Relative Prices

- Some Definitions
- Real Exchange Rates: The PPP Puzzle
  - PPP - Validity, Econometric Testing, Non-Linearities/Heterogeneity vs. Price Discrimination
  - Border Effects
- Nominal Exchange Rates:
  - Forecastability, Beating the Random Walk
  - Interest Parities - Covered, Uncovered. The Forward Discount Puzzle. Risk Premia, Expectations, Limited Rationality
Defiition

- **Nominal Exchange Rate** $E$: relative price of currencies in two countries. E.g. Euro/Dollar

- The number quoted in Financial Times and The Economist.

- Technically, this is the rate at which bank deposits in different currencies are exchanged. Exchange rates for coins or notes slightly different as they include transport costs.

- Bilateral exchange rates are cross-rates, telling the price of one currency relative to another. 1 Euro = 1.25 Dollar

- Effective exchange rates are weighted averages of all bilateral exchange rates. Weights are chosen to reflect trade weights, i.e. share of total trade accounted for by particular country. No obvious unit, so normalize to 100 in an arbitrary year.

- Useful to evaluate overall value of a currency over time.
The Real Exchange Rate $\varepsilon$ captures the cost of purchasing the same bundle of goods in different countries. Connected to "competitiveness".

Define

$$\varepsilon = \frac{E \cdot P^*}{P}$$

where: $P$ denotes the domestic price of a bundle of goods, $P^*$ is its foreign counterpart, and $E$ is labelled in *units of domestic currency per unit of foreign* (e.g. £/$)

An *increase* in $\varepsilon$ means a gain in domestic competitiveness according to that measure. A *real depreciation*.

Note that an *increase* in $E$ means a *nominal depreciation* of the domestic currency relative to the foreign currency.
PPP stands for Purchasing Power Parity. Conjectures there are arbitrage forces that bring back to parity international relatives prices of (all?) goods.

Formally, the simplest form of PPP imposes

$$\varepsilon = 1$$

Bundles of goods cost the same across borders. If they do not, buy where cheap to sell where expensive - brings back parity.

This is also the strongest form of parity, or "Absolute PPP". Requires instantaneous arbitrage, and infinite amounts of resources taking advantage of price differences.

Related to "Law of One Price" (LOOP), which imposes parity at the individual's good level. Note that Absolute PPP is stronger than LOOP, since it requires LOOP holds for all goods and the bundles of goods have the same composition across borders. Not true with home bias and/or non-traded goods.

Quite obvious LOOP or PPP do not hold at every point in time.
Long Run PPP: $/£
In the long run PPP appears to hold but can be long term deviations in exchange rates from their PPP value

Long Run PPP: the Franc

Fig. 2.—Logarithm of franc-sterling real exchange rate, 1805–1990 (1900 = 0; increase = sterling appreciation).
In fact, Absolute PPP is very difficult to test. To have a meaningful measure of $\varepsilon$, need meaningful measures on the LEVELS of $P$ and $P^*$

But $P$ and $P^*$ are typically price indices, such as CPI or PPI

Price indices are normalized to 100 (e.g.) on a given year. Their levels are typically meaningless.

The graphs just shown conveyed the general notion that relative prices do not explode out of proportion, and there are forces preventing explosive price differences in the long run. But they do not say to what LEVELS relative prices converge back to.

For that, need data on price LEVELS. Not that much available: International Price Comparisons (UN, every 5 years), Scanner’s data (but proprietary and incomplete), Economist Intelligence Unit (but again incomplete)
Long run data on relative prices display some tendency to reversal. What could be behind that? Suppose domestic inflation very high - much higher than foreign one, and over a long time..

ε would fall following a long run trend - a real appreciation. Unless $E$ adjusts. How? A nominal depreciation will offset the increase in prices in domestic currency.

Relative PPP implies high inflation economies see their nominal exchange rate increase, i.e. depreciate. There is mean reversion in the RER.

That can be explored econometrically rigorously. Assume an autoregressive process (of order one) drives real exchange rates, and let $q_t$ denote $\ln \varepsilon_t$:

$$q_t = \alpha + \beta q_{t-1} + \eta_t$$

$\beta < 1$ implies there is some reversion to the mean. What mean? Solve for a time-invariant level of $q$:

$$q = \frac{\alpha}{1 - \beta}$$
Low Power: it is difficult to distinguish with certainty values of $\beta$ that are very close to 1 from $\beta$ effectively taking value 1.

Unfortunately, the difference in economic terms is central. $\beta = 0.99$ still means some mean reversion, if slow, and thus some predictability.

Way out: in the 90s, increase time dimension (as in previous graphs). Or use panel approaches, i.e. introduce an international dimension. Estimation becomes

$$q_{it} = \alpha_i + \beta q_{it-1} + \eta_{it}$$

with $i$ denoting a certain country's exchange rate. Note that $\alpha_i$ now controls for the (unobserved) level of international price differences, which may be different for different countries. E.g. France-Germany as compared with Japan-Sierra Leone. It is also controlling for the arbitrary normalizations in price indices.

Consensus finding: quite some certainty that $\beta < 1$. BUT it is so close to 1 that implied rate of mean reversion unconscionably slow.
In particular, so-called "half-life" of a given disturbance to international relative prices between 3-5 years. In an AR(1) case, half-life $H$ verifies

$$\beta^H = \frac{1}{2} \Rightarrow H = \frac{-\ln 2}{\ln \beta}$$

No closed form for AR(p). Typically use simulated impulse-response function.

