Wirtschaftswissenschaftliches Zentrum (WWZ) der Universität Basel

December 2011

The Balassa-Samuelson Effect Reversed: New Evidence from OECD Countries

WWZ Discussion Paper 2011/09

Matthias Gubler, Christoph Sax

UNI BASEL

The Authors:

Matthias Gubler
Faculty of Business and Economics
Monetary Macroeconomics
University of Basel
Peter Merian-Weg 6
CH - 4002 Basel
phone: +41(0)61 267 33 56

matthias.gubler@unibas.ch

Christoph Sax, Dr.
Faculty of Business and Economics
International Trade and European Integration
University of Basel
Peter Merian-Weg 6
CH - 4002 Basel
phone: +41(0)61 267 33 74

c.sax@unibas.ch

A publication of the Center of Business and Economics (WWZ), University of Basel.

The Balassa-Samuelson Effect Reversed: New Evidence from OECD Countries*

Matthias Gubler   Christoph Sax

Abstract

This paper explores the robustness of the Balassa-Samuelson (BS) hypothesis. We analyze a panel of OECD countries from 1970 to 2008 and compare three different datasets on sectoral productivity, including a newly constructed database on total factor productivity. Overall, our DOLS estimation results do not support the BS hypothesis. For the last two decades, we find a very robust negative relationship between the productivity in the tradable sector and the equilibrium real exchange rate, in contrast to BS. Earlier supportive findings depend strongly on the choice of the dataset. Except for the terms of trade, the explanatory power of other variables is weak.

JEL Classifications: F14, F31, F41
Keywords: Real Exchange Rate, Balassa-Samuelson Hypothesis, Panel Data Estimation, Terms of Trade

*Without implication, we would like to thank Markus Ludwig, Christian Rutzer, Linda Walter, Rolf Weder and seminar participants at the University of Basel for extensive comments and suggestions; and the OECD for kindly sharing two essential datasets on sectoral productivity. Hermione Miller-Moser provided excellent editorial assistance. Matthias Gubler is indebted to Peter Kugler for his help and supervision. Christoph Sax acknowledges support from the WWZ Forum Foundation.

1 Introduction

The Balassa-Samuelson (BS) hypothesis is one of the most widespread explanations for structural deviations from purchasing power parity (PPP) (Dornbusch, 1985).[1] The BS hypothesis was stated by both Balassa (1964) and Samuelson (1964) simultaneously, but has a research precedent in the work of Harrod (1933). According to the hypothesis, differences in the productivity differential between the non-tradable and the tradable sector lead to differences in price levels between countries, when converted to the same currency.[2] Ceteris paribus, a productivity increase in tradables raises factor prices, i.e., wages, which in turn leads to higher prices of non-tradables and thus to an appreciation of the real exchange rate. In contrast, when the relative productivity of non-tradables increases, marginal cost cuts result in a lower price level.

The empirical evaluation of the BS hypothesis has gained a great deal of attention. As stated in a survey by Tica and Družić (2006), the major share of the evidence supports the BS model, but the strength of the results depends on the nature of the tests and set of countries analyzed.

There are several studies based on a disaggregation of the tradable and non-tradable sector that find empirical support for the BS hypothesis (see, e.g., Lee et al. (2008), Choudhri and Khan (2005) or Calderón (2004)). In particular, since sector-specific data for OECD countries on total factor productivity (TFP) has become available, various studies have tested and confirmed the BS hypothesis using OECD panel data (MacDonald and Ricci, 2007; Chinn and Johnston, 1996; De Gregorio and Wolf, 1994). All these studies are based on the discontinued International Sectoral Database (ISDB) provided by the OECD. According to this literature, countries with higher productivity differentials between tradables and non-tradables exhibit

[1] According to the absolute PPP theory, a unit of currency must have the same purchasing power in the foreign economy as in the domestic economy, once it is converted into the same currency. However, in his seminal work establishing the so-called PPP puzzle, Rogoff (1996) argues that the speed of adjustment of real exchange rates is too slow to be in line with the PPP theory. Recent studies that stress the importance of nonlinear adjustments (Taylor, 2003) or dynamic aggregation bias (Imbs et al., 2005) challenge this finding. Indeed, there is empirical evidence for the failure of PPP, although the results are mixed (for reviews see Taylor (2003) or Froot and Rogoff (1996)).

[2] The hypothesis assumes that the law of one price for tradable goods holds. However, there are empirical contributions to the literature that find deviations from this law (see, e.g., Engel (1999)) as well as theoretical explanations for these deviations (for a discussion see MacDonald and Ricci (2007) and the references therein).

higher price levels expressed in a single currency, i.e., a stronger real exchange rate.

This paper estimates the long-run relationship between the real exchange rate and key explanatory variables, including the TFP differential between tradables and non-tradables. We apply a panel cointegration model that manages to treat the non-stationarity of the variables correctly. Recently, the OECD has provided a new database called PDBi with sector-specific TFP data from 1985 to 2008.

With this new dataset, our estimations cannot confirm the findings of previous research based on the ISDB.[3] In fact, the results point to a negative relationship between tradable productivity and the real exchange rate. This finding is the *opposite* of what is claimed by the BS hypothesis. Furthermore, we can confirm this result when TFP is replaced by labor productivity (LP) using the OECD Structural Analysis (STAN) database, which covers more countries and a longer time period, from 1970 to 2008. On the other hand, the connection between non-tradable productivity and the real exchange rate is not robust. Finally, with the exception of the terms of trade, our estimation results indicate that the explanatory power of control variables discussed in the literature is weak or not robust.

In order to detect the causes of the conflicting results, we systematically compare the three datasets and their implications regarding the estimation results. Our robustness tests reveal that severe outlier dependency exists for the traditional pro-Balassa-Samuelson finding regarding *non-tradables*. In particular, Japanese labor productivity in the non-tradable sector strongly weakens the estimated BS effect. For the time period from 1970 to 1992, the coefficient even significantly changes its sign once Japan is included, a finding that is robust against several variations of the model specification. For the time period after 1992, the connection strongly depends on wether the United States are included in the country sample.

However, the negative relationship between the productivity of *tradables* and the real exchange rate is very robust. A rigorous analysis of the tradable sector reveals that this reversal is robust for the last two decades against the choice of the country-sample, the precise start of the sample period, the exact model specification, and the inclusion of additional explanatory

[3]However, we are indeed able to replicate the results in favor of the BS hypothesis with data from the ISDB.

variables. In Gubler and Sax (2011), we provide a static general-equilibrium framework with skill-based technological change (SBTC) that can explain the new relationship. We show that SBTC may increase the relative amount of physical capital used by the tradable industry, thus releasing low-skilled labor; and finally, leading to lower wages for low-skilled labor and lower prices in the economy. An increase in tradable productivity may thus be connected to a lower real exchange rate.

Overall, we conclude from the results of our analysis that the presence of the Balassa-Samuelson effect is not robust for a panel of major OECD countries. In fact, during the last two decades an increase in the productivity of tradables, all else being equal, has given rise to a depreciation of the real exchange rate.

The remainder of this paper is organized as follows. Section 2 presents the data. We outline the methodology in Section 3 and show the results in Section 4. Section 5 concludes.

2 Data

The data for the 18 major OECD countries included in our dataset stem from different databases of the IMF, OECD, World Bank and the Penn World Tables. Depending on the estimation, the country-sample has to be reduced, because we aim to replicate the results of MacDonald and Ricci (2007) or because not all data are available.[4]. A detailed description of all variables is given in Table 1 and in Appendix A.2.

In order to test the BS hypothesis, we condition the real exchange rate on productivity measures for both the tradable and the non-tradable sector, as well as on control variables. The choice of the dependent variable is discussed in Section 2.1. Due to its importance and complexity, productivity data are separately examined in Section 2.2. All other exogenous variables are discussed in Section 2.3. The time series properties of the variables are assessed in the final Section 2.4.

