TI 2023-013/II
Tinbergen Institute Discussion Paper

Analyzing and forecasting economic crises with an agent-based model of the euro area

Cars Hommes[1]
Sebastian Poledna[2]

1 Bank of Canada, University of Amsterdam and Tinbergen Institute

2 International Institute for Applied Systems Analysis, Institute for Advanced Study

Analyzing and forecasting economic crises with an agent-based model of the euro area

Cars Hommes
Bank of Canada, University of Amsterdam, Tinbergen Institute

Sebastian Poledna
International Institute for Applied Systems Analysis, Institute for Advanced Study

7 March 2023

Abstract
We develop an agent-based model for the euro area that fulfils widely recommended requirements for next-generation macroeconomic models by i) incorporating financial frictions, ii) relaxing the requirement of rational expectations, and iii) including heterogeneous agents. Using macroeconomic and sectoral data, the model includes all sectors (financial, non-financial, household, and a general government) and connects financial flows and balance sheets with stock-flow consistency. The model, moreover, incorporates many features considered essential for future policy models, such as a financial accelerator with debt-financed investment and a complete GDP identity, and allows for non-linear responses. We first show that the agent-based model outperforms dynamic stochastic general equilibrium and vector autoregression models in out-of-sample forecasting. We then demonstrate that the model can help make sense of extreme macroeconomic movements and apply the model to the three recent major economic crises of the euro area: the Financial crisis of 2007-2008 and the subsequent Great Recession, the European sovereign debt crisis, and the COVID-19 recession. We show that the model, due to non-linear responses, is capable of predicting a severe crisis arising endogenously around the most intense phase of the Great Recession in the euro area without any exogenous shocks. By analysing the COVID-19 recession, we further demonstrate the model for scenario analysis with exogenous shocks. Here we show that the model reproduces the observed deep recession followed by a swift recovery and also captures the persistent rise in inflation following the COVID-19 recession. (JEL: E70, E32, E37)

Keywords: agent-based models, behavioural macro, macroeconomic forecasting, microdata, financial crisis; inflation and prices; coronavirus disease (COVID- 19).

Acknowledgments: The views expressed in this paper are solely those of the authors and may differ from official Bank of Canada views. No responsibility for them should be attributed to the Bank. The authors have no relevant interests to disclose.

E-mail: c.h.hommes@uva.nl (Hommes); poledna@iiasa.ac.at (Poledna)

1. Introduction

In this paper, we develop an agent-based model (ABM) for the euro area that fulfils widely recommended requirements for next-generation macroeconomic models and use it for scenario analysis of three recent major crises: the Financial crisis of 2007-2008 and the subsequent Great Recession, the European sovereign debt crisis, and the COVID-19 recession. These economic crises, as terrible as they were, have also presented an opportunity for the development of next-generation macroeconomic models. While the COVID-19 recession caused a larger contraction, the Financial crisis of 2007-2008 and the subsequent Great Recession had arguably the most profound influence on the economics profession in recent history (Haldane and Turrell, 2018). The benchmark model of Smets and Wouters (2007) at that time, which shares many features with models used by central banks and large international institutions, did not give any warning of the emergence of a crisis in 2008 and had difficulty explaining both the depth and the slow recovery of the Great Recession (Lindé et al., 2016). In the wake of the crisis, Vines and Wills (2018) collected requirements for next-generation macroeconomic models from several prominent voices in the economics profession.[1] These requirements include relaxing rational expectations; introducing heterogeneous agents; underpinning the model with more appropriate micro-foundations; incorporating financial frictions rather than assuming that financial intermediation is costless, all of which have been active areas of research

1. For earlier critiques, see, e.g. Canova and Sala (2009), Colander et al. (2009), Kirman (2010), Edge and Gurkaynak (2010), Krugman (2011), Stiglitz (2011, 2018), Blanchard (2016), and Romer (2016). See also the response defending DSGE models by Christiano et al. (2018).

in the DSGE literature in the recent past.[2,3,4] Complementary to DSGE models, research is also conducted on approaches with different microfoundations, such as behavioural macro models[5] and ABMs, which are a promising complement to the current crop of macroeconomic models, especially when making sense of the crises the world has witnessed for the past decade (Haldane and Turrell, 2018). ABMs explain the evolution of an economy by simulating the micro-level behaviour of individual agents to give a macro-level picture and have started to attract increasing attention from the early 2000s onwards (Dawid and Delli Gatti, 2018).[6]

The contribution of this paper is, first, to develop an ABM that fulfils these requirements for next-generation macroeconomic models and fits the micro and macroeconomic data of a large economy,

2. In the wake of the financial crisis, DSGE models with financial frictions experienced a revival and credit market imperfections were quickly incorporated into the more modern vintages of models. However, work that empirically evaluates the performance of these models finds that the inclusion of financial frictions does not necessarily improve the performance of the benchmark model (see, e.g. Brzoza-Brzezina and Kolasa (2013), Lindé et al. (2016) and Del Negro et al. (2016)).

3. The DSGE literature has also increasingly considered deviations from the rational expectations hypothesis and informational frictions as important for understanding macroeconomic dynamics. While the assumption of full information rational expectations (FIRE) has been the workhorse approach for the past several decades, bounded rationality, particularly adaptive learning, has increasingly been implemented in DSGE models. Examples of estimated medium-scale macroeconomic models that relax the rational expectations assumption find that some of the persistence observed in macroeconomic time series is no longer attributed to structural parameters but to the expectations formation process and that models with rational expectations can be outperformed (c.f. Milani (2007, 2012), Slobodyan and Wouters (2012), and Hommes et al. (2022b)).

4. Another vivid area of economic research explores the effects of agent heterogeneity in a general equilibrium framework, which has led to the development of heterogeneous agent New Keynesian (HANK) models. A non-exhaustive list of prominent examples includes Kaplan et al. (2018), Kaplan and Violante (2014), McKay and Reis (2016), Khan and Thomas (2008), and Chatterjee et al. (2007). HANK models have been used to show that household and firm heterogeneity affect macroeconomic aggregate but have rarely been used to forecast economic aggregates and representative agent New Keynesian (RANK) models have, so far, remained the benchmark (Kaplan and Violante, 2018; Christiano et al., 2018; Del Negro and Schorfheide, 2013).

5. In recent years, a large body of literature on behavioural macro models with boundedly rational agents and heterogeneous expectations has appeared. See, e.g. the recent contribution of De Grauwe and Ji (2020) or the recent overview in Hommes (2021).

6. Main developments in macroeconomic ABMs during the last decade are discussed in Dawid and Delli Gatti (2018). More recently, in the wake of the COVID-19 pandemic, ABMs experienced increased attention again. Several researchers used the flexibility of ABMs to forecast the economics of the COVID-19 pandemic and assess the economic impact of lockdowns (Poledna et al., 2020; Basurto et al., 2022; Delli Gatti and Reissl, 2022; Pangallo et al., 2022).

the euro area. Second, we assess the forecast performance of the model and show that the out-of-sample forecast performance of the ABM improves upon that of unconstrained VAR models and the benchmark DSGE model of Smets and Wouters (2007). As ABMs are better placed to produce conditional forecasts (Haldane and Turrell, 2018), we also compare the conditional forecasting performance against standard models. Finally, we demonstrate the good fit of the model in explaining the Great Recession and the European sovereign debt crisis—two crises that DSGE models struggled to predict and have difficulties explaining (Lindé et al., 2016; Lindé, 2018; Haldane and Turrell, 2018)—as well as the COVID-19 recession. We show with out-of-sample forecasting exercises that the model predicts an endogenous crisis around the most intense phase of the Great Recession in the euro area, albeit with lower severity because there is no global downturn, which is exogenous to the model. With conditional forecasts, which include an exogenous shock on exports from the global downturn, we demonstrate that the model captures both the severity and the slow recovery of the Great Recession. By analysing the COVID-19 recession, we further demonstrate the potential of the model for scenario analysis with exogenous shocks. Here we show that by implementing an industry-specific shock caused by the restrictions of economic activities due to the lockdown measures, the model reproduces the observed deep recession followed by a swift recovery. By additionally adding an export demand and import supply shock from the global supply chain crisis, as well as a fiscal shock from increased government spending, we demonstrate that the model also captures the persistent rise of inflation following the COVID-19 recession.

The model developed in this paper is based on Poledna et al. (2023), who have developed the first ABM that fits micro and macroeconomic data of a small open economy and allows out-of-sample forecasting of the aggregate macro variables, such as GDP (including its components), inflation and interest rates. This ABM was recently adopted by the Bank of Canada to complement their current suite of in-house core macro models (Hommes et al., 2022a). The model of Poledna et al. (2023) is, in turn, based on Assenza et al. (2015), who have developed a stylized ABM with households, firms (upstream and downstream) and a bank, which replicates a number of stylized facts. Following Poledna et al. (2023), the model in this paper includes all sectors (financial, non-financial, households, and a general government) which are populated by heterogeneous agents. In general, "stocks affect flows" in the model and financial flows are connected to balance sheets with stock-flow consistency. The model includes a complete GDP identity with all GDP components of the production (or output or value added) approach, income approach, and expenditure approach. Markets are fully decentralized and characterized by a continuous search and matching process that allows for trade frictions. The model is based on data from national accounts, sector accounts, input-output tables, government statistics, and structural business statistics. Model parameters are either taken directly from data or are calculated from national accounting identities. For exogenous processes such as government spending and exports of the euro area, parameters are estimated.

The model in this paper includes a financial accelerator and debt-financed investment with collateralized loans where firms' ability to borrow depends on the market value of their assets and their net worth[7]. Thus, the model allows for central problems of finance-bankruptcy, debt, and asymmetric information—which cannot arise in a representative agent model (Stiglitz, 2018). Agents' expectations are modelled by adaptive learning, where agents act as econometricians who estimate the parameters of their model and make forecasts using their estimates (Evans and Honkapohja, 2001). To forecast economic growth and inflation, agents learn the optimal parameters of vector autoregressive models with exogenous variables (VARX) with time-varying coefficients.[8] With adaptive learning, we relax the requirement of the law of iterated expectations necessary to derive DSGE models (Hendry and Muellbauer, 2018; Lindé, 2018). Adaptive learning with autoregressive models with exogenous variables is further suitable for policy analysis and allows agents to consider future policies, such as austerity, during the European debt crisis. Using autoregressive models with time-varying coefficients in adaptive learning allows for non-linear responses, which are underestimated by linearized DSGE models (Lindé, 2018). Together with financial frictions, balance sheet effects, and heterogeneity, non-linear responses allow for the possibility of endogenous economic

7. Lindé et al. (2016) extend the Smets and Wouters (2007) DSGE model with a financial accelerator. Together with estimating the model allowing explicitly for the zero lower bound constraints on nominal interest rates and introducing time-variation in the volatility of the exogenous disturbances, they find that these extensions go some way in accounting for features of the Great Recession and its aftermath. However, Lindé et al. (2016) do not find that the financial accelerator adds much propagation of other macroeconomic shocks.

8. VAR-based expectations are also used in the FRB/US model (for a discussion of the role of expectations in FRB/US, see Brayton et al. (1997)). More generally, one may consider linear AR(p) or VAR(p) forecasting models (Hommes and Zhu, 2014). Poledna et al. (2023) use simple parsimonious AR(1) rules. Xiao and Xu (2014) study learning and predictions with an AR(1) and a VAR(1) model in a two-dimensional New Keynesian model. Slobodyan and Wouters (2012) and Hommes et al. (2022b), furthermore, estimated the Smets and Wouters (2007) DSGE model with expectations modelled by simple AR(1) and AR(2) forecasting rules under time-varying beliefs and show that this leads to an improvement in the empirical fit of the model and its ability to capture the short-term momentum in the macroeconomic variables.

crises to occur in the model without exogenous shocks. The model includes a government sector, which collects taxes and provides social transfers, and a housing market where households invest in dwellings. Thus, the model includes the automatic stabilizing mechanisms associated with a large welfare state, which sustained household consumption and investment with government transfers during the Great Recession, while disposable income was buffered by falling tax receipts (Krugman, 2018).

The remainder of this manuscript is structured as follows. Section 2 describe the most important modelling choices and mechanisms of the model. Section 3 provides an overview of the data sources and the calibration of the model to the euro area. Section 4 assesses the forecast performance of the ABM, where we compare the ABM against unconstrained VAR models and the benchmark DSGE model. In Sections 5 to 7, we conduct a scenario analysis of three recent major crises: the Financial crisis of 2007-2008 and the subsequent Great Recession (Section 5), the European sovereign debt crisis (Section 6), and the COVID-19 recession (Section 7). Section 8 concludes the paper.

2. An agent-based model for the euro area

In this section, we describe the most important modelling choices and mechanisms of the model with a focus on the differences with respect to the model developed in Poledna et al. (2023). For a complete description and discussion of the model, see Poledna et al. (2023). Model codes and programs are provided in the data repository. There are three main differences with respect to the model developed in Poledna et al. (2023). First, the model by Poledna et al. (2023) of a small open economy is generalized to a large open economy setting for the euro area. To this end, the policy rate is determined by a forward-looking Taylor rule, where the central bank agent learns the optimal parameters, and imports are driven by domestic demand, assuming a large open economy setting where world prices develop in line with domestic price levels. Second, the model in this paper includes a financial accelerator and debt-financed investment with collateralized loans where firms' ability to borrow depends on the market value of their assets and their net worth. Additionally, the model includes an inventory cycle caused by the accumulation and selling of inventories. Third, agents' expectations are modelled by adaptive learning, where agents learn the optimal parameters of a VARX(1) rule with time-varying coefficients. Adaptive learning with VARX rules lets agents consider future policies, and time-varying coefficients allow for non-linear responses.

As in Poledna et al. (2023), the model economy is structured following the conventions of the European System of Accounts (ESA) into four mutually exclusive domestic institutional sectors: (1) non-financial corporations (firms), (2) financial corporations and the central bank (the financial sector), (3) the general government, and (4) households. The four sectors make up the total domestic economy of the euro area and interact with (5) the rest of the world through imports and exports. Each sector is populated by heterogeneous agents who represent natural persons or legal entities (households, corporations, and government entities). We use a scale of 1:1000 between the model and data so that each agent in the model represents a thousand individuals or businesses in the euro area. All individual agents have separate balance sheets depicting assets, liabilities, and ownership structures. Balance sheets of the agents, and economic flows between them, fit micro and macroeconomic data of the euro area.

2.1. Firms

The firm sector is populated by heterogeneous agents that represent all firms from every industry in the euro area. The firm population of each industry s is obtained from structural business statistics, and the firm size distribution follows a power law that approximately corresponds to firm-level data of the euro area. Each firm produces industry-specific goods g by means of labour, capital, and intermediate inputs from other industries. We assume a production function with Leontief technology and separate nests for intermediate goods ($M_i(t-1)$), labour ($N_i(t)$) and capital ($K_i(t-1)$), respectively—which all represent upper limits to production. The production of the i-th firm at time t is

$$Y_i(t) = \min\left(Q_i^s(t), \beta_i M_i(t-1), \alpha_i(t) N_i(t), \kappa_i K_i(t-1)\right), \tag{1}$$

where $\alpha_i(t)$ is the productivity of labour of firm i and β_i and κ_i are productivity coefficients for intermediate inputs and capital, respectively, which are calibrated to input-output tables. Production by firm i may not equal the desired scale of activity ($Q_i^s(t)$) and could be limited by labour market or trade frictions that may result in a lack of intermediate goods, available labour force or capital stock. In these cases, the firm has to scale down activity.

Firms are subject to fundamental uncertainty regarding their future sales, market prices, cash flow, etc. and form expectations about the future. Firm agents' expectations are modelled by adaptive learning, where agents act as econometricians who estimate the parameters of their model and make forecasts using their estimates (Evans and Honkapohja, 2001). Forecasts are based on the trends of economic growth, inflation and the policy rate as well as on (forward-looking) expectations for government spending and exports of the euro area.[9] To forecast economic growth and inflation, agents learn the optimal parameters of a VARX(1) rule with time-varying coefficients,

$$\gamma^e(t) = \mathrm{e}^{a^{\gamma,\gamma}(t)\gamma(t-1)+a^{\gamma,\pi}(t)\pi(t-1)+a^{\gamma,\bar{r}}(t)\bar{r}(t-1)+b^{\gamma,G}(t)\gamma^G(t)+b^{\gamma,E}(t)\gamma^E(t)+c^{\gamma}(t)+\epsilon^{\gamma}(t)} - 1 \tag{2a}$$
$$\pi^e(t) = \mathrm{e}^{a^{\pi,\gamma}(t)\gamma(t-1)+a^{\pi,\pi}(t)\pi(t-1)+a^{\pi,\bar{r}}(t)\bar{r}(t-1)+b^{\pi,G}(t)\gamma^G(t)+b^{\pi,E}(t)\gamma^E(t)+c^{\pi}(t)+\epsilon^{\pi}(t)} - 1\,, \tag{2b}$$

where the matrices $a(t)$ and $b(t)$, as well as the vectors of intercepts $c(t)$ and the standard deviations of the normally distributed error terms $\epsilon(t)$ are re-estimated during simulations every period on the time series of growth $\gamma(t')$, inflation $\pi(t')$, and the 3-month Euribor $\bar{r}(t')$ as well as the exogenous variables, growth of government consumption $\gamma^G(t')$ and exports $\gamma^E(t')$ where $t' = -T', -T'+1, -T'+2, \ldots, 0, 1, 2, \ldots, t-1$[10]. As initial conditions for $t' = -T', -T'+1, -T'+2, \ldots, 0$, we use the log differences of real GDP, the GDP deflator (inflation), real government consumption and real exports of the euro area as well as the 3-month Euribor from national accounting data. Thus, agents use econometric modelling with unfiltered time series (Hendry and Muellbauer, 2018; Wren-Lewis, 2018). With adaptive learning, we relax the requirement of the law of iterated expectations necessary to derive DSGE models (Hendry and Muellbauer, 2018; Lindé, 2018). Adaptive learning with VARX rules is further suitable for policy analysis and allows agents to consider future policies by letting them observe shocks at the beginning of period t. Time-varying coefficients, furthermore, allow for non-linear responses of the agents, which are underestimated by linearized DSGE models (Lindé, 2018).

