

# Through the Looking Glass: A WARPed View of Real Exchange Rate History\*

Douglas L. Campbell<sup>†</sup>  
New Economic School (NES)

November, 2014

## Abstract

Commonly used trade-weighted real exchange rate indices are computed as indices-of-indices, and thus do not adequately account for growth in trade with developing countries. Weighted Average Relative Price (WARP) indices solve this problem but do not control for productivity differences, as developing countries are observed to have lower price levels via the Balassa-Samuelson effect. I remedy these problems in two ways. First I propose a Balassa-Samuelson productivity adjustment to Weighted Average Relative Price indices (BS-WARP). Secondly, I introduce a Weighted Average Relative Unit Labor Cost index (WARULC) for manufacturing and show that this measure does a much better job predicting trade imbalances and declines in manufacturing employment than the IMF's Relative ULC measure created as an index-of-indices. The new series reveal that for many countries currently mired in liquidity traps, relative prices reached historic highs heading into the financial crisis of 2008. I document that in 2002 – during the surprisingly sudden collapse in US manufacturing – US relative prices had not been that high relative to trading partners since the worst year of the Great Depression.

***JEL Classification:*** F10, F31, N70, C43

***Keywords:*** Real Exchange Rate Indices, Relative Unit Labor Cost Indices, Weighted Average Relative Prices, Balassa-Samuelson, Trading Partner Substitution Bias

---

\*I am indebted to comments received from seminar participants at UC Davis, Colby College, the New Economic School, Santa Clara, and at the All-UC Economic History conference at Berkeley. I would also like to thank Ju Hyun Pyun, Paul Bergin, Robert Feenstra, Chris Meissner, Kim Ruhl, and John Devereux for their suggestions. Special thanks to Barry Eichengreen and the Berkeley Economic History Lab for providing access to data resources. I would also like to thank the hardworking public servants at the BLS, the BEA, and the OECD for responding to data inquiries.

<sup>†</sup>Visiting scholar, Berkeley Economic History Lab. New Economic School, Moscow Russia Tel.: 7-925-629-6600, e-mail: [dolcampb@gmail.com](mailto:dolcampb@gmail.com), Homepage: [dougcampbell.weebly.com](http://dougcampbell.weebly.com).

# 1 Introduction

One of the most important prices in any open economy is the real rate of exchange. Trade-weighted real exchange rate indices thus provide a useful guide to both policy-makers and academic economists as rough measures of the competitiveness of a currency in international trade.<sup>1</sup> In this paper I examine the methodology used to create these indices with the central goal of producing theoretically-appropriate measures of real exchange rate indices that an applied researcher could use to gauge the impact of RER movements on trade. I argue that real exchange rate history needs to be viewed through the appropriate looking glass. And what one finds there in this distorted world is that many key events in economic history—the Asian Financial Crisis, the swift decline of American manufacturing, the Great Depression, and the “Lesser Depression”, as well as the ongoing structural US trade deficit—are cast in new light.

The most commonly used real exchange rate indices are constructed by the Federal Reserve, the IMF, and the OECD as indices-of-indices. The levels of these series thus are not internationally comparable and they suffer from what I call a “trading partner substitution bias” problem, as they do not adequately account for growth in trade with developing countries. This problem is analogous to the “outlet substitution bias” problem in the CPI (Reinsdorf 1998) and to what is called either the “offshoring bias” (Houseman *et al.* 2011) or “input-sourcing bias” (Inklaar 2013 and Reinsdorf and Yuskavage 2014).<sup>2</sup> India and China are assigned the same base value in these price indices as are Switzerland and Germany, even though the latter have much higher prices for all years, which becomes problematic when trade increases with India and China relative to countries with higher price levels. In a seminal contribution, Fahle, Marquez, and Thomas (2008) rewrote the prior 20 years of US real exchange rate history by showing that a simple Weighted Average Relative Price (WARP) index implies that the dollar

---

<sup>1</sup>I began this project while doing research on the impact of exchange rate movements on various economic variables. I soon discovered, as Fahle *et al.* (2008) did, that the real exchange rate indices created by the Fed, the IMF, and the OECD, which have appeared widely in academic research, are not suitable for many tasks for which they are often employed. In addition, there are no appropriate indices which are publicly available for easy downloading, even for the modern era, much less historically. Any economist or policymaker who wants to consult a real exchange rate index must choose between plotting a series likely to mislead (often unwittingly), or else engage in the time-consuming task of creating a series from scratch. Thus most central bank presidents and heads of state, even in severely depressed economies such as Ireland, have never seen a real exchange rate index for their own country that accounts for compositional changes in trade for the simple reason that none exist. Thus, part of the value-added of this paper is that I provide these indices for many countries on my website for free, easy downloading.

<sup>2</sup>Diewert *et al.* (2014) provides a nice overview of the general issue, which they call “sourcing substitution bias” for the context of changing intermediate input sources.

appreciated substantially more from 1990 to 2006 compared to “divisia” based indices-of-indices produced by the Federal Reserve Board and the IMF. Fahle *et al.* (2008) also find that a geometric WARP index does a much better job of explaining trade balances from the period 1970-2006 than divisia-based alternatives.

First, I extend WARP to the period 1950-2011 using version 8.0 of the Penn World Tables, which includes changes in terms of trade, and show that compared to WARP constructed using version 7.1 of the PWT, WARP v8.0 implies that US prices appreciated nearly 16% more over the period 1990-2002 relative to trading partners. This feature can help explain the rise in the structural current account deficit and sudden collapse in tradables sector employment over that period. By 2011, according to the new version of WARP, the price level in the US was 10% higher than the price level of US trading partners. In addition, compared to v7.1, v8.0 of the PWT implies lower US prices relative to trading partners in all periods but is much more pronounced before the late 1990s. It shows less of a dollar appreciation in the 1980s, and for the Bretton Woods era, WARP lines up more closely with the Federal Reserve Board’s Broad Trade Weighted Real Exchange Rate Index, which I also extend back to 1950 using the Fed’s methodology.

One problem with using WARP as a measure of competitiveness is that poor countries should theoretically have lower price levels according to the Balassa-Samuelson effect. Having a price level twice that of Japan in 1946 has very different implications for bilateral competitiveness than having a price level twice that of Japan in 1986. A straightforward resolution to this problem is to make a Balassa-Samuelson adjustment to WARP (BS-WARP). Increased trade with less-developed countries will only result in a stronger dollar index if these countries are undervalued relative to their level of development. The index is conceptually similar to the Balassa-Samuelson residuals used by Rodrik (2008) and many others in the literature on real exchange rates and growth, except that the index proposed in this paper is a trade-weighted average of the difference between the US residual and the residuals of US trading partners.

The level of the BS-WARP index indicates a substantially more competitive dollar relative to WARP for all years from 1950 to 2011, with the dollar actually 3% *undervalued* by 2011. This finding was not anticipated and is counterintuitive given the large structural trade deficit. However, after the dollar’s dramatic rise in the 1980s, it also took several years after the dollar depreciated before trade was balanced, giving rise to an academic literature on hysteresis. The US BS-WARP index had fallen below unity because the US Balassa-Samuelson residual had fallen close to zero by 2011 and US trade is biased toward countries which also have richly-valued currencies such as Canada, Japan, and the Euro Area. That the US Balassa-Samuelson residual itself indicates that the

US price level is not overvalued (given US productivity) may in part be a function of relatively low US value-added taxes and tariffs, is distinct from the relative unit labor cost data, and could be revised in the next round of revisions of the Penn World Tables.<sup>3</sup>

In the US case, directional changes in BS-WARP are broadly similar to the directional changes in WARP (the differences are far more pronounced for countries growing or contracting quickly, such as Ireland, Korea, and Poland). The similarity between WARP and BS-WARP for the US after 2002 was not easily anticipated – the Balassa-Samuelsan adjustment lowers the RER for countries growing quickly, such as China, so it could have been expected that after 2002, the BS-WARP index would show a more moderate depreciation as trade with fast-growing China increased. Using PWT v7.1, the BS-WARP index does show a more moderate depreciation after 2002, and was still 20% overvalued as of 2010. However, PWT version 8.0 marked up the growth in Chinese prices after 2005 and thus marked down the growth rate of Chinese GDP per capita by 21% over this period, partly moderating the impact.

Of course, it has long been recognized that real exchange rate indices need to be adjusted for productivity. This is why economists have generally preferred to use real exchange rate indices computed using unit labor costs in manufacturing rather than those based on other measures, such as consumer prices, as the key measures of competitiveness in international trade. Commonly used real exchange rate indices computed by the IMF and the OECD using relative unit labor costs are also computed as indices-of-indices and thus suffer from trading partner substitution bias. In addition, they use fixed trade weights and do not include China. I propose a simple geometric Weighted Average Relative Unit Labor Cost index (WARULC), computed as total labor income in manufacturing converted to the local currency at exchange rates and total manufacturing output converted to the local currency at manufacturing PPP. I compute manufacturing PPP using PWT v8.0 methodology described in Feenstra *et al.* (2013), applying the Geary-Khamis indexing method to the manufacturing basic headings of all six publicly available International Comparison Program (ICP) benchmark years, and interpolating using manufacturing value-added growth rates reported by country specific sources for the years in between. The index I create shows a much greater dollar appreciation over time than the IMF or OECD indices, and by 2001 stood 32% higher than the IMF's index relative to 1975. This index appears to do a superior job predicting trade imbalances and periods when relatively more import-competing manufacturing sectors experience relative declines in employment compared with existing series. Reassuringly, this series

---

<sup>3</sup>The next version will include the 2011 ICP, and will be available in the fall of 2013. Subsequent drafts of this paper after that time will update to the most recent version of the PWT.

yields broadly similar conclusions about US competitiveness over time as WARP and BS-WARP, and thus solves the puzzle discussed by Chinn (2006) about the different implications of the IMF RULC index vs. other measures of the RER.

As of 2009, while China employed about 9 times as many man-hours in manufacturing than did the US to produce slightly more, Chinese hourly wages in dollars were just \$1.74 compared to \$35.18 for the US.<sup>4</sup> Thus I calculate that Chinese unit labor costs were about 37% of US unit labor costs in 2009. Although full Chinese data on employment and hours worked was unavailable through 2011, Chinese hourly wages went from being 5% of US wages to 7% of US wages in those two years alone, while production rose 24% in China versus just 10% for the US. Thus, while the gap appears to be closing, the picture that emerges of competitiveness from relative unit labor costs in the US vs. China is different from what emerges with the Balassa-Samuelson adjustment.

To the extent possible, I extend all indices over both space and time. For the US, I extend both “divisia” and WARP indices for the US historically for the period 1820-2010. The Thomas *et al.* WARP series spans 1970-2006, while the Fed’s broad trade-weighted real exchange rate index starts in 1973. The Fed’s series commences at an inopportune time as it misses the large depreciation at the end of the Bretton Woods period. I extend both series back to 1950 using the same sample of countries, trade-weighting scheme and indexing methodology as the Federal Reserve. I also extend these series back to 1922 on a consistent sample of 30 countries, and back to 1820 for a sample of five countries. Compared to divisia, WARP implies a lower US price level in the period before WWII relative to the Bretton Woods period and exhibits a slightly sharper dollar appreciation during the Great Depression, with a difference from 1928-1932 of 3%. Additionally, I propose an improvement to the Federal Reserve Board’s trade weights, but find that this leaves all indices little-changed, resulting in an increased dollar appreciation from 1992-2002 of an additional 1%.

Internationally, I produce WARP, BS-WARP, and WARULC indices for major European nations. I find that for Italy, Greece, the UK, and the Russian Federation, the WARP and BS-WARP indices reveal a much greater real appreciation since 1990 than do the IMF’s divisia-based series. For example, in 2010 the BS-WARP series for Italy stood more than 20% higher than the IMF’s series relative to 1990, and in 2007, the

---

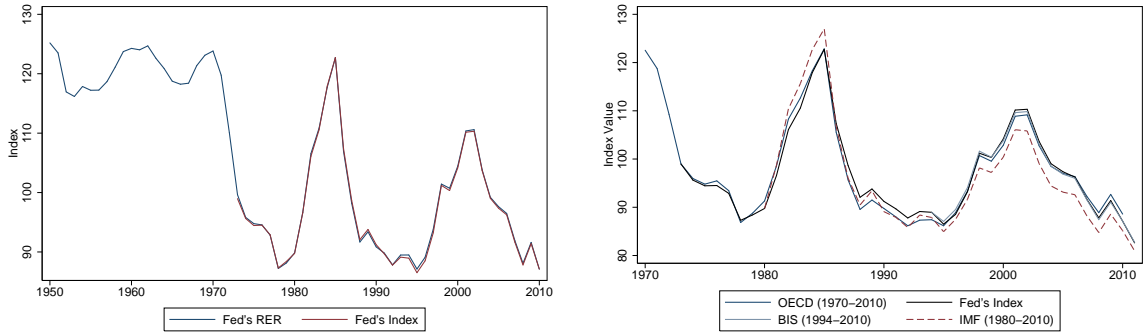
<sup>4</sup>These estimates use OECD data on US manufacturing employment and hours, which are based on household survey data for the US which are used for international comparability, and government data for Chinese employment. The better-known manufacturing employment numbers in the US come from the establishment survey, which shows 2 million workers in manufacturing. Chinese manufacturing output from the World Bank was converted into dollars at manufacturing PPP estimates, but would not be substantially different in 2009 converting at exchange rates. The hourly wage data comes from the BLS.

UK's WARP index also stood 20% higher relative to 1990 than the IMF's RER index. Conversely, Germany's BS-WARP index is similar to its IMF CPI-based real effective exchange rate. This revision of relative prices is not merely an academic curiosity given the economic problems now facing Europe. It accentuates the difficulties the European Central Bank faces in devising one monetary policy for countries with very disparate trends in relative prices measured relative to trading partners.

The Balassa-Samuelson adjusted index also reveals a substantial appreciation for the relative price level of Greece, demonstrating that currency appreciation since 1990 cannot merely be explained by Greece's convergence in GDP as is often thought. In addition, I show that WARP and BS-WARP indices for Iceland appreciated much more rapidly than did the IMF's measures leading up to the financial crises in 2008, and that these measures have also depreciated more markedly since. By contrast, I find that WARP and BS-WARP imply a more gradual appreciation for relative prices in the Russian Federation than the IMF's REER index in the 2000s.

I provide WARP and BS-WARP series for China, Korea, and Japan. Once again, these indices are substantially different from commonly used divisia-based indices. I find that China's real exchange rate was undervalued by 45% in 2005, but by 2011 it was undervalued by just 21% on a Balassa-Samuelson-adjusted basis. In 2005 China's price level was nearly 60% lower than its trading partners, but by 2011 this difference had fallen to just 35%. For Korea, the WARP index appreciated by roughly 14% more than the OECD's divisia-based real exchange rate series from 1990 to 1996, the period leading up to the Asian Financial Crisis. Since then, as Korea's trade with China continued to grow, the WARP index continued to appreciate relative to "divisia" based series, and in 2010 stood 73% higher than the OECD's index relative to 1990. As Korea has been growing fast during this period, the BS-WARP index for Korea shows a more muted difference, as it was just 49% higher than the divisia series relative to 1990. Japan's two decades spent mired in a liquidity trap have been accompanied by a domestic price level on average 95% higher than that of its trading partners, with an increase about 17% larger from 1990-2000 than the IMF's divisia-based counterpart.

The rest of the paper proceeds as follows: First I extend the Divisia and WARP indices to 1820 for the US, and then I introduce a Balassa-Samuelson adjustment to WARP and a Weighted Average Relative Unit Labor Cost (WARULC) measure. Then I test WARULC using trade and manufacturing employment data for the US. Lastly, I present indices for many other countries internationally.