2.1 Dependent Variable: Real Exchange Rate

We use the logarithm of the unweighted real exchange rate (RER) as the dependent variable in our estimation equations. In principle, the real ex-

[4] All country-samples featured in our estimations are presented in Appendix A.1

Table 1: Description and Construction of the Variables

Abbr.	Name	Definition	Source
RER	Real Exchange Rate	log(CPI / Nominal Exchange Rate to USD)	IMF, IFS
TFP_T	TFP of Tradables	Solow Residual	OECD, EO
TFP_NT	TFP of Non-Tradables	Solow Residual	OECD, EO
LP_T	LP of Tradables	log(Value Added / Hours-Worked)	OECD, EO
LP_NT	LP of Non-Tradables	log(Value Added / Hours-Worked)	OECD, EO
CA	Current Account	as % of GDP	OECD, EO
$DPOP$	Population Growth	Δ log(Population)	PWT
GDP	Real GDP per capita	log(Real GDP per capita)	PWT
GOV	Government Spending	as % of GDP	OECD, EO
NFA	Net Foreign Assets	as % of GDP	WB, WDI
RI	Long-Term Real Int. Rate	Gov. bond yield long term - CPI	IMF, IFS
TOT	Terms of Trade	log(Export-Prices / Import-Prices)	OECD, EO

change rate can only be computed towards a reference country. However, instead of defining such a reference country, we circumvent the problem by using time fixed effects throughout our analysis. The real exchange rate can thus be thought of as a deviation from the sample mean, which, in this case, acts as the reference country ("average OECD country"). Proceeding this way allows us to keep as many observations as possible.

An extensive body of the empirical literature uses *effective* real exchange rates (see, e.g., Lee et al. (2008), Calderón (2004) or De Gregorio and Wolf (1994)), that are weighted by the share of exports. Effective real exchange rates have the advantage that there is no need to specify a specific reference country. While effective real exchange rates are a useful measure for competitiveness, the share of exports seems not only irrelevant in our context but also misleading. If, for example, a country changes its export destinations to countries with a weaker real exchange rate, effective real exchange rates would indicate a real appreciation, while, in fact, the country still has the same relative price level towards all countries.

2.2 Productivity Data

We use data on sectoral productivity from three datasets provided by the OECD: The first is a new dataset on sectoral total factor productivity (TFP) computed by the OECD, called PDBi (Productivity Database, i represents the specific sector). The dataset contains annual sector-specific TFP numbers and covers the time period from 1985 to 2008. Sectoral TFP is calculated as Solow residuals by the same method for all countries, using sectoral data on production, employment, capital stock and the labor share of income. Capital stocks are estimated applying the permanent inventory method, where streams of investments are added, and a certain fraction of depreciation is subtracted each year.

A second database, STAN, includes yearly data on sectoral production and employment, and thus on labor productivity, but not on capital stock or TFP. As the only dataset, STAN covers a long time range from 1970 to 2008 for many OECD countries.

In order to compare our findings with existing studies, in particular with the results of MacDonald and Ricci (2007), sectoral productivity data from the discontinued International Sectoral Database (ISDB) have been used as well. This old database contains annual values on labor and total factor productivity, in principle from 1970 to 1997, but has been discontinued before 1997 for most countries.

STAN and PDBi data are improvements to the ISDB. In the old dataset, output, employment and capital stocks were based on data from an old system of national accounts, SNA68. For social services, these changes in the measurement of output may have been especially important, as estimates of real value added growth for the public sector in the ISDB have simply been based on labor inputs, so that estimates of productivity had a very limited meaning. Moreover, in the ISDB, volumes were calculated using constant prices instead of chained linking. Finally, capital stock estimates may have been calculated differently and in a non-standardized way in the ISDB.[5]

The classification of subsectors into tradable and non-tradable is made according to the following scheme: Agriculture, manufacturing and transport, storage and communications are classified as tradables; utilities (energy, gas and water), construction, and social services (community, social, personal services) as non-tradables. Our division of the subsectors into tradable or non-tradable sectors follows De Gregorio and Wolf (1994), who defined a subsector as tradable if its share of exports in the total production exceeds

[5]While these are very general observations about the evolution of the system of national accounts, it would be desirable if international organizations could provide more information about the changes over time. As in the present case, that would tremendously facilitate the task of replicating earlier results.

Figure 1: Sectoral Productivity Data Coverage

Notes: For each country, the first row describes the coverage span of the STAN database, the second of the PDBi and the third of the ISDB. If all six sectors are available, the line is drawn *black*, if some sectors are available, it is drawn *grey*. The STAN database covers the broadest range of the three databases.

10% and as non-tradable otherwise.[6] While no division has become standard in the field (Tica and Družić, 2006), studies based on data from OECD countries usually refer to the division proposed by De Gregorio and Wolf (1994) (see, e.g., MacDonald and Ricci (2007), Chinn and Johnston (1996)). Like MacDonald and Ricci (2007), we exclude distribution, mining and financial subsectors due to classification difficulties (MacDonald and Ricci, 2005), or data availability.

The tradable and non-tradable sectors, classified this way, are roughly equal in terms of value added. Each sector comprises 50% of the total value added produced by these six subsectors. Within the tradable sector, manufacturing is by far the largest subsector, representing 64% of value added, whereas both agriculture and transport, storage and communications amount to 11% and 24%, respectively. Among the non-tradables, social services (70%) outweigh construction (20%) and utilities (9%). Figure 1 displays the data availability in each of the three datasets.

Table 2 shows the correlations between the three datasets. LP and TFP values from the ISDB are similar to the two newer datasets only in the tradable subsectors. In the non-tradable sectors, the correlations are lower (construction and utilities) or virtually non-existent (social services). To a lesser extent, this is also true for employment and value added. Possible reasons for these divergences have been discussed earlier in this section. On the other hand, data from the PDBi on TFP are highly correlated with labor productivity from the STAN database. These correlations are present in all subsectors, although the values are somewhat lower in the non-tradable subsectors.

We consider TFP as the preferred measure for productivity. As pointed out by De Gregorio and Wolf (1994), average labor productivity grows much faster during economic downturns, and, hence, it is not a reliable indicator of sustainable productivity growth which can affect the economy in the medium or long term. Nevertheless, there are some advantages of LP, and we will use the measure to check the robustness of our TFP results.[7]

[6] Adjustments of the threshold value to 5% and 20% leave the division virtually unchanged (De Gregorio and Wolf, 1994).

[7] The advantages are summarized by Canzoneri et al. (1999): First, labor productivity data are available for more countries and over a longer time period than TFP numbers. Second, the calculation of LP figures does not require an estimation of the capital stock and the income share of labor, with both estimations likely to be imprecise. Third, the BS hypothesis holds for more technologies than the Cobb-Douglas production function

Table 2: Median Correlations across Subsectors

	AGR	IND	TSC	EGW	CST	SOC
PDBi (TFP), STAN (LP)	0.95	0.97	0.92	0.95	0.93	0.84
ISDB (TFP), STAN (LP)	0.90	0.91	0.93	0.75	0.76	**0.28**
ISDB (TFP), ISDB (LP)	0.99	0.98	0.98	0.96	0.94	0.97
ISDB (LP), STAN (LP)	0.90	0.88	0.88	0.72	0.77	**0.27**
ISDB (EMP), STAN (EMP)	0.91	0.98	0.91	0.89	0.99	**0.45**
ISDB (VA), STAN (VA)	0.91	0.95	0.89	0.72	0.93	**0.45**

Notes: The table contains median correlation coefficients between variables in the three datasets for all six subsectors. The values are based on all countries for which a correlation coefficient can be calculated. AGR: Agriculture; IND: Manufacturing; TSC: Transport, Storage and Communications; EGW: Energy, Gas and Water; CST: Construction; SOC: Community, Social, Personal Services. The first three rows show the median correlations between TFP from the PDBi or the ISDB and LP from the STAN database or the ISDB. Median correlations between LP from the STAN database and the ISDB are reported in the fourth row. The last two rows contain the median correlation values between EMP from the ISBD and the STAN database, and between VA from the same sources. Due to a very low number of time-overlapping observations, no comparison between the PDBi and the ISDB is presented.

2.3 Control Variables

We include several control variables along with the productivity variables in our estimations: The importance of the terms of trade (TOT) has been proposed to be an important determinant of the long-run real exchange rate (e.g., De Gregorio and Wolf (1994); Sax and Weder (2009)). An improvement in the terms of trade allows a country to raise its imports for a given amount of factor inputs in the export sector. Hence, a change in TOT may be interpreted analogous to a change in the productivity in the tradable sector (De Gregorio and Wolf, 1994).

Several authors point out the importance of demand-side factors for the determination of the long-run real exchange rate. Therefore, we consider the government spending share (GOV), net foreign assets (NFA), the current account (CA) and real GDP per capita (GDP) as control variables.