3-5 years to take advantage of international price differences???? Rather dense arbitragists...

First element of "PPP Puzzle". Possible to rationalize in model with price stickiness (i.e. prices do not change for a while). But then (i) required calibrated stickiness way too long (no more than 3-4 quarters in the data), and (ii) impossible to reproduce the RER volatility we observe in the data.
The PPP Puzzle

Three explanations:

- Time Aggregation: frequency of observation is inadequate to observe arbitrage
- Non-Linearities: there are inaction bands, for instance because of trade costs
- Heterogeneity: some goods are easy to arbitrage, others are not.
Suppose arbitrage occurs within weeks, but you only have (say) quarterly data, which is an AVERAGE of higher frequency data.

No problem if frequency of observation is different - problem is (moving) averaging.

Then the true model is

$$q_t = \alpha + \beta q_{t-1} + \eta_t$$

but we only observe

$$q_t^* = \frac{1}{P} \left( q_s + q_{s+1} + ... + q_{s+P-1} \right)$$

where $P$ is the period of sampling. For instance, there are $P = 12$ weeks in a quarter.
The dynamics of $q_t^*$ are fundamentally different. For instance, suppose we estimate

$$q_t^* = \alpha^* + \beta^* q_{t-1}^* + \eta_t^*$$

Then, OLS estimator of $\beta^*$ given by

$$\hat{\beta}^* = \frac{\text{cov}(q_t^*, q_{t-1}^*)}{\text{var}q_{t-1}^*}$$

and the corresponding half-life

$$H^* = \frac{P \ln \frac{1}{2}}{\ln \beta^*}$$

Now we seek to establish that $H^* > H$ and $\frac{H^*}{H}$ increases with $P$ and $H^*$ explodes as $P \to \infty$. 
Now

\[ \hat{\beta}^* = \frac{\text{cov} \left( q_s + q_{s+1} + \ldots + q_{s+P-1}, q_{s-P} + q_{s-P+1} + \ldots + q_{s-1} \right)}{\text{var} \left( q_s + q_{s+1} + \ldots + q_{s+P-1} \right)} \]

\[ = \frac{n(P)}{d(P)} \]

We know the true process is

\[ q_t = \alpha + \beta q_{t-1} + \eta_t \]

so that

\[ \text{cov} \left( q_t, q_{t-k} \right) = \beta^k \text{var} q \]
Define

\[ N(P) = \frac{n(P)}{\text{varq}} = \beta + \beta^2 + \ldots + \beta^P \]
\[ + \beta^2 + \beta^3 + \ldots + \beta^{P+1} \]
\[ + \ldots \]
\[ + \beta^P + \beta^{P+1} + \ldots + \beta^{2P-1} \]
\[ = \beta \left( 1 + \beta + \beta^2 + \ldots + \beta^{P-1} \right)^2 \]
\[ = \beta \left( \frac{1 - \beta^P}{1 - \beta} \right)^2 \]
Time Aggregation

• And

\[
D(P) = \frac{d(p)}{varq} = D(P - 1) + 1 + 2\beta + \ldots + 2\beta^{P-1}
\]

\[
= D(P - 1) + 2\frac{1 - \beta^P}{1 - \beta} - 1
\]

since \( var(A + B + C) = var(A) + var(B + C) + 2cov(A, B + C) \).

• Notice \( D(1) = 1 \) and solve recursively to get

\[
D(P) = \frac{P(1 - \beta^2) - 2\beta \left(1 - \beta^P\right)}{(1 - \beta)^2}
\]

• So finally,

\[
\beta^* = \frac{\beta \left(1 - \beta^P\right)^2}{P \left(1 - \beta^2\right) - 2\beta \left(1 - \beta^P\right)}
\]
The (biased) half life is given by

\[ H^* = \frac{P \ln 2}{\ln \frac{P (1-\beta^2) - 2\beta (1-\beta^P)}{\beta (1-\beta^P)^2}} \]

The true half life, in turn, verifies

\[ H = \frac{-\ln 2}{\ln \beta} \Rightarrow \beta = 2^{-\frac{1}{H}} \]

Replace:

\[ H^* = \frac{P \ln 2}{\ln \frac{P \left(1-2^{-\frac{2}{H}}\right) - 2^{1-\frac{1}{H}} \left(1-2^{-\frac{P}{H}}\right)}{2^{-\frac{1}{H}} \left(1-2^{-\frac{P}{H}}\right)^2}} \]
Finally:

\[
\frac{H^*}{H} = \frac{-P \ln \beta}{\ln \frac{P(1-\beta^2)-2\beta(1-\beta^P)}{\beta(1-\beta^P)^2}}
\]

It is easy to show that

\[
\lim_{P \to \infty} \frac{H^*}{H} = \lim_{P \to \infty} \frac{-P \ln \beta}{\ln \frac{P(1-\beta^2)}{\beta}} = \lim_{P \to \infty} \frac{-P \ln \beta}{\ln P} = \infty
\]
The gap between $\beta^*$ and $\beta$ increases in $P$, i.e. a bias becomes larger as the gap between observed and actual frequency increases. In other words, trying to identify daily convergence using monthly (yearly!) data is particularly biased.