(a) Fed Benchmark vs. Actual Fed RER Index

(b) Comparison of CPI-Based RER Indices

Figure 1: Benchmarking

## 2 Benchmarking the Fed, with Historical Extensions

### 2.1 Post-War Benchmark

In Figure 1(a), I benchmark the Fed’s Broad Trade-Weighted Real Exchange Rate Index using data collected from various sources detailed in Appendix Table 1. Using the same methodology and largely the same sample as the Fed, I extend the Federal Reserve Board’s series back to 1950, capturing the large depreciation after the end of Bretton Woods. While the reconstruction does not mimic the Fed’s index exactly, it never deviates by more than seven-tenths of one percent. The difference could be due to differing data sources, as I used national sources when WDI or IMF data were missing, such as for Taiwan. There also may be slight differences due to the Fed’s annual series being an average of underlying monthly data, or from special issues arising from the creation of the Euro or end of the Soviet Union which are not discussed in detail in the Fed’s literature on the creation of its index.

For the period 1973-2010, I use the Federal Reserve’s trade weights, which are based on trade net of oil. For the period prior, I calculate the trade-weights using the Federal Reserve’s trade-weighting methodology, trade data from the IMF DOTS, and oil trade data from the UN’s Comtrade database. The bilateral oil trade data is not available before 1962, so I assumed that the share of each country’s trade represented by oil for the period 1950-1961 is the same as the average for the period from 1962-1966. Even ostensibly major revisions to the weights during this period seem to have minor impacts on the overall index—if I use total trade with no adjustment for oil before 1962, the series is little changed.

In Figure 1(b), I show that other CPI-based effective exchange rate indices, created

by the IMF, the OECD, and the Bank of International Settlements, all use very similar methodology as the Fed and yield very similar results. Of these, the IMF is the outlier, perhaps because it uses constant trade-weights for the entire period (see Bayoumi *et al* 2005).

## 2.2 Historical Series

Figure 2 below extends the Fed's series using historical data for 30 US trading partners, and compares it to the Fed's post-war broad trade-weighted real exchange rate index (indexed to be equal in 2010). Six partners have data from 1820, eight from 1830, 12 by 1861, and 19 by the 1880s; the sample reaches its max of 30 in 1922 (the historical sample of countries with the starting date for each country is in the appendix where it is compared to the Fed's sample). This series builds on the scholarship of Taylor (2002) and Fratzscher, Mehl, and Vansteenkiste (2011), who provide real exchange rate indices for the US from 1880, and Catao and Solomou (2005), who provide trade-weighted RER indices for a variety of countries from 1870-1914. Taylor's series uses a simple average as historical trade data was not then available, while Fratzscher *et al.*'s series includes six trading partners.<sup>1</sup>

The historical trade weights are computed using trade data from the IMF DOTS, Comtrade, Jacks, Meissner, and Novy (2011), Barbieri (2002), and the Historical Statistics of the United States. Barbieri (2002) and the Jacks, Meissner, and Novy (2011) data begin in 1870 for a broad sample of countries, while the HSUS data begin in 1790 for a smaller sample. The prewar trade weights are simply the share of imports and exports in total US trade, while post-WWII I also include third country weights using the methodology which I discuss in Section 3.2 (I plot the weights over time for major trading partners when I introduce historical WARP in section 3.4). To extend the coverage of countries in the trade-weighted sample, for three countries I imputed trade shares for several decades based on the earliest recorded trade and changes in each country's share of world GDP as estimated by the most recent revision to the Maddison project.

---

<sup>1</sup> Taylor (2002), notes that a trade-weighted index would be ideal, suggests it as a direction for future research, but writes that it would be a "significant undertaking".



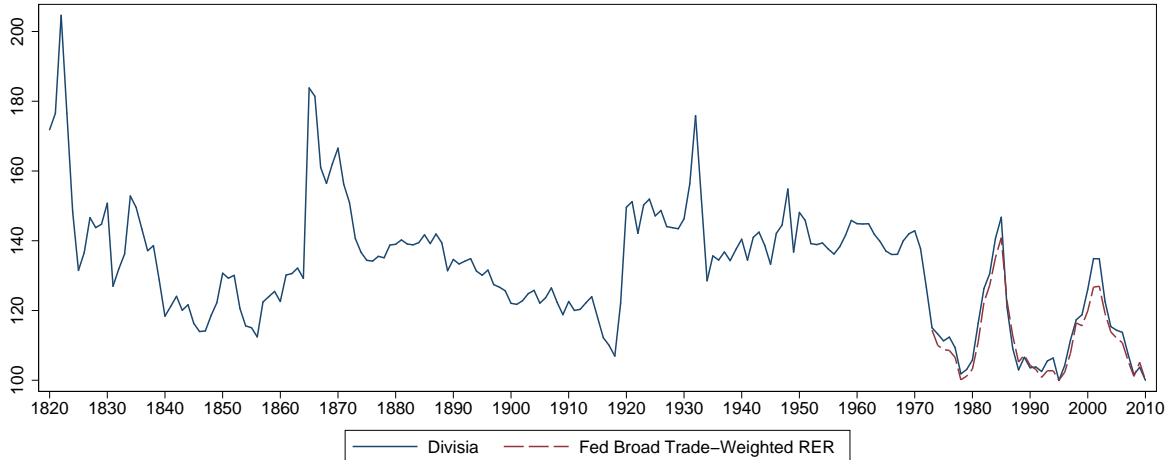


Figure 2: Historical Index Benchmarked to the Post-War Fed Index

### 3 Indexing Methods

#### 3.1 A Review of Divisia vs. WARP

The Fed’s Broad Real Exchange Rate Index is computed as a weighted average of changes in underlying bilateral real exchange rate indices (this method is called “divisia”), where the base year value of each bilateral index is arbitrary. This is the appropriate construction of a nominal exchange rate index, as nominal exchange rates only contain relevant information when movements are plotted over time or when they are compared to relative prices. Real exchange rates, however, do contain information, as they are an indication of the relative price of a basket of goods. As noted in Fahle, Marquez, and Thomas (2008), this information is lost in the Fed’s approach, which is only informative when changes in the index values are plotted over time.

The Fed’s real exchange rate index is:

$$I_t^d = I_{t-1} \times \prod_{j=1}^{N(t)} \left( \frac{e_{j,t} p_t / p_{j,t}}{e_{j,t-1} p_{t-1} / p_{j,t-1}} \right)^{w_{j,t}}. \quad (3.1)$$

Where  $e_{j,t}$  is the price of a dollar in terms of the currency of country  $j$  at time  $t$ ,  $p_t$  is the US consumer price index at time  $t$ ,  $p_{j,t}$  is the consumer price index of country  $j$  at time  $t$ ,  $N(t)$  is the number of countries in the basket, and  $w_{j,t}$  is the trade weight of country  $j$  at time  $t$ . The base year is set at an arbitrary level, both for the index and for each bilateral real exchange rate. The trade weight is a weighted average of each country’s share of imports, exports, and the degree of competition in third markets (trade weights are discussed later in this section).

Note that while directional changes in real exchange rates will affect the index, changes in trade weights between countries with different levels of real exchange rates will not. An issue arises when there is a shift in trade from countries with similar price levels to countries with very different price levels. Table 1 below describes a case in point. It compares several possible real exchange rate indices: a simple weighted average, a geometric average (used by Fahle *et. al.* (2008)), and the Fed's method. In this example, the bilateral real exchange rate for country A varies without trend, while the real exchange rate for country B appreciates substantially over the period relative to the home country. Reflecting this, the Fed's indexing method (also a geometric average) reveals a substantial depreciation. However, at the beginning of the period, the home country is trading mostly with country A (87% with country A implying 13% with country B), which has a similar price level, while at the end of the period a majority of trade is with country B, which has a much lower price level. This is reflected in a simple weighted average, or in the geometric average, which both show that by the end of the period the home country's currency is much higher vs. a weighted average of its trading partners than it was at the beginning of the period.

In addition, the simple weighted average of real exchange rates has an intuitive interpretation. For example, its value of 1.28 in the first year means that the price level is 28% higher at home than in a weighted average of its trading partners, and about 20% lower than in the eighth year. By contrast, the value of 100 for year one using the Fed's method is just an arbitrary number with no economic meaning by itself.<sup>5</sup>

---

<sup>5</sup>Note that while the arithmetic average appears to be easier to intuit than the geometric average, which is less affected by extreme values, instead of using the price of goods in the US relative to countries A and B, I could have used the prices of goods in those countries relative to the US. Inverting the results using the arithmetic mean (which would be the harmonic mean), would yield very different results. By contrast, with a geometric mean, inversion yields the same results. Otherwise I might favor the arithmetic mean, since from a competitive perspective, having a currency overvalued by 20% with respect to one trading partner is probably more damaging than having your currency overvalued by 1% with respect to 20 countries. The arithmetic average will yield the same results for these two cases, while the geometric average will yield a lower value for the first scenario.

**TABLE 1**  
**Comparing Indexing Methods**

<b>Year</b>	<b>Share of Trade with Country A</b>	<b>RER A</b>	<b>RER B</b>	<b>Weighted Average</b>	<b>Geometric Average</b>	<b>Fed's Indexing Method</b>
1	87.0%	0.92	3.68	1.28	1.11	100
2	77.9%	0.97	3.06	1.44	1.26	100.1
3	74.2%	0.89	3.11	1.46	1.23	94.1
4	71.3%	1.11	2.92	1.63	1.46	107.9
5	63.8%	1.25	2.98	1.88	1.71	117.3
6	53.5%	0.97	2.73	1.79	1.57	98.2
7	49.8%	0.85	2.09	1.47	1.34	80.5
8	43.9%	0.94	2.01	1.54	1.44	82.4

The choice of indexing method has dramatically differing implications for the behavior of the index in the special example above, but do examples such as this happen in practice? In fact, the data in Table 1 are real: “country A” is the Euro Area, “country B” is China, the trade shares are taken from the actual trade-weighting scheme used by the Federal Reserve in the creation of its broad trade-weighted real exchange rate series. The real exchange rates are computed using relative prices (specifically, the ratio of the relative price of output-side real GDP via the Penn World Tables v8.0) for every several years between 1990 and 2010.<sup>6</sup> In 1990, the US traded a small amount with China relative to the Euro Area, but by 2010 the US traded more with China, which has a comparatively inexpensive price level. As a result the Fed’s method and a simple weighted average yield dramatically diverging series.

This example illustrates why divisia-based indices can present a very misleading view of competitiveness, since the relative price levels should theoretically matter more for competitiveness than changes in price levels from arbitrary base year values. When a large country with a low price level, such as China, goes from autarky to free trade in a single year, but has a price level that is initially fixed due to capital controls or government control of the nominal exchange rate in concert with nominal rigidities, both intuition and trade models would suggest that this country would pose a competitive threat. This threat is reflected in the WARP index, but it would not be reflected in the divisia index. And while it may be possible to write down a model where the divisia index is more relevant than a simple Weighted-Average of Relative Prices, most standard trade models would favor WARP. And, in fact, China’s opening of trade has posed a competitive threat (see Autor, Dorn, and Hanson, 2013 and Pierce and Schott, 2014),

<sup>6</sup>Note that this is the PWT v8.0 equivalent of the measure of relative prices that Fahle *et al.* (2008) use, as the variable definitions changed slightly in version 8.0 of the PWT as PPP was no longer included by itself.

and Fahle *et al.* (2008) show that WARP seems to do a better job explaining US trade flows.

However, the problem with WARP is particularly salient in the example with China: as China has developed rapidly, we would expect Chinese relative prices to rise as per the Balassa-Samuelson effect. Thus WARP could over or understate the extent of the dollar's appreciation in the 1990s from a practical perspective.

### 3.2 Trade Weights: Fixed Import and Export Share Weights vs. Adjustable

Another issue in choosing an exchange rate index is the choice of trade weights. The Fed's trade-weighting scheme is a weighted average of three measures – import exposure, destination-market export exposure, and the degree of competition in third markets. It attaches weights symmetrically – 50% to imports and 25% each to destination export markets and competition in third-country markets. Of course, since U.S. trade is often unbalanced, a preferable approach, used by the Bank of International Settlements (Klau and Fung, 2006, and also preferred by Chinn, 2006), is to decide the import weight based on the share of imports in total trade. The Federal Reserve's own documentation admits that its trade-weighting scheme is arbitrary on the grounds that moderate adjustments in the direction of plausibility add complication but seem to have little impact (Loretan, 2005).

The Fed's trade weight for country  $j$  at time  $t$  is:

$$w_{j,t}^{FED} = \frac{1}{2}\mu_{us,j,t} + \frac{1}{2}\left(\frac{1}{2}\epsilon_{us,j,t} + \frac{1}{2}\tau_{us,j,t}\right). \quad (3.2)$$

Where  $\mu_{j,t}$  is country  $j$ 's share of merchandise imports into the US, equal to:

$$\mu_{us,j,t} = M_{us,j,t} / \sum_{k=1}^{N(t)} M_{us,k,t}. \quad (3.3)$$

$M_{us,j,t}$  are imports from country  $j$  for  $N(t)$  markets at time  $t$ . The bilateral export share,  $\epsilon_{us,j,t}$ , is defined as:

$$\epsilon_{us,j,t} = X_{us,j,t} / \sum_{k=1}^{N(t)} X_{us,k,t}. \quad (3.4)$$

Where  $X_{us,j,t}$  are exports from the US to country  $j$  at time  $t$ . Since US exports also compete with the exports of country  $j$  in third-market economies,  $\tau_{us,j,t}$  measures the

competitiveness in third markets:

$$\tau_{us,j,t} = \sum_{k \neq j, k \neq us}^{N(t)} \epsilon_{us,k,t} \mu_{k,j,t} / (1 - \mu_{k,us,t}). \quad (3.5)$$

Where  $\mu_{k,j,t}$  is the fraction of country  $k$ 's merchandise imports from country  $j$  in year  $t$  and the multiplicative factor  $(1 - \mu_{k,us,t})$  ensures that the weights sum to 1.

Instead, I propose using the share of imports in total trade, denoted  $m_t$ , for the import weight. I then follow the Federal Reserve Board in splitting the export weight between the destination and third-country markets.

$$w_{j,t}^{Alt} = m_t \mu_{us,j,t} + (1 - m_t) \left( \frac{1}{2} \epsilon_{us,j,t} + \frac{1}{2} \tau_{us,j,t} \right). \quad (3.6)$$

When the alternative trade-weights are used which reflect the actual composition of imports and exports in trade, I confirm Loretan's (2005) finding that the results are little-changed, and thus for simplicity report all results other than the benchmarks using these adjusted trade weights. The main reason why the results are not systematically different is because the differences in the trade-weights between these two methods are fairly minor. However, this does not imply that using fixed trade weights instead would also not matter. In that case, the trade weights and results would be substantially altered, particularly for the class of Weighted-Average Relative (WAR) exchange rate indices.

A very prudent second critique is that the Fed's trade-weights measure trade in goods rather than trade in value-added. Bems and Johnson (2012) show that for the US, the differences in trade shares using value-added measures of trade seems to make little difference.<sup>7</sup> For example, they find that the US trade share with China shrinks by just -.2% in 2005 when they use value-added measures of trade as compared with aggregate trade (ex-oil).

A third critique was mounted by Ho (2012), who proposed using GDP weights instead of trade weights, and found some support that in many cases (although not for the US), the GDP weights do a better job of explaining real exports using cointegration analysis. Thus, following Ho (2012), I also provide GDP-weighted versions of my index, which actually differ more substantially than trade-weighted indices for the class of weighted-

---

<sup>7</sup>While there is very good logic for why one would want to use trade in value-added rather than actual trade, the downside is that it effectively ignores trade in intermediate inputs. Thus the domestic value-added share of a highly fragmented production processes could see greater volatility in value-added for a given movement in the exchange rate since multiple decisions must be made whether to import or buy domestically.

average relative indices proposed here.

A fourth critique I owe to an anonymous referee, who critiqued the methodology on the grounds that theoretically, we should use physical trade *volumes* rather than aggregate trade values. However, there is a reason why the Federal Reserve, the IMF, the World Bank, the OECD, the BIS, and Fahle *et al.* all used trade values instead of volumes: you would still need some way to compare apples and oranges, and using prices is a very natural choice of weights, and would seem to make more sense than physical weight.