De Gregorio and Wolf (1994) show theoretically that an increase in government spending reduces the equilibrium real exchange rate if capital mobility across countries is restricted. This increase affects the relative price of tradable and non-tradable goods negatively, because government spending tends to fall more heavily on non-tradables. Hence, government spending is widely used as an additional explanatory variable (see, e.g., Lee et al. (2008),

generally employed to determine TFP.

Chinn and Johnston (1996) or Sax and Weder (2009)).

Private demand may affect the real exchange rate as well. It is likely that a higher income is associated with a higher demand for non-tradables. The associated rise in the relative price of non-tradables gives rise to a higher overall price level (De Gregorio and Wolf, 1994). Furthermore, trade deficits or surpluses could affect the demand for non-tradables, by increasing or decreasing the amount of tradables that are available for consumption. As a permanent trade deficit can only be sustained in the presence of net foreign assets, several authors have emphasized the importance either of the net foreign assets or the current account deficit for the determination of the real exchange rate (Lee et al., 2008; Lane and Milesi-Ferretti, 2004; Krugman, 1990).

Finally, two other macroeconomic variables, the real interest rate (RI), and the population growth rate ($DPOP$), are taken into account. Their importance for the determination of RER has been discussed in theoretical and empirical contributions to the literature. According to the theoretical model provided by Stein and Allen (1997), a higher real interest rate is associated with an appreciated long-run real exchange rate because of portfolio adjustments and capital inflows. Rose et al. (2009) show in an overlapping generation model that a country experiencing a decline in its fertility rate will also experience a real exchange rate depreciation.

2.4 Assessing the Time Series Properties of the Variables

The panel unit root tests proposed by Levin et al. (2002) (LLC) and Im et al. (2003) (IPS) have been conducted for all variables (Table 3). In order to obtain reliable results, the test statistics are based on all available information for both time and cross-sectional dimensions. Thus, although the full length of a series is not used for the estimation, all observations are nevertheless used for the unit root test.

As described in Section 2.1, we do not compute the real exchange rate towards a specific country (or towards a basket of countries, such as in the effective real exchange rate). Instead, time-specific dummy variables are included in all estimations. In order to assess the times series properties, the real exchange rate is calculated towards the average of the sample (denoted RER_AVG).

Overall, we find strong evidence for non-stationary behavior for all vari-

Table 3: IPS and LLC Panel Unit Root Test Results

		Det. Trend	IPS	LLC	No. of Countries	Time Period	Obs.
CA			0.933	0.994	18	1970-2008	587
$DPOP$			-4.269***	-2.837***	18	1970-2007	626
GDP		x	1.010	1.591	18	1970-2007	656
GOV		x	3.091	0.130	18	1970-2008	632
NFA			3.920	5.589	18	1970-2006	611
RER_AVG		x	-1.172	-1.116	18	1970-2008	665
RI			-0.500	-0.331	18	1970-2008	621
TOT			0.233	0.214	18	1970-2008	640
LP_T	STAN	x	1.282	-1.540*	18	1970-2008	559
LP_NT	STAN	x	1.651	1.131	18	1970-2008	550
TFP_T	PDBi	x	-0.021	-1.537*	14	1985-2008	198
TFP_NT	PDBi	x	1.782	0.077	13	1985-2008	192
LP_T	ISDB	x	2.923	2.906	14	1970-1997	325
LP_NT	ISDB	x	1.909	1.103	14	1970-1997	322
TFP_T	ISDB	x	1.360	0.886	14	1970-1997	314
TFP_NT	ISDB	x	1.720	0.614	14	1970-1997	307

Notes: x indicates the inclusion of a deterministic trend. As all estimations contain time-specific dummy variables, the real exchange rate of each country is computed with respect to the average sample country for the unit root tests (RER_AVG). IPS: Lag length selection by modified SIC (Ng and Perron, 2001); LLC: Lag length selection by modified SIC; Bartlett kernel, Newey-West bandwidth. The panel is unbalanced: The time period marks the maximum years available */**/*** denotes significance at 10%, 5% and 1% level, respectively ables with the exception of the population growth rate, $DPOP$. As $DPOP$ is the first difference of the logarithm of population, this result is not surprising. The total factor productivity in the tradable sector (TFP_T) from the PDBi and labor productivity in the tradable sector (LP_T) from the STAN database show ambiguous results. However, non-stationarity of these variables is confirmed by the Fisher-type augmented Dickey-Fuller (ADF) panel unit root test proposed by Maddala and Wu (1999) (not shown) and is theoretically founded in macroeconomic models (see, e.g., Lindé (2009), Galí (1999) or King et al. (1991)). All results are also in line with the results found in similar empirical studies (see, e.g., Lee et al. (2008), MacDonald and Ricci (2007) or Calderón (2004)).

3 Methodology: Cointegration Tests and Panel DOLS

The number of observations for each country is limited, given the length of the sample (23 years in our benchmark model) and the annual data frequency. Therefore, we pool the data and apply a panel estimation technique to improve the power of our results. We are primarily interested in the long-run relationship between the real exchange rate and its determinants described in Section 2 and summarized in Table 1.

In order to estimate this relationship, we employ a panel cointegration model that treats the non-stationarity of the variables correctly. Furthermore, the dynamic speed of adjustment of the real exchange rate to its long-run equilibrium is determined.

Our estimation results are based on the dynamic ordinary least squares (DOLS) method. Several methods to estimate a panel cointegration model are discussed in the literature. However, Kao and Chiang (2000) show that the DOLS approach developed by Stock and Watson (1993) outperforms the panel OLS or the fully modified OLS (FMOLS) procedures in the sense that the DOLS estimator is less biased in finite samples. In addition, the choice of this method facilitates comparison with the results from similar studies, e.g., MacDonald and Ricci (2007). Our estimation equation has the following form:

$$RER_{it} = \delta_t + \alpha_i + X_{it}\beta + \sum_{j=-p}^{j=k} \Delta X_{it+j}\gamma_j + \epsilon_{it} \tag{1}$$

where RER_{it} denotes the real exchange rate at time t of country i, α_i is a country fixed effect, δ_t is a time fixed effect, X_{it} is a vector containing the explanatory variables, β is the cointegration vector, k and p are the maximum and minimum lag lengths, respectively, γ_j are the $k+p+1$ vectors containing the coefficients of the leads and lags of changes in the explanatory variables, and ϵ_{it} represents the error term.

The inclusion of the leads and lags solves the potential endogeneity problem by orthogonalizing the error term.[8] We choose the number of leads and lags to be one ($k = p = 1$). A rising number of leads and lags further constrains the number of observations. This may be a caveat particularly in subsamples with reduced numbers of years.[9]

Both time and country fixed effects are added in order to reduce omitted variable bias, and because some variables are indices, and hence, their levels are not comparable across countries. Furthermore, as described in Section 2.1, time fixed effects are necessary, because our real exchange rate is not computed towards a reference country.

We report standard errors developed by Driscoll and Kraay (1998) that are robust to very general forms of spatial and temporal dependence. For the computation, we follow Cribari-Neto (2004), who proposed an estimator (called HC4) that is reliable when the data contain influential observations.[10]

To ensure that what we find is indeed a long-run relationship between the real exchange rate and the set of explanatory variables, we test for cointegration using two methods. First, we follow MacDonald and Ricci (2007), who apply the standard unit root test of Levin et al. (2002) to the estimated residuals.[11] Second, we employ the Kao (1999) panel cointegration test. Since this test requires a balanced panel, some observations have to be dropped, and, therefore, the test is mainly applied to check the robustness of the first test results.

In order to capture the short-run dynamic adjustment of the real exchange rate to temporary disequilibria, an error correction specification is

[8]The leads and lags remove the correlation between the error term and the stationary component of the non-stationary variables.

[9]However, our main conclusions are robust to an increased number of leads and lags.

[10]As a robustness check, we employ the HC3 estimator proposed by Long and Ervin (2000). The conclusions do not change.

[11]For the theoretical foundation of this methodology, see Pedroni (2004). The conclusions do not change if the residuals are corrected by the estimated leads and lags.

applied to the data. The model can be written as follows:

$$\Delta RER_{it} = \Theta_i + \Theta_t + \eta\, gap_{it-1} + \sum_{j=0}^{j=1} \phi_j \Delta RER_{it-j} + \sum_{j=0}^{j=1} \Delta X_{it-j}\omega_j + \mu_{it} \quad (2)$$

where

$$gap_{it} = RER_{it} - \delta_t - \alpha_i - X_{it}\beta \quad (3)$$

and where η represents the adjustment speed coefficient. The gap_{it} is computed using the results of Equation (1).

