AGGREGATE VS. DISAGGREGATE EURO-AREA MACRO-MODELLING

by Libero Monteforte* and Stefano Siviero*

This version: February 2003

Abstract

The available empirical evidence suggests that non-negligible differences in economic structures persist among euro area countries. Because of those asymmetries, an area-wide (aggregate) modelling approach is in principle less reliable than a multi-country (disaggregate) one. This paper examines the aggregate/disaggregate modelling trade-off from both a statistical and an economic viewpoint, using two simple models (an aggregate one and a disaggregate one) for the three largest economies in the euro area. From a statistical viewpoint, we find that standard aggregation bias criteria and tests signal that the degree of structural heterogeneity among euro area economies is such that the loss of information entailed by an aggregate modelling approach is not trifling. To tackle the area-wide/multi-country modelling trade-off from an economic viewpoint, we investigate the following issue: Are those statistically detectable asymmetries of any practical relevance when it comes to supporting monetary policy decision-making? To provide an answer to this question, we compute optimal monetary policy reaction functions on the basis of either the aggregate model or the disaggregate one, and compare the associated welfare losses. The results suggest that the welfare under-performance of an area-wide-model-based rule is not only not negligible, but also systematic, significant and robust with respect to a number of sensitivity analyses.

* Bank of Italy, Research Department. Email: monteforte.libero@insedia.interbusiness.it; siviero.stefano@insedia.interbusiness.it.

1. Introduction and main findings[1]

Following the introduction of the single currency in a number of European countries on 1st January 1999 and the establishment of a single monetary authority for the newly-born euro area, a number of novel challenges had to be faced by the European policymakers, the economic scholars and the practitioners alike.

Strictly from a modeller's viewpoint, the most fundamental new challenge is arguably posed by the following question: Do significant/relevant structural differences exist among euro-area economies? One's answer to this question will have major consequences in terms of modelling choices. Specifically, evidence of significant structural heterogeneity should recommend adopting a disaggregate (multi-country) modelling approach, consisting of specifying and estimating separate models for each economy (or groups of economies) in the area, allowing for possible interactions among them (models of this kind —e.g., the IMF's Multimod and OECD's Interlink models of the world economy— typically include a detailed description of trade linkages). If, by contrast, no significant indications of heterogeneity can be detected, then one may just as well follow a more parsimonious (and less costly) modelling approach (aggregate, area-wide), consisting of treating aggregate data as if they referred to one single, large and sufficiently homogeneous economy.

A question naturally arises here: Why there seems to be no comparable interest for aggregation issues in the case of other monetary unions? The answer rests, we believe, with the fact that heterogeneity among euro-area countries is widely presumed to be much more pronounced than in other monetary unions or federal states (the US being the most obvious comparison), largely because of differences in the institutional frameworks of participating countries, which are expected to persist, at least to a certain extent, for some time into the future (one just needs to think about the rather mild degree of co-ordination of fiscal policies). It is thus natural to conjecture that the potential information loss associated with using aggregate econometric tools is likely to be larger for the euro area than for other economies.

---

[1] We thank the late Albert Ando, Andrea Brandolini, Fabio Busetti, Riccardo Cristadoro, Balazs Horvath, Marco Lippi, Daniele Terlizzese and participants in workshops at the Banque de France and University of Rome "La Sapienza" for many helpful comments and suggestions on previous versions of this work. The usual disclaimers apply. The views expressed in this paper are those of the authors and do necessarily represent those of the Bank of Italy.

A casual look at the available empirical evidence indeed suggests that non-negligible differences in the economic structure of euro-area countries still persist to this day.[2] However, not much evidence is available that formally establishes the significance and relevance of those differences.

This paper tackles the aggregate/disaggregate euro-area modelling trade-off from both a statistical and an economic viewpoint.

To address the trade-off from a statistical viewpoint we explore —using a number of criteria and tests proposed in the literature— whether signs of aggregation bias are detectable in euro-area economic data. We find that they are.

To examine that trade-off from an economic viewpoint, we observe that any statistical evidence of aggregation bias does not necessarily *per se* imply that the use of aggregate econometric tools should result in unreliable analyses and insight and hence in significantly sub-optimal economic decisions. To explore this issue we test, specifically, whether the performance of an hypothetical monetary policy-maker relying on an aggregate model of the euro area would be significantly different from (particularly, worse than) that of a (similarly hypothetical) policy-maker whose decisions rest on a disaggregate model. Our approach may thus be viewed as proposing a policy-effectiveness-based metric for assessing the economic relevance of structural asymmetries across euro-area economies.[3] Should this analysis suggest that no big welfare losses are at the stake, an area-wide modelling approach might remain the preferred option, regardless of what the standard statistical checks of aggregation bias may indicate. Our results suggest, in short, that the underperformance of aggregate-model-based euro-area monetary policy-making is not only not negligible, but also systematic and significant.

---

[2] A (very) partial list of recent works that have a bearing on this issue includes Dornbusch, Favero and Giavazzi (1998), Ramaswamy and Sloek (1997), Guiso, Kashyap, Panetta and Terlizzese (1999), Hughes Hallett and Piscitelli (1999), Dedola and Lippi (2000), Clements, Kontolemis and Levy (2001), Ciccarelli and Rebucci (2002), Fabiani and Morgan (2003), Mayes and Virén (2000), and the papers presented at a recent ECB conference ("Monetary Policy Transmission in the Euro Area", ECB, Frankfurt, 18-19 December 2001). Not much effort has yet been devoted to trying to identify the structural determinants underlying the observed asymmetries. Fragmentary evidence may be found in van Els, Locarno, Morgan and Villetelle (2001).

[3] From the viewpoint of the debate on robust rules (see the contributions in Taylor (1999) and, more recently, Levin and Williams (2002)), the paper may be viewed as focussing on one particular type of robustness, i.e., robustness of rules based on aggregate models of the euro-area economy with respect to the assumption of aggregability.

The paper is organized as follows. Section 2 shortly describes the two basic modelling options faced by euro-area modellers and presents and compares the main features of the stylised aggregate and disaggregate models used in the remainder of the paper. Despite their sketchy nature, those models mimic some key properties of medium- to large-size models maintained by other institutions (most notably from our viewpoint, they mimic the properties that are presumably most relevant when it comes to the aggregate/disaggregate trade-off). Section 3 briefly describes the criteria and test statistics we use to investigate the aggregation bias issue and presents the empirical results. Section 4 presents the approach followed to assess the economic relevance of structural heterogeneity across euro-area countries and offers a quantification of the additional welfare losses that would be incurred by the euro-area monetary policy-maker should her/his decisions rely on an aggregate model rather than on a disaggregate one. We furthermore estimate the significance of those additional losses and explore how the results would likely be affected should euro-area economies tend to converge in the future. The concluding section draws some tentative conclusions as to what we believe our results imply concerning the choice of the appropriate modelling approach when it comes to supporting euro-area monetary policy-making.

2. Two simple models of the euro area

A modeller wishing to build empirical tools for forecasting and policy analysis purposes in the euro area faces two basic options: as a first alternative, one could build a disaggregate, or multi-country, model, i.e., a model that describes the functioning of the economic mechanisms in the individual countries of the area and the inter-linkages amongst them. Area-wide developments would then be derived by aggregating the individual country results. In a model of this kind, any country-specific features may be reflected by either the structure of the model and/or the value of its parameters. As a second, much less onerous, alternative, one may first aggregate the individual country data[4] and model the latter as if they referred to one single, large and homogeneous economy (aggregate, or area-wide, model).[5]

---

[4] Labhard, Weeken and Westaway (2001) argue that the actual choice of the aggregating function is unlikely to affect the properties of the model in any significant way. The aggregating functions used in this paper are briefly described in footnote 8.

[5] It is worthwhile emphasising that, while one's intuition might be that a disaggregate model must always be at least as reliable as the corresponding aggregate one, the econometric literature does not univocally predict that that will be the case. In particular, Grunfeld and Griliches (1960) show that the opposite ranking of the two approaches is possible in the event of measurement errors and/or misspecification of the disaggregate

Both approaches are being pursued in practice, even by the same institutions. For instance, the European Central Bank (ECB) maintains an aggregate model of the euro-area economy (Area-Wide Model, AWM; see Fagan, Henry and Mestre (2001)) and is in the process of building a disaggregate one (Multi-Country Model, MCM). Also, the Eurosystem projections, which have been published by the ECB since December 2000, are the result of a multi-staged process that involves aggregating country-specific projections (mostly based on the national models of participating NCBs, but also on some of the national blocks of the MCM) while also using information derived from the AWM, to come to one single, consistent picture (see ECB (2001)).[6]

The practical advantages of adopting an area-wide approach are obvious: an area-wide model is more parsimonious, less costly, more readily available, arguably more transparent. Unfortunately, snappiness often comes at a cost. Assessing the size of that cost is precisely one of the main purposes of this paper.

The next paragraphs describe the two models (aggregate and disaggregate) used in the remainder of this paper and present their main properties.[7]

2.1 *The aggregate (area-wide) model*

The Aggregate Euro Area Model (AEAM) is a simple two-equation model estimated using aggregate data for the three largest economies in the euro area (Germany, France and Italy, jointly accounting for over 70 per cent of the area GDP). It includes an aggregate supply equation (also referred to as Phillips curve) and an aggregate demand equation (also referred to as IS curve). The first equation determines inflation as a function of lagged inflation and the output gap. The sum of the coefficients on lagged inflation is constrained to unit (the

---

relationships.

[6] The foregoing description should have made it clear that we by no means intend to suggest that the process through which Eurosystem's decisions are formulated corresponds to either of the two extreme hypothetical cases that, for the sake of the argument, we contrast in this paper.

[7] The simple models presented below are entirely backward-looking; their parameters cannot be given a structural interpretation in terms of "deep" underlying parameters relating to preferences and technology. Hence, both models are potentially affected by the well-known difficulties associated with the evaluation of policy changes on the basis of behavioural relationships found to hold under a different policy set-up (Lucas (1976)). There are, however, several general reasons to believe that the Lucas Critique may in practice be less disruptive than is widely held to be (for a summary of the literature, see e.g. Altissimo, Siviero and Terlizzese (2002)). Also, the empirical evidence presented below overwhelmingly supports the hypothesis of structural stability, even for the most recent period, when, arguably, a major shock occured in the policy regime.

restriction cannot be rejected), so that the Phillips curve is of the accelerationist type. The second equation relates the output gap to its own lagged values and the real interest rate.

A general-to-specific modelling approach was followed in searching for a satisfactory empirical specification, starting with 6 lags for all variables on the right-hand-side of the two equations. The final specification is the following:

$$\pi_{t+1} = \alpha_1\pi_t + (1-\alpha_1)\pi_{t-3} + \eta y_t + u_{t+1}$$
$$y_{t+1} = \theta y_t + \psi(i_{t-1} - 4\cdot\pi_{t-1}) + v_{t+1}$$

where:[8]

- π_{t+1} = quarter-on-quarter consumer inflation rate;
- y_{t+1} = output gap;
- i_{t+1} = short-term interest rate;
- $i_{t+1-k} - 4\cdot\pi_{t+1-k} = r_{t+1}$ is thus a measure of the ex-post real interest rate.

The model was estimated with SURE, thus allowing for the possibility of correlation between the residuals of the two equations. The sample period extends from 1978.Q1 to 1998.Q4; 84 quarterly observations were therefore available for estimation. The estimation results are presented in Table 1.

2.2 *The disaggregate (multi-country) model*

The Disaggregate Euro Area Model (DEAM) includes, for each of the three largest euro-area countries, the same set of equations as the AEAM. The specification of both the aggregate supply and the aggregate demand equations is similar to the one adopted in the AEAM but, in addition, it allows for cross-country linkages. Specifically, inflation in any given country depends not only on its own lagged values and on the corresponding output gap,

---

[8] The source of data is the ESA-95 National Accounts for inflation and the output gap, and the BIS data-bank for the short-term interest rate. Inflation is measured by the quarter-on-quarter rate of change of the (seasonally adjusted) households' consumption deflator. Potential output was estimated by applying a band-pass filter (see Baxter and King (1995) for details) to the (log) GDP (selecting frequency components of 32 quarters and higher, with a truncation of 16 quarters). National variables were aggregated using a fixed-weight procedure, similar to the one adopted by the ECB. For inflation, 1999 PPP consumer spending shares (as computed by the ECB) were used. For output gap, the weights are given by 1999 PPP real GDP shares (again, the source of the shares is the ECB). For interest rates, the weights are the PPP nominal GDP shares computed by the OECD. The GDP and consumer spending weights are, respectively, 0.43 and 0.44 for Germany, 0.29 and 0.27 for France, 0.28 and 0.29 for Italy.

but also, at least in principle, on inflation "imported" from the other two countries (imported inflation is given, in estimation, by the sum of inflation in the foreign country and the rate of change of the relevant bilateral exchange rate; the hypothesis of equality of the corresponding coefficients could not be rejected). As in the AEAM, the sum of the coefficients on lagged and imported inflation is constrained to be 1 (the restriction is accepted by the data for all countries). The output gap in any of the three countries depends on its own lagged values and on the corresponding real interest rate, as in the AEAM; in addition, it may react to the output gap in the other two countries, reflecting trade linkages.

The DEAM also comprises two identities for euro area inflation and output gap (with the same weights as those underlying AEAM data).

As the model set-up allows for instantaneous cross-country linkages, 3SLS were used to estimate its parameters. The sample period extends from 1978.Q1 to 1998.Q4, as for the AEAM. For most of the sample period, the exchange rates among Germany, France and Italy, though constrained by the ERM of the EMS, were not fixed. Accordingly, the measure of "inflation imported in country j from country i" was constructed, as mentioned earlier, as the sum of the inflation rate in country i and the quarter-on-quarter percentage change in the bilateral exchange rate (units of currency of country j needed for 1 unit of country i's currency). In theory, full 3SLS estimation would require the model to include a set of equations for bilateral exchange rates. Given the well-known difficulty of finding satisfactory empirical specifications for the exchange rate, no attempt was made to augment the model with exchange rate equations; instead, lagged values of all variables included in the model were used as instruments for the exchange rates. At any rate, in the experiments presented below the percentage change of the exchange rate was set identically equal to zero, consistently with the introduction of the single currency as of January 1, 1999.

While the model set-up allows for instantaneous cross-country linkages, so that a simultaneous system estimation strategy is required, we chose to assume that the real interest rate affects the output gaps only with a lag (which, incidentally, is consistent with most available evidence). Hence, estimation could be carried out without augmenting the model with country-specific interest rate reaction functions for the three countries.