### 3.3 Post-War WARP for the United States

Weighted average relative prices (WARP) are computed as a geometric weighted average using trade-weights,  $w_{j,t}$ , of the nominal exchange rate,  $e_{j,t}$ , divided by purchasing power parity,  $PPP_{j,t}$ :<sup>8</sup>

$$I_t^{WARP} = \prod_{j=1}^{N(t)} \left( \frac{e_{j,t}}{PPP_{j,t}} \right)^{w_{j,t}} = \prod_{j=1}^{N(t)} \left( RER_{j,t} \right)^{w_{j,t}}. \quad (3.7)$$

The nominal exchange rate divided by PPP (the ratio of the local currency price of a basket of goods) yields the ratio of the price level in the US to the price level in the comparison country, also called the relative price. Thus if the nominal exchange rate between the US and China is 8 to 1, but the cost of a given basket of goods costs 2 Yuan in China but only a dollar in the US, then the price level in the US will be four times the price level in China.

In Figure 3 I compare WARP using version 8.0 of the PWT to four alternatives: (1) WARP using version 7.1 of the PWT, (2) WARP using GDP weights as suggested by Ho (2012), (3) a divisia index using CPIs with the Federal Reserve’s Methodology, and (4) divisia using the PWT. WARP version 8.0 has lower US relative prices in every period compared to version 7.1, with the difference becoming less pronounced in the 1990s (as might have been expected). Consequently, version 8.0 has a larger relative price increase for the US for the period 1992-2002, which registered as a much larger shock to relative prices than the 1980s dollar appreciation, which appears to be more moderate in version

---

<sup>8</sup>Older versions of the PWT include a PPP variable, but version 8.0 contains price levels (PPP divided by the nominal exchange rate). Thus it is easier to compute relative prices directly using the ratio of the price-level of output-side US GDP to the price level of output-side GDP in other countries (using expenditure-side GDP instead leads to very similar results). I thus use the same measure of relative prices at Fahle *et al.* Also note that the PWT creates estimates of the price level annually which are based on the ICP benchmark years with years in between interpolated using country-specific deflators.

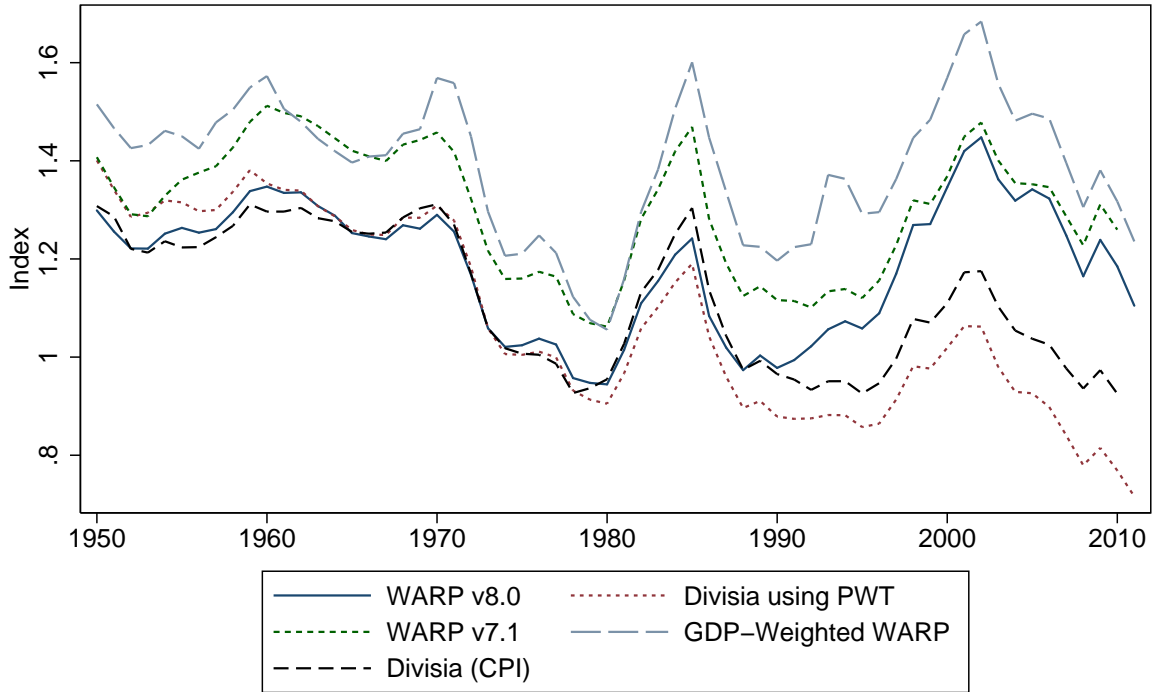


Figure 3: WARPs vs. Divisias

8.0. The more recent version also does not include the substantial dollar appreciation during Bretton Woods that is seen in the previous version of the PWT.

I also find that divisia using PWT implies a further dollar depreciation than divisia using country CPIs, matching the finding of Fahle *et al.* (2008). For ease of comparison the Divisia using the CPI (essentially the Fed's series) is multiplied by a scaling factor so that it begins at the same level as the WARP in 1973, which gives the Fed's series base year an intuitive economic meaning – in 1973, the U.S. price level was about 30% higher than a (geometric) weighted average of U.S. trading partners. WARP v8.0 approximates the Fed's index up until the dollar appreciation in the 1980s, when it shows less of an appreciation (this was much less apparent in version 7.1 of the PWT). Since the early 1990s, the WARP index reveals a much larger appreciation relative to the Fed's index, appreciating 26% more from 1990-2002. From 1990-2011 WARP appreciated by 12.9% versus a 9% *depreciation* according to divisia. The divisia index computed using the PPP of output from PWT v8.0 is very similar to that using expenditure-based PPP, and also very similar to using World Bank GDP deflators, as used in the construction of value-added exchange rates (Bems and Johnson, 2012, and Bayoumi *et al.*, 2013).

### 3.4 A WARPed View of US Real Exchange Rate History

This paper is the first to plot weighted average relative prices for the U.S. before 1970, adding 150 years of data to the Fahle *et. al.* (2008) series. This series uses PWT v8.0 as the basis of all bilateral relative prices from 1950, and before that extrapolates historically using data sources detailed in Table A.3.

How does WARP change our view of history? The major difference is that in the WARP series, the price level was lower in the interwar period and in the period before World War I than divisia relative to the Bretton Woods period. In the divisia series, the interwar price level was generally higher. Another difference is that the divisia series exhibits a more dramatic depreciation in the period following the Civil War and just before the return of the US to the gold standard.

Outside of the period around World War II, when trading partners rapidly evolved and the sample temporarily shrinks, the largest disparity between divisia and WARP comes after 1990. This implies that the recent rise of China and its impact on average relative prices is unprecedented in US history. During Japan's more methodical rise up until the 1980s (see the trade shares in Figure 7), the US did not experience the same large increase in relative prices. One small difference is that the WARP view of history reveals a slightly sharper appreciation during the Great Depression, with a 25.2% appreciation from 1929-1932 vs. a 22.2% appreciation using a divisia-based method from 1928-1932. In figure 2 in the previous section, one of the striking features of the long run history of the dollar is that prices in the U.S. in 2010 were as low relative to trading partners as they had ever been. This is no longer the case with the WARP index, as prices in the US prior to the Civil War were much cheaper relative to trading partners than they are today, although extreme caution is warranted when interpreting historical data that result from using country-specific deflators over long periods of time.

The historical estimates of price-levels used to create the WARP index were created by starting with Penn World Table estimates and then extrapolating backward using the best historical estimates available. For the most important trading partners such as the UK, I then checked relative prices with studies conducted by economic historians (Devereux and Ward, 2003 and 2006) in benchmark years. Reassuringly, for most of this period, my relative price estimates never differ by more than 10%. However, there is a difference in the late 19th century that widens up until World War I, with Devereux and Ward (2003) finding higher US prices. For Japan, my price estimates (Table A.3 in the appendix) are very similar to Allen *et al.* (2011).

My estimate for the real exchange rate between the US and the UK in 1925 is also



reassuringly very close to John Maynard Keynes's estimate in that year. As Chancellor of the Exchequer, Winston Churchill decided to return to the gold standard at the prewar parity, citing wholesale PPP data which reportedly showed the price differentials between America and Britain were similar. Keynes's alternative PPP series based on retail prices and wages showed that a return to gold at the prewar parity would result in a pound overvalued by about 12% (Skidelsky 2005). My series estimates that the pound was overvalued by just 9%.

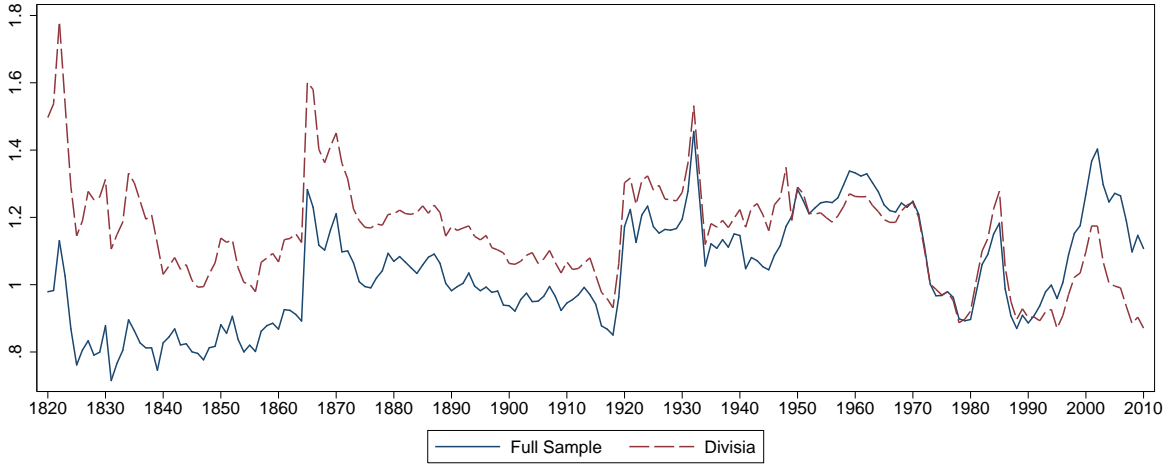


Figure 4: WARP vs. Divisia, 1820-2010

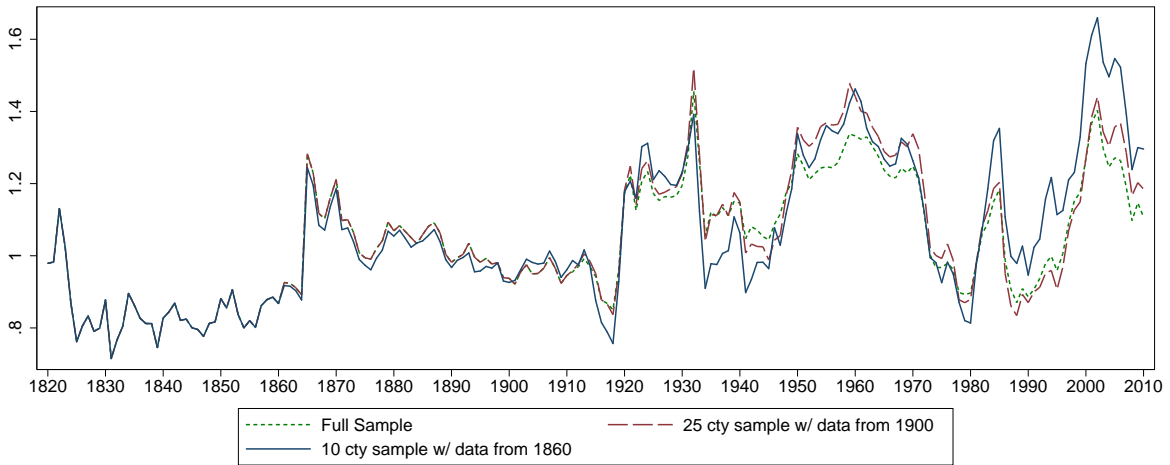


Figure 5: WARP, Full Sample vs. Smaller Balanced Samples

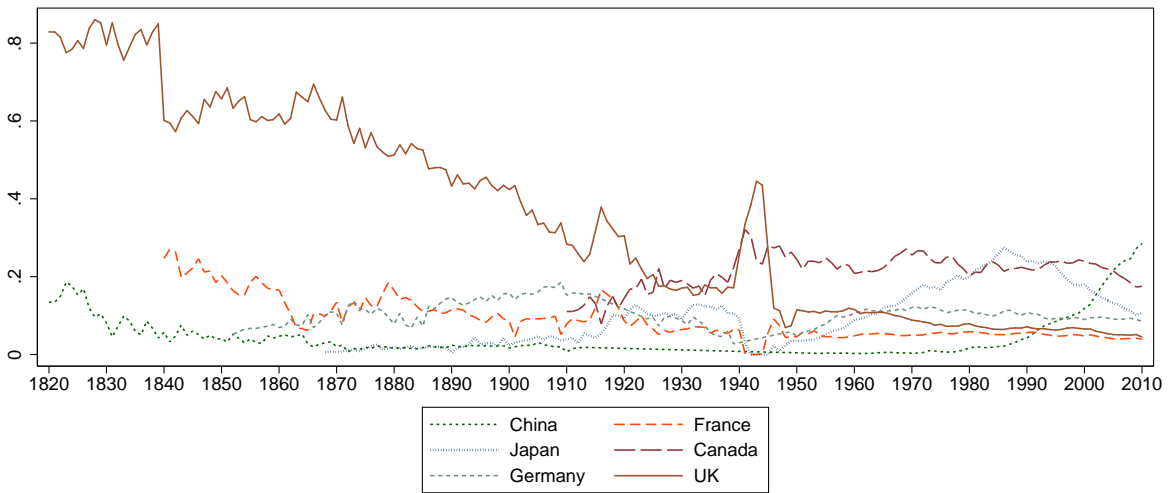


Figure 6: Trade Shares in Sample, 1820-2010

## 4 Balassa-Samuelson Productivity Adjustment

The WARP index, while likely preferable to the Fed’s series for the purposes of measuring the competitiveness of U.S. goods and services in international trade, may not be the optimal method since it only factors in prices and not productivity. The Balassa-Samuelson (or the “Penn”) effect implies that if traded sectors have relatively fast productivity growth, then the productivity differentials between rich and poor countries will be greater in the tradable sectors. In each country wages in the non-traded sector will be bid to equality with wages in the traded sector, which depend on productivity, and hence non-traded prices in less developed countries will be lower reflecting their lower productivity.

Intuitively, if Afghanistan has low prices, but also low productivity, then its low price level does not necessarily indicate a competitive threat. The original motivation for this paper was thinking about Japan in 1985 vs. 1960. Japan was a much more productive economy in 1985, but the Japanese price level relative to the US had not appreciated much from the 1960s. Given its higher productivity, it is clear that Japan in 1985 should have been a larger competitive threat than Japan in 1960, which it was. Thus, what matters for competitiveness is the level of real exchange rates relative to some measure of productivity, such as per capita GDP.