4 Empirical Results

In order to explore the validity of the Balassa-Samuelson (BS) hypothesis, we estimate various DOLS model specifications. This section presents the results for the relationship between the equilibrium real exchange rate and relative productivity as well as control variables. In addition, the results of the cointegration tests described in Section 3 are reported. Finally, we provide an extensive robustness analysis of our main findings.

4.1 Comparison of the Productivity Datasets

In a first step, we examine the validity of the BS hypothesis using the newest sectoral productivity data from the PDBi. For the purpose of comparability, we restrict our sample to the same set of countries and control variables as MacDonald and Ricci (2007). Therefore, the real exchange rate (RER) is conditioned on total factor productivity or labor productivity of tradables (TFP_T, LP_T) and non-tradables (TFP_NT, LP_NT), net foreign assets in percent of GDP (NFA), and the long-term real interest rate (RI). The cross-section dimension is reduced to the countries listed in sample (i) in Appendix A.1.

The results in column (1) of Table 4 are based on the model with TFP data from the new Productivity Database (PDBi) and the sample period lasting from 1985 to 2008.[12] There is a statistically significant negative

[12] Notice that Japan is not covered by the PDBi database.

Table 4: Comparison of the Datasets (DOLS)

Dependent Variable: RER	(1)	(2)	(3)	(4)	(5)
TFP_T	-0.427***	0.119			
	(0.108)	(0.280)			
TFP_NT	0.260	-0.661***			
	(0.227)	(0.101)			
LP_T			-0.114***	-0.114	-0.196***
			(0.039)	(0.161)	(0.058)
LP_NT			0.580***	1.058***	0.445
			(0.290)	(0.170)	(0.383)
NFA	0.0004	-0.002	0.0002	-0.001	0.0001
	(0.001)	(0.002)	(0.001)	(0.004)	(0.001)
RI	0.006	0.007	0.008	0.014***	-0.007
	(0.009)	(0.007)	(0.010)	(0.004)	(0.012)
LLC Test	-6.385***	-6.632***	-6.522***	-5.388***	-5.934***
Kao Test	-3.647***	-4.043***	-3.030***	-2.588***	-3.663***
η	-0.46	-0.24	-0.18	-0.21	-0.47
Half Lifetime (years)	1.1	2.5	3.4	2.9	1.1
Sample Period	1985-2008	1970-1992	1970-2008	1970-1992	1992-2008
Obs.	112	188	289	163	113

Notes: See Table 1 for the definitions of the variables. All FE estimator regressions include country-specific and time-specific dummy variables as well as first differences of each explanatory variable (1 lead/lag). Country-sample (Appendix A.1): Sample (i). Japan is not included in (1) (Footnote 12). The productivity data stem from the PDBi (1), the ISDB (2), and the STAN database (3)-(5). Robust standard errors proposed by Driscoll and Kraay (1998) are reported in parentheses. LLC test: Cointegration test following MacDonald and Ricci (2007): t-statistic of Levin et al. (2002) (Lag length selection by SIC; Bartlett kernel, Newey-West bandwidth). Kao test: Cointegration test proposed by Kao (1999): t-statistic (Lag length selection by SIC; Bartlett kernel, Newey-West bandwidth). η is obtained from Equation (2). Half lifetime of deviations of the real exchange rate from estimated relation (years): $\ln(0.5)/(\ln(1+\eta))$. */**/*** denotes significance at 10%, 5% and 1% level, respectively.

impact of TFP_T on RER. A 10% increase in the TFP of tradables implies a 4% depreciation of the real exchange rate. The coefficient on TFP_NT is positive but not significant. Overall, this result contradicts not only the BS hypothesis, but also the usual conclusions drawn from empirical studies analyzing Balassa-Samuelson in a panel of OECD countries.

However, most of the related literature obtains sectoral productivity data from the older International Sectoral Database (ISDB). For a comparison with the existing literature, in particular MacDonald and Ricci (2007), column (2) reports the estimation results with TFP data from the ISDB for the period from 1970 to 1992. Except for NFA, the results are now qualitatively equal to the findings of MacDonald and Ricci (2007). In particular, the signs of the coefficients related to both TFP variables are consistent with the BS hypothesis. Quantitatively, the coefficients on both coefficients are smaller than in MacDonald and Ricci (2007), and the coefficient on TFP_T is statistically insignificant. These differences disappear to a great extent once we follow MacDonald and Ricci (2007) and increase the number of leads and lags to three. The results are shown in column (1) of Table 8 in Appendix B. Hence, we are able to replicate the results in favor of the BS theory with data from the ISDB.

The successful confirmation of the BS hypothesis may depend either on the productivity data source or the sample period or both. Unfortunately, the two datasets ISDB and PDBi contain only very few overlapping observations. Therefore, we are not able to distinguish the *time* from the *source* effect. In order to verify our finding, we estimate the model with labor productivity (LP) data from STAN, as this database covers both periods.[13]

Column (3) displays the results for the whole sample from 1970 to 2008. Column (4) shows the results of the subsample from 1970 to 1992 and thus covers the same period as the estimations with productivity data from the ISDB (column 2). The second subsample (column 5) ranges from 1992 to 2008.

The coefficients on LP_T are negative in all estimations and statistically significant for the whole sample and in the second subsample, confirming the results from column (1). The coefficients on LP_NT are positive in all estimations and statistically significant for the whole sample and for the first subsample, again confirming the results from column (1). Thus, labor productivity data from the STAN database lead to similar results as total factor productivity data from the PDBi, but both results contradict the BS hypothesis and differ from the findings using the ISDB. This result is further analyzed in Section 4.4. As the sign of the coefficients on LP_T and LP_NT is the same across both subsamples, the differences between columns (1) and (2) are likely to be determined by the dataset rather than by the sample period. Note that the essential difference remains when we substitute LP for TFP from the ISDB. The coefficients are displayed in column (2) of Table 8 in Appendix B.

The control variable NFA has the correct sign in columns (1), (3) and

[13] Notice that due to lack of data for some years, the coverage is not exactly the same. See Figure 1 for more details.

(5). However, the coefficients are statistically insignificant and the economic effect is considerably smaller compared to the results of Lee et al. (2008) and Lane and Milesi-Ferretti (2004). The model with labor productivity data from the STAN database estimated from 1970 to 1992 has a significant positive coefficient on RI consistent with the theory (Stein and Allen, 1997).

Our results with ISDB productivity data point to a half-life of deviations of the real exchange rate from its estimated long-run relationship of 2.5 to 2.7 years (column (2), Table 4, and column (2), Table 8, in Appendix B). This result is in line with the existing literature (see, e.g., Lee et al. (2008)), but larger than in MacDonald and Ricci (2007). In recent times, the adjustment speed has accelerated to about one year. These results are reported in columns (1) and (5) of Table 4. As our main interest concerns the cointegration relationship, we do not discuss this issue further.

4.2 Full Country-Sample Estimations

The estimations in Table 4 are based on a reduced set of countries. We now re-estimate the model with all countries available (sample (ii) and (iii), Appendix A.1). In addition, we drop the variables NFA and RI, since neither variable seems to have considerable explanatory power for the long-run real exchange rate. Instead, we use the terms of trade (TOT) as a control variable in the baseline model, as TOT turns out to be an important and robust determinant of the equilibrium real exchange rate.[14]

Table 5 summarizes the results. Compared to Table 4, the coefficients on TFP_T and LP_T are smaller, but remain negative and significant with a single insignificant exception (column 3). However, for the same first period subsample, the coefficient is also insignificant (but negative) with the reduced country-sample (column (4), Table 4). A 10% increase in the TFP of tradables from the new PDBi would imply an almost 2% depreciation of the real exchange rate. A similar result emerges with LP data from the STAN database for the sample period from 1992 to 2008 (column 4). Thus, the negative relationship between the productivity of tradables and the real exchange rate persists when all countries are included. We will further explore

[14]The inclusion of TOT raises the concern about possible endogeneity. We conduct a very simple exercise to check for reverse causation by substituting the contemporaneous value by the one-year-lagged value of TOT. The results are robust to this modification. Therefore, we conclude that this potential endogeneity problem is not of a major concern in our analysis. The results are shown in Table 9 in Appendix B.

