Improving the Big Mac Index

F.E. van der Schans
302933

Bachelor thesis
Economics

Under supervision of professor J.M.A. Viaene

Table of Contents

1. Introduction

In macroeconomics and international economics the theory of Purchasing Power Parity (PPP) plays a major role. It states that the prices of a certain basket of goods and services in two countries should be the same, once corrected for the foreign exchange rate. One of the fundaments of this theory is the *Law of One Price*; this law says that two identical goods should sell for the same price everywhere. This statement sounds rather straightforward but as I shall clarify further on in my paper, due to numerous factors, this is not the case. Normally PPP is put into practice for empirical testing with the comparison of a certain basket of goods and services that represents the consumer purchasing behaviour in two countries. A typical indicator used to determine the general price level via a basket of goods is the Consumer Price Index (CPI).

However, to determine consumer spending behaviour and the (development of) prices of thousands of items in different countries takes a considerable amount of time and thus is very expensive. As a much cheaper alternative for comparison, *The Economist* introduced its famous Big Mac Index (BMI) in 1986 and has been publishing it since. They make use of the differences in the prices of McDonald's Big Mac burger around the globe as a much simpler substitute for the complicated basket of goods, to analyse the exchange rates and their movements. It is very interesting (and perhaps even illuminating for those unfamiliar with the theory) to look closer at this relatively simple index as it captures almost all of the principles and limitations of the theory of PPP. The BMI assumes fully flexible prices, so that all economic changes in a certain country have been absorbed immediately into the price of the burger. It therefore assumes that if inflation is really high / low in any particular country, it will cause its Big Mac prices to rise a lot/little. But if we take a closer look at the literature on the matter of price stickiness, we find that prices do not always adjust instantaneously and thus prices are not fully flexible. This could mean, that the BMI might actually use burger prices that need to be augmented by the inflation. The exchange rate implied by the PPP or the BMI is thus not correct as a tool for comparison with the actual exchange rate, because prices do not have time to adjust.

I believe the BMI can be enhanced by correcting for inflation in the preceding period, the length of which depends on the price rigidity. In this way, a country's inflation is incorporated in the index, thus potentially making the index a more accurate tool in the sense that it will approach the actual exchange rate more closely (assuming PPP is a valid theory). And maybe it proves to be a handy tool for exchange rate predictions.

First of all, I shall analyse the relevant literature covering the main theories and concepts that I will use in my research subject. Furthermore I will explain how these theories and concepts are linked to the Big Mac Index and why they are so vital. During this literature review I shall simultaneously discuss the theoretical framework used to conduct my research. In short, it will contain a discussion of Purchasing Power Parity, the Law of One Price, price rigidity and how this all relates to the Big Mac Index and especially to the adjusted version. I shall also dedicate some words as to why the BMI is just as flawed as comparable indices based on more sophisticated baskets of goods and services. Thereafter the statistical research part will commence and I will discuss the data and methodology used to create the adjusted BMI and how the variables were computed to measure its ability to predict. I shall also elaborate on why certain techniques were chosen and others rejected and how the results compare with the original BMI. Then I will present the results obtained and discuss their relevance extensively. Finally, I shall conclude my thesis by discussing the overall relevance and validity of the outcomes.

2. Relevant Literature and Theoretical Framework

2.1. Purchasing Power Parity and the Law of One Price

Various versions of the PPP theory have been proposed since it first appeared in documents of the Spanish Salamanca School in the 16th century. Thereafter it has been mentioned in many different forms, but the main shaping came by the hands and minds of famous economists such as Marshall, Mill, Ricardo and Goschen throughout the 18th century. However, Cassel (1918) was the one to formulize the theory, based on empirical evidence. Several years later (Cassel 1928) he acknowledged that some “disturbances” might be present, which interfered with PPP. Such as expected inflation, international shifts in movement of capital and all variations of obstacles to international trade. While acknowledging these disturbances, he didn’t think of them to be significant enough because of their small impact. Later on we shall see that at least some of these forces at work gravely distort PPP and are not so insignificant after all. But first I shall discuss PPP as it evolved during the decades after World War I and after Cassel. Afterwards I will explain more in depth on how to mathematically define the theory.

In its essence, PPP is a theory that describes foreign exchange rates or in which direction they will move. The theory states that the difference in purchasing power of two national currencies determines the exchange rate between those countries. Moreover, any differences between them opens the window for arbitrage when the exchange rate between the countries does not neutralize the nominal difference. Even though it is intuitively straightforward and one of the major theories in macroeconomics, it remains a controversial one. The major concern about strict versions of PPP is that they do not hold empirically, while soft versions of the theory fail to be useful.

I will now look into versions of the theory, as Dornbusch (1985) and Pakko and Pollard (2003) had in mind. Strict versions of the PPP theory are based primarily on the Law of One Price. This states that - ignoring disturbing factors such as transportation costs, taxes and tariffs - the price of a particular good must be the same in all countries when converted to a common currency.

Let p_i be the price of good i at home and p^*_i the price of good i abroad, both depicted in their respective national currencies. In the case of the Big Mac these goods would be the separate ingredients to create the burgers such as lettuce, buns, beef, cheese etc. Furthermore, let e be the bilateral exchange rate quoted as the amount of domestic currency per unit of foreign currency. Also, let P and P^* be the domestic and foreign general price levels quoted in their national currency. Usually these general price levels are calculated using complicated baskets consisting of various goods, which should accurately reflect the overall consumer behavior in a country, most commonly the Consumer Price Index. However, it consumes a considerable amount of time to put together a basket like such, making it outdated to work with, when it is finally completed. Therefore, The Economist created the Big Mac Index to create a simple substitute for this basket. The reason why their index can be so comprehensible and yet still effective is that the Big Mac is composed out of exactly the same ingredients in almost all countries. It therefor serves as a good mean of comparison between them.

This being said, we can now formulate the equation for the Law of One Price. Given that we abstract from the disturbing factors mentioned above, assume a competitive market and quote prices in their respective domestic currency:

$$p_i = e \times p_i^* \tag{2.1}$$

which says that the price of good *i* at home, should cost exactly as much as abroad when we convert it with the foreign exchange rate.

Now, lets consider the domestic general price level:

$$P = f(p_1, p_2, p_i, \dots p_n) \tag{2.2}$$

and the foreign general price level:

$$P^* = g(p_1^*, p_2^*, p_i^*, \dots p_n^*) \tag{2.3}$$

these are weighted averages of a particular set of goods and services, most commonly consumed in a single country. Earlier referred to as the basket of goods. The change in percentage of which over a certain timespan is commonly known as inflation. Inflation, being the outcome of a rise in prices, brings about a loss of purchasing power, other things being equal. Hence, the name Purchasing Power Parity. In an efficient market a deviation from this parity cannot be a stable equilibrium due to arbitrage opportunities.