Theoretically, the key result from the Balassa-Samuelson model (borrowing the setup from Obstfeld and Rogoff), in a 2 sector model, with tradables and non-tradables sector, and with a tradables sector price which is the same everywhere due to the law of one price, is that the ration of the price of non-tradables to tradables is the ratio of tradable to non-tradable productivity.

$$p_N/p_T = \theta_T/\theta_N \tag{4.1}$$

Thus, countries with relatively higher tradable-sector productivities will have higher non-tradable prices. The overall price level includes both tradable and non-tradable goods, and so the relative price of US goods to foreign goods for country  $i$  can be written as:

$$RP = \frac{p_T s_T + p_N (1 - s_T)}{e p_T^* s_T + e p_N^* (1 - s_T)}, \tag{4.2}$$

where  $e$  is the nominal exchange rate, or the dollar price of foreign currency,  $s_T$  is the share of tradables, and foreign prices are denoted with an asterisk. We can rewrite the

equation for relative prices using equation (4.1):

$$RP = \frac{(s_T + \theta_T/\theta_N(1 - s_T))(p_T)}{(s_T + \theta_T^*/\theta_N^*(1 - s_T))(ep_T^*)} \quad (4.3)$$

In equilibrium, the price of tradables are the same in the US and abroad ( $p_T = ep_T^*$ ), and thus the term on the right of equation 4.3 will cancel. This relative price, which depends on the ratio of relative tradables-to-nontradables productivity levels for the US vs. trading partners, has historically been very strongly correlated with relative per capita GDP. However, the nominal exchange rate moves much more sharply, even within the course of a day, than local retail prices. Thus, if the nominal exchange rate moves from its equilibrium value, but local-currency price of goods were rigid, then the relative price could also deviate from its equilibrium level. Thus, a sudden dollar appreciation (or a decline in the dollar price of a foreign currency) would lead to US prices which are also high relative to their equilibrium value, indicating overvaluation.

This is essentially the logic behind a large body of research which studies the impact of exchange rates on the economy using the Balassa-Samuelson residual, including Rodrik (2008), Johnson, Ostry, and Subramanian (2008), and Cheung, Chinn, and Fujii (2010), among many others. While these papers use only a residual, I propose the following trade-weighted average of differences in Balassa-Samuelson residuals:

$$I_t^{BSWARP} = \prod_{j=1}^{N(t)} (\exp(\epsilon_{US,t} - \epsilon_{j,t}))^{w_{j,t}}. \quad (4.4)$$

Where  $w_{j,t}$  are trade weights, and  $\epsilon_{US,t}$  and  $\epsilon_{j,t}$  are the residuals for the US and country j from the Balassa-Samuelson regression (used by Rodrik, 2008):

$$\ln RER_{j,t} = \alpha + \beta * \ln RGDP_{j,t} + \sum_{t=1950}^{2010} f_t + \epsilon_{j,t}. \quad (4.5)$$

Where  $RER_{j,t}$  is the real exchange rate vs. the dollar for each country in the world (in this case, the RER is defined such that larger numbers indicate a higher price level for country j relative to the US),  $RGDP_{j,t}$  is the real GDP per capita, and  $f_t$  are year fixed effects. The regression yields a coefficient on log GDP per capita of .133 for 186 countries for the period 1973-2011 (a smaller, balanced sample yields a similar estimate). The residual  $\epsilon_{US,t}$  has a simple economic meaning – it tells us how over or under-valued the dollar is relative to where it should be based on US GDP per capita. This number is then adjusted based on the relative valuation of US trading partners. The advantages of this adjustment should be obvious – the US trades much more with

countries such as Canada and Mexico, and thus the US's valuation relative to its largest trading partners should matter more than the residual by itself. If the US and each of its trading partners were to lie on the Balassa-Samuelson regression line, then the index would be zero, indicating that the dollar is fairly valued.

One can see the relationship between divisia, WARP, and BS-WARP by totally differentiating the log of the BS-WARP index:

$$d \ln I_t^{WARP} = \underbrace{-\sum_i w_{it} d \ln(RER_{j,t})}_{\text{Divisia}} + \underbrace{\sum_i d w_{i,t} \ln(RER_{j,t})}_{\text{WARP}} \quad (4.6)$$

$$-\beta d \ln(RGDPPC_{US,t}) + \beta \sum_i w_{i,t} d \ln(RGDPPC_{i,t})$$

Thus the BS-WARP index has four distinct channels which can change the index. The first term tells us that when bilateral real exchange rates move, this affects the index. This term is equivalent to the changes in the divisia index. The second term indicates that shifts in trade patterns also affect the index. The first two terms combined yield the changes from WARP. The third and fourth terms are new—they indicate that when US real GDP per capita increases, this will reduce the BS-WARP index holding the real exchange rate equal, and when real GDP per capita rises in US trading partners, the index will appreciate if prices do not also adjust as expected based on the Penn effect.

The Balassa-Samuelson-adjusted weighted average relative price (BS-WARP) index for the US using is plotted in Figure 8 below vs. the Fed's benchmark. The magnitude of the dollar's appreciation from 1990-2002 is much larger than the Fed's index, with an appreciation of 44.8% for the BS-WARP index to just 21.7% for the Fed's index (and 48% for WARP). Over the period 1990-2011, the BS-WARP index appreciated about 10.8% vs. 12% for WARP, while the Fed's index fell by 9.4%. The major difference between the BS-WARP and WARP indices for the US is that the BS-WARP index implies a much lower valuation for the dollar in all periods, with the dollar on average just 4% more richly valued than trading partners vs. an average of 19% using WARP for the period 1950-2011. This implies that while US prices are higher on average than prices in US trading partners at present, this is accounted for by high relative GDP per capita in the US, as shown by the Balassa-Samuelson adjusted series.

Figure 9 decomposes the degree of over or undervaluation in exchange rates according to the Balassa-Samuelson regression between the US, US trading partners, and China. In the 1992-2002 period, the US residual and the average residual of US trading partners both appreciated significantly. China became much less undervalued in this period,

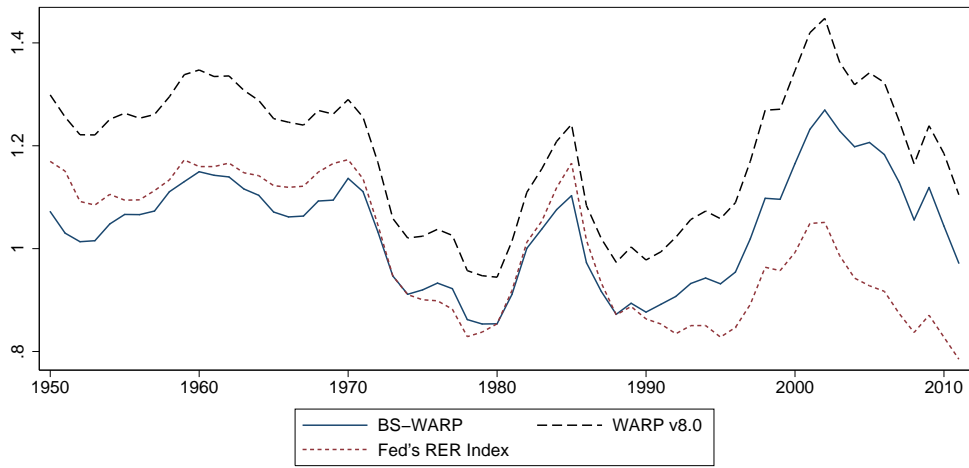


Figure 7: BS-WARP vs. WARP vs. Divisia

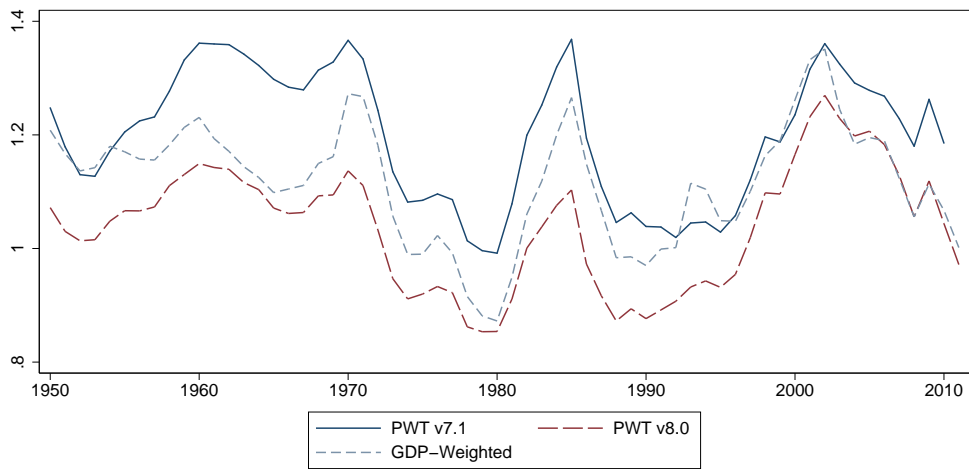


Figure 8: Variations of Balassa-Samuelson Adjusted WARP

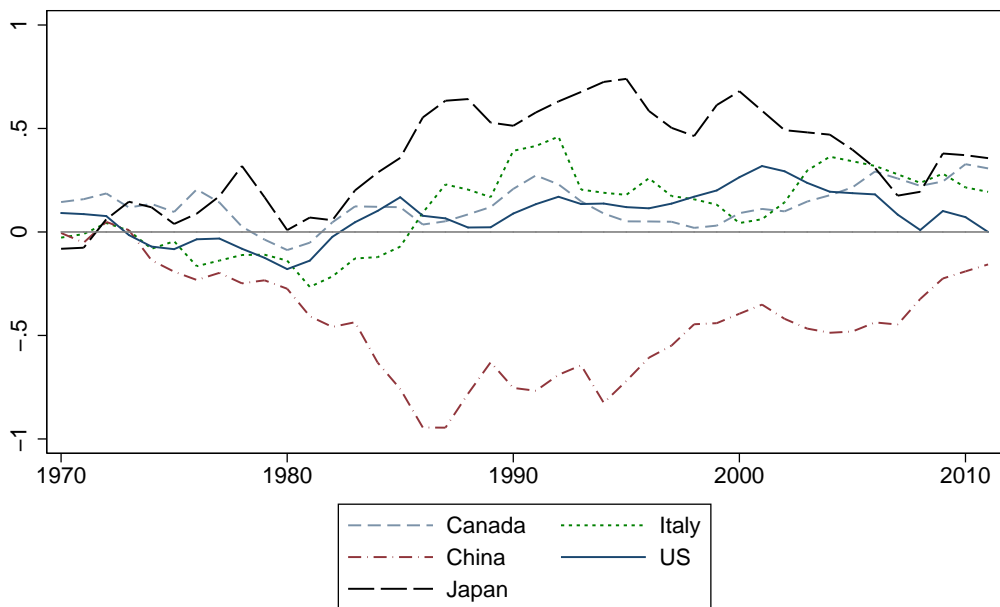


Figure 9: US Balassa-Samuelson Coefficient vs. Trading Partners

but rapid GDP growth and a generally weak dollar meant that while the Renminbi appreciated significantly vs. the dollar after 2000, China's Balassa-Samuelson residual was essentially unchanged until 2008.

One alternative to the BS-WARP index with a fixed Balassa-Samuelson coefficient would be to let the Balassa-Samuelson coefficient vary by year, as implemented by Johnson, Ostry, and Subramanian (2008). A compelling reason to discount this method *a priori* is that Feenstra, Inklaar, and Timmer *et al.* (2013) convincingly show that the apparent growing Balassa-Samuelson effect discovered by Bergin, Glick, and Taylor (2006) using version 6 of the PWT is a statistical artifact of extrapolated data. Data from benchmark years of the International Comparison Project always show a similarly-sized Balassa-Samuelson effect, although there is year-to-year variation. For this reason, and since the results are similar to the constant-coefficient BS-WARP index since 1970, I would refer interested readers to previous drafts of this paper.

An alternative to using indices based on PPP data from the Penn World Tables would be to use data on export prices relative to import prices, which is also called the “terms of trade”.<sup>9</sup> The BEA's terms of trade index for the US does suffer from same index numbers problem as other RER indices, although this problem should theoretically be corrected for export and import price indices provided in PWT v8.0, at least for the period 1984 to 2007, as prices are computed in this period using “EKS aggregation” (Feenstra *et al* 2013). However, the goal in this paper is to create an index that summarizes the change in competitiveness due to shocks to the real exchange rate, which primarily arise from changes in the nominal exchange rate. Even sharp changes in the nominal exchange rate, which do tend to impact trade flows, do not necessarily affect the terms of trade, as a dollar appreciation could make imports and exports cheaper in dollars. Traditionally, for the US, it has been changes in oil prices which have had the largest affect on the terms of trade, while many users of RER indices are interested in the impact on trade ex-oil. For these reasons, the terms of trade are a less than ideal measure of competitiveness, while, for the US at least, there is generally no clear relation between aggregate trade flows and the terms of trade.

Other possibilities would be to use producer or wholesale prices rather than consumer prices, which are what the Penn World Tables provides based on ICP data. Producer or wholesale prices, in fact, might be preferable *ceteris paribus* since, as Chinn (2006) notes, they exclude non-tradable retail services inputs, but the problem is comparable data availability. While one can get country-specific producer price indices for many

---

<sup>9</sup>I am indebted to an anonymous referee for suggesting this.

countries, these series are indices and so do not contain necessary information on levels.<sup>10</sup> And even if they did contain information on levels for a single year, when compared over long-time horizons with only a single benchmark year these indices will become biased over time, suffering from the same problem as versions of the Penn World Tables which predated v8.0. I am not aware of any large cross-country producer price comparison projects which seriously rival the ICP in terms of breadth and coverage.<sup>11</sup>

## 5 Unit Labor Costs

### 5.1 Theory

It is generally thought to be preferable on theoretical grounds to use wages rather than prices when measuring competitiveness, particularly for the manufacturing sector (Turner and Van't Dack, 1993). Intuitively, this is because labor is perhaps the largest non-traded input into the manufacturing sector, and labor tends to be at least somewhat immobile. Thus, the level of wages should naturally be one key determinate for firms deciding where to produce. And, indeed, the levels of wages matter for competitiveness in almost all standard trade models, including Heckscher-Ohlin, the “Specific Factors” Model, the Melitz (2003) model, and in Ricardian Models of trade, including Eaton and Kortum (2002), among many others.<sup>12</sup>

This explains the popularity of using unit labor costs in manufacturing to gauge competitiveness. Similar to the CPI-based REER produced using “divisia”, the ULC indices produced by the IMF and the OECD also are computed as indices of bilateral unit labor cost indices and thus also suffer from trading partner substitution bias. In addition, manufacturing PPP for each country is deflated using country-specific deflators, exactly as older versions of the Penn World Tables, which predated version 8.0, were computed. The series I propose is thus a simple Weighted Average of Relative Unit Labor Costs (WARULC) rather than of the unit labor cost *indices* – analogous to WARP. In my series, I compute manufacturing PPP using PWT v8.0 methodology described in Feenstra *et*

---

<sup>10</sup>I thank an anonymous referee for suggesting I use producer prices, and a discussant, Tadashii Ito, for suggesting using sectoral producer data collected by Sato *et al.* (2012). However, the Sato *et al* data covers just 7 years and 26 countries, and does not include data on relative price levels.

<sup>11</sup>It is also unlikely that there are any other cross-country consumer price data which could seriously rival the ICP. For example, MIT’s billion prices project does not contain information on quantities and thus also suffers from substitution bias.

<sup>12</sup>I am indebted to an anonymous referee for the Review of Economic Studies who argued that wages should not matter for competitiveness, only output prices, as this demonstrated that this point should be explained in more detail.



*al.* (2013). When I also expand the sample to include developing countries such as China and time-varying trade-weights, the differences in the underlying index become substantial. This is due to China’s systematically lower unit labor costs and growing weight in manufacturing trade over time.