The general form of the two-equation sub-model for country j is the following:[9]

$$\begin{aligned}\pi^j_{t+1} &= \sum_{k=1}^{p}\alpha_{j,k}\pi^j_{t+1-k} + \sum_{i\neq j}\sum_{k=0}^{p}\beta_{j,i,k}(\pi^i_{t+1-k} + \dot{e}^{i,j}_{t+1-k}) + \sum_{k=0}^{p}\eta_{j,k}y^j_{t+1-k} + u^j_{t+1}\\ y^j_{t+1} &= \sum_{k=1}^{p}\theta_{j,k}y^j_{t+1-k} + \sum_{i\neq j}\sum_{k=0}^{p}\varphi_{j,i,k}y^i_{t+1-k} + \sum_{k=1}^{p}\psi_{j,k}(i^j_{t+1-k} - 4\cdot\pi^j_{t+1-k}) + v^j_{t+1}\end{aligned}$$

where:

- π^j_{t+1} = quarter-on-quarter consumer inflation rate in country j;
- $\dot{e}^{i,j}_{t+1-k}$= quarter-on-quarter rate of change of the exchange rate between country i and country j (units of country j's currency for 1 unit of country i's currency; in the experiments of Section 4 this variable is identically zero, consistently with the introduction of the single currency in January 1999);
- y^j_{t+1} = output gap in country j;
- i^j_{t+1} = short-term interest rate in country j (country-specific short-term interest rates were used in estimation; b contrast, in the experiments of Section 4 it is imposed that the nominal interest rate be the same for all countries, i.e, $i^j_{t+1} = i_{t+1}$ for all j's);
- $i^j_{t+1-k} - 4\cdot\pi^j_{t+1-k} = r^j_{t+1}$ is thus a measure of the *ex-post* real interest rate in country j.

The starting specification included on the right-hand-side of each estimated equation the first 6 lags of all relevant variables. Joint 3SLS estimation of the three sub-models resulted, after dropping all insignificant lags, in a much more parsimonious specification (see Table 2, where the exchange rates have been omitted, as they play no role in the version of the model used in the remainder of the paper).

2.3 *Stability testing*

Just like any other model of the euro area currently in use, the DEAM and AEAM were estimated with data for the period pre-dating the introduction of the euro. It may thus be feared that, notwithstanding their performance in the estimation period, they might be affected by structural discontinuities following the introduction of the single currency. Exploring whether this is, or is not, the case is the objective of this section.

---

[9] In keeping with the approach followed in similar literature, the model used for the welfare simulation in Section 4 does not include any constant terms, i.e., it may be taken to provide a description of the functioning of the euro area economy in the neighborhood of equilibrium. This amounts to implicitly assuming that the same equilibrium values apply to all countries, a condition that does not hold in the sample period.

While the euro was officially introduced only on January 1st, 1999, one may argue that, at least since late 1996, the monetary policies for the three countries we consider had been tightly constrained: the bilateral exchange rates remained basically constant at about the same level as the irrevocable exchange rates with which those countries joined the euro area two years later;[10] the financial markets considered it to be highly probable that those countries would participate in the single courrency (with the exception, for 1997, of Italy); moreover, fiscal policies were also tightly constrained by the convergence process. Taking 1997 to be the beginning of the euro era allows us to use a reasonably sized sample (20 quarterly observations) to test for stability. Accordingly, both models were re-estimated using pre-euro data as defined just above (1978.Q1 to 1996.Q4). For both models, the parameter estimates are basically the same as those found with the original sample (1978.Q1 to 1998.Q4).

The results of out-of-sample stability testing are shown in Tables 3 and 4 and Figures 1 and 2. For both models, the empirical evidence overwhelmingly rejects the hypothesis of parameter instability; the figures show no detectable signs of convergence of the DEAM parameters (actually, the cross-country dispersion of any parameter —say, the autoregressive term in the AD equations— is, if anything, slightly higher using the whole sample up to 2001 rather than using the samples up to 1996 or 1998, the only exception being the parameter of the interest rate in the AD equations). Although one cannot rule out the possibility of sizeable changes in the future, these results at least indicate that no such change is detectable yet, even though there is scarcely any doubt that the introduction of the euro represented a major breakpoint in the policy framework.

2.4 *Impulse responses of AEAM and DEAM*

Figures 3, 4 and 5 show the impulse responses of both models to a number of shocks. Since the Phillips curve is vertical in both models, neither of them would be stable if they were not augmented with a stabilizing policy rule. To compute impulse responses, both models were supplemented with the same monetary policy reaction function (a Taylor-type rule with coefficients 1.5, 0.5 and 0.5 for current inflation, the output gap and the lagged interest rate,

---

[10] In particular, Italy, having abandoned the ERM of the EMS in September 1992, re-joined it in late 1996 at the same bilateral exchange rate with the DM as the one irrevocably fixed when the euro was introduced in 1999.

respectively). As shown by the results reported in the figures, both models are stable, although even temporary shocks may result in very persistent deviations from equilibrium.[11]

The results show a number of similarities between the AEAM and the DEAM. First, in both models the effects of the shocks are rather long-lasting. Second, a shock to the aggregate supply equation induces a (dampened) oscillatory response of both inflation and the nominal interest rate. Third, the general pattern of responses is similar across models: e.g., a Phillips curve shock induces a contraction of output that reaches its maximum, in both models, in the third and fourth years after the shock; similarly, a (temporary) increase in the policy-controlled interest rate results in a temporary contraction of output that reaches its maximum at the end of the first year after the shock (moreover, the size of the contraction is not too dissimilar in the two models). Fourth, the response of inflation to a monetary policy shock comes with a further lag with respect to the reaction of output (the lag is somewhat more pronounced in the case of the DEAM).

The results, however, also signal several relevant differences. First, according to the DEAM the economy takes a longer time to get back to equilibrium after being hit by a shock. Second, the size of the responses is usually larger for the DEAM model (e.g., while the contractionary effect of an aggregate supply shock reaches a maximum, for both models, in the third and fourth years after the shock, the reaction of output in the DEAM is about three times as large as in the AEAM; also, the DEAM is more reactive to monetary policy as far as inflation is concerned, while it is somewhat less sensitive than the AEAM if one considers the effects on the output gap). Third, because of the overall more pronounced impact of aggregate supply and aggregate demand shocks on the economy, monetary policy is more activist in the DEAM, notwithstanding the fact that both models were augmented with exactly the same Taylor-type rule.

2.5 *Comparison with larger models*

The AEAM and DEAM representation of the functioning of the (three largest economies in the) euro area is admittedly rather crude. Would our conclusions below be dramatically different if larger models were considered that include a more detailed description of the entire euro-area economy (in particular, of the monetary policy transmission channels)? This

[11] For both the aggregate demand and aggregate supply equations, the shock amounts to one standard deviation of the corresponding estimation residuals. In the case of a monetary policy shock, the short-term interest rate is raised (for just one period) by 100 basis points.

section provides a tentative answer, albeit only an indirect one, to this question. It does so by exploring whether the main features of the AEAM and DEAM are in accordance with those of some of the main models of the euro area in use with policy-making and economic analysis institutions.[12]

The comparison focuses on the properties of the AEAM and DEAM that are most likely responsible for the results presented in Sections 3 and 4. From our viewpoint it is therefore of particular interest to ascertain whether the differences between the main properties of the AEAM and DEAM may be deemed representative of the effects of aggregation in larger and more detailed models.

From a qualitative viewpoint, the features of the AEAM and DEAM are reasonably similar to those of the (average of the) other models we consider. In most models, the full effects of a monetary policy shock on demand, output and prices unfold fully only with some lag. The impact is initially stronger on demand and production (reaching its maximum intensity in the course of the first two years); inflation tends to react more slowly (the largest fall occurring, in general, in the course of years 2 and 3). In the AEAM and DEAM, while the effects of the shock take about 1 year longer to unfold fully, the lag between the reaction of the output gap and that of inflation is about the same. Moreover, according to most disaggregate models the asymmetries in the individual-country responses to shocks are far from trifling (and are in fact sizeable according to both the Mark III model and the results reported in van Els, Locarno, Morgan and Villetelle (2001)), the only exception being the Quest.

To add some quantitative evidence to our analysis, let us focus on the IMF's Mark III (which includes a disaggregate euro-area block) and Mark IIIb (aggregate) models only. This choice is motivated by two main reasons: First, the Mark III and Mark IIIb were developed by the same modelling team and thus presumably share the same theoretical underpinnings and estimation techniques. Therefore, any difference between those two models may be interpreted

---

[12] Specifically, the discussion in the text reflects a comparison of the AEAM and DEAM with the following models: the ECB's Area Wide Model (Fagan, Henry and Mestre (2001); Dieppe and Henry (2002)); the IMF's Multimod Mark III (disaggregate) and Mark IIIb (aggregate) versions (Hunt and Laxton (2002)); the European Commission's Quest (Roeger and in't Veld (2002)); the National Institute's NiGem (Barrell, Gottschalk, Hurst, and Welsum, (2002)). Furthermore, the results presented in van Els, Locarno, Morgan and Villetelle (2001) — based on the models of the individual euro-area economies developed and maintained by the respective NCBs— were also taken into consideration. Since the information available is considerably less detailed than what would be needed for a systematic model comparison exercise (a notoriously difficult and tricky task), the evidence below should be viewed as being only indicative.

—more safely than it would be the case of other models— as largely stemming from what the data themselves indicate rather than, say, from differences in the theoretical framework or in the way the empirical models are specified and estimated. Second, while the Mark III and Mark IIIb differ in the way the euro area is modelled, there are only minor differences in the way the blocks for all other countries or regions are modelled. This is not the case for the rest of the models (e.g., while some of them include a description of the rest of the world, others do not).

Table 5 reports the effects of a 4-quarter 100 basis points nominal interest rate shock in the AEAM, DEAM, Mark IIIb and Mark III models.[13] Comparing the results for the latter pair of models, the effects of the shock on euro-area real GDP is initially stronger in the aggregate model; from year 3 onwards the differences between Mark III and Mark IIIb are negligible. By contrast, the fall in inflation is higher in the Mark III (disaggregate) than in the Mark IIIb (aggregate), the average difference between the two being between -0.05 (euro-area shock only) and -0.07 per cent (world-wide shock). Exactly the same pattern is found in the case of our models: the decline in the output gap is initially more pronounced in the AEAM; the differences between the AEAM and DEAM become negligible from year 3 onward. By contrast, the effects on inflation are sensibly more marked in the DEAM (by 0.04 per cent on average). While the comparison also highlights some differences (partly attributable to the fact that the effects of the shock in the Mark III and Mark IIIb models are by construction stronger and more front-loaded, at least as far as inflation is concerned, than in the AEAM and DEAM), the salient features associated with aggregate and disaggregate modelling approaches clearly

---

[13] In interpreting the results, it should be noted that the effects of the shock in the Mark III and Mark IIIb models are *a priori* likely to be stronger, at least as far as inflation goes, than in the AEAM and DEAM, as indeed confirmed by the figures (for the real economy effects, the sign of the distortion is not obvious). The main differences between the two sets of simulation results are the following: (a) both Mark III and Mark IIIb include a description of the rest-of-the-world economy, implying a number of spillover and feedback mechanisms that are absent in the AEAM and DEAM; (b) the real economy variable in our models is the output gap, while for the Mark III and Mark IIIb models only data for real GDP are available; (c) the simulation results for the Mark III and Mark IIIb refer to the case of endogenous euro exchange rates. More specifically, Hunt and Laxton (2002) report the effects of two monetary policy simulations: (i) shock to the euro-area policy interest rate only (resulting in a considerable initial appreciation of the effective euro exchange rate); (ii) shock to the world policy interest rate (the reaction of the euro exchange rate is here relatively muted; however, the monetary policy shock itself is implicitly stronger than in simulation (i), because it occurs world-wide). In either case, the effects of the shock are likely to be more pronounced in the Mark III and Mark IIIb models than in the DEAM and AEAM; (d) the two pairs of models are supplemented with different monetary policy reaction functions. Indeed, if one corrects the Mark III and Mark IIIb outcomes on the basis of the effects of exchange rate movements as estimated in other models (e.g., the simulation experiments in van Els, Locarno, Morgan and Villetelle (2001)), the numerical results commented in the text become very similar across the two pairs of models.

go in the same direction in both pairs of models: a disaggregate model tends to result in more pronounced effects of monetary policy on inflation, while the opposite applies to output.

3. Direct aggregation bias investigation

To introduce the issue of aggregation bias, we assume, following Theil's (1954) seminal work, that n observations on a set of variables (y, X) are available for m micro (disaggregate) units, and that, for each micro unit i, the following model holds:[14]

$$y_i = \beta_i X_i + e_i, \qquad i = 1, ..., m, \tag{1}$$

where e_i is a stochastic disturbance, with zero mean and variance $\sigma^2_{e_i}$. Let us consider the corresponding equation for the aggregate variables $y = \sum_{i=1}^{m} y_i$ and $X = \sum_{i=1}^{m} X_i$:

$$y = \beta X + e. \tag{2}$$

To derive the relationship linking the aggregate disturbance e to the micro disturbances e_i's it is convenient to introduce, as suggested by Theil (1954), a set of auxiliary regressions (being the projections of the micro exogenous variables on the aggregate ones):

$$X_i = b_i X + \eta_i \qquad i = 1, ..., m.$$

Substituting these expressions into eqs. (1) and (2) above and rearranging, the aggregate equation error can be written as:

$$e = \sum_{i=1}^{m} (e_i + \beta_i \eta_i). \tag{3}$$

The aggregate equation error may thus be viewed as being made up to of two components: the micro-equation disturbances and the aggregation error ("aggregation bias" in Theil's (1954) terminology).

---

[14] For the time being, the micro-equations are assumed to be correctly specified.

The variance of the aggregate equation error is given by:

$$\sigma_e^2 = \sum_{i=1}^{m} \sigma_{e_i}^2 + \sum_{i=1}^{m}\sum_{J=1}^{m} \sigma_{e_i\, e_j} + \sum_{i=1}^{m} \beta_i^2 \sigma_{\eta_i}^2$$

where $\sigma_{e_i\, e_j}$ is the covariance between the disturbances in the i-th and j-th disaggregate models; this expression may also be re-written as follows:

(4) $$\sigma_e^2 = \sum_{i=1}^{m} \sigma_{e_i}^2 + \sum_{i \neq j} \sigma_{e_i\, e_j} + \sum_{i=2}^{m} (\beta_i - \beta_1)^2 \sigma_{\eta_i}^2$$

Theil (1954) discusses two special cases in which the aggregation bias is nil: (i) micro homogeneity; (ii) compositional stability. Micro homogeneity holds when the parameters are identical in all micro equations; if this is the case, then the last term in eq. (4) vanishes. By compositional stability one means the case in which the ratio between each micro exogenous variable and the corresponding aggregate one is constant over time, which implies that the variance of the auxiliary regression errors in always exactly zero, and hence the last term in eq. (4) is zero as well. If neither of those two conditions hold, the aggregate relationship will have an additional error component as compared with the disaggregate ones (under the assumption of correct specification of the latter).

