Heterogeneous dynamics, aggregation and the persistence of economic shocks.

Laura Mayoral*

Institute for Economic Analysis, (CSIC) and Barcelona GSE

June 9, 2012

Abstract

Estimates of shock persistence based on disaggregate or on aggregate data are frequently very different. These discrepancies are often attributed to the existence of heterogeneity at the disaggregate level, which has been argued to cause estimates of shock persistence based on aggregate data to be significantly larger than those derived from its disaggregate counterpart.

This paper takes a step towards reconciling this apparent disconnect between micro- and macro-based estimates of shock response. To this end, a fairly general disaggregate model with heterogeneous dynamics is examined and the values of several measures of persistence are compared across aggregation levels. It is shown that, while the average of the individual impulse response functions (IRFs) is identical to the aggregate IRF, averages of other popular persistence measures, such as the sum of the autoregressive coefficients (SAC) among others, tend to be larger the higher the aggregation level. I argue, however, that this should not be interpreted as evidence in favor of a persistence increase but, rather, as an undesirable property of these measures. The theoretical results are illustrated with two applications that use U.S. and European inflation data.

Key words: Heterogeneous dynamics, aggregation, persistence, impulse response function, sum of the autoregressive coefficients, inflation persistence, nominal rigidities.

*E-mail: laura.mayoral@iae.csic.es. I am very grateful to the editor and three anonymous referees. I am also indebted to Manuel Arellano, Lola Gadea, Albert Marcet, Enrique Sentana and participants in seminars at the Universities of Harvard, MIT, Glasgow, Southampton, Pompeu Fabra, CEMFI and IAE. Financial support from the Fundación Ramón Areces, the Spanish Ministry of Education through grant ECO2011-25293, the AXA research fund and Recercaixa is gratefully acknowledged. The usual disclaimer applies.

1. INTRODUCTION

Estimates of the persistence of shocks based on disaggregate data versus those based on aggregate data are often difficult to reconcile. For instance, one of the conclusions of the Inflation Persistence Network, created by the European Central Bank with the aim of analyzing the patterns of European inflation persistence, was that there is clear evidence of large differences of inflation persistence across sectors and *"that measures of the degree of inflation persistence increase with the level of aggregation. Individual or highly disaggregate price series are, on average, much less persistent than aggregate ones"* (see Angeloni, Aucremanne, Ehrmann, Galí, Levin and Smets, 2006, and the references therein).[1] A similar type of tension has also been found in estimates of the persistence of real exchange rates (Imbs, Mumtaz, Ravn and Rey, 2005), and in estimates of the duration of nominal rigidities (Imbs, Jondeau and Pelgrin, 2011), among others.

These discrepances are often attributed to a large amount of heterogeneity at the disaggregate level. Two types of arguments have been employed to explain how individual heterogeneity can create this apparent disconnect between micro- and macro-based estimates. On the one hand, building on the results of Pesaran and Smith (1995), it has been argued that when the dynamics of the disaggregate variables are heterogeneous, the estimates computed with aggregate data are biased and that the sign of the bias is positive (Imbs et al., 2005). On the other hand, influenced by the results in Robinson (1978) and Granger (1980), some authors have suggested that the aggregation of heterogeneous dynamic processes is not an innocuous operation and that it may tend to increase overall shock response (Altissimo et al., 2009 among others).

In contrast, some authors have reported very similar persistence values across aggregation levels in highly heterogeneous datasets (Crucini and Shintani, 2008, Gadea and Mayoral, 2009 and Mayoral and Gadea, 2011). The conclusions of all these articles are typically drawn by comparing averages of individual persistence measures with their corresponding values computed with aggregate data. However, different papers employ different tools to measure persistence and, therefore, it is not clear whether these findings are related to properties of the data or whether the use of a particular persistence measure can systematically bias the conclusions towards a specific direction.

[1] Similar findings have also been reported for U.S. inflation (Clark, 2006)

This paper takes a step towards reconciling the apparent disconnet between micro- and macro-based estimates of shock persistence. To this effect, it examines the relationship among measures of persistence of aggregate shocks computed at different levels of aggregation. We consider a general model at the disaggregate level and use some of the most popular measures to establish the comparison; namely, the impulse response function and other popular scalar measures such as the sum of the autoregressive coefficients, the largest autoregressive root and the half-life.

Our results demonstrate that not all the measures routinely employed in applications fare equally well. It is shown that the response to an aggregate shock, as measured by the impulse response function (IRF) of the aggregate model or by the average of the individual IRFs, is the same on all horizons with or without individual dynamic heterogeneity. This implies that, according to the IRF, the aggregation of heterogeneous units does not magnify the *average* response to a shock. The intuition of this result is simple: both the aggregate process and the IRF can be defined as two expectations. Under standard assumptions, the order of these expectations can be interchanged in such a way that identical results are obtained if one aggregates first and then computes the IRF or if the IRFs of the individual processes are obtained first and their average is computed in a second step.

In contrast, other popular persistence measures, such as the sum of the autoregressive coefficients (SAC), are not invariant to aggregation when individual heterogeneity is allowed for, being typically larger, the higher the level of aggregation. However, this should be interpreted as an undesirable property of these measures in this particular context, rather than as a sign of different average persistence across aggregation levels. The reason is that the SAC is a nonlinear transformation of the IRF and, therefore, the order in which expectations are taken matters. A straightforward application of Jensen's inequality shows that the average SAC tends to increase systematically with the level of aggregation. Thus, by relying on the latter measure, one could conclude that the average response to a shock increases with the level of aggregation when, in fact, a more thorough analysis of the IRF would suggest the opposite. Similar problems appear when other nonlinear measures, such as the largest autoregressive root (LAR), are employed for analogous purposes.

To illustrate some of the pitfalls involved in the micro-macro comparisons, two empirical illustrations are provided, one dealing with the estimation of U.S. inflation persistence across aggregation levels and the other, with the estimation of the duration of nominal rigidities using sectoral and aggregate French inflation.

The structure of this paper is as follows. Section 2 establishes the relation between micro- and macro-based IRFs. Section 3 considers other persistence measures that are frequently employed in applications. Section 4 presents two empirical illustrations using French and U.S. inflation data at different levels of aggregation, and Section 5 concludes. The Appendix presents other materials not included in the main text.

