

Did purchasing power parity hold in medieval Europe?

Article

Accepted Version

Bell, A. R. ORCID: <https://orcid.org/0000-0003-4531-0072>,
Brooks, C. ORCID: <https://orcid.org/0000-0002-2668-1153> and
Moore, T. K. (2017) Did purchasing power parity hold in
medieval Europe? *The Manchester School*, 85 (6). pp. 682-
709. ISSN 1467-9957 doi: <https://doi.org/10.1111/manc.12167>
Available at <https://centaur.reading.ac.uk/65800/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1111/manc.12167>

Publisher: Wiley

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online



Adrian R. Bell, Chris Brooks and Tony K. Moore, Did Purchasing Power Parity Hold in Medieval Europe?, *The Manchester School* (forthcoming)

AUTHORS' FINAL TEXT

Abstract

This paper employs a unique, hand-collected dataset of exchange rates for five major currencies (the *lira* of Barcelona, the pound sterling of England, the *pond groot* of Flanders, the florin of Florence and the *livre tournois* of France) to consider whether the law of one price and purchasing power parity held in Europe during the late fourteenth and early fifteenth centuries. Using single series and panel unit root and stationarity tests and cointegration analysis on ten real exchange rates between 1383 and 1411, we show that the parity relationship held for the pound sterling and some of the Florentine florin series individually and for almost all of the groups that we investigate. Our findings add to the weight of evidence that trading and arbitrage activities stopped real exchange rates deviating permanently from fair values. This research extends the results reported in other studies back more than 600 years.

JEL Classifications: F31, N13, N23.

Keywords: Law of one price, purchasing power parity, medieval markets, historical finance.

Acknowledgements: We are grateful to the Leverhulme Trust for funding this research under grant RPG-193 and to two anonymous referees for useful comments on a previous version of this paper. We also thank Francesco Bettarini and Helen Bradley for research assistance in collecting the exchange rate data.

1. Introduction

Two German ships arrived [in Ibiza] on the 13th from Flanders, and load salt for Germany... all here marvel at their coming - so long a road, only for salt!

(letter of 15th April 1404 from Giovanni di Gennaio in Ibiza to Francesco di Marco Datini & co. in Barcelona, translation in Origo, 1963 p.97).

Contrary to popular conception, the laws of supply and demand were well known in the Middle Ages. For instance, the French theologian Richard de Menneville (d.1302) proposed a thought experiment involving two countries, one in which grain was plentiful but wine scarce and the other in which wine was abundant and grain in short supply. As ‘the same commodity when plentiful is less appreciated than when it is scarce’, so ‘corn in country A will be cheaper than in country B, while conversely wine in country A will be dearer than in country B.’ Menneville argued that both countries would therefore benefit from exchanging their respective surpluses. Fortunately, ‘it is natural for the business of trade to equalize supply’ as ‘the merchant ‘buys corn cheap in country A and sells it at the higher price that is ruling in country B’, and vice versa for wine. Importantly, such intermediaries could justly profit from buying at the lower market price in one country and selling at the higher market price in the other (translated in Beer, 1938 pp.42-3). It should be stressed, however, that this falls somewhat short of the full modern concept because Menneville did not consider the impact of this trade on the market price or exchange rates in both countries, although this is implicit in his premises.

Then, as now, academics were not always ahead of market practitioners. Indeed, it is clear from the surviving mercantile correspondence that medieval merchants were eager to keep up to date with the prices of goods in different markets across Europe as well as to track exchange rate movements. A particularly illustrative example can be found in a letter of 11 July 1393 sent by Alamanno Mannini in London to Francesco di Marco Datini and Andrea di Bonanno in Genoa.¹ At the time of writing, the price of pepper in London was 13d but a ship had just arrived in Southampton with a large cargo of pepper, and Mannini planned to start buying once the price had dropped to 10d. However, the ship had not brought any ginger, which was then in short supply in both England and Flanders. As a result, good quality ginger was worth 25d per lb and was expected to rise to more than 30d. Meanwhile, Mannini continued, his correspondents in Venice had sent news that a Catalan ship had left Beirut in March bound for Barcelona, carrying a large consignment of ginger. Its arrival should cause the price of ginger in Barcelona to fall and, so long as no new supplies of ginger arrived in England or Flanders in the meantime, then Mannini thought Datini and Bonanno could profit from buying ginger in Spain and dispatching it north.

On the one hand, this example and that quoted above of the German ships loading up with salt in Ibiza is a clear demonstration of arbitrage in action. But on the other, it is not clear that Datini and Bonanno were able to capitalise on Mannini’s advice, while the fact that the

¹ *Archivio di Stato di Prato*, 777/37/313003. Translation by Helen Bradley.

merchants in Ibiza ‘marvel[ed] at the coming’ suggests that such arbitrage may not have been routine.

In modern financial theory, the same underlying concept is formalised in the notion of the ‘law of one price’ which asserts that the forces of arbitrage should ensure a single good (or asset) cannot simultaneously trade at different prices in different locations. Purchasing power parity (PPP) is an aggregate version of the law of one price – a theory purporting that a representative basket of goods and services should cost the same whatever country it is purchased in, once it is converted to a common currency. Unfortunately, testing for whether PPP holds in practice is fraught with difficulties, even with the long, relatively high frequency datasets available in the modern context, and while PPP appeals to economists’ senses of how markets should behave, the data often have other ideas. For example, even a cursory examination of the data seems to show long-lived deviations from parity and real exchange rates are found to be just as noisy as their nominal counterparts (see Rogoff, 1996), demonstrating that relative price levels do little to explain the latter’s volatility.

In this paper we contribute to the debate on the empirical validity of the law of one price using a rich, hand-collected dataset of medieval exchange rates for a period of thirty years during the late fourteenth and early fifteenth centuries. Importantly, and unlike previous studies, we are able to compare nominal prices using contemporary exchange rates rather than relying on precious metal equivalents derived from the supposed intrinsic value of the coinage. In addition, while there are now several studies using very long spans of data going back to the late fourteenth and early fifteenth centuries in some cases, it seems implausible that series of such length are not subject to structural breaks or regime shifts, which may result in a spurious finding against PPP (see for example Perron, 1989; or Leybourne, Mills, and Newbold, 1998). Finally, these long-run studies have generally concentrated on just one currency pair whereas we investigate the inter-relationships between five of the most important currencies of medieval Europe. In summary, our research adds to a growing body of evidence regarding the sophistication of early financial markets.

The remainder of this paper develops as follows. Section 2 provides an overview of the historical and economic literature. In Section 3 we define the notation and the key models used in the literature on purchasing power parity. The sources and nature of the unique dataset that we employ are discussed in detail in Section 4. Sections 5 and 6 explain the econometric approaches used to test for PPP in the single-series and panel data contexts respectively, and the results from applying those methods are presented and discussed. Finally, Section 7 includes some concluding comments and offers suggestions for further research.

2. Literature review

Within the historical literature, there is a long-standing debate on the extent to which historical markets were integrated (Federico, 2012), and especially the grain market (Persson, 1999). This work is clearly of relevance to the debate about whether PPP held, although such research typically does not directly address the issue. Although the bulk of such work focuses

on the period post-1500, as the necessary price and exchange rate data is more abundant, there is a significant body of research investigating medieval markets.² We may identify three strands of particular relevance to the current study.

First, there is a substantial body of research on price history going back to the thirteenth and fourteenth centuries. An important strand in the current historiography, following Allen (2001), has explored the ‘Great Divergence’ in real wages between Europe and the rest of the world and the ‘Little Divergence’ between the north and the south of Europe. On a smaller scale, van der Wee (1993) and Munro (2003) compared wages and prices in England and the Low Countries and Malanima (2013) in London and Florence, while Álvarez-Nogal and de la Escosura (2013) compiled a Consumer Price Index (CPI) for Spain. By contrast, there is relatively little work on French wages and prices during the Middle Ages, with the partial exception of Miskimin’s (1963, 1984) studies of the relationship between the frequent changes in the standard of the coinage and grain prices.

Second, there is a lively debate about the extent of integration in the medieval wheat market. Unfortunately, there is very little quantitative evidence for medieval trade flows, either locally, regionally or internationally, so more recent work has instead used the correlation between prices in different cities to assess integration. From this evidence, it has been argued that wheat markets were at least regionally unified in England (Galloway, 2000; Clark, 2014), the Low Countries (Van der Wee, 1963) and Italy (Epstein, 2000). In terms of international integration, Hybel (2002) suggests that there was an active grain trade between England, the Low Countries, Scandinavia and the Baltic as far back as the thirteenth century, although the lack of reliable price data means that this cannot be demonstrated quantitatively. From the fourteenth century, Söderberg (2006) finds evidence of integration in the wheat markets around the North Sea and also in Northern Italy but not in the wider Mediterranean area (including Spain). It should be noted that these results have been challenged by Unger (2007), who argues that most cities were supplied by their immediate hinterlands and that grain markets were integrated locally but not regionally or internationally.