Now, if the Law of One Price holds for all the goods and services in this basket and thus have the same price among countries (once depicted in a common currency); and if there is an equal share or weight of all the goods in the product baskets then the absolute version of PPP will hold. If we again consider the Big Mac we can see that this holds perfectly as every Big Mac has exactly the same ingredients. In this specific case the Law of One Price is transferred from the individual goods to the aggregate price levels. The absolute version of PPP then takes this form.

$$e = \frac{P}{P^*} = \frac{\text{€ } price\ of\ standard\ basket}{\$\ price\ of\ standard\ basket} \quad \text{or} \quad e = \frac{\text{€ } price\ of\ Big\ Mac\ in\ Europe}{\$\ price\ of\ Big\ Mac\ in\ the\ USA} \qquad (2.4)$$

The implication of this absolute version of PPP is that whatever real or monetary disturbances in the economy, the prices of the standard basket in two countries, when quoted in the same exchange rate will be equal $\frac{P}{eP^*} = 1$ at all times. This because it assumes that a situation of disequilibrium is immediately offset by costless arbitrage. There also exists a weaker version of the theory; relative Purchasing Power Parity, which will not be discussed here, as it plays no role in my research.

Theoretically, the formula above is sound and straightforward. But like many theories, the empirical validity is questionable at least. There are numerous factors disturbing the immediate and costless flows of international trade to ensure arbitrage. Also other impediments exist that distort the empirical validation of the absolute version of PPP theory. Imagine the real life possibility to benefit from arbitrage opportunities, by buying a Big Mac in the one country and selling it with a profit in the next. These disturbances basically can be split in the following categories; trade barriers, non-traded goods and market pricing. I will discuss them extensively below.

2.2. Why does absolute PPP not hold?

Looking at the dollar prices of Big Macs around the world, you will find major price differences. However, according to absolute PPP these differences should have been evened out immediately because an arbitrage opportunity arose. As mentioned above, there are huge practical restrictions to trading in physical hamburgers. But the ingredients of the Big Mac are subject to everyday trading around the globe. So apparently other impediments exist, which prevent the smoothing of burger prices.

2.2.1. Trade barriers

The first obvious reason for this deviation is that there are different barriers to trade. For instance, transporting goods isn't costless, as absolute PPP –the Law of One Price – requires. For the sesame seeds on top, for instance, this might not be significant compared to the actual commodity price. But for more perishable goods, such as lettuce and other vegetables, transport costs are far greater because of cooling systems to prevent rotting on the way. Thus a gap is created between those ingredients' prices in different countries. Hummels (2001) estimated an added 16% on prices of vegetables in the U.S.A. because of transport; this is a significant influence on the price of the burger. Cassel (1921) himself, just three years after formulating the theory, also argued that trade restrictions, such as quotas and tariffs, could interfere with PPP. He said that in a two-country model, one of which has relatively more import restrictions, that country would have an overvalued currency according to PPP. So if we assume no other disturbing factors on PPP, one could argue that a country with an currency that is overvalued according to the BMI, must have relatively high import restrictions on e.g. the agricultural sector and vice versa. In addition to the two factors already described (i.e. cost of transport, legal restrictions), consumer taxes also play a role in explaining deviations from PPP. Consumer taxes can differ severely among countries. As these taxes are included in the prices of e.g. Big Macs, they will influence the outcome of the foreign exchange rate e in the PPP theory and thus disturb the outcome.

2.2.2. Non-traded goods

Besides the trade barriers discussed above, there is the disturbance of non-traded goods. The name itself already violates one of the key assumptions for PPP to work properly, or at least to enable arbitrage and restore PPP when in disequilibrium. In case of the Big Mac many factors determine the price at the counter besides the costs of the ingredients for the actual burger. Some come easily to mind, like the rent or mortgage costs of the building, electricity, gas, etc. Due to their nature, these are so called non-traded goods. Salaries for employees and franchise costs are in economic terms also included in this category of non-traded goods. So if rent and utilities, i.e. all non-traded goods, make up for a significant part of producing a Big Mac, this will be reflected in the price. This in turn would be a very good explanation as to why these major deviations from PPP pertain. Ong (1997) even estimated the price of the Big Mac to consist of non-traded goods for a staggering 94%.

The notion, that non-traded goods are responsible for systematic deviations from PPP, was formalized independently by Balassa (1964) and Samuelson (1964). Both argued that relatively high-income countries would have overvalued currencies to relatively low-income countries as a result of the fact that non-tradable goods are included in price indices (like the BMI or general price levels). Their theory is supported on the base of the empirically tested idea, that differences in per capita income between countries are strongly correlated with differences in labor productivity in those countries. This means that a low-income country's labor force is less productive than that of a high-income country. These differences are most notable in the traded goods sector. The higher productivity of high-income countries in this sector is reflected in higher wages. As companies in both the traded- and non-traded goods sector compete over the labor force, the wages in the non-traded goods sector have to rise as well. So the relatively high wages in the non-traded goods industry, e.g. the service sector, of high-income countries, compared to low-income countries, ensure higher price levels even when the prices of tradable goods are similar globally. Due to which the currencies in high-income countries should be overvalued relative to the currencies of low-income countries. If we apply this knowledge to the Big Mac Index, note that it would be hard to find any differences in productivity between high- and low-income countries when it comes to preparing burgers. However, the wages between those countries that serve Big Macs do differ greatly. Ashenfelter and Jurajda (2001) showed that the average McDonald's worker in the year 2000 earned $0,42 per hour in China, in Poland this was $1,14 and in the U.S.A. they earned $6,50 (all converted at historic exchange rates). These huge gaps in wages can then explain why the yuan and zloty have been structurally undervalued against the dollar, according to the BMI.

While this theory can explain the income differentials between high- and low-income countries, it still fails to explain why large PPP deviations exist between countries with similar per capita income levels. Froot and Rogoff (1995) argue that a difference in government spending also contributes to deviations from Purchasing Power Parity. This follows the same line of reasoning as Ashenfelter and Jurajda (2001), because governments usually spend a bigger portion of their money on non-traded goods than the private sector does. The price of non-traded goods will thus rise in a country where the government increases her spending and as such the overall price level within that country. If beforehand PPP holds, an increase in government spending by, for instance, the Chinese authorities will cause the Yuan to be overvalued afterwards compared to its value under PPP.

Krugman (1990) argues there is another way through which a macroeconomic phenomenon distorts PPP. He states that a country that runs a deficit on their current account spends, relative to other countries, more on traded goods, thus lowering the relative price of non-traded goods in that country. This would result in a lower Big Mac price, i.e. general price level and thus an undervalued currency.

2.2.3. Market pricing

After having discussed the first two causes for deviation from PPP, trade barriers and non-traded goods, it is now time to discuss the last cause; market pricing. A key assumption for Purchasing Power Parity is perfect competitiveness of the markets. If this is not the case and firms have market power to some extent on certain markets, this assumption of perfect competition is violated. According to standard economic theory, firms with market power will adjust their prices in such a way that they make full use of the market elasticity to maximize their profits. Krugman (1987) found empirical evidence to support this theory. He concluded this when he saw that firms in different countries charged different prices for traded goods. Dornbusch (1987) found the same price differential. Up to which what level companies can price to the market, and so disturb PPP, is influenced mainly by the ease of which the good can be resold in countries abroad. Again, in case of the Big Mac, reselling in another country surely cannot be done. All the ingredients involved in preparing a Big Mac, however, are resold more easily and one could buy all the ingredients necessary at the lowest price available and so, at least in theory, create a burger that competes with McDonald's burger.

Even though various empirical studies showed why absolute PPP does not hold in practice (and therefore the BMI might be flawed in some ways), it is still valuable as it provides us with a certain guideline by which exchange rates should, in theory, behave themselves in the long term.