Supply choice of firm i is made based on the expected rate of economic growth ($\gamma^e(t)$), the previous period's demand for its product $Q_i^d(t-1)$ and the level of inventories $S_i(t-1)$:

$$Q_i^s(t) = \max\left(Q_i^d(t-1)(1+\gamma^e(t)) - S_i(t-1), 0\right). \tag{3}$$

9. Thus agents incorporate, for example, expectations for future government austerity or export demand shocks in their behaviour. Alternatively, agents could also take expectations on future policy rate changes into account. However, we chose to set the policy rate according to a forward-looking Taylor rule, for details see Equation (10).

10. Observed data of exogenous variables are treated as deterministic future responses that are known in advance, for example, set by policy.

Along with the expected rate of inflation $\pi^e(t)$ ("built-in inflation"), the change from the previous period's price factors in the cost structure faced by the firm $\pi_i^c(t)$ ("cost-push inflation"):

$$P_i(t) = P_i(t-1) \cdot \underbrace{(1+\pi_i^c(t))}_{\substack{\text{Cost-Push}\\ \text{Inflation}}} \cdot \underbrace{(1+\pi^e(t))}_{\substack{\text{Built-In}\\ \text{Inflation}}}, \tag{4}$$

where

$$\pi_i^c(t) = \underbrace{\frac{(1+\tau^{SIF})\bar{w}_i}{\bar{\alpha}_i}\left(\frac{\bar{P}^{HH}(t-1)}{P_i(t-1)} - 1\right)}_{\text{Increase of unit labour costs}} + \underbrace{\frac{1}{\beta_i}\left(\frac{\sum_g a_{sg}\bar{P}_g(t-1)}{P_i(t-1)} - 1\right)}_{\text{Increase of unit material costs}} + \underbrace{\frac{\delta_i}{\kappa_i}\left(\frac{\bar{P}^{CF}(t-1)}{P_i(t-1)} - 1\right)}_{\text{Increase of unit capital costs}}. \tag{5}$$

Here, $\bar{\alpha}_i$ indicates the average productivity of labour, w_i are gross wages indexed by the consumer price index $\bar{P}^{HH}(t)$, and including employers' contribution to social insurance charged with a rate τ^{SIF}; $\frac{1}{\beta_i}\sum_g a_{sg}$ are unit expenditures (in real terms) on intermediate input by industry s on good g weighted by the average product price index for good g ($\bar{P}_g(t)$), δ_i/κ_i are unit capital costs due to depreciation (δ_i is the firm's capital depreciation rate and κ_i the productivity coefficient for capital), $\bar{P}^{CF}(t)$ is the average price of capital goods, for details see Poledna et al. (2023).

The model includes a financial accelerator and debt-financed investment with collateralized loans where firms' ability to borrow depends on the market value of their assets and their net worth. Firms may need external financial resources to finance current or future expenditures. To ascertain the need for external finance, each firm i forms an expectation of its future cash flow based on the expected profit ($\Pi_i^e(t)$) and the schedule of debt instalments ($\theta L_i(t-1)$),

$$\Delta D_i^e(t) = \underbrace{\Pi_i^e(t)}_{\text{Exp. profit}} - \underbrace{\theta L_i(t-1)}_{\text{Debt instalment}} - \underbrace{\tau^{FIRM}\max(0,\Pi_i^e(t))}_{\text{Corporate taxes}} - \underbrace{\theta^{DIV}(1-\tau^{FIRM})\max(0,\Pi_i^e(t))}_{\text{Dividend payout}}, \tag{6}$$

where $\Pi_i^e(t)$ is profit expected by firm i based on the firms' current input cost structure as well as their expectations on inflation and interest rates; θ is the rate of debt instalment on firm i's outstanding loans $L_i(t-1)$, τ^{FIRM} is the corporate tax rate, and θ^{DIV} is the dividend payout ratio. If the internal financial resources of a firm are not enough to finance its expenditures, firms ask for a bank loan to cover its financing gap, $\Delta L_i^d(t) = \Delta D_i^e(t) - D_i(t-1)$. The availability of credit depends on the financial condition of the firm and also the capitalization of the banking sector, see Poledna et al. (2023) for details. The bank will grant a loan to firm i up to a fixed loan-to-value ratio ζ^{LTV}. Due to fundamental uncertainty, the bank forms expectations on both the total outstanding debt of firm i ($L_i^e(t) = (1-\theta)L_i(t-1) + \Delta L_i(t)$), as well as on the market value of firm i's collateral

$(K_i^e(t) = K_i(t-1) - \frac{\delta_i}{\kappa_i} Q_i^s(t) + I_i^d(t))$. Thus, funds provided by the bank will be limited by the expected market value of the collateral and the total outstanding debt,

$$\Delta L_i(t) \leq \zeta^{LTV} \bar{P}^{CF}(t-1)(1+\pi^e(t))K_i^e(t) - (1-\theta)L_i(t-1)\,, \tag{7}$$

and firm i's expenditures in period t are constraint by the availability of internal and external funds,

$$\Delta D_i(t) \leq D_i(t-1) + \zeta^{LTV} \bar{P}^{CF}(t-1)(1+\pi^e(t))K_i^e(t) - (1-\theta)L_i(t-1)\,. \tag{8}$$

Therefore, a fall in asset prices results in a deterioration of the ability of firms to borrow, which has a negative impact on their investment. Decreased economic activity further reduces asset prices, which leads to a feedback cycle of falling asset prices, tightening of financial conditions and further declining economic activity.

Firm investment is conducted according to the expected wear and tear of capital and the availability of internal and external funds. If firm i has a funding gap, i.e. the difference between requested and granted external funding $(\Delta L_i^d(t) - \Delta L_i(t))$, the firm investment is reduced, The desired investment in capital stock in period t is

$$I_i^d(t) = \frac{\delta_i}{\kappa_i} \min\left(Q_i^s(t), \kappa_i K_i(t-1)\right) - \frac{\Delta L_i^d(t) - \Delta L_i(t)}{\bar{P}^{CF}(t-1)(1+\pi^e(t))}\,. \tag{9}$$

where δ_i is the firm's capital depreciation rate. Realized investment $(I_i(t))$ is the outcome of a search-and-matching process that is subject to trade frictions. Thus, realized investment may be lower than desired investment.

2.2. The financial sector

For reasons of simplicity, we assume that there is one representative financial intermediary for the euro area. This bank takes deposits from firms as well as households and extends loans to firms. Interest rates for loans are set by a fixed markup on the policy rate. The provision of loans is conditional on a minimum loan-to-value (LTV) ratio and credit creation is limited by minimum capital requirements. The capital of the banking sector grows or shrinks according to bank profits or losses and the write-off of bad debt.

The central bank sets the policy rate according to a Taylor rule, provides advances to the banking sector, and takes deposits in the form of bank reserves. The central bank purchases external assets (government bonds) and thus acts as a creditor to the central government. The policy rate is determined by a forward-looking Taylor rule, where the central bank agent learns the optimal parameters. Following Blattner and Margaritov (2010), we use a forward-looking "growth" rule specification where the output gap does not enter the equation:[11]

$$\bar{r}(t) = \rho(t)\bar{r}(t-1) + (1-\rho(t))(r^*(t) + \pi^* + \xi^\pi(t)(\log(\pi^e(t)+1) - \pi^*) + \xi^\gamma(t)\log(\gamma^e(t)+1))\,, \quad (10)$$

where $\rho(t)$ is a measure for gradual adjustment of the policy rate, $r^*(t)$ is the real equilibrium interest rate, π^* is the inflation target by the central bank, $\xi^\pi(t)$ is the weight put on inflation targeting, and $\xi^\gamma(t)$ the weight placed on economic growth, respectively. $\rho(t)$, $r^*(t)$, $\xi^\pi(t)$, and $\xi^\gamma(t)$ are re-estimated every period on time series of growth $\gamma(t')$, inflation $\pi(t')$, and $\bar{r}(t')$ where $t' = -T', -T'+1, -T'+2, \ldots, 0, 1, 2, \ldots, t-1$. As initial conditions for $t' = -T', -T'+1, -T'+2, \ldots, 0$, we use the log differences of real GDP, the GDP deflator (inflation), as well as the 3-month Euribor from national accounting data.

2.3. The general government

The government sector is modelled after the European social model with a large welfare state. A central government collects taxes and social security contributions and distributes social as well as other transfers. Government revenues are composed of taxes: on wages (income taxes), capital income (income and capital taxes), firm profit income (corporate taxes), household consumption (value-added tax), other products (sector-specific, paid by industry sectors), firm production (sector-specific), as well as on exports and capital formation, social security contributions by employees and employers, and of other net transfers such as property income, investment grants, operating surplus, as well as proceeds from government sales and services. Government expenditures consist of final government consumption, interest payments on government debt, social benefits other than social benefits in kind, subsidies and other current expenditures. A government deficit adds to the sovereign debt and increases interest payments in subsequent periods.

Government consumption is modelled by individual government entities that purchase on the consumption market. The amount and trend of government spending are obtained from government statistics and national accounts. Demand for final consumption expenditure of the

11. Here, we rely on empirical evidence and statements by leading central bankers reported in Blattner and Margaritov (2010) implying that the concept of an output gap does not seem to influence the behaviour of the European Central Bank (ECB) to a large extent.

general government ($C^G(t)$) is assumed to follow an autoregressive process of lag order one (AR(1)):

$$C^G(t) = \mathrm{e}^{\alpha^G \log(C^G(t-1))+\beta^G+\epsilon^G(t)} , \tag{11}$$

where $\epsilon^G(t)$ is normally distributed with standard deviation σ^G. Realized government consumption ($C_j(t)$) is then another outcome of a search-and-matching process that is subject to trade frictions.

2.4. Households

The household sector consists of all employed, unemployed, and inactive persons residing in the euro area. The respective populations are obtained from auxiliary national accounting data. Each person receives income according to his or her occupation. Employed persons supply labour to firms and earn wages. Unemployed persons search for employment and receive unemployment benefits, which are a fraction of previously received wages. Investors obtain dividend income from firm ownership. Inactive persons (e.g. students or pensioners) do not participate in the labour market and receive social benefits. All households receive additional social transfers (e.g. childcare payments) from the general government. Each household spends a fraction of its expected income on consumption:

$$C^d_h(t) = \frac{\psi Y^e_h(t)}{1+\tau^{VAT}} , \tag{12}$$

and on investment:

$$I^d_h(t) = \frac{\psi^H Y^e_h(t)}{1+\tau^{CF}} , \tag{13}$$

where ψ, ψ^H are propensities to consume and invest out of expected income; τ^{VAT}, τ^{CF} are value-added and investment tax rates. Realized household consumption ($I_h(t)$) and investment ($C_h(t)$) is another outcome of a search-and-matching process that is subject to trade frictions. Total household consumption allocated to goods g according to fixed coefficients from input-output tables, analogous to firm investment above.

2.5. Imports and Exports

We model the euro area as a large open economy where a segment of the firm sector participates in international trade. A representative foreign firm for each sector imports goods from the rest of the world and supplies them to domestic markets. Thus the m-th, ($m = 1, 2, \ldots, S$), foreign firm representing an industry s imports the principal product g:[12]

$$Y_m(t) = Q^d_m(t-1)(1+\gamma^e(t)) \,. \tag{14}$$

Prices for import goods are assumed to develop in line with domestic price levels. The foreign firm thus sells its products at the inflation-adjusted average sectoral domestic price level. Consequently,

$$P_m(t) = \bar{P}_g(t-1)(1+\pi^e(t)) \,, \tag{15}$$

where m produces the principal product g. This corresponds to the assumption of a fixed relation between the domestic and international price levels, i.e. the same inflation rate at home and abroad. Sales of imports are then the realized demand as an outcome of the search-and-matching process on the goods markets:

$$Q_m(t) = \min(Y_m(t), Q^{\mathrm{d}}_m(t)) \,, \tag{16}$$

where $Q^{\mathrm{d}}_m(t)$ is subject to a search-and-matching process that is subject to trade frictions.

Export demand from the rest of the world is obtained from exogenous projections based on national accounts. Thus, we suppose the demand for exports to be exogenously given. Analogous to government consumption, real export demand ($C^E(t)$) is assumed to follow an autoregressive process of lag order one (AR(1)):

$$C^E(t) = \mathrm{e}^{\alpha^E \log(C^E(t-1)) + \beta^E + \epsilon^E(t)} \,, \tag{17}$$

where $\epsilon^E(t)$ is normally distributed with standard deviation σ^E. Realized exports ($C_l(t)$) depend then on domestic supply and is an outcome of a search-and-matching process that is subject to trade frictions.

2.6. Macroeconomic aggregates

Finally, GDP in our model is calculated by aggregating the value of all final goods and services produced and purchased by agents in the model in a given period. The GDP and its components in

12. As for domestic firms, we suppose there is a one-to-one correspondence between the sets of industries s and products g, meaning that the n-th sector produces only the n-th good and $S = G$.

constant prices of each period t as defined by the expenditure approach is:

$$
\begin{aligned}
GDP(t) = & \underbrace{\sum_h (1+\tau^{\text{VAT}}) \frac{C_h(t)}{\bar{P}_h(t)}}_{\text{Household consumption}} + \underbrace{\sum_j (1+\tau^{\text{G}}) \frac{C_j(t)}{\bar{P}_j(t)}}_{\text{Government consumption}} + \underbrace{\sum_h (1+\tau^{\text{CF}}) \frac{I_h(t)}{\bar{P}^{\text{CF}}_h(t)} + \sum_i I_i(t)}_{\text{Gross fixed capital formation}} \\
& + \underbrace{\sum_i \Delta S_i(t) + \left(\Delta M_i(t) - \frac{1}{\beta_i} Y_i(t)\right)}_{\text{Changes in inventories}} + \underbrace{\sum_l (1+\tau^{\text{EXPORT}}) \frac{C_l(t)}{\bar{P}_l(t)}}_{\text{Exports}} - \underbrace{\sum_m Q_m(t)}_{\text{Imports}} .
\end{aligned}
$$

Similarly, inflation, which is measured by the GDP deflator, is the economy-wide average price of all goods and services produced and sold, where again all individual prices and sales are determined on the agent level by search and matching mechanisms:

$$
\begin{aligned}
& \textit{GDP deflator}(t) \\
& = \frac{\sum_i \tau^{\text{Y}}_i P_i(t) Y_i(t) + \sum_h \tau^{\text{VAT}} C_h(t) + \sum_h \tau^{\text{CF}} I_h(t) + \sum_j \tau^{\text{G}} C_j(t) + \sum_l \tau^{\text{EXPORT}} C_l(t) + \sum_i (1-\tau^{\text{Y}}_i) P_i(t) Y_i(t) - \sum_i \frac{1}{\beta_i} \bar{P}_i(t) Y_i(t)}{\sum_i \tau^{\text{Y}}_i Y_i(t) + \sum_h \tau^{\text{VAT}} \frac{C_h(t)}{\bar{P}_h(t)} + \sum_h \tau^{\text{CF}} \frac{I_h(t)}{\bar{P}^{\text{CF}}_h(t)} + \sum_j \tau^{\text{G}} \frac{C_j(t)}{\bar{P}_j(t)} + \sum_l \tau^{\text{EXPORT}} \frac{C_l(t)}{\bar{P}_l(t)} + \sum_i (1-\tau^{\text{Y}}_i) Y_i(t) - \sum_i \frac{1}{\beta_i} Y_i(t)}
\end{aligned}
$$

. Additionally, GDP, its components, and the respective deflators of each period t can be defined by the production and income approach, for details see Poledna et al. (2023).

3. Calibration to the euro area

This section provides a general overview of the data sources and the calibration of the model to the euro area. We start with an overview of the calibration procedure of model parameters, similar to that from Poledna et al. (2023), followed by a discussion of the initial conditions. The model is calibrated to the euro area to 60 reference quarters from the first quarter of 2005 to the last quarter of 2019, just before the COVID-19 pandemic started. Several subsamples of these reference quarters are considered to evaluate the forecast performance of the model around the Great Recession (Sections 4 to 6). To analyze the COVID-19 recession (Section 7), the model is calibrated to the last quarter of 2019. For each of these reference quarters, a wide range of parameters and initial conditions are calibrated to euro area economic statistics so that the model reproduces exactly the state of the economy in that quarter in terms of aggregate GDP, GDP components, and industry sizes. Parameter values and initial conditions for each reference quarter and all data are provided in the data repository.