The third research theme has focused more specifically on financial, rather than goods, market integration. Much of this is based on the possibility for arbitrage between the nominal values of different coinages and their precious metal contents (Boerner and Volckart, 2011, Chilosì and Volckart, 2011). Transaction costs bands that precluded such arbitrage trading were around 7% in Basel between 1365 and 1429 (Kugler, 2011) and still of the order of 6% in sixteenth century Spain (Bernholz and Kugler, 2011). By comparison, Bignon, Chen and Ugoli (2013), employing a threshold autoregressive (TAR) model, show that the transaction costs thresholds in European exchange rate markets had fallen to well below 1% by the mid-nineteenth century. Perhaps surprisingly, Li (2012), using exchange rate data from the Datini archive (the same source as used in this paper), found that triangular arbitrage kept the direct exchange rate for bills of exchange between Venice and Paris and the cross-exchange rate

² Federico (2012) only identifies two works on market integration before 1550 (Froot, Kim and Rogoff, 1995 and Bateman, 2011). This is because he restricts his survey to papers available online while many of the medieval studies discussed in this paper are only available in print.

Venice-Bruges-Paris within 1% of each other. However, it should be noted that arbitrage did not operate so effectively between Venice-Barcelona and Venice-Bruges-Barcelona, where the threshold bounds were closer to 8%.³

This suggests that some financial centres and trading routes may have been more closely integrated than others during the Middle Ages. In particular, there seems to be greater integration along the north-south axis between England, the Low Countries and Italy than east-west across the Mediterranean. Moreover, the level of integration in the most closely-linked areas during the later Middle Ages may even have been comparable with the early nineteenth century. For instance, Bateman (2011) finds there existed a period of convergence in European wheat markets at the end of the Middle Ages, followed by disintegration during the late sixteenth and early seventeenth centuries before a recovery subsequently took place. The idea of an integration crisis in the seventeenth century is also supported by Jacks (2004) and may plausibly be associated with the Thirty Years War. As a result, market integration c.1800 may only have recovered to the level reached before c.1500 (at least in certain areas). Nonetheless, that the markets functioned so well before railways and telecommunications, or modern economic theory, is perhaps surprising. This paper will further investigate this precocious integration by testing for the validity of PPP in medieval Europe, using the price data described above and our hand-collected dataset of exchange rates.

Focusing on PPP specifically, with a few exceptions, we can broadly categorise there having been three phases in the development of the literature on testing for it. The first generation of approaches applied unit root tests to time-series of real exchange rates of modest length (typically 15-30 years, as in our sample period). The idea is that if real exchange rates contain a unit root then PPP does not hold while the reverse conclusion holds if they are mean-reverting. These tests mainly failed to reject the null hypothesis that the series are random walks, containing a unit root, thus refuting the theory (see Roll, 1979; Alder and Lehmann, 1983; Meese and Rogoff, 1988 and Taylor, 1988). However, it is widely known that unit root tests lack power in finite samples, and this may partially explain the early evidence against mean-reversion. In the present context, this argument is put forward by Frankel (1986), who argues that the samples used in previous research had been insufficient to reliably capture the slow reversions back to parity in exchange rate and pricing systems.

The second generation of tests uses longer runs of data and slightly less restrictive versions of the tests. Such tests are often based on the presence or otherwise of cointegration between the (logs of) exchange rates and domestic and foreign price levels rather than unit root tests on real exchange rates, equivalent to the restriction that there is a cointegrating vector $[1 \ 1 \ -1]$ applying to the three series (see Rogoff, 1996). Perhaps as a result, this kind of research typically finds evidence in favour of long-run PPP (e.g., Edison, 1987; Glen, 1992; Steigerwald, 1996).

³These were based, however, on financial instrument transactions rather than bullion.

Within this second generation of PPP tests exist studies in both the historical and economics literature employing several hundred years of data from one or two specific exchange rate series. However, given the inherent difficulties in compiling sufficiently long and detailed datasets to examine PPP using historical data, it is hardly surprising that the number of such studies is very small. Lothian and Taylor (1996) examine almost 200 years of data for the US dollar – British pound and French franc – pound going back to around 1800. They find support for PPP over the long run, a conclusion that is disputed by Cuddington and Liang (2000). The latter study, while using the same data and broad testing approach as Lothian and Taylor, refute PPP when different lag lengths or trends are included in the test regressions.

More recent research by Kugler (2013) employs a threshold-type model to examine the extent to which arbitrage worked in the Dutch guilder – British pound sterling exchange rate during the 1600-1912 period; however, it is not the objective of his study to consider price levels or to test PPP. Lothian and Devereux (2011) employ the same series over a longer period extending to 2009. They argue, on the basis of a unit root test on the real exchange rate, in favour of long-run PPP holding. The guilder – sterling rate is again the focus of a test, but of the narrower law of one price, using an even longer span of data (c. 1273-1991) in Froot, Kim and Rogoff (1995). Although they do not formally test PPP, they conduct separate analyses using six food commodities. They document the time-series properties of the deviations from the law of one price, showing that they are highly correlated across series and stable over time, bearing strong similarities to the properties of their modern-day counterparts. However, they lacked exchange rate data for the medieval period, and instead their analysis expressed prices in terms of grams of silver based on the ostensible metallic content of the two coinages.

The third generation of PPP studies employs more recently developed and sophisticated econometric techniques involving panel unit root and cointegration models. These tests are generally more powerful than single equation approaches and as such, have resulted in more support for PPP than had existed previously – see, for example, Frankel and Rose (1996) and Coakley and Fuertes (1997). However, while the weight of evidence from panel data supports PPP, these favourable findings are far from universal (see Pedroni, 2001 or Xu, 2003 for contrary results). There are significant differences in the conclusions reached according to the precise test, sample period and even base currency employed.

Econometrically, even panel techniques have not been above criticism. In particular, Banerjee, Marcellino and Osbat (2005) demonstrate using a Monte Carlo simulation that panel unit root tests not linked to possible cointegrating relationships between the series can be horrendously over-sized, finding in favour of PPP even when the null hypothesis of a unit root is correct. They argue that this size-distortion may be responsible for the increasing tendency to favour the PPP relationship since these tests have been adopted. Of the panel unit root tests available, they recommend the Levin and Lin approach over Maddala-Wu (1999) as being the least oversized.

In addition to the more focused research discussed above that uses specific techniques and datasets, there are several comprehensive surveys of the literature on the empirical support for PPP, including Dornbusch (1985), MacDonald (1995), Rogoff (1996), and more recently Sosvilla-Rivero and García (2003) and Taylor and Taylor (2004). However, it is probably fair to say that the increasing growth in the variety and sophistication of available econometric tests for PPP has merely served to further cloud the conclusions about which version of the theory holds and to what extent (if at all).

3. Models

Let S_t denote the nominal exchange rate at time t measured in local (domestic) currency units per unit of a foreign currency, $P_{i,t}$ be the local price of good i and its foreign price be $P_{i,t}^*$. The law of one price states that the exchange rate should be set so that domestic and foreign prices are equal when measured in a common currency

$$P_{i,t} = S_t P_{i,t}^* \quad (1)$$

Aggregating across goods i and assuming that the law of one price holds for each of them leads to the absolute version of purchasing power parity. Notationally, if we let P_t denote the aggregate local price level and P_t^* be the aggregate foreign price level for the same basket of goods and if we further let lower case letters denote logarithms of these quantities, we could write the real exchange rate at time t , e_t , as

$$e_t \equiv s_t + p_t^* - p_t \quad (2)$$

If PPP holds, then the real exchange rate in (2) should be constant over time. This is sometimes known as long-run PPP, with short-run PPP being violated at any point when the current real exchange rate does not accord with its equilibrium value (see Abuaf and Jorion, 1990). In practice, it is unreasonable to assume that e_t will be completely time-invariant since even if the forces of arbitrage work well, prices are likely to be sticky in the short-run. Arguably, a more sensible test is whether it has a constant mean, constant variance, and constant autocovariance structure – in other words, whether it is covariance stationary. Thus a considerable body of research has employed the Dickey-Fuller or Phillips-Perron tests, described below, for non-stationarity or the KPSS stationarity test.