2.3. Price Rigidity

One of the assumptions of the PPP theory is that inflationary forces are immediately absorbed in the prices of goods and services. Therefore the differences in inflationary forces between countries are the reason for the exchange rate to vary. In The Economist this is also assumed, as the BMI uses the prices quoted at McDonald's restaurants at the very moment the index is composed. However, much literature on this subject has been written that states that some form of price rigidity exists (for example Bils and Klenow (2002), Taylor (1999) and Baharad & Eden (2004)). Consequently, the price of a Big Mac presented in a McDonald's restaurant has not taken the previous period's inflation of the ingredients into its price, due to this rigidity.

The theory of sticky prices is very important, in macroeconomics above all. As it helps to explain the phenomenon that markets do not reach their equilibrium levels instantaneously and thus are not fully efficient. Some reasons as to why prices do not react immediately are the following:

- Menu costs as described by Sheshinski and Weiss (1977). According to this theory, some costs are associated with changing one's prices. The theory is called after the cost of printing an actual menu incurred by restaurants when they change their prices. In general the broader concept comprises of updating databases and computer systems etc.;
- Irving Fischer's money illusion, as proposed in his book *The Money Illusion* (1928). John Maynard Keynes originally invented this concept, but the theory is most commonly known through the book of Fischer. It states that economic agents have the tendency to think of money in nominal- rather than real terms. This means people confound the nominal value of money for its purchasing power. Empirical evidence for the effect was found by behavioral economist Amos Tversky et al. (1997). They also provided psychological arguments as to why the effect arises.
- Imperfect information. This is not so much a theory, as it is a disputation of one of the most frequently made assumptions in economic science. Perfect or complete information is a popular assumption, commonly in combination with rationality, because it allows for mathematical

derivation of formulas etc. Though widely spread through economics, it is an assumption very unlikely to hold empirically, as Nobel laureate Joseph E. Stiglitz (Stiglitz, 2002) showed.

After establishing the notion of price stickiness, the question arises how we can use this to refine the Big Mac Index. If we adjust Big Mac prices for recent inflation, we might be able to come to a more accurate approximation of PPP. But how much time passes until Big Mac prices have absorbed recent inflation? In other words: how sticky are Big Mac prices?

Blinder et al. (1998) did a survey among more than 200 participants and found that firms change the prices of their products once a year on average. A survey by Hall et al. (2000) of 654 firms concluded approximately the same. They found that the median firm claimed to change their prices roughly every year. Research more specifically focused on the restaurant branch found that in restaurants the median running time of prices is little over ten months. Bils and Klenow (2002) found similar results, but further differentiated their results in breakfast (price changes once every 11,6 months), lunch (10,7 months) and dinner (10,6).

Considering all of the research discussed above and the magnitude of the global corporate business McDonald's is, an adjustment period of one year would be the most realistic and comprehensible option to adjust the BMI. Thus follows that it is assumed to take one year for the price of a Big Mac to fully absorb all inflationary forces.

3. Data

The Economist's BMI compares burger prices across countries to measure a currency's over- or undervaluation under Purchasing Power Parity. This theory states that any nominal differences in the prices in local currencies should be smoothened by the foreign exchange rate. As clarified above, the hypothesis is that, due to price stickiness, the comparison of Big Mac prices across countries is out-dated as a measurement tool for PPP. At least, a better approximation is expected, once accounted for the inflation of consumer prices.[1] First, I computed values of over- or undervaluation of the BM prices and compared them with the original values of over- or undervaluation as composed by The Economist. I computed these deviations from the actual exchange rate for the publication date, as well as for sixty days thereafter.[2] Supported by the example of Switzerland the above will be illustrated.

Also, the theory implies that an overvaluation at time *t* will result in a movement of the exchange rate towards parity at *t+1.* So secondly, the means to test this predictive ability will be discussed.

In 2011 you paid SFr 6.5 for a Big Mac in Switzerland, while in the United States it only cost $4.07. According to PPP the Swiss franc/dollar exchange rate then should have been $\frac{6.5}{4.07} = 1.60$. In reality, the exchange rate was quite different at the time; 0.81 SFr/$. This means that a Swiss BM was relatively expensive compared to one from the U.S.A., converted to a dollar price ($\frac{6.5}{0.81} = \$8.06$). Of course it is impossible for any given person from Switzerland to buy a Big Mac in San Francisco instead of Geneva and benefit from this price differential. But it does tell us something about the bilateral exchange rate, or the currency's valuation. If burgers in both countries are composed of exactly the same ingredients, they should roughly cost the same via arbitrage. So if PPP holds, the Swiss franc must be relatively 'expensive', i.e. overvalued.

---

[1] http://data.worldbank.org/indicator/FP.CPI.TOTL.ZG

[2] Exchange rates at specific dates collected from: http://www.oanda.com/currency/historical-rates/

Indeed, The Economist calculated the Swiss currency to be overvalued by a staggering 98.00% in 2011. This means that the implied PPP (through the BM prices) deviates by this percentage from the actual foreign exchange rate. Once we take the inflation of consumer prices into account, the implied PPP of the adjusted BM prices falls to a deviation of 'only' 95.34% from the actual exchange rate. I will test whether the actual exchange rate can be approximated significantly better by making use of the adjusted BMI, for all the countries included in the original version.

An overvaluation like the Swiss currency would imply a relatively weak competitive position of the Swiss economy against the United States'. According to PPP theory, arbitrage opportunities will be exploited and Swiss exports will fall, lowering demand for the Swiss franc. This should establish an upward pressure on the bilateral exchange rate. To test if the (adjusted) BMI has any form of predictive power, I looked at the exchange rates again, sixty days afterwards. Looking at the huge overvaluation of the Swiss currency, you can expect a depreciation. And in this case, the exchange rate sixty days later has indeed risen to 0.9058 SFr/$. The adjusted BMI now deviates 74.68%, where the standard BMI differs as much as 76.64% from the actual exchange rate. This of course varies across countries, so I will also test if the adjusted BMI approaches the actual exchange rate significantly better than does the original BMI sixty days afterwards.

To test these hypotheses statistically, the results were paired per country. This way a Paired Samples T Test (PSTT) can be performed. This test compares the means of two variables. In this case it compares the over- or undervaluation according to the standard BMI, with the over- or undervaluation according to the adjusted BMI. First, the difference between the two selected variables is computed for each country. Then the average difference is tested to see whether it differs significantly from zero.

Because of the way the PSTT is computed, (the same applies to the Wilcoxon Signed Rank Test, which I shall discuss below) the data had to be altered to fit the test. The PSTT subtracts one variable from the other. But it does not look at the sign of the numbers. In other words, it looks at the lowest value instead of the value closest to zero (the closer to zero, the better the respective

BMI approaches the real exchange rate under PPP). Hence, every negative number is multiplied by -1 to make the data fit the test.

However, another problem arises with doing a T Test on this data. The assumption that the data has a normal distribution is not met. Some currencies have been structurally over- or undervalued (e.g. Norwegian krone, Chinese renminbi). This results in non-normality of the data. Two typical Q-Q Plots are added in the appendix to illustrate this (graphs 1 and 2). Besides the plots, two tests on normality were performed (Kolmogorov-Smirnov & Shapiro-Wilk), that both confirm this claim statistically. The results of these tests can be found in table 2. Obviously, normality of the data cannot be assumed. Although the Paired Samples T Test is more robust for this assumption than the unpaired version (Moore, McCabe, Duckworth & Alwan, 2009), a non-parametric test is performed for completeness. The Wilcoxon Signed Rank Test (WSRT) is the best fit, because it takes the magnitude of the difference between the two variables into account. Although the validity of the test is higher, it comes at a cost. The power of the test has decreased, because it does not make use of the information about the underlying distribution.