The IMF’s RULC index, documented by Desruelle and Zanello (1997), is computed as:

$$I_{US,t}^{RULC} = \prod_{i=1} \left( \frac{C_{US}^I R_{US}^I}{C_i^I R_i^I} \right)^{w_i} \quad (5.1)$$

Where  $C_i^I$  is the normalized unit labor cost index for country  $i$ , computed as the ratio of nominal sectoral wages to real productivity,  $R_i$  is the nominal exchange rate index, and  $w_i$  are the time invariant trade weights. One intuitive proposed alternative specification would be to replace the unit labor cost indices with the same unindexed unit labor costs, and actual nominal exchange rates. However, the relative unit labor costs using deflated real productivity will depend on the base year used to deflate productivity. To circumvent this problem, I convert nominal productivity into dollars using the PPP exchange rate conversion for the manufacturing sector, following the method Ceglowski and Golub (2007) implement for just the US and China, while converting nominal wages into dollars at the nominal exchange rate.<sup>13</sup>

For this index, I used OECD data created specifically for the construction of ULC indices where available, and supplemented this with data from the BLS, the Chinese government, the World Bank’s WDI, and UNIDOs. The manufacturing PPP data for benchmark years come from the relevant manufacturing ICP headings and were computed using PWT methodology (described in Section 3 of Feenstra *et al.* 2013), and were interpolated in the intervening years using country-specific deflators from either the OECD, or country-specific sources (in the case of China). The methodology and formulas for the PPP interpolation were also borrowed from Feenstra *et al.* (2013). *E.g.*, after the last benchmark year in 2005, the series are extended based on country growth rates for country  $i$ :

$$P_{i,2006} = P_{i,2005}^{ICP} * \frac{P_{i,2006}^{i,deflator}}{P_{i,2005}^{i,deflator}}, \quad (5.2)$$

where  $P_{i,t}^{i,deflator}$  is the country-specific deflator at time  $t$ .

---

<sup>13</sup>Since the index is of *relative* unit labor costs, and as Rudiger Dornbusch used to say, two nominals make a real, the use of nominal wages converted at exchange rates is not problematic.

For the years in between ICP benchmarks, a weighted average was used. For example, for the years between 1996 and 2005, the formula is:

$$P_{i,t} = P_{i,1996}^{ICP} * \frac{P_{i,t}^{i,deflator}}{P_{i,1996}^{i,deflator}} \left( \frac{2005-t}{2005-1996} \right) + P_{i,2005}^{ICP} * \frac{P_{i,t}^{i,deflator}}{P_{i,2005}^{i,deflator}} \left( \frac{t-1996}{2005-1996} \right). \quad (5.3)$$

Data on manufacturing trade to create trade weights  $\omega_{i,t}$  is computed from bilateral manufacturing data at the SITC 4 level from Feenstra *et al.* (2005), and with updated data through 2008 via direct communication with Feenstra. Manufacturing trade data from 2009 and 2010 were taken from the OECD.

The weighted average relative unit labor cost index (WARULC) is computed as:

$$I_{US,t}^{WARULC} = \prod_{i=1} \left( \frac{C_{US,t}}{C_{i,t}} \right)^{\omega_{i,t}} = \prod_{i=1} \left( \frac{\frac{w_{US,t}}{e_{US,t}} / \frac{Y_{US,t}}{PPP_{US,t}}}{\frac{w_{i,t}}{e_{i,t}} / \frac{Y_{i,t}}{PPP_{i,t}}} \right)^{\omega_{i,t}}. \quad (5.4)$$

Where  $w_{i,t}$  are manufacturing wages of country  $i$  at time  $t$ ,  $e_{i,t}$  is the nominal exchange rate to convert to dollars,  $Y_{i,t}$  is manufacturing production, and is divided by PPP for the manufacturing sector. Thus the  $C$ 's in this equation are actual unit labor costs rather than indexes of unit labor costs.

## 5.2 Data

When the Weighted Average Relative Unit Labor Cost (WARULC) index is compared with the official IMF RULC index (indexed to start at the same value in 1975) and an index using my data but the IMF's index-of-indices method, the results are strikingly different, with the difference much larger than the disparity between WARP and divisia computed with CPIs. The series are roughly similarly until the late 1980s, but by 2001, the WARULC index is 32% higher than the IMF's index, and 44% higher in 2008. The IMF benchmark index constructed here using the IMF's index-of-index method is similar to the IMF's index, despite the fact that I used time-varying manufacturing trade weights, a larger sample of countries (including China), and I compute manufacturing value-added using PPP. The IMF instead uses an index of real output measured in the home currency, and so it is striking that the benchmark is similar to both the IMF and the OECD indexes (the latter is not shown but also similar). I have also plotted a WARULC series which uses manufacturing PPP computed using only a single benchmark year and country deflators (short maroon dashes in figure 10). The series without multiple benchmarks displays a downward trend relative to my preferred series

with multiple benchmarks.

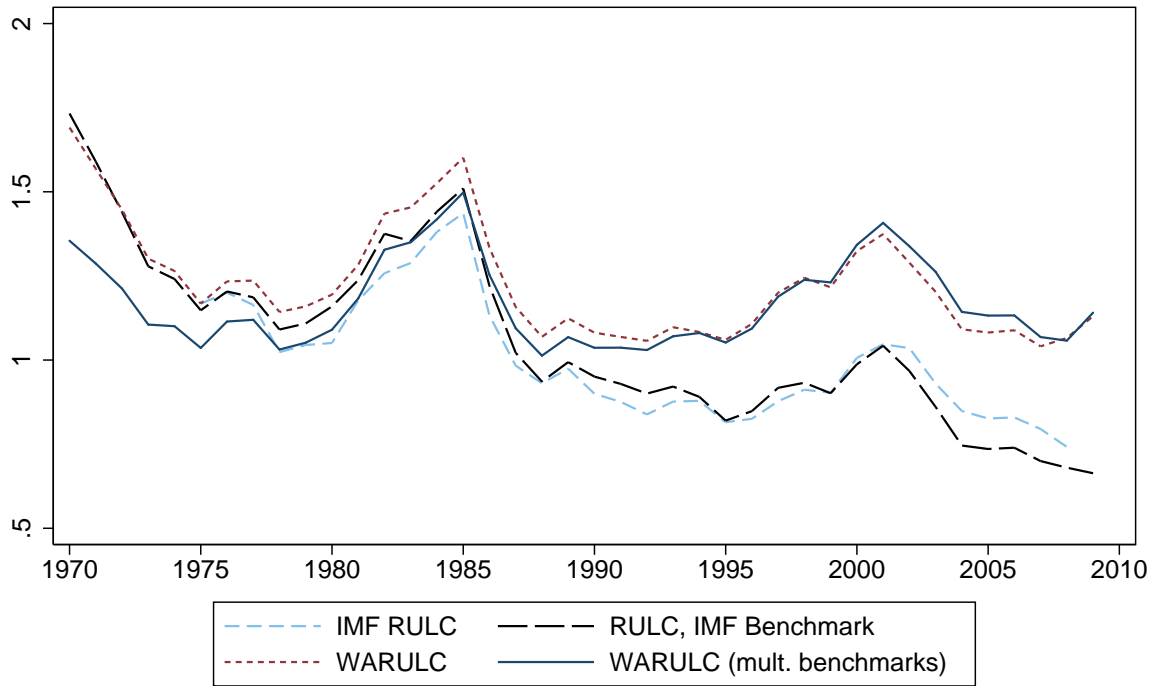


Figure 10: IMF Method vs. WARULC

And, just as with WARP, the difference between WARULC and the IMF’s index is largely China, as evidenced in Figure 11(a). In 11(b) I compare WARP, BS-WARP, and WARULC, and find that they are all broadly similar, with the exception being that the WARULC index displays a sharper depreciation after 2001, and that the BS-WARP index implies that the dollar has generally been less over-valued as compared to WARP. Note that there aren’t necessarily any pure theoretical grounds to prefer any of these three indices, as they all have their own strengths and weaknesses. Comparing WARP to BS-WARP, one might prefer BS-WARP on the grounds that it controls for productivity, but WARP on the grounds that it is more elegant (in any case, it is a distinct measure). WARULC controls for productivity explicitly, and may be better designed for the manufacturing sector itself (though not for other tradable sectors), but given that it does not include other non-traded inputs, including services inputs, into manufacturing aside from labor, it can not be said to be a better measure *a priori*.

Figure 12 details estimates of relative hourly productivity, wages, and ULCs for the US relative to China. These results are very similar to Ceglowski and Golub (2012) for the 1998-2009 period. The ratio of hourly wages has indeed fallen dramatically since the early 1990s, but not much more quickly than the convergence in productivity. Relative

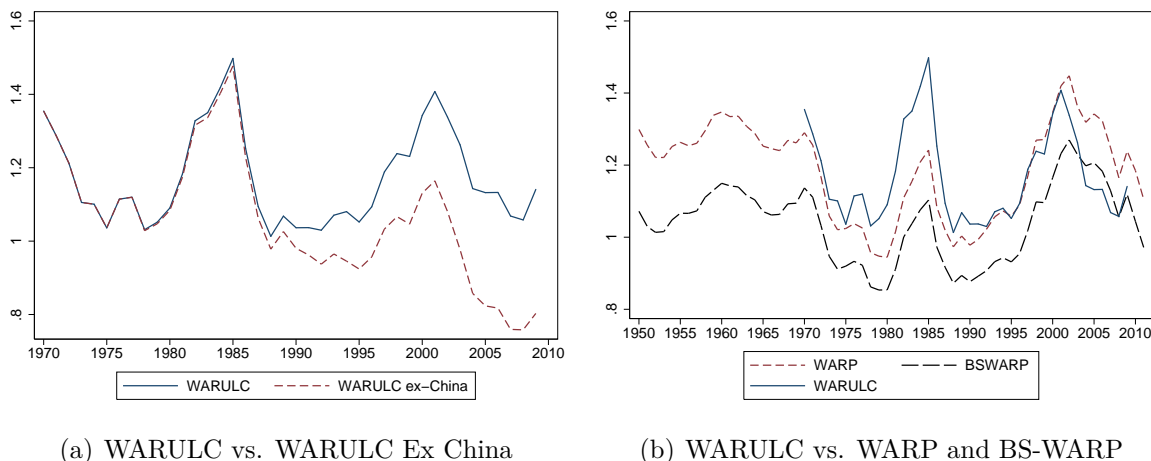


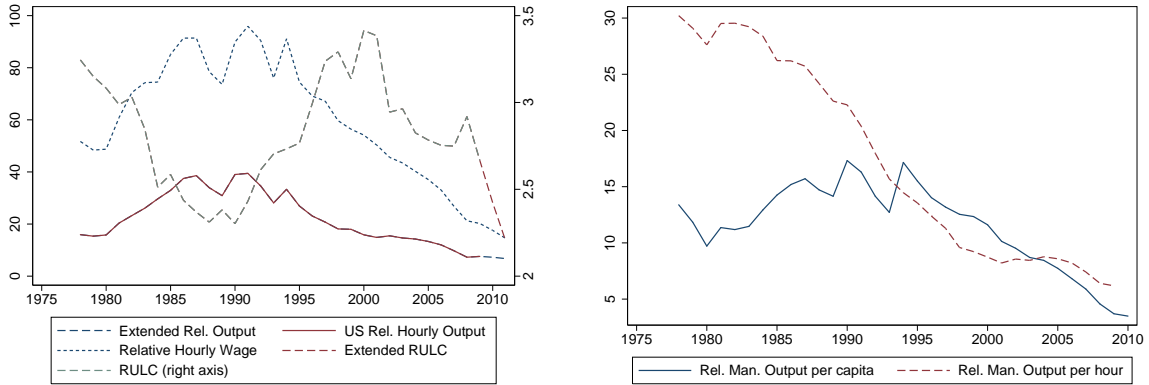
Figure 11: Comparing WARULC

unit labor costs spiked in the late 1990s and early 2000s, during the collapse of US manufacturing employment concentrated heavily in China-competing industries. As of 2009, US manufacturing wages were still approximately 20 times larger than Chinese manufacturing wages, while unit labor costs in the US were about 2.6 times higher. In 2010, using data on just output and wages, with the strong assumption that hours worked stayed constant, I can estimate that the RULC did narrow substantially.

It is worth noting that the hiatus in the downward trend in relative productivity between China and the US during the 2000-2004 period (Figure 12(b)) came amidst a collapse in employment in manufacturing in the US biased toward low-productivity firms and sectors in the US and was accompanied by a large growth in relative Chinese manufacturing output. This is apparent in Figure 12(b), where it is clear that relative manufacturing output *per capita* for the US vs. China fell at a steady pace during the period in which hourly relative output was stable. One solution to this problem would be to do ULCs by sector, but this would require at the very least sectoral output and wage data for China, and also PPP data and sectoral deflators for specific manufacturing sectors. In addition, it is likely that even within narrowly defined manufacturing sectors, low-productivity firms were more likely to lose workers, and even within firms, low-productivity workers were more likely to lose their jobs.

In June of 2013, the Boston Consulting Group released a report on relative unit labor costs in manufacturing between China and Mexico, finding that unit labor costs in China exceeded unit labor costs in Mexico beginning in 2011.<sup>14</sup> However, as of 2009, I find that

<sup>14</sup>BCG press release, 6/28/2013, available here: <http://www.bcg.com/media/PressReleaseDetails.aspx?id=tcm:12-139022>.



(a) Relative Wages and Productivity

(b) Output vs. Productivity

Figure 12: US vs. China

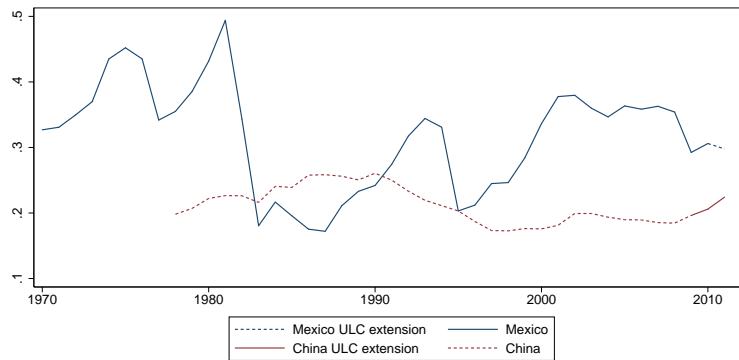


Figure 13: ULCs, Mexico vs. China

ULCs in Mexico were still about 50% larger than the ULCs in China (Figure 13). While I do not have access to complete data on manufacturing employment and hours worked through 2011, hourly dollar wages in China were just \$2.52 vs. \$6.48 in Mexico in 2011, and total Chinese manufacturing production rose 27% vs. just 17% for Mexico from 2009-2011. If there was no change in relative hours worked, admittedly a very strong assumption, then ULCs did converge a bit between 2009 and 2011, but Mexican ULCs were still roughly 33% higher than Chinese ULCs in 2011. Thus Mexican manufacturing workers, in the aggregate, are still substantially more productive than their Chinese counterparts, although also better paid relative to productivity.

## 6 WARULC vs. IMF RULC: Empirical Tests

The main point of this paper is to introduce new and improved measures of relative prices. I would argue that theoretical and intuitive concerns should dictate the choice of RER index, and that the choice should not necessarily depend on which index matches the data the most closely.<sup>15</sup> Thus, correcting the “trading partner substitution bias” problem and the problems arising from country-specific deflators, and including China are more persuasive reasons to prefer WARULC to the IMF’s RULC index rather than which index better predicts trade flows. However, for WARP, BSWARP, and WARULC, each series has strengths and weaknesses, while all three series happen to be quite similar for the US, and so testing between them would not be a useful exercise. Nevertheless, the class of WAR indices was created in order to be useful, and so in this section I briefly discuss the predictive value of these indices.

Fahle *et al.* (2008) have already shown that relative price indices based on WARP seem to do a better job predicting US trade imbalances than do divisia-based indices using CPIs. While it would be nice to extend this evidence using WARP for the period 1820 to 1970, most of the large movements in relative prices during this period were associated with major wars or the Great Depression, and so exchange rates were unlikely to be the major determinate of trade flows as they arguably were in the post-Bretton Woods period. Instead, in this section, I focus on the relative unit labor cost index, presenting evidence that WARULC does a better job predicting aggregate manufacturing trade for the US than does the IMF’s index, and also that it does a remarkably good job predicting the timing of the collapse of import-competing manufacturing sectors.

First, in Figure 14 I show that the level of WARULC seems to do a reasonably good job of predicting changes in the import share of manufacturing trade not due to changes in GDP.<sup>16</sup> The IMF’s RULC index, by contrast, implies a steadily more competitive US manufacturing sector over time, which seems to be at odds with the large import share of trade the US has experienced since the late 1990s, and at odds with the realized collapse in US manufacturing in the early 2000s concentrated in sectors more exposed to international trade (Campbell 2014) and Chinese import competition (Autor *et al.* 2013).