It also increases in $\beta$: a positive bias is larger for very persistent processes. In other words, if true mean reversion is somewhat slow, it will appear even slower if the frequency of the data is not adequate.
### Table 1: Temporal-Aggregation Bias for Half-Life Estimates

<table>
<thead>
<tr>
<th>temporal aggregation ($P$)</th>
<th>true half-life ($H$, days)</th>
<th>estimated half-life ($H^*$, days)</th>
<th>bias factor ($H^*/H$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>1</td>
<td>1.0</td>
<td>2.0</td>
<td>7.0</td>
</tr>
<tr>
<td>2</td>
<td>1.4</td>
<td>2.7</td>
<td>9.4</td>
</tr>
<tr>
<td>7</td>
<td>2.2</td>
<td>3.7</td>
<td>11.0</td>
</tr>
<tr>
<td>14</td>
<td>3.3</td>
<td>4.7</td>
<td>11.8</td>
</tr>
<tr>
<td>30</td>
<td>5.5</td>
<td>7.0</td>
<td>13.8</td>
</tr>
<tr>
<td>90</td>
<td>12.8</td>
<td>15.1</td>
<td>22.6</td>
</tr>
<tr>
<td>180</td>
<td>22.3</td>
<td>25.8</td>
<td>35.5</td>
</tr>
<tr>
<td>365</td>
<td>40.1</td>
<td>45.6</td>
<td>59.5</td>
</tr>
<tr>
<td>730</td>
<td>72.3</td>
<td>81.1</td>
<td>102.0</td>
</tr>
</tbody>
</table>

|                           | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| 2                         | 1.41 | 1.37 | 1.34 | 1.34 | 1.34 | 1.33 | 1.33 | 1.33 | 1.33 |
| 7                         | 2.25 | 1.83 | 1.57 | 1.53 | 1.50 | 1.49 | 1.49 | 1.49 | 1.49 |
| 14                        | 3.30 | 2.33 | 1.69 | 1.59 | 1.54 | 1.51 | 1.50 | 1.50 | 1.50 |
| 30                        | 5.53 | 3.52 | 1.97 | 1.71 | 1.59 | 1.53 | 1.51 | 1.51 | 1.50 |
| 90                        | 12.76| 7.57 | 3.22 | 2.27 | 1.81 | 1.59 | 1.54 | 1.52 | 1.51 |
| 180                       | 22.32|12.91 | 5.07 | 3.23 | 2.21 | 1.70 | 1.59 | 1.54 | 1.52 |
| 365                       | 40.15|22.81 | 8.50 | 5.12 | 3.12 | 1.95 | 1.70 | 1.59 | 1.54 |
| 730                       | 72.32|40.53 |14.57 |8.50 | 4.88 | 2.52 | 1.95 | 1.70 | 1.59 |

**Notes:** See text. Given $P$ and $H$, $H^*$ is calculated from (7).
There are "bands of inaction". Arbitrage will only occur provided price differences are high enough (to cover for transport costs).

This can explain permanent international differences in prices. Can it also explain slow convergence?

Consider the true model

\[
q_t = \begin{cases} 
\alpha + \beta (q_{t-1} - \alpha) + \eta_t & \text{if } q_{t-1} > \alpha \\
q_{t-1} + \eta_t & \text{if } -\alpha \leq q_{t-1} \leq \alpha \\
-\alpha + \beta (q_{t-1} + \alpha) + \eta_t & \text{if } q_{t-1} < -\alpha 
\end{cases}
\]

Here \( \alpha \) captures the transport costs: price deviations have to exceed that amount for reversal to occur (with speed captured by \( \beta \)). Rest of the time, real exchange rate is effectively *not* reverting to its mean.
Suppose you ignore this non-linearity and estimate a linear setup instead

\[ q_t^{**} = \alpha^{**} + \beta^{**} q_{t-1}^{**} + \eta_t^{**} \]

How is that going to affect your estimates of \( \beta \)?

The OLS estimate of \( \beta^{**} \) will be given by

\[
\hat{\beta}^{**} = \frac{\text{cov} (q_t^{**}, q_{t-1}^{**})}{\text{var} q_{t-1}^{**}} = \frac{E (q_t^{**} q_{t-1}^{**} I_{IN}(q_{t-1}^{**})) + E (q_t^{**} q_{t-1}^{**} I_{OUT}(q_{t-1}^{**}))}{Eq_{t-1}^{**2}}
\]

Define

\[
\theta_{IN} = EI_{IN}(q_{t-1}^{**}) \frac{\text{var} (q_{t-1}^{**} I_{IN}(q_{t-1}^{**}))}{\text{var} q_{t-1}^{**}}
\]

\[
\theta_{OUT} = EI_{OUT}(q_{t-1}^{**}) \frac{\text{var} (q_{t-1}^{**} I_{OUT}(q_{t-1}^{**}))}{\text{var} q_{t-1}^{**}}
\]

Clearly, \( \theta_{IN} + \theta_{OUT} = 1 \)
And we have

\[
\hat{\beta}^{**} = EI_{IN}(q_{t-1}^{**}) \frac{E(q_t^{**} q_{t-1}^{**} l_{IN}(q_{t-1})}{E q_{t-1}^{**}} \\
+ EI_{OUT}(q_{t-1}^{**}) \frac{E(q_t^{**} q_{t-1}^{**} l_{OUT}(q_{t-1}))}{E q_{t-1}^{**}} \\
= \theta_{IN} \frac{E(q_t^{**} q_{t-1}^{**} l_{IN}(q_{t-1}))}{E(q_{t-1}^{**} l_{IN}(q_{t-1}))} \\
+ \theta_{OUT} \frac{E(q_t^{**} q_{t-1}^{**} l_{OUT}(q_{t-1}))}{E(q_{t-1}^{**} l_{OUT}(q_{t-1}))} \\
= \theta_{IN} \cdot 1 + \theta_{OUT} \cdot \beta
\]