2. THE RELATION BETWEEN MICRO AND MACRO IRFS UNDER HETEROGENEITY

For the sake of clarify, this section begins by illustrating the main result of the paper using a simple model. Then, this result is extended to a more general framework.

Consider the problem of assessing the impact of an aggregate shock when micro data is available and the different units are heterogeneous. A simple but frequently postulated model for microeconomic behavior that allows for heterogeneous dynamics is the Random Coefficients Model. For each individual (or sector of the economy) i it is given by[2]

$$y_{it} = a_i y_{it-1} + b_i' x_{it} + \nu_{it}, \; i = 1, ..., N, \; t = 1, ...T, \tag{1}$$

$$\nu_{it} = \rho_i u_t + \varepsilon_{it,} \tag{2}$$

where t denotes time, y_{it} and x_{it} are observable variables, a_i, b_i and ρ_i are unknown coefficients that are assumed to be particular draws from some random variables $a = \bar{a} + \eta^a$, $b = \bar{b} + \eta^b$ and $\rho = 1 + \eta^\rho$, where η^k, for $k \in \{a, b, \rho\}$, are mutually independent, zero-mean random variables with variance σ_k^2, and $(\bar{a}, \bar{b})$ are some constants. The distribution of a has bounded support in the interval (-1,1]. The innovation ν_{it} is the sum of one common shock, u_t, and one idiosincratic, ε_{it}. The processes u_t and $\{\varepsilon_{it}\}_{i=1}^N$ are orthogonal, zero-mean martingale difference sequences. If x_{it} is just a constant, then (1) is simply the first-order autoregressive model.

Suppose now that, at time t, a unitary aggregate shock occurs. For each unit i, the impact of this shock h periods ahead can be evaluated through the IRF, defined as the difference between two conditional expectations (see Koops et al., 1996)

$$IRF_i(t, h) = E\left(y_{it+h} | u_t = 1; z_{it-1}\right) - E\left(y_{it+h} | u_t = 0; z_{it-1}\right), \tag{3}$$

[2]For some macroeconomic applications see Imbs et al. (2005), Crucini and Shintani (2008) or Mayoral and Gadea (2011).

where the operator E(.|.) denotes the best mean squared error predictor and $z_{it-1} = \begin{pmatrix} y_{it-1}, & y_{it-2}, & \dots & x_{it-1}, & x_{it-2}\dots \end{pmatrix}'$. Application of this definition to (1) yields

$$IRF_i(t,h) = \rho_i a_i^h, \text{ for } h \geq 0. \tag{4}$$

The average response to this aggregate shock can be computed as the expected value of (4) over the distribution of units, denoted as IRF$_{dis}$,

$$IRF_{dis}(t,h) = E_I\left(IRF_i(t,h)\right), \text{ for } h \geq 0, \tag{5}$$

where $E_I(.)$ denotes expectation across the distribution of units. Assuming that $\mathrm{E}_I\left(a^h\right)$ exists for all h, (5) is given by

$$IRF_{dis}(t,h) = E_I\left(a^h\right), \text{ for } h \geq 0. \tag{6}$$

In a representative agent economy, the relation between the IRF computed from aggregate data and the average of the individual impulse responses defined in (5) is straightforward. This is because the individual and the aggregate models share the same dynamics so the IRFs derived from each model are also the same. When the individuals are heterogeneous, however, the dynamics of the aggregate process are different from those of the individual units. Thus, in order to compute the aggregate IRF, it is first necessary to consider the aggregation of model (1). This problem has been addressed by Lewbel (1994), who followed the "stochastic" approach to aggregation (Stoker, 1984). The latter author defines an aggregate function as the expected value over the distribution of agents of the micro relations (see Stoker, 1984, Definition 2).[3]

To simplify the aggregation problem, we further assume that the process x_{it} is equal to 1 for all i and t, and that a is independent of the distribution of (b, ν_{it}). In addition, $B = E_I(b)$ is assumed to exist. So,

$$Y_t = B + E_I(ay_{t-1}) + u_t, \tag{7}$$

where $Y_t = E_I(y_{it})$. As shown by Lewbel (1994), expression (7) can be written as

$$Y_t = B + \sum_{s=1}^{\infty} A_s Y_{t-s} + u_t, \tag{8}$$

[3] In applications, aggregate data is typically constructed as the average of the individual processes. The Appendix summarizes the conditions for the holding of a law of large numbers relating the above-mentioned average with the corresponding expectation.

for constants $A_1,\ A_2,...,$ defined as $A_s = E(\alpha_s)$, where $\alpha_1 = a$ and $\alpha_s = (\alpha_{s-1} - A_{s-1})\,a$ for $s > 1$.

Using (8) it is easy to compute the (aggregate) IRF in a similar way as before,

$$IRF_{AG}(t,h) = E\left(Y_{t+h}|u_t = 1; Z_{t-1}\right) - E\left(Y_{t+h}|u_t = 0; Z_{t-1}\right), \tag{9}$$

where $Z_{t-1} = (Y_{t-1},\ Y_{t-2}, ...)$. Application of this definition to (8) yields

$$IRF_{AG}(t,h) = \begin{cases} IRF_{AG}(t,0) = 1, \\ \sum_{j=1}^{h} A_j IRF_{AG}(h-j), \text{ if } h \geq 1. \end{cases} \tag{10}$$

In order to simplify this expression, notice that the coefficients A_s can easily be shown to satisfy the equation

$$A_s = m_s - \sum_{r=1}^{s-1} m_{s-r} A_r, \tag{11}$$

where $m_s = E_I(a^s)$. It follows from equation (11) that $m_s = \sum_{r=0}^{s-1} m_r A_{s-r}$. Iterating (10) from $h = 1$, it is straightforward to check that

$$\begin{aligned} IRF_{AG}(t,\ 1) &= A_1 = m_1 = E_I(a) = IRF_{dis}(t,\ 1) \\ IRF_{AG}(t,\ 2) &= A_1^2 + A_2 = m_2 = E_I\left(a^2\right) = IRF_{dis}(t,\ 2) \\ IRF_{AG}(t,\ 3) &= A_1(A_1^2 + A_2) + A_2 A_1 + A_3 = m_3 = E_I\left(a^3\right) = IRF_{dis}(t,\ 3) \\ &\cdots \\ IRF_{AG}(t,\ h) &= m_h = E_I\left(a^h\right) = IRF_{dis}(t,\ h). \end{aligned} \tag{12}$$

Equation (12) is interesting since it establishes a link between the micro and macro response to an aggregate shock: the IRF computed in the aggregate model is just the expected value of the individual IRFs and, as will be shown below, this is also true in a more general setting than the one considered above. Thus, the aggregation of heterogeneous processes like (1) does not amplify the *average* response to a given shock.[4]

[4] The result above does not imply that other properties of the micro processes are invariant to aggregation. As is well known, the aggregation of short-memory stationary processes may induce long memory or nonstationary behavior in the aggregate variable (see the Appendix for details). Similarly, if some of the individual processes are I(1) while the others are strictly stationary, the aggregate will also be an I(1) process. Nevertheless, even in this case, the above-described relationship among IRFs across different aggregation levels still applies.