The most natural criterion to assess the relevance of the aggregation bias consists of comparing the sum of squared residuals associated with the disaggregate equations with that of the aggregate one, as first proposed by Grunfeld and Griliches (1960):

$$e_d^{'} e_d \leq e^{'} e, \qquad \text{where } e_d = \sum_{i=1}^{m} e_i.$$

Pesaran, Pierse and Kumar (1989) propose a slightly different criterion, correcting for small sample bias in the sum of square residuals. The criterion consists of comparing the aggregate and disaggregate standard errors; the former is given by:

$$s_a = e_a{'} e_a / (n - k_a);$$

for the latter, the estimator proposed by Pesaran, Pierse and Kumar (1989) is:

$$s_d = \sum_{i=1}^{m}\sum_{j=1}^{m}[n - k_i - k_j + tr(A_i A_j)]^{-1} e_i\prime e_j$$

where $A_l = X_l(X_l\prime X_l)^{-1}X_l'$; k_l = number of regressors in the l-th equation; n = number of observations for every micro unit.[15]

If the micro equations are correctly specified, s_d is less or equal to s_a by construction. In practice, s_d may be greater than s_a: such an occurrence may be taken as an indication of misspecification errors in the micro equations.

Lippi and Forni (1990) explicitly allow for dynamics —assuming an ARMAX representation to hold for both the aggregate equation and the disaggregate ones— and propose a more general representation of the aggregation bias. A detailed discussion of their results is beyond the scope of this section; suffice it to say that, in their set-up, aggregation is feasible under more general conditions than the two special cases identified by Theil (1954). Adopting a set-up similar to the one in Lippi and Forni (1990), Monteforte (2002) developed a Factor-Analysis-based approach for aggregation bias testing. In short, this approach rests on identifying the idiosyncratic components of the micro equations: intuitively, the larger those components, the less appropriate is the hypothesis of aggregability.

Grunfeld and Griliches's (1960) and Pesaran, Pierse and Kumar's (1989) criteria for the AEAM and DEAM are shown in Table 6. Clear signs of aggregation bias emerge on the basis of both criteria, particularly in the case of the aggregate supply (Phillips Curve) equations. Similar indications are provided by the criteria based on the estimated factor models (Table 7): the idiosyncratic components of the micro equation appears to be far from trifling.

Table 8 presents the average RMSE's of 1- to 8-steps-ahead forecasts for both the AEAM and the DEAM.[16] On the basis of the entire available sample, the DEAM sharply outperforms the AEAM, especially in the case of the aggregate supply equation. Out-of-sample results

[15] In addition to proposing the criterion presented in the text, Pesaran, Pierse and Kumar (1989) also derived a formal test for aggregation bias. A drawback of that test is that it may be computed only when the number of micro units is relatively large (this condition is not satisfied in the empirical application of this paper).

[16] The RMSE's were computed both using the whole sample (1978:Q1 to 2001:Q4) and only with out-of-sample data (1999.Q1 to 2001.Q4). Assessing the relevance of the aggregation bias on the basis of the relative forecast performance of the aggregate and disaggregate models has been advocated by Deutsch and Radler(1990) and Baltagi, Griffin and Xiong (2000).

are mixed: in particular, for relatively long forecast horizons the AEAM aggregate demand equation performs slightly better than the corresponding DEAM equations; in the case of aggregate supply equations, by contrast, the performance of the DEAM remains consistently better than that of the AEAM (note, however, that the number of out-of-sample observations is very small, and that there are reasons to believe that in-sample tests are more reliable than out-of-sample ones; see Inoue and Kilian (2002)).

Finally, Table 9 reports the results of formal tests of forecast encompassing, consisting of projecting the actual historical values of the aggregate endogenous variables of our models (the inflation rate and the output gap) over the respective projections based on the AEAM and DEAM, and testing whether the information provided by either model is encompassed by that supplied by its competitor. The hypothesis that the DEAM encompasses the AEAM cannot be rejected for the inflation rate equations, at least for relatively short forecast horizons; by contrast, the opposite hypothesis (the AEAM encompasses the DEAM) is always sharply rejected. The results are mixed for the output gap equations (e.g., for 1-step-ahead forecasts, the encompassing restriction are accepted for both models, although more markedly so for the DEAM).

To sum up, all available evidence points in the direction of rejecting the hypothesis of no aggregation bias —although with different degree of "sharpness"— especially in the case of the aggregate supply equations.[17]

4. Indirect aggregation bias investigation: The economic consequences of aggregate euro-area modelling

4.1 *Experimental design*

Having established that a disaggregate modelling approach of the euro area is statistically sounder than its alternative, it remains to be ascertained whether it is also preferable from an economic viewpoint. To do this, we compare the performance of two hypothetical European monetary policy-makers. The first policy-maker is assumed to rely on the AEAM; specifically, s/he is assumed to react to the state of the economy according to a reaction function whose parameters are optimised, given a standard specification of the loss function, under the set of

[17] Our results are consistent with those of other works that address the issue of aggregability of euro-area Phillips curves: Mayes and Virén (2000) argue that asymmetries across euro-area aggregate supply curves are very pronounced; Fabiani and Morgan (2003) also find evidence of asymmetry.

constraints given by the AEAM. By contrast, the second policy-maker's optimal reaction is computed on the basis of the DEAM.

Since the vector of state variables is different for the two models, the corresponding optimal instrument rules (i.e., rules that exploit all the information provided by the whole set of state variables, which we label FO rules) would not be easily comparable. We thus impose that both rules belong to the Taylor-type family (i.e., the arguments of both rules are the current area-wide annualised quarterly inflation, the average output gap and the lagged value of the policy instrument only); for the sake of making the comparison as fair as possible, we further require that the DEAM-based rule only respond to area-wide aggregates.[18]

In both cases, a standard time-separable quadratic loss function is assumed, its arguments being the euro area average annual inflation rate, the output gap and a term that attaches a cost to the volatility of the policy instrument; i.e.:

$$L_t = E_t \sum_{\tau=0}^{\infty} \delta^{\tau} [(\overline{\pi}_{t+\tau})^2 + \lambda \cdot y_{t+\tau}^2 + \mu \cdot (\Delta i_{t+\tau})^2] \tag{5}$$

where δ is a discount factor, and λ and μ are parameters that reflect the policy-maker's preferences (the weight on deviations of inflation from its target is normalized to 1); $\overline{\pi}_{t+1}$ is the (euro-area average) year-on-year consumer inflation rate (i.e., $\overline{\pi}_{t+1} = \pi_{t+1} + \pi_t + \pi_{t-1} + \pi_{t-2}$); y_{t+1} is the (average) output gap; i_{t+1} is the short-term policy-controlled interest rate. It is worth stressing that our specification of the loss function implies that the euro-area policy-maker is only interested in euro-area average outcomes, and hence is consistent with the official Eurosystem's view of the monetary policy objective and strategy.

For $\delta \to 1$ the sum in eq. (5) becomes unbounded; however, following Rudebusch and Svensson (1999), p. 215, "the value of the inter-temporal loss function approaches the infinite sum of the unconditional means of the period loss function"; this implies that one can "interpret the inter-temporal loss function as the unconditional mean of the period loss function," which is given by the weighted sum of the unconditional variances of the target variables:

$$L_t = \text{var}[\overline{\pi}_t] + \lambda \cdot \text{var}[y_t] + \mu \cdot \text{var}[\Delta i_t] \tag{6}$$

---

[18] The performance of rules that allow the policy-maker to respond to country-specific variables is investigated in Angelini, Del Giovane, Siviero and Terlizzese (2002).

In the following we adopt the loss function defined as in eq.(6). The quest for optimal policy was repeated with a wide range of values for λ and μ, ranging from a case in which the monetary policy-maker is only interested in inflation ($\lambda = \mu = 0$) to the opposite extreme, in which the policy-maker attaches a comparatively high cost to deviations of the output gap from its equilibrium value (zero) and to the volatility of the policy-controlled interest rate ($\lambda = 5; \mu = 3$).[19]

The two competing rules may thus be synthetically described as follows:[20]

AEAM-based rule

$$\min_{\gamma_1^A,\gamma_2^A,\gamma_3^A} E_t \sum_{\tau=0}^{\infty} L_{t+\tau} = \min_{\gamma_1^A,\gamma_2^A,\gamma_3^A} E_t \sum_{\tau=0}^{\infty} [(\overline{\pi}_{t+\tau})^2 + \lambda \cdot y_{t+\tau}^2 + \mu \cdot (\Delta i_{t+\tau})^2]$$

s.to:
- AEAM (see Par. 2.1)
- $i_t = \gamma_1^A \cdot (\overline{\pi}_t) + \gamma_2^A \cdot y_t + \gamma_3^A \cdot i_{t-1}$

and:

DEAM-based rule

$$\min_{\gamma_1^M,\gamma_2^M,\gamma_3^M} E_t \sum_{\tau=0}^{\infty} L_{t+\tau} = \min_{\gamma_1^M,\gamma_2^M,\gamma_3^M} E_t \sum_{\tau=0}^{\infty} [(\overline{\pi}_{t+\tau})^2 + \lambda \cdot y_{t+\tau}^2 + \mu \cdot (\Delta i_{t+\tau})^2]$$

s.to:
- DEAM (see Par. 2.2)
- $i_t = \gamma_1^M \cdot (\overline{\pi}_t) + \gamma_2^M \cdot y_t + \gamma_3^M \cdot i_{t-1}$

Let us now tackle the crucial issue of how the performance of these two rules may be compared.

The statistical evidence presented in Section 3 indicates that the DEAM provides a more reliable description of the functioning of the euro-area economy than the AEAM. Consistently with that evidence, we assess the welfare function values associated with the optimised DEAM-based and AEAM-based rules using the DEAM and the corresponding variance-covariance matrix of residuals.

---

[19] The ranges chosen for the loss function parameters are similar to the ones typically assumed in the literature; see, e.g., the papers collected in Taylor (1999).

[20] In our benchmark experiments, the loss function and the rules are specified in terms of annual quarterly inflation, as in much of the literature (see, e.g., the contributions in Taylor (1999)). Our results below do not change much if annualised quarterly inflation is used in the loss function and/or in the rules (actually, our conclusions would be even somewhat sharper).

Given the way in which the two rules are compared, it is clear that the AEAM-based rule cannot outperform the DEAM-based one by construction. However: (i) it is clearly the case that, for any comparison to be sensible, the performance of the two rules must be computed on the basis of a common framework. Given the empirical evidence presented in Section 2.5, the DEAM stands out as a more reasonable choice than the AEAM; (ii) we do not simply rank the two rules, but provide a measure of the “welfare distance” between the two (”By how much does the DEAM-based rule outperform the AEAM-based one?”); while the ranking of the two rules is implicit in the experimental design, there is no *a priori* reason why the under-performance of the AEAM-based rule should not be trifling; (iii) furthermore, we not only compute the welfare distance between the two rules but we also assess its significance; (iv) finally, the assumption that the DEAM provides a perfectly accurate description of the functioning of the euro-area economy will be relaxed below, formulating the weaker alternative assumption that the DEAM only provides a statistically acceptable representation of the economy (thus, we not only estimate the size of the difference in the performances of the two rules, but also its robustness with respect to changing the assumption that the data generating process coincides with the DEAM).

As a benchmark, we also compute, on the basis of the DEAM, the FO rule and the associated optimised target variances and welfare loss.[21] This third set of results is used to compare the gains attainable with the DEAM-based rule with the (larger) ones that could be achieved by relying on the rule that, by definition, performs best within the DEAM.

4.2 *The results*

4.2.1 *Basic findings*

The main results of our experiments are shown in Figures 6 and 7 and Table 10.

Let us focus first on the final outcome of the two competing rules (Figure 6). The top chart of the figure reports the percentage increase in the optimized value of the objective function if the AEAM-based rule is followed instead of the DEAM-based one. The welfare losses are far from negligible, being smallest (just abowe 10 per cent) in the neighborhood

---

[21] For this purpose, we first derive the state-space representation of the DEAM, and then solve a standard stochastic linear regulator problem (see Chow (1970), Sargent (1987), and, for an application to the issue of optimal monetary policy design, Rudebusch and Svensson (1999)). For the sake of brevity, we omit the technical details here.

of pure inflation targeting. In the case of pure inflation targeting, the AEAM-based rule does not actually succeed in stabilizing the DEAM, thus resulting in an explosive unconditional variance-covariance matrix, so that the welfare loss due to following the AEAM-based rule is infinite. The loss exceeds 20 per cent if the weights of either the output gap or the volatility of the policy instrument are relatively high. The key message given by the figure is that ignoring the structural differences among the euro area economies, and so adopting a model that treats them as a single and homogeneous "whole," may lead to a sizeable worsening of the performance of monetary policy, particularly if the policy-maker only cares about inflation.

The bottom chart of Figure 6 reports the same results in relative terms, using the FO rule as a benchmark;[22] the chart suggests that the hypothetical policy-maker relying on the AEAM would go a long way towards further worsening the distance (measured in terms of welfare) between the DEAM-based and the FO rules. Specifically, for $\mu = 0$ the AEAM-based rule implies an additional loss comprised between almost 25 and over 50 per cent. For most values of λ and μ the additional loss amount to at least 20 per cent. Thus, not only is the size of the gains that can be attained with a multi-country modelling approach not negligible, but, adopting the true optimum rule as a benchmark, those gains are considerable.

The results can be assessed directly in terms of the optimised unconditional standard deviations of inflation, the output gap and interest rate changes. This is done in Figure 7, showing the optimal inflation/output gap frontier (in terms of optimized standard deviations of those variables) for the AEAM-based, DEAM-based and FO rules. The frontiers have been computed, for given μ, by letting λ take a grid of values between 0 (north-west) and 5 (south-east). While the frontier associated with the FO rule is positioned considerably to the south-west with respect to the frontier associated with the DEAM-based rule, the latter consistently attains a combination of inflation and output gap volatility that is sizeably better than that of the AEAM-based rule. For no combination of preference parameters do the performances of the DEAM-based and AEAM-based rules come close to one another.

Can one trace these outcomes back to the properties of the different optimal rules, and in particular to the optimised parameters on inflation, the output gap and the lagged interest rate in the monetary policy reaction functions? A selection of the latter are presented in Table

---

[22] The FO rule differs from either of the other two, in that the arguments of the rule are the current and lagged quarter-on-quarter country-specific inflation rates, rather than the average year-on-year inflation rates.