Now, I'll discuss the adjusted BMI's ability to predict future movements of exchange rates. In theory, an over- or undervaluation should result in a movement in the opposing direction. So first, I computed a dichotomous variable (did / did not move towards parity) to test for this. This variable, called 'goedfout*year* / goedfout*'yy'yy.dd*[3]', produced a 1 when moving towards parity and 0 when not. I will perform a Binomial Test to see if this variable can predict exchange rate movements.

This previous variable does not take into account the magnitude of the movement. So I computed another variable that measured the amount of percentage points the over- or undervaluation changed (PP*year* / PP*'yy'yy.dd*[4]). A positive value indicating a movement towards-, a negative value away from

[3] I.e. 'goedfout2008' for the year 2008, etc. / 'goedfout0809.30' for the movement between the publication dates in 2008 and 2009, thirty days after publishing, etc.

[4] i.e. 'PP2008' for the change in 2008, etc. & 'PP0809.30' for the change between the publication dates in 2008 and 2009, thirty days after publishing, etc.

parity. This was also tested to see if there are any predictions to be made with the adjusted BMI.

4. Results

For the obtained numeric results, I would like to refer to the appendix with SPSS output. The only side note is that SPSS performs the test with the null hypothesis stating no difference between the means (for both the PSTT and WSRT) and the alternative hypothesis stating that there is a difference. As such, SPSS looks at the two-sided significance of the test results. Here, the hypothesis is that the adjusted version approaches the actual exchange rate better than the standard version of the BMI. So we are only interested in whether or not its mean is significantly closer to zero, thus one-sided. Thus, all P-values in the output have to be divided by two. For purposes of comprehensability, the insignificant results were underlined.

The hypotheses for the Paired Samples T Test were as follows:

$$H_0: \mu_{\text{standard}} = \mu_{\text{adjusted}}$$

$$H_a: \mu_{\text{standard}} > \mu_{\text{adjusted}}$$

the significance level has been set at α = 0.05. Looking at table 1, one can derive that the adjusted BMI approaches the actual exchange rate significantly better in three out of eight comparisons than does the standard BMI. An argument can be made about the naivety of a confidence interval of 95% for variables that have proven to be so difficult to predict. The bold results in the table indicate significance using
α = 0.10%. Now for five out of the eight dates a significant improvement is found.

A shortcoming of the Paired Samples T Test is that it assumes a normal distribution of the data. Our data has no normal distribution, as can be obtained from graphs 1 and 2 in the appendix. Clearly, these graphs only provide information on the adjusted Big Mac prices in 2009. But due to the systematic behaviour of the deviations, the insight it provides is useful enough. Additionally, two tests have been performed to statistically establish the non-normality. The SPSS output of the tests is stated in table 2. Especially the low P-values for the Shapiro-Wilk Test make the assumption of normality questionable. As expected, a Wilcoxon Signed Rank Test is in order to test without the assumption of a normal distribution.

Our hypotheses for the WSRT were the same as under the PSTT:

$$H_0: \mu_{\text{standard}} = \mu_{\text{adjusted}}$$

$$H_a: \mu_{\text{standard}} > \mu_{\text{adjusted}}$$

the significance level has been set at $\alpha = 0.05$. As can be seen in table 3, a systematic difference between the two versions of the BMI is established for five of the eight dates. Again, this is at a significance level of $\alpha = 0.05$. At a significance level of $\alpha = 0.10$ (bold in the table), we would obtain a systematic difference for all but one date.

In conclusion some final remarks on the tests that have been used. First, the Paired Samples T Test is a more powerful test than the Wilcoxon Signed Rank Test. But this only applies to data that meet the assumptions of the test. The assumption of normality is violated in this instance. Although the PSTT is more robust than a non-paired T Test, a case can be made for the more valid WSRT, which does not assume normality. Either way, the WSRT showed the adjusted BMI approaches the actual exchange rate significantly better than the original one.

Now, we will look at the adjusted BMI's ability to predict movements in bilateral exchange rates. As mentioned earlier, an over- or undervaluation of a currency at time *t* should result in a movement in the opposite direction at *t+1.* So in the first place, I computed a dichotomous variable (did / did not move towards parity) for every country in all four years, to test for this. The benchmarks for the predictions were set at thirty and sixty days after the BMI was composed to check for the short-/ medium term. For the long-term adjustment, we looked at the gap of about a year[5] between two publications of the BMI, as well as the one year gap thirty and sixty days later. The variable produced a '1' for a good prediction and a '0' for a bad one.

[5] 2010's BMI was published in April instead of August. Consequently goedfout0910/PP0910 were all four months shorter and goedfout1011/PP1011 all four months longer.

The appropriate test for a dichotomous variable like this is the Binomial Test. This tests whether the probability of producing either outcome is significantly different from 0.5, which would imply a random movement of the exchange rate. First the short- and medium term will be discussed, as their results are quite similar. As can be derived from tables 4 (30 days) and 5 (60 days), the null hypothesis of equal probabilities cannot be rejected at a significance level of $\alpha = 0.05$. As a result, the BMI cannot be said to be a good predictor of the exchange rate in these periods on the basis of these tests.

The second variable I computed measured the number of percentage points the exchange rate moved. That is, the difference between the over- or undervaluation at composing date and the over- or undervaluation thirty and sixty days thereafter in percentage points. The variable subtracted the valuation at composing date, from the future dates. So a positive value indicated a movement towards PPP, whereas a negative value a movement away from parity and finally a value of zero when the exchange rate did not move. This was also tested to see if there are any predictions to be made with the adjusted BMI. This way a measure of magnitude was implied in the variable.

First a T Test was performed to ascertain whether the mean differed significantly from zero. For both dates (tables 6 and 8 respectively) it did not vary significantly from zero. Although, when the loosened significance level of $\alpha = 0.10$ is applied, 2008 yields significant results after thirty and sixty days. But as mentioned above, a One Sample T Test (OSTT) is more sensitive for non-normality. And again (see tables 7 and 9) for both dates the data has no normal distribution. So in this instance a non-parametric test is even more advisable. Again, I chose Wilcoxon Signed Rank Test. The results of which can be found in tables 10 and 11. For not a single year or date a significant result was acquired. This means, that as a predictor of future exchange rate movements the adjusted BMI performs very poorly on the short- and medium run, according to these tests. Apparently, the disturbing forces distort PPP too severely for the theory to hold as a tool for prediction.

Now supposedly, these forces distort PPP in the short run, but are less influential in the long run. Thus it was tested whether the predictive abilities were better in the long run. The same tests were performed as before[6], the results of which can be found in tables 12–15, respectively. Yet again, significant results were scarce. When the significance level was changed to $\alpha = 0.10$, the results in rows 1, 7, 8 and 9 appear to be significant. However, the variables of rows 7, 8 and 9 contain more 0's than 1's. Which means PPP predicted significantly in the wrong (according to the theory) direction in those cases. This also applies to table 15. Here, rows 1 and 2 prove to predict significantly and in the right direction though.