Naive linear estimation is a weighted average of persistence within the band (1) and persistence outside of it (\(\beta\)). The weights \(\theta\) are given by the proportion of time spent in (\(\theta_{IN}\)) and outside of the band (\(\theta_{OUT}\)). The bias will be larger the more time is spent within the band, i.e. the larger \(\alpha\).
Non Linearities

- In terms of half life:

\[
H^* = \frac{-\ln 2}{\ln \hat{\beta}^{**}} = \frac{-\ln 2}{\ln \hat{\beta}} \frac{\ln \hat{\beta}}{\ln \hat{\beta}^{**}} = H \frac{\ln \beta}{\ln (\beta + \theta_{IN}(1 - \beta))}
\]
## Table 2: Linear-Specification Bias for Half-Life Estimates

<table>
<thead>
<tr>
<th>threshold parameter ( (k) )</th>
<th>true half-life ( (H, \text{ days}) )</th>
<th>( 1 )</th>
<th>( 2 )</th>
<th>( 7 )</th>
<th>( 14 )</th>
<th>( 30 )</th>
<th>( 90 )</th>
<th>( 180 )</th>
<th>( 365 )</th>
<th>( 730 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 0.00 )</td>
<td>fraction of observations outside the band</td>
<td>( 1.00 )</td>
<td>( 1.00 )</td>
<td>( 1.00 )</td>
<td>( 1.00 )</td>
<td>( 1.00 )</td>
<td>( 1.00 )</td>
<td>( 1.00 )</td>
<td>( 1.00 )</td>
<td>( 1.00 )</td>
</tr>
<tr>
<td>( 0.10 )</td>
<td></td>
<td>( 0.93 )</td>
<td>( 0.94 )</td>
<td>( 0.97 )</td>
<td>( 0.98 )</td>
<td>( 0.98 )</td>
<td>( 0.99 )</td>
<td>( 0.99 )</td>
<td>( 1.00 )</td>
<td>( 1.00 )</td>
</tr>
<tr>
<td>( 0.50 )</td>
<td></td>
<td>( 0.71 )</td>
<td>( 0.77 )</td>
<td>( 0.85 )</td>
<td>( 0.89 )</td>
<td>( 0.92 )</td>
<td>( 0.97 )</td>
<td>( 0.97 )</td>
<td>( 0.98 )</td>
<td>( 0.98 )</td>
</tr>
<tr>
<td>( 1.00 )</td>
<td></td>
<td>( 0.54 )</td>
<td>( 0.61 )</td>
<td>( 0.73 )</td>
<td>( 0.80 )</td>
<td>( 0.83 )</td>
<td>( 0.92 )</td>
<td>( 0.92 )</td>
<td>( 0.96 )</td>
<td>( 0.99 )</td>
</tr>
<tr>
<td>( 2.00 )</td>
<td></td>
<td>( 0.35 )</td>
<td>( 0.42 )</td>
<td>( 0.60 )</td>
<td>( 0.67 )</td>
<td>( 0.72 )</td>
<td>( 0.82 )</td>
<td>( 0.90 )</td>
<td>( 0.89 )</td>
<td>( 0.90 )</td>
</tr>
<tr>
<td>( 5.00 )</td>
<td></td>
<td>( 0.18 )</td>
<td>( 0.23 )</td>
<td>( 0.38 )</td>
<td>( 0.47 )</td>
<td>( 0.59 )</td>
<td>( 0.73 )</td>
<td>( 0.73 )</td>
<td>( 0.79 )</td>
<td>( 0.83 )</td>
</tr>
<tr>
<td>( 0.00 )</td>
<td>estimated half-life ( (H^*, \text{ days}) )</td>
<td>( 1.0 )</td>
<td>( 2.0 )</td>
<td>( 7.0 )</td>
<td>( 14.0 )</td>
<td>( 30.0 )</td>
<td>( 90.0 )</td>
<td>( 180.0 )</td>
<td>( 365.0 )</td>
<td>( 730.0 )</td>
</tr>
<tr>
<td>( 0.10 )</td>
<td></td>
<td>( 1.1 )</td>
<td>( 2.1 )</td>
<td>( 7.2 )</td>
<td>( 14.3 )</td>
<td>( 30.6 )</td>
<td>( 90.6 )</td>
<td>( 181.3 )</td>
<td>( 366.8 )</td>
<td>( 733.3 )</td>
</tr>
<tr>
<td>( 0.50 )</td>
<td></td>
<td>( 1.6 )</td>
<td>( 2.7 )</td>
<td>( 8.3 )</td>
<td>( 15.8 )</td>
<td>( 32.6 )</td>
<td>( 94.0 )</td>
<td>( 185.9 )</td>
<td>( 376.6 )</td>
<td>( 744.6 )</td>
</tr>
<tr>
<td>( 1.00 )</td>
<td></td>
<td>( 2.2 )</td>
<td>( 3.5 )</td>
<td>( 9.7 )</td>
<td>( 17.6 )</td>
<td>( 36.3 )</td>
<td>( 97.9 )</td>
<td>( 196.0 )</td>
<td>( 381.4 )</td>
<td>( 739.5 )</td>
</tr>
<tr>
<td>( 2.00 )</td>
<td></td>
<td>( 3.6 )</td>
<td>( 5.3 )</td>
<td>( 11.9 )</td>
<td>( 21.2 )</td>
<td>( 41.7 )</td>
<td>( 109.4 )</td>
<td>( 199.6 )</td>
<td>( 411.5 )</td>
<td>( 813.3 )</td>
</tr>
<tr>
<td>( 5.00 )</td>
<td></td>
<td>( 7.2 )</td>
<td>( 10.0 )</td>
<td>( 19.1 )</td>
<td>( 30.1 )</td>
<td>( 51.4 )</td>
<td>( 123.9 )</td>
<td>( 245.4 )</td>
<td>( 459.4 )</td>
<td>( 880.5 )</td>
</tr>
<tr>
<td>( 0.00 )</td>
<td>bias factor ( (H^*/H) )</td>
<td>( 1.00 )</td>
<td>( 1.00 )</td>
<td>( 1.00 )</td>
<td>( 1.00 )</td>
<td>( 1.00 )</td>
<td>( 1.00 )</td>
<td>( 1.00 )</td>
<td>( 1.00 )</td>
<td>( 1.00 )</td>
</tr>
<tr>
<td>( 0.10 )</td>
<td></td>
<td>( 1.11 )</td>
<td>( 1.07 )</td>
<td>( 1.03 )</td>
<td>( 1.02 )</td>
<td>( 1.02 )</td>
<td>( 1.01 )</td>
<td>( 1.01 )</td>
<td>( 1.00 )</td>
<td>( 1.00 )</td>
</tr>
<tr>
<td>( 0.50 )</td>
<td></td>
<td>( 1.57 )</td>
<td>( 1.36 )</td>
<td>( 1.19 )</td>
<td>( 1.13 )</td>
<td>( 1.09 )</td>
<td>( 1.04 )</td>
<td>( 1.03 )</td>
<td>( 1.03 )</td>
<td>( 1.02 )</td>
</tr>
<tr>
<td>( 1.00 )</td>
<td></td>
<td>( 2.19 )</td>
<td>( 1.75 )</td>
<td>( 1.38 )</td>
<td>( 1.25 )</td>
<td>( 1.21 )</td>
<td>( 1.09 )</td>
<td>( 1.09 )</td>
<td>( 1.04 )</td>
<td>( 1.01 )</td>
</tr>
<tr>
<td>( 2.00 )</td>
<td></td>
<td>( 3.56 )</td>
<td>( 2.65 )</td>
<td>( 1.70 )</td>
<td>( 1.51 )</td>
<td>( 1.39 )</td>
<td>( 1.22 )</td>
<td>( 1.11 )</td>
<td>( 1.13 )</td>
<td>( 1.11 )</td>
</tr>
<tr>
<td>( 5.00 )</td>
<td></td>
<td>( 7.20 )</td>
<td>( 4.99 )</td>
<td>( 2.73 )</td>
<td>( 2.15 )</td>
<td>( 1.71 )</td>
<td>( 1.38 )</td>
<td>( 1.36 )</td>
<td>( 1.26 )</td>
<td>( 1.21 )</td>
</tr>
</tbody>
</table>