The intuition behind this result relies on the fact that both the IRF and the aggregate process are defined as the expected value of some expressions (see (3) and (7)). Thus, under standard assumptions, the order of the expections can be interchanged in such a way that identical results are obtained if one aggregates first and then computes the IRF, or if the IRFs of the individual processes are obtained and their average is computed next.[5]

This implies that the relation in (12) can also be established in a more general framework than the one considered above. Consider now a group of AR(p) processes that may present heterogeneous dynamics

$$y_{it} = b_i + a_{1i} y_{it-1} + ... + a_{pi} y_{it-p} + \nu_{it}, \quad (13)$$

$$\nu_{it} = \rho_i u_t + \varepsilon_{it}, \quad (14)$$

where $\{a_{ji}\}_{j=1}^{p}$, b_i, ρ_i and ν_{it} are defined as before. The aggregate model $Y_t = E_I(y_{it})$ can be written as a linear combination of lagged values of Y_t where the coefficients are complex functions of the moments of the individual parameters (see Lewbel (1994) for details).

The following theorem states that the relation between sectoral and aggregate IRFs also holds for the general AR(p) case.

Theorem 1 *Let $\{y_{it}\}_{i=1}^{N}$ be a group of heterogeneous processes defined as in (13) and (14). Under the previous assumptions it follows that*

$$IRF_{AG}(t,\ h) = IRF_{dis}(t,\ h),\ for\ h \geq 0. \quad (15)$$

The proof of this theorem is a straightforward application of Fubini's theorem, that allows to interchange the order of the expectations that define the IRFs and the aggregate process (see the Appendix).

An interesting consequence of (15) is that estimation of the aggregate IRF can be carried out using either aggregate or sectoral data. The Appendix includes some Monte Carlo

[5] If all the individual processes are stationary and the aggregate process is also so (which is not simply implied by the stationary of the units, see Zaffaroni 2004 for the relevant conditions), there is an even simpler way of looking at this result. From the MA representation of the individual processes, $y_{it} = (1 + a_i L + a_i^2 L^2 + ...)(\rho_i u_t + \varepsilon_t)$, one can easily obtain the individual IRF to an aggregate shock, $IRF_i(t,h) = a_i^h \rho_i$. The aggregate process can be computed by taking expectations over the distribution of units, $Y_t = E(y_t) = (1 + E(a)L + E(a^2)L^2 + ...)u_t$. It follows that $IRF_{AG}(t,h) = E(a^h)$, which coincides with the average of the individual IRFs, $IRF_{dis}(t,h) = E(a^h \rho) = E(a^h)$.

experiments that estimate the impulse response to an aggregate shock using both types of data. Two main conclusions stand out. Firstly, they show that good approximations to the population quantities can be obtained using either type of data and standard estimation techniques. Secondly, important efficiency gains are obtained when disaggregate data is employed. This implies that even when the focus of the analysis is exclusively on aggregate outcomes considering disaggregate information can yield much more efficient estimators.

Finally, it is common practice to apply transformations to the data, typically logs, prior to modelling them. As pointed out by a referee, this might induce discrepancies between micro and macro IRF estimates in some cases, for instance, when the individual processes are modelled in logs but the aggregate series is defined as the log of the average of the units. Clearly, the log of the average differs from the average of the logs and, therefore, their associated IRFs will also be different. However, from an empirical perspective, the resulting discrepancy could be potentially small. In a Monte Carlo experiment, similar to that reported in Section 2 in the Appendix, this is indeed what is found. In that exercise, the log of the individual units follow (heterogeneous) AR(1) processes, for different distributions of the heterogeneous AR parameter (described in Table A1 in the Appendix). These processes have been aggregated in two ways, Y_t and Y_t^*, defined as $Y_t = \sum_{i=1}^{N} \log y_{it}/N$ and $Y_t^* = \log(\sum_{i=1}^{N} y_{it}/N)$. The IRFs associated to Y_t and to Y_t^* and their associated MSE were very similar in all cases.[6] Of course, this should not be taken as a general result since the choice of data generating process could affect the magnitude of the discrepancies. Further research is needed to establish this result in a general setting.

3. OTHER PERSISTENCE MEASURES

Since the IRF is an infinite vector of numbers, it is a rather unwieldy measure of persistence. For this reason scalar measures are frequently preferred. This section evaluates the properties of some of the most popular scalar measures when used to compare persistence across aggregation levels. In particular, the cumulated impulse response (CIR), the half life (HL), the sum of the autoregressive coefficients (SAR), and the largest autoregressive root (LAR) are examined. It is shown that not all of these measures perform equally well. As will be illustrated in the following section, this is one of the reasons why discrepances between aggregate and disaggregate measures have been found repeatedly in the literature.

[6] This experiment is not reported in the Appendix but it is available upon request.

The *cumulated impulse response* (CIR) evaluates the total cumulative effect of a shock over time. Its average value at the micro level can be computed as

$$CIR_{dis} = \sum_{h=0}^{\infty} IRF_{dis}(t, h). \tag{16}$$

The *half life* (HL) is usually defined as the largest value of h^* that verifies $IRF(t, h^*-1) \geq 0.5$ and $IRF(t, h^*+1) <0.5$ (see Kilian and Zha, 2002). There are several ways of defining the average HL based on disaggregate data. One way is to apply the HL definition to IRF$_{dis}$, i.e., the *disaggregate* HL (HL$_{dis}$) is the largest value of h^* that verifies $IRF_{dis}(t, h^*-1) \geq 0.5$ and $IRF_{dis}(t, h^*+1) <0.5$. A second way to define it is to compute the HL associated to each IRF_i and, then, compute their average value, $\overline{HL}$.