3.1. Parameters

Model parameters are calibrated to micro and macro data from the euro area from national accounts, sector accounts, input-output tables, government statistics, and structural business statistics. They are calibrated so that a period t represents a quarter. At a scale of 1:1000 between the model and data, each agent in the model represents either a thousand natural persons (all persons who are permanently settled in the euro area) or euro area legal entities, such as corporations, government entities, or other institutions. In general, model parameters are obtained directly from data or are calculated from national accounting identities. Parameters associated with exogenous processes, such as exports and government consumption, are estimated from national accounts.

Population-related parameters are taken directly from auxiliary national accounting data and structural business statistics and are scaled as a representative sample of the euro area. Parameters concerning the numbers of firms in the industries are calibrated to the respective numbers in structural business statistics aggregated at the two-digit NACE level, with a total of 64 industries. Similarly, the total number of active and inactive persons is derived from auxiliary national accounting data on population and employment. In the scaling of 1:1000, the model is composed of approximately 342,650 households and 22,650 firms in 2019.

Parameters concerning productivity and technology coefficients, as well as capital formation and consumption coefficients, are taken directly from input-output tables or are derived from them. The parameters vary by industry (NACE classification) and are calibrated to the annual values for each reference quarter of a calendar year. For parameters that are calibrated based on the data of sectoral accounts and input-output tables (such as the productivity coefficients for labour and capital, depreciation rates, and average wage rates, we use cross-classification tables and structural business statistics (business demography) that link information by industry to each sector.

Tax rates and marginal propensities to consume or invest are calculated from national accounting identities. These rates are set to match the financial flows observed in input-output tables, government statistics, and sector accounts. Households' marginal propensity to consume and invest is calibrated such that consumption out of disposable income equals actual household consumption and investment in dwellings as obtained from input-output tables for the euro area. Capital ratios and the monetary authority's inflation target are set according to statutory guidelines, financial regulation (Basel III), and banking practices. Since the statutory guidelines and regulations did not change during the calibration period, these parameters are assumed to be constant for all reference quarters.

For exogenous processes such as exports and government consumption, parameters are estimated from national accounts (main aggregates). The growth rates of exports and the final consumption expenditure of the general government are assumed to follow an autoregressive process of lag order one (AR(1)). The coefficients of the respective AR(1) models are estimated from the observable time series of the euro area's real exports and real government consumption. These parameters are estimated over the sample from the first quarter of 1996 to the respective reference quarter of the calibration.

3.2. Initial conditions

Initial conditions, like model parameters, are set according to the procedure from Poledna et al. (2023) to represent the euro area economy. As model parameters, initial conditions are set for 60 reference quarters from the first quarter of 2005 to the last quarter of 2019. For each of these reference quarters, initial conditions are set such that stock variables in the model, including assets and liabilities of firms and households, as well as the general government, correspond with national balance sheets of the euro area in the respective quarter.

Initial conditions are set according to national balance sheets for non-financial and financial assets. National balance sheets are not available at the two-digit NACE level in official statistics and can only be obtained at the aggregate level for the institutional sectors, non-financial corporations (firms), financial corporations, the central bank (the financial sector), the general government, and households. Thus, the procedure to set the initial stock variables of the model involves distributing sector aggregates over individual agents.

For the firm sector, the only stock variable available at the two-digit NACE level in official statistics is non-financial fixed assets. These assets are available at the two-digit NACE level for a few euro area countries and have to be scaled up to the euro area as a whole. We thus first distribute industry-specific non-financial fixed assets (capital) onto the firm level and then use the non-financial assets as a proxy to allocate financial assets.

The distribution of firm sizes in industrial countries is well known to be highly skewed, with large numbers of small firms coexisting with small numbers of large firms (Ijiri and Simon, 1977; Axtell, 2001). Initial employment of firm i ($N_i(0) \quad \forall i \in I_s$) is therefore drawn from a power law distribution with exponent -2 (where $\sum_{i \in I_s} N_i(0) = N_s$ and $N_i(0) > 0$), which approximately corresponds to firm size distribution in the euro area. To determine initial production $Y_i(0)$ of the i-th firm, we use the initial employment by firm $N_i(0)$, and compute the corresponding amount of production by the productivity of labour per unit of output $\bar{\alpha}_i$:

$$Y_i(0) = Q_i^{\mathrm{d}}(0) = \bar{\alpha}_i N_i(0) \ .$$

The initial capital of firm i, $K_i(0)$, (i is part of industry s) is then obtained by dividing firm i's initial level of production $Y_i(0)$ by the productivity of capital κ_i and the desired rate of capacity utilization $\omega = .85$:

$$K_i(0) = \frac{Y_i(0)}{\kappa_i \omega} \ .$$

Thus, it is the share of the capital of the i-th firm in sector s as measured by production, accounting for the reserve capacity of its capital stock targeted by firm i.

Since a breakdown of financial and current assets is not available at the two-digit NACE level, we calibrate initial debt $L_i(0)$ to i-th individual firms by disaggregating total firm debts according to the share of the firms' capital stock $K_i(0)$ in the total capital stock $\sum_i K_i(0)$:

$$L_i(0) = L^{\mathrm{I}} \frac{K_i(0)}{\sum_i K_i(0)} \ ,$$

where the total amount of firm debt L^{I} is obtained from national balance sheet accounts. The total initial liquidity (deposits) of all firms as an aggregate, D^{I}, is set according to national balance sheet accounts. This aggregate is broken down onto single firms by the share of firm i's operating surplus in the overall operating surplus, where we assume that firm liquidity (deposits) moves in line with its production as a liquid form of working capital used for current expenditures:

$$D_i(0) = D^{\mathrm{I}} \frac{\max(\bar{\pi}_i Y_i(0), 0)}{\sum_i \max(\bar{\pi}_i Y_i(0), 0)} \ ,$$

where $\bar{\pi}_i = 1 - \left(1 + \tau^{\mathrm{SIF}}\right) \frac{\bar{w}_i}{\bar{\alpha}_i} - \frac{\delta_i}{\kappa_i} - \frac{1}{\beta_i} - \tau_i^{\mathrm{K}} - \tau_i^{\mathrm{Y}}$ is the operating margin. Thus overall, profitable firms have more cash at hand than unprofitable ones, and firms with a large capital stock are more in debt than less capital-intensive ones.

The initial equity of the financial sector ($E_k(0)$) is obtained from quarterly national financial balance sheets, and the initial government debt ($L^{\mathrm{G}}(0)$) is set according to the consolidated gross debt of the euro area members. Initial central bank's equity ($E^{\mathrm{CB}}(0)$) is the residual on the central bank's passive side, obtained by deducting initial bank reserves held ($D_k(0)$) and the initial net creditor/debtor position with the rest of the world ($D^{\mathrm{RoW}}(0) = 0$) from the central bank's assets (initial government debt ($L^{\mathrm{G}}(0)$)).

Initial personal assets (deposits) of the h-th person ($D_h(0)$) are obtained from national balance sheet accounts, which are disaggregated onto the individual level according to the share of each person's income of total income as a proxy for a person's wealth:

$$D_h(0) = D^{\mathrm{H}} \frac{Y_h(0)}{\sum_h Y_h(0)} ,$$

where D^{H} are the initial personal assets (deposits) of the household sector and $Y_h(0)$ is the initial income of the h-th person. The initial capital stock (residential structures such as dwellings) of the h-th person ($K_h(0)$) is obtained from national balance sheet accounts (residential structures of the household sector) and is again disaggregated onto the individual level according to the share of each person's income of total income as a proxy for the person's wealth:

$$K_h(0) = K^{\mathrm{H}} \frac{Y_h(0)}{\sum_h Y_h(0)} ,$$

where K^{H} is the initial lump-sum capital (dwellings) of the household sector.

4. Forecast performance

In this section, we validate the forecast performance of the ABM against standard models in a series of traditional out-of-sample root mean squared error (RMSE) forecast exercises along the lines of Poledna et al. (2023).[13] First, we benchmark the performance of the ABM against an unconstrained VAR model. Second, we compare the out-of-sample forecast performance of the ABM to that of a benchmark DSGE model. Third, we evaluate the conditional forecast performance of the ABM. Finally, we test the sectoral out-of-sample forecast performance of the ABM.

4.1. Comparison with an unconstrained VAR model

We start by comparing the out-of-sample forecast performance of the ABM with that of an unconstrained VAR model. To determine the optimal lag length of the VAR model, we use the Bayesian Information Criterion (BIC). For the entire period from 2005:Q1 to 2016:Q4, VAR models of lag order one minimize the BIC. For the optimized log-likelihoods and the forecast performance

13. The root mean squared error is defined as: $RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{T}(\hat{x}_t - x_t)^2}$, where $\hat{x_t}$ is the forecast and x_t the observed data point for time period t.

of VAR models of different lag orders, see Tables B.1 and B.2 in the appendix. Observable time series include the log difference of real GDP, real government consumption, and real exports of the euro area, as well as the log difference of the GDP deflator (inflation) and the 3-month Euribor. For this exercise, the unconstrained VAR model is initially estimated over the sample 1996:Q1 to 2005:Q1 (the sample 1995:Q2 to 1995:Q4 is used as a presample period). The VAR model is then used to forecast the five time series from 2005:Q2 to 2019:Q4, whereby the model is re-estimated every quarter. ABM forecasts are constructed analogously to the VAR: the ABM is calibrated to 48 different reference quarters of the calibration period 2005:Q1 to 2016:Q4. Then, we let the model run for 12 quarters from each of these starting points (i.e., in the last simulation up until 2019:Q4), where we average the results of 500 Monte Carlo simulations before we evaluate the forecasting accuracy. To test whether the ABM forecasts are significantly different in accuracy than the VAR(1) forecasts, we conduct (modified) Diebold-Mariano tests (Harvey et al., 1997), correcting for the overall length of the forecasting horizon.

Table 1. Out-of-sample forecast performance

	GDP	Inflation	Euribor	Government consumption	Exports
VAR(1)	*RMSE-statistic for different forecast horizons*				
1q	0.74	0.21	0.09	0.31	2.1
2q	1.63	0.21	0.18	0.48	4.88
4q	3.59	0.23	0.39	0.88	10.73
8q	6.98	0.25	0.71	1.75	20.46
12q	7.72	0.22	0.7	2.61	22.47
ABM	*Percentage gains (+) or losses (-) relative to VAR(1) model*				
1q	3.1 (0.31)	8.5 (0.10)	1.4 (0.47)	-98.4 (0.90)	-3.4 (0.54)
2q	11 (0.11)	4.6 (0.29)	9.1 (0.27)	-64.2 (0.85)	18.5 (0.24)
4q	30.4 (0.12)	9.6 (0.18)	26.1 (0.12)	-9 (0.82)	36.4 (0.13)
8q	45.3 (0.13)	14.6 (0.15)	46.7 (0.09)	-17.3 (0.82)	52.4 (0.11)
12q	38.1 (0.14)	9.1 (0.12)	42.5 (0.03)	-19.8 (0.84)	52.8 (0.07)
ABM (with financial frictions)	*Percentage gains (+) or losses (-) relative to VAR(1) model*				
1q	8.1 (0.40)	3.1 (0.43)	-2.3 (0.53)	-18.8 (0.96)	-4.5 (0.56)
2q	21.2 (0.21)	11.6 (0.26)	10.8 (0.30)	-4.6 (0.73)	18.6 (0.24)
4q	35.1 (0.14)	3.1 (0.33)	32.3 (0.13)	-7 (0.86)	36.6 (0.13)
8q	53.9 (0.14)	18 (0.14)	53.1 (0.10)	-14.1 (0.84)	51.9 (0.11)
12q	52.5 (0.16)	4.9 (0.29)	57.6 (0.03)	-12.4 (0.92)	52.2 (0.07)

Note: The VAR(1) model is estimated starting from 1996:Q1 to 2005:Q1 (1995:Q2 to 1995:Q4 is used as a presample period) and is then re-estimated each quarter from 2005:Q2 to 2016:Q4. The forecast period is 2005:Q2 to 2019:Q4. ABM results are obtained as an average of 500 Monte Carlo simulations. The values in brackets indicate the p-values of Diebold and Mariano (1995) tests, where we test whether the ABM forecasts are significantly different in accuracy than the VAR(1) forecasts (the null hypothesis of the test is that there is no difference between two competing forecasts).

Table 1 reports the out-of-sample RMSEs for different forecast horizons of 1, 2, 4, 8, and 12 quarters over the period 2005:Q2 to 2019:Q4. In parentheses, we show the p-values of (modified) Diebold-Mariano tests, where we test whether the ABM forecasts are significantly different in accuracy than the VAR(1) forecasts (the null hypothesis of the test is that the ABM is less accurate than the VAR(1)). The out-of-sample forecast statistic shows a good forecast performance of the ABM relative to the VAR(1) model. For GDP, the 3-month Euribor and exports of the euro area, the ABM does better than the VAR(1) model by a considerable margin for almost all horizons. For most horizons, these margins are significant at or close to the ten per cent level. The bad empirical performance of unconstrained VARs is not too surprising, as it is known that overparameterized models typically perform poorly in out-of-sample forecast exercises (Smets and Wouters, 2007).

4.2. Comparison with a benchmark DSGE model

Next, we compare the out-of-sample forecast performance of the ABM to that of a benchmark DSGE model. As a benchmark DSGE model, we use the model of Smets and Wouters (2007). We estimate the benchmark DSGE model for the euro area on the following set of seven variables: the log difference of real GDP, real household consumption, real fixed investment, and the real wage, as well as detrended log hours worked, the log difference of the GDP deflator (inflation), and the 3-month Euribor. As variables for this comparison, we choose real GDP and the main components—real household consumption and real fixed investment—as well as log differences of the GDP deflator (inflation) and the 3-month Euribor. To provide a benchmark for the ABM and DSGE forecasts, we estimate AR models on the log difference of these variables. As above, we use the BIC to determine the optimal lag length for the AR models. For the optimized log-likelihoods and the forecast performance of AR models of different lag orders, see Tables B.3 and B.4 in the appendix. The DSGE and AR models are initially estimated over the sample 1996:Q1 to 2005:Q1 (the sample 1995:Q2 to 1995:Q4 is used as a presample period). The models are then used to forecast the five time series from 2005:Q2 to 2019:Q4, whereby the models are re-estimated each quarter. The DSGE model is estimated using Bayesian methods.[14] Again, ABM forecasts are constructed analogously to the comparison with the VAR model in Section 4.1.

14. DSGE estimations are done with Dynare, see http://www.dynare.org/. A sample of 250,000 draws was created (neglecting the first 50,000 draws).

Table 2. Out-of-sample forecast performance in comparison to DSGE model

	GDP	Inflation	Euribor	Household consumption	Investment
AR(1)	*RMSE-statistic for different forecast horizons*				
1q	0.71	0.2	0.1	0.36	2.15
2q	1.54	0.21	0.18	0.68	3.12
4q	3.35	0.21	0.28	1.42	5.35
8q	7.5	0.21	0.37	2.82	8.81
12q	12.36	0.2	0.4	4.11	11.44
DSGE	*Percentage gains (+) or losses (-) relative to AR(1) model*				
1q	9.5 (0.37)	-5.4 (0.81)	14.2 (0.08)	-15 (0.93)	-0.4 (0.53)
2q	22.1 (0.22)	-0.6 (0.55)	1.3 (0.46)	-19.4 (0.84)	4 (0.37)
4q	34 (0.21)	-5.5 (0.71)	-21 (0.90)	-18.9 (0.77)	5.8 (0.40)
8q	53.8 (0.18)	-27.6 (0.90)	-56.3 (0.94)	-7.7 (0.62)	13.7 (0.32)
12q	69.4 (0.14)	-40.6 (0.88)	-85.9 (0.94)	6.3 (0.39)	23.9 (0.20)
ABM	*Percentage gains (+) or losses (-) relative to AR(1) model*				
1q	0.3 (0.18)	1.4 (0.03)	14.4 (0.08)	-95.4 (0.83)	17.6 (0.05)
2q	5.6 (0.19)	0.2 (0.42)	7.6 (0.06)	-97.4 (0.81)	12.9 (0.07)
4q	25.4 (0.16)	1.3 (0.05)	-2.2 (0.65)	-32.8 (0.78)	8.3 (0.11)
8q	49.1 (0.16)	0.8 (0.16)	-1 (0.56)	10.3 (0.20)	5.9 (0.20)
12q	61.3 (0.14)	0 (0.49)	-1.6 (0.57)	9.8 (0.20)	8.1 (0.12)
ABM (with financial frictions)	*Percentage gains (+) or losses (-) relative to AR(1) model*				
1q	5.4 (0.42)	-4.4 (0.62)	11.2 (0.04)	-17.7 (0.90)	11.7 (0.13)
2q	16.4 (0.26)	7.6 (0.26)	9.4 (0.14)	-4.4 (0.59)	6.9 (0.16)
4q	30.4 (0.18)	-5.9 (0.75)	6.4 (0.22)	3.9 (0.44)	18.1 (0.18)
8q	57.1 (0.16)	4.8 (0.30)	11.1 (0.07)	17.7 (0.30)	22.6 (0.27)
12q	70.3 (0.15)	-4.6 (0.78)	25.1 (0.04)	27.4 (0.29)	18.4 (0.31)

Note: The AR(1) and the DSGE model are estimated starting from 1996:Q1 to 2005:Q1 (1995:Q2 to 1995:Q4 is used as a presample period) and are then re-estimated each quarter from 2005:Q2 to 2016:Q4. The forecast period is 2005:Q2 to 2019:Q4. ABM results are obtained as an average of 500 Monte Carlo simulations. The DSGE model is estimated using Bayesian methods where a sample of 250,000 draws was created (neglecting the first 50,000 draws). The values in brackets indicate the p-values of Diebold and Mariano (1995) tests, where we test whether the ABM forecasts are significantly different in accuracy than the AR(1) forecasts (the null hypothesis of the test is that there is no difference between two competing forecasts).