Notes: We generated 10,000 observations of the TAR process (8) for given \( \rho = 2^{-1/H}, \alpha = 1, \) and \( k = c. \) We then estimated the standard model (1) on this generated data. We derived \( H^* \) from the estimated \( \rho \) and compared it with the true \( H. \)
### Table 3: Combined Bias for Half-Life Estimates

<table>
<thead>
<tr>
<th>temporal aggregation (P)</th>
<th>threshold parameter (k)</th>
<th>true half-life (H, days)</th>
<th>bias factor (H^*/H)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>0.00</td>
<td>1.39</td>
<td>1.37</td>
</tr>
<tr>
<td></td>
<td>0.10</td>
<td>1.57</td>
<td>1.45</td>
</tr>
<tr>
<td></td>
<td>0.50</td>
<td>2.27</td>
<td>1.92</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>3.52</td>
<td>2.51</td>
</tr>
<tr>
<td></td>
<td>2.00</td>
<td>6.67</td>
<td>4.59</td>
</tr>
<tr>
<td></td>
<td>5.00</td>
<td>23.67</td>
<td>13.79</td>
</tr>
<tr>
<td>7</td>
<td>0.00</td>
<td>2.24</td>
<td>1.75</td>
</tr>
<tr>
<td></td>
<td>0.10</td>
<td>2.38</td>
<td>1.94</td>
</tr>
<tr>
<td></td>
<td>0.50</td>
<td>3.15</td>
<td>2.36</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>4.47</td>
<td>3.27</td>
</tr>
<tr>
<td></td>
<td>2.00</td>
<td>8.29</td>
<td>5.37</td>
</tr>
<tr>
<td></td>
<td>5.00</td>
<td>31.38</td>
<td>16.99</td>
</tr>
<tr>
<td>14</td>
<td>0.00</td>
<td>3.40</td>
<td>2.31</td>
</tr>
<tr>
<td></td>
<td>0.10</td>
<td>3.30</td>
<td>2.63</td>
</tr>
<tr>
<td></td>
<td>0.50</td>
<td>4.35</td>
<td>2.90</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>5.32</td>
<td>3.50</td>
</tr>
<tr>
<td></td>
<td>2.00</td>
<td>9.33</td>
<td>6.11</td>
</tr>
<tr>
<td></td>
<td>5.00</td>
<td>31.08</td>
<td>18.17</td>
</tr>
<tr>
<td>30</td>
<td>0.00</td>
<td>5.25</td>
<td>3.38</td>
</tr>
<tr>
<td></td>
<td>0.10</td>
<td>4.84</td>
<td>3.90</td>
</tr>
<tr>
<td></td>
<td>0.50</td>
<td>6.79</td>
<td>4.52</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>7.09</td>
<td>5.00</td>
</tr>
<tr>
<td></td>
<td>2.00</td>
<td>11.44</td>
<td>7.30</td>
</tr>
<tr>
<td></td>
<td>5.00</td>
<td>34.37</td>
<td>18.89</td>
</tr>
<tr>
<td>90</td>
<td>0.00</td>
<td>7.14</td>
<td>7.23</td>
</tr>
<tr>
<td></td>
<td>0.10</td>
<td>10.60</td>
<td>8.55</td>
</tr>
<tr>
<td></td>
<td>0.50</td>
<td>16.32</td>
<td>8.57</td>
</tr>
<tr>
<td></td>
<td>1.00</td>
<td>17.72</td>
<td>9.12</td>
</tr>
<tr>
<td></td>
<td>2.00</td>
<td>21.66</td>
<td>12.20</td>
</tr>
<tr>
<td></td>
<td>5.00</td>
<td>43.39</td>
<td>23.14</td>
</tr>
</tbody>
</table>