10. Since the arguments of the FO rule are the quarter-on-quarter inflation rates, as opposed to the year-on-year inflation rates of the AEAM- and DEAM-based rules, the latter were re-computed under the assumption that the policy rate reacts to quarterly inflation and aims at stabilizing annualised quarter-on-quarter inflation. Thus, the FO, AEAM-based and DEAM-based rules reported in Table 10 are fully comparable. The FO rule depends on the complete set of the 15 state variables in the DEAM: the latter set comprises inflation and output gap in the various countries for different lags. For ease of comparison, the coefficient on inflation reported in Table 10 is given, for the FO rule, by the sum of the value of all coefficients that the rule assigns to inflation in all countries and for all lags; similarly for the output gap. At leat two features are noteworthy in that table: First, the optimized parameters of the DEAM-based rule come generally much closer to the corresponding optimized parameters in the optimal instrument rule, while those of the AEAM-based rule are often distant. Consider, for instance, the first set of loss function weights ($\lambda = \mu = 0.1$): the fully optimal parameter on inflation is 2.93; for the DEAM-based rule, the corresponding value is 2.46, while for the AEAM-based rule it is as low as 1.66. Similarly for the output gap, and for all other combinations of loss function weights. Second, the AEAM-based rule is consistently not "reactive" enough to either inflation or the output gap compared with the other two rules.

4.3 *Testing the significance of the results*

The results presented so far suggest that, were the euro area policy-maker to formulate her/his decisions on the basis of the indications of an aggregate area-wide model, s/he would likely incur non-negligible welfare losses as opposed to the case in which s/he relied on a multi-country tool. However, while the size of the welfare gains that are at the stake are *prima-facie* rather large, it remains to be established whether they are significantly so.

To tackle this issue (which, as far as we know, has never been addressed in the literature on optimal monetary rules), we perform two stochastic simulation exercises. In the first we compute, for a (large) number of realizations of the stochastic disturbances (drawn from the distribution of the estimation residuals), the value of the objective function conditional on following the AEAM-based or the DEAM-based rule, alternatively.[23] The second exercise is similar, except that we sample from the stochastic distribution of the estimated parameters.

Focusing on the first exercise, we extract 1,000 replications from the set of estimated residuals and simulate the model, for each replication, under either one or the other of the two competing rules. Each replication consists of 800 realizations of the shocks for the six stochastic equations in the model, one realization per period. Although the model is simulated for 800 periods, only the average outcomes in the last 400 periods are used to evaluate the objective function. This is done for the purpose of preventing the results from being biased by the initial conditions (we begin simulating the model from a situation of equilibrium; by scrapping the first 400 results, the simulated variance of the objective variables should, and indeed do, provide a reasonable approximation of their unconditional variance). This we repeat for several preference parameter combinations.

For all combinations of preference parameters, the DEAM-based rule delivers a better outcome than the alternative in the overwhelming majority of replications (always at least 80 per cent; see the top chart of Figure 8). Hence, not only is the gain large on average, but is also systematic. The bottom chart of Figure 8 also shows that, for most combinations of preference parameters, the welfare gain associated with the DEAM-based rule amounts to at least 20 per cent of the loss associated with the AEAM-based rule in 20 to 40 per cent of all replications, with the exception of a neighborhood around (but not including) $\lambda = \mu = 0$ (the lowest figure being just over 10 per cent).

We also formally tested the hypothesis that the average welfare loss associated with following the DEAM-based rule is lower than the average loss with the AEAM-based rule (the test is a one-sided test based on comparing the averages of the objective function outcomes associated with either one or the other of the two rules for all 1,000 replications).[24] The results are overwhelmingly supportive of the hypothesis: for all combinations of policy parameters the tail probability of the test is virtually zero (the values of the test for all λ's and μ's are shown in Figure 9, together with the 1 per cent critical value).

---

[23] A different experiment could consist of sampling from the error distribution, re-estimating the model for each replication, and re-computing the rules each time. Such an experiment, however, would by construction result in an under-performing AEAM-based rule for each and any replication, which is not necessarily the case here.

[24] The test is based on the standard statistic for the equality of the means of normally distributed variables.

Overall, these results indicate that the gain associated with adopting the DEAM-based rule is not only large, but also significantly so and, moreover, systematic.

The second exercise explicitly accounts for the stochastic nature of the estimated model coefficients. In the previous section, as well as in virtually all the literature on policy rules, the model used to derive and appraise the optimal rules is assumed to describe accurately the functioning of the economy, and the stochastic nature of the estimated model parameters is systematically ignored. Actually, the most one could argue is that with a certain probability the "true" model parameters lie in the neighborhood of the estimated ones. It could then be that their variance-covariance matrix is so "large" as to make whatever differences one finds between the performances of competing rules statistically irrelevant. This exercise can be interpreted as a check on the robustness of our main results: we check whether the latter would still hold were the "true" model somewhat different from the one used to derive the two rules. The need for such a check is particularly acute in the case at hand, since we compare the performance of the DEAM-based and AEAM-based rules by computing the respective loss functions under the assumption that the DEAM is the true model, an assumption that, while justified by the empirical findings of Section 3, has a clear implication as to the ranking of the two rules (although, as remarked above, it says nothing about their distance).

In more detail, to account for the variability of the estimated coefficients we extract 5,000 replications from the empirical distribution of the estimated DEAM coefficients and, without re-computing the DEAM-based and AEAM-based rules, we compute, for each replication of the model coefficients, the associated loss function (almost half of the replications had to be discarded, as they produced explosive estimates of the unconditional variance-covariance matrix with either the DEAM-based or the AEAM-based rules, and in general with both; see below for more details). We then examine the distribution of the loss function under the two rules. These steps are repeated for 77 combinations of values of the preference parameters λ and μ.

A first set of results is shown in Figure 10 (top chart). It can be seen that in (almost) 70 to 85 per cent of all the "alternative worlds" that are plausible given the estimate of the DEAM, the DEAM-based rule does strictly better than the AEAM-based one for any combination of preference parameters. Hence, coefficient variability is not such as to jeopardize our conclusions above. For 20 to 40 per cent of the replications (depending on the particular combination of preference parameters) the DEAM-based rule delivers a reduction of the loss function of at least 20 per cent (bottom chart of Figure 10). Overall, it seems safe to conclude that the results are systematic across "alternative worlds" (as long as the latter are statistically compatible with the estimated DEAM), and the gains are large relatively often.

Finally, as in the exercise above, we formally test the hypothesis that the average (across replications) welfare loss associated with the DEAM-based rule is lower than the average loss obtainable with the AEAM-based one (Figure 11). With a confidence level of 5 per cent, only for 2 combinations of the preference parameters (less than 3 per cent of the cases) one is not able to accept the null hypothesis (and even then is the rejection only marginal). For most combinations of preference parameters (62 out of 77, i.e., over 80 per cent) the null hypothesis is accepted at the confidence level of 1 per cent. The few (marginal) rejections reflect the fact that some of the individual drawings of the parameters of the model result in extreme outcomes, which are arguably not fully realistic (e.g. the policy instrument becomes virtually ineffective). Indeed, if those additional (few) outliers are eliminated, the tail probability of the test is always much lower than 1 per cent for all preference parameters.

Overall, these results clearly indicate that, whatever the "true" data generating process, the DEAM-based rule tends to be significantly better than the AEAM-based alternative (provided that our multi-country model is a reasonable approximation of the data generating process, or at least a more reasonable one than the AEAM). Not only is the welfare loss associated with the AEAM-based rule large, but it is also statistically significant and generally "robust" to parameter uncertainty.

4.4 *What could be ahead?*

Despite the evidence presented in Section 2.3 (showing no signs of instability in the DEAM and AEAM models in the recent past), in this section we explore how the comparison between the AEAM-based and DEAM-based rules would be affected were more symmetry of stochastic disturbances to prevail among the euro area countries than detected in the past.

According to the results in Section 2.3, not much convergence of structures can be observed in the recent past. By contrast, the DEAM variance-covariance matrix computed with the 20 out-of-sample observations (from 1997.Q1 to 2001.Q4) suggests that the symmetry of shock has somewhat risen since 1996, in particular for the disturbances in the AD equations (note, however, that the number of out-of-sample observations is relatively small). This variance-covariance matrix was used to compute the value of the objective function associated with the AEAM-based and DEAM-based rules obtained on the basis of the data available up to the end of 1996 (this amounts to investigating what would have happened if optimal monetary policy rules computed on the basis of pre-euro models had been used in the early stages of the euro era). The results tend to be about as unfavorable to the AEAM-based rule as in the benchmark experiment, with the exception of a very small neighborhood around $\lambda = \mu = 0$ (but not including the latter point). To sum up, the recent evidence supports the claim that the gains associated with using the DEAM-based rule, as opposed to the AEAM-based rule, have not started to diminish yet. One tends to conjecture that they will remain non-negligible at least in the near future.

But what could happen in the far future? To answer this question, we assume that some sort of convergence in the stochastic processes that generate the disturbances of the DEAM will occur in the future.

Of course, convergence of disturbances might occur (if at all) in many different ways: all countries' shocks could become similar to some average of what they are now; the stochastic structure of the shocks of smaller countries could become more and more similar to that of the largest one; or the final outcome of the convergence process could well be something that does not at all resemble the current situation. In fact, there is no reason why convergence should necessarily take place; moreover, there is no compelling evidence that much convergence has taken place in the long run-up to the euro area.[25]

This leaves us with many (indeed, too many) ways to model convergence, and we have no clear-cut criterion to offer as to which of them could be more plausible. Nevertheless, we believe that exploring the sensitivity of our results to some form of convergence can be informative, even if the eventual convergence process were to follow a different path.

To proceed, we will assume that countries that become more intimately tied to one another tend to share the same shocks, and influence those common shocks proportionately to

[25] Eichengreen (1997) and Demertzis and Hughes Hallett (1998), have tackled the issue of the symmetry of the shocks to the European economies, or lack thereof; their empirical evidence shows that, although the European economies have followed rather similar policies in recent years, there is little evidence of a strengthening of the degree of symmetry of the disturbances affecting the various economies.

their relative size (the largest country exerting a comparatively stronger effect on the common shocks than the other two, and so on).

More in detail, we take full convergence of aggregate demand shocks to mean that the disturbances in the aggregate demand equation become exactly the same in all countries (hence, the cross-country correlation equals 1). As in De Grauwe and Piskorski (2001), we assume that, once full convergence has been reached, the common variance (as well as covariances) is given by the square of a weighted average of the historical estimated standard deviations:

$$\sigma^2_{y|FC} = (\omega_{y_G}\sigma_{y_G} + \omega_{y_F}\sigma_{y_F} + \omega_{y_I}\sigma_{y_I})^2 \tag{7}$$

where $\sigma^2_{y|FC}$ denotes the variance of the common AD shock under convergence; $\sigma_{y_G}, \sigma_{y_F}, \sigma_{y_I}$ are the estimated standard deviation of AD disturbances in the three countries; $\omega_{y_G}, \omega_{y_F}, \omega_{y_I}$ are the GDP weights of the three countries.

We also consider the possibility of partial convergence, which we assume to be parameterized by ξ_{AD}, ranging from 0 (no convergence) to 1 (full convergence). For any given choice of the ξ_{AD} parameter, the corresponding elements of the variance-covariance matrix of the disturbances are given by:

$$\sigma^2_{y_i|PC} = \xi_{AD}\sigma^2_{y|FC} + (1 - \xi_{AD})\sigma^2_{y_i} \tag{8}$$

$$\sigma_{y_j y_i|PC} = \xi_{AD}\sigma_{y_i|PC}\sigma_{y_j|PC} \tag{9}$$

for all i, j, so that the correlation of shocks among countries is given by ξ_{AD} itself.[26]

Full and partial convergence of aggregate supply disturbances are defined in a similar way, with the convergence process now parameterized by ξ_{AS}.

Turning to the results, under the extreme assumption that there are only two stochastic processes in the euro area (specifically, one stochastic process driving Phillips curve shocks, and one driving aggregate demand shocks, common to all countries), the under-performance of the AEAM-based rule is considerably attenuated. Figure 12 reports, for the case $\lambda = \mu = 1$

[26] It would, of course, be possible to introduce the further complication that the speed of convergence is not the same for all countries. However, for the sake of simplicity we ignore that possibility. Let us just remark that our concept of partial convergence tends to make cross-country heterogeneity disappear more smoothly than it would be conceivably possible.

(the results are similar for all other values of the policy parameter), the additional welfare loss entailed by adopting the AEAM-based rule rather than the DEAM-based one, as the degree of similarity of supply- and demand-side shocks across countries increases (in the figure, the loss found above for the case of no convergence is set equal to 100).[27] With a high degree of convergence, there remains virtually no scope at all for using the DEAM-based rule. Note, however, that a non-negligible degree of (uniform) convergence is needed before the loss associated with using the AEAM-based rule becomes relatively small. Examining what happens if the pace of convergence differs on the supply- and demand-sides (i.e., looking at the off-diagonal elements in the figure), one concludes that demand-side convergence without supply-side convergence would not be very effective in reducing the additional welfare loss associated with the AEAM-based rule (with $\xi_{AS} = 0$, the loss would not even halve even if demand-supply shocks were to become exactly identical in all countries). Whether these features are empirically robust seem worth investigating further in future work.

5. Concluding remarks: What implications for euro area econometric modelling?

The results presented in this paper support the conclusion that heterogeneity in the economic structures of the countries participating in the euro area is not only statistically detectable but, perhaps more importantly, economically relevant. Specifically, monetary policy in the euro area is likely to be more effective if the econometric tools used to help monetary policy decisions acknowledge the structural differences among the various economies in the area, and so do not model aggregate euro area data as if they referred to one single, relatively homogeneous economy.

The welfare losses associated with adopting and aggregate-model-based rule are not only sizeable but also highly significant.[28] Moreover, our results are generally robust with

---

[27] All rules perform less satisfactorily than in the set of experiments where the historical variance-covariance matrix was assumed to hold, the worsening being, in general, more pronounced for the DEAM-based rule (and for the optimal instrument one) than for the AEAM-based rule. A general worsening of the optimized losses should indeed be expected: in the latter experiment the shocks are perfectly correlated, while the historical ones are virtually independent, and hence do not tend to reinforce each other.

[28] Moreover, according to Angelini, Del Giovane, Siviero and Terlizzese (2002), the optimized value of the loss function could be further reduced if the single monetary policy were to exploit fully the available national information (by not simply relying on a DEAM, but also reacting to national information). Combining these results with ours, one can appreciate the full distance between a “pure aggregate approach” (using an AEAM model and computing an AEAM-based rule) and a “full multi-country one” (using a DEAM and allowing for the policy instrument to react to country-specific variable): the total reduction in the optimized value of the loss function is always in the neighborhood of 50 per cent or more.

respect to model parameter variability. Finally, while our investigation of possible instabilities of the model in the most recent past does not suggest that euro area economies are becoming increasingly similar to one another, we nevertheless probe what could happen if convergence occurred in the future. We find that sizeable convergence has to occur before our conclusions no longer apply.