All in all, the results are not convincing at all. To label PPP with predictive abilities would be naïve for the short-, medium- and long term.

[6] Binomial Test on 'goedfout*'yy'yy.dd*'. OSTT, Kolmogorov-Smirnov & Shapiro-Wilk and WSRT on 'PP*'yy'yy.dd*'.

5. Conclusion

Over the decades various theories have been formulated, which try to explain exchange rates movements and price differentials between countries. Purchasing Power Parity has been one of the most commonly known theories explaining exchange rates, with the support of The Economist's Big Mac Index. According to PPP the exchange rate evens out price differentials between countries, when denominated in their respective currencies. But looking at prices, a lot of distortions to this equilibrium are present. The theory of price stickiness provides a reasonable argument as to why these differences exist and using this information, predictions could be made about where exchange rates are heading in the short-, medium- and long run.

However intuitively sound, empirically the theory does not live up to expectations. There are various practical factors disturbing Purchasing Power Parity. Trade barriers, non-traded goods and price setting by companies have been discussed as examples of these factors. Moreover, because it is put to use, with a market basket of goods, which takes a vast amount of time to compute, the predictive power of the theory is not quite useful. To surpass this and make the theory more comprehensible, The Economist put it to practice using Big Mac prices. Theoretically, the Big Mac's homogeneity across countries makes it a rather good replacement. Also, the prices are retrieved much faster than it takes to compute the general price level at any given time.

My aim was to improve on The Economist's BMI, as computed by them since 1986. To control for price rigidity, the prices were adjusted for the one-year's inflation. With these newly computed series of prices, I tested whether the index adjusted for inflation approximated the actual exchange rate significantly better than the standard version when PPP was implied.

The results from the Paired Samples T Test were not conclusive. But as the data was clearly not normally distributed, an argument could be made for a non-parametric test, the Wilcoxon Signed Rank Test. Although a PSTT is not as sensitive for non-normality as a One Sample T Test, the WSRT's results were more clear-cut. A systematic difference was obtained for five out of eight comparisons at a significance level of $\alpha = 0.05$. All but one proved to be systematically different, when a looser significance level of $\alpha = 0.10$ was applied. All in all, the Big Mac Index was indeed improved on as a closer approximation of the actual exchange rate. A research performed on a larger sample size qua years would be recommendable to improve on external validity.

As a predictor of future movements in the exchange rate, the adjusted BMI performed rather poorly. Even when a measure of magnitude was incorporated, the index didn't beat the odds. The disturbing forces, such as trade barriers, apparently interact too much with PPP. Even in the long run, no conclusive results were obtained, as some of the significant results pointed in the wrong direction.

In the end, The Economist's Big Mac Index was improved on in the sense that with the adjustment for inflation, it did approach the actual exchange rate better. As a tool for predictions it did not perform too well. But as Mark Twain is reputed to have said: 'predicting remains a very difficult job, especially when it regards the future'.

6. Bibliography

Ashenfelter, O., & Jurajda, S. 2001. Cross-Country Comparisons of Wage Rates: The Big Mac Index. Unpublished manuscript, Princeton University and CERGE-EI/Charles University, October 2001.

Baharad, E., & Eden, B. 2004. Price rigidity and price dispersion: evidence from micro data. *Review of Economic Dynamics*. Volume 7, Issue 3, July 2004, Pages 613–641.

Balassa, B. 1964. The Purchasing-Power Parity Doctrine: A Reappraisal. *Journal of Political Economy*, December 1964, 72(6), pp. 584-96.

Bils, M., & P.J. Klenow. 2002. Some Evidence on the Importance of Sticky Prices. Working Paper No. 9069. *National Bureau of Economic Research.*

Blinder, A.S., Canetti, E.R.D., Lebow D.E., & Rudd, J.B. 1998. *Asking About Prices: A New Approach to Understanding Price Stickiness*. New York: Russell Sage Foundation.

Cassel, G. 1918. Abnormal Deviations in International Exchanges. *The Economic Journal*, Vol. 28, No. 112, December, 1918 pp. 413-415.

Cassel, G. 1921. *The World's Monetary Problems*. London: Constable and Co.

Cassel, G. 1928. The Rate of Interest, the Bank Rate, and the Stabilization of Prices. *The Quarterly Journal of Economics* (1928) 42 (4): 511-529.

Data of Big Mac Indices: as published by The Economist

Data of exchange rates: http://www.oanda.com/currency/historical-rates/ (downloaded 1st April 2012 and 5th June 2012)

Data of inflation: http://data.worldbank.org/indicator/FP.CPI.TOTL.ZG (downloaded 1st April 2012)

Dornbusch, R. 1985. Purchasing Power Parity. Working Paper No. 1591. *National Bureau of Economic Research.*

Dornbusch, R. 1987. Exchange Rates and Prices. *American Economic Review*, March 1987, 77(1), pp. 93-106.

Fisher, I. 1928. *The Money Illusion*. New York: Adelphi Company.

Froot, K.A., & Rogoff, K. 1995. Perspectives on PPP and Long-Run Real Exchange Rates. In G. Grossman & K. Rogoff, (ed.), *Handbook of International Economics*. Volume 3. Amsterdam: North Holland Press.

Hall, S., Walsh, M., & Yates, A. 2000. Are UK Companies' Prices Sticky?. *Oxford Economic Papers* 52 (3), 425-446.

Historical exchange rates: http://www.oanda.com/currency/historical-rates/ (downloaded 3rd April 2012)

Hummels, D. 2001. Toward a Geography of Trade Costs. Unpublished manuscript, Purdue University, September 2001.

Krugman, P.R. 1987. Pricing to Market when the Exchange Rate Changes. In S.W. Arndt, & J.D. Richardson, (ed.), *Real-Financial Linkages among Open Economies*. Cambridge, MA: MIT Press, 1987 pp. 49-70.

Krugman, P.R. 1990. Equilibrium Exchange Rates. In W.H. Branson, J.A. Frenkel, & M.Goldstein, (ed.), *International Policy Coordination and Exchange Rate Fluctuations*. Chicago: University of Chicago Press, 1990, pp. 159-87.

Moor, D.S., McCabe, G.P., Duckworth, W.M. & Alwan, L.C. 2009. *The Practice of Business Statistics.* New York: W.H. Freeman and Company

Ong, L.L. 2003. *The Big Mac Index: Applications of Purchasing Power Parity*. New York: Palgrave MacMilan.

Pakko, M.R., & Pollard, P.S. 2003. Burgernomics: A Big MacTM Guide to Purchasing Power Parity. *Federal Reserve Bank of St. Louis Review*, November/December 2003 Vol. 85, No. 6, pp. 9-28.

Samuelson, P.A. 1964. Theoretical Notes on Trade Problems. *Review of Economics and Statistics*, May 1964, 46(2), pp. 145-54.

Sheshinski, E., & and Weiss, Y. 1977. Inflation and Costs of Price Adjustment. *The Review of Economic Studies*, Vol. 44, No. 2 (Jun., 1977), pp. 287-303

Shafir, E., Diamond, P. A. & Tversky, A. 1997. On Money Illusion. *Quarterly Journal of Economics,* 112 (2), pp. 341–374

Stiglitz, J.E. 2002. Information and the Change in the Paradigm in Economics. *The American Economic Review,* Vol. 92, No. 3 (Jun., 2002), pp. 460-501.