At the aggregate level, the CIR and the HL are defined as $CIR_{AG} = \sum_{h=0}^{\infty} IRF_{AG}(t, h)$ and as the largest value of h^* such that $IRF_{AG}(t, h^*-1) \geq 0.5$ and $IRF_{AG}(t, h^*+1) <0.5$, respectively. Clearly, the equality between IRF$_{AG}$ and IRF$_{dis}$, established in Theorem 1, implies that $CIR_{dis} = CIR_{AG}$ and $HL_{dis} = HL_{AG}$. Notice however that, since the HL is a non-linear function, $\overline{HL}$ would be in general different from both HL_{dis} and HL_{AG}.[7]

In many applications average shock persistence is evaluated using the *sum of the autoregressive coefficients* (SAC) and the *largest autoregressive root*. Typically, these measures are computed for each individual time series and, next, averages (or distribution quantiles) are reported. Then, these measures are compared with the SAC or LAR obtained from aggregate data. See Altissimo et al. (2006), Bilke (2005) or Clark (2006) for some applications.

The SAC was originally introduced by Andrews and Chen (1994) because of its relation with the CIR through the expression

$$SAC = 1 - CIR^{-1}, \tag{17}$$

and so, *"different values of the SAC can be interpreted easily in terms of persistence because they correspond straightforwardly to different values of the CIR"* (Andrews and Chen, 1994). Thus, the SAC is a non linear function of the CIR (and of the IRF). This nonlinearity implies that taking expectations (aggregating) before or after applying this transformation will yield

[7] There is not a clear relation between HL_{dis} and $\overline{HL}$. Even in the case where the individual processes are all AR(1), HL_{dis} can be larger or smaller than $\overline{HL}$, depending on the distribution of the autoregressive coefficient.

different results. To see this, consider the simple model described in (1) . Notice that CIR$_{dis}$ is given by

$$CIR_{dis} = 1 + E_I(a) + E_I(a^2) + ... = E_I\left(\frac{1}{1-a}\right) \quad (18)$$

and recall that $CIR_{dis} = CIR_{AG}$. On the other hand, $SAC_{dis} = E_I(a)$. Since $\frac{1}{1-a}$ is a convex function, Jensen's inequality implies that the macro-based SAC, SAC_{AG}, is strictly larger than the average of the individual SACS, SAC$_{dis}$, because

$$SAC_{AG} = 1 - \left(E_I\left(\frac{1}{1-a}\right)\right)^{-1} > 1 - \left(\frac{1}{1-E_I(a)}\right)^{-1} = E_I(a) = SAC_{dis}, \quad (19)$$

unless there is no heterogeneity, in which case both measures are equal. Notice, however, that this result does not imply that average persistence increases with the level of aggregation: it only implies that under individual heterogeneity, SAC$_{dis}$ is a lower bound of the CIR$_{dis}$ and therefore, is a poor summary of this measure, since (17) does not hold anymore.

The LAR suffers from similar problems to the SAC. For instance, for the AR(1) model in (1), the average of the sectoral LARs, LAR$_{dis}$=SAC$_{dis}$ $= E_I(a)$. The aggregate value of the LAR (LAR$_{AG}$) will be, in general, different from this quantity. To illustrate this, consider a particular distribution for the autoregressive parameter. Assume that a follows a U(0,1) distribution. In this case, LAR$_{dis}$ =SAC$_{dis}$ $= 0.5$. As for the corresponding aggregate values, notice that the aggregate autoregressive polynomial, A(L), verifies A(L)=$M(L)^{-1}$, where $M(L) = \sum_{h=0}^{\infty} E(a^h) L^h$. The moments of the uniform distribution, $E(a^j) = (j+1)^{-1}$ are not summable, which implies that M(1)=∞. Therefore, A(1)=0 or, in other words, 1 is a root of the autoregressive polynominal. Thus, LAR$_{AG}$=1. Also, SAC$_{AG}$ $= 1 - A(1) = 1$.

As the former example illustrates, values of the SAC and the LAR can be very different across aggregation levels, even when the average effect of the aggregate shocks is identical.

An alternative measure of sectoral persistence is employed by Imbs et al. (2005) (IMRR henceforth), who report very different estimates of (average) persistence of real exchange rates when computed with sectoral or aggregate data. Gadea and Mayoral (2009) (GM henceforth) have analyzed in depth the methodology followed in that paper and have shown that IMRR's measures of persistence based on sectoral data systematically underestimate (average) persistence. The source of the bias is precisely the definition of the sectoral impulse response function used by these authors. Instead of computing the individual impulse responses and averaging them in order to produce an estimate of the average sectoral impulse response, they first estimate the mean value of the (heterogeneous) model coefficients in a panel of countries and, then, use this value to estimate their 'average' impulse response function, as if the model was one of homogeneous coefficients given by the mean value of the heterogeneous AR coefficients. As shown in GM, averaging the IRFs may yield very different results to averaging the AR coefficients and then computing the IRF. In fact, Jensen's inequality ensures that, for most empirically relevant cases, the former measure is always larger than the latter. Using the same data set and the same estimation strategy as those employed in IMRR's paper, GM have quantified the size of the bias that affects IMRR's measures of sectoral persistence. It turns out that the bias is substantial and that, once it is corrected, sectoral persistence estimates increase considerably: the classical result of 3-5 year half-life of PPP deviations is recovered. Moreover, those estimates are largely compatible with the ones obtained when aggregate real exchange rates are employed, as the theoretical results in this paper predict.

4. EMPIRICAL ILLUSTRATION

Among the central issues in macroeconomics is the nature of short and long-run inflation dynamics. After a great deal of research in this area, comparisons of estimates of important parameters based on disaggregate versus on aggregate inflation often produce a conflicting picture. Several authors have reported estimates of the degree of inflation persistence that systematically increase with the level of aggregation (Angeloni et al., 2006). Estimates of the extent of nominal price rigidity display a similar tension. For instance, Imbs, Jondeau and Pelgrin (2011, IJP henceforth) have estimated sectoral and aggregate New Keynesian Phillips curves (NKPCs) and have found that estimates of the duration of nominal rigidities estimated using aggregate data almost double those obtained with sectoral data.