Our conclusions are apparent in our simplified model for the three main countries. Arguably they would be all the more supported by an analysis that were to include all 12 economies in the area —possibly with a more sophisticated and detailed description of their functioning than is provided by the simple aggregate demand-Phillips curve models we use. In particular, a fully-fledged model for each individual country could pay closer attention to country-specific institutional features, labor market arrangements, tax structures, etc., thereby presumably resulting in a more pronounced degree of asymmetry amongst country models. In this respect, it seems legitimate to conjecture that the reduction in the welfare losses that we measure is likely to be a lower bound estimate.

Our results make a clear case for relying on a multi-country modelling approach when offering advice in support of the single monetary policy, and suggest that a line of research worth pursuing is a systematic investigation of the aggregation bias (both its size and its nature) that is likely to affect aggregate (area-wide) estimated relationships and their effects on optimal policies.

References

Altissimo, F., S. Siviero and D. Terlizzese (1999), “How Deep are the Deep Parameters,” Banca d’Italia, Temi di Discussione, No. 354.

Angelini, P., P. Del Giovane, S. Siviero and D. Terlizzese (2002), “Monetary Policy Rules for the Euro Area: What Role for National Information?,” Banca d’Italia, Temi di Discussione, No. 457.

Angeloni, I. and L. Dedola (1998), “From the ERM to the Euro: New Evidence on Economic and Policy Convergence among EU Countries,” ECB, Working Paper No. 5.

Angeloni, I., A. Kashyap, B. Mojon and D. Terlizzese, (2002), “Monetary Transmission in the Euro Area: Where Do We Stand?,” ECB, Working Paper No. 114.

Barker, T.S. and M.H. Pesaran (1990), “Disaggregation in Econometric Modelling,” London and New York, Routledge.

Barrell, R, Gottschalk, S. Hurst, I. and D. V. Welsum, (2002) ”Macroeconomic Policy in Europe” paper presented at the CEPR/ZEI Conference ”Empirical Models of the Euro Economy Euroconference1: Macro Performance”, Bonn, 7/8 June.

Baxter, M. and R.G. King (1995), “Measuring Business Cycles Approximate Band-Pass Filters for Economic Time Series,” National Bureau of Economic Research, Working Paper, No. 5022.

Bayoumi, T. and B. Eichengreen (1993), “Shocking Aspects of European Monetary Integration,” in Torres, F, and F, Giavazzi (eds), ”Adjustment and Growth in the European Monetary Union,” Cambridge University Press, Cambridge.

Beyer, A., J.A. Doornik and D.F. Hendry (2000), “Reconstructing Aggregate Euro-Zone Data,” Journal of Common Markets Studies, Vol. 38, No. 4.

Chow, G.C. (1970), “Analysis and Control of Dynamic Economic Systems,” Wiley, New York.

Ciccarelli, M. and A. Rebucci (2002), “Asymmetries in the Transmission Mechanism of European Monetary Policy: is There Heterogeneity? Is it Changing over Time?,” IMF Working Paper, No. 02/54 .

Clements, B., Z.G. Kontolemis and J. Levy (2001), “Monetary Policy under EMU: Differences in the Transmission Mechanism?,” IMF Working Paper, WP/01/102, August.

Dedola, L. and F. Lippi (2000), “The Monetary Transmission Mechanism: Evidence from the Industries of Five OECD Countries,” Banca d’Italia, Temi di Discussione, No. 389.

De Grauwe, P., and T. Piskorski (2001), “Union-Wide Aggregates versus National Data Based Monetary Policies: Does it Matter for the Eurosystem?,” CEPR Discussion Paper Series, No. 3036.

Demertzis M., A. Hughes Hallett and O. Rummell (1998), “Is a 2-Speed System in Europe the Answer to the Conflict between Germany and the Anglo-Saxon Models of Monetary Control?,” in Black., S. and M. Moersch (eds), “Competition and convergence in financial markets - The German and Anglo-American models,” Elsevier, North-Holland, New York.

Dieppe, A., and J., Henry (2002), ”The ECB Model of the Euro Economy”, paper presented at the CEPR/ZEI Conference ”Empirical Models of the Euro Economy Euroconference1: Macro Performance”, Bonn, 7/8 June.

Dornbusch, R., C. Favero and F. Giavazzi (1998), “A Red Letter Day,” CEPR Discussion Paper, No. 1804.

ECB (2001), “A Guide to Eurosystem Staff Macroeconomic Projection Exercises,” June.

Eichengreen, B. (1997), “European Monetary Unification: Theory, Practice, and Analysis,” Cambridge, Massachussetts, MIT Press.

van Els, P., A. Locarno, J. Morgan and J.P. Villetelle, (2001), “Monetary Policy Transmission in the Euro Area: what do Aggregate and National Structural Models Tell Us?,” ECB Working Paper, No. 94.

Fabiani, S. and J. Morgan (2003), “Aggregation and Euro-Area Phillips Curves,” ECB Working Paper, forthcoming.

Fagan, G., J. Henry and R. Mestre (2001), “An Area-Wide Model (AWM) for the Euro Area,” ECB Working Paper No. 84.

Farmer, R. (1991), “The Lucas Critique, Policy Invariance and Multiple Equilibria,” Review of Economic Studies, pp.321-32.

Favero, C.A. and D.F. Hendry (1992), “Testing the Lucas Critique: a Review,” Econometric Reviews, pp.265-66.

Grunfeld, Y. and Z. Griliches (1960), “Is Aggregation Necessarily Bad?,” The Review of Economics and Statistics, Vol. 17, No. 1.

Guiso, L., A. Kashyap, F. Panetta and D. Terlizzese (1999), “Will a Common European Monetary Policy Have Asymmetric Effects?,” Federal reserve Bank of Chicago, Economic Perspective, pp. 56-75.

Hendry, D.F. (1988), “Testing Feed-Back vs. Feedforward Econometric Formulations,” Oxford Economic Papers, pp.312-49.

Hughes Hallett, A. and L. Piscitelli (1999), “EMU in Reality: the Effect of a Common Monetary Policy on Economies with Different Transmission Mechanism,” CEPR Discussion Paper Series, No. 2068.

Hunt, B. and D. Laxton (2002), ”The euro economy in MULTIMOD”, paper presented at the CEPR/ZEI Conference ”Empirical Models of the Euro Economy Euroconference1: Macro Performance”, Bonn, 7/8 June.

Inoue, A. and L. Kilian (2002), ”In-Sample or Out-Of-Sample Tests of Predictability: Which one Should We Use?”; CEPR Discussion Paper N.3671.

Labhard, V., O. Weeken and P. F. Westaway (2001), “Characterising Macroeconomic Behaviour in the Euro Area: Are Statisticians More Influential than Econometricians?,” paper presented at the Banca d’Italia/CEPR Conference “Monitoring the Euro Area Business Cycle,” Rome, 7/8 September.

Levin, A. and J.C. Williams (2002), “Robust Monetary Policy with Competing Reference Models,” mimeo.

Lippi, M. and M. Forni (1990), “On the Dynamic Specification of Aggregated Models,” in Barker, T.S. and M.H. Pesaran (1990).

Lucas, R.E. (1976), “Econometric Policy Evaluation: A Critique,” in K. Brunner and A. Meltzer (eds), “The Phillips Curve and Labor Market,” Amsterdam, North-Holland.

Mayes, D.G. and M. Virén (2000), “Asymmetry and the Problem of Aggregation in the Euro Area,” Bank of Finland Discussion Paper No. 11/2000.

Monteforte, L. (2002), “Cross Country Aggregation of Econometric Models: The Case of the Euro Area,” Banca d’Italia, mimeo.

Monteforte, L. and S. Siviero (2002), “The Economic Consequences of Euro Area Modelling Shortcuts,” Banca d’Italia, Temi di Discussione, No. 458.

Pesaran, H.M., R. G. Pierse and M.S. Kumar (1989), “Econometric Analysis of Aggregation in the Context of Linear Prediction Models,” Econometrica, Vol. 57, No. 4.

Ramaswamy, R. and T. Sloek (1997), “The real effects of monetary policy in the European Union: what are the differences?,” IMF Working Paper, No. 97/160.

Roeger, W. and J. in ’t Veld (2002), ”Some selected simulation experiments with the European Commission’s QUEST model”, paper presented at the CEPR/ZEI Conference ”Empirical Models of the Euro Economy Euroconference1: Macro Performance”, Bonn, 7/8 June.

Rudebusch G. D. and L. E. O.Svensson (1999), “Policy Rules for Inflation Targeting,” in J. B. Taylor (1999).

Sargent T. J. (1987), “Dynamic Macroeconomic Theory,” Harvard University Press, Cambridge.

Sims, C. (1982), “Policy Analysis with Econometric Models,” Brookings Paper on Economic Activity, pp.106-152.

Taylor, J. B., ed. (1999), “Monetary Policy Rules,” The University of Chicago Press, Chicago.

Theil, H. (1954), “Linear Aggregation of Economic Relations,” Amsterdam, North-Holland.

Winder, C. C. A. (1997), “On the Construction of European Area-Wide Aggregates – A Review of the Issues and Empirical Evidence,” Irving Fisher Committee on Central-bank Statistics, IFC Bulletin, No. 1.

Table 1

ESTIMATE OF THE AEAM

Input from:	Equation for: π	Equation for: y
π	0.652 [-1] (0.075) 0.348 [-4] (restr.)	
y	0.088 [-1] (0.035)	0.769 [-1] (0.060)
r		$-.050$ [-2] (0.022)
R^2	0.874	0.715
$\bar{R}^2$	0.869	0.704
σ	0.286	0.487
DW	2.209	1.800

In parentheses: standard error of the coefficients.

In brackets: lag with which the variables enter the equations.

Table 2

ESTIMATE OF THE DEAM

		Equations for: Germany		Equations for: France		Equations for: Italy	
Input from:		π	y	π	y	π	y
Germany	π	0.292 [-1] (0.089) 0.600 [-4] (0.069)		0.063 [0] (restr.)		0.036 [0] (restr.)	
	y	0.095 [-1] (0.036)	0.785 [-1] (0.062)				0.173 [0] (0.058)
	r		-0.073 [-2] (0.038)				
France	π	0.108 [0] (restr.)		0.937 [-1] (0.044)			
	y			0.022 [-2] (0.012) 0.022 [-3] (0.012) 0.022 [-4] (0.012) 0.022 [-5] (0.012)	0.838 [-1] (0.052)		
	r				-0.036 [-2] (0.015)		
Italy	π					0.964 [-1] (0.010)	
	y					0.064 [0] (0.028)	0.657 [-1] (0.061)
	r						-0.038 [-1] (0.016)
	R^2	0.514	0.635	0.902	0.730	0.960	0.752
	$\overline{R}^2$	0.483	0.622	0.894	0.720	0.958	0.740
	σ	0.411	0.799	0.332	0.443	0.259	0.490
	DW	2.160	2.059	2.050	1.888	2.024	1.815

In parentheses: standard error of the coefficients.

In brackets: lag with which the variables enter the equations.

Table 3

OUT_OF SAMPLE STABILITY, AEAM, 1997.Q1-2001.Q4

Equation	F-value	Tail probability
AS	0.91	57.49
AD	0.57	92.21

Table 4

OUT_OF SAMPLE STABILITY, DEAM, 1997.Q1-2001.Q4

Equation		F-value	Tail probability
Germany	AS	0.89	60.12
	AD	0.37	99.23
France	AS	0.67	84.46
	AD	0.76	75.38
Italy	AS	0.52	95.12
	AD	0.68	83.59

Table 5

Comparison of MARK III/MARK III B models and DEAM/AEAM

(100 b.p. shock to policy short-term interest rate for 4 quarters)

Models \ Years	Inflation					Real activity [1]				
	1	2	3	4	5	1	2	3	4	5
AEAM	0.00	-0.09	-0.14	-0.14	-0.11	-0.09	-0.16	-0.05	0.00	0.00
DEAM	-0.01	-0.08	-0.16	-0.21	-0.21	-0.08	-0.12	-0.05	0.00	0.02
DEAM-AEAM	-0.01	0.01	-0.02	-0.07	-0.10	0.01	0.04	0.00	0.00	0.02
Shock to euro-area interest rate [2]										
MARK III B	-0.18	-0.04	-0.07	-0.09	-0.09	-0.38	-0.18	-0.03	-0.03	-0.01
MARK III	-0.14	-0.12	-0.16	-0.16	-0.13	-0.20	-0.14	-0.02	-0.04	-0.02
MARK III - MARK III B	0.04	-0.08	-0.09	-0.07	-0.04	0.18	0.04	0.01	-0.01	-0.01
Shock to world interest rate [2]										
MARK III B	-0.05	-0.10	-0.12	-0.11	-0.10	-0.35	-0.21	-0.01	0.00	0.01
MARK III	-0.11	-0.18	-0.20	-0.18	-0.14	-0.17	-0.22	-0.02	-0.02	-0.01
MARK III - MARK III B	-0.06	-0.08	-0.08	-0.07	-0.04	0.18	-0.01	-0.01	-0.02	-0.02

Notes: (1) Output gap for AEAM and DEAM; real GDP for MARK III and MARK III B
(2) Exchange rates endogenous.

Table 6

AGGREGATION CRITERIA

	$DEAM$	$AEAM$
AGGREGATE SUPPLY		
GG criterium	3.856	6.523
PPK criterium	0.225	0.286
AGGREGATE DEMAND		
GG criterium	15.948	18.961
PPK criterium	0.449	0.487

GG : Grunfeld and Griliches (1960)

PPK: Pesaran, H.M., R. G. Pierse and M.S. Kumar (1989).

Table 7

IDIOSYNCRATIC COMPONENTS OF MODEL REGRESSORS

(percentage of the standard deviation explained
by the idiosyncratic components across countries)

Country	Output Gap	Interest Rates
Germany	75.7	91.2
France	77.2	52.7
Italy	40.7	32.0

Note: The factor models have one common component specified as AR(2) in the case of the output gap and AR(1) for the interest rates. The estimation algorithm is Kalman filter, solved with Berndt, Hall, Hall, and Hausman (BHHH) optimization method; starting conditions for the AR coefficients in the common components are imposed to be equal to the OLS estimation of AR models.