*Notes:* See text and Table 2. Note that, in principle, the results for \( k = 0 \) should correspond exactly to those in Table 2. However, these results are generated by simulation in finite samples and not from an exact formula. Thus, there are small discrepancies between the two tables, especially for larger \( P \), where aggregation means a shorter span in the finite-sample averaged data.
Some goods are easier to arbitrage than others. (gold or wheat vs. bed mattresses or houses)

Speed of mean reversion should be allowed to differ, as in

\[ q_{st} = \alpha_s + \beta_s q_{st-1} + \eta_{st} \]

where \( s \) denotes a good. Key is that \( \beta_s \) differs for different goods.

Suppose now \( \beta_s = \beta + \varepsilon_s \), i.e. each good has a speed of mean reversion that differs additively from the "true" average \( \beta \).

Then a naive estimation that ignores this heterogeneity rewrites (even if it allows for FE):

\[ q_{st} = \alpha_s + \beta q_{st-1} + (\eta_{st} + \varepsilon_s \cdot q_{st-1}) \]

Basic econometrics tells us this estimation will be biased, because its residuals \( \eta_{st} + \varepsilon_s \cdot q_{st-1} \) unavoidably depends on the regressor.

Instrumentation will not solve anything, since an instrument must be correlated with the regressor \( q_{st-1} \), but not with the residuals \( \eta_{st} + \varepsilon_s \cdot q_{st-1} \). That is impossible.
Can say something about the direction of the bias. Suppose $\beta_{FE}$ is the fixed effects estimator. Can show that

$$p_{N \to \infty, T \to \infty} \beta_{FE} = \beta + \Delta$$

$$\Delta = \sum_{i=1}^{N} (\beta_s - \beta) \delta_s$$

$$\delta_s = \frac{\sigma^2_s}{1 - \beta_s^2} / \sum_{i=1}^{N} \left( \frac{\sigma^2_s}{1 - \beta_s^2} \right)$$

Necessary and sufficient condition for a positive bias is $\text{cov} (\tilde{\beta}, \tilde{\delta})$.

$\Delta$ also increases in $\beta_s - \beta$, i.e. the cross-sectional dispersion in persistence.
The same is true of *aggregate* regression in the presence of heterogeneity.

Consider country-level regression

\[
q_{st} = \alpha_s + \beta_s q_{st-1} + \eta_{st}
\]

\[
q_t = \sum_s \omega_s q_{st}
\]

and \( \omega \) are (CPI) expenditures weights.

Note that in reality, if \( q_{st} \) follows an AR(1), the aggregate RER verifies

\[
\left[ \prod_{s=1}^{N} (1 - \beta_s L) \right] q_t = \sum_{s=1}^{N} \omega_s \left[ \prod_{s=1}^{N} (1 - \beta_s L) \right] \eta_{st}
\]

I.e an ARMA process - potentially of infinite order. Any attempt to account for heterogeneity estimating that will be an approximation. Dead end.
Heterogeneity

- Now estimating

$$q_t = \alpha + \beta q_{t-1} + \eta_t$$

will imply $\beta = \sum_s \omega_s \beta_s$ and $\eta_t = \sum_s \omega_s \eta_{st} + \sum_s \omega_s q_{st-1}$. So that the same pathology prevails.

- Can sign that bias again. Consider the OLS estimate obtained from aggregate data $\beta_{OLS}$. By definition,

$$\beta_{OLS} = \frac{Eqt q_{t-1}}{Eq^2_t}$$

- Substitute disaggregated components of numerator and denominator, to obtain:

$$\beta_{OLS} = \beta + \Delta'$$

$$\Delta' = \sum_{i=1}^{N} (\beta_s - \beta) \gamma_s$$
Heterogeneity

And we have

\[ \gamma_s = \frac{\omega_s \sigma_s^2}{1 - \beta_s^2} + \sum_{s \neq t} \frac{\omega_s \omega_t \sigma_{st}}{1 - \beta_s \beta_t} \]

As before, bias positive for positive values of \( \text{cov}(\tilde{\beta}, \tilde{\gamma}) \).