Taylor, J.B. 1999. Chapter 15 Staggered price and wage setting in macroeconomics. *Handbook of Macroeconomics*, Volume 1, Part B, 1999, Pages 1009–1050.

Appendix

*Table 1**

Comparison of the BMI with the adjusted (for inflation) BMI for every year as well as sixty days thereafter with a Paired Samples T Test

Paired Samples Test									
		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	**2008 % Over or Undervalued - 2008 % Over or Undervalued adjusted**	**0.00420%**	**0.03141%**	**0.00473%**	**-0.00535%**	**0.01375%**	**.886**	**43**	**.380**
Pair 2	2008 % Over or Undervalued (60) - 2008 % Over or Undervalued adjusted (60)	0.00771%	0.03058%	0.00461%	-0.00159%	0.01701%	1.673	43	.102
Pair 3	2009 % Over or Undervalued - 2009 % Over or Undervalued adjusted	0.01559%	0.04348%	0.00655%	0.00237%	0.02880%	2.378	43	.022
Pair 4	2009 % Over or Undervalued (60) - 2009 % Over or Undervalued adjusted (60)	0.02852%	0.08573%	0.01292%	0.00245%	0.05458%	2.206	43	.033
Pair 5	2010 % Over or Undervalued - 2010 % Over or Undervalued adjusted	0.00917%	0.04073%	0.00621%	-0.00337%	0.02170%	1.476	42	.147
Pair 6	**2010 % Over or Undervalued (60) - 2010 % Over or Undervalued adjusted (60)**	**0.00727%**	**0.04143%**	**0.00632%**	**-0.00548%**	**0.02002%**	**1.151**	**42**	**.256**
Pair 7	**2011 % Over or Undervalued Raw Index - 2011 % Over or Undervalued adjusted**	**0.00569%**	**0.03047%**	**0.00508%**	**-0.00462%**	**0.01600%**	**1.120**	**35**	**.270**
Pair 8	2011 % Over or Undervalued (60) - 2011 % Over or Undervalued adjusted (60)	0.00799%	0.02820%	0.00470%	-0.00155%	0.01754%	1.701	35	.098

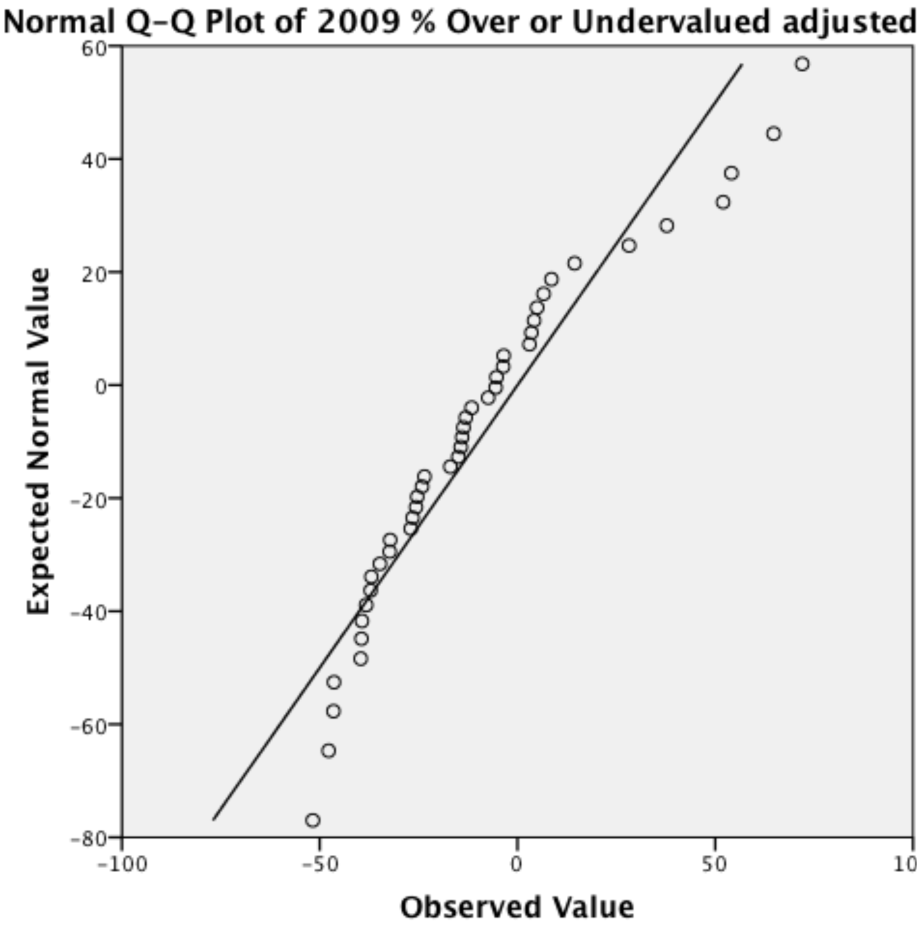
*Table 3**
Comparison of the BMI with the adjusted (for inflation) BMI for every year as well as sixty days thereafter with a Wilcoxon Signed Rank Test

Test Statistics[a]

	2008 % Over or Undervalued adjusted - 2008 % Over or Undervalued	2008 % Over or Undervalued adjusted (60) - 2008 % Over or Undervalued (60)	2009 % Over or Undervalued adjusted - 2009 % Over or Undervalued	2009 % Over or Undervalued adjusted (60) - 2009 % Over or Undervalued (60)	2010 % Over or Undervalued adjusted - 2010 % Over or Undervalued	2010 % Over or Undervalued adjusted (60) - 2010 % Over or Undervalued (60)	2011 % Over or Undervalued adjusted - 2011 % Over or Undervalued Raw Index	2011 % Over or Undervalued adjusted (60) - 2011 % Over or Undervalued (60)
Z	**-.969[b]**	-1.762[b]	-2.579[b]	-3.198[b]	-1.715[b]	-1.606[b]	-1.634[b]	-2.702[b]
Asymp. Sig. (2-tailed)	**.333**	.078	.010	.001	.086	.108	.102	.007

a. Wilcoxon Signed Rank Test
b. Based on positive ranks.