Table 8

RMSE OF N-STEP AHEAD ERRORS

(1978:1 2001:4)

DEAM	Aggregate supply	Aggregate demand
1	0.219	0.463
2	0.273	0.607
3	0.303	0.687
4	0.344	0.764
8	0.470	0.846

AEAM	Aggregate supply	Aggregate demand
1	0.283	0.468
2	0.327	0.620
3	0.370	0.701
4	0.394	0.771
8	0.552	0.834

OUT OF SAMPLE RMSE OF N-STEP AHEAD ERRORS

(1999:1 2001:4)

DEAM	Aggregate supply	Aggregate demand
1	0.278	0.289
2	0.331	0.460
3	0.258	0.601
4	0.366	0.676
8	0.339	0.540

AEAM	Aggregate supply	Aggregate demand
1	0.357	0.277
2	0.382	0.439
3	0.325	0.538
4	0.407	0.582
8	0.439	0.459

Table 9

FORECAST ENCOMPASSING TEST REGRESSIONS

Aggregate supply	Step ahead of the prediction (quarters) π_1	π_2	π_3	π_4	π_8
$cost$	0.027 (0.048)	0.057 (0.070)	0.068 (0.083)	0.099 (0.093)	0.211 (0.127)
S_AEAM	0.095 (0.109)	0.169 (0.159)	0.113 (0.164)	0.200 (0.234)	0.158 (0.347)
S_DEAM	0.887 (0.119)	0.792 (0.185)	0.841 (0.203)	0.731 (0.294)	0.658 (0.431)
$\bar{R}^2$	0.92	0.88	0.85	0.81	0.63
$AEAM\ encompass^*$					
$F-statistic$ $p-value\%$	32.46 0	12.17 0	13.89 0	7.07 0	4.87 0.3
$DEAM\ encompass^{**}$					
$F-statistic$ $p-value\%$	0.51 67.7	0.67 57.0	0.45 72.0	0.71 54.9	0.92 43.3

Aggregate demand	Step ahead of the prediction (quarters) y_1	y_2	y_3	y_4	y_8
$cost$	−0.016 (0.053)	−0.025 (0.090)	−0.022 (0.120)	−0.006 (0.141)	0.015 (0.170)
S_AEAM	0.244 (0.423)	0.088 (0.567)	0.143 (0.715)	0.369 (0.811)	0.852 (0.746)
S_DEAM	0.767 (0.418)	0.929 (0.562)	0.853 (0.754)	0.502 (0.918)	−0.310 (0.924)
$\bar{R}^2$	0.72	0.52	0.39	0.26	0.12
$AEAM\ encompass^*$					
$F-statistic$ $p-value\%$	1.13 34.0	0.92 43.6	0.51 67.6	0.52 66.9	0.94 42.7
$DEAM\ encompass^{**}$					
$F-statistic$ $p-value\%$	0.19 90.0	0.06 98.1	0.04 99.0	0.14 93.5	1.25 29.7

In parentheses: heteroskedasticity and autocorrelation consistent (Newey West) standard error of the coefficients.

* Test of the restrictions : coeff(AEAM)=1; coeff(DEAM)=0; constant=0.

** Test of the restrictions : coeff(AEAM)=0; coeff(DEAM)=1; constant=0.

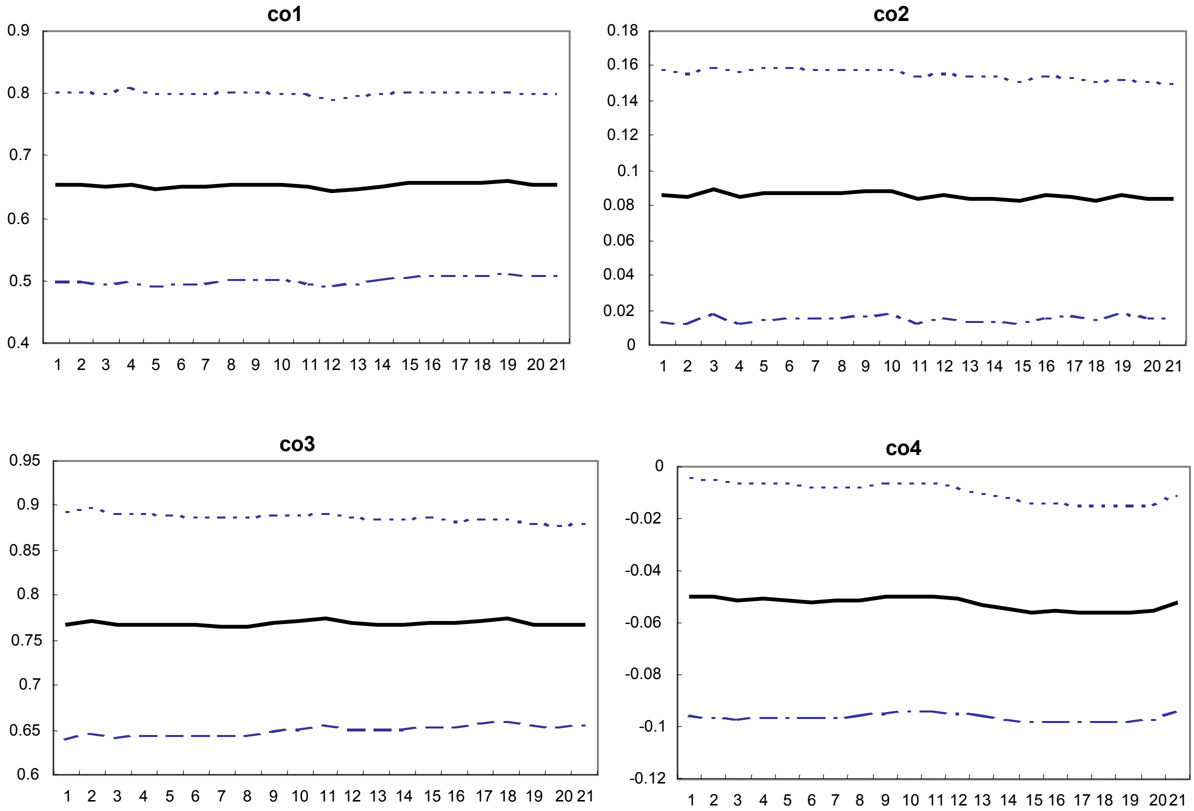
Table 10

Reaction function coefficients and loss values for the optimal, the AEAM-based and the DEAM-based rules

Parameter values in the loss function:		Type of rule	Coefficients on:			Standard deviation of:		Loss
			Annualized Inflation	Output gap	Lagged interest rate	Annualized Inflation	Output gap	
$\lambda = 0.1$	μ=0.1	FOR	2.93	2.86	0.49	1.79	1.37	4.06
		DEAM-based	2.46	1.75	0.69	2.18	1.68	6.41
		AEAM-based	1.66	1.27	0.75	2.42	1.62	6.84
	μ=1	FOR	1.18	1.23	0.67	2.18	1.29	6.41
		DEAM-based	0.99	0.86	0.82	2.81	1.57	11.34
		AEAM-based	0.69	0.65	0.85	3.21	1.50	12.29
$\lambda = 1$	μ=0.1	FOR	2.90	3.31	0.41	1.83	1.26	5.59
		DEAM-based	2.54	1.98	0.64	2.24	1.52	8.67
		AEAM-based	1.64	1.42	0.70	2.63	1.42	9.50
	μ=1	FOR	1.19	1.34	0.64	2.20	1.25	7.86
		DEAM-based	1.02	0.92	0.81	2.84	1.50	13.46
		AEAM-based	0.69	0.68	0.83	3.37	1.42	14.91
$\lambda = 2$	μ=0.1	FOR	2.88	3.68	0.35	1.89	1.20	7.10
		DEAM-based	2.61	2.19	0.61	2.32	1.43	10.84
		AEAM-based	1.65	1.57	0.66	2.85	1.32	12.18
	μ=1	FOR	1.20	1.44	0.62	2.22	1.22	9.38
		DEAM-based	1.06	0.98	0.79	2.88	1.45	15.64
		AEAM-based	0.70	0.72	0.82	3.54	1.35	17.70

Figure 1

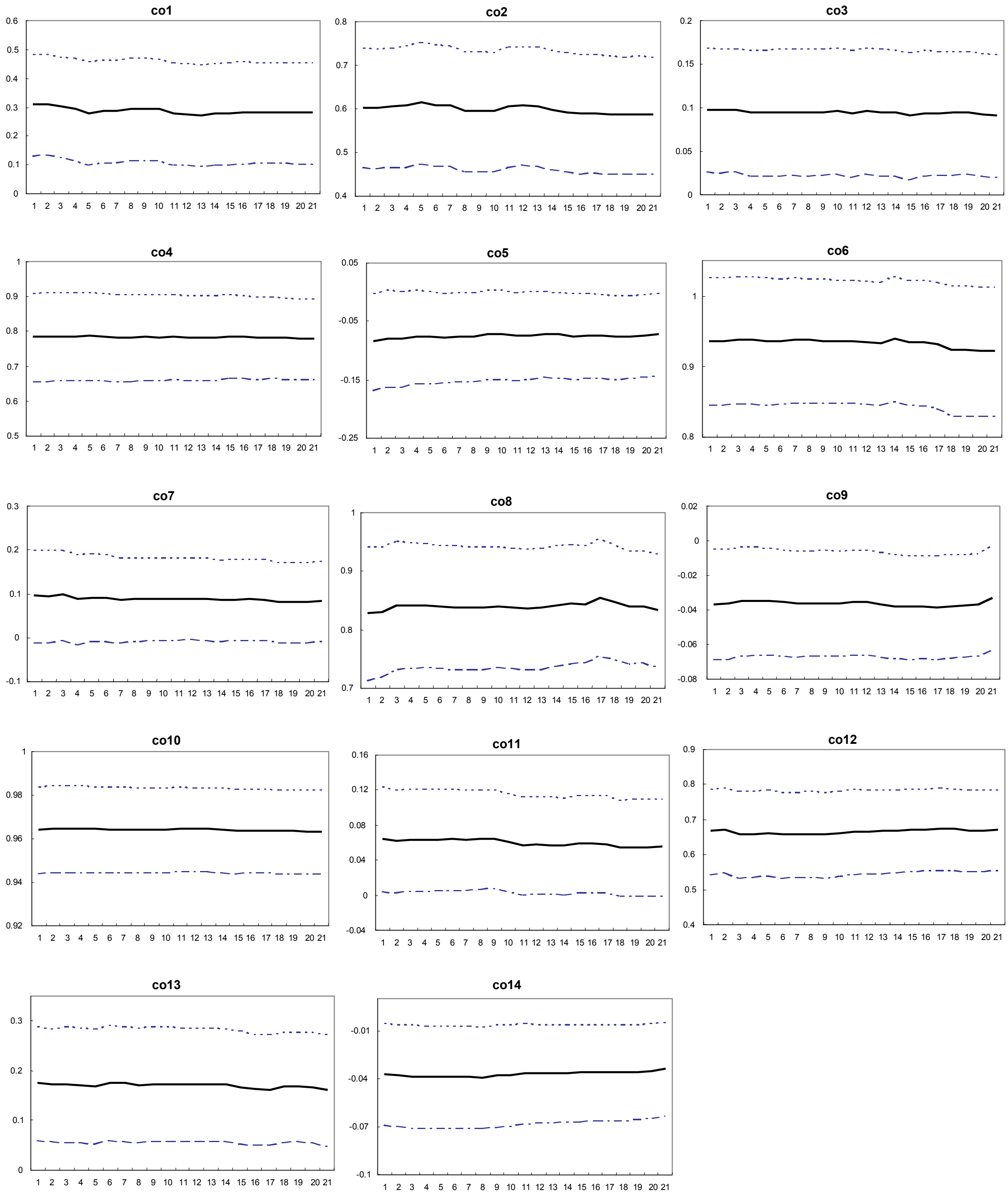
RECURSIVE ESTIMATES OF AEAM, 1996.Q4 – 2001.Q4



Legenda: co1: coeff. inflation (lag 1) in AS curve; co2: coeff. of output gap (lag 1) in AS curve; co3: coeff. of output gap (lag 1) in AD curve; co4: coeff. of real interest rate (lag 2)in AD curve

Figure 2

RECURSIVE ESTIMATES OF DEAM COEFFICIENTS, 1996.Q4 – 2001.Q4

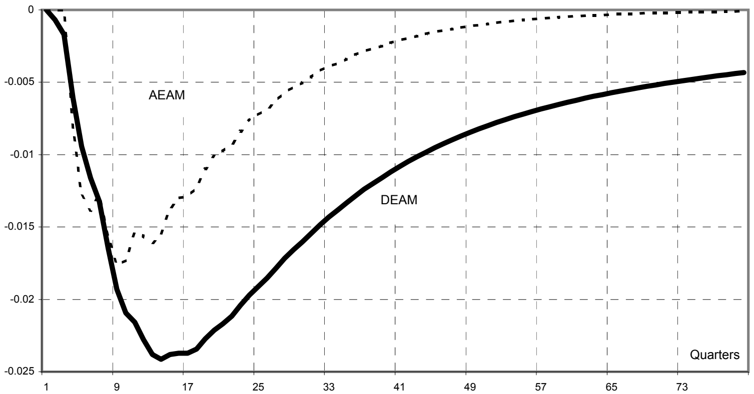


Legenda: co1: coeff. of German inflation (lag 1) in German AS curve; co2: coeff. of German inflation (lag 4) in German AS curve; co3: coeff. of German output gap (lag 1) in German AS curve; co4: coeff. of German output gap (lag 1) in German AD curve; co5: coeff. of real interest rate in German AD curve; co6: coeff. of French inflation (lag 1) in French AS curve; co7: coeff. of French output gap (average of lags 2-5) in French AS curve; co8: coeff. of French output gap (lag 1) in French AD curve; co9: coeff. of real interest rate in French AD curve; co10: coeff. of Italian inflation (lag 1) in Italian AS curve; co11: coeff. of Italian output gap (lag 1) in Italian AS curve; co12: coeff. of Italian output gap (lag 1) in Italian AD curve; co13: coeff. of German output gap in Italian AD curve; co14: coeff. of real interest rate in Italian AD curve

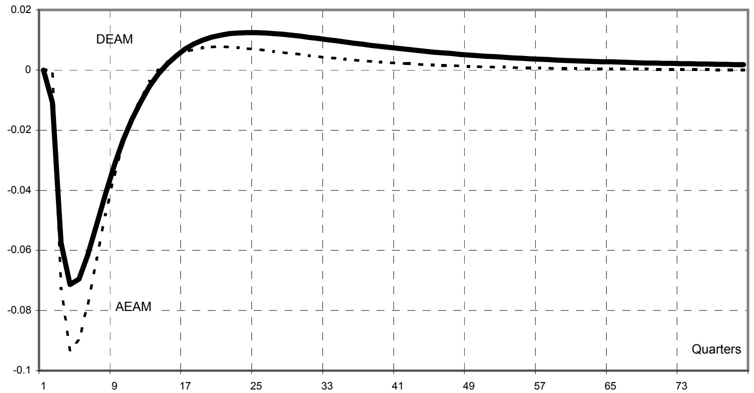
Figure 3

Impulse responses to a temporary monetary policy shock (+100 b.p.)