Importantly, this is not *necessarily* requiring that high \( \beta \) sectors have high CPI expenditures weights.
Figure I

Check on Conditions for a Positive Aggregation Bias
### TABLE II

Persistence Estimates using Aggregate Data

\[ q_{ct} = \gamma_c + \sum_{p=1}^{P} \beta_p q_{ct-p} + e_t \]

<table>
<thead>
<tr>
<th>Model</th>
<th>P</th>
<th>( \sum_{p=1}^{P} \beta_p )</th>
<th>Half-Life</th>
<th>LAR</th>
<th>CIR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed Effects</td>
<td>18</td>
<td>0.98</td>
<td>46</td>
<td>0.97</td>
<td>64.38</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(31.57)</td>
<td>(0.962, 0.981)</td>
<td></td>
</tr>
<tr>
<td>Anderson-Hsiao</td>
<td>11</td>
<td>0.99</td>
<td>72</td>
<td>0.96</td>
<td>109.68</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(33, \infty)</td>
<td>(0.941, 1.05)</td>
<td></td>
</tr>
<tr>
<td>Arrelano Bond</td>
<td>18</td>
<td>0.99</td>
<td>54</td>
<td>0.98</td>
<td>75.57</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(46.75)</td>
<td>(0.975, 0.989)</td>
<td></td>
</tr>
</tbody>
</table>

\(^{a}H_0 : \rho_c = \rho\) \(-0.4046\) (1.0000) \(^{c}H_0 : E(\gamma_c, X) = 0\) \(25.856\) (0.0021)

\(^{b}H_0 : \rho_c = \rho\) \(70.96\) (0.9999) \(^{d}H_0 : \gamma_c = 0\) \(9.8714\) (0.000)

The estimates are based on real exchange rates from 11 countries over the period 1981:01-1995:12. The choice of P is based on general to specific lag selection procedure with a maximum lag of 20 for all models, except AH where it was restricted to 12. At each choice of P, the impulse response was examined and the specification was only selected if the IRF was continuous around 0.5. For the GMM estimator two lags of the levels of relative prices were used as instruments. The confidence intervals in the parentheses were estimated using non-parametric bootstrap with 500 replications. Note that the bootstrap for the Arellano and Bond estimator was carried out using the methods described in Brown and Newey [2001]. “LAR” denotes the largest autoregressive root. “CIR” denotes the cumulated impulse response. “a” is the Hausman test for homogeneity, while “b” denotes the Swami test for this hypothesis. “c” and “d” are the Hausman test for random effects and an F-test for fixed effects, respectively.
### TABLE III

**Persistence Estimates using Disaggregated Data**

\[ q_{it} = \gamma_c + \sum_{k=1}^{K} \rho_{tk} q_{it-k} + \varepsilon_{it} \]

<table>
<thead>
<tr>
<th>Model</th>
<th>P</th>
<th>( \sum_{k=1}^{K} \rho_{tk} )</th>
<th>Half-Life</th>
<th>LAR</th>
<th>CIR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed Effects</td>
<td>12</td>
<td>0.98</td>
<td>36</td>
<td>0.97</td>
<td>46.71</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(21.47)</td>
<td>(0.961, 0.981)</td>
<td></td>
</tr>
<tr>
<td>Fixed Effects (SURE)</td>
<td>12</td>
<td>0.98</td>
<td>34</td>
<td>0.97</td>
<td>44.30</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(27.43)</td>
<td>(0.958, 0.978)</td>
<td></td>
</tr>
<tr>
<td>Fixed Effects (CCE)</td>
<td>12</td>
<td>0.99</td>
<td>58</td>
<td>0.99</td>
<td>104.20</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(10.91)</td>
<td>(0.980, 0.995)</td>
<td></td>
</tr>
<tr>
<td>Mean Group</td>
<td>19</td>
<td>0.97</td>
<td>26</td>
<td>0.95</td>
<td>33.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(14.28)</td>
<td>(0.903, 0.973)</td>
<td></td>
</tr>
<tr>
<td>Mean Group (SURE)</td>
<td>20</td>
<td>0.96</td>
<td>22</td>
<td>0.96</td>
<td>29.48</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(17.27)</td>
<td>(0.945, 0.968)</td>
<td></td>
</tr>
<tr>
<td>Mean Group (CCE)</td>
<td>12</td>
<td>0.95</td>
<td>11</td>
<td>0.95</td>
<td>20.51</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(7.12)</td>
<td>(0.924, 0.963)</td>
<td></td>
</tr>
</tbody>
</table>

\( aH0 : \rho_i = \rho \)

| 98.15               | 4353.4          | 485.02          | 2194698  |
| (0.0000)            | (0.0007)        | (0.0022)        | (0.0000) |

\( bH0 : \rho_i = \rho \)

\( dH0 : E(\gamma_c, X) = 0 \)

\( eH0 : \gamma_c = 0 \)

\( fH0 : E(\eta_i, X) = 0 \)

\( fLM \)