The purpose of this section is to illustrate some of the pitfalls involved in the above-mentioned comparisons in light of the results presented in Sections 2 and 3.

4.1. Inflation persistence

Most of the empirical studies that have analyzed the persistence of inflation shocks across different aggregation levels have relied on (averages of) the SAC and/or the LAR, see Clark (2006), Altissimo et al. (2006) and the references therein, for examples considering US and European inflation data, respectively. The general consensus in these papers is that inflation persistence increases with the level of aggregation.

This section compares estimates of U.S. inflation persistence computed at different levels of aggregation. In addition to the SAC and the LAR, cross-sectional averages of the IRFs are considered as measures of shock response.

A similar data set to Clark (2006) has been employed. Price indexes and nominal expenditures for all components of consumption, as measured in the NIPA accounts, have been obtained from the webpage of the Bureau of Economic Analysis (BEA). This dataset permits breakdowns at various levels of aggregation. We focus on core inflation, which excludes food and energy prices. Then, the aggregate variable, denoted as Level 1, is core inflation. We also report results for data broken into several levels of disaggregation, each spanning all the core inflation. The most disaggregate level (that we will refer to as Level 4) contains 109 disaggregate prices. Level 3 and Level 2 aggregate these 109 series into 46 and 11 categories, respectively. See Clark (2006) for details on the construction of these variables. The data is quarterly and covers the period 1976 to 2002.[8]

Sectoral inflation is highly heterogeneous. For the series in Level 4, mean inflation over the period considered ranges from -1.91 to 7.46 with a mean (median) of 3.70 (4.04) and a standard deviation of 1.5. Individual dispersion also shows important disparities across series, as shown by the minimum and maximum values (2.25 and 29, respectively). More importantly, the dynamics of inflation series, as measured by the SAC, also seem to be very heterogeneous. Values of the SAC (bias-corrected SAC) range from -0.24 (-0.14) to 0.95 (1.00), with a mean of 0.66 (0.77) and a standard deviation of 0.21 (0.18). See the Appendix for more details.

To compute impulse response functions and other persistence measures, AR(k) processes have been fitted to the data.[9] The order k has been chosen according to the GTS and to

[8] Clark (2006) analyzed the period 1959-2002. Nevertheless, in order to avoid problems derived from the existence of structural breaks around the 1973 crisis, which would have a great impact on persistence estimates (see Perron, 1989), we have preferred to avoid this period by considering only post-crisis data.

[9] As regards the estimation of aggregate data, difficulties arise because the corresponding models contain an infinite number of parameters. Berk (1974) has shown that $\sqrt{T}$-consistent and asymptotically normally distributed estimates can be obtained by approximating the AR(∞) process by an AR(k) one, where k does not increase too quickly or too slowly (k verifies an upper and a lower bound condition).Thus, the choice of this parameter is key. Standard selection criteria (AIC or BIC) choose values of k, $\hat{k}$, that are too small and this may lead to severe finite sample biases (Ng and Perron, 1995). The general-to-specific (GTS) approach, however, can be used to produce a data-dependent selection rule such that k verifies the

the AIC.[10] Table I reports several measures of persistence computed at different levels of aggregation (Levels 1 to 4). The first four rows of Table I present the sum of the first h values of the IRFs relative to aggregation levels 1 to 4, for $h = \{4, 8, 12, 16 \text{ and } 20\}$, that is, the cumulated response of inflation from 1 to 5 years after the shock occurs. The bottom rows of Table I report averages of the SAC and the LAR for the four aggregation levels. Confidence intervals (CIs) have been computed using bootstrap methods following Kilian (1998). Figure I, in turn, depicts the estimated IRFs associated with aggregation levels 1 to 4.[11]

(Table I around here)

In agreement with the theoretical results, Table I shows that impulse responses computed from aggregate and sectoral data are very close at all the considered horizons, especially when the GTS selection method is employed. Using the CIs associated with the CIRs computed in Level 1, it is not possible to reject that the estimates obtained in Level 4 are identical to those corresponding to Level 1. Notice that the CIs associated with Level 4's estimates are strictly included in those of Level 1. As can be seen from Figure I, the four IRFs present approximately the same values and the same pattern of decay, so they imply a very similar degree of shock persistence, in line with the results presented in Table I. It is also remarkable that the GTS performs better than the AIC, as the theory predicts, in obtaining more homogeneous values across aggregation levels, stressing the importance of fitting long autoregressions when individual heterogeneity is present.

The values of the SAC and the LAR, however, vary considerably across aggregation levels, reproducing what has been found in previous studies. Using the GTS method, values of the SAC range from 0.74 corresponding to Level 4 (with a CI of (0.69, 0.78)), to 0.89 for Level 1 (with a C.I. of (0.82, 1.05)). Notice that the confidence intervals associated with these estimates do not even overlap. Identical conclusions can be drawn from estimates of

two bound conditions and the parameters obtained in the AR(k) model are consistent and asymptotically normal for the parameters of the underlying AR(∞) model (Kuersteiner, 2005).

[10] The maximum number of lags was set to 20, 16, 12 and 8, for aggregation levels 1 to 4, respectively. The significance level for applying the general-to-specific criterion was 10%. Small sample-bias corrected estimates have also been computed but they are not reported since results are qualitatively identical.

[11] The GTS method has been employed in the calculations. Confidence bands at the 5% significance level have been obtained using bootstrap and correspond to $\text{IRF}_{Level\ 4}$.

the LAR or when the AIC is employed to select k.

Therefore, from only looking at LAR and SAC figures, one would conclude that the response to a shock is higher, the higher the level of aggregation at which it is measured. However, a more thorough analysis of the evolution of the shock as described by the IRF suggest the opposite conclusions.