Table 5: Full Country-Sample Estimation Results (DOLS)

Dependent Variable: RER	(1)	(2)	(3)	(4)
TFP_T	-0.176***			
	(0.061)			
TFP_NT	-0.767**			
	(0.314)			
LP_T		-0.0493**	0.143	-0.164***
		(0.024)	(0.143)	(0.062)
LP_NT		0.596***	0.407***	-0.110
		(0.185)	(0.125)	(0.151)
TOT	0.265*	0.302**	0.132	0.263**
	(0.145)	(0.126)	(0.156)	(0.127)
LLC Test	-8.160***	-8.766***	-8.468***	-9.002***
Kao Test	-2.052**	1.285*	1.352*	1.622*
Sample Period	1985-2008	1970-2008	1970-1992	1992-2008
Obs.	181	532	251	258

Notes: See Table 1 for the definitions of the variables. All FE estimator regressions include country-specific and time-specific dummy variables as well as first differences of each explanatory variable (1 lead/lag). Country-samples (Appendix A.1): Sample (ii) for (1), sample (iii) for (2) and (4), and sample (iv) for (3). The productivity data stem from the PDBi (1), and the STAN database (2)-(4). Robust standard errors proposed by Driscoll and Kraay (1998) are reported in parentheses. LLC test: Cointegration test following MacDonald and Ricci (2007): t-statistic of Levin et al. (2002) (Lag length selection by SIC; Bartlett kernel, Newey-West bandwidth). Kao test: Cointegration test proposed by Kao (1999): t-statistic (Lag length selection by SIC; Bartlett kernel, Newey-West bandwidth). */**/*** denotes significance at 10%, 5% and 1% level, respectively.

the robustness of this relationship in Section 4.4.

The effect of non-tradable productivity is less robust. Compared to Table 4, the coefficients on the productivity of non-tradables switch signs in two estimations: In column (1), the coefficient on TFP_NT drops from 0.26 to -0.77 and becomes statistically significant (compared to column (1), Table 4), in column (4) the coefficient on LP_NT turns negative, but remains insignificant (compared to column (5), Table 4). We will explore this *lack* of robustness in Section 4.4.

TOT is statistically and economically significant with the correct sign in columns (1), (2) and (4). On average, a 10% increase in the terms of trade leads to a 2% appreciation of the real exchange rate.

Table 6: Control Variables (DOLS)

Dependent Variable: *RER*				
Variables	(1)	(2)	(3)	(4)
TFP_T	-0.140	-0.083	-0.211**	-0.239***
	(0.087)	(0.090)	(0.084)	(0.060)
TFP_NT	-0.696**	-0.336	-0.701**	-0.845***
	(0.317)	(0.281)	(0.355)	(0.286)
TOT	0.231**	0.396**	0.177*	0.254
	(0.109)	(0.170)	(0.106)	(0.163)
GOV	0.001			
	(0.002)			
CA		-0.010***		
		(0.003)		
GDP			0.254***	
			(0.089)	
DPOP				-15.8
				(9.93)
LLC Test	-8.316***	-8.150***	-7.166***	-7.871***
Kao Test	-1.817**	-2.595***	-3.102***	-2.521***
Obs.	181	181	174	169

Notes: See Table 1 for the definitions of the variables. All FE estimator regressions include country-specific and time-specific dummy variables as well as first differences of each explanatory variable (1 lead/lag). Country-sample (Appendix A.1): Sample (ii). The productivity data stem from the PDBi. Robust standard errors proposed by Driscoll and Kraay (1998) are reported in parentheses. LLC test: Cointegration test following MacDonald and Ricci (2007): t-statistic of Levin et al. (2002) (Lag length selection by SIC; Bartlett kernel, Newey-West bandwidth). Kao test: Cointegration test proposed by Kao (1999): t-statistic (Lag length selection by SIC; Bartlett kernel, Newey-West bandwidth). */**/*** denotes significance at 10%, 5% and 1% level, respectively.

4.3 Control Variables

The impact of additional explanatory variables on the long-run real exchange rate is analyzed in Table 6. In line with the previous results, both coefficients on the productivity variables are negative and predominantly significant in all models. For the tradable sector productivity, this is the opposite effect of what is claimed by the BS hypothesis. Additionally, the significant positive impact of the terms of trade on the price level remains.

The selection of the explanatory variables is discussed in Section 2.3. In line with the theory, government spending (GOV) has a positive but insignificant effect on RER (column 1). Moreover, a current account surplus (CA) has a statistically significant positive effect, as predicted (column 2); however, the very small coefficient points to a limited economic significance.

Real GDP per capita (GDP) affects RER significantly positive (column 3) and confirms the hypothesis that the income level affects the consumption pattern: A 10% increase in GDP implies a 3% increase appreciation of the real exchange rate. Finally, contrary to the theory, in our sample of OECD countries, there is no significant connection between the population growth rate ($DPOP$) and RER (column 4).

4.4 Robustness Analysis

From the estimation results presented in the previous sections, we conclude that an overall stable Balassa-Samuelson effect cannot be found. The coefficients on productivity of tradables and, in particular, on the productivity of non-tradables are not robust against various sample variations, such as the source of the productivity data, the time period, or the set of control variables.

As we have seen, TFP from the PDBi in the non-tradable sector changes its coefficient from 0.26 to -0.77 and becomes highly significant once we employ the full country-sample (column (1), Table 4, and column (1), Table 5). By repeatedly re-estimating this specification and each time omitting one of the countries of sample (ii) (Appendix A.1), we are able to identify the United States as a critical outlier. As soon as the country is omitted, the coefficient changes from being significantly negative to being insignificantly positive. But if we continue the exclusion exercise without the United States, we find that omitting Italy switches the sign of the coefficient again, this time from positive to negative. Then, excluding France changes the coefficient to positive. Next, dropping Norway leads to another sign reversal. Although the coefficients are insignificant, at least for the non-tradable sector, the result crucially depends on the country-sample chosen.

The estimation of individual slope coefficients on non-tradable productivity confirms this finding. While the effect is positive for Austria, Denmark, Greece, Italy and Norway, the contrary holds for Belgium, Finland, France, Germany, Netherlands, Spain, Sweden and the United States. Again, only the coefficient on the United States is significant. Therefore, the relationship between non-tradable productivity and the real exchange rate seem to differ across the countries making a panel approach for analyzing this relationship questionable.

Furthermore, Japan seems to be an outlier that critically affects the

Table 7: The Impact of Japan on the Results (DOLS)

Dependent Variable: *RER*			
Variables	(1)	(2)	(3)
LP_T	-0.070**	-0.020	-0.189***
	(0.027)	(0.113)	(0.051)
LP_NT	0.304***	-0.334**	-0.256*
	(0.112)	(0.179)	(0.133)
TOT	0.380***	0.397***	0.165**
	(0.075)	(0.090)	(0.075)
LLC Test	-9.230***	-8.561***	-9.832***
Kao Test	0.350	2.223**	1.125
Obs.	497	230	245

Notes: See Table 1 for the definitions of the variables. All FE estimator regressions include country-specific and time-specific dummy variables as well as first differences of each explanatory variable (1 lead/lag). Country-samples (Appendix A.1): Sample (iii) without Japan for (1) and (3), sample (iv) without Japan for (2). The productivity data stem the STAN database. Robust standard errors proposed by Driscoll and Kraay (1998) are reported in parentheses. LLC test: Cointegration test following MacDonald and Ricci (2007): t-statistic of Levin et al. (2002) (Lag length selection by SIC; Bartlett kernel, Newey-West bandwidth). Kao test: Cointegration test proposed by Kao (1999): t-statistic (Lag length selection by SIC; Bartlett kernel, Newey-West bandwidth). */**/*** denotes significance at 10%, 5% and 1% level, respectively.

results of the estimations using labor productivity data from the STAN database. Re-estimating the specifications in columns (2)-(4) of Table 5 without Japan produces different results, reported in Table 7. In particular, while an increase in labor productivity in the non-tradable sector gives rise to a significant real exchange rate appreciation in the full country-sample from 1970 to 1992 (column (3), Table 5), the contrary is true if Japan is omitted (column (2), Table 7).