The good forecast performance of the ABM relative to the unconstrained VAR is confirmed by comparison with the benchmark DSGE model reported in Table 2. The out-of-sample forecast statistic shows a good forecast performance of the ABM and the benchmark DSGE model relative to the AR(1) models. This performance is in line with the good forecast performance of the benchmark DSGE model for the U.S. economy (Smets and Wouters, 2007). Again, for GDP, the 3-month Euribor, as well as household consumption, and fixed investment, the ABM does better than the AR(1) models by a considerable margin for almost all horizons. As above, for longer forecast horizons, these margins are significant at or close to the ten per cent level. Equally, by a considerable margin, the benchmark DSGE model improves forecasts of GDP on the AR models. Surprisingly, however, the benchmark DSGE model performs worse than the simple AR(1) model for inflation and the 3-month Euribor.

4.3. Conditional forecasts

As a further exercise, we compare the conditional forecast performance of the ABM with that of an unconstrained VAR model with exogenous variables (VARX). In this exercise, forecasts of the models are conditional on the observed data for real government consumption and real exports of the euro area, which are exogenous variables in the ABM.[15] Observed data of the exogenous variables are treated as deterministic future responses that are known in advance, for example, set by policy. Thus agents in the ABM have, for example, (forward-looking) expectations for government austerity.[16] To determine the optimal lag length of the VARX model, we use the Bayesian Information Criterion (BIC) again. For the optimized log-likelihoods and the forecast performance of VARX models of different lag orders, see Tables B.5 and B.6 in the appendix. In the VAR model, we enter GDP, the GDP deflator, government consumption, and exports of the euro area in log levels. As above, the VARX model is initially estimated over the sample 1996:Q1 to 2005:Q1 (the sample 1995:Q2 to 1995:Q4 is used as a presample period). The VARX model is then used to forecast real GDP, inflation and the 3-month Euribor conditional on observed real government consumption and real exports of the euro area from 2005:Q2 to 2019:Q4, whereby the model is re-estimated each quarter. Again, ABM forecasts are constructed analogously to the comparison with the VAR model in Section 4.1.

Similarly to the good out-of-sample forecast performance of the ABM relative to the unconstrained VAR, the ABM also improves on conditional forecasts relative to the VARX(1) model reported in Table 3. Here the ABM does better than the VARX(1) model by a considerable margin for GDP and inflation for almost all horizons. As above, for most horizons, these margins are significant at or close to the ten per cent level. This is not too surprising as ABMs are better placed to produce conditional forecasts (Haldane and Turrell, 2018).

4.4. Sectoral forecasts

Finally, we test the sectoral out-of-sample forecast performance of the ABM. In this exercise, we decompose ABM forecasts from Sections 4.1 and 4.2 for different sectors by economic activities.

---

15. Thus, in this setup, we replace Equations (11) and (17) of the model and set government consumption and exports according to observed data.

16. For exercises that explicitly analyse export demand shocks during the Great Recession and government austerity during the European sovereign debt crisis see Sections 5 and 6.

Table 3. Conditional forecast performance

	GDP	Inflation	Euribor
VARX(1)	*RMSE-statistic for different forecast horizons*		
1q	0.63	0.22	0.1
2q	1.07	0.27	0.17
4q	1.65	0.37	0.3
8q	2.21	0.35	0.59
12q	2.62	0.26	0.85
ABM (conditional forecasts)	*Percentage gains (+) or losses (-) relative to VARX(1) model*		
1q	-4.4 (0.81)	4.3 (0.39)	-2.3 (0.56)
2q	9.9 (0.06)	18.3 (0.14)	2.3 (0.44)
4q	19.4 (0.04)	34.7 (0.08)	15.5 (0.13)
8q	13.2 (0.20)	24.8 (0.03)	39.9 (0.02)
12q	6.1 (0.24)	-6.8 (0.72)	56.2 (0.00)
ABM (with financial frictions)	*Percentage gains (+) or losses (-) relative to VARX(1) model*		
1q	-2.8 (0.72)	4.9 (0.37)	-2 (0.55)
2q	21.3 (0.01)	19 (0.13)	2.2 (0.44)
4q	34.1 (0.02)	33 (0.08)	13.2 (0.13)
8q	34.5 (0.09)	23.9 (0.03)	31.6 (0.02)
12q	44.1 (0.07)	-5.5 (0.69)	38.7 (0.00)

Note: Forecasts are conditional on real government consumption and real exports from the euro area. The VARX(1) model is estimated starting from 1996:Q1 to 2005:Q1 (1995:Q2 to 1995:Q4 is used as a presample period) and is then re-estimated each quarter from 2005:Q2 to 2016:Q4. The forecast period is 2005:Q2 to 2019:Q4. ABM results are obtained as an average of 500 Monte Carlo simulations. The values in brackets indicate the p-values of Diebold and Mariano (1995) tests, where we test whether the ABM forecasts are significantly different in accuracy than the VARX(1) forecasts (the null hypothesis of the test is that there is no difference between two competing forecasts).

Specifically, sectoral gross value added (GVA) is disaggregated for ten economic activities according to the statistical classification of economic activities in the European Community (NACE Rev. 2). See Table B.7 in the appendix for details. As a benchmark, we use AR models estimated on the log differences of sectoral GVA. As above, AR models are initially estimated over the sample 1996:Q1 to 2005:Q1 (the sample 1995:Q2 to 1995:Q4 is used as a presample period) and are then used to forecast the ten time series from 2005:Q2 to 2019:Q4, with the model being re-estimated every quarter for the periods 2005:Q2 to 2016:Q4. Again, ABM forecasts are constructed analogously to the comparison with the VAR model in Section 4.1.

Table 4 shows the sectorally disaggregated forecast performance of the ABM in comparison to AR(1) models. Similarly to the good out-of-sample forecast performance for real GDP and the main components, the ABM also improves on sectorial forecasts relative to AR(1) models. For most sectors and forecast horizons, the ABM does better than the AR(1) models by a considerable margin. These performance gains are, in general, significant or close to the ten per cent significance threshold according to the (modified) Diebold-Mariano tests.

Table 4. Out-of-sample forecast performance of sectoral gross value added (GVA)

	A	B, C, D and E	F	G, H and I	J	K	L	M and N	O, P and Q	R and S
AR(1)	*RMSE-statistic for different forecast horizons*									
1q	2.7	1.66	1.56	0.76	0.93	1.29	0.84	1.09	0.49	0.53
2q	4.81	3.55	2.73	1.4	1.56	2.14	1.56	1.93	0.8	0.92
4q	8.33	7.35	5.1	2.52	3.05	3.41	2.9	3.5	1.5	1.76
8q	10.53	14.67	9.64	4.45	6.13	5.47	5.05	5.68	2.92	3.36
12q	9.06	22.16	14.01	5.81	9.17	7.48	6.51	7.11	4.32	5.08
ABM	*Percentage gains (+) or losses (-) relative to AR(1) model*									
1q	-6.2 (0.79)	11.5 (0.28)	6.6 (0.06)	-3.3 (0.58)	3.4 (0.44)	-13.4 (0.93)	6.5 (0.37)	19.9 (0.18)	-63.9 (0.86)	-23.8 (0.70)
2q	-2.3 (0.60)	19 (0.19)	7.7 (0.14)	-9.6 (0.69)	6.6 (0.30)	-17.2 (0.83)	-0.7 (0.52)	15.1 (0.20)	-22.7 (0.70)	-25.1 (0.75)
4q	1.5 (0.44)	28.9 (0.17)	8.8 (0.20)	-0.4 (0.54)	19.1 (0.04)	-7.5 (0.69)	6.2 (0.32)	8.5 (0.26)	29.2 (0.01)	3.7 (0.36)
8q	9.4 (0.12)	46.6 (0.15)	1.5 (0.36)	-1.3 (0.70)	21.1 (0.11)	14.3 (0.00)	20.1 (0.12)	8.8 (0.29)	26.2 (0.06)	21 (0.03)
12q	1.1 (0.48)	57.9 (0.13)	2.6 (0.26)	-3.8 (0.81)	23.5 (0.15)	16.9 (0.00)	32.8 (0.01)	13.8 (0.26)	24.9 (0.11)	21.8 (0.03)
ABM (with financial frictions)	*Percentage gains (+) or losses (-) relative to AR(1) model*									
1q	-7.3 (0.82)	7.5 (0.37)	10.3 (0.01)	6 (0.12)	15.1 (0.07)	-0.3 (0.53)	15.1 (0.07)	13.2 (0.19)	-7.7 (0.65)	17.4 (0.14)
2q	-3.4 (0.66)	19.3 (0.20)	11.9 (0.05)	7 (0.18)	23 (0.03)	3 (0.32)	14 (0.19)	12.9 (0.24)	21.2 (0.12)	29.4 (0.02)
4q	3.3 (0.35)	31.4 (0.17)	12.9 (0.12)	5.3 (0.29)	24.3 (0.11)	12.9 (0.06)	14.2 (0.25)	8.3 (0.34)	35.2 (0.02)	32.7 (0.02)
8q	12.4 (0.04)	53.4 (0.15)	35.8 (0.13)	19.3 (0.21)	40 (0.17)	21.1 (0.02)	33.3 (0.10)	22.5 (0.29)	29.2 (0.05)	47.7 (0.01)
12q	11.2 (0.31)	64.4 (0.13)	35.4 (0.14)	22 (0.26)	41.8 (0.21)	36.1 (0.00)	46.2 (0.01)	22.1 (0.34)	30.2 (0.08)	52 (0.01)

Note: GVA is shown for the sectors Agriculture, forestry and fishing (A); Industry (except construction) (B, C, D and E); Manufacturing (C); Construction (F); Wholesale and retail trade, transport, accommodation and food service activities (G, H and I); Information and communication (J); Financial and insurance activities (K); Real estate activities (L); Professional, scientific and technical activities, as well as administrative and support service activities (M and N); Public administration, defence, education, human health and social work activities (O, P and Q); Arts, entertainment, and recreation, as well as other service activities (R and S). AR(1) models are estimated starting from 1996:Q1 to 2005:Q1 (1995:Q2 to 1995:Q4 is used as a presample period) and are then re-estimated each quarter from 2005:Q2 to 2016:Q4. The forecast period is 2005:Q2 to 2019:Q4. ABM results are obtained as an average of 500 Monte Carlo simulations. The values in brackets indicate the p-values of Diebold and Mariano (1995) tests, where we test whether the ABM forecasts are significantly different in accuracy than the AR(1) forecasts (the null hypothesis of the test is that there is no difference between two competing forecasts).

5. The Financial crisis of 2007-2008 and the Great Recession

In this section, we assess the performance of our ABM during the Financial crisis of 2007-2008 and the subsequent Great Recession. First, we compare the forecast performance of the ABM against standard models in a series of traditional RMSE forecast exercises during the most intense phase of the Great Recession in the euro area. Second, we attempt to shed some light on the causes of the recession and the slow recovery with the ABM. Finally, we break down ABM simulation for different sectors by economic activities and demonstrate structural change in the euro area economy as a consequence of the Great Recession.

5.1. Comparison with an unconstrained VAR model

We start by comparing the forecast performance of the ABM against an unconstrained VAR model during the most intense phase of the Great Recession in the euro area in a similar setup as in Section 4.1. Here we choose a forecast period from 2007:Q2 to 2011:Q4. Forecasts covering these quarters are of interest as they are carried out during the Financial crisis of 2007-2008 and cover the peak of the Great Recession as well as the recovery in the euro area. For this exercise, the unconstrained VAR model is initially estimated over the sample 1996:Q1 to 2007:Q1 (the sample 1995:Q2 to 1995:Q4 is used as a presample period). The VAR model is then re-estimated each quarter from the subsample from 2007:Q2 to 2008:Q4. ABM forecasts are constructed analogously to the comparison with the VAR model in Section 4.1: the ABM is calibrated to 12 different reference quarters of the calibration period 2007:Q1 to 2008:Q4. Then, we let the model run for 12 quarters from each of these starting points (i.e., in the last simulation up until 2011:Q4), where we average the results of 500 Monte Carlo simulations before we evaluate the forecasting accuracy.

Table 5. Out-of-sample forecast performance during the Great Recession

	GDP	Inflation	Euribor	Government consumption	Exports
VAR(1)	*RMSE-statistic for different forecast horizons*				
1q	0.72	0.21	0.11	0.33	2.19
2q	1.72	0.28	0.19	0.45	5.15
4q	3.66	0.31	0.34	0.65	10.69
8q	6.35	0.29	0.5	0.96	17.75
12q	7.99	0.28	0.52	1.62	20.99
ABM	*Percentage gains (+) or losses (-) relative to VAR(1) model*				
1q	-19.5 (0.88)	6.6 (0.16)	-34.9 (0.82)	-86.2 (0.88)	-50.2 (0.87)
2q	-11.5 (0.90)	8.2 (0.17)	-36.7 (0.81)	-43.2 (0.93)	-20.4 (0.86)
4q	-7.7 (0.85)	6.2 (0.11)	-31.3 (0.82)	9.9 (0.37)	1.3 (0.38)
8q	-4.9 (0.91)	4.8 (0.25)	-20.1 (0.92)	-26.5 (0.76)	17.5 (0.11)
12q	-5.1 (0.90)	0.4 (0.40)	-6.1 (0.89)	-45.1 (0.97)	29.5 (0.00)
ABM (with financial frictions)	*Percentage gains (+) or losses (-) relative to VAR(1) model*				
1q	-58.8 (0.88)	-4.4 (0.57)	-51.4 (0.85)	-18.9 (0.94)	-59.2 (0.87)
2q	-26.9 (0.87)	28.2 (0.13)	-51.9 (0.84)	-10.2 (0.73)	-25 (0.86)
4q	-5.9 (0.89)	8.9 (0.08)	-38.3 (0.85)	4.6 (0.44)	-0.7 (0.58)
8q	32.5 (0.03)	15.4 (0.17)	-12.2 (0.99)	-20.3 (0.97)	16 (0.12)
12q	52.8 (0.00)	5.6 (0.13)	14.1 (0.01)	-13.9 (0.92)	29.8 (0.00)

Note: The VAR(1) model is estimated starting from 1996:Q1 to 2007:Q1 (1995:Q2 to 1995:Q4 is used as a presample period) and is then re-estimated each quarter from the subsample from 2007:Q2 to 2008:Q4. The forecast period is 2007:Q2 to 2011:Q4. ABM results are obtained as an average of 500 Monte Carlo simulations. The values in brackets indicate the p-values of Diebold and Mariano (1995) tests, where we test whether the ABM forecasts are significantly different in accuracy than the VAR(1) forecasts (the null hypothesis of the test is that there is no difference between two competing forecasts).

Table 5 reports the out-of-sample RMSEs for different forecast horizons of 1, 2, 4, 8, and 12 quarters over the period 2007:Q2 to 2011:Q4, covering the Financial crisis of 2007-2008, the peak of the Great Recession as well as the recovery in the euro area. The out-of-sample forecast statistic shows a good forecast performance of the VAR(1) model up to the one-year-ahead horizon. Over longer horizons, up to three years, however, the ABM does considerably better than the VAR model. For inflation, the ABM does better than the VAR(1) model by a considerable margin for almost all horizons. The good forecasting performance of the ABM over longer horizons is not too surprising, as the ABM allows for non-linear responses and the possibility of endogenous economic crises to occur in the model without exogenous shocks.

5.2. Comparison with a benchmark DSGE model

Next, we compare the out-of-sample forecast performance of the ABM to that of the benchmark DSGE model during the most intense phase of the Great Recession in the euro area in a similar setup as in Section 4.2. As in Section 4.2, we estimate AR models as a benchmark for the DSGE model and the ABM. As above, the forecast period is from 2007:Q2 to 2011:Q4 to cover the peak of the Great Recession as well as the recovery in the euro area. For this exercise, the benchmark DSGE model and the AR models are initially estimated over the sample 1996:Q1 to 2007:Q1 (the sample 1995:Q2 to 1995:Q4 is used as a presample period). All models are then re-estimated each quarter from the subsample from 2007:Q2 to 2008:Q4. ABM forecasts are constructed analogously to the comparison with the VAR model in Section 5.1.