Other research has argued that such non-stationarity tests are strict evaluations of PPP. Dornbusch (1985), for instance, notes that it is implausible that the same goods will always trade at the same price in different locations due to transport costs and so on, even if they are perfectly homogeneous commodities. This motivates a consideration of the changes in the log of the real exchange rate as a test for “relative” (or “weak”, as Dornbusch terms it) PPP, which is a generalisation of absolute PPP requiring only that changes in the exchange rate move in proportion to the difference between the inflation rates of the two countries. Thus, if absolute PPP holds, relative PPP must also hold but not vice versa. An advantage of relative

PPP as an empirical proposition is that it does not require the baskets of goods in the home and foreign countries to be the same – merely that they do not change over time. Thus relative PPP holds as a response to a monetary shock which moves all goods prices and the exchange rate proportionally. Therefore the basket weights play no role and they can vary between countries and over time without affecting the PPP property. Moreover, the law of one price does not necessarily hold for all goods; by contrast, absolute PPP requires the same basket weights and the law of one price for every individual good included in the basket.⁴ Thus relative PPP is a useful relaxation since, in practice, national price indices will only include a sample of goods and services in an economy and will differ internationally reflecting variations in purchasing behaviour and accounting conventions.

4. Data

This paper focuses on the real exchange rates between five European currencies – the florin of Florence, the *livre tournois* of France, the pound sterling of England, the *lira* of Barcelona and the *pond groot* of Flanders during the period c.1383 to 1411. The currencies and date range chosen reflect the coincidence of data on both commodity prices and exchange rates.

The substantial literature on price history has been introduced above. We employ two sets of price series in parallel – first, a general price level based on a Consumer Price Index (CPI) and second, a more specific focus on the price of one particular commodity, wheat. The single commodity serves as a medieval counterpart to the Economist’s ‘Big Mac’ index.

Our CPIs are based on those constructed by Munro (2003) for Bruges and London, Malanima (2013) for Florence, and Álvarez-Nogal and de la Escosura (2013) for Spain (Barcelona). Allen and Munro’s indices are based on baskets of goods representing a subsistence budget for a worker. They are broadly comparable; Allen’s basket uses prices for bread and beer (derived from the wheat and barley prices for this period) while Munro uses wheat, rye and barley. Both baskets also contain allowances for meat, dairy, fuel and textiles. Álvarez-Nogal and de la Escosura, by contrast, use a *divisia* index weighted (75%) towards agricultural products (wheat, barley, rye, oats, straw, wine, olive oil, chicken, mutton, rabbit, and cheese) with a lesser weighting (25%) for industrial products (timber, plaster, lime, tiles, nails as building materials; coal and wood for fuel; cloth, linen and silk as textiles; and paper, parchment and wax). As a result, the Spanish *divisia* index is largely composed of the same basic commodities as the subsistence baskets. There is no existing CPI index for France and, given the nature of the available sources, it is not currently feasible to construct one. Allen’s series for Paris only starts in 1431, when it is derived from Baulant’s (1968) prices for wheat in Paris. We follow a similar approach and use the wheat price in Douai (see below) as a proxy for changes in French CPI.⁵

⁴ We are grateful to an anonymous referee for providing this clarification of the implications of absolute versus relative PPP.

⁵ Thus the price series for the basket and for wheat are the same in the French case. We prefer this solution to the alternative of dropping the country from the analysis altogether. While we should note the limitation of the use of wheat as a basket proxy in this case, the likely outcome is that the tests for PPP for exchange rate pairings including France should be made more conservative – in other words, it is more plausible that a parity relationship that existed is not discernible than a spurious relationship created where it did not exist in reality.

Although these indices differ in their precise contents between countries based on local conditions, they are likely to be more comparable than modern CPI baskets. As subsistence baskets, they largely contain basic foodstuffs: bread (from wheat), beer (from barley) or wine, dairy products, meat, and fuel. They do not include rent or services, which make up a majority of modern baskets, or luxury goods. Also, taxation and regulation probably had less impact on medieval prices than they do today.

As a further test, we also calculate real exchange rates using wheat prices. Wheat is chosen because the data are both accessible and relatively easy to compare, and there is evidence that wheat was an internationally traded good. This analysis will also contribute to the debate over the extent to which the medieval wheat market was integrated, by comparing nominal prices rather than silver equivalents. Wheat prices for London and Bruges are taken from Munro (2003) and Florence from Pinto (1993). There are no wheat prices from Barcelona or Paris during this period. For Barcelona, wheat prices were taken from the nearby town of Lleida (Argilés, 2010). The French price data is especially scattered and inconsistent. The most complete set of data from this period, which we employ, is from Douai in northern France (Mestayer, 1963).

The key innovation in this paper is the use of market rates for bills of exchange as quoted in mercantile letters to calculate real exchange rates. There are difficulties in compiling evidence about exchange rates, as discussed below, and previous studies have compared prices in terms of silver equivalents using the ostensible metallic contents of the coinage. As Beveridge (1939) pointed out, ‘to describe silver and gold equivalents as prices is to ignore the nature of money and to confuse barter with exchange by the use of money’. Further, it is now generally accepted that medieval coins circulated by face value and not by weight (Munro, 2012). Changes to the monetary standards certainly affected exchange rates and prices but such adjustments were not immediate (Spufford, 2012). This lag between debasement of/enhancement to the metallic content of the coinage and the adjustment of prices is another reason why market exchange rates are preferable to silver equivalents (which essentially assume that this process was instantaneous) as a basis for comparison.

The real constraint is thus the relative paucity of standardised and comparable exchange rates between our currencies. The standard reference for exchange rates during the Middle Ages is the *Handbook of Medieval Exchange* (Spufford, 1986). However, these rates were taken from a variety of different sources, leading to major issues with comparability. Moreover, the *Handbook* quotes an average of one or two rates for each year. This presents a real problem because there are strong seasonal fluctuations in the exchange rates (Bell, Brooks and Moore, 2013) and intra-year variations are usually significantly greater than changes from year to year. There are also frequent gaps in the series for each currency. Fortunately, at least for modern historians, when the ‘merchant of Prato’ Francesco di Marco Datini died in 1411, he left his money and property to a charitable foundation. Incidentally, this bequest also included his business records, comprising nearly 600 account books and 150,000 business letters, mostly dating from 1383, when Datini entered the merchant banking business, and

continuing until shortly after his death in 1410 (Melis, 1962). These letters have now been catalogued and images can be consulted online.⁶ Many of these business letters provide market prices of goods and finish by quoting the current exchange rates for bills of exchange. Thus they are the precursors of the early modern price currents, the financial press of the nineteenth and twentieth centuries, and today's Bloomberg or Thompson-Reuters terminals.

We have extracted our exchange rate dataset from this source, using an expanded version of the exchange rates from Barcelona and Bruges collected by de Roover (1968) and our own hand-collected dataset for Florence, London and Paris. This source avoids many of the problems with the *Handbook* discussed above. We use one type of exchange rate, namely quotations of market rates for bills of exchange rather than spot or official rates, taken from one source the Datini correspondence. Moreover, the relatively high frequency of the observations allows us to smooth out the seasonal variations in the exchange rates. For each currency pair, we first calculate an average exchange rate for each month from all the observations for that pair. We then calculate an annual exchange rate as the average of all the monthly rates from that year. The data are summarised in Table 1.

An additional complication is that the modern PPP literature uses spot exchange rates. The speed of modern communication and the efficiency of the foreign exchange (FX) market today means that, in effect, the exchange rate between the pound sterling and the Euro is the same in London as in Paris or Frankfurt. This was not the case in the Middle Ages. For instance, the pound sterling-*pond groot* exchange rate quoted at London differed from that quoted at Bruges. In part, this reflected the slower speed of medieval communications. Postal times between London and Bruges or Bruges and Paris averaged four to six days during the period under study, compared to 20-23 days between Barcelona and Paris/Bruges/Florence or Florence and Paris/Bruges, and 27-30 days between London and Barcelona/Florence (Spufford, 1986). As a result, there was always a lag between a change in one city and the arrival of this information in another. Furthermore, each bill of exchange was only payable after a set period, either calculated from when the bill was drawn or when it was presented to the payor. This was known as the *usance* period and varied depending on the distance between city pairs. The *usance* periods between the cities in our sample is shown in Table 2, and was usually between two weeks and two months.

