We thank the referees for their detailed reading of our paper, and for highlighting a number of interesting issues. We believe that we would be able to take account of the referees' comments in a revision, and offer the following remarks as responses to the referees (and as ways in which we would modify the paper.)

Responses to Referee 1

We preface the responses to this referee with three general comments:

a)   all the suggestions about the re-writing issues (discussion/conclusion) can be taken into account, although we would favour a light rewriting of the paper, given that it is already quite long and the process of adding more detail (which many of the remarks call for) would lead to a yet more detailed paper. We can also streamline the paper a bit more in order to offset this, but without changing its current structure.

b)   concerning the issue of break dating - under the null hypothesis, i.e. no cointegration, there is no break. The use of the infimum operator is in order to make the test of the null hypothesis of no cointegration not dependent on the timing of the unknown break (that may exist, **but only under the alternative**.) Thus, as in Gregory and Hansen's much simpler framework, consistent estimation of the break date under the alternative is not needed to make consistent inference about cointegration (or no cointegration) under the null. However, the model must allow for the possibility of a break in order for the tests to have the correct properties. Gregory and Hansen have a detailed discussion of this issue, and we could summarize the arguments in a brief paragraph if this is felt to be helpful.

c)   much of the world trade is invoiced in dollars (though this does not guarantee dollar pricing - as the prices can be adjusted to local conditions). An interesting issue that we could try to tackle in a future paper is the "missing" type of pricing, i.e. not PCP, not LCP but USD-pricing (see Gopinath et al).

We move now to specific responses to the questions raised.

Question 1: The authors argue on page 5 that the standard estimation method in the literature is "single-equation autoregressive distributive lag models". This does not characterize the empirical literature appropriately. Following the work of Feenstra (1989) many papers in the literature have foresaken the distributed lag models for simpler models which impose less structure on the lags. This statement sets up the ARDL model as a sort of straw man, and should be rewritten.

We do not quite understand what is meant by 'foresaking the distributed lag models for simpler models which impose less structure on the lags'. That is, we do not understand the need for simpler models, or indeed what structure is imposed on the lags by following an ARDL framework. Perhaps this is a question of interpretation and we have misunderstood the comment. Although much of what we have read has this ARDL structure – including Feenstra (1989) and recent papers by Campa, Goldberg and co-authors - we could try to re-work this part of the discussion to be more representative of the literature.

Question 2: The discussion and conclusions should be split into two sections. At present, it makes for a rambling and unwieldy conclusion.

```
We could do this easily.
```

Question 3: Work by Bill Alterman has shown that the use of unit values significantly distorts pass-through coefficients in standard pass-through regressions. What are the implications of his work for the paper's findings? I suggest testing the same long-run cointegrated relationship for an industry in which you observe prices, as a robustness check.

```
This would imply writing a different paper.  In addition, we would lose the
benefits of a comparative analysis since it is very unlikely that similar data
would be available for all countries. Our paper is based on the IUVs, which is
not uncommon in the literature, and is very careful to mention the associated
drawbacks.  These drawbacks do not however change the fact that IUVs are among
the best data available, taking the various considerations into account - i.e.
they are the most widely available, with reasonable comparability, aggregation
and a decent time dimension).
```

Question 4 The authors' argument that the structural break they identify is not a data artifact should be elaborated.

```
We are certain that the existence of breaks is not an artefact of the method.

One may wonder about the negative correlation between breaks in constant and
slope. However, allowing for a break in both constant and slope does not bring
about such a result, the higher slope does not "offset" the lower intercept.

This is clear if one considers Table 5 and 6.
    • In Table 5, comparing column (1) and (2) reveals that the overall slope
      coefficient is not always higher when a break in intercept is allowed
      (36 cases out of 90 are constant or lower, with 6 in Germany and 5 in
      Finland, but 7 in Greece). The fact that the evidence is not
      overwhelming in either direction gives some comfort that the
      methodology is not flawed.
    • The negative intercept is robust across specifications, since it
      appears already in Table 5 (51 significant downward breaks in intercept
      out of 90) and does not emerge as a result from table 6. Indeed there
      are even fewer significant breaks in intercept (35 out of 90) when
      allowing for a break in both slope and intercept.
```

Question 5: The authors assert the superiority of their empirical approach without addressing some of the weaknesses of the structural break methods they employ. More discussion is needed of problems these techniques have: What if the timing of the structural break is wrong, or if there are multiple structural breaks over the sample period? The one sentence discussion of this issue on p. 30 is insufficient. A fuller development of structural-break models and how they work is needed.

```
We could mention the theoretical issues in a little more detail if the referee
wishes.  The framework we use (and to which we refer the reader) is Banerjee
```

and Carrion-i-Silvestre, 2006, and we felt that we had included as much information as was necessary. But it may indeed be a key issue to highlight the fact that a break is not present under the null, and that the infimum procedure adopted makes the procedure not to depend on the presence (or consistent estimation) of breaks under the alternative. Cointegration under the alternative could occur with or without breaks – the procedures developed need to be robust to all possibilities.

Multiple structural breaks could also be allowed for very easily – especially since trend changes in the cointegrating are not considered here (because they are not thought to be empirically relevant). The methods are robust to multiple changes in the constant, since upon differencing the data (and extracting the factors), breaks in constant behave like impulse dummies which do not have an impact upon the densities of the test statistics, while impulse dummies disappear upon differencing. Thus the same critical values (based on the null model, which do not have any breaks) would apply to the models which allow for multiple breaks.

Question 6: A longer discussion is needed on how this analysis will inform the debate and what needs to be done with the pass-through models that are currently being used by central banks, the IMF, and so on.