The replication of the exclusion exercise with country-sample (iv) in the absence of Japan shows that the coefficient on LP_NT remains negative and becomes insignificant only when Norway or the United States are excluded. Therefore, the positive relationship between Japanese non-tradable productivity and RER dominates the overall negative relationship in the other countries.

This result is also robust against moving the end date from 1985 to 1995, although the coefficient shrinks and loses its significance for the years 1993 until 1995. Furthermore, the sign remains unchanged with additional explanatory variables taken into account (Table 10 in Appendix B). Finally, the increase in the number of leads and lags of up to three (MacDonald and

Ricci, 2007) does not alter the conclusion.

There is a second, truly robust finding: The relationship between productivity in the tradable sector and the real exchange rate is negative, since the beginning of the 1990s (columns (1) and (4), Table 5). While this result contradicts the BS hypothesis, it is robust against variations in the country-sample, the sample period, the exact model specification and the set of control variables. The finding is also independent of whether we use labor or total factor productivity.

The exclusion exercise reveals that the negative sign is persistent against the omission of any country. In rare cases, the coefficient becomes statistically insignificant (if Italy, Sweden (with LP), or Norway (with LP and TFP) are excluded). Varying the start point of the sample shows that the negative coefficient is significantly negative from 1988 to 1995, independent of the productivity data source. The relationship remains significantly negative when up to three leads and lags are included into the model. Finally, the finding is robust against the inclusion of additional explanatory variables (Tables 11 and 12 in Appendix B).

Once we increase the starting point of the sample from 1985 to 1992, the variables CA and GDP are not significant anymore (Table 12 in Appendix B). Of all additional explanatory variables, it is thus only the terms of trade that is robust against a sample variation.

The relationship between productivity in the tradable sector and the real exchange rate is illustrated in Figure 2. The left panel contains the productivity of tradables in relation to the real exchange rate, both adjusted by country-specific and time-specific effects. In the right panel, the real exchange rate is additionally adjusted by the productivity of non-tradables and the terms of trade. The small differences between the left and the right panel indicate that the relationship does not depend on whether control variables are used. In line with the DOLS estimation results, all scatter plots show the significant negative relationship.

Having found a robust negative relationship between tradable productivity and the real exchange rate, we conclude that sectoral productivity panel data from the OECD offer no support for the BS hypothesis for a panel of major OECD countries.

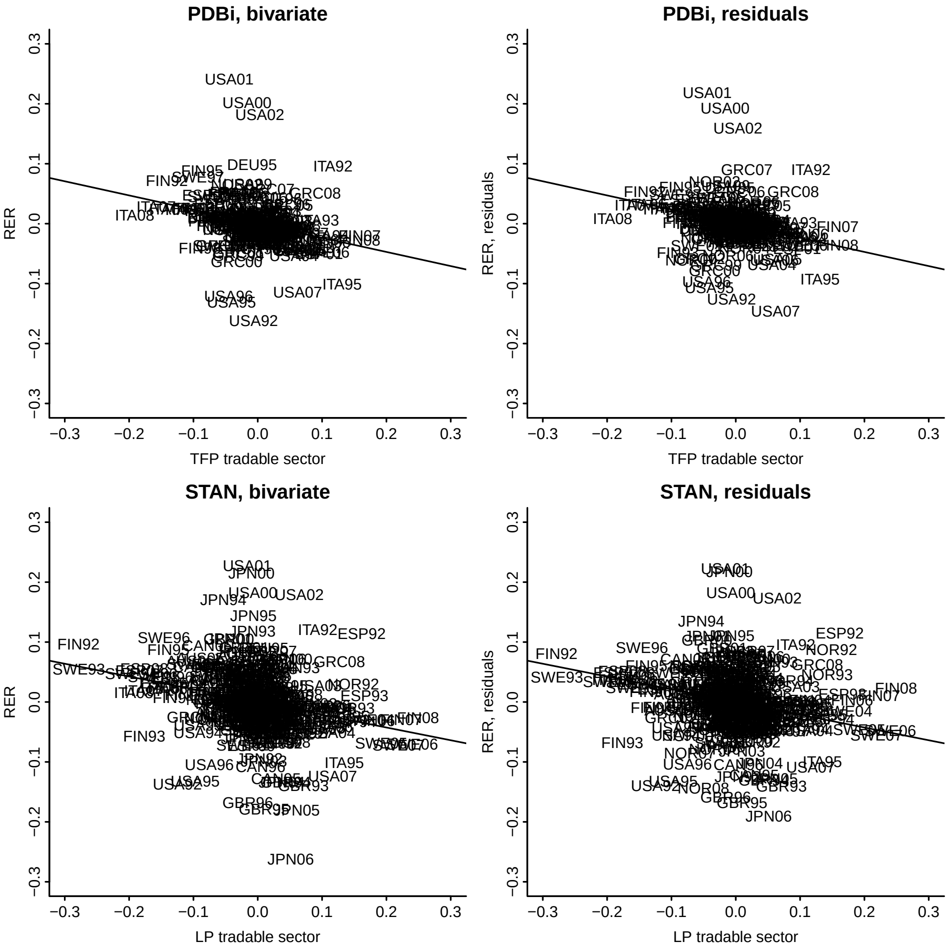

Figure 2: Tradable Productivity and the Real Exchange Rate since 1992

5 Summary and Conclusions

This paper explores the robustness of the Balassa-Samuelson (BS) hypothesis. We analyze a panel of OECD countries from 1970 to 2008 and compare three different datasets on sectoral productivity provided by the OECD, including a newly constructed database on total factor productivity (TFP).

Overall, we cannot find support for the BS hypothesis. In contrast, our DOLS estimations point to a very robust negative relationship between productivity in the tradable sector and the equilibrium real exchange rate during the last two decades. We find this negative relationship with respect to TFP from the new Productivity Database (PDBi) as well as with sectoral labor productivity (LP) from the STAN database. The finding not only contradicts the BS hypothesis, but also the results of previous empirical research that is based on the older International Sectoral Database (ISDB).

Results from estimations with LP indicate that the difference in the findings from studies using TFP data from the ISDB (in favor of BS) and the PDBi (against BS) are due to the data source and not due to a change of the relationship over time.

An extensive robustness analysis shows that the negative relationship does not depend on the choice of the country-sample, the precise start of the time period, the exact model specification, the inclusion of additional explanatory variables or the non-tradable productivity. On the other hand, the relationship between the productivity in the non-tradable sector and the long-run real exchange rate during the last two decades is strongly affected by the choice of the country-sample.

Prior to 1992, the robustness tests further reveal a strong dependency of the results on a single outlier: The coefficient on non-tradable labor productivity significantly changes the sign once Japan is included. Without Japan, we find a robust negative relationship between non-tradable productivity and the real exchange rate, in line with the BS hypothesis.

Finally, we examine the explanatory power of control variables whose importance for the real exchange rate determination has been discussed in the literature. The results indicate that, with the exception of the terms of trade, their explanatory power is weak or not robust against the chosen time period.

The fact that we find a robust negative relationship between tradable productivity and the real exchange rate is puzzling. According to the Balassa-Samuelson hypothesis, we would expect a higher productivity to be connected with higher wages and thus with a higher price level. Why is the opposite the case? In Gubler and Sax (2011), we show that skill-based technological change may lead to an increase in productivity *and* a lower demand for low-skilled labor, and thus to lower prices in the economy. Of course, other explanations are equally possible.

Our findings potentially facilitate future empirical research on the determination of the equilibrium real exchange rate in OECD countries. As the Balassa-Samuelson hypothesis does not contribute to an explanation of real exchange rate movements, sectoral productivity data do not have to serve necessarily as a control variable. This should bring a major gain in data availability. Not only more countries but more years can be included without running into potential omitted variable bias.

References

Balassa, Bela, "The Purchasing-Power Parity Doctrine: A Reappraisal," *The Journal of Political Economy*, December 1964, *72* (6), 584–596.

Calderón, C.A., "Real Exchange Rates in the Long and Short Run: A Panel Cointegration Approach," *Revista de Análisis Económico*, 2004, *19*, 41–83.

Canzoneri, M.B., R.E. Cumby, and B. Diba, "Relative labor productivity and the real exchange rate in the long run: evidence for a panel of OECD countries," *Journal of International Economics*, 1999, *47* (2), 245–266.

Chinn, Menzie D. and Louis Johnston, "Real Exchange Rate Levels, Productivity and Demand Shocks: Evidence from a Panel of 14 Countries," NBER Working Paper 5709, National Bureau of Economic Research, Cambridge, MA, USA August 1996.