A further consequence is that all bills of exchange necessarily involved the extension of credit; the buyer of a bill was effectively a lender and the seller a borrower. The 'time value of money' was incorporated into the spread between the exchange rates as quoted in different cities. Thus the exchange rate between the pound sterling and *pond groot* quoted in London not only differed from that in Bruges but it was almost invariably more favourable to the buyer, on average by 2.16% between 1392 and 1406, which, taking into account the *usance* period of one month in each direction, equates to a compounded annualised interest rate of 13.0%. Using the exchange rate data from the Datini letters for seven financial centres, it has

⁶ These letters have now been catalogued and images can be consulted online at <http://datini.archiviodistato.prato.it/>

been possible to demonstrate consistent interest rates of between 10% and 15% per annum, varying with the depth of the local money market and the term of the loan (Bell, Brooks and Moore, 2015). Indeed, it has been argued that merchants used exchange transactions to circumvent the church prohibition on usury (Bell, Brooks and Moore, 2009).

This has a further advantage for the current paper. One element of an arbitrage trade is the ‘cost of carry’, which includes either the interest paid on money borrowed to fund the position, or the opportunity cost of investing funds in a non-interest bearing commodity. As explained above, however, an element of interest is already included within the market rates for bills of exchange quoted in the Datini letters. Moreover, for a would-be arbitrageur in London, the ‘bill’ rate was arguably more relevant than either a notional spot rate or the relative metallic content of the coins since he had the option either to buy goods and ship them to Bruges for sale or to buy a bill of exchange payable in Bruges. In the latter case, the bill rate would almost always be more favourable than either the spot rate or the intrinsic value of the coinage based on its metallic contents because it included an element of interest. Using bill exchange rates should therefore be a more rigorous test for the presence of arbitrage opportunities.

These features of the medieval FX market raise an interesting possibility; since the exchange rates at the two places could, to some extent, move independently depending on local monetary conditions and interest rates, it would be theoretically possible for PPP to hold in one direction but not in the other. Finally, it should be noted that we do not have both quotations at both ends for all currency pairs. For example, letters from Florence routinely quote exchange rates for Bruges, London and Paris but none of the latter three quotes direct exchange rates for Florence. A merchant wishing to move funds from Bruges to Florence would have presumably had to transfer his money via a third place, most likely Genoa or Venice. Our analysis is limited to those currency pairs for which we have direct quotations: we do not attempt to estimate an exchange rate between, for example, Barcelona and London by using cross-rates via Bruges. This is not to rule out the possibility of such transactions but using estimated exchange rates would complicate our analysis as it would rely on assumptions about the effectiveness of arbitrage that have not yet been proven.

Although we have exchange rate data on a monthly basis, we do not have access to monthly data on prices, since such information is not reliably available from the medieval period. While it would be interesting to explore the possibility for intra-year arbitrage, this is not currently feasible. Thus we work on an annual frequency. However, evidence suggests that the length of the sample is more important than the frequency to cover as many cycles of deviations from and reversions to PPP; Abuaf and Jorion (1990) argue that such cycles may typically be of three years’ duration and Perron (1991) presents the econometric arguments underpinning this proposition.

5. Tests on Individual Real Exchange Rates

The core of our analysis focuses on the use of unit root and stationarity tests on the real exchange rates as defined in (2). We focus on the absolute version of PPP since we find the

results sufficiently clear that it is unnecessary to examine the weaker relative version. We first employ the standard Dickey-Fuller (DF) and Augmented Dickey-Fuller (ADF) tests using a constant but not a trend in the test regressions – see Dickey and Fuller (1979) and Fuller (1976). We employ Schwarz’s Bayesian information criterion to select the number of lags to use in the ADF test. Given our modest sample size, it selects a zero lag order as optimal for all of the series except for the London-Bruges rate when wheat is used as the price variable, and in that case three lags are chosen. Cognisant of the limited power of unit root tests in finite samples noted above, in parallel we employ the KPSS (see Kwiatkowski, Phillips, Schmidt, Shin, 1992) test. While both the DF and ADF tests have non-stationarity under the null hypothesis, the KPSS test has stationarity under the null hypothesis. Thus we are able to perform a confirmatory data analysis that may be able to shed some light on whether the findings of one approach are robust to reversing the null and alternative hypotheses (i.e., PPP is supported in both cases or not supported in both cases) or whether they may be due to a spurious rejection/non-rejection that leads to contradictory results.

The results are presented in Table 3, with the first panel containing the results from constructing the real exchange rates based on baskets of goods, while the second contains the results from constructing them based on wheat only. For each panel, the first row reports the ADF test statistics, while the second contains the corresponding critical values which vary by sample size. The third row gives the KPSS test, and thus PPP is supported when the null hypothesis is rejected in the first row but not rejected in the third. The KPSS critical values are given in the notes underneath the table.

Considering first the real exchange rates based on baskets, evidence in favour of PPP is found for the Bruges-London, Florence-Bruges, Florence-Paris and Paris-Bruges rates, while no support for PPP is found for Barcelona-Bruges, Bruges-Barcelona and Florence-Barcelona. In the case of Florence-Bruges, the rejection of the random walk null of the Dickey-Fuller test occurs at the 1% level, indicating overwhelming evidence for mean-reversion in these series. These results are corroborated by the strong non-rejection that occurs with the KPSS test for the same series, with a statistic of 0.21 compared with a 10% critical value of 0.35. The reverse is true for the Barcelona-Bruges and Bruges-Barcelona rate where the DF statistics are well below their corresponding critical values, even at the 10% level but the KPSS test shows rejections at somewhere between the 5% and 1% levels. Finally, in the case of the Bruges-Paris rate, the results are mixed with both the ADF and KPSS tests showing rejections at the 5% level.

This illustrates the need for a nuanced interpretation of medieval market integration. Overall, our results show that there was a greater degree of integration in what has been described as the ‘blue banana’, that is the arc of cities curving south to north from Northern Italy through the Low Countries to Southern England than there was east to west across the Mediterranean between Italy, France and Spain (Spufford, 2000). The ties between London and Florence via Bruges may also reflect the role of the latter as a ‘node in networks that stretched from the Levant to Ireland’ (Murray, 2005, pp.229). At the same time, the lack of evidence for PPP holding between Bruges and Florence on the one hand and Barcelona on the other is

noteworthy given the triangular trading and financial relationships between those three cities, which is clearly evident from the Datini correspondence itself (de Roover, 1968).

Turning to the wheat-only measures of price (Panel B of Table 3), there are now more conflicting results between the tests having non-stationarity and stationarity under the null hypothesis. For instance, both the DF/ADF and KPSS tests for the Barcelona-Bruges and Florence-Barcelona rates reject the null hypotheses of non-stationarity and mean-reversion respectively at the 10% level or lower; similar conflicting findings arise for the Bruges-Paris rate (both tests reject) and the Bruges-Paris and Florence-Paris rates (neither test rejects).

Table 3 also presents the half-lives of each series, calculated from the slope parameter in a first order autoregressive representation in each case. These are mostly within the range of one to four years, figures that are comparable to those observed from modern data (e.g. Lothian and Taylor, 1996), indicating that the impact of a unit shock will die down relatively quickly. This represents yet another feature of foreign exchange series that appears to have changed remarkably little over several hundred years in addition to their volatilities (see Froot, Kim and Rogoff, 1995).

In order to obtain more information about the speed of adjustment of any deviations from equilibrium back to the long run solutions, we also implement a set of two-step single equation cointegration tests. In the first step, a regression of the log of the exchange rate is run on a constant, the log of the domestic price series and the log of the foreign price series with a unit root test being conducted on the residuals. In the cases where there is evidence that the residuals are stationary, we proceed to the second stage where a regression of the change in the log of the exchange rate is regressed on a constant, the change in the domestic price series, the change in the foreign price series and the residual series (deviation from equilibrium). The results are presented in Table 4, with again Panel A showing price levels measured using a basket and Panel B where prices are measured using wheat (except for France where a broader basket is unavailable and therefore we also use wheat).

The results are mostly consistent with Table 3, showing evidence in favour of cointegration and therefore that PPP holds for Florence-Barcelona, Florence-Bruges, Florence-London, Florence-Paris, and Bruges-London when prices are measured using a basket and only Florence-Barcelona, Florence-Paris and London-Bruges when prices are measured using wheat. The final rows in Panels A and B both show the speed of adjustment parameters from the error correction models for the cases where cointegration holds. Given the way that we have run the second stage regressions, a positive sign would be expected and the parameter estimates are mostly around typically of the order of 0.4-0.6, indicating that around half of any disequilibrium is corrected within a year. Adjustment is particularly quick in the case of the Florence series (Florence-Barcelona, Florence-London and Florence-Paris with highly significant parameter estimates of 0.62, 0.72 and 0.68 respectively) when the basket is used. These results indicating the efficiency of the market are consistent with the half-life figures reported above.