Again, the good forecast performance over longer horizons of the ABM relative to the unconstrained VAR model during the Great Recession is confirmed by comparison with the benchmark DSGE model reported in Table 6. As in Table 2, the benchmark DSGE model and the ABM show a good out-of-sample forecast performance relative to the AR models during the Great Recession. This good performance is somewhat surprising as the benchmark DSGE model failed to predict the crisis and has a clear tendency to forecast a quick recovery (Lindé et al., 2016). Nonetheless, both theoretical models do considerably better than the AR models over longer horizons of up to three years for all variables. The ABM, again, improves the forecasts of household consumption and investment by a considerable margin for almost all horizons.

Next, we compare the out-of-sample forecast performance of the ABM to that of the benchmark DSGE model estimated on data up to the fourth quarter of 2006. Similarly, we calibrate the ABM with data up to the fourth quarter of 2006 and make forecasts for twelve quarters ahead up to the fourth quarter of 2009. Forecasts covering these quarters are of particular interest as they are starting well before the Financial crisis of 2007-2008 and contain the peak of the Great Recession in the euro area. In the euro area, GDP plummeted in 2008:Q4 (about -6.6 per cent at an annualized quarterly rate) and in 2009:Q1 (about -11.4 per cent at an annualized rate), while the recovery started in

Table 6. Out-of-sample forecast performance during the Great Recession in comparison to DSGE model

	GDP	Inflation	Euribor	Household consumption	Investment
AR(1)	*RMSE-statistic for different forecast horizons*				
1q	0.86	0.2	0.18	0.55	2.32
2q	1.87	0.26	0.3	1.01	4.65
4q	3.81	0.3	0.48	2.07	8.86
8q	6.56	0.28	0.6	3.95	15.29
12q	8.16	0.28	0.55	5.76	19.5
DSGE	*Percentage gains (+) or losses (-) relative to AR(1) model*				
1q	-20.1 (0.80)	-7.5 (0.69)	18 (0.07)	20.1 (0.10)	14.1 (0.12)
2q	3.1 (0.43)	-7.9 (0.74)	13.6 (0.04)	40.1 (0.10)	19.3 (0.20)
4q	21.6 (0.24)	-7.6 (0.71)	8.8 (0.04)	39.3 (0.13)	26.1 (0.23)
8q	39.7 (0.00)	-6.7 (0.98)	2.5 (0.20)	38.1 (0.00)	38.5 (0.03)
12q	50.4 (0.00)	-2.9 (0.85)	-10.1 (0.99)	37.4 (0.00)	47 (0.00)
ABM	*Percentage gains (+) or losses (-) relative to AR(1) model*				
1q	0.9 (0.09)	1.9 (0.15)	15 (0.14)	12.4 (0.20)	17.7 (0.09)
2q	-2.5 (0.83)	1.6 (0.09)	13.2 (0.15)	18.6 (0.18)	15.9 (0.14)
4q	-3.2 (0.76)	1.1 (0.01)	6.9 (0.20)	24 (0.11)	9.9 (0.19)
8q	-1.6 (0.65)	1.1 (0.23)	0.4 (0.44)	11.6 (0.00)	9.6 (0.00)
12q	-2.9 (0.84)	-0.9 (0.77)	-0.5 (0.68)	4.2 (0.05)	11.4 (0.00)
ABM (with financial frictions)	*Percentage gains (+) or losses (-) relative to AR(1) model*				
1q	-31.6 (0.87)	-9.6 (0.65)	4.6 (0.20)	26.6 (0.03)	4.5 (0.16)
2q	-16.6 (0.84)	23.1 (0.11)	3.6 (0.26)	30.9 (0.09)	7.4 (0.19)
4q	-1.5 (0.60)	4 (0.29)	2 (0.37)	27.8 (0.13)	29.4 (0.19)
8q	34.6 (0.00)	12 (0.13)	7 (0.29)	43.1 (0.00)	69.8 (0.01)
12q	53.8 (0.00)	4.4 (0.20)	18.6 (0.00)	61.8 (0.00)	86.4 (0.00)

Note: The AR(1) and the DSGE model are estimated starting from 1996:Q1 to 2007:Q1 (1995:Q2 to 1995:Q4 is used as a presample period) and are then re-estimated each quarter from the subsample from 2007:Q2 to 2008:Q4. The forecast period is 2007:Q2 to 2011:Q4. ABM results are obtained as an average of 500 Monte Carlo simulations. The DSGE model is estimated using Bayesian methods where a sample of 250,000 draws was created (neglecting the first 50,000 draws). The values in brackets indicate the p-values of Diebold and Mariano (1995) tests, where we test whether the ABM forecasts are significantly different in accuracy than the AR(1) forecasts (the null hypothesis of the test is that there is no difference between two competing forecasts).

2009:Q3 (about 1.3 per cent at an annualized quarterly rate) and in 2009:Q4 (about 2.2 per cent at an annualized rate).

Figure 1 shows the distribution of GDP growth rates for the ABM, an AR(1) model estimated on data up to the fourth quarter of 2006, and observed Eurostat data for the euro area from 2006 to 2009. For the AR(1) model (blue bars), GDP growth rates are normally distributed, while in contrast, for ABM (grey bars), the distribution of GDP growth rates clearly has a fat tail. However, while the ABM shows the probability of a serious crisis arising endogenously, the projected GDP drops are less pronounced, as in the observed data for the euro area (black bars). Figures 2 to 4 show out-of-sample forecasts of the ABM, AR(1) models, and the benchmark DSGE model estimated on data up to the fourth quarter of 2006. In Figure 2, we show GDP growth, the inflation rate, and the

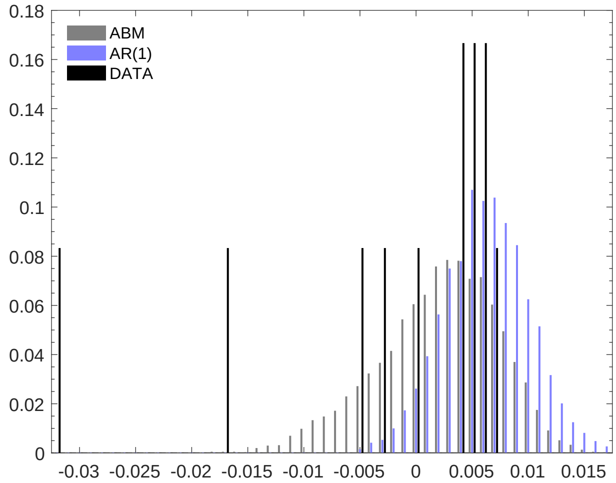

FIGURE 1. Distribution of GDP growth rates from 2006 to 2009. GDP growth rates are shown for the ABM (grey bars), the AR(1) (blue bars), and observed Eurostat data for the euro area (black bars).

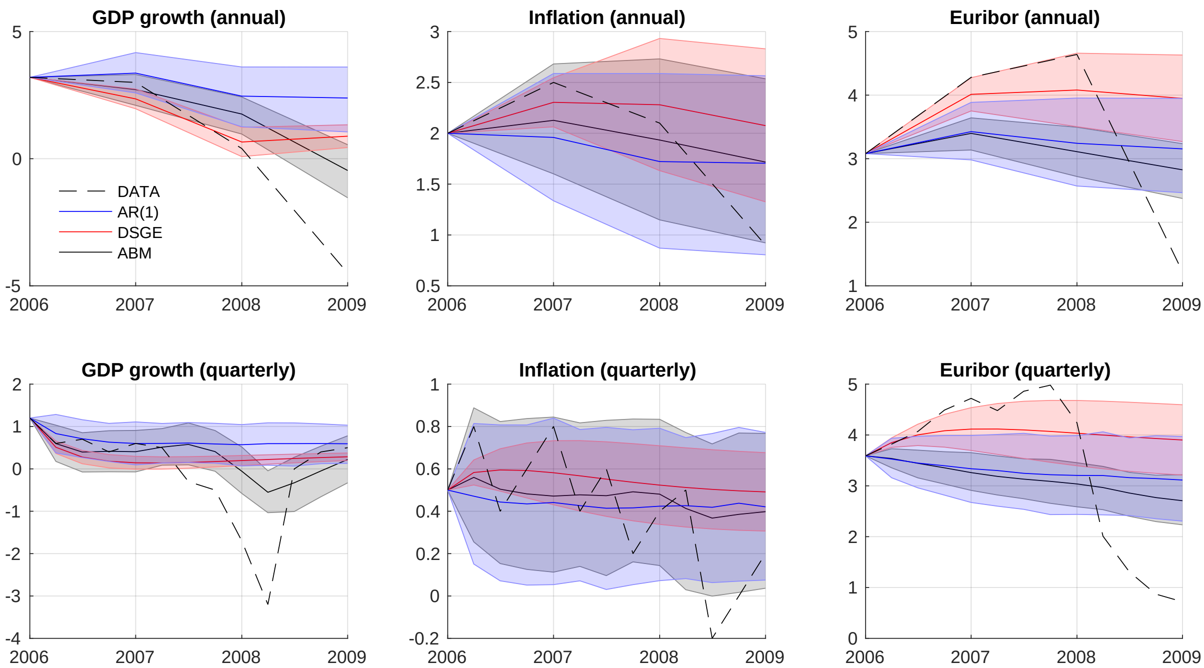

FIGURE 2. Out-of-sample forecasts estimated on data up to the fourth quarter of 2006. Top figures show GDP growth, inflation and the 3-month Euribor on an annualized basis, bottom figures depict quarterly growth, inflation and interest rates for ABM (black line), DSGE (red line), AR(1) (blue line), and observed Eurostat data for the euro area (dashed line). A 90 per cent confidence interval is plotted around the mean trajectory. Model results are obtained as an average over 500 Monte Carlo simulations. The DSGE model is estimated using Bayesian methods where a sample of 250,000 draws was created (neglecting the first 50,000 draws).

3-month Euribor—annually (top) and quarterly (bottom). Clearly, the linear models—the linearized DSGE model (red line) and the AR(1) model (blue line)—do not predict the recession. In particular, the linearized DSGE model does not include a financial sector and thus, cannot be used to explicitly model the effects of a financial crisis (Lindé et al., 2016; Lindé, 2018). Thus, for the linearized DSGE model, the Great Recession was only a highly unlikely tail event (Del Negro and Schorfheide, 2013; Lindé et al., 2016; Lindé, 2018). In contrast, the ABM explicitly models the financial sector and includes a financial accelerator and debt-financed investment. In ABM simulations, clearly, an endogenous crisis occurs (Figure 2 (black line)). This endogenous crisis is mainly driven by credit rationing and the financial conditions of firms in the euro area. Because export demand follows the trend from exogenous projections and an export demand shock does not occur in ABM simulations (Figures 3 and 4 (black line)), the recession is less pronounced, as in the observed data for the euro area (see also Section 5.3). Forecasts of the inflation rate and the 3-month Euribor are more or less similar for the three model classes, whereby the ABM seems to capture the falling inflation somewhat better.

In Figures 3 and 4, we compare forecasts for the levels of GDP and the main components (household consumption, government consumption, investment, exports and imports) of the ABM and the benchmark DSGE model. While Figure 3 shows annual levels, in Figure 4, we show the respective quarterly levels for GDP and the main components. Evidently, the benchmark DSGE model (red line) improves on forecasts of the AR(1) model (blue line) for GDP, household consumption and investment.[17] While the AR(1) model merely extrapolates the short-term trends, the benchmark DSGE model accurately forecasts the long-term trends. Forecasts of the ABM (black line) are, in general, more accurate, as the ABM predicts the recession. However, the recession is less pronounced as in the observed data because export demand follows a trend from exogenous projections, and an export demand shock does not occur (Figures 3 and 4). ABM forecasts for investment are particularly accurate.

17. We do not show forecasts for government consumption, exports and imports of the benchmark DSGE model and the AR(1) because the DSGE does not include a full GDP identity.

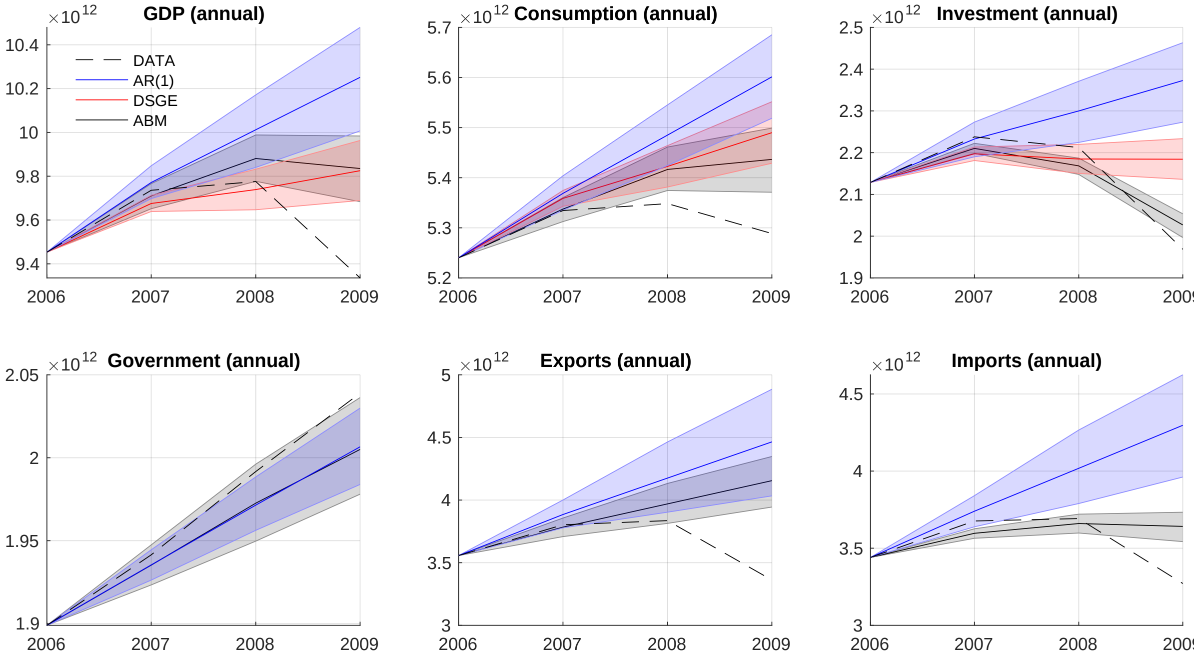

FIGURE 3. Out-of-sample forecasts estimated on data up to the fourth quarter of 2006. Figures show GDP (annually, in Euro and in real terms with base year 2010), household consumption (annually, in Euro and in real terms with base year 2010), fixed investment (annually, in Euro and in real terms with base year 2010), government consumption (annually, in Euro and in real terms with base year 2010), exports (annually, in Euro and in real terms with base year 2010), and imports (annually, in Euro and in real terms with base year 2010) for ABM (black line), DSGE (red line), AR(1) (blue line), and observed Eurostat data for the euro area (dashed line). A 90 per cent confidence interval is plotted around the mean trajectory. Model results are obtained as an average over 500 Monte Carlo simulations. The DSGE model is estimated using Bayesian methods where a sample of 250,000 draws was created (neglecting the first 50,000 draws).