We adopt the same kind of specification as in the US where issues on trade deficit and responsiveness of prices, hence of volume to exchange rate, are important. A number of policy issues are already mentioned in the paper, including the optimal response of monetary policy to exchange rate, and the forecasting performance of models with and without changes in the long run. We are reluctant to push the policy conclusions too far, given the limitations with our data and the fact that theoretical models often have conclusions at variance with each other. In addition this would imply expanding the paper that the referees consider to be relatively long.

Question 7: The literature review, especially as it relates to microeconomic pass-through models, is thin. The discussion on pages 30 and 31 may want to tie this paper's results to previous work by Goldberg and Verboven (2001), Campa and Goldberg (2006), Goldberg and Tille (2006), and Gopinath, Itskhoki, and Rigobon (2007). Some discussion should be mentioned regarding the debate over whether pass-through has fallen in recent years. For example, the footnote that mentions the Federal Reserve Marazzi et al's specification including commodity prices should also mention the critique of this paper by Hellerstein, Daly, and Marsh (2006), which is consistent with this paper's results, as are some of the results in Campa and Goldberg (2006).

We thank the referee for pointing this out and agree that adding the reference to Hellerstein et al. (2006), Campa and Goldberg (2006) is a good idea. Indeed, CGM for the euro area, and Campa and Goldberg yield similar results to our paper. Our innovation is rather to reach the same conclusion with what we see as a more robust empirical procedure. The paper of Hellerstein et al is interesting since it challenges the idea of fall in PT. It mentions the role of intra-firm trade (that we cannot control for in our database, although our reading of the literature for the US indicates that evidence is mixed), as well as a commodity channel which creates a downward bias in the estimation of PT. Please see general comment for Referee 1.

Referee 2:

Question 1. The paper is very long and not very focused. In particular at times it seems as a comment on Campa and Gonzalez Minguez and several data sets are used. To my mind it would be more natural to either make it a comment on CM and send to EER, or to focus more squarely on your own results. It is hard to see what is the main punch line of the paper – there is a break in most series, cointegration holds or CM are wrong? I think you could more or less jump directly to section 6 without loosing much.

We see a number of difficulties not only with CM or CGM but, as we argue, with much of the empirical literature on pass through. However, the criticisms could be softened yet further as they are not the primary purpose of this paper, which is indeed to argue about the efficacy of our methodology and not to serve as comment on Campa and Gonzalez Minguez.

The main result of the paper is evidence of the existence of a properly estimated long run relationship once a more flexible testing procedure is allowed, as stated in the introduction. When we allow varying lag lengths in and introduce a break in the levels relationship, we find strong evidence for a theory-backed cointegration relationship. The benchmarking against CGM is only there to motivate the paper

Question 2. Writing as a comment on CM would be particularly appealing given the strong words – CMs techniques are described as "inappropriate" (p.5 for instance). Given the highly aggregated price indexes it is not obvious to me that we should expect a cointegrating relation. Detailed price studies in for instance Gopinath and Rigobon find that there is little pass-through, also in the long run for a class of very disaggregated prices. Thus if this line is taken I think it would be useful to motivate more why their technique is inappropriate.

See response to Question 1 above.

The problem with very disaggregated prices is two-fold. Firstly, there are very few data series available in terms of the number of countries, and the length of the series. Secondly, the frequency available at lower levels is rarely monthly or higher which may not matter so much for the long run per se, but if we intend to produce proper calculations of the long run with the intent of comparing these with appropriate calculations of the short run, it would be somewhat awkward to talk about short run effects of exchange rate changes with annual frequency. Finally, quality changes may be more abrupt with annual data.

Gopinath and Rigobon provide a careful analysis since they look at individual prices and compute exactly the PT between 2 price changes. But, from our point of view it is

A. Difficult to have the same data for all countries.

B. Their analysis focuses on the US which appears as very specific with LCP for imports and PCP for exports. It is unlikely to be the case, by construction, for the euro area countries (high level of pass through for the other countries, since we have PCP at least from the US).

C. In addition, their analysis may be plagued by other problems, in particular since they use consumption prices to measure cost of foreign exporters. In

addition, they use some correspondence between the sectoral breakdown for consumption prices and import prices due to Feenstra, while we use an homogeneous sector breakdown.

D. Another paper by Gopinath and Rigobon (with Itskhoki) shows pass through is much higher for goods not priced in home currency (USD), which should also apply to euro area countries.

Question 3. Why not use the same data as CM in p. 11.

We actually do have a full set of results the 1989-2001 data (which was included in a working paper version of the paper) in order to compare the results. We chose not to include these results here for reasons of space constraints but are very happy to report them if the referee wishes us to do so. As the two sets of data (1989-2001 and 1995-2005 are not compatible, due to revisions in the data and the absence of bridging series, we decided not to merge them. We then had to decide on which of the two to focus (given that the paper is already rather lengthy). For a number of reasons, including the more up-to-date status of the 1995-2005 series, or the fact that the 1989-2001 data set could not have dealt accurately with the transition to the euro, we chose the more recent series, while reworking all the results for the 1989-2001 series.

Question 4: The paper is very long for an academic paper. If nothing else they could streamline the language quite a bit in many places. For instance p. 5, third row have investigated the issue of exchange … "the issue" could easily be skipped. Also sure statements that are outside the study should be skipped, for instance "hence truly creating a single market for exporters".

We would do our best to provide a more streamlined discussion, although there is a depth of detail – on data, on economic and econometric theory and on estimation – that would be a shame to lose. We consider this as an important part of the richness of our analysis.

Question 5: Similarly, of they keep section 5, could skip directly to equation (13).

We would feel very reluctant to do this. Skipping Section 5 up to equation 13 would make the exposition less clear, while saving only about one page.