6. Panel Tests on Groups of Exchange Rates

Altogether, the lack of consensus in particular in the wheat case as presented in the previous section motivates a consideration of a panel approach which may shed light on whether the individual series results arise from a lack of information in the sample, notably for the rates with shorter available sample periods. Further, as Engel (2000) notes, it is possible that the unit root tests will be over-sized and the KPSS tests lacking power, jointly leading to a spurious conclusion that PPP holds. In the context of the historical period and available data, we believe that, whilst noting their limitations, the panel approach seems appropriate as the increase in power from the use of several series in combination justifies the additional complexity and concerns of Banerjee, Marcellino and Osbat (2005). The panel we employ, spanning almost 30 years in some cases, at least implies that structural breaks are unlikely to be an issue while holding considerably more information than a univariate series of similar length. From a historical perspective, the use of panels also allows us to test for multi-lateral arbitrage between groups of cities rather than just bilateral arbitrage between two cities.

Within the more modern panel PPP testing literature, it is standard to group the exchange rates in the sample by base currency (e.g., a set of rates all denominated in US dollars) but in our case, as described above, we have a number of cross-rates and also multiple trading venues. We therefore employ several groups of real exchange rates that will each constitute a separate panel and test for PPP: an English sterling group traded outside of England (Bruges-London, Florence-London); a Barcelona *lira* group traded outside of Barcelona (Bruges-Barcelona, Florence-Barcelona); a florin group traded in Florence (Florence-Barcelona, Florence-Bruges, Florence-London, Florence-Paris); and two groups for the Flemish *pond groot*, one traded in Bruges (Bruges-Barcelona, Bruges-London, Bruges-Paris) and the other traded outside of Bruges (Barcelona-Bruges, Florence-Bruges, London-Bruges, Paris-Bruges). The latter two groups will allow us to explore the effects of one of the peculiarities of the medieval FX market, at least to modern eyes, whereby the same exchange rate was quoted at different rates at either end of the currency pair.

For each group, we employ a battery of panel approaches which we now describe closely summarising Brooks (2014). The first is the test due to Levin, Lin and Chu (2002, hereafter LLC) which is a simple extension of the augmented Dickey-Fuller test

$$\Delta y_{i,t} = \alpha_i + \theta_t + \delta_i t + \rho_i y_{i,t-1} + \sum \gamma_j \Delta y_{i,t-j} + v_{i,t}, \quad t=1, 2, \dots, T; \quad i = 1, 2, \dots, N \quad (3)$$

As for the Dickey-Fuller tests, the results focus on the ρ_i parameter. LLC assumes a common root under the alternative hypothesis so that $\rho_i \equiv \rho = 0 \forall i$ under the null hypothesis and $\rho < 0$ under the alternative (the so-called “homogeneity assumption”). In the results presented below, we employ only the intercepts, although the results are not qualitatively altered if deterministic trends or time-fixed effects are also included.

The second test we employ is due to Breitung (2000), who develops a modified version of the LLC test which standardises the residuals from the auxiliary regression in a more sophisticated fashion.

Under the LLC and Breitung approaches, only evidence against the non-stationary null in one series is required before the joint null will be rejected. Breitung and Pesaran (2008) suggest that the appropriate conclusion when the null is rejected is that “a significant proportion of the cross-sectional units are stationary.” The homogeneity assumption may also be invalid since there is no theory suggesting that all of the series have the same autoregressive dynamics and thus the same value of ρ . This difficulty led Im, Pesaran and Shin (2003, hereafter IPS) to propose an alternative approach where, given equation (3) as above, the null and alternative hypotheses are now $H_0: \rho_i = 0 \forall i$ and $H_1: \rho_i < 0, i = 1, 2, \dots, N_1; \rho_i = 0, i = N_1+1, N_1+2, \dots, N$.⁷ As noted above, the Monte Carlo study by Banerjee, Marcellino, and Osbat (2005) highlights the possibility of size distortions in the panel unit root tests but with the LLC test being the least affected.

Finally, Maddala and Wu (1999) and Choi (2001) develop a slight variant on the IPS approach based on an idea dating back to Fisher (1932), where unit root tests are again conducted separately on each series in the panel, and the p -values associated with the test statistics are then combined. If we call these p -values $pv_i, i = 1, 2, \dots, N$, then under the null hypothesis of a unit root in each series, the pv_i will be distributed uniformly over the $[0,1]$ interval and hence the following will hold for given N as $T \rightarrow \infty$

$$\lambda = -2 \sum_{i=1}^N \ln(pv_i) \sim \chi_{2N}^2 \quad (4)$$

The results from the various panel approaches are reported in Table 5. Comparing these findings with those from Tables 3 and 4, it is clear that they are now much more consistent across testing approaches and that there is now considerably more evidence in favour of PPP holding. For the basket prices in Panel A, PPP is found for the sterling group, the group traded in Florence, and the *pond groot* group traded outside of Bruges. For the sterling group, the p -values for the five tests are all less than 0.5%, indicating very strong evidence in favour of mean-reversion – the same conclusion as for the individual series but to an even greater extent. The test statistics also show rejections at the 1% level in all cases for the *pond groot* group traded outside of Bruges and for the currencies traded in Florence, albeit slightly less strongly than the sterling group. No support for PPP is found within the Barcelona *lira* group, with none of the five testing approaches showing rejections of the random walk null hypothesis even at the 10% level in these instances. This corroborates the conclusion from the tests on the individual component series and again highlights that Spain may not have been particularly well-integrated into wider European markets. The results for the group

⁷ We also consider the Hadri (2000) test, which is the panel analogy to the KPSS test, having stationarity under the null. However, it has been argued that the test suffers from very severe size distortions when the series under consideration are stationary but heavily autocorrelated or heavily cross-correlated (see, for example, Demetrescu, Hassler and Tarcolea, 2010, for a discussion of these issues), which might well be the case here and we thus elect not to use this approach.

traded at Bruges present a conflicting picture between the tests, with the LLC and related Breitung approaches not rejecting the non-stationary null while the IPS test rejects at the 5% level and the two Fisher statistics all show rejections at below the 1% level. We would thus tentatively conclude in this case that PPP does not hold. This is a very interesting result, given that PPP appears to hold for the *pond groot* group traded outside of Bruges.

When wheat is used as the price measure, PPP is found to hold for all series under all tests, with rejections at well below the 1% level in most cases except for the sterling and Bruges groups when Breitung is used. In this case the results are more conclusive than those on the component series and very strongly supportive of the law of one price holding broadly with regard to wheat. This result would seem to suggest that the markets for particular commodities may have been more integrated than others. Thus, PPP holds for wheat even in cases where it does not for a broader basket of goods, indicating that some of the components of such baskets (possibly perishables like dairy or bulky/low-value goods like firewood) may not have been so widely-traded.

Finally, for robustness Table 6 repeats the analysis of Table 5 for the three groups where Paris was previously included. This additional analysis is undertaken in order to investigate the sensitivity to incorporation of the Paris series where a prices basket is not available. For all three groups, the main findings from Table 5 are unaltered. Thus for the group traded in Florence and the *Pond groot* group, strong evidence in favour of PPP holding still applies while there is much less support for PPP in the case of the group traded in Bruges.

7. Conclusions

This study has employed a range of approaches based on unit root and stationarity testing in order to determine whether the law of one price with respect to wheat and/or purchasing power parity held for a unique panel of medieval real exchange rates. When individual series are employed, the results are mixed but still around half of the real exchange rate series are best characterised as stationary, mean-reverting processes. However, when a panel approach is used, the weight of evidence falls decisively in favour of PPP holding, with at least some support for it from four of the five groupings investigated. This finding echoes the corresponding results of Abuaf and Jorion (1990), Lothian and Taylor (1996), Lothian and Devereux (2011) and numerous others for the more recent period. Certainly the support for PPP is no weaker here than in such recent work, and this requires us to carefully reconsider whether the features that we believed were necessary for the price-setting mechanism to function effectively across international borders are really so. It appears that abaci, commercial intuition and handwritten letters delivered by couriers were sufficient and that computers, pricing models, telephones, and railways are not the pre-requisite they may seem today.

Our research therefore contributes to the emerging consensus that markets in the later Middle Ages were surprisingly well integrated and that the process of globalisation of markets had already begun. This includes research into the efficiency of the wool forward market in medieval England (Bell, Brooks and Dryburgh, 2007) and the domestic exchange market in

Florence (Booth and Umit, 2008). Equally, we must be careful not to overstress the development of medieval markets. Notably, there is significantly less evidence for PPP holding for east-west relationships such as Barcelona and Paris/Florence and Paris and Bruges than for the north-south axis between London, Bruges and Florence. This suggests that further historical research into how the financial and trading connections between these places differed would be fruitful. Moreover, the rise of markets was not an inexorable or inevitable progress, and indeed, the Middle Ages were followed by a period of dis-integration during the later sixteenth and seventeenth centuries.