5.3. Main Driving Forces of the Great Recession

After having shown that the forecast performance of the ABM improves on standard models, we use it to better understand the Great Recession in the euro area. In the ABM, GDP and other macroeconomic aggregates are an emergent property of the model. Therefore, it is not immediate to detect the main mechanisms that drive the macroeconomic aggregates. In this section, we try to shed some light on these mechanisms. Natural candidates for the main driving forces during the Great Recession in the euro area are a credit crunch—the reduction in the general availability of loans—and a shock from the rest of the world as the financial crisis and the subsequent recession did not originate

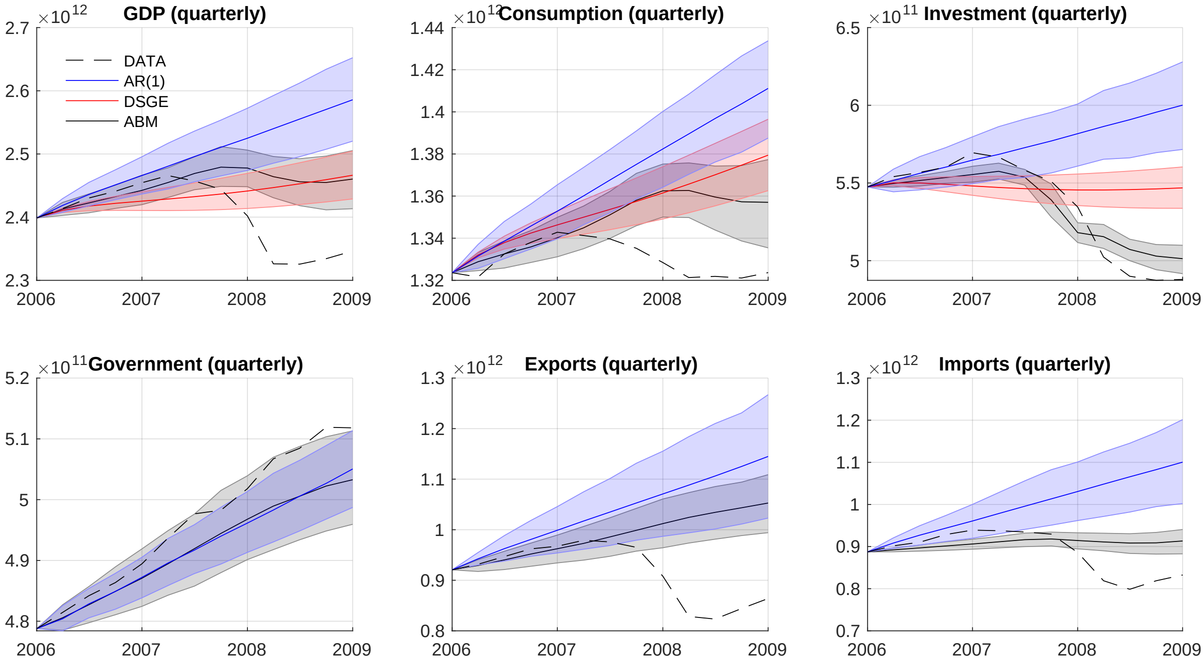

FIGURE 4. Out-of-sample forecasts estimated on data up to the fourth quarter of 2006. Figure show GDP (quarterly, in Euro and in real terms with base year 2010), household consumption (quarterly, in Euro and in real terms with base year 2010), fixed investment (quarterly, in Euro and in real terms with base year 2010), government consumption (quarterly, in Euro and in real terms with base year 2010), exports (quarterly, in Euro and in real terms with base year 2010), and imports (quarterly, in Euro and in real terms with base year 2010) for ABM (black line), DSGE (red line), AR(1) forecasts (blue line), and observed Eurostat data for the euro area (dashed line). A 90 per cent confidence interval is plotted around the mean trajectory. Model results are obtained as an average over 500 Monte Carlo simulations. The DSGE model is estimated using Bayesian methods where a sample of 250,000 draws was created (neglecting the first 50,000 draws).

in the euro area.[18] Thus in this exercise, we analyze the Great Recession with three different setups of the ABM focusing on credit tightening and a shock from the rest of the world on the euro area. First, we consider the ABM from Poledna et al. (2023) without a financial accelerator mechanism. Second, the model from Section 2 with a financial accelerator and debt-financed investment with collateralized loans where firms' ability to borrow depends on the market value of their assets and their net worth. Third, the model from Section 2 with the financial accelerator and, additionally, an

18. For example, the peak of the Great Recession in the U.S. was in the last quarter of 2008 (roughly -9.8 per cent at an annualized quarterly rate), while in the euro area, the height of the recession was in the first quarter of 2009 (about -11.4 per cent at an annualized rate).

exogenous shock from the rest of the world. In this setup, ABM simulations are conditional on the observed data for real exports of the euro area.

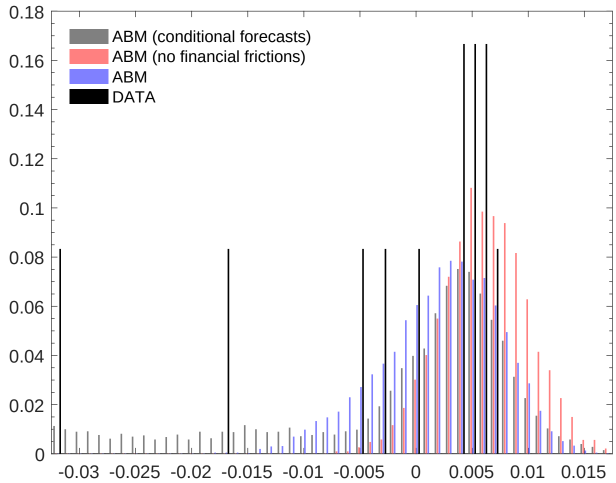

FIGURE 5. Distribution of GDP growth rates from 2006 to 2009. GDP growth rates are shown for the ABM with financial frictions and exogenous export demand shocks (grey bars), the ABM without financial frictions (red bars), the ABM with financial frictions (blue bars), and observed Eurostat data for the euro area (black bars).

In Figures 5 to 8, we show the effects of the different setups of the model on the main macroeconomic aggregates and on the emergence of an endogenous crisis. As in Figures 1 to 4, we calibrate the ABM on data up to the fourth quarter of 2006 and make forecasts for twelve quarters ahead up to the fourth quarter of 2009. In the first setup, where we do not allow for financial frictions, an endogenous crisis does not occur, as is shown in Figure 6 (red line). In this setup, forecasts are very similar to forecasts of the benchmark DSGE model, as shown in Figure 2 (red line). As a comparison, we show forecasts of the ABM with the financial accelerator mechanism from Section 5.2 (Figures 2 to 4). Here, clearly, an endogenous crisis occurs (Figure 6 (blue line)). Hence the endogenous crisis, which is shown in Figure 6 (blue line), is mainly driven by credit rationing and the financial conditions of firms in the euro area. In this setup, export demand from the rest of the world is assumed to follow an autoregressive process of lag order one (AR(1)). Thus export demand follows a trend from exogenous projections, and an export demand shock does not occur (Figures 7 and 8 (blue line)). Therefore, the recession in the out-of-sample forecast is less pronounced, as in the observed data for the euro area. In the third setup, where we consider the model from Section 2 with the financial accelerator and an exogenous export demand shock, exports from the euro area correspond closely to the observed data, as is shown in Figures 7 and 8 (black line). In this setup,

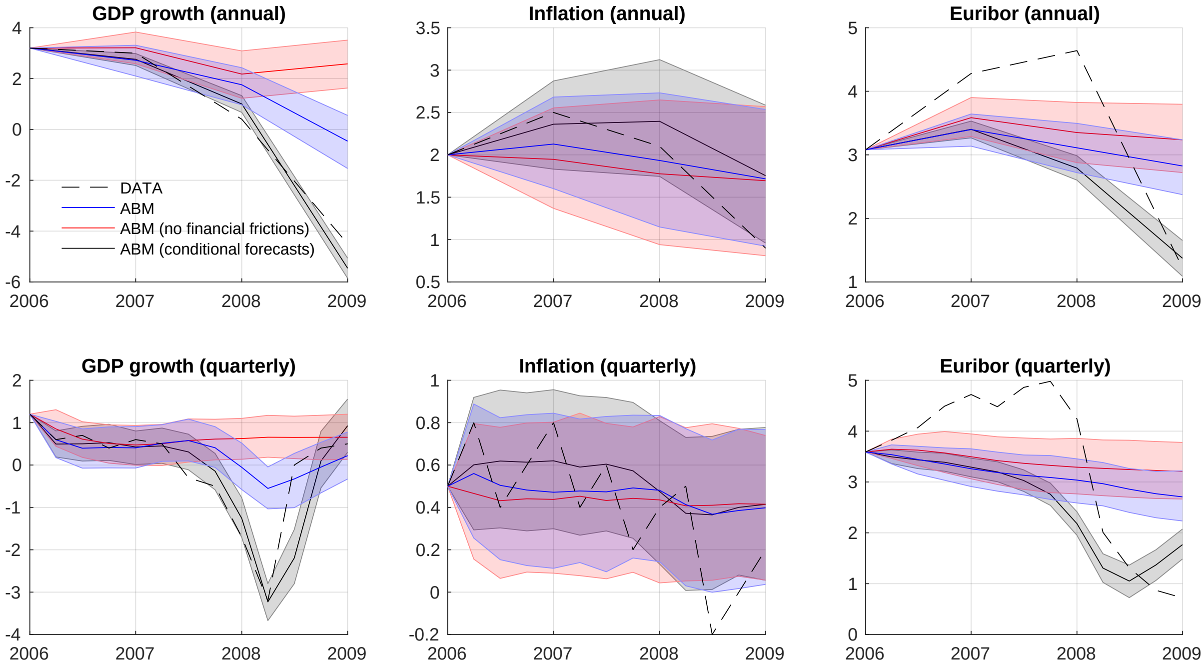

FIGURE 6. Forecasts estimated on data up to the fourth quarter of 2006 with and without financial frictions and exogenous export demand shocks. Conditional forecasts are conditional on real exports from the euro area (export demand shock). Top figures show GDP growth, inflation and the 3-month Euribor on an annualized basis, bottom figures depict quarterly growth, inflation and interest rates for ABM with financial frictions and exogenous export demand shocks (black line), ABM without financial frictions (red line), ABM with financial frictions (blue line), and observed Eurostat data for the euro area (dashed line). A 90 per cent confidence interval is plotted around the mean trajectory. Model results are obtained as an average over 500 Monte Carlo simulations. The DSGE model is estimated using Bayesian methods where a sample of 250,000 draws was created (neglecting the first 50,000 draws).

agents in the model have (forward-looking) expectations for a decline in exports of the euro area. With this setup, forecasts for GDP and the main components (household consumption, government consumption, investment, exports and imports) are almost spot-on for annual levels Figure 7 (black line) and very close for quarterly levels Figure 8 (black line).

5.4. Impact of the Financial crisis of 2007-2008 differentiated by industries and components of GDP

In this section, we break down ABM simulations by industries and components of GDP to shed some light on the sectoral impact of the Financial crisis of 2007-2008 in the ABM. Figure 9 shows the gross value added (GVA) from ABM simulation with the financial accelerator from Figure 4 disaggregated

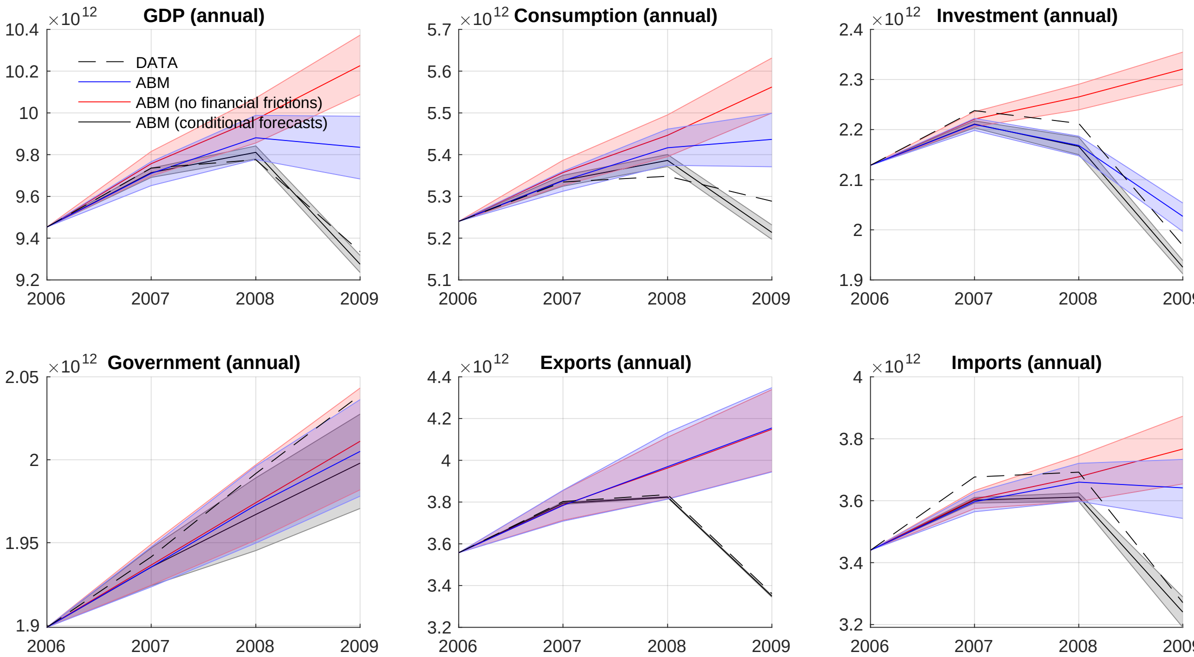

FIGURE 7. Forecasts estimated on data up to the fourth quarter of 2006 with and without financial frictions and export demand shocks. Conditional forecasts are conditional on real exports from the euro area (export demand shock). Figures show GDP (annually, in Euro and in real terms with base year 2010), household consumption (annually, in Euro and in real terms with base year 2010), fixed investment (annually, in Euro and in real terms with base year 2010), government consumption (annually, in Euro and in real terms with base year 2010), exports (annually, in Euro and in real terms with base year 2010), and imports (annually, in Euro and in real terms with base year 2010) for ABM with financial frictions and exogenous export demand shocks (black line), ABM without financial frictions (red line), ABM with financial frictions (blue line), and observed Eurostat data for the euro area (dashed line). A 90 per cent confidence interval is plotted around the mean trajectory. Model results are obtained as an average over 500 Monte Carlo simulations. The DSGE model is estimated using Bayesian methods where a sample of 250,000 draws was created (neglecting the first 50,000 draws).

for economic activities (industries).[19] GVA is disaggregated for ten economic activities (NACE*10) according to the statistical classification of economic activities in the European Community (NACE Rev. 2) (see Table B.7 in the appendix for details). ABM out-of-sample forecasts in Figure 9 show a good forecast performance for most sectors. Noticeable observations that provide insights into the ABM are: the recession is less pronounced as in the observed data because export demand follows a trend from exogenous projections, and an export demand shock does not occur, see also Figures 7 and 8 (blue line). In particular, this can be seen in the forecasts for the export-oriented industrial sectors (B, C, D and E). A noticeable exception where the recession is almost as deep as

19. Note the varying scales for the sectors of different sizes.

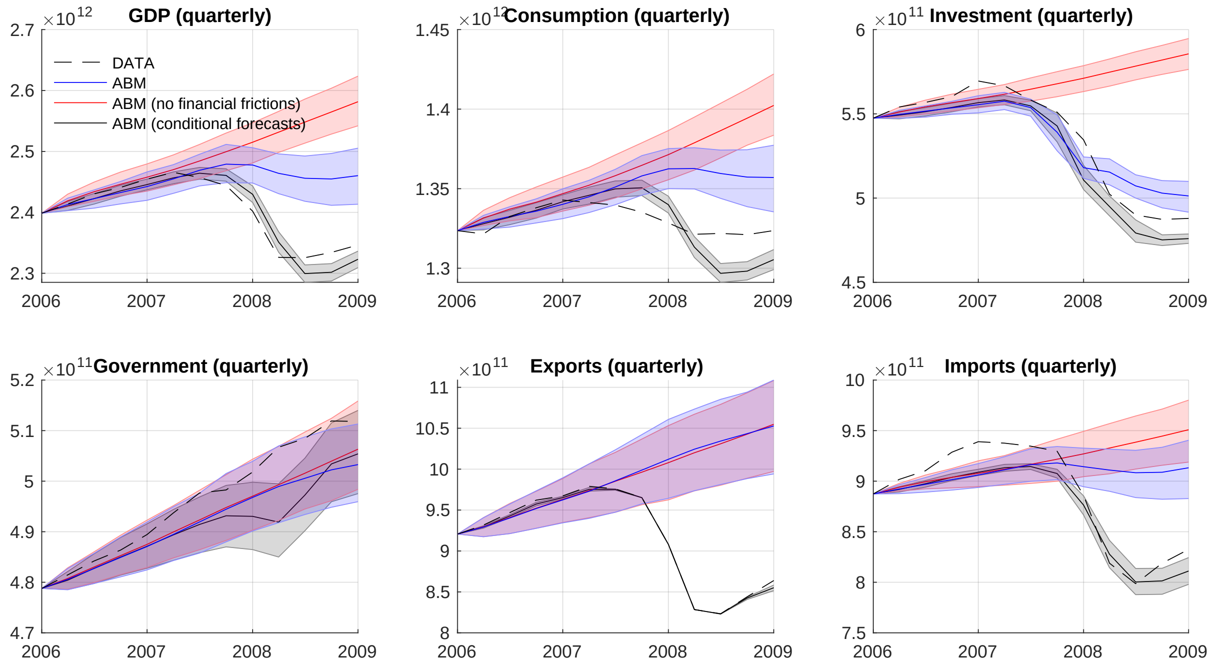

Figure 8. Forecasts estimated on data up to the fourth quarter of 2006 with and without financial frictions and export demand shocks. Conditional forecasts are conditional on real exports from the euro area (export demand shock). Figure show GDP (quarterly, in Euro and in real terms with base year 2010), household consumption (quarterly, in Euro and in real terms with base year 2010), fixed investment (quarterly, in Euro and in real terms with base year 2010), government consumption (quarterly, in Euro and in real terms with base year 2010), exports (quarterly, in Euro and in real terms with base year 2010), and imports (quarterly, in Euro and in real terms with base year 2010) for ABM with financial frictions and exogenous export demand shocks (black line), ABM without financial frictions (red line), ABM with financial frictions (blue line), and observed Eurostat data for the euro area (dashed line). A 90 per cent confidence interval is plotted around the mean trajectory. Model results are obtained as an average over 500 Monte Carlo simulations. The DSGE model is estimated using Bayesian methods where a sample of 250,000 draws was created (neglecting the first 50,000 draws).

in the observed data is the construction sector (F), confirming the good forecasting performance for investment (Figures 3 and 4). ABM forecasts for the financial sector (K) do not include all losses associated with the Financial crisis of 2007-2008. Sectors with heavy government involvement are less affected by the recession. Figure 10 shows the sectoral gross value added (GVA) from ABM simulation with the financial accelerator and the exogenous export demand shock from Figure 7, where GVA is again disaggregated for ten economic activities. In this setup, forecasts are conditional on the observed data for real exports from the euro area. With this setup, sectorial forecasts show a good forecast performance for almost all sectors. Noticeable exceptions are only the financial sector (K) and the agricultural sector (A), which is relatively small.