(Figure 1 about here)

4.2. Duration of nominal rigidities

IJP use French data on 16 economic sectors to estimate sectoral NKPCs allowing for heterogeneity at the sectoral level. The (log) price level of sector j at time t is defined as

$$p_{j,t} = \alpha_j p_{j,t-1} + (1-\alpha_j)\, p^*_{j,t},\ \ j = 1, ...16, \tag{20}$$

where α_j is the probability that a firm in sector j does not adjust its price at t and $p^*_{j,t}$ is the price that is set when firms are allowed to change their prices. The latter is given by $p^*_{j,t} = \omega_j p^b_{j,t} + (1-\omega_j) p^f_{j,t}$, where $p^b_{j,t}$ ($p^f_{j,t}$) denote the price set by backward (forward) looking firms (see IJP for details), and ω_j is the proportion of backward looking firms in sector j. Following Galí and Gertler (1999), they derive the following sectoral Phillips curves

$$\pi_{j,t} = \delta_{1j}\pi_{j,t-1} + \phi_{1j} h_j s_{j,t} + \phi_{2j} h_j s_{j,t-1} + \varepsilon_{j,t}, \tag{21}$$

where $\pi_{j,t}$ and $s_{j,t}$ denote sectoral inflation and average real marginal costs, respectively, h_j is a correction term, and the parameters δ_{1j}, ϕ_{1j} and ϕ_{2j} are (nonlinear) functions of the structural parameters ω_j and α_j (see IJP for details). IJP compare the estimates of the duration of nominal rigidities obtained in the sectoral specification to those obtained in an aggregate Phillips curve that does not allow for heterogeneity, similar to that considered by Galí and Gertler (1999), i.e.,

$$\pi_t = \delta_1 \pi_{t-1} + \phi_1 h s_t + \phi_2 h s_{t-1} + \varepsilon_t, \tag{22}$$

where π_t and s_t denote aggregate inflation and marginal costs, respectively. The key parameters for the computation of the duration of nominal rigidities are the $\alpha_j s$ and, therefore, the problem has a similar structure to that studied in the previous sections.

IJP find that '*sectoral estimates are in the vicinity of two quarters, as opposed to a little less than one year in the aggregate*'. In addition, they argue that their sectoral estimates are consistent with estimates of nominal rigidities based on microeconomic surveys of price duration, whereas aggregate estimates imply significantly longer nominal rigidities.

IJP's estimates of sectoral duration are based on the value of α corresponding to the *'representative sector'*. To estimate this parameter, denoted by $\hat{\alpha}^{RS}$, IJP estimate (21) for $j = 1, ..., 16$ and average across sectors the reduced-form estimates $(\hat{\delta}_{1j},\ \hat{\phi}_{1j},\ \hat{\phi}_{2j})$. The structural parameters are *'inferred'* on the basis of these averages. Next, the duration of nominal rigidities is derived from (20) according to the formula $\left(1/\left(1-\hat{\alpha}^{RS}\right)\right)$. Panel A in Table 2 summarizes their sectoral results.[12] The estimated value of $\hat{\alpha}^{RS}$ equals 0.508 which, according to the above-mentioned formula, implies a duration of nominal rigidities equal to 2.03 quarters.

Nevertheless, although averages of $(\hat{\delta}_{1j},\ \hat{\phi}_{1j},\ \hat{\phi}_{2j})$ provide consistent estimates of their corresponding population means, IJP's procedure does not yield consistent estimates of the means of the structural parameters $(\alpha_j,\ \omega_j)$, since the relation between the reduced-form and the structural parameters is highly nonlinear. Thus, it is unclear what $\hat{\alpha}^{RS}$ is actually estimating.

IJP provide sector by sector calculation of all the relevant parameters and, in particular, of α_j (Tables 2 and 3). Thus, it is possible to obtain a consistent estimator of $\mathrm{E}(\alpha)$, $\widehat{E(\alpha)}$, by averaging the sectoral estimates, $\hat{\alpha}_j$. The average of $\hat{\alpha}_j$, $j = 1, ...16$, is 0.688 (Panel B in Table 2).[13] Interestingly, this figure is almost identical to the estimate of α obtained by IJP when aggregate data is employed. The corresponding aggregate value is $\hat{\alpha}_{Agg} = 0.684$, (Panel D in Table 2).[14] Based on the latter values, IJP provide an (aggregate) estimate of the duration of nominal rigidities equal to $1/(1\text{-}\hat{\alpha}_{Agg}) = 3.17$ quarters. If ones uses $\widehat{E(\alpha)}$ in a similar formula, a duration equal to 3.21 quarters is obtained.

However, neither $1/\left(1-\alpha^{RS}\right)$ nor $1/1-E(\alpha)$ are appropriate measures of the average duration of nominal rigidities. The relevant quantity to be estimated is $E\left(1/1-\alpha\right)$ (that is, the CIR as defined in Section 3) since sectoral durations are given by $1/1-\alpha_j$. As equation (19) shows, $1/1-E(\alpha)$ is just a lower bound of the true average of nominal rigidities. The

[12] In the following, only figures estimated with ML-SURE are reported. Very similar figures are obtained using ML.

[13] Simple averages are employed throughout. Very similar results are obtained if GDP weights are employed to compute the averages.

[14] The coincidence of these values does not seem to be by pure luck. IJP also present Monte Carlo simulations suggesting that the estimated value of α based on aggregate data is very close to $\mathrm{E}(\alpha)$, see Table 5, Panel C in their paper.

difference between $1/1 - E(\alpha)$ and $E(1/1 - \alpha)$ can be important if some of the individual series are very persistent, as in this case. The average of 1/1-$\hat{\alpha}_j$, $j = 1, ...16$ (using the values of α_j in Table 3 in IJP) equals 252 quarters.

The latter estimate has to be taken with caution: the number of sectors is very small so the influence of a few very persistent ones (4 out of the 16 sectors have a root very close to 1) is very large. However, the fact that 1/4 of the sectors have $\alpha_j s$ so close to 1 is by itself at odds with the existing microeconomic evidence of price durations.[15]

Thus, as opposed to IJP's conclusions, sectoral NKPC do not seem to match well with French inflation data after all. Further research has to be undertaken to check whether this conclusion can be reversed by considering more disaggregate data at the sectoral level and/or more flexible representations. In addition, it would be interesting to be able to compare sectoral estimates with those obtained in an aggregate NKPC that explicitly allows for heterogeneity at the sectoral level. As mentioned above, aggregate NKPCs as that in (22) seem to yield estimates of α that are close to $E(\alpha)$. This implies that the estimates of the durations of nominal rigidities obtained from that specification seem to be close to 1/1-E(α) , which is a lower bound of the durations really implied by these models, unless there is no hetegeneity at the sectoral level. Further research is also needed to clarify this last issue.