Our results also have implications for the mechanism by which adjustment will take place to correct deviations from purchasing power parity. The theory suggests that if such deviations exist, market forces will intervene to remove them; however, the theory does not indicate whether adjustment will occur through changes in exchange rates or in actual prices. Medieval exchange rates were fairly fixed around their intrinsic metallic values (Volckart and Wolf, 2006), which implies that the bulk of any adjustment must have fallen on the prices of goods (Froot, Kim and Rogoff, 1995), a setup which may also occur for some pegged currencies today. This situation contrasts with that applying for the majority of modern economies which have floating exchange rates and paper currencies, where adjustment would be expected through the exchange rate rather than to the sticky prices of actual goods. In fact, Bergin, Glick and Wu (2012) have suggested that the half-life of shocks was shorter during the Bretton Woods regime of relatively fixed exchange rates than it has been since 1973. This tallies with our finding that PPP still held in the medieval context, despite the narrower range of channels available to correct disequilibria.

References

- Abuaf, N. and Jorion, P. (1990) Purchasing Power Parity in the Long Run *Journal of Finance* 45(1), 157-174.
- Alder, M. and Lehman, B. (1983) Deviations from Purchasing Power Parity in the Long Run *Journal of Finance* 38(5), 147-187.
- Allen, R.C. (2001) The Great Divergence in European Wages and Prices from the Middle Ages to the First World War *Explorations in Economic History* 38, 411-447.
- Álvarez-Nogal, C. and de la Escosura, L.P. (2013) The Rise and Fall of Spain (1270-1850) *Economic History Review* NS 66(1), 1-37.
- Argilés, C. (2010) *Una Ciutat Catalana de Època de Crisi: Lleida, 1358-1500 : el Treball, els Salaris, la Producció Agrícola i els Preus a Través dels Llibres d'Obra de la Seu Vella* Institut d'Estudis Ilerdencs, Lleida.
- Banerjee, A., Marcellino M. and Osbat C. (2005) Testing for PPP: should we use panel methods? *Empirical Economics* 30, 77-91.
- Bateman, V. (2011) The Evolution of Markets in Early Modern Europe, 1350-1800: A Study of Wheat Prices *Economic History Review* NS 64(2), 447-471.
- Baulant, M. (1968) Le Prix des Grains à Paris de 1431 à 1788 *Annales Histories, Sciences Sociales* 23(3), 520-540.
- Beer, M. (1938) *Early British Economics from the XIIIth to the Middle of the XVIIIth Century* G. Allen & Unwin, London.
- Bell, A.R., Brooks, C. and Dryburgh, P. (2007) Interest Rates and Efficiency in Medieval Wool Forward Contracts *Journal of Banking and Finance* 31(2), 361-80.
- Bell, A.R, Brooks, C. and Moore, T.K. (2009) Interest in Medieval Accounts: Examples from England, 1272-1340, *History* 94, 411-33.
- Bell, A.R., Brooks, C. and Moore, T.K. (2013) Medieval Foreign Exchange: A Time Series Analysis, in Casson, M. and Hashimzade, N. (eds.) *Large Databases in Economic History: Research Methods and Case Studies* Routledge, London, 97-123.
- Bell, A.R., Brooks, C. and Moore, T.K. (2016) The 'Buying and Selling of Money for Time': Foreign Exchange and Interest Rates in Medieval Europe, *Economic History Review*, forthcoming
- Bergin, P., Glick, R. and Wu, J.-L. (2012) 'Mussa Redux and Conditional PPP' *NBER Working Paper* No.18331
- Bernholz, P. and Kugler, P. (2011) Financial Market Integration in the Early Modern Period in Spain: Results from a Threshold Error Correction Model *Economics Letters* 110, 93-96.

- Beveridge, W. (1939) *Prices and Wages in England From the Twelfth to the Nineteenth Centuries: Volume I* Longmans, Green and co., London.
- Bignon, V., Chen, J. and Ugolini, S (2013) Beneath the Gold Points: Financial Market Integration in Europe, 1820-1870 *Working Paper*, Banque de France.
- Boerner, L. and Volckart, O. (2011) The Utility of a Common Coinage: Currency Unions and the Integration of Money Markets in Late Medieval Central Europe *Explorations in Economic History* 48(1), 53-65.
- Booth, G.G. and Umit, U.G. (2008) Volatility Clustering and the Bid-ask Spread: Exchange Rate Behaviour in Early Renaissance Florence *Journal of Empirical Finance* 15, 131-144.
- Breitung, J. (2000) The Local Power of Some Unit Root Tests for Panel Data in Baltagi, B. (ed.) *Nonstationary Panels, Panel Cointegration and Dynamic Panels* Advances in Econometrics 15, 161-178, JAI Press, Amsterdam.
- Breitung, J. and Pesaran, M.H. (2008) Unit Roots and Cointegration in Panels, in Matyas, L. and Sevestre, P. (eds.) *The Econometrics of Panel Data*, third edition, Springer-Verlag, Berlin.
- Brooks, C. (2014) *Introductory Econometrics for Finance*, 3rd edition, Cambridge University Press, Cambridge.
- Chilosi, D. and Volckart, O. (2011) Money, States and Empire: Financial Integration and Institutional Change in Central Europe, 1400-1520 *Journal of Economic History* 71(3), 762-791.
- Choi, I. (2001) Unit Root Tests for Panel Data *Journal of International Money and Finance* 20, 249-272.
- Clark, G. (2014) Markets Before Economic Growth? The Grain Market of Medieval England *Cliometrica* 9(3), 265-287.
- Coakley, J. and Fuertes, A-M. (1997) New Panel Unit Root Tests of PPP *Economics Letters* 57, 17-22.
- Cuddington, J.T. and Liang, H. (2000) Purchasing Power Parity over Two Centuries? *Journal of International Money and Finance* 19, 753-757.
- Demetrescu, M., Hassler, U. and Tarcolea, A.I. (2010) Testing for Stationarity in Large Panels with Cross-dependence, and US evidence on Unit Labor Cost *Journal of Applied Statistics* 37(8), 1381-1397.
- Dickey, D.A. and Fuller, W.A. (1979) Distribution of Estimators for Time Series Regressions with a Unit Root *Journal of the American Statistical Association* 74, 427-431
- Dornbusch, R. (1985) Purchasing Power Parity *NBER Working Paper* No.1591.

- Edison, H.J. (1987) Purchasing Power Parity in the Long Run: A Test of the Dollar/Pound Exchange Rate (1890-1978) *Journal of Money, Credit and Banking* 19(3), 376-387.
- Engel, C. (2000) Long-run PPP May not Hold after All *Journal of International Economics* 51(2), 243-273.
- Epstein, S.R. (2000) *Freedom and Growth: the Rise of States and Markets in Europe, 1350-1700* Routledge, London.
- Federico, G. (2012) What do we Know about Market Integration in Europe? *Economic History Review* 65(2), 470-97.
- Fisher, R.A. (1932) *Statistical Methods for Research Workers* fourth edition, Oliver & Boyd, Edinburgh.
- Frankel, J.A. (1986) International Capital Mobility and Crowding Out in the US Economy: Imperfect Integration of Financial Markets or Goods Markets? In Hafer, R.W. (ed.) *How Open is the US Economy?* Lexington Books, Lexington.
- Frankel, J.A. and Rose, A. (1996) A Panel Project on Purchasing Power Parity: Mean-reversion Within Countries and Between Countries *Journal of International Economics* 40, 209-224.
- Froot, K.A., Kim, M. and Rogoff, K. (1995) The Law of One Price over 700 Years *NBER Working Paper* No. 5132.
- Fuller, W.A. (1976) *Introduction to Statistical Time Series* Wiley, New York.
- Galloway, J.A. (2000) One Market or Many? London and the Grain Trade of England, in Galloway, J.A. (ed.), *Trade, Urban Hinterlands and Market Integration c.1300-1600* Centre for Metropolitan History, Institute of Historical Research, London, 23-42.
- Glen, J.D. (1992) Real Exchange Rates in the Short, Medium and Long Run *Journal of International Economics* 33(1/2), 147-166.
- Hadri, K. (2000) Testing for Stationarity in Heterogeneous Panel Data *Econometrics Journal* 3, 148-161.
- Hybel N. (2002) The Grain Trade in Northern Europe before 1350 *Economic History Review* NS 55(2), 219-47.
- Im, K.S., Pesaran, M.H. and Shin, Y. (2003) Testing for Unit Roots in Heterogeneous Panels *Journal of Econometrics* 115, 53-74.
- Jacks, D. (2004) Market Integration in the North and Baltic Seas, 1500–1800 *Journal of European Economic History* 33, 285–329.
- Kugler, P. (2011) Financial Market Integration in Late Medieval Europe: Results from a Threshold Error Correction Model for the Rhinegulden and Basle Pound, 1365-1429 *Swiss Journal of Economics and Statistics* 147(3), 337-357.