Second, I postulate a simple model based on intuition and the foundation of a simple

---

<sup>15</sup>An alternative method to measuring competitiveness using prices and wages would be a measure based on trade flows. This is a useful, but quite different, exercise.

<sup>16</sup>*I.e.*, they do a good job of explaining the residuals of regression:  $M_{US,t} = \alpha + \rho M_{US,t-1} + \beta_0 \ln \Delta RGDPO_t + \beta_1 \ln \Delta RGDPO_t^* + \beta_2 I_{t-1}^j + \epsilon_t$ , with the lagged value of WARULC times the coefficient on the lagged value of WARULC added back in.

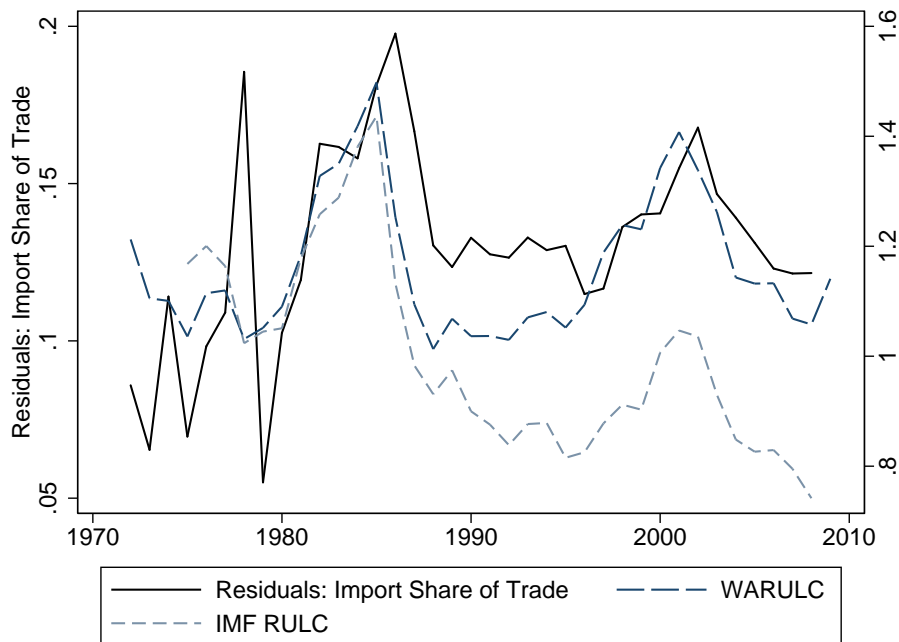


Figure 14: RULC Indices vs. Import Share of Manufacturing Trade

gravity model to predict the share of manufactured imports of total manufacturing trade:<sup>17</sup>

$$M_{US,t} = \alpha + \rho M_{US,t-1} + \beta_0 \ln \Delta RGDP O_t + \beta_1 \ln \Delta RGDP O_t^* + \beta_2 I_{t-1}^j + \epsilon_t, \quad (6.1)$$

j = WARULC, IMF RULC Index.<sup>18</sup>

This equation supposes that the import share of manufacturing trade ( $M_{US,t}$ ) grows when home GDP growth is higher ( $RGDP O_t$ ), falls when foreign GDP growth is faster ( $RGDP O_t^*$ ), and rises when the lagged level of the real exchange rate is higher ( $I_{t-1}^j$ ). The results are displayed in Table 1, which shows that each coefficient is of the theoretically predicted sign. I find that the R-squared is higher using WARULC than it is using the IMF's RULC index, and also that the coefficient on the RER is more highly significant. This simple estimation strategy based on intuitive priors is not without trouble, as in the

<sup>17</sup>The solution to a CES expenditure function (*e.g.*, from Feenstra (2004) used to derive a gravity equation:  $x_{ij} = \frac{Y_j p_i^{1-\sigma} \tau_{ij}^{1-\sigma}}{P_j^{1-\sigma}}$ , where  $x_{ij}$  are exports from  $i$  to  $j$ ,  $Y_j$  is GDP,  $\tau_{ij}$  are trade costs, and  $p_i$  is the price level in country  $i$ , while  $P_j$  is the CES price index for country  $j$ . This equation tells us that prices and GDP matter for trade flows. Thus, home and foreign GDP, and the relative price levels would matter for the import share of total trade.

<sup>18</sup>I do not have enough observations in this case to do an error correction model. Note that I start in 1975 because this is when the IMF's index begins.

regression using the IMF's index, the coefficients on GDP and the lagged import share of trade likely change in a way that would be unlikely to hold out of sample, in order to counteract the strong negative trend in the IMF's index. An additional problem with using a lagged dependent variable is Nickell (1981) bias, where the bias on the lagged dependent variable will be equal to  $(1+\rho)/(T-1)$ , which should be on the order of 5% with  $t=33$  in this case, which I argue is not material.

Table 1: Predicting the Import Share of Manufacturing Trade 1975-2008

	(1)	(2)
	Import Share of Trade	Import Share of Trade
L.Import Share of Trade	0.691*** (0.0814)	0.871*** (0.100)
ln $\Delta$ RGDP	0.372* (0.200)	0.480** (0.234)
ln $\Delta$ Foreign RGDP	-0.0113 (0.156)	-0.144 (0.194)
L.WARULC	0.116*** (0.0290)	
L.IMF RULC		0.0539* (0.0281)
Observations	33	33
r2	0.845	0.784

\* $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . The dependent variable is the manufacturing import share of manufacturing trade. Foreign GDP is the average of the G7 economies.

Campbell (2014) finds that the class of Weighted-Average Relative (WAR) price indices developed in this paper can predict declines in more open manufacturing sectors relative to less open manufacturing sectors, which, he argues, are less exposed to international trade. I confirm and strengthen this evidence, by showing (Figure 15) that an import-Weighted Average Relative Unit Labor Cost (iWARULC) index does a remarkably good job predicting years when manufacturing sectors with a higher share of import-penetration suffered declines in employment.<sup>19</sup> The coefficient on import penetration, plotted with two standard deviation error bounds, came from the regression:

<sup>19</sup>This index is constructed in the same manor as WARULC, only using only import weights instead of overall trade weights.



$$\ln(L_{ht}/L_{h,t-1}) = \alpha_t + \beta_0 MPPen_{h,t-1} + \beta_2 \ln(D_{h,t}/D_{h,t-1}) + \beta_3 \ln((TFP)_{h,t}/(TFP)_{h,t-1}) + \epsilon_{ht}, \quad (6.2)$$

where  $h$  is one of 353 balanced 4-digit SIC manufacturing industries (data from the BEA's Annual Survey of Manufactures), and the years span 1973 to 2009.  $MPPen$  is import penetration, defined as imports divided by "domestic consumption", which is equal to domestic production plus imports minus exports.  $D_{ht}$  is defined as demand in sector  $h$  at time  $t$ , and  $TFP$  is a measure of 4-factor productivity (the details of this measure, which come from the NBER-CES manufacturing database, are described in detail by Bartelsman and Gray (1996)). Once again, a RULC index derived using the IMF's data and methods, except using time-varying import weights, implies that US RULCs in manufacturing were no higher in the early 2000s than they had been in the late-1980s or late-1970s, periods which were not associated with adverse performance of import-competing manufacturing sectors. By contrast, the import-Weighted Average Relative Unit Labor Cost (iWARULC) implies that US RULCs were very high from the late-1990s to the mid-2000s, a period associated with a relatively worse performance of more open manufacturing sectors (this graph shows employment, but this is also true of value-added and labor productivity).

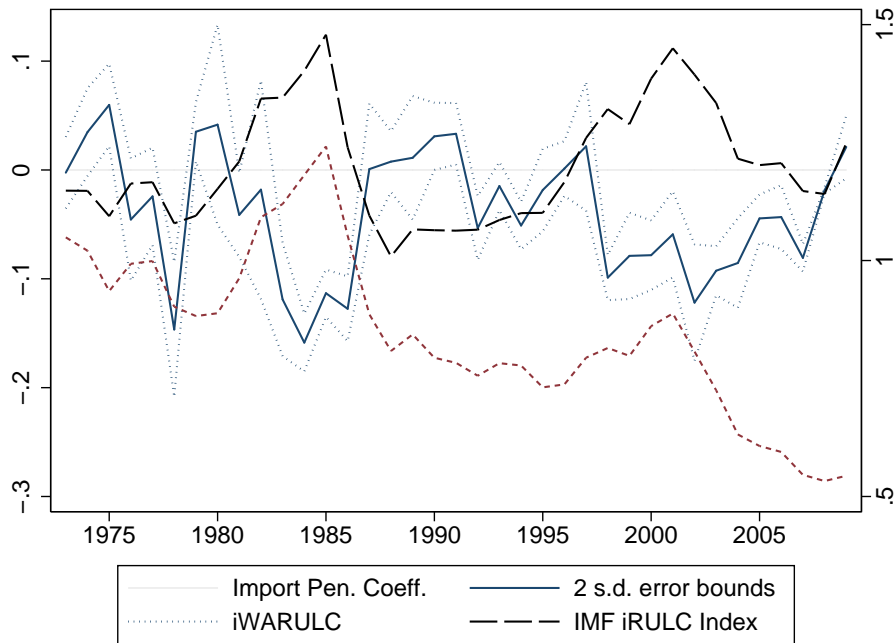


Figure 15: The Coefficient on Import Penetration vs. Two Measures of RULCs

Thus the evidence presented in this section is strongly suggestive that WARULC may be more relevant for many empirical applications than divisia-based RULC indices.

## 7 International Extensions

In this section, I create Weighted-Average Relative (WAR) price and unit labor cost indices for major Asian and Euro-Area nations, using trade weights computed from each country's 40 largest trading partners, with time-varying import shares as in equation 3.6. First, in Figure 16, I show the results for trade-weighted WARP for France, Germany, Greece, Italy, Spain and the Netherlands. This paints a picture of a relatively steady appreciation of the southern European countries vs. France and Germany since 1980, with Italy and Greece now having as strong an exchange rate as Germany. On a Balassa-Samuelson adjusted basis, since 1990, the picture looks broadly similar, except that once I control for level of development, Germany's real exchange rate is now weaker than that of Italy, Greece, and France, although the difference with Figure 16 is perhaps not as dramatic as may have been expected. Remarkably, the WARULC indices suggest more dramatic appreciations for Spain and Italy relative to Germany, and also suggest that the Euro area as a whole had high relative ULCs going into the Financial Crisis and later Euro debt crisis.