5. CONCLUSION

This paper examines the relations among shock persistence measures computed at different levels of aggregation in a context where individual dynamic heterogeneity may exist. It is shown that the average of the individual IRFs equals the aggregate IRF at all horizons. A similar relationship also holds for any scalar measure that is a linear transformation of the IRF, such as the CIR. However, non-linear transformations, such as the SAC, do not verify this relationship. In fact, they tend to be larger the higher the level of aggregation considered. Since these measures are the most employed in applications, it is not surprising that different average shock behavior has been reported in many empirical papers. Two empirical illustrations, using U.S. and French inflation data at different levels of aggregation,

[15] According to the microdata collected by Baudry et al (2004), the average duration of nominal rigidities in France is around 10 months. On the other hand, the duration associated to the 'average' α, that is, 1/1-E(α) is around 5 months.

have been provided.

REFERENCES

Altissimo, F., B. Mojon and P. Zaffaroni (2009), "Can aggregation explain the persistence of inflation?," *Journal of Monetary Economics* 56, 231-241.

Altissimo, F., L. Bilke, A. Levin, T. and B. Mojon (2006), "Sectoral and Aggregate Inflation Dynamics in the Euro," *Journal of the European Economic Association* 4, 585-593.

Andrews, D. and H. Chen (1994), "Approximately median-unbiased estimation of autoregressive models," *Journal of Business and Economic Statistics* 12, 187-204.

Angeloni, I., L. Aucremanne, M. Ehrmann, J. Galí, A. Levin and Frank Smets (2006), "New Evidence on Inflation Persistence and Price Stickiness in the Euro Area: Implications for Macro Modelling," *Journal of the European Economic Association* (Papers and Proceedings) 4, 562-574.

Baudry, L., H. Le Bihan, P. Sevestre and S. Tarrieu (2004), "Price Rigidity in France - Evidence from Consumer Price Micro-Data," *ECB Working Papers Series*, 384.

Berk, K.N. (1974), "Consistent autoregressive spectral estimates," *The Annals of Statistics* 2, 489-502.

Bilke, L. (2005), "Break in the mean and persistence of inflation: a sectoral analysis of French CPI," ECB Working Paper 463.

Clark, T. (2006), "Disaggregate evidence on the persistence of consumer price inflation", *Journal of Applied Econometrics* 21, 563-587.

Crucini, M. J. and M. Shintani (2008), "Persistence in Law-of-One-Price Deviations: Evidence from Micro-data," *Journal of Monetary Economics* 55, 629-644.

Gadea, L. and L. Mayoral (2009), "Aggregation is not the solution: the PPP puzzle strikes back," *Journal of Applied Econometrics* 24, 875-894.

Granger, C.W.J. (1980), "Long memory relationships and the aggregation of dynamic models", *Journal of Econometrics* 14, 227-238.

Imbs, J., F. Pelgrin and E. Jondeau (2011), "Sectoral Phillips Curves and the Aggregate Phillips Curve", *Journal of Monetary Economics,* (forthcoming).

Imbs, J, A. Muntaz, M.O. Ravn and H. Rey (2005), "PPP Strikes Back: Aggregation and the Real Exchange Rate," *Quarterly Journal of Economic*s 120, 1-43.

Kilian, L. and T. Zha (2002), "Quantifying the Uncertainty about the Half-Life of Deviations from PPP," *Journal of Applied Econometrics* 17, 107-125.

Kuersteiner, G. M. (2005), "Automatic Inference for Infinite Order Vector Autoregressions", *Econometric Theory*, 21, 85-115.

Lewbel, A. (1994), "Aggregation and Simple Dynamics," *American Economic Review* 84, 905-918.

Mayoral, L. and L. Gadea (2010), "Aggregate real exchange rates persistence through the lens of sectoral data", *Journal of Monetary Economics,* 58.

Ng, S. and P. Perron, (1995), "Unit root tests in ARMA models with data-dependent methods for the selection of the truncation lag," *Journal of the American Statistical Association* 90, 268-281.

Perron, P. (1989) , "The great crash, the oil price shock and the unit root hypothesis", *Econometrica* 58, 1361-1401.

Pesaran, M. H. and R. Smith (1995), "Estimating long-run relationships from dynamic heterogeneous panels", *Journal of Econometrics* 68, 79-113.

Robinson, P.M. (1978), "Statistical inference for a random coefficient autoregressive model", *Scandinavian Journal of Statistics* 5, 163-168.

Stoker, T. M. (1984), "Completeness, distribution restrictions and the form of the aggregate function," *Econometrica* 52, 887-907.

Teräsvirta,T. (2006), "Forecasting economic variables with non-linear models," in G. Elliot, C.W.J. Granger and A. Timmermann, eds, *Handbook of Economic Forecasting*, Vol 1, Amsterdam: Elsevier, 413-457.

TABLES AND FIGURES

Table I. U.S Inflation Persistence

	Level 1		Level 2		Level 3		Level 4	
IRFs at different Aggregation Levels								
	$\sum_{i=1}^{h} IRF_{Level\ 1}(h)$		$\sum_{i=1}^{h} IRF_{Level\ 2}(h)$		$\sum_{i=1}^{h} IRF_{Level\ 3}(h)$		$\sum_{i=1}^{h} IRF_{Level\ 4}(h)$	
	GTS	AIC	GTS	AIC	GTS	AIC	GTS	AIC
$h = 4$	1.77 (1.25, 2.28)	2.08 (1.66, 2.52)	2.05 (1.86, 2.28)	1.98 (1.83 2.17)	1.89 (1.78, 2.03)	1.88 (1.76, 2.01)	1.85 (1.75, 1.96)	1.82 (1.73, 1.93)
$h = 8$	3.06 (1.94, 4.63)	3.70 (2.94, 4.74)	3.28 (2.86, 3.91)	3.08 (2.78, 3.58)	3.07 (2.85, 3.45)	3.04 (2.81, 3.44)	2.97 (2.77, 3.30)	2.94 (2.70, 3.22)
$h = 12$	4.01 (2.47, 7.27)	4.77 (3.54, 6.55)	4.08 (3.44, 5.23)	3.78 (3.30, 4.60)	3.97 (3.64, 4.72)	3.84 (3.50, 4.65)	3.81 (3.52, 4.48)	3.78 (3.38, 4.30)
$h = 16$	4.58 (2.54, 9.82)	5.56 (4.04, 8.03)	4.67 (3.84, 6.43)	4.24 (3.70 5.43)	4.70 (4.31, 5.94)	4.48 (4.08, 5.57)	4.46 (4.05, 5.58)	4.44 (3.94, 5.41)
$h = 20$	4.86 (2.62, 12.28)	6.13 (4.49, 9.44)	5.20 (4.21, 7.78)	4.65 (4.05, 6.30)	5.19 (4.89, 7.19)	4.99 (4.50, 6.72)	5.01 (4.47, 6.67)	4.97 (4.38, 6.48)
SAC at different Aggregation Levels								
	0.89 (0.82, 1.05)	0.87 (0.81, 0.95)	0.82 (0.73, 0.91)	0.80 (0.73, 0.87)	0.79 (0.72, 0.85)	0.77 (0.71, 0.83)	0.74 (0.69, 0.78)	0.72 (0.67, 0.76)
LAR at different Aggregation Levels								
	0.97 (0.95, 1.02)	0.94 (0.90, 0.98)	0.95 (0.93, 0.98)	0.93 (0.91, 0.96)	0.91 (0.90, 0.94)	0.89 (0.87, 0.92)	0.89 (0.87, 0.91)	0.84 (0.82, 0.87)

