Timmer, Marcel, O'Mahony, Mary and van Ark, Bart (2007). The EU KLEMS Growth and Productivity Accounts: An Overview. *International Productivity Monitor* **14** 71–85.

APPENDIX A: DATA COVERAGE

A.1. *Countries Covered*

This study covers the following OECD and Central and Eastern European countries for which consistent data from 1995 to 2005 are available both in the EU KLEMS data base and in the Penn World Tables (* indicates that consistent data are available from 1970 to 2005): Australia*; Austria*; Belgium*; Cyprus; Czech Republic; Denmark*; Estonia; Finland*; France*; Germany*; Greece*; Hungary, Ireland*; Italy*; Japan; Korea (Republic of)*; Latvia; Lithuania; Luxembourg*; Malta; Netherlands*; Poland; Portugal*; Slovakia; Slovenia; Spain*; Sweden*; United Kingdom*; United States*.

A.2. *Sectors Covered*

Table VI list the sectors covered by this study. The sectoral breakdown is based on a two-digit NACE classification (Nomenclature statistique des activités économiques dans la Communauté Européenne / Statistical Classification of Economic Activities in the European Community).

Table VI also reports whether based on table 4 ("Share of Total Employment by Tradable/Non-Tradable") of Jensen and Kletzer (2006) a sector is classified as tradable or non-tradable. Not all two-digit NACE sectors are covered due to limited data availability, difficulties of mapping NAICS [the North American Industry Classification System used by Jensen and Kletzer (2006)] to NACE or almost equal tradability and non-tradability scores reported by Jensen and Kletzer (2006) [cf. Eichengreen and Gupta (2009) who also map Jensen and Kletzer (2006)'s division of traded and non-traded sectors to the NACE classification].

TABLE VI

NACE Sectors covered

NACE code	description	tradability
1	agriculture	tradable
2	forestry	tradable
B	fishing	tradable
10	mining of coal and lignite; extraction of peat	tradable
11	extraction of crude petroleum and natural gas	tradable
12	mining of uranium and thorium ores	tradable
13	mining of metal ores	tradable
14	other mining and quarrying	tradable
15	food and beverages	tradable
16	tobacco	tradable
17	textiles	tradable
18	wearing apparel, dressing and dying of fur	tradable
19	leather, leather products and footwear	tradable
20	wood and products of wood and cork	tradable
21	pulp, paper and paper products	tradable
22	printing, publishing and reproduction	tradable
23	coke, refined petroleum and nuclear fuel	tradable
24	chemicals and chemical products	tradable
25	rubber and plastic products	tradable
26	other non-metallic mineral products	tradable
27	basic metals	tradable
28	fabricated metals	tradable
29	machinery not elsewhere covered (n.e.c.)	tradable
30	office, accounting and computing machinery	tradable
31	electrical machinery and apparatus n.e.c.	tradable
32	radio, television and communication equipment	tradable
33	metal, precision and optical instruments	tradable
34	motor vehicles, trailers and semi-trailers	tradable
35	other transport equipment	tradable
36	manufacturing n.e.c.	tradable
37	recycling	tradable
40	electricity and gas	non-tradable
41	water supply	non-tradable
F	construction	non-tradable
H	hotels and restaurants	non-tradable
64	post and telecommunications	tradable
65	financial intermediation, except insurance and pension funding	tradable
66	insurance and pension funding, except compulsory social security	tradable
67	activities related to financial intermediation	tradable
70	real estate activities	tradable
71	renting of machinery and equipment	tradable
L	public administration and defence; compulsory social security	non-tradable
M	education	non-tradable
N	health and social work	non-tradable
90	sewage and refuse disposal, sanitation and similar activities	non-tradable
91	activities of membership organizations n.e.c.	non-tradable
92	recreational, cultural and sporting activities	non-tradable
93	other service activities	non-tradable
P	private households with employed persons	non-tradable

Notes: Tradability based on table 4 ("Share of Total Employment by Tradable/Non-Tradable") of Jensen and Kletzer (2006); sectors 50 (sale, maintenance and repair of motor vehicles and motorcycles; retail sale of fuel), 51 (wholesale trade and commission trade, except motor vehicles and motorcycles), 60 (other inland transport), 61 (other water transport), 62 (other air transport), 63 (other supporting and auxiliary transport activities; activities of travel agencies), 72 (computer and related activities), 73 (research and development), 74 (other business activities) and Q (extra-terrestrial organizations and bodies) are not covered.