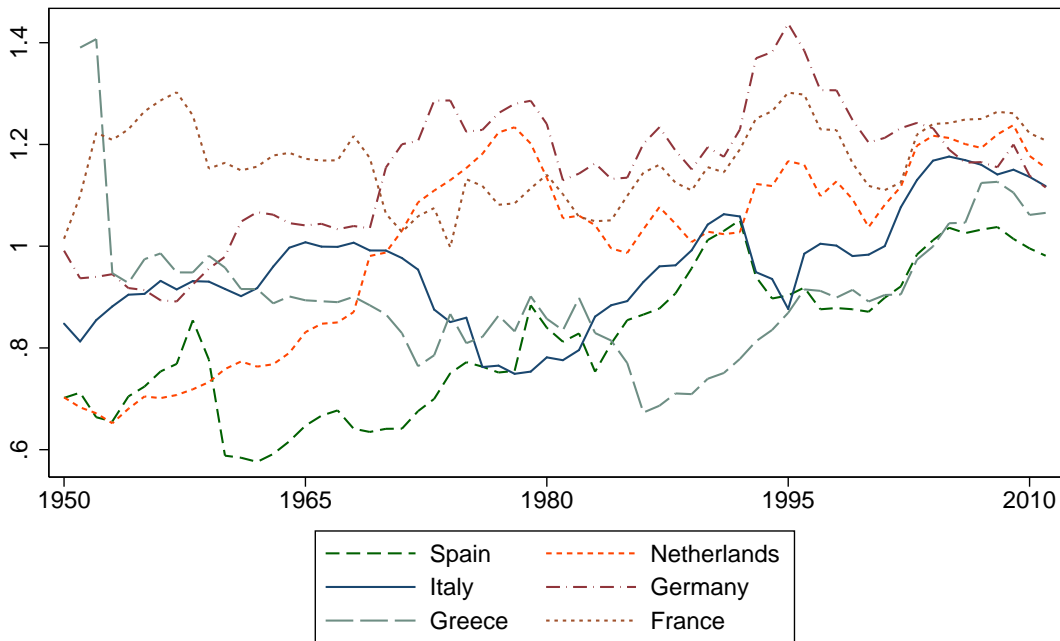


Figure 16: WARP, Major Euro Area Countries

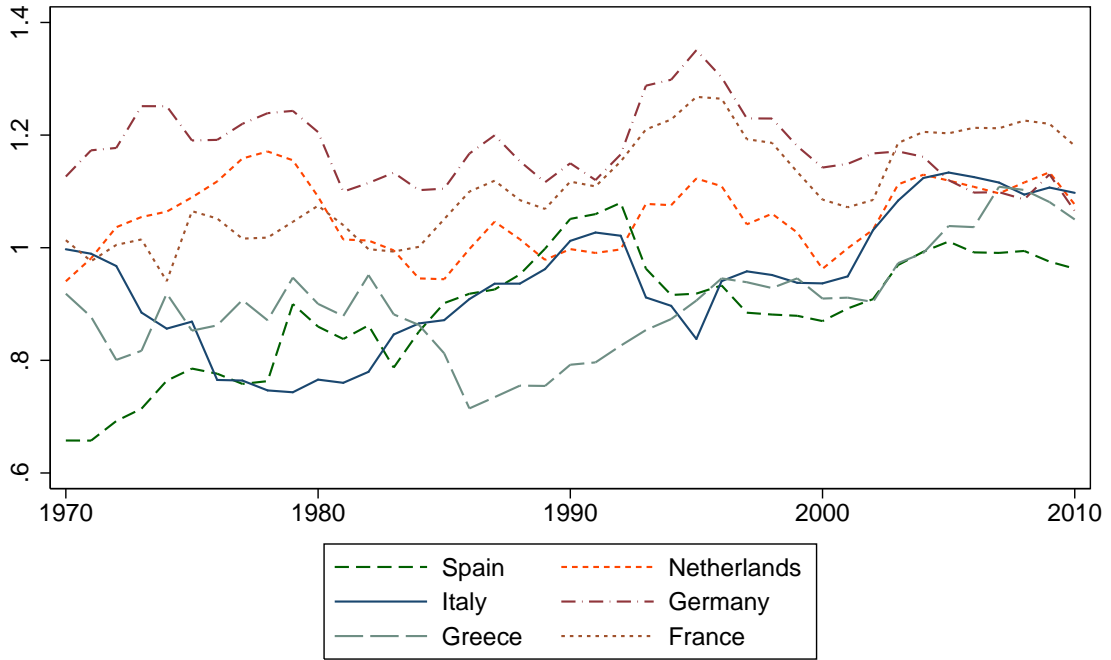


Figure 17: BS-WARP, Major Euro Area Countries

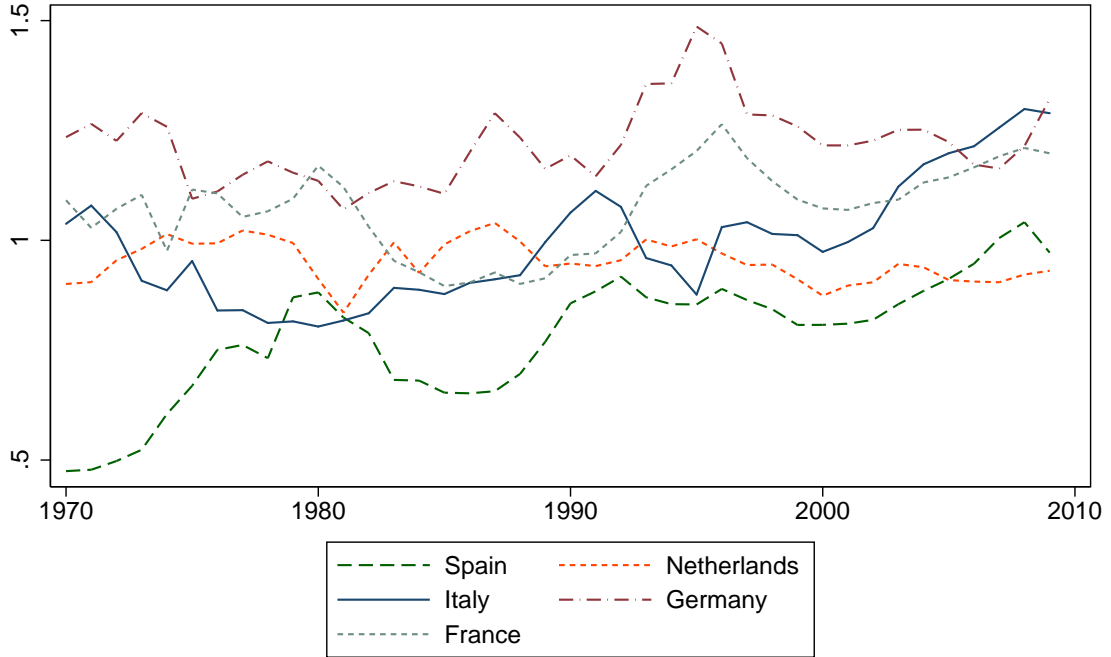


Figure 18: WARULC, Major Euro Area Countries

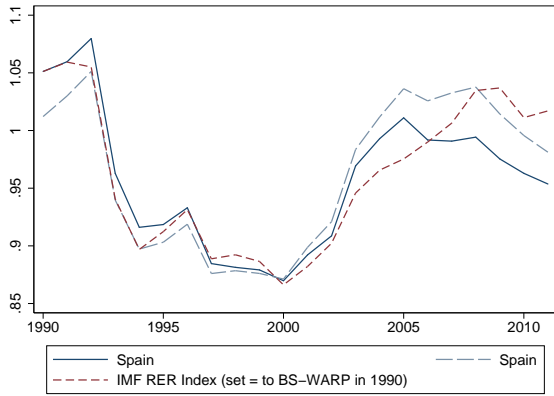
In Figure 19, I plot both the WARP and the Balassa-Samuelson adjusted series (which are similar to WARP for Europe) vs. the divisia-based indices produced by the IMF. The IMF and OECD indices are indexed to be equal to the BS-WARP series in 1990. For Italy, Greece, and France, the WARP and BS-WARP indices reveal substantially greater appreciations than the IMF's RER series.

In Figure 20, I plot indices for a handful of other European countries, several of which have plans to join the Euro and several which are still deciding. Interestingly, Ireland's WARP index reveals a slower appreciation than its IMF series, which is an artifact of large changes made to Ireland's prices in version 8.0 of the PWT. (Arguably, Ireland's BS-WARP index is unfairly penalized, since a relatively large portion of Irish GDP is income paid to foreign companies located in Ireland.) For Iceland and the UK, the WARP and BS-WARP indices reveal much sharper appreciations leading up to the financial crisis of 2008. The IMF's measure implies a substantially faster appreciation for Poland since 1990 than the WARP index does, since Poland has moved toward trading more with high-priced western European countries.

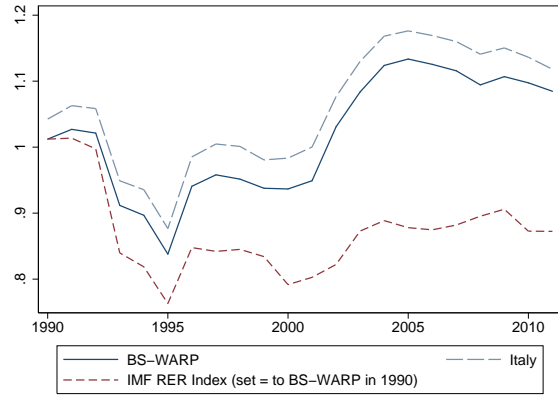
Large discrepancies between WAR indices vs. divisia-based series also exist for several large Asian countries and Russia. It should be noted that in Figure 21(c), the scale for Korea's real exchange rate varies much more dramatically than for the other economies. In the period leading up to the Asian Financial Crisis, 1990-1996, the WARP index for Korea appreciated roughly 8% more than the OECD's index (although, as in Chinn 2000, these indices still imply that Korea was undervalued before the crisis). As China has continued to rise as Korea's top trading partner, it is no surprise to find that Korea's WARP index deviates from its divisia-based counterparts more than any other country, as its WARP is now 86% higher than its OECD REER index relative to 1989. In recent years, Korea has come under criticism for amassing large amounts of foreign reserves despite a dramatically weaker trade-weighted real exchange, but Figure 21 indicates that this weakening of Korea's divisia real exchange rate was more than offset by changes in Korean trade patterns toward China and other economies with low price levels.

As of 2011, the price level in the People's Republic of China was just 35% less than prices in China's trading partners, and only 21% less on a Balassa-Samuelson adjusted basis. Quite surprisingly, China's BS-WARP has appreciated just as fast as its WARP index since the mid-1990s, and despite the fact that China's Balassa-Samuelson residual itself has been flat since around 2000. Hence, China's appreciation in this measure was driven instead by a shrinking in the Balassa-Samuelson residual of China's trading partners – most notably the United States.

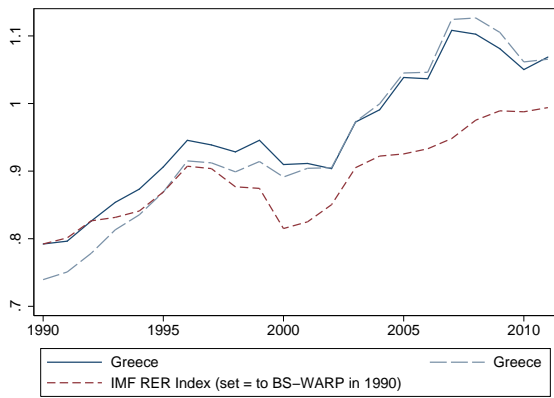
Japan's WARP more than doubled from the early 1980s into the mid-1990s, outpac-



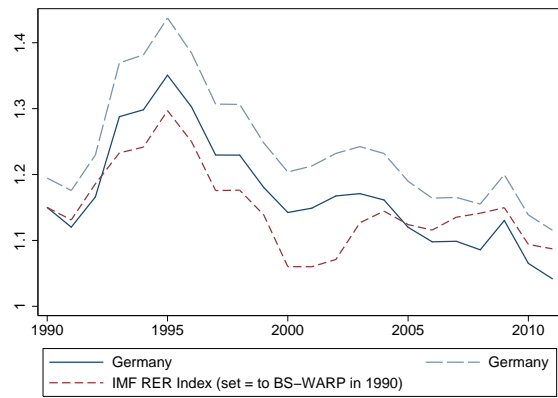
(a) Spain



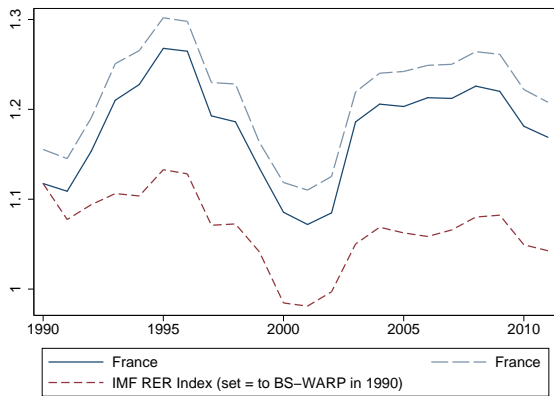
(b) Italy



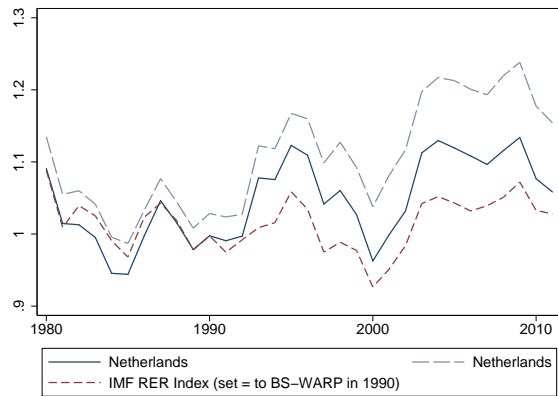
(c) Greece



(d) Germany

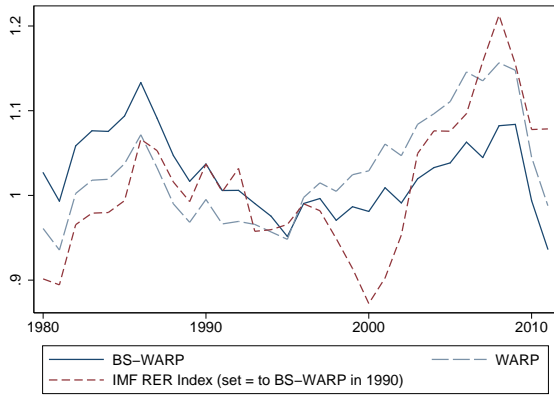


(e) France

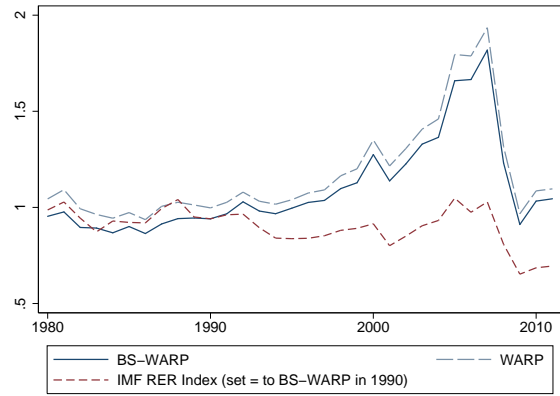


(f) Netherlands

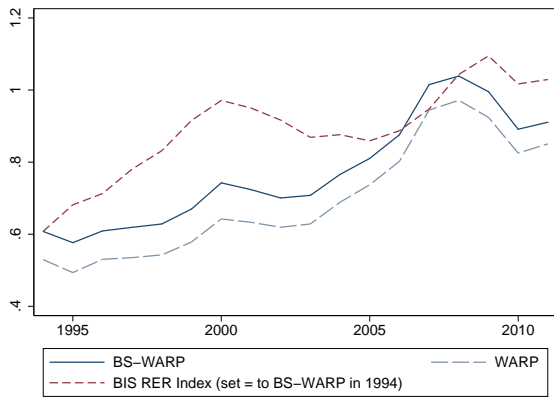
Figure 19: BS-WARP and WARP vs. IMF Divisia RER Index



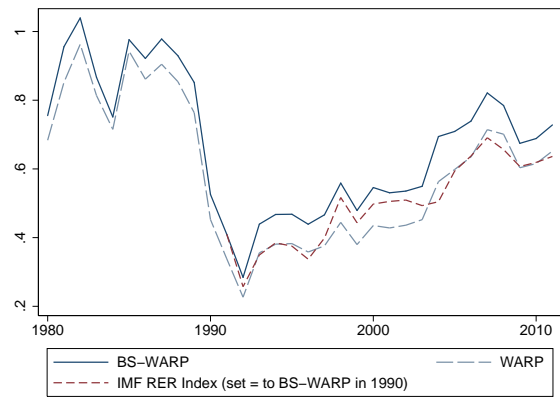
(a) Ireland



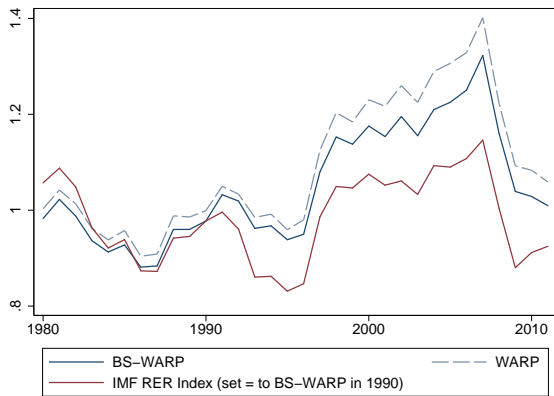
(b) Iceland



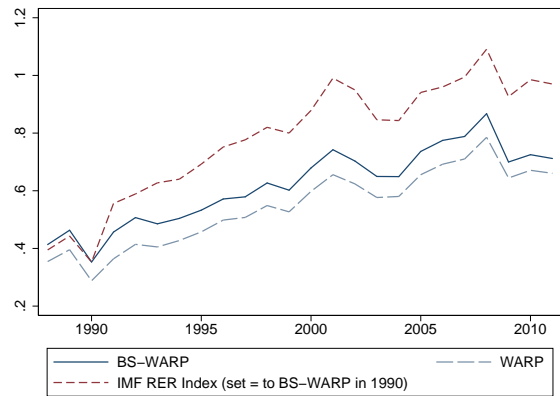
(c) Latvia



(d) Romania

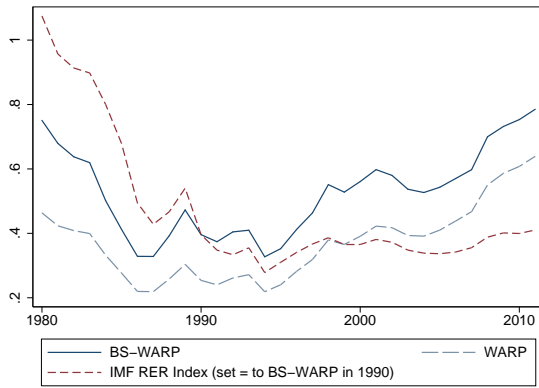


(e) UK

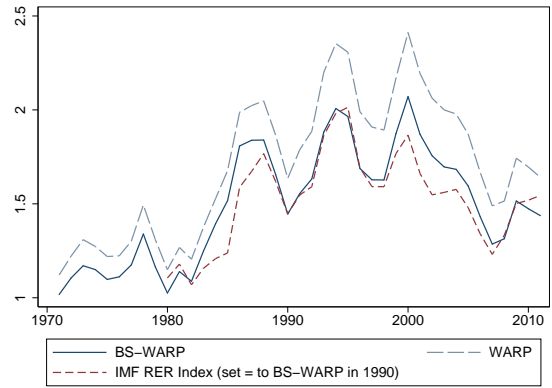


(f) Poland

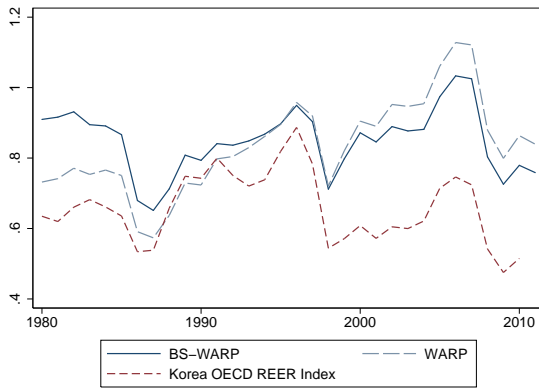
Figure 20: BS-WARP and WARP vs. IMF Divisia RER Index



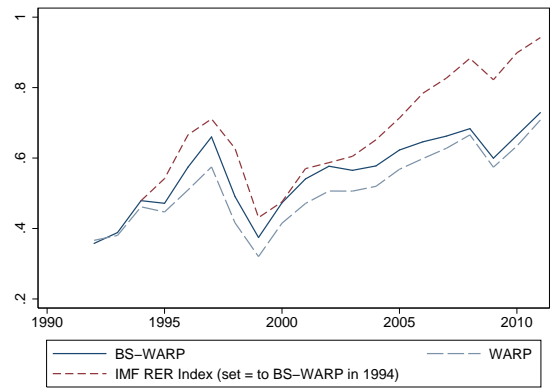
(a) People's Republic of China



(b) Japan



(c) Korea



(d) Russian Federation

Figure 21: BS-WARP and WARP vs. IMF Divisia RER Index

ing even the swift appreciation of the IMF's trade-weighted REER index. For much of the past two decades, just as Japan has been immersed in a deflationary liquidity trap, its price level has been on average more than twice than that of its trading partners. Japan's BS-WARP index has mimicked the IMF's REER series much more closely, and implies that as of 2010, Japan's Balassa-Samuelson residual was still about 50% higher than the residuals of her trading partners.