*Graph 1**
Typical QQ Plot of the distribution of the (adjusted) over- or undervalued data



*Graph 2**
Typical detrended QQ Plot of the distribution of the (adjusted) over- or undervalued data

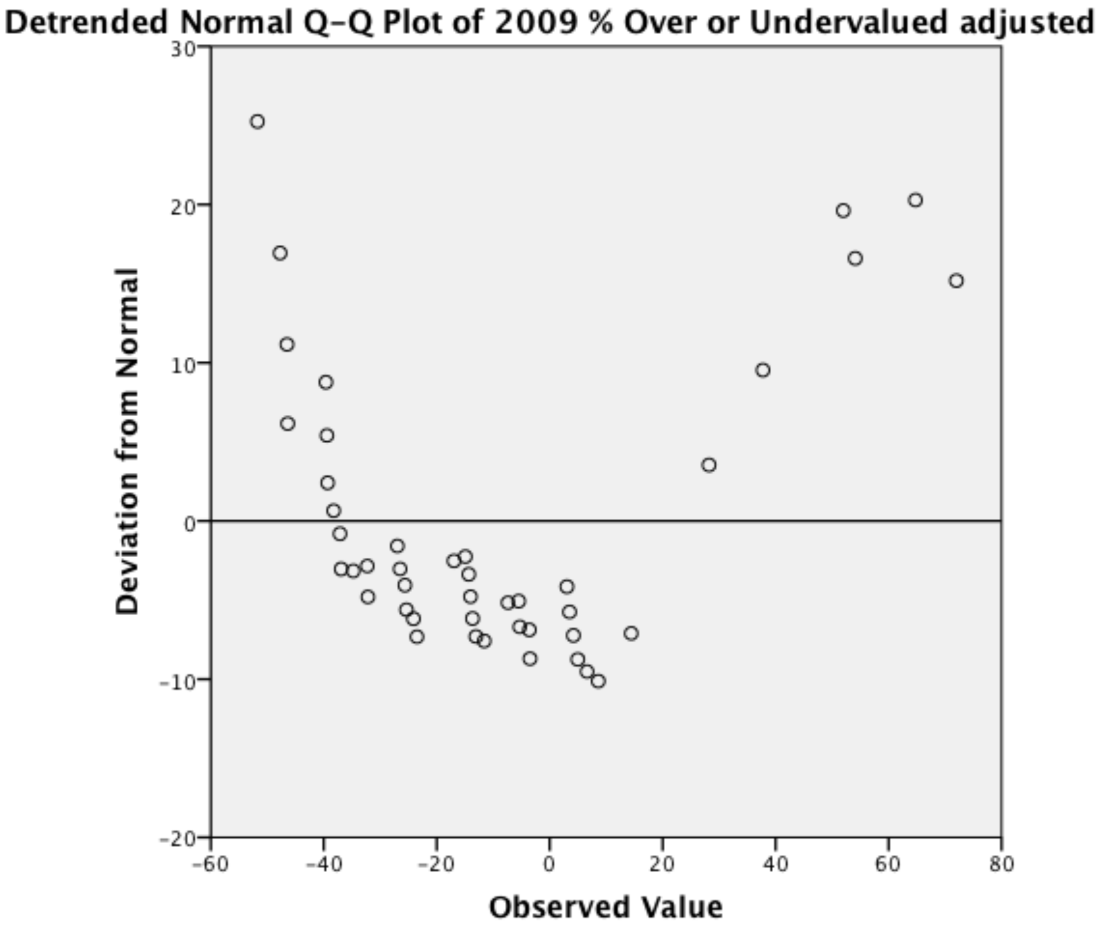


*Table 2**

Results from tests on normality of distribution of the percentages over- or undervaluation

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
2008 % Over or Undervalued	.138	44	.036	.911	44	.002
2008 % Over or Undervalued adjusted	.135	44	.043	.925	44	.007
2008 % Over or Undervalued (60)	.142	44	.026	.912	44	.003
2008 % Over or Undervalued adjusted (60)	.134	44	.045	.926	44	.008
2009 % Over or Undervalued	.125	44	.082	.900	44	.001
2009 % Over or Undervalued adjusted	.119	44	.131	.906	44	.002
2009 % Over or Undervalued (60)	.136	44	.041	.905	44	.002
2009 % Over or Undervalued adjusted (60)	.117	44	.150	.914	44	.003
2010 % Over or Undervalued	.138	43	.038	.876	43	.000
2010 % Over or Undervalued adjusted	.125	43	.090	.887	43	.001
2010 % Over or Undervalued (60)	.121	43	.124	.890	43	.001
2010 % Over or Undervalued adjusted (60)	.125	43	.087	.901	43	.001
2011 % Over or Undervalued Raw Index	.135	36	.097	.883	36	.001
2011 % Over or Undervalued adjusted	.131	36	.125	.889	36	.002
2011 % Over or Undervalued (60)	.154	36	.031	.883	36	.001
2011 % Over or Undervalued adjusted (60)	.157	36	.026	.888	36	.002

a. Lilliefors Significance Correction

Table 4†

Binomial Test on predictions of exchange rate movements by the adjusted BMI for ***30 days*** *later*

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The categories defined by goedfout2008 = 1.00 and 0.00 occur with probabilities 0.5 and 0.5.	One-Sample Binomial Test	.451	Retain the null hypothesis.
2	The categories defined by goedfout2009 = 0.00 and 1.00 occur with probabilities 0.5 and 0.5.	One-Sample Binomial Test	.651	Retain the null hypothesis.
3	The categories defined by goedfout2010 = 0.00 and 1.00 occur with probabilities 0.5 and 0.5.	One-Sample Binomial Test	1.000	Retain the null hypothesis.
4	The categories defined by goedfout2011 = 1.00 and 0.00 occur with probabilities 0.5 and 0.5.	One-Sample Binomial Test	.243	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Table 5†

Binomial Test on predictions of exchange rate movements by the adjusted BMI for ***60 days*** *later*

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The categories defined by goedfout2008 = 1.00 and 0.00 occur with probabilities 0.5 and 0.5.	One-Sample Binomial Test	.291	Retain the null hypothesis.
2	The categories defined by goedfout2009 = 0.00 and 1.00 occur with probabilities 0.5 and 0.5.	One-Sample Binomial Test	.291	Retain the null hypothesis.
3	The categories defined by goedfout2010 = 0.00 and 1.00 occur with probabilities 0.5 and 0.5.	One-Sample Binomial Test	.222	Retain the null hypothesis.
4	The categories defined by goedfout2011 = 1.00 and 0.00 occur with probabilities 0.5 and 0.5.	One-Sample Binomial Test	.243	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Table 6 §

*One Sample T Test on predictions of exchange rate movements by the adjusted BMI for **30 days** later*

One-Sample Test

	Test Value = 0					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
PP2008	-1.583	43	.121	-.01357	-.0308	.0037
PP2009	.777	43	.441	.00502	-.0080	.0181
PP2010	.108	42	.914	.00072	-.0127	.0141
PP2011	-.254	35	.801			

Table 7 §

*Results from tests on normality of distribution of difference in percentage points over- or undervaluation between composing date and **30 days** later*

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
PP2008	.217	44	.000	.935	44	.016
PP2009	.196	44	.000	.901	44	.001
PP2010	.188	43	.001	.795	43	.000
PP2011	.184	36	.003	.924	36	.017

a. Lilliefors Significance Correction

Table 8 §

*One Sample T Test on predictions of exchange rate movements by the adjusted BMI for **60 days** later*

One-Sample Test

	Test Value = 0					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
PP2008	-1.452	43	.154	-.02045	-.0489	.0080
PP2009	.852	43	.399	.00614	-.0084	.0207
PP2010	.070	42	.944	.00047	-.0129	.0138
PP2011	-.390	35	.699	-.00611	-.0379	.0257

Table 9 §

*Results from tests on normality of distribution for difference in percentage points over- or undervaluation between composing date and **60 days** later*

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
PP2008	.246	44	.000	.905	44	.002
PP2009	.233	44	.000	.928	44	.009
PP2010	.210	43	.000	.789	43	.000
PP2011	.165	36	.015	.940	36	.049

a. Lilliefors Significance Correction

Table 10 §

*Results from tests on whether the adjusted BMI predicted the correct movement of the exchange rate after **30 days***