The estimates are based on relative prices on a maximum of 19 goods from 11 countries over the period 1981:01-1995:12. The choice of P is based on general to specific lag selection procedure with a maximum lag of 20 for all models, except AH where it was restricted to 12. At each choice of P, the impulse response was examined and the specification was only selected if the IRF was continuous around 0.5. The confidence intervals in the parenthesis were estimated using non-parametric bootstrap with 500 replications. "LAR" denotes the largest autoregressive root. "CIR" denotes the cumulated impulse response. "a" is the Hausman test for homogeneity (allowing for correlated residuals), while "b" denotes the Swami test for this hypothesis. "c" is the Pudney [1978] test for the null of no correlation between the random coefficients and the error term. "d" and "e" are the Hausman test for random effects and an F-test for fixed effects, respectively, while "f" is a Breusch-Pagan test for the diagonality of the covariance matrix.
### Table IV

**Persistence Estimates using Disaggregated Data (Bias Corrected)**

\[ q_{tet} = \gamma_t + \sum_{k=1}^K \rho_{tk} \hat{q}_{tet-k} + \epsilon_{tet} \]

<table>
<thead>
<tr>
<th>Model</th>
<th>P</th>
<th>Half-Life (Indirect)</th>
<th>Half Life (Direct)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Group</td>
<td>19</td>
<td>41</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(17, 64)</td>
<td>(17, 57)</td>
</tr>
<tr>
<td>Mean Group (SURE)</td>
<td>5</td>
<td>43</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(18, 105)</td>
<td>(16, 65)</td>
</tr>
<tr>
<td>Mean Group (CCE)</td>
<td>5</td>
<td>20</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(11, 28)</td>
<td>(11, 28)</td>
</tr>
</tbody>
</table>

The Bias Correction is carried out via the Kilian [1998] bootstrap method. "Indirect" refers to a method where $\rho$ is corrected and the half life is estimated on the basis of $\rho^*$. In the "direct" case, the half-life is corrected directly. In each case, the bootstrap uses 500 replications. For the Mean group model $N=204$, and $T=1981:01$ to 1995:12. For the other two models the cross section in Chen and Engel [2004] is used to ensure non-singularity of covariance matrices. In addition, the time series is restricted to 1981:06 to 1994:09 in order to produce a balanced panel. This helps to decrease computation time and has little impact on the underlying (uncorrected) estimates. The confidence intervals are calculated via a double bootstrap procedure. That is, at each replication bootstrap samples are drawn using the mean estimates from the models in the table and the generated data is used to estimate the models via Kilian's bootstrap using 100 replications (50 for Mean Group (SURE)). This is repeated 100 times (50 for Mean Group (SURE)) and the 95% confidence intervals are calculated.
Well known pathology. The way out is to estimate sector (good?) specific $\beta_s$, and then aggregate these up to the country level.

How to aggregate them? Several estimators are available: Mean Group simply takes an arithmetic average, Random Coefficient takes a weighted average, where weights reflect accuracy of estimation of each coefficient.

Key insight is that aggregate estimates is fundamentally different from aggregating data. Reasoning on aggregated data implicitly assumes all $\beta$ are the same for all goods forming the Real Exchange Rate.

$\beta$ computed as a weighted average of $\beta_s$ is much smaller - at least on European data. In fact, so small that half-life of mean reversion down to 10 months. Compatible with plausible nominal rigidities.
So What to Make of PPP?

- These results are not uncontroversial. For instance, debates exist about the standard errors around these estimates.

- Two remaining puzzles:
  - The Border Effect: relative prices are much, much more volatile around borders, than within the same country.
  - The Traded vs Non Traded decomposition: it would make sense that prices be different between countries because of non-traded goods. A house in London is not a house in Tokyo or Dacca. Surprisingly however, virtually all of the persistence in international relative prices arises in (what is usually construed as) traded goods.

- Both puzzles point to the alternative view that there is "pricing to market", i.e. market power on the part of international firms, who price discriminate across different (international) markets.
The "Engel" Puzzle

- Consider a two-country model with both traded and non-traded goods. The (log) ideal price indices are:

\[ p_t = (1 - \alpha) p^T_t + \alpha p^N_t \]
\[ p^*_t = (1 - \beta) p^{*T}_t + \beta p^{*N}_t \]

- The RER is

\[ q_t = s_t + p^*_t - p_t \]
\[ = \left( s_t + p^{*T}_t - p^T_t \right) + \beta \left( p^{*N}_t - p^{*T}_t \right) - \alpha \left( p^N_t - p^T_t \right) \]

- The first term pertains to arbitrage, presumably.
- The second and third ones pertain to the within-country relative price of traded vs. non traded goods - and whether that difference is the same across countries.
- Engel performs this decomposition and finds the bulk of volatility in \( q \) does NOT come from the former.
Fig. 1.—MSE decomposition: CPI data. MSE ratio estimates and 95 percent confidence intervals: $\frac{\text{MSE}(x_{i+k} - x_i)}{[\text{MSE}(x_{i+k} - x_i) + \text{MSE}(y_{i+k} - y_i)]}$, January 1962 to December 1995. a, Canadian/U.S. real exchange rate. b, Japanese/U.S. real exchange rate. c, France, Germany, and Italy.
So What to Make of PPP?

- Most prominent approach is to compute the degree of "exchange rate pass-through", i.e. how much changes in the nominal exchange rate are passed into final prices. Imperfect pass-through requires some market power, since it means firm can choose its level of profit.

- Enormous empirical industry in current international macroeconomics. Seems to be some (very recent) evidence that all three refinements (time aggregation, non-linearities, heterogeneity) are present in estimates of border effects, traded vs. non-traded decomposition and exchange rate pass-through.

- Reversion to price parity still a powerful force in determining international relative prices. Especially in disaggregated (goods? sectors?) data.