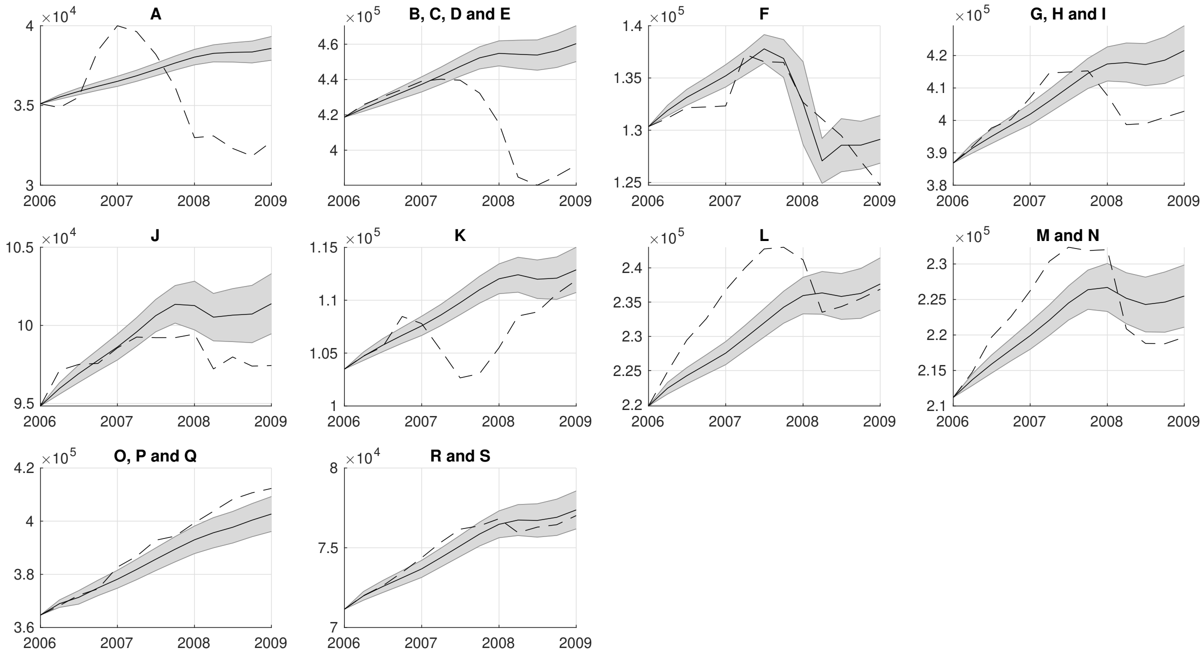

FIGURE 9. Sectoral out-of-sample forecasts estimated on data up to the fourth quarter of 2006. Figures show sectoral gross value added (GVA) (quarterly, in Euro) for ABM (solid line, a 90 per cent confidence interval is plotted around the mean trajectory) and observed Eurostat data for the euro area (dashed line). GVA is disaggregated for 10 economic activities (NACE*10) according to the statistical classification of economic activities in the European Community (NACE Rev. 2), see Table B.7 in the appendix for details. Model results are obtained as an average over 500 Monte Carlo simulations.

Finally, we further break down the ABM simulation with the financial accelerator and the exogenous export demand shock from Figure 7 for components of GDP. Here we use that the model includes a complete GDP identity, which is, together with a financial sector, an important criterion for future macroeconomic models (Lindé, 2018). In Figure 11, we show the components of GDP for the production, income and expenditure approach. For details on the calculation of GDP according to the different approaches, see Poledna et al. (2023). Coloured areas show the components from ABM simulations, while the dashed line refers to the corresponding values from observed data for the euro area. Clearly, ABM simulations demonstrate a good forecast performance for nearly all components. Most noticeably, a major structural change occurred after the height of the recession in the first quarter of 2009: while wages and social contributions are less affected by the recession, tax revenues are stagnating (income approach). Similarly, with the expenditure approach, household and government consumption is less affected by the recession, as capital formation.

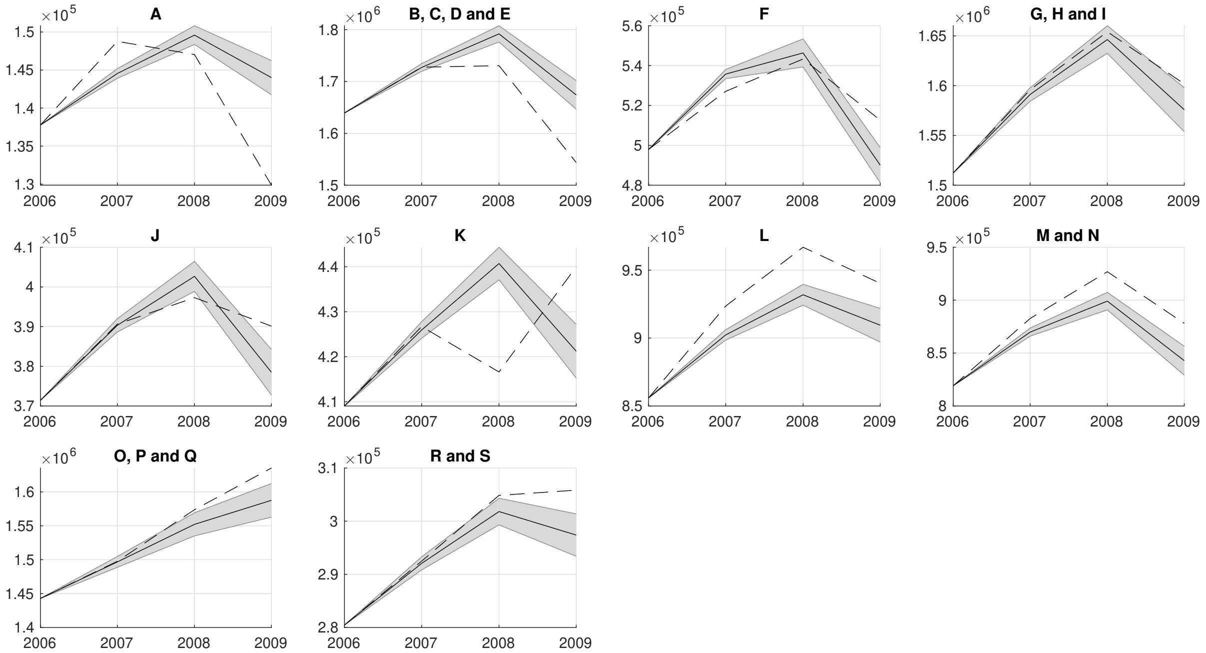

FIGURE 10. Sectoral conditional forecasts estimated on data up to the fourth quarter of 2006. Figures show sectoral gross value added (GVA) (annually, in Euro) for ABM (solid line, a 90 per cent confidence interval is plotted around the mean trajectory) and observed Eurostat data for the euro area (dashed line). GVA is disaggregated for 10 economic activities (NACE*10) according to the statistical classification of economic activities in the European Community (NACE Rev. 2), see Table B.7 in the appendix for details. Model results are obtained as an average over 500 Monte Carlo simulations.

6. The European debt crisis

After having demonstrated the ABM during the Financial crisis of 2007-2008 and the subsequent Great Recession, we use it to study the European debt crisis—the second major crisis in the recent history of the euro area.[20] This crisis has had significant adverse economic effects on the euro area and is interesting as standard models—apart from failing to predict the recession in the first place—also

20. The European sovereign debt crisis was a multi-year debt crisis in the euro area and the wider European Union commencing in 2010. Although only a few euro area countries (Greece, Portugal, Ireland, Spain and Cyprus) were unable to repay or refinance their government debt, it was perceived as a problem for the euro area as a whole. The causes for the debt crises were mostly a mix of weak actual and potential growth, competitive weakness, large pre-existing debt-to-GDP ratios from bank bailouts, and considerable liabilities of the private sector. To fight the crisis, many governments have adopted austerity policies contributing to social unrest and a debate among economists, with many advocating for greater deficits when economies are struggling. The crisis has had significant adverse economic effects and was blamed for low economic growth in the euro area and the European Union as a whole.

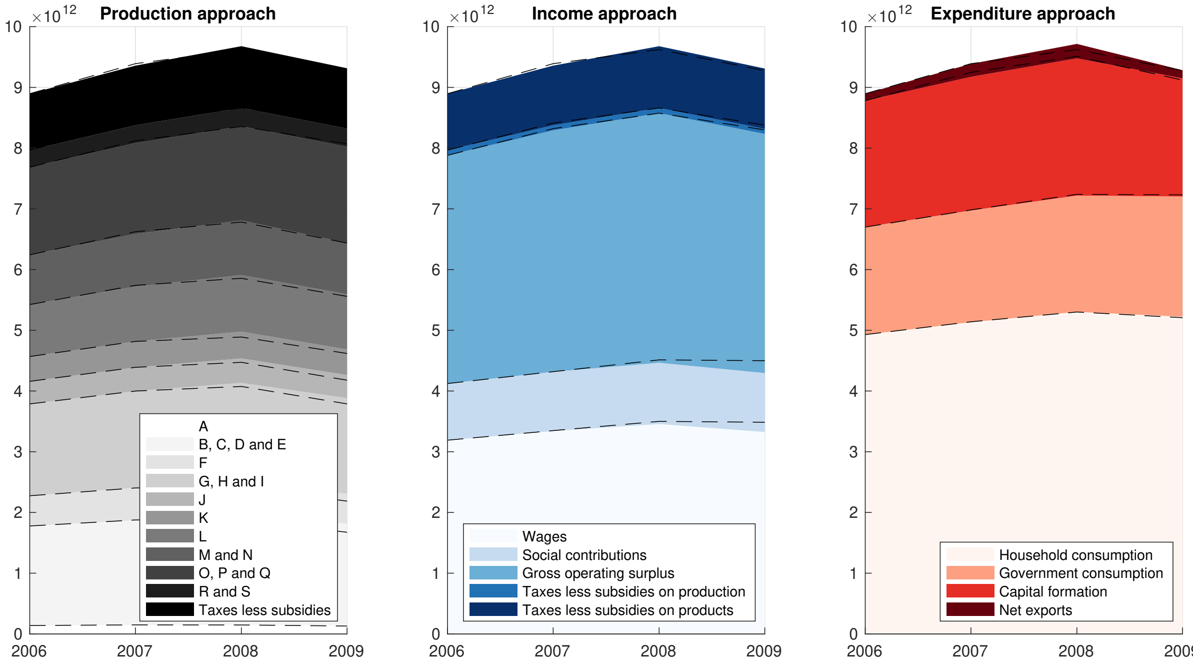

FIGURE 11. Composition of GDP according to production, income and expenditure approaches. Colored areas show conditional forecasts of the ABM estimated on data up to the fourth quarter of 2006, again as an average over 500 Monte Carlo simulations. The dashed line shows the corresponding values from observed data for the euro area.

have a clear tendency to forecast a quick recovery (Lindé et al., 2016). We follow the same approach as in Section 5 and start by comparing the forecast performance of the ABM against standard models in a series of traditional RMSE forecast exercises during the height of the European debt crisis. In the second part, we attempt to shed some light on the causes and the adverse economic effects of the European debt crisis with the ABM.

6.1. Comparison with an unconstriant VAR model

We start again by comparing the forecast performance of the ABM against an unconstrained VAR model during the height of the European debt crisis in a similar setup as in Sections 4.1 and 5.1. For this exercise, the unconstrained VAR model is initially estimated over the sample 1996:Q1 to 2010:Q1 (the sample 1995:Q2 to 1995:Q4 is used as a presample period). The VAR model is then re-estimated each quarter from the subsample from 2010:Q2 to 2012:Q4. The forecast period is 2010:Q2 to 2015:Q4. ABM forecasts are constructed analogously to the comparison with the VAR model in

Sections 4.1 and 5.1: the ABM is calibrated to 12 different reference quarters of the calibration period 2010:Q1 to 2012:Q4. Then, we let the model run for 12 quarters from each of these starting points (i.e., in the last simulation up until 2015:Q4), where we average the results of 500 Monte Carlo simulations before we evaluate the forecasting accuracy.

TABLE 7. Out-of-sample forecast performance during the European sovereign debt crisis

	GDP	Inflation	Euribor	Government consumption	Exports
VAR(1)	*RMSE-statistic for different forecast horizons*				
1q	0.54	0.14	0.05	0.37	1.35
2q	1.19	0.1	0.1	0.63	2.9
4q	2.84	0.15	0.23	1.24	7.02
8q	6.15	0.2	0.54	2.52	14.53
12q	7.77	0.21	0.67	3.55	18.32
ABM	*Percentage gains (+) or losses (-) relative to VAR(1) model*				
1q	28 (0.06)	13.3 (0.19)	13.9 (0.08)	41.4 (0.00)	-33.1 (0.77)
2q	38.7 (0.00)	8.3 (0.26)	13.1 (0.22)	9.2 (0.19)	13.7 (0.26)
4q	45.7 (0.00)	-3 (0.56)	32.4 (0.00)	-14.9 (0.94)	50.1 (0.00)
8q	56.4 (0.05)	7.4 (0.22)	66.9 (0.04)	-17.3 (0.86)	69.5 (0.00)
12q	59.8 (0.00)	8.4 (0.02)	71.4 (0.00)	-10.6 (0.98)	64.4 (0.00)
ABM (with financial frictions)	*Percentage gains (+) or losses (-) relative to VAR(1) model*				
1q	21.3 (0.19)	-29.9 (0.85)	-13.6 (0.87)	51.2 (0.00)	-26.1 (0.72)
2q	40.5 (0.00)	-42.1 (0.83)	-10.4 (0.65)	18.5 (0.03)	16 (0.22)
4q	56.8 (0.00)	-29.9 (0.99)	26.3 (0.11)	-4.4 (0.72)	50.3 (0.00)
8q	72.3 (0.04)	-11.7 (0.86)	77.1 (0.03)	-11.5 (0.82)	69.6 (0.00)
12q	74.6 (0.00)	-6.6 (0.86)	86.7 (0.00)	-4.5 (0.84)	63.9 (0.00)

Note: The VAR(1) model is estimated starting from 1996:Q1 to 2010:Q1 (1995:Q2 to 1995:Q4 is used as a presample period) and is then re-estimated each quarter for the subsample from 2010:Q2 to 2012:Q4. The forecast period is 2010:Q2 to 2015:Q4. ABM results are obtained as an average of 500 Monte Carlo simulations. The values in brackets indicate the p-values of Diebold and Mariano (1995) tests, where we test whether the ABM forecasts are significantly different in accuracy than the VAR(1) forecasts (the null hypothesis of the test is that there is no difference between two competing forecasts).

Table 7 reports the out-of-sample RMSEs for different forecast horizons of 1, 2, 4, 8, and 12 quarters over the period 2010:Q2 to 2015:Q4, covering the European debt crisis as well as the subsequent recovery in the euro area. In general, the out-of-sample forecast statistic shows a good forecast performance of the ABM relative to the VAR(1) model for this subsample. As for the entire sample (Table 1), for GDP, the 3-month Euribor, and exports of the euro area, the ABM does better than the VAR(1) model by a substantial margin for almost all horizons. The good forecasting performance of the ABM during the European debt crisis is again not too surprising as the ABM accounts for some of the main driving forces during that debt crisis (see also Section 6.3).

6.2. Comparison with a benchmark DSGE model

Next, we compare the out-of-sample forecast performance of the ABM to that of the benchmark DSGE model during the height of the European debt crisis in a similar setup as in Sections 4.2 and 5.2. As in Sections 4.2 and 5.2, we estimate AR models as a benchmark for the DSGE model and the ABM. For this exercise, the benchmark DSGE model and the AR models are initially estimated over the sample 1996:Q1 to 2010:Q1 (the sample 1995:Q2 to 1995:Q4 is used as a presample period). All models are then re-estimated each quarter from the subsample from 2010:Q2 to 2012:Q4. The forecast period is, as above, from 2010:Q2 to 2015:Q4 to cover the height of the European debt crisis with persistently low growth. ABM forecasts are constructed analogously to the comparison with the VAR model in Section 6.1.