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The median of PP2008 equals 0.00.	One-Sample Wilcoxon Signed Rank Test	.364	Retain the null hypothesis.
2	The median of PP2009 equals 0.00.	One-Sample Wilcoxon Signed Rank Test	.914	Retain the null hypothesis.
3	The median of PP2010 equals 0.00.	One-Sample Wilcoxon Signed Rank Test	.671	Retain the null hypothesis.
4	The median of PP2011 equals 0.00.	One-Sample Wilcoxon Signed Rank Test	.823	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Table 11 §

Results from tests on whether the adjusted BMI predicted the correct movement of the exchange rate after ***60 days***

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The median of PP2008 equals 0.00.	One-Sample Wilcoxon Signed Rank Test	.374	Retain the null hypothesis.
2	The median of PP2009 equals 0.00.	One-Sample Wilcoxon Signed Rank Test	.625	Retain the null hypothesis.
3	The median of PP2010 equals 0.00.	One-Sample Wilcoxon Signed Rank Test	.881	Retain the null hypothesis.
4	The median of PP2011 equals 0.00.	One-Sample Wilcoxon Signed Rank Test	.993	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Table 12❖

Binomial Test on predictions of exchange rate movements by the adjusted BMI for ***a year*** *later*

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The categories defined by goedfout0809 = 0.00 and 1.00 occur with probabilities 0.5 and 0.5.	One-Sample Binomial Test	.097	Retain the null hypothesis.
2	The categories defined by goedfout0809.30 = 0.00 and 1.00 occur with probabilities 0.5 and 0.5.	One-Sample Binomial Test	.291	Retain the null hypothesis.
3	The categories defined by goedfout0809.60 = 0.00 and 1.00 occur with probabilities 0.5 and 0.5.	One-Sample Binomial Test	.175	Retain the null hypothesis.
4	The categories defined by goedfout0910 = 1.00 and 0.00 occur with probabilities 0.5 and 0.5.	One-Sample Binomial Test	.880	Retain the null hypothesis.
5	The categories defined by goedfout0910.30 = 1.00 and 0.00 occur with probabilities 0.5 and 0.5.	One-Sample Binomial Test	.651	Retain the null hypothesis.
6	The categories defined by goedfout0910.60 = 1.00 and 0.00 occur with probabilities 0.5 and 0.5.	One-Sample Binomial Test	.880	Retain the null hypothesis.
7	The categories defined by goedfout1011 = 0.00 and 1.00 occur with probabilities 0.5 and 0.5.	One-Sample Binomial Test	.097	Retain the null hypothesis.
8	The categories defined by goedfout1011.30 = 0.00 and 1.00 occur with probabilities 0.5 and 0.5.	One-Sample Binomial Test	.097	Retain the null hypothesis.
9	The categories defined by goedfout1011.60 = 0.00 and 1.00 occur with probabilities 0.5 and 0.5.	One-Sample Binomial Test	.010	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Table 13 ‡

One Sample T Test on predictions of exchange rate movements by the adjusted BMI for ***a year*** *later*

One-Sample Test

	Test Value = 0					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
PP0809	-1.868	43	.069	-.04389	-.0913	.0035
PP0809.30	-1.895	43	.065	-.02811	-.0580	.0018
PP0809.60	-1.610	43	.115	-.02236	-.0504	.0056
PP0910	.450	42	.655	.00840	-.0292	.0460
PP0910.30	.577	42	.567	.00898	-.0224	.0404
PP0910.60	.251	42	.803	.00421	-.0296	.0380
PP1011	1.946	34	.060	.03249	-.0014	.0664
PP1011.30	1.704	34	.098	.02626	-.0051	.0576
PP1011.60	2.501	34	.017	.03214	.0060	.0583

Table 14 ‡

Results from tests on normality of distribution of difference in percentage points over- or undervaluation between composing date and ***a year*** *later*

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
PP0809	.149	44	.015	.972	44	.342
PP0809.30	.067	44	.200*	.981	44	.655
PP0809.60	.082	44	.200*	.976	44	.472
PP0910	.151	43	.015	.948	43	.052
PP0910.30	.151	43	.015	.937	43	.021
PP0910.60	.158	43	.009	.947	43	.046
PP1011	.095	35	.200*	.918	35	.013
PP1011.30	.142	35	.070	.879	35	.001
PP1011.60	.090	35	.200*	.941	35	.061

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.

Table 15 ‡

*Wilcoxon Signed Rank Test on predictions of exchange rate movements by the adjusted BMI for **a year** later*

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The median of PP0809 equals 0.00.	One-Sample Wilcoxon Signed Rank Test	.086	Retain the null hypothesis.
2	The median of PP0809.30 equals 0.00.	One-Sample Wilcoxon Signed Rank Test	.058	Retain the null hypothesis.
3	The median of PP0809.60 equals 0.00.	One-Sample Wilcoxon Signed Rank Test	.116	Retain the null hypothesis.
4	The median of PP0910 equals 0.00.	One-Sample Wilcoxon Signed Rank Test	.986	Retain the null hypothesis.
5	The median of PP0910.30 equals 0.00.	One-Sample Wilcoxon Signed Rank Test	.961	Retain the null hypothesis.
6	The median of PP0910.60 equals 0.00.	One-Sample Wilcoxon Signed Rank Test	.980	Retain the null hypothesis.
7	The median of PP1011 equals 0.00.	One-Sample Wilcoxon Signed Rank Test	.145	Retain the null hypothesis.
8	The median of PP1011.30 equals 0.00.	One-Sample Wilcoxon Signed Rank Test	.179	Retain the null hypothesis.
9	The median of PP1011.60 equals 0.00.	One-Sample Wilcoxon Signed Rank Test	.033	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

*

'. . . . *Over or Undervalued*…' are variables that measure the deviation of the implied exchange rate (by comparison of Big Mac prices) under PPP from the actual exchange rate. The numbers in front of them indicate which year; "*adjusted'* indicates whether the BM prices were adjusted for inflation; '*(60)*' indicates whether the deviation was measured sixty days after the composing date of the standard Big Mac Index.

†

All variables ('goedfout….') in both tables produced '1' when the exchange rate moved in the direction the adjusted BMI predicted (good prediction) after respectively thirty and sixty days. They produced '0' for movements in the other direction (bad prediction). This was calculated for every year in the data (e.g. 'goedfout***2008***').

§

'PP. . . .' is a variable that measures the difference between the over- or undervaluation on composing date and a later point in time in percentage points. This means it looks at how the exchange rate moved in that certain period. How long this period is varies per table and is denoted in the title of the table in bold. The variable was computed for 2008-2011 (e.g. PP***2008***).

❖

All variables ('goedfout….') in both tables produced '1' when the exchange rate moved in the direction the adjusted BMI predicted (good prediction) between composing date and a year later. They produced '0' for movements in the other direction (bad prediction). This was calculated for every year between composing dates (e.g. 'goedfout***0809***') and thirty and sixty days thereafter (e.g. 'goedfout0809***.30***').

‡

'PP. . . .' is a variable that measures the difference between the over- or undervaluation on composing date and a year later. This means it looks at how the exchange rate moved in that year. It was calculated for every year between composing dates (e.g. 'PP***0809***') and thirty and sixty days thereafter (e.g. 'PP0809***.30***').