Notes. Level 1 is aggregate core inflation. Levels 2-4 contain 11, 46 and 109 series, respectively. $IRF_{Level\ i}$: Impulse response function (IRF) computed as the average of the IRFs of the series in Level i. LAR: Largest autoregressive root; SAC: sum of autoregressive coefficients. AR(K) models have been fitted to the series in Levels 1 to 4, where K has been chosen according to the Akaike Information Criterion (AIC) or to the General-to-Specific (GTS) approach. IRFs, SACs and LARs have been obtained on the bases of these estimates. 5% confidence intervals have been computed using bootstrap techniques. Small sample bias corrections as in Kilian (1998) have been performed to compute CIs.

Table 2. Estimates of the Duration of Nominal Rigidities

Sectoral data					Aggregate data	
Panel A		Panel B		Panel C	Panel D	
$\hat{\alpha}_{RS}$	$\frac{1}{1-\hat{\alpha}_{RS}}$	$\widehat{E(\alpha)}$	$\frac{1}{1-\widehat{E(\alpha)}}$	$E(\widehat{1/1-\alpha})$	$\widehat{\alpha_{Agg}}$	$\frac{1}{1-\widehat{\alpha_{Agg}}}$
0.508	2.03	0.688	3.21	252	0.684	3.17

Notes. All calculations are based on the estimates provided by IJP based on ML-SURE. Panels A and D reproduce their estimates (see Table 5, Panels A and B, respectively, in their paper). Estimates from Panels B and C are based on the figures provided in Table 3 in IJP's paper.

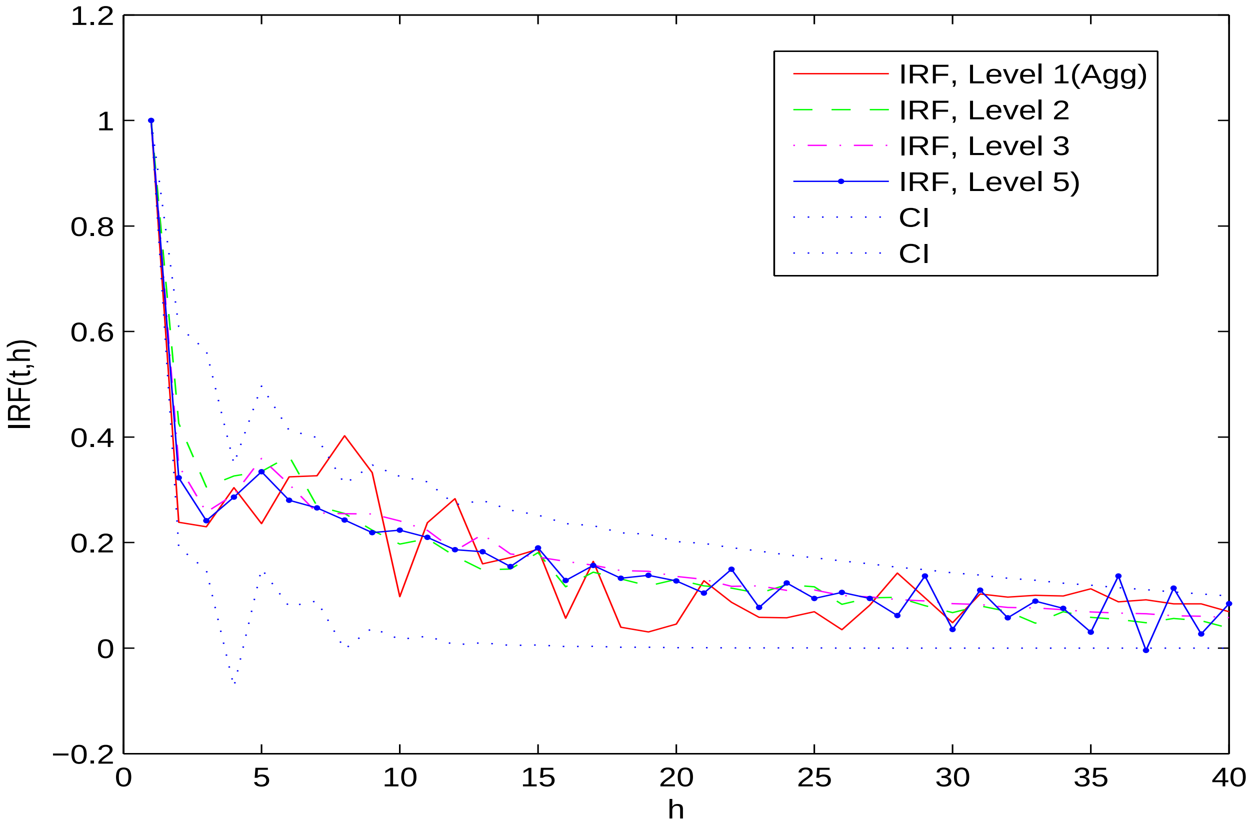

FIG 1. Impulse responses of U.S. inflation computed at different aggregation levels