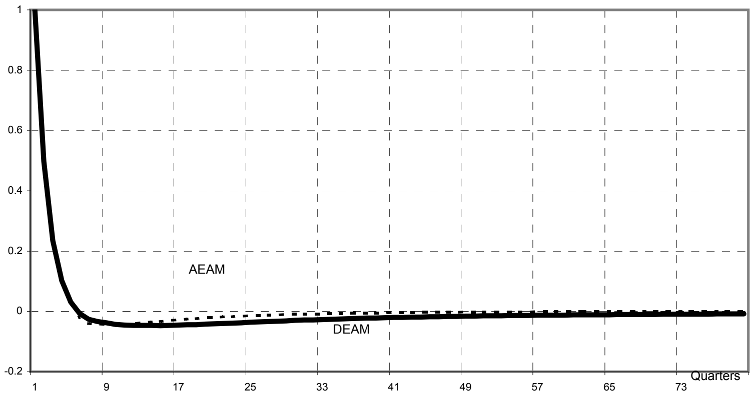
(a) Response of euro area inflation rate



(b) Response of euro area output gap



(a) Response of euro area real interest rate



(d) Response of euro area nominal interest rate

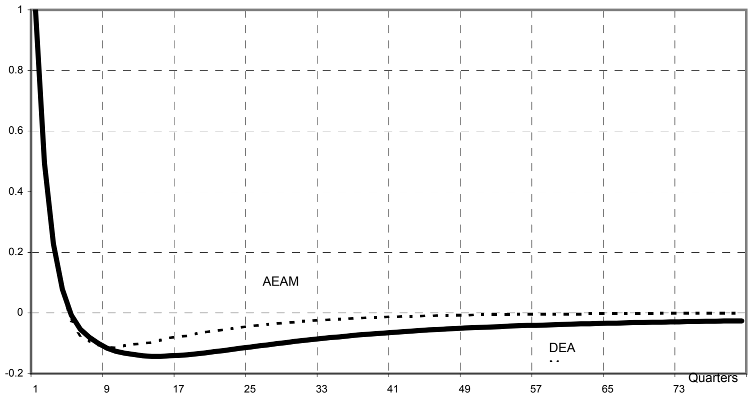
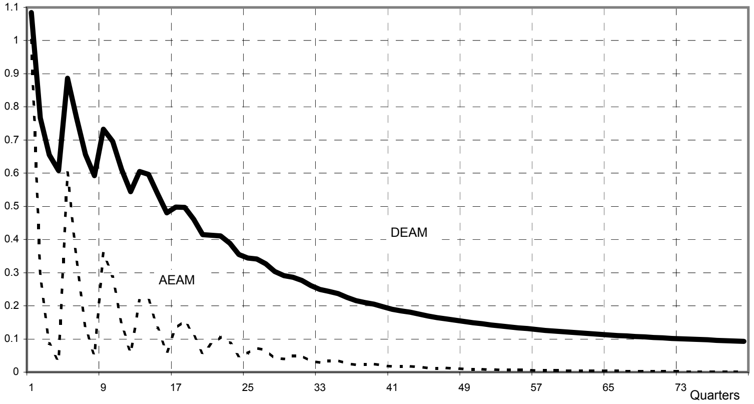


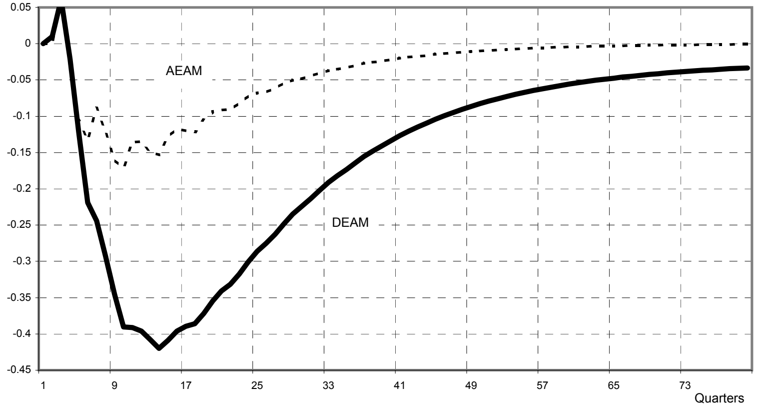
Figure 4

Impulse responses to a temporary Phillips curve shock (+1 per cent)

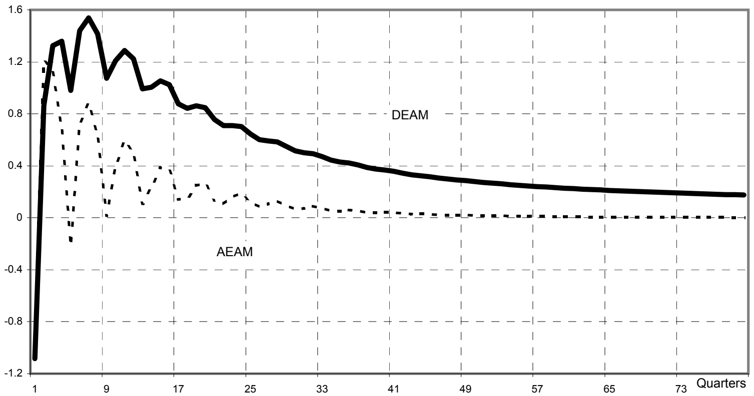
(a) Response of euro area inflation rate



(b) Response of euro area output gap



(a) Response of euro area real interest rate



(d) Response of euro area nominal interest rate

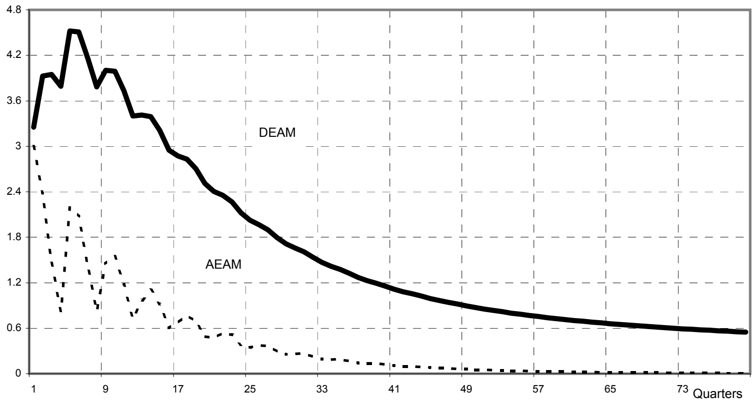
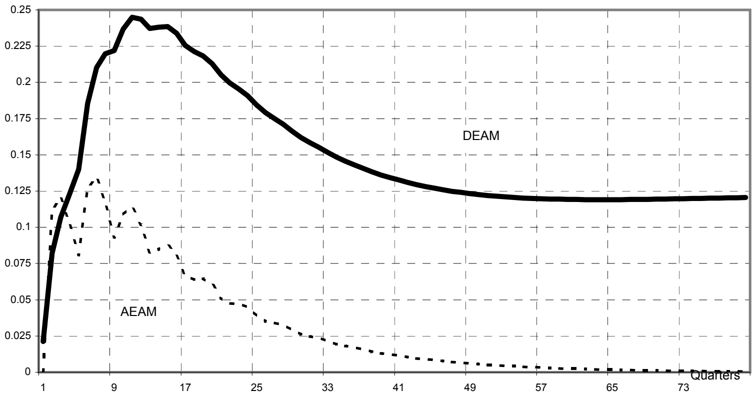


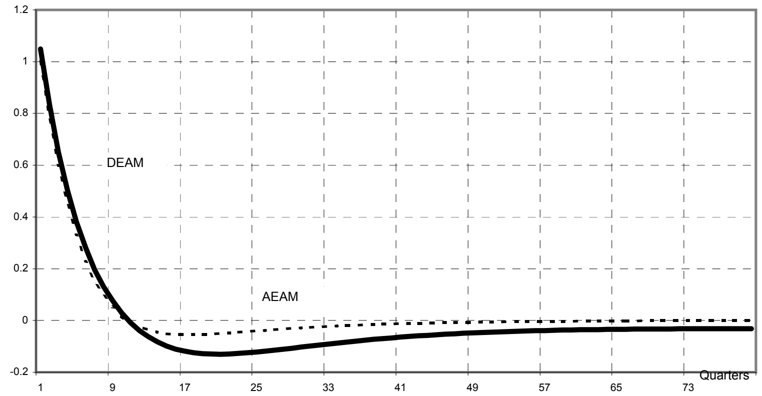
Figure 5

Impulse responses to a temporary aggregate demand shock (+1 per cent)

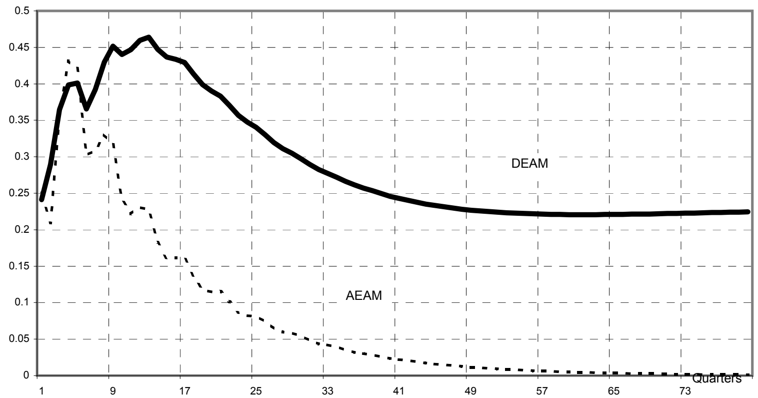
(a) Response of euro area inflation rate



(b) Response of euro area output gap



(a) Response of euro area real interest rate



(d) Response of euro area nominal interest rate

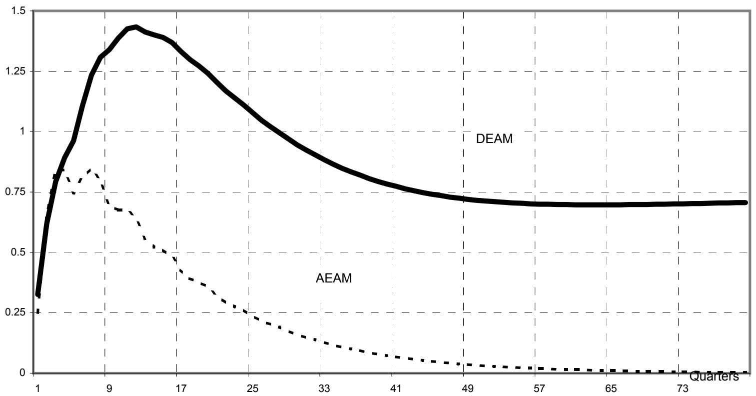
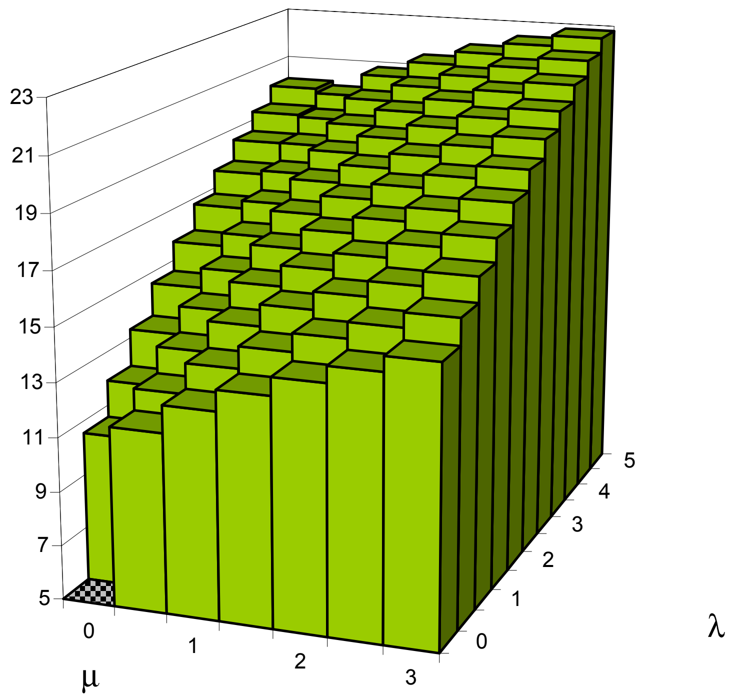


Figure 6

Percentage additional loss, AEAM-based rule vs. DEAM-based rule



Percentage additional loss, AEAM-based rule vs. DEAM-based rule
(as a share of the additional loss with the DEAM-based rule vs. the loss with the FO rule)