TABLE 8. Out-of-sample forecast performance during the European sovereign debt crisis in comparison to DSGE model

	GDP	Inflation	Euribor	Household consumption	Investment
AR(1)	*RMSE-statistic for different forecast horizons*				
1q	0.39	0.13	0.06	0.4	1.22
2q	0.73	0.1	0.12	0.81	1.84
4q	1.56	0.16	0.2	1.82	3.42
8q	3.08	0.19	0.29	3.68	5.88
12q	3.83	0.2	0.36	4.75	7.1
DSGE	*Percentage gains (+) or losses (-) relative to AR(1) model*				
1q	-43.2 (1.00)	-4.5 (0.57)	-1 (0.55)	-50.4 (1.00)	-21.2 (0.93)
2q	-63.1 (0.98)	-0.7 (0.51)	-24.9 (0.99)	-59.2 (1.00)	-48.9 (0.99)
4q	-64.9 (0.96)	28.7 (0.05)	-58.9 (0.99)	-51.1 (1.00)	-59.7 (0.97)
8q	-50 (1.00)	-10.3 (0.83)	-105.1 (1.00)	-33.8 (1.00)	-69 (1.00)
12q	-37.3 (1.00)	-18 (0.98)	-105 (1.00)	-22.3 (1.00)	-61.3 (1.00)
ABM	*Percentage gains (+) or losses (-) relative to AR(1) model*				
1q	1 (0.36)	0.8 (0.21)	29.6 (0.08)	15.8 (0.17)	12 (0.14)
2q	0.6 (0.45)	6.8 (0.09)	27.4 (0.18)	26.3 (0.03)	5.8 (0.32)
4q	1.1 (0.46)	1.1 (0.41)	20.9 (0.27)	22.1 (0.00)	2.5 (0.39)
8q	13.1 (0.10)	-0.5 (0.68)	37.8 (0.00)	19.5 (0.08)	3.3 (0.31)
12q	18.5 (0.01)	2.9 (0.01)	46.7 (0.00)	22.6 (0.00)	4.6 (0.18)
ABM (with financial frictions)	*Percentage gains (+) or losses (-) relative to AR(1) model*				
1q	-8.2 (0.66)	-48.7 (0.97)	7.2 (0.37)	37 (0.04)	10.6 (0.13)
2q	3.5 (0.44)	-44.4 (0.85)	7.7 (0.42)	47.8 (0.01)	12.3 (0.10)
4q	21.3 (0.19)	-24.7 (0.82)	13.8 (0.40)	50.8 (0.00)	20.6 (0.03)
8q	44.9 (0.00)	-21.2 (0.82)	57.1 (0.00)	45.8 (0.03)	49.4 (0.00)
12q	48.6 (0.01)	-13 (0.92)	75.1 (0.00)	47.6 (0.00)	28.2 (0.09)

Note: The AR(1) and the DSGE model are estimated starting from 1996:Q1 to 2010:Q1 (1995:Q2 to 1995:Q4 is used as a presample period) and are then re-estimated each quarter for the subsample from 2010:Q2 to 2012:Q4. The forecast period is 2010:Q2 to 2015:Q4. ABM results are obtained as an average of 500 Monte Carlo simulations. The DSGE model is estimated using Bayesian methods where a sample of 250,000 draws was created (neglecting the first 50,000 draws). The values in brackets indicate the p-values of Diebold and Mariano (1995) tests, where we test whether the ABM forecasts are significantly different in accuracy than the AR(1) forecasts (the null hypothesis of the test is that there is no difference between two competing forecasts).

Again, the good forecast performance of the ABM relative to the unconstrained VAR model during the European debt crisis is confirmed by comparison with the benchmark DSGE model reported in Table 8. For the subsample of the European debt crisis, the out-of-sample forecast statistic shows a good forecast performance of the ABM relative to the AR models and the benchmark DSGE model. Here the ABM does better than the different AR models by a considerable margin for all horizons and all variables with the exception of inflation. The benchmark DSGE model, on the other hand, does perform worse for almost all horizons and all variables. This is not too surprising as, apart from failing to predict the recession in the first place, the benchmark DSGE also has a clear tendency to forecast a quick recovery (Lindé et al., 2016).

Next, we compare the out-of-sample forecast performance of the ABM to that of the benchmark DSGE model estimated on data up to the fourth quarter of 2010. Here, we compare forecasts for 12 quarters ahead up to the fourth quarter of 2013. Forecasts covering these quarters are of particular interest as they include the height of the European debt crisis and contain the entire period with persistently low growth. For eight consecutive quarters—from the second quarter of 2011 to the first quarter of 2013—growth in the euro area was zero or negative (Figure 12). In Figures 12 to 14, we show the out-of-sample forecasts of the ABM, the benchmark DSGE model, and the AR(1) model estimated on data up to the fourth quarter of 2010. Figure 12 shows GDP growth, the inflation rate, and the 3-month Euribor—annually (top) and quarterly (bottom). Clearly, the linear models—the linearized DSGE model (red line) and the AR(1) model (blue line)— did not only fail to predict the recession (Figure 2) but also forecast a much faster recovery, as observed in the euro area. This is particularly evident in quarterly GDP growth, where the euro area growth was zero or negative for eight consecutive quarters from the second quarter of 2011 to the first quarter of 2013. However, both the linearized DSGE model (red line) and the AR(1) model (blue line) forecast growth rates of around half a per cent for these eight quarters. On the other hand, the ABM, which takes the financial conditions of firms into account, accurately captures persistently lower growth rates throughout the European debt crisis. Similarly, for the 3-month Euribor, the linear models merely extrapolate the trend for the interest rate, while the ABM more accurately captures lower interest rates in the euro area. Forecasts of the inflation rate are similar for the three model classes.

Figures 13 and 14 compare forecasts for the levels of GDP and the main components (household consumption, government consumption, investment, exports and imports) of the ABM and the benchmark DSGE model. As above, in Figure 13, we show annual levels, and Figure 14 shows

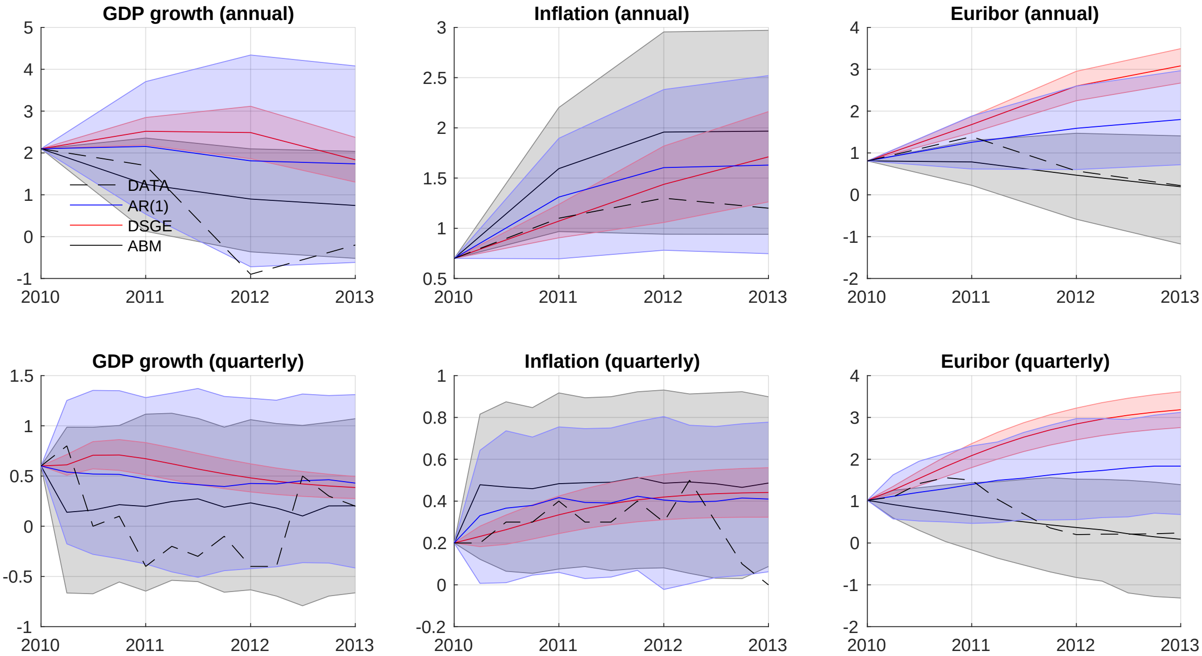

FIGURE 12. Out-of-sample forecasts estimated on data up to the fourth quarter of 2010. Top figures show GDP growth, inflation and the 3-month Euribor on an annualized basis, bottom figures depict quarterly growth, inflation and interest rates for ABM (black line), DSGE (red line), AR(1) (blue line), and observed Eurostat data for the euro area (dashed line). A 90 per cent confidence interval is plotted around the mean trajectory. Model results are obtained as an average over 500 Monte Carlo simulations. The DSGE model is estimated using Bayesian methods where a sample of 250,000 draws was created (neglecting the first 50,000 draws).

the respective quarterly levels of GDP and the main components. Evidently, forecasts of the linear models—the linearized DSGE model (red line) and the AR(1) model (blue line)—are very similar for GDP, household consumption and investment. While the linear models forecast a swift recovery visible in all variables, the ABM accurately captures the long-term trends of all variables with the only exception of government consumption. The reason for this exception is that government spending is an external variable in the ABM and follows a trend from exogenous projections. Thus, government austerity does not occur, as can be seen in Figures 13 and 14. For a setup of the ABM that includes government austerity as an exogenous fiscal shock, see Section 6.3.

6.3. Main Driving Forces of the European debt crisis

After having shown that the ABM can be used to shed some light on the main driving forces during the Great Recession, we use it to better understand the European debt crisis. Natural

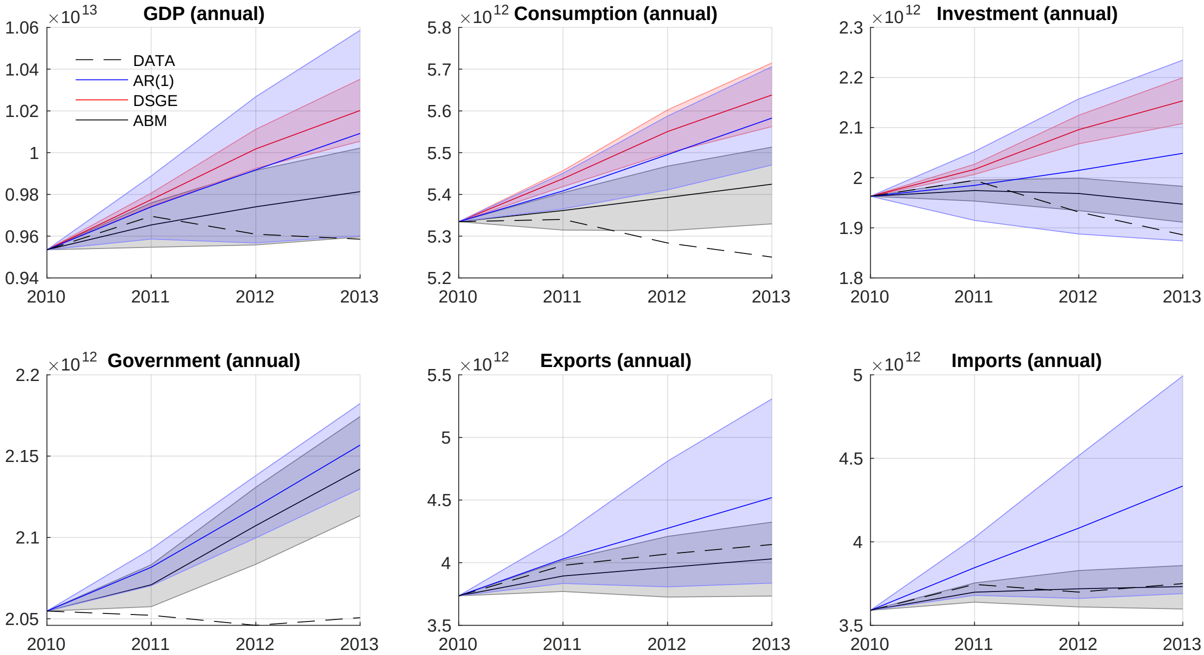

FIGURE 13. Out-of-sample forecasts estimated on data up to the fourth quarter of 2010. Figures show GDP (annually, in Euro and in real terms with base year 2010), household consumption (annually, in Euro and in real terms with base year 2010), fixed investment (annually, in Euro and in real terms with base year 2010), government consumption (annually, in Euro and in real terms with base year 2010), exports (annually, in Euro and in real terms with base year 2010), and imports (annually, in Euro and in real terms with base year 2010) for ABM (black line), DSGE (red line), AR(1) (blue line), and observed Eurostat data for the euro area (dashed line). A 90 per cent confidence interval is plotted around the mean trajectory. Model results are obtained as an average over 500 Monte Carlo simulations. The DSGE model is estimated using Bayesian methods where a sample of 250,000 draws was created (neglecting the first 50,000 draws).

candidates for the main driving forces are government austerity and credit tightening resulting from the deteriorating financial conditions of firms in the euro area. Thus, in this exercise, we analyze the European debt crisis with three different setups of the ABM, focusing on government austerity and credit tightening, and compare the forecast performance of these different setups of the ABM. First, we consider the ABM from Poledna et al. (2023) without a financial accelerator mechanism and assume no government austerity. Thus, in this setup, credit rationing does not occur, and sufficient funds are available to firms in the model. In the second setup, we use the model from Section 2 with a financial accelerator and debt-financed investment with collateralized loans where firms' ability to borrow depends on the market value of their assets and their net worth. In this setup, we allow for credit tightening but assume no government austerity in the euro area (government consumption is assumed to follow an autoregressive process of lag order one (AR(1))). In the third setup, we

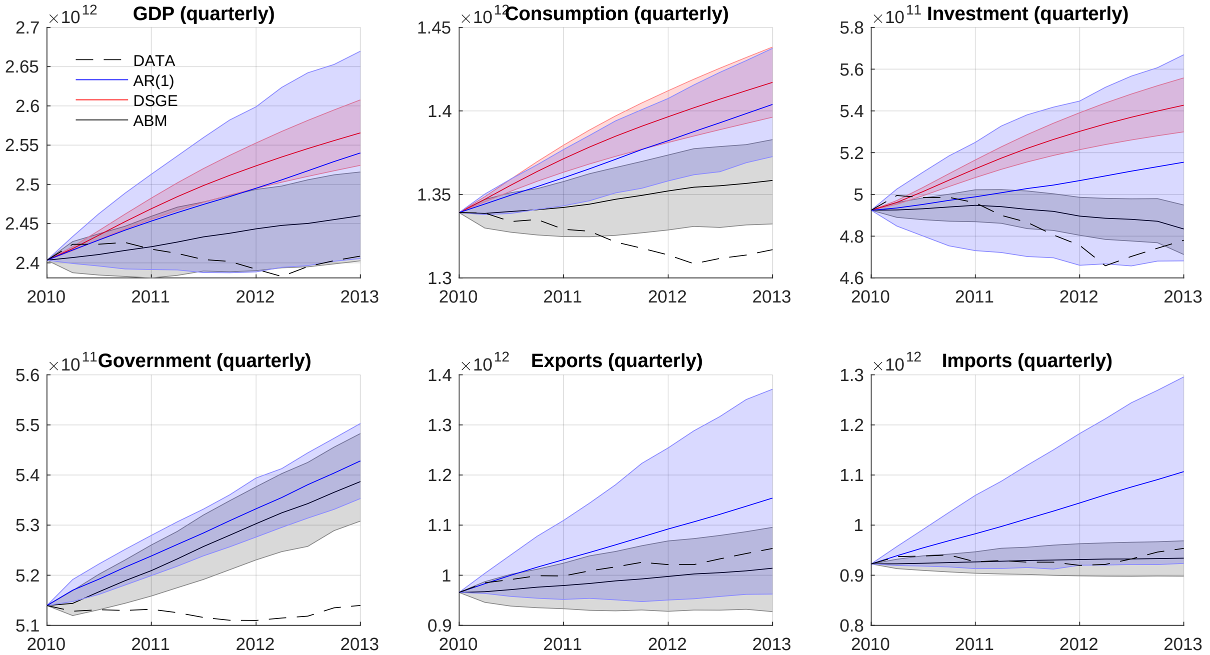

FIGURE 14. Out-of-sample forecasts estimated on data up to the fourth quarter of 2010. Figure show GDP (quarterly, in Euro and in real terms with base year 2010), household consumption (quarterly, in Euro and in real terms with base year 2010), fixed investment (quarterly, in Euro and in real terms with base year 2010), government consumption (quarterly, in Euro and in real terms with base year 2010), exports (quarterly, in Euro and in real terms with base year 2010), and imports (quarterly, in Euro and in real terms with base year 2010) for ABM (black line), DSGE (red line), AR(1) forecasts (blue line), and observed Eurostat data for the euro area (dashed line). A 90 per cent confidence interval is plotted around the mean trajectory. Model results are obtained as an average over 500 Monte Carlo simulations. The DSGE model is estimated using Bayesian methods where a sample of 250,000 draws was created (neglecting the first 50,000 draws).

additionally assume government austerity in the euro area. In this setup, forecasts are conditional on the observed data for real government consumption of the euro area, and agents have (forward-looking) expectations for government austerity.

Figures 15 to 17 show the effects of the three different setups of the model on the main macroeconomic aggregates. In Figure 15, we show GDP growth, inflation rates and the 3-month Euribor—annually (top) and quarterly (bottom) for the different setups of the model. In the first setup (red lines), as a comparison, we assume no government austerity in the euro area and sufficient funds are available to firms in the model. Clearly, with this setup, the European debt crisis does not occur in the model, as is shown in Figure 15 (red lines). Surprisingly, with this setup, forecasts are very similar to the forecasts of the benchmark DSGE model during the European debt crisis shown in Figure 12 (red lines). In the second setup (blue lines), we allow for credit tightening but