Kugler, P. (2013) The Changing Pattern of Arbitrage between the Amsterdam and London Foreign Exchange Market 1600–1912 *Working Paper*, WWZ/Universität Basel.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., and Shin, Y. (1992) Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root *Journal of Econometrics* 54, 159-178.

Levin, A., Lin, C., and Chu, C. (2002) Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties *Journal of Econometrics* 108, 1-24.

Leybourne, S.J., Mills, T.C., and Newbold, P. (1998) Spurious Rejections by Dickey-Fuller Tests in the Presence of a Break under the Null *Journal of Econometrics* 87, 191-203.

Li, L.-F. (2012) Bullion, Bills and Arbitrage: Exchange Markets in Fourteenth- to Seventeenth Century Europe, Unpublished London School of Economics PhD thesis.

Lothian, J.R. and Devereux, J. (2011) Exchange Rates and Prices in the Netherlands and Britain over the Past Four Centuries *Working Paper*, Bank of Greece No. 135.

Lothian, J.R. and Taylor, M.P. (1996) Real Exchange Rate Behavior: The Recent Float from the Perspective of the Past Two Centuries *Journal of Political Economy* 104(3), 488-509.

MacDonald, R. (1995) Long Run Exchange Rate Modelling: A Survey of Recent Evidence *IMF Staff Papers* 42, 437-489.

Maddala, G.S. and Wu, S. (1999) A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test *Oxford Bulletin of Economics and Statistics* 61, 631-652.

Malanima, P. (2013) When did England Overtake Italy? Medieval and Early Modern Divergence in Prices and Wages *European Review of Economic History* 17, 45-70.

Meese, R.A. and Rogoff, K.S. (1988) Was it Real? The Exchange Rate Interest Differential over the Modern Floating Period *Journal of Finance* 43(4), 933-948.

Mestayer, M. (1963) Les Prix du Ble et de l'Avoine a Douai de 1329 a 1793 *Revue du Nord* 45, 167-76.

Melis, F. (1962) *Aspetti della Vita Economica Medievale: Studi nell'Archivio Datini di Prato Monte dei Paschi di Siena*, Siena.

Melis, F. (1985) Intensità e Regolarità nella Diffusione dell'Informazione Economica Generale nel Mediterraneo e in Occidente alla Fine del Medioevo in L. Frangioni (ed.) *I Trasporti e le Comunicazioni nel Medioevo* Le Monnier, Florence, 185-93.

Miskimin, H.A. (1963) *Money, Prices, and Foreign Exchange in Fourteenth-Century France* Yale University Press, New Haven.

Miskimin, H.A. (1984) *Money and Power in Fifteenth-Century France* Yale University Press, New Haven.

- Munro, J.H.A. (2003) Wage Stickiness, Monetary Changes, and Real Incomes in Late-Medieval England and the Low Countries, 1300-1500: Did Money Matter? *Research in Economic History* 21, 185-297.
- Munro, J.H.A. (2012) The Technology and Economics of Coinage Debasements in Medieval and Early Modern Europe: With Special Reference to the Low Countries and England, in Munro, J.H.A. (ed.) *Money in the Pre-Industrial World: Bullion, Debasements and Coin Substitutes* Pickering and Chatto, London.
- Murray, J.M. (2005) *Bruges: Cradle of Capitalism, 1280-1390* Cambridge University Press, Cambridge.
- Origo, I.M. (1963) *The Merchant of Prato: Francesco di Marco Datini, 1335-1410* revised edition, Penguin, Harmondsworth.
- Pedroni, P. (2001) Purchasing Power Parity Tests in Cointegrated Panels *Review of Economics and Statistics* 83, 727-731.
- Perron, P. (1989) The Great Crash, the Oil Price Shock and the Unit Root Hypothesis *Econometrica* 57, 1361-1401.
- Perron, P. (1991) Test Consistency with Varying Sampling Frequency *Econometric Theory* 7, 341-368.
- Persson, K. G. (1999) *Grain markets in Europe 1500-1900: Integration and Deregulation* Cambridge University Press, Cambridge.
- Pinto, G. (1993) I Livelli di Vita dei Salarinati Fiorentini (1380–1430), in *Toscana Medievale. Paesaggi e Realtà Sociali* Le Lettere, Florence, 113–52.
- Rogoff, K. (1996) The Purchasing Power Parity Puzzle *Journal of Economic Literature* 34, 647-668.
- Roll, R. (1979) Violations of Purchasing Power Parity and their Implications for Efficient International Commodity Markets, in Sarnat, M. and Szego, G. (eds.) *International Finance and Trade*, Ballinger, Cambridge, MA.
- de Roover, R.A. (1968) *The Bruges Money Market Around 1400 (with a Statistical Supplement by Hyman Sardy)* Paleis der Academiën, Brussels.
- Söderberg, J. (2006) Grain Prices in Cairo and Europe in the Middle Ages *Research in Economic History* 24, 189-216.
- Sosvilla-Rivero, S. and García, E. (2003) Purchasing Power Parity Revisited *FEDEA Working Paper* 2003-20.
- Spufford, P. (1986) *Handbook of Medieval Exchange* Royal Historical Society, London.

Spufford, P. (2000) Trade in Fourteenth-Century Europe *New Cambridge Medieval History Volume VI: c.1300-c.1415* Cambridge University Press, Cambridge, 155-208.

Spufford, P. (2012) Debasement of the Coinage and its Effects on Exchange Rates and the Economy: in England in the 1540s, and in the Burgundian-Habsburg Netherlands in the 1480s, in Munro, J.H.A (ed.) *Money in the Pre-Industrial World: Bullion, Debasements and Coin Substitutes* Pickering and Chatto, London.

Steigerwald, D.G. (1996) Purchasing Power Parity, Unit Roots and Dynamic Structure *Journal of Empirical Finance* 2, 343-357.

Taylor, A.M. and Taylor, M.P. (2004) The Purchasing Power Parity Debate *Journal of Economic Perspectives* 18, 135-158.

Taylor, M.P. (1988) An Empirical Examination of Long-run Purchasing Power Parity using Cointegration Techniques *Applied Economics* 20(10), 1369-1381.

Unger, R.W. (2007) Thresholds for Market Integration in the Low Countries and England in the Fifteenth Century, in Armstrong, L., Elbl I., and Elbl, M. M. (eds.) *Money, Markets and Trade in Medieval Europe: Essays in Honour of John H.A. Munro* Brill, Leiden, 349-380.

Van der Wee, H. (1963) *The Growth of the Antwerp Market and the European Economy, Fourteenth-Sixteenth Centuries* 3 volumes, Martinus Nijhoff, The Hague.

Van de Wee, H. (1993) Prices and Wages as Development Variables: A Comparison Between England and the Southern Netherlands, 1400-1700 in *The Low Countries in the Early Modern World* Cambridge University Press and Variorum, Cambridge and New York, 223-41.

Volckart, O. and Wolf, N. (2006) Estimating Financial Integration in the Middle Ages: What Can We Learn from a TAR Model? *Journal of Economic History* 66, 122-39.

Xu, Z. (2003) Purchasing Power Parity, Price Indices and Exchange Rate Forecasts *Journal of International Money and Finance* 22, 105-130.

Table 1: Summary Information for Real Exchange Rate Series

	Barcelona- Bruges	Bruges- Barcelona	Bruges- London	Bruges- Paris	Florence- Barcelona	Florence- Bruges	Florence- London	Florence- Paris	London- Bruges	Paris- Bruges
Location of Trading	Barcelona	Bruges	Bruges	Bruges	Florence	Florence	Florence	Florence	London	Paris
Currency Units	d of Barcelona per écu of 22d <i>groot</i>	d of Barcelona per écu of 22d <i>groot</i>	d sterling per écu of 24d <i>groot</i>	d <i>groot</i> per <i>livre</i> <i>tournois</i>	d of Barcelona per florin	d <i>groot</i> per florin	d sterling per florin	<i>Livre</i> <i>tournois</i> per florin	d sterling per écu of 24d <i>groot</i>	d <i>groot</i> per <i>livre</i> <i>tournois</i>
Number of FX observations	700	904	742	820	2147	2577	1103	2196	175	729
Year Range	1383-1411	1391-1411	1391-1411	1391-1411	1383-1411	1383-1411	1394-1411	1384-1411	1392-1406	1384-1410
Number of Annual Observations	29	21	21	21	29	29	18	28	15	23

Table 2: *Usance* (in days) between key financial centres

City From	City To				
	Florence	Barcelona	Paris	Bruges	London
Florence		d60	d60	d60	d90
Barcelona	d60		N/A	s30 (53)	N/A
Paris	d60	N/A		s10/12 (14-16)	N/A
Bruges	d60	s30 (53)	d30		d30
London	N/A	N/A	N/A	d30	

Source: Melis 1984. d indicates the bill was payable a set number of days after drawing, and s after sight (the figure in brackets includes the average postal time).