Finally, the BS-WARP and WARP indices show a much more gradual appreciation than does the IMF RER index for the Russian Federation.

## 8 Conclusion

In this paper, I sought to improve the methodology used to create trade-weighted real exchange rate indices, proposing several new series, including a Balassa-Samuelson adjusted Weighted Average Relative Price index, and a Weighted Average Relative Unit Labor Cost index, which are designed to solve the trading partner substitution bias problem and control for productivity as per the Penn Effect. These new indices provide a seriously warped view of economic history, particularly recently, for many countries. I demonstrate that, for many countries currently mired in liquidity traps, including Japan, the US, and for the Euro area, prices and wages had grown to high levels relative to trading partners heading into the financial crisis of 2008. For the US, I demonstrate that in 2002 – during the surprisingly swift collapse of US manufacturing – relative prices had not been that high since the worst year of the Great Depression, when the golden straitjacket led to an appreciated currency, collapsing commodity prices, rising real debt levels and a financial and economic crisis. Anyone wishing to understand the past 20 years of economic history would do well to start with an investigation of relative prices.

## References

- ALLEN, R. C., J.-P. BASSINO, D. MA, C. MOLL-MURATA AND J. L. VAN ZANDEN, “Wages, Prices, and Living Standards in China, 1738–1925: In Comparison with Europe, Japan, and India,” *The Economic History Review* 64 (2011), 8–38.
- AUTOR, D., D. DORN AND G. H. HANSON, “The China Syndrome: The Local Labor Market Effects of Import Competition in the US,” *American Economic Review* 103 (2013), 2121–68.
- BARBIERI, K., *The Liberal Illusion: Does Trade Promote Peace?* (University of Michigan Press, 2002).
- BAYOUMI, T., S. JAYANTHI AND J. LEE, “New Rates from New Weights,” (2006).
- BAYOUMI, T., M. SAITO AND J. TURUNEN, “Measuring Competitiveness: Trade in Goods or Tasks?,” (2013).
- BEMS, R. AND R. C. JOHNSON, “Value-Added Exchange Rates,” Working Paper 18498, National Bureau of Economic Research, October 2012.



- BERGIN, P. R., R. GLICK AND A. M. TAYLOR, “Productivity, Tradability, and the Long-Run Price Puzzle,” *Journal of Monetary Economics* 53 (2006), 2041–2066.
- CAMPBELL, D. L., “Relative Prices, Hysteresis, and the Decline of American Manufacturing,” (2014).
- CATAO, L. AND S. SOLOMOU, “Effective exchange rates 1879–1913,” *European Review of Economic History* 4 (2000), 361–382.
- CEGLOWSKI, J. AND S. GOLUB, “Just How Low are China’s Labour Costs?,” *The World Economy* 30 (2007), 597–617.
- CEGLOWSKI, J. AND S. S. GOLUB, “Does China Still Have a Labor Cost Advantage?,” *Global Economy Journal* 12 (2012).
- CHEUNG, Y.-W., M. D. CHINN AND E. FUJII, “China’s Current Account and Exchange Rate,” in *NBER Volume: China’s Growing Role in World Trade* (University of Chicago Press, 2010), 231–271.
- CHINN, M. D., “Before the Fall: Were East Asian Currencies Overvalued?,” *Emerging Markets Review* 1 (2000), 101–126.
- , “Incomes, Exchange Rates and the US Trade Deficit, Once Again\*,” *International Finance* 7 (2004), 451–469.
- , “A Primer on Real Effective Exchange Rates: Determinants, Overvaluation, Trade Flows and Competitive Devaluation,” *Open Economies Review* 17 (2006), 115–143.
- DESRUELLE, D. AND A. ZANELLO, “A Primer on the IMF’s Information Notice System,” *IMF Working Paper No. 97/71* (1997).
- DEVEREUX, J. AND M. WARD, “Measuring British Decline: Direct Versus Long-span Income Measures,” *The Journal of Economic History* 63 (2003), 826–851.
- , “Relative British and American Income Levels During the First Industrial Revolution,” *Research in Economic History* 23 (2006), 249–286.
- DIEWERT, W. E., A. O. NAKAMURA, J. S. GREENLEES, L. I. NAKAMURA AND M. B. REINSDORF, “Sourcing Substitution and Related Price Index Biases,” (2014).

- EATON, J. AND S. KORTUM, “Technology, Geography, and Trade,” *Econometrica* 70 (2002), 1741–1779.
- FAHLE, S., J. MARQUEZ AND C. THOMAS, “Measuring US International Relative Prices: A WARP View of the World,” *FRB International Finance Discussion Paper* (2008).
- , “Measures of International Relative Prices for China and the USA,” *Pacific Economic Review* 14 (2009), 376–397.
- FEENSTRA, R. C., R. INKLAAR AND M. TIMMER, “The Next Generation of the Penn World Table,” (2011).
- FEENSTRA, R. C., R. E. LIPSEY, H. DENG, A. C. MA AND H. MO, “World Trade Flows: 1962-2000,” Technical Report, National Bureau of Economic Research, 2005.
- FEENSTRA, R. C., H. MA, J. PETER NEARY AND D. PRASADA RAO, “Who Shrank China? Puzzles in the Measurement of Real GDP,” *The Economic Journal* (2012).
- FRATZSCHER, M., A. MEHL AND I. VANSTEENKISTE, “130 Years of Fiscal Deficits and Currency Crashes in Advanced Economies,” (2011).
- GOLUB, S. AND J. CEGLOWSKI, “Does China Still Have a Labor Cost Advantage?,” (2011).
- GOLUB, S. S. AND A. G. TURNER, “Towards a System of Multilateral Unit Labor Cost-Based Competitiveness Indicators for Advanced, Developing, and Transition Countries,” Technical Report, International Monetary Fund, 1997.
- HO, L. S., “Globalization, Exports, and Effective Exchange Rate Indices,” *Journal of International Money and Finance* 31 (2012), 996–1007.
- HOUSEMAN, S., C. KURZ, P. LENGERMANN AND B. MANDEL, “Offshoring Bias in U.S. Manufacturing,” *Journal of Economic Perspectives* 25 (Spring 2011), 111–32.
- INKLAAR, R., “Offshoring Bias in Manufacturing Growth Across Advanced and Emerging Economies,” (2012).
- IRWIN, D. A., “Trade Restrictiveness and Deadweight Loses from US Tariffs, 1859-1961,” Technical Report, National Bureau of Economic Research, 2007.

- JACKS, D., C. MEISSNER AND D. NOVY, “Trade Booms, Trade Busts, and Trade Costs,” *Journal of International Economics* (2006).
- JOHNSON, S., J. D. OSTRY AND A. SUBRAMANIAN, “Prospects for Sustained Growth in Africa: Benchmarking the Constraints.,” *IMF Staff Papers* 57 (2010).
- LORETAN, M., “Indexes of the Foreign Exchange Value of the Dollar,” *Fed. Res. Bull.* 91 (2005), 1.
- MELITZ, M. J., “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity,” *Econometrica* 71 (November 2003), 1695–1725.
- NICKELL, S., “Biases in dynamic models with fixed effects,” *Econometrica: Journal of the Econometric Society* (1981), 1417–1426.
- PIERCE, J. R. AND P. K. SCHOTT, “The Surprisingly Swift Decline of U.S. Manufacturing Employment,” Working Paper 18655, National Bureau of Economic Research, December 2012.
- REINSDORF, M. AND R. YUSKAVAGE, “Offshoring, Sourcing Substitution Bias and the Measurement of US Import Prices, GDP and Productivity,” (2014).
- REINSDORF, M. B., “Formula Bias and Within-Stratum Substitution Bias in the US CPI,” *Review of Economics and Statistics* 80 (1998), 175–187.
- RODRIK, D., “The Real Exchange Rate and Economic Growth,” *Brookings Papers on Economic Activity* 2008 (2008), 365–412.
- SATO, K., J. SHIMIZU, N. SHRESTHA AND S. ZHANG, “Industry-Specific Real Effective Exchange Rates for Japan,” Technical Report, RIETI Discussion Paper, 12-E-044, 2012.
- SCHMITZ, M., M. DE CLERCQ, M. FIDORA, B. LAURO AND C. PINHEIRO, “Revisiting the Effective Exchange Rates of the Euro,” *ECB Occasional Paper* (2011).
- SKIDELSKY, R., *John Maynard Keynes: Volume 2: The Economist as Savior, 1920-1937*, volume 2 (Penguin Books, 1995).
- TAYLOR, A. M., “A Century of Purchasing-Power Parity,” *Review of Economics and Statistics* 84 (2002), 139–150.

TURNER, P. AND J. VAN'T DACK, "Measuring International Price and Cost Competitiveness," *Bank for International Settlements, Monetary and Economic Department* (1993).

## 9 Appendix

**Table A.1: Historical Relative Price Levels (vs. US)**

Year	<u>UK</u>		<u>Japan</u>	
	Ward and Devereux	Our Estimates	Allen <i>et al.</i>	Our Estimates
1831	156	143.9		
1839	122	134.6		
1849	123	133.6		
1859	112	124.7		
1869	110	105.6		
1872	108	105.8		
1874	113	112.5		
1878	112	110.3		
1884*	94	106.6	200.5	194.2
1891*	99	113.0	147.3	147.3
1905	83	121.2	121.5	112.7
1930	83	95.5		
1950	71	56.4		
1955	82	64.5		
1967	83	75.9		
1970	72	67.2		
1973	84	76.4		

\*For Japan, the comparisons are for 1886 and 1892.

**Table A.2: Historical US vs. Federal Reserve Samples**

---

Country	Historical Sample Start	Post-war Sample Start
1. China	1820	1950
2. Netherlands	1820	1950
3. Norway	1820	1950
4. Spain	1820	1950
5. Sweden	1820	1950
6. UK	1820	1950
7. Belgium	1830	1950
8. Brazil	1830	1950
9. France	1840	1950
10. Germany	1852	1950
11. Australia	1861	1950
12. Italy	1861	1950
13. Denmark	1864	
14. Japan	1868	1950
15. India	1873	1950
16. Chile	1880	1950
17. Austria	1884	1950
18. Switzerland	1885	1950
19. Uruguay	1885	
20. Egypt	1895	
21. South Africa	1895	
22. Peru	1900	
23. Venezuela	1900	1950
24. Philippines	1902	1950
25. Canada	1910	1950
26. New Zealand	1914	
27. Colombia	1919	1950
28. Finland	1920	1950
29. Greece	1922	2001
30. Turkey	1922	
31. Hong Kong		1950
32. South Korea		1950
33. Singapore		1950
34. Russia		1993
35. Taiwan		1950
36. Malaysia		1950
37. Mexico		1950
38. Euro Area		1999
39. Israel		1950
40. Indonesia		1950
41. Saudi Arabia		1950

---

**Table A.3**  
**Data Sources**

Data	Source	URL
Bilateral Trade Data (total flows)	IMF Dots	<a href="http://elibrary-data.imf.org/">http://elibrary-data.imf.org/</a>
Bilateral Trade in Oil	UN Comtrade	<a href="http://comtrade.un.org/">http://comtrade.un.org/</a>
Federal Reserve Trade Weights	Federal Reserve Board of Governors	<a href="http://www.federalreserve.gov/releases/h10/weights/">http://www.federalreserve.gov/releases/h10/weights/</a>
Fed's Broad Trade-Weighted Real Exchange Rate, Korean Inflation	Federal Reserve Economic Data	<a href="http://research.stlouisfed.org/fred2/">http://research.stlouisfed.org/fred2/</a>
PPP, Real GDP, Exchange Rates	Penn World Tables, v7.1 and v8.0	<a href="http://www.rug.nl/research/ggdc/data/penn-world-table">http://www.rug.nl/research/ggdc/data/penn-world-table</a>
Historical Exchange Rates	Bordo	<a href="https://sites.google.com/site/michaelbordo/home4">https://sites.google.com/site/michaelbordo/home4</a>
Historical Inflation, Trade Data	Historical Statistics of the United States	<a href="http://hsus.cambridge.org/">http://hsus.cambridge.org/</a>
Historical Trade Data	Barbieri (2002)	<a href="http://people.cas.sc.edu/barbierk/databases.html">http://people.cas.sc.edu/barbierk/databases.html</a>
Historical Trade and Tariff Data	Jacks, Meissner, Novy (2009)	N/A
Average MFN Applied Tariff Rates in Developing and Industrial Countries	World Bank	<a href="http://econ.worldbank.org">http://econ.worldbank.org</a>
Historical Tariff Data	Irwin (2007)	<a href="http://www.nber.org/papers/w13450">http://www.nber.org/papers/w13450</a>
Post-war Taiwan Inflation	Taiwanese Government	<a href="http://eng.stat.gov.tw/mp.asp?mp=5">http://eng.stat.gov.tw/mp.asp?mp=5</a>
Post-War Tariff Data	UN Comtrade	<a href="http://comtrade.un.org/">http://comtrade.un.org/</a>
Manufacturing Data (Employment, etc.)	BEA, Annual Survey of Manufactures	<a href="http://www.census.gov/manufacturing/asm">www.census.gov/manufacturing/asm</a>
Historical Exchange Rates	Historical Volumes of Mitchell	N/A
Historical Exchange Rates	<i>Wahrungen der Welt</i> , Vol. 1-10	N/A
Historical Price Indices	Lindert & Arroyo Abad, Global Price and Income History Group	<a href="http://gpih.ucdavis.edu">gpih.ucdavis.edu</a>
British Prices From 1750	Office for National Statistics (ONS)	<a href="http://www.ons.gov.uk">http://www.ons.gov.uk</a>
Chinese Historical Prices	Allen et al. (2011)	via <a href="http://gpih.ucdavis.edu">gpih.ucdavis.edu</a>
IMF Trade Weights	Bayoumi <i>et al.</i> (2006)	<a href="http://www.imf.org/external/pubs/ft/wp/wp9771.pdf">http://www.imf.org/external/pubs/ft/wp/wp9771.pdf</a>
Tariff Data, 1984-1995	UNCTAD -- TRAINS	<a href="http://www.unctad.info/en">www.unctad.info/en</a>
Trade Data by SITC4, 1962-2008	Feenstra, Lipsey et al. 2005	<a href="http://cid.econ.ucdavis.edu">cid.econ.ucdavis.edu</a>
Manufacturing Labor Costs, Productivity	UNIDOs INDSTAT	<a href="http://data.un.org/">http://data.un.org/</a>
Manufacturing Labor Costs, Productivity	BLS	<a href="http://www.bls.gov/fls/">http://www.bls.gov/fls/</a>
Relative Price Levels in Manufacturing	World Bank, International Comparison Project 2005	<a href="http://databank.worldbank.org/data/databases.aspx">http://databank.worldbank.org/data/databases.aspx</a>
Relative Price Levels in Manufacturing	Groningen Growth and Development Centre Productivity Level Database	<a href="http://www.rug.nl/research/ggdc/data/">http://www.rug.nl/research/ggdc/data/</a>
Manufacturing Labor Costs, Productivity, Trade	OECD	<a href="http://stats.oecd.org/">http://stats.oecd.org/</a>
Canada, Manufacturing Deflator	Statistics Canada	<a href="http://www.statcan.gc.ca">http://www.statcan.gc.ca</a>