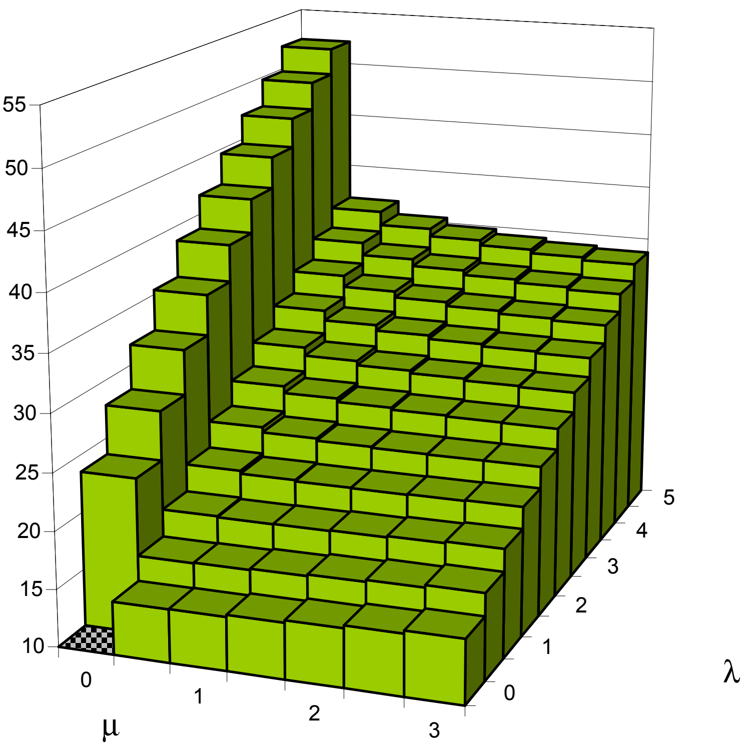


Fig. 7

Inflation - output gap optimal frontiers

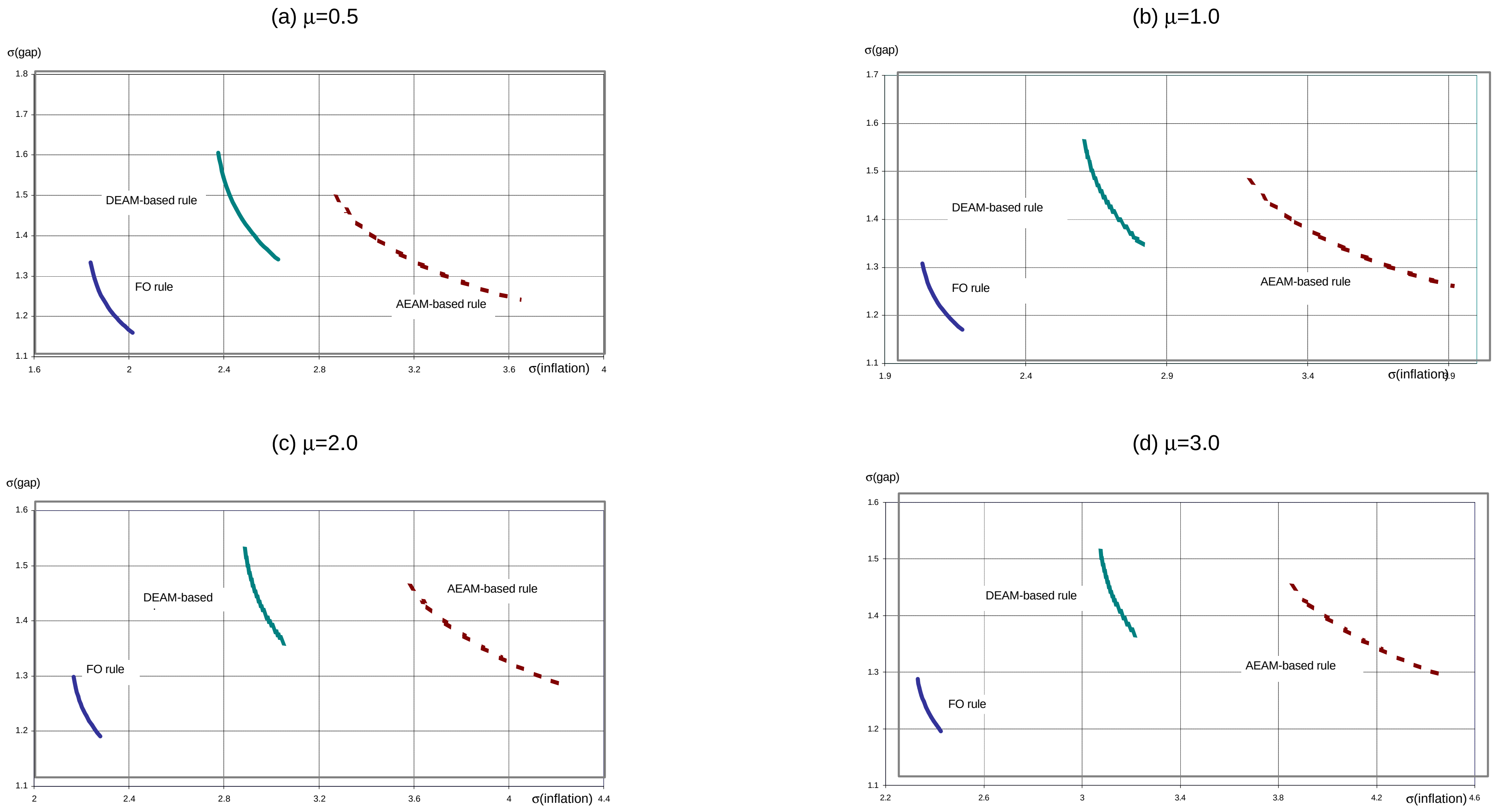
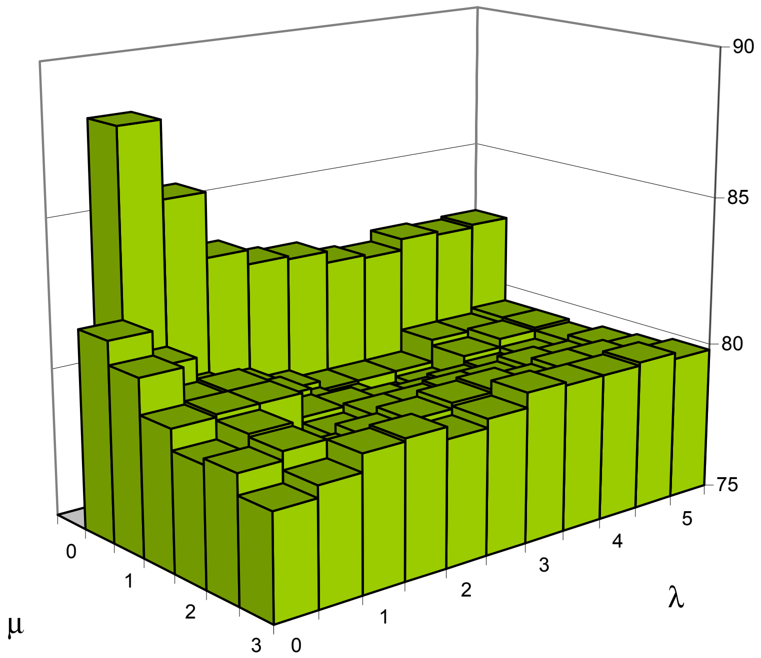


Figure 8

Random drawings from distribution of estimation residuals, DEAM-based rule vs. AEAM-based rule

Percentage of cases in which the DEAM-based rule outperforms the AEAM-based rule



Percentage of cases in which the DEAM-based rule outperforms the AEAM-based rule by at least 20 per cent of the optimised loss function associated with the latter

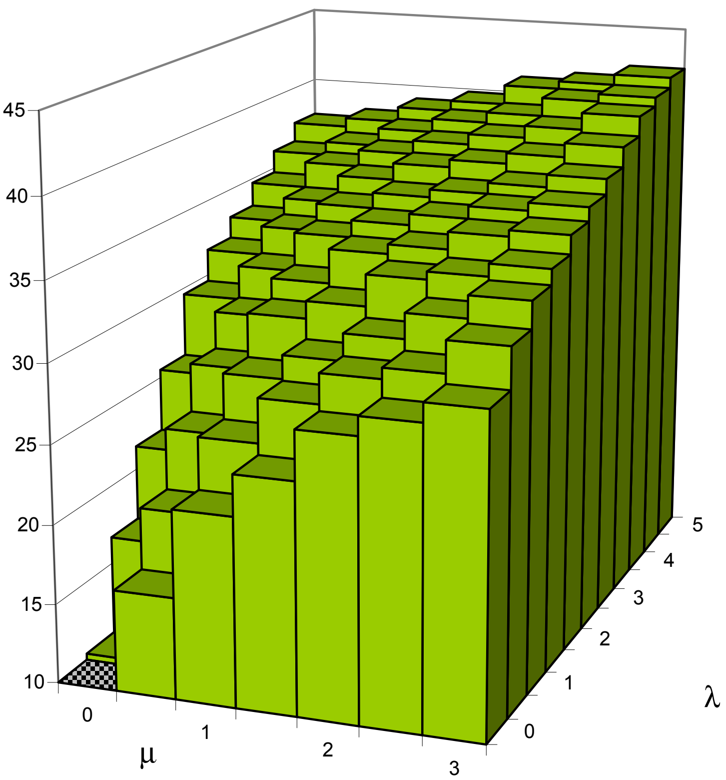


Figure 9

Random drawings from distribution of estimated DEAM residuals
Testing the significance of the underperformance of the AEAM-based rule

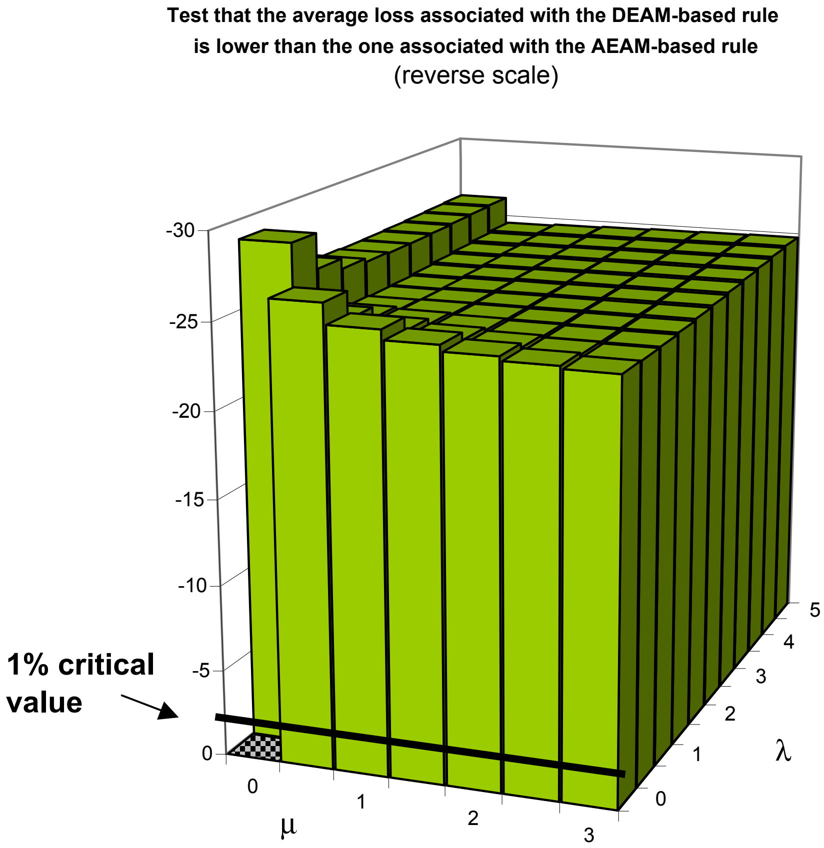
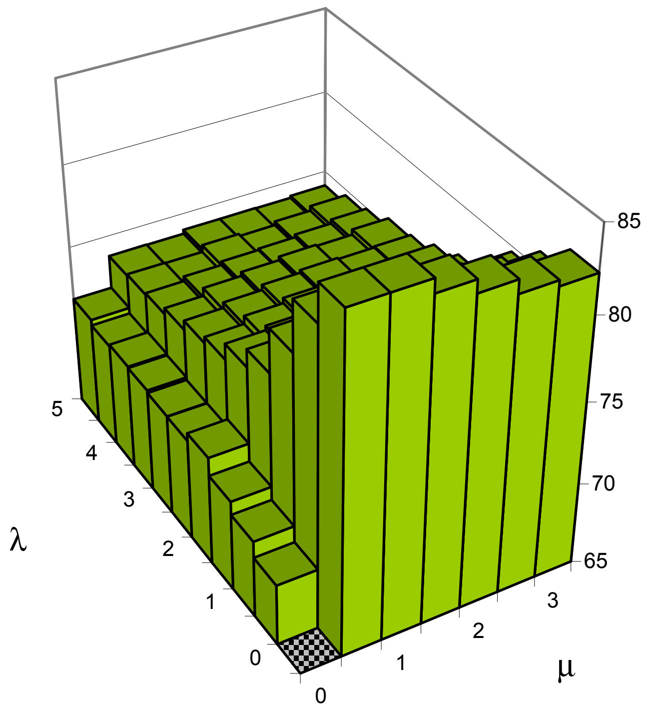


Figure 10

Random drawings from distribution of estimated DEAM parameters, DEAM-based rule vs. AEAM-based rule

Percentage of cases in which the DEAM-based rule outperforms the AEAM-based rule



Percentage of cases in which the loss function associated with the AEAM-based rule is higher than the one associated with the DEAM-based rule by at least 20 per cent

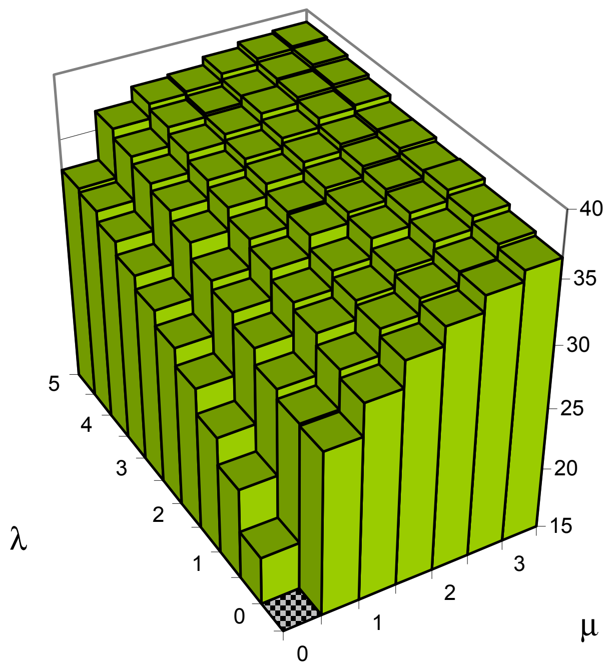


Figure 11

Random drawings from distribution of estimated DEAM parameters
Testing the significance of the underperformance of the AEAM-based rule

Test that the average loss associated with the DEAM-based rule is lower than the one associated with the AEAM-based rule
(tail probability)

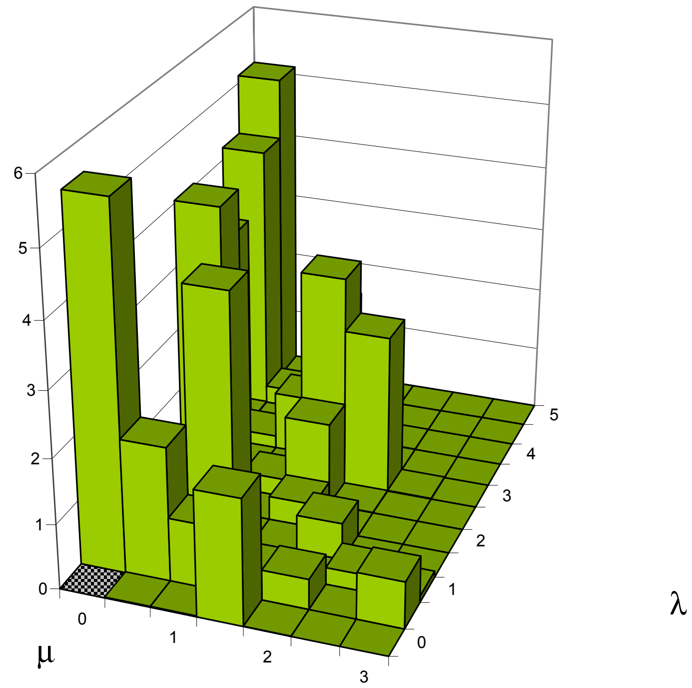


Figure 12

Additional welfare loss, AEAM-based rule vs. DEAM-based rule

(with gradual convergence of Phillips curve and aggregate demand stochastic processes, $\lambda = \mu = 1$;
the additional loss in the case of no convergence is set=100)