Choudhri, E.U. and M.S. Khan, "Real Exchange Rates in Developing Countries: Are Balassa-Samuelson Effects Present?," *IMF Staff Papers*, 2005, *52* (3), 387–409.

Cribari-Neto, F., "Asymptotic inference under heteroskedasticity of unknown form," *Computational Statistics & Data Analysis*, 2004, *45* (2), 215–233.

De Gregorio, José and Holger C. Wolf, "Terms of Trade, Productivity, and the Real Exchange Rate," NBER Working Paper 4807, National Bureau of Economic Research, Cambridge, MA July 1994.

Dornbusch, Rudiger, "Purchasing Power Parity," NBER Working Paper 1591, National Bureau of Economic Research, Cambridge, MA March 1985.

Driscoll, J.C. and A.C. Kraay, "Consistent covariance matrix estimation with spatially dependent panel data," *Review of Economics and Statistics*, 1998, *80* (4), 549–560.

Engel, C., "Accounting for US Real Exchange Rate Changes," *Journal of Political Economy*, 1999, *107* (3).

Froot, Kenneth A. and Kenneth Rogoff, "Perspectives on PPP and Long-Run Real Exchange Rates," NBER Working Paper 4952, National Bureau of Economic Research, Cambridge, MA April 1996.

Galí, Jordi, "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?," *American Economic Review*, 1999, *89* (1), 249–271.

Gubler, Matthias and Christoph Sax, "Skill-Biased Technological Change and the Real Exchange Rate," Unpublished Manuscript, University of Basel 2011.

Harrod, Roy F., *International Economics*, London: Nisbet & Cambridge University Press, 1933.

Im, K.S., M.H. Pesaran, and Y. Shin, "Testing for unit roots in heterogeneous panels," *Journal of Econometrics*, 2003, *115* (1), 53–74.

Imbs, J., H. Mumtaz, M.O. Ravn, and H. Rey, "PPP Strikes Back: Aggregation and the Real Exchange Rate," *Quarterly Journal of Economics*, 2005, *120* (1), 1–43.

Kao, C., "Spurious regression and residual-based tests for cointegration in panel data," *Journal of Econometrics*, 1999, *90* (1), 1–44.

_ **and M. H. Chiang**, "Nonstationary panels, cointegration in panels and dynamic panels: A survey," *Nonstationary Panels, Panel Cointegration, and Dynamic Panels*, 2000, *15*, 179–222.

King, R.G., C.I. Plosser, and J.H. Stock, "Stochastic Trends and Economic Fluctuations," *The American Economic Review*, 1991, *81* (4), 819–840.

Krugman, Paul, "Equilibrium Exchange Rates," in William H. Branson, Jacob A. Frenkel, and Morris Goldstein, eds., *International Policy Coordination and Exchange Rate Fluctuations*, The University of Chicago Press, 1990.

Lane, P.R. and G.M. Milesi-Ferretti, "The transfer problem revisited: Net foreign assets and real exchange rates," *Review of Economics and Statistics*, 2004, *86* (4), 841–857.

Lee, J., G.M. Milesi-Ferretti, and L.A. Ricci, "Real Exchange Rates and Fundamentals: A Cross-Country Perspective," *IMF Working Papers*, 2008.

Levin, A., C.F. Lin, and C.S. James Chu, "Unit root tests in panel data: asymptotic and finite-sample properties," *Journal of Econometrics*, 2002, *108* (1), 1–24.

Lindé, Jesper, "The effects of permanent technology shocks on hours: Can the RBC-model fit the VAR evidence?," *Journal of Economic Dynamics and Control*, March 2009, *33* (3), 597–613.

Long, J.S. and L.H. Ervin, "Using heteroscedasticity consistent standard errors in the linear regression model," *American Statistician*, 2000, pp. 217–224.

MacDonald, R. and L.A. Ricci, "The Real Exchange Rate and the Balassa–Samuelson Effect: The Role of the Distribution Sector," *Pacific Economic Review*, 2005, *10* (1), 29–48.

_ **and** _ , "Real exchange rates, imperfect substitutability, and imperfect competition," *Journal of Macroeconomics*, 2007, *29* (4), 639–664.

Maddala, G.S. and S. Wu, "A comparative study of unit root tests with panel data and a new simple test," *Oxford Bulletin of Economics and Statistics*, 1999, *61* (S1), 631–652.

Ng, Serena and Pierre Perron, "Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power," *Econometrica*, 2001, *69* (6), 1519–1554.

Pedroni, P., "Panel cointegration: asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis," *Econometric Theory*, 2004, *20* (03), 597–625.

Rogoff, Kenneth, "The Purchasing Power Parity Puzzle," *Journal of Economic Literature*, 1996, *19*, 647–668.

Rose, A.K., S. Supaat, and J. Braude, "Fertility and the real exchange rate," *Canadian Journal of Economics/Revue canadienne d'économique*, 2009, *42* (2), 496–518.

Samuelson, Paul, "Theoretical Notes on Trade Problems," *Review of Economics and Statistics*, May 1964, *46* (2), 145–154.

Sax, C. and R. Weder, "How to Explain the High Prices in Switzerland?," *Swiss Journal of Economics and Statistics*, 2009, *145* (IV), 463–483.

Stein, J.L. and P.R. Allen, *Fundamental determinants of exchange rates*, Oxford University Press, USA, 1997.

Stock, J.H. and M.W. Watson, "A simple estimator of cointegrating vectors in higher order integrated systems," *Econometrica*, 1993, *61* (4), 783–820.

Taylor, M.P., "Purchasing power parity," *Review of International Economics*, 2003, *11* (3), 436–452.

Tica, J. and I. Družić, "The Harrod-Balassa-Samuelson Effect: A Survey of Empirical Evidence," *University of Zagreb, Faculty of Economics and Business, FEB–Working Paper Series (No. 06-7/686)*, 2006.

A Data Appendix

A.1 Country-samples

This section contains all country-samples used in the estimation models:

i Belgium (BEL), Denmark (DNK), Finland (FIN), France (FRA), Germany (DEU), Italy (ITA), Japan (JPN), Norway (NOR) and Sweden (SWE)

ii Austria (AUT), Belgium (BEL), Denmark (DNK), Finland (FIN), France (FRA), Germany (DEU), Greece (GRC), Italy (ITA), Netherlands (NLD), Norway (NOR), Spain (ESP), Sweden (SWE) and the United States (USA)

iii Australia (AUS), Austria (AUT), Belgium (BEL), Canada (CAN), Denmark (DNK), Finland (FIN), France (FRA), Germany (DEU), Great Britain (GBR), Greece (GRC), Italy (ITA), Japan (JPN), Netherlands (NLD), Norway (NOR), Portugal (PRT), Spain (ESP), Sweden (SWE) and the United States (USA)

iv Austria (AUT), Belgium (BEL), Canada (CAN), Denmark (DNK), Finland (FIN), France (FRA), Great Britain (GBR), Greece (GRC), Italy (ITA), Japan (JPN), Netherlands (NLD), Norway (NOR), Portugal (PRT), Spain (ESP) and the United States (USA)

A.2 Data Sources

i IMF, International Financial Statistics

We gained the following IFS variables via Datastream:

- BOND YIELD (AUY61... etc.)
- CPI (AUY64...F etc.)
- EXCHANGE RATE, US$ PER LC (AUOCFEXR etc.)

ii OECD, Economic Outlook

The data are from Economic Outlook No 88., available on `http://www.oecd-ilibrary.org/`. These variables were used:

- Imports of goods and services, deflator, national accounts basis (PMGSD)
- Exports of goods and services, deflator, national accounts basis (PXGSD)
- Current account balance, as a percentage of GDP (CBGDPR)
- Total disbursements, general government, as a percentage of GDP

iii OECD, STAN Database for Structural Analysis

The data are from `oecd.stat` and have been downloaded as a single ASCII file. The series for Germany have been retropolated with the former West-Germany series.

iv OECD, PDBi, Sectoral Productivity Database

A new dataset provided by the OECD (we used a pre-released version of the dataset).