Table 3: Unit Root and Stationarity Tests on Real Exchange Rates 1383-1411

	Barcelona- Bruges	Bruges- Barcelona	Bruges- London	Bruges- Paris	Florence- Barcelona	Florence- Bruges	Florence- London	Florence- Paris	London- Bruges	Paris- Bruges
<i>Panel A: Prices based on a Basket</i>										
DF/ADF statistic	-1.417	-1.593	-3.087**	-2.109	-1.273	-3.958***	-3.389**	-3.498**	-2.336	-3.934***
DF/ADF 5% critical value	-2.966	-3.011	-3.004	-3.004	-2.966	-2.966	-3.040	-2.971	-3.082	-2.985
Half-life (in years)	4.1	3.34	1.0	1.4	5.4	0.9	0.9	1.9	1.0	3.8
KPSS statistic	0.571**	0.546**	0.134	0.093	0.564**	0.210	0.090	0.196	0.193	0.435*
<i>Panel B: Prices based on Wheat</i>										
DF/ADF statistic	-2.704*	-3.381**	-2.846*	-2.541	-3.435**	-3.590***	-3.497**	-2.642	-5.388***	-2.704*
DF/ADF 5% critical value	-3.011	-3.004	-3.004	-2.966	-2.966	-3.040	-2.971	-3.148	-2.985	-3.011
Half-life (years)	1.2	1.0	0.7	1.6	1.1	0.8	1.5	0.4	0.6	1.2
KPSS statistic	0.398*	0.190	0.179	0.187	0.478**	0.085	0.291	0.189	0.172	0.398*

Notes: *, ** and *** denote significance at the 10%, 5% and 1% levels respectively; the KPSS critical values are 0.347, 0.463, and 0.739 at the 10%, 5% and 1% levels respectively; for the price series in France, no basket figures are available and so wheat prices were used as a proxy.

Table 4: Engle-Granger Cointegration Tests on Real Exchange Rates 1383-1411

	Barcelona- Bruges	Bruges- Barcelona	Bruges- London	Bruges- Paris	Florence- Barcelona	Florence- Bruges	Florence- London	Florence- Paris	London- Bruges	Paris- Bruges
<i>Panel A: Prices based on a Basket</i>										
DF/ADF statistic	-2.671*	-2.563	-1.0428	-1.795	-3.368**	-3.586**	-3.995***	-3.948***	-3.985***	-1.933
Cointegration?	No	No	No	No	Yes	Yes	Yes	Yes	Yes	No
Speed of adjustment	0.378**	-	-	-	0.618***	0.404**	0.720**	0.677***	0.204*	-
<i>Panel B: Prices based on Wheat</i>										
DF/ADF statistic	-2.113	-2.491	-2.702	-2.356	-3.069**	-2.638	-2.536	-3.412**	-3.112**	-1.693
Cointegration?	No	No	No	No	Yes	No	No	Yes	Yes	No
Speed of adjustment	-	-	-	-	0.587***	-	-	0.594***	0.180**	-

Notes: *, ** and *** denote significance at the 10%, 5% and 1% levels respectively; speed of adjustment parameters are not presented in cases where there is no cointegration

Table 5: Panel Unit Root Tests on Real Exchange Rates 1383-1411

Group	Sterling group	Barcelona lira group	Group traded in Florence	Group traded in Bruges	Pond groot group
Component exchange rates	Bruges-London, Florence-London	Bruges-Barcelona, Florence-Barcelona	Florence-Barcelona, Florence-Bruges, Florence- London, Florence-Paris	Bruges-Barcelona, Bruges-London, Bruges- Paris	Barcelona-Bruges, Florence-Bruges, London-Bruges, Paris-Bruges
<i>Panel A: Prices based on a Basket</i>					
Levin, Lin and Chu t -statistic	-2.786 (0.003)***	0.398 (0.655)	-2.165 (0.015)**	-0.913 (0.181)	-2.331 (0.010)***
Breitung t -statistic	-3.505 (0.000)***	-1.229 (0.110)	-1.753 (0.040)**	-0.287 (0.387)	-2.898 (0.002)***
Im, Pesaran and Shin W -statistic	-3.332 (0.000)***	1.224 (0.889)	-3.292 (0.000)***	-2.097 (0.018)**	-2.893 (0.002)***
ADF Fisher χ^2 statistic	17.527 (0.000)***	0.899 (0.925)	28.175 (0.000)***	18.282 (0.006)***	26.187 (0.001)***
Phillips-Perron Fisher χ^2 statistic	17.501 (0.000)***	0.1837 (0.933)	28.102 (0.000)***	18.231 (0.006)***	26.927 (0.001)***
<i>Panel B: Prices based on Wheat</i>					
Levin, Lin and Chu t -statistic	-2.866 (0.002)***	-2.274 (0.012)**	-3.036 (0.001)***	-4.706 (0.000)***	-3.781 (0.000)***
Breitung t -statistic	-0.626 (0.266)	-2.470 (0.007)***	-2.822 (0.002)***	-0.185 (0.426)	-2.790 (0.003)***
Im, Pesaran and Shin W -statistic	-2.809 (0.000)***	-2.195 (0.014)**	-3.291 (0.001)***	-4.114 (0.000)***	-3.332 (0.000)***
ADF Fisher χ^2 statistic	14.631 (0.006)***	11.588 (0.021)**	24.504 (0.002)***	26.956 (0.000)***	25.507 (0.001)***
Phillips-Perron Fisher χ^2 statistic	14.631 (0.006)***	11.776 (0.019)**	23.937 (0.002)***	30.793 (0.000)***	24.432 (0.002)***

Notes: The tests include an intercept but no trend, except the Breitung test which includes both an intercept and a trend. p -values in parentheses; *, ** and *** denote significance at the 10%, 5% and 1% levels respectively; for the price series in France, no basket figures are available and so wheat prices were used as a proxy.

Table 6: Panel Unit Root Tests on Real Exchange Rates Excluding France 1383-1411

Group	Group traded in Florence	Group traded in Bruges	<i>Pond groot</i> group
Component exchange rates	Florence-Barcelona, Florence-Bruges, Florence- London	Bruges-Barcelona, Bruges-London,	Barcelona-Bruges, Florence-Bruges, London-Bruges
<i>Panel A: Prices based on a Basket</i>			
Levin, Lin and Chu t -statistic	-0.198 (0.422)	-0.456 (0.324)	-1.222 (0.111)
Breitung t -statistic	-1.185 (0.118)	-1.943 (0.026)**	-1.897 (0.029)**
Im, Pesaran and Shin W -statistic	-1.692 (0.045)**	-1.019 (0.154)	-2.095 (0.018)**
ADF Fisher χ^2 statistic	13.246 (0.039)**	10.157 (0.038)**	15.441 (0.017)**
Phillips-Perron Fisher χ^2 statistic	14.247 (0.027)**	10.131 (0.038)**	20.412 (0.002)***
<i>Panel B: Prices based on Wheat</i>			
Levin, Lin and Chu t -statistic	-3.131 (0.001)***	-2.932 (0.002)***	-4.419 (0.000)***
Breitung t -statistic	-2.646 (0.004)***	-1.008 (0.157)	-2.671 (0.004)***
Im, Pesaran and Shin W -statistic	-3.250 (0.001)***	-2.335 (0.010)**	-3.551 (0.000)***
ADF Fisher χ^2 statistic	20.785 (0.002)***	12.270 (0.016)**	22.992 (0.001)***
Phillips-Perron Fisher χ^2 statistic	20.594 (0.002)***	12.184 (0.016)**	22.682 (0.001)***

Notes: The tests include an intercept but no trend, except the Breitung test which includes both an intercept and a trend. p -values in parentheses; *, ** and *** denote significance at the 10%, 5% and 1% levels respectively.