Both for the STAN database and the PDBi, tradable and non-tradable productivity is calculated the following way:

$$P_{NT} = \frac{S_{7599} \cdot P_{7599} + S_{4041} \cdot P_{4041} + S_{4500} \cdot P_{4500}}{S_{7599} + S_{4041} + S_{4500}},$$
$$P_{T} = \frac{S_{0105} \cdot P_{0105} + S_{1537} \cdot P_{1537} + S_{6064} \cdot P_{6064}}{S_{0105} + S_{1537} + S_{6064}},$$

where P denotes labor productivity in the STAN case and total factor productivity in the PDBi case. S is the share of the subsector.

v OECD, ISDB, Sectoral Productivity Database

A vintage dataset provided by the OECD.

Tradable and non-tradable total factor productivity is calculated the following way (again, P denotes labor or total factor productivity and S the share of the subsector):

$$P_{NT} = \frac{S_{SOC} \cdot P_{SOC} + S_{EGW} \cdot P_{EGW} + P_{CST} \cdot S_{CST}}{S_{SOC} + S_{EGW} + S_{CST}},$$
$$P_{T} = \frac{S_{AGR} \cdot P_{AGR} + S_{MAN} \cdot P_{MAN} + S_{TRS} \cdot P_{TRS}}{S_{AGR} + S_{MAN} + S_{TRS}}.$$

vi Penn World Tables

The data are from `http://pwt.econ.upenn.edu/php_site/pwt_index.php`. These variables were used:

- Real GDP-per-capita (USD of 2005) (RGDPL)
- Population (in 1000) (POP)

The population growth rate is calculated as the first difference of the logarithm of POP.

vii World Bank, World Development Indicators

The following variables are extracted from the WDI CD-ROM:

- Net foreign assets

The share of net foreign assets (NFA in the text) is calculated in the following way:

$$NFA = \frac{NFA_{Level}}{GDP \cdot 1000000}$$

where NFA_{Level} are the net foreign assets as taken from WDI and GDP denotes nominal GDP taken from the OECD Economic Outlook. The missing value of NFA_{Level} for Belgium and France for the year 1998 is replaced by a linearly interpolated value. The results do not change.

B Additional Estimation Results

Table 8: Additional Results with ISDB Productivity Data (DOLS)

Dependent Variable: *RER*	(1)	(2)
TFP_T	0.958***	
	(0.328)	
TFP_NT	-1.070***	
	(0.122)	
LP_T		0.430**
		(0.182)
LP_NT		-0.166**
		(0.074)
NFA	0.0004	0.004*
	(0.003)	(0.002)
RI	-0.010	0.009
	(0.010)	(0.008)
LLC Test	-6.146***	-6.632***
Kao Test	-4.043***	-1.285*
η	-0.23	-0.23
Half Lifetime (years)	2.7	2.7
Sample Period	1970-1992	1970-1992
Obs.	152	188

Notes: See Table 1 for the definitions of the variables. All FE estimator regressions include country-specific and time-specific dummy variables as well as first differences of each explanatory variable. (1) includes 3 leads/lags; (2) includes 1 lead/lag. Country-sample (Appendix A.1): Sample (i). The productivity data stem from the ISDB. Robust standard errors proposed by Driscoll and Kraay (1998) are reported in parentheses. LLC test: Cointegration test following MacDonald and Ricci (2007): t-statistic of Levin et al. (2002) (Lag length selection by SIC; Bartlett kernel, Newey-West bandwidth). Kao test: Cointegration test proposed by Kao (1999): t-statistic (Lag length selection by SIC; Bartlett kernel, Newey-West bandwidth). η is obtained from Equation (2). Half lifetime of deviations of the real exchange rate from estimated relation (years): $\ln(0.5)/(\ln(1+\eta))$. */**/*** denotes significance at 10%, 5% and 1% level, respectively.

Table 9: Full Country-Sample Estimation Results with $TOT(-1)$ (DOLS)

Dependent Variable: RER	(1)	(2)	(3)	(4)
$TFP\_T)$	-0.147**			
	(0.070)			
$TFP\_NT)$	-0.893***			
	(0.296)			
$LP\_T)$		-0.046*	0.140	-0.219***
		(0.024)	(0.142)	(0.049)
$LP\_NT)$		0.527***	0.374**	-0.282*
		(0.194)	(0.180)	(0.146)
$TOT(-1)$	0.285	0.377***	0.210	0.271**
	(0.175)	(0.121)	(0.150)	(0.132)
Sample Period	1985-2008	1970-2008	1970-1992	1992-2008
Obs.	168	514	237	240

Notes: See Table 1 for the definitions of the variables. All FE estimator regressions include country-specific and time-specific dummy variables as well as first differences of each explanatory variable (1 lead/lag). Country-samples (Appendix A.1): Sample (ii) for (1) and sample (iii) for (2)-(4); Australia, Germany and Sweden are not included in (3) due to missing data. The productivity data stem from the PDBi (1), and the STAN database (2)-(4). Robust standard errors proposed by Driscoll and Kraay (1998) are reported in parentheses. */**/*** denotes significance at 10%, 5% and 1% level, respectively.

Table 10: Control Variables: Japan Omitted (DOLS)

Dependent Variable: *RER*				
Variables	(1)	(2)	(3)	(4)
LP_T	-0.053	-0.179	-0.089	-0.046
	(0.107)	(0.141)	(0.127)	(0.153)
LP_NT	-0.177	-0.107	-0.257**	-0.329**
	(0.125)	(0.188)	(0.118)	(0.138)
TOT	0.412***	0.480***	0.390***	0.427***
	(0.122)	(0.044)	(0.086)	(0.085)
GOV	0.001			
	(0.002)			
CA		-0.0138***		
		(0.004)		
GDP			0.726***	
			(0.136)	
DPOP				-2.71
				(2.31)
Obs.	229	195	230	224

Notes: See Table 1 for the definitions of the variables. All FE estimator regressions include country-specific and time-specific dummy variables as well as first differences of each explanatory variable (1 lead/lag). Country-sample (Appendix A.1): Sample (iv) without Japan. The productivity data stem from the STAN database. Sample period: 1970-1992. Robust standard errors proposed by Driscoll and Kraay (1998) are reported in parentheses. */**/*** denotes significance at 10%, 5% and 1% level, respectively.

Table 11: Control Variables: Estimations with LP (DOLS)

Dependent Variable: *RER*				
Variables	(1)	(2)	(3)	(4)
LP_T	-0.142**	-0.083	-0.184***	-0.209***
	(0.064)	(0.079)	(0.071)	(0.045)
LP_NT	-0.083	-0.071	-0.050	-0.206
	(0.149)	(0.150)	(0.156)	(0.162)
TOT	0.249**	0.321***	0.258***	0.319***
	(0.113)	(0.118)	(0.098)	(0.089)
GOV	0.002			
	(0.002)			
CA		-0.006***		
		(0.002)		
GDP			0.296**	
			(0.120)	
DPOP				-7.78
				(8.8)
Obs.	258	258	247	231

Notes: See Table 1 for the definitions of the variables. All FE estimator regressions include country-specific and time-specific dummy variables as well as first differences of each explanatory variable (1 lead/lag). Country-sample (Appendix A.1): Sample (iii). The productivity data stem from the STAN database. Sample period: 1992-2008. Robust standard errors proposed by Driscoll and Kraay (1998) are reported in parentheses. */**/*** denotes significance at 10%, 5% and 1% level, respectively.

Table 12: Control Variables: Estimations with TFP (DOLS)

Dependent Variable: *RER*				
Variables	(1)	(2)	(3)	(4)
TFP_T	-0.222***	-0.221**	-0.377***	-0.284***
	(0.083)	(0.089)	(0.067)	(0.070)
TFP_NT	-0.912**	-0.886**	-0.954*	-0.655
	(0.433)	(0.408)	(0.530)	(0.405)
TOT	0.168***	0.195**	0.131*	0.036
	(0.060)	(0.089)	(0.068)	(0.059)
GOV	0.001			
	(0.003)			
CA		-0.002		
		(0.002)		
GDP			0.256	
			(0.167)	
DPOP				-22.6*
				(13.7)
Obs.	146	146	139	134

Notes: See Table 1 for the definitions of the variables. All FE estimator regressions include country-specific and time-specific dummy variables as well as first differences of each explanatory variable (1 lead/lag). Country-sample (Appendix A.1): Sample (ii). The productivity data stem from the PDBi. Sample period: 1992-2008. Robust standard errors proposed by Driscoll and Kraay (1998) are reported in parentheses. */**/*** denotes significance at 10%, 5% and 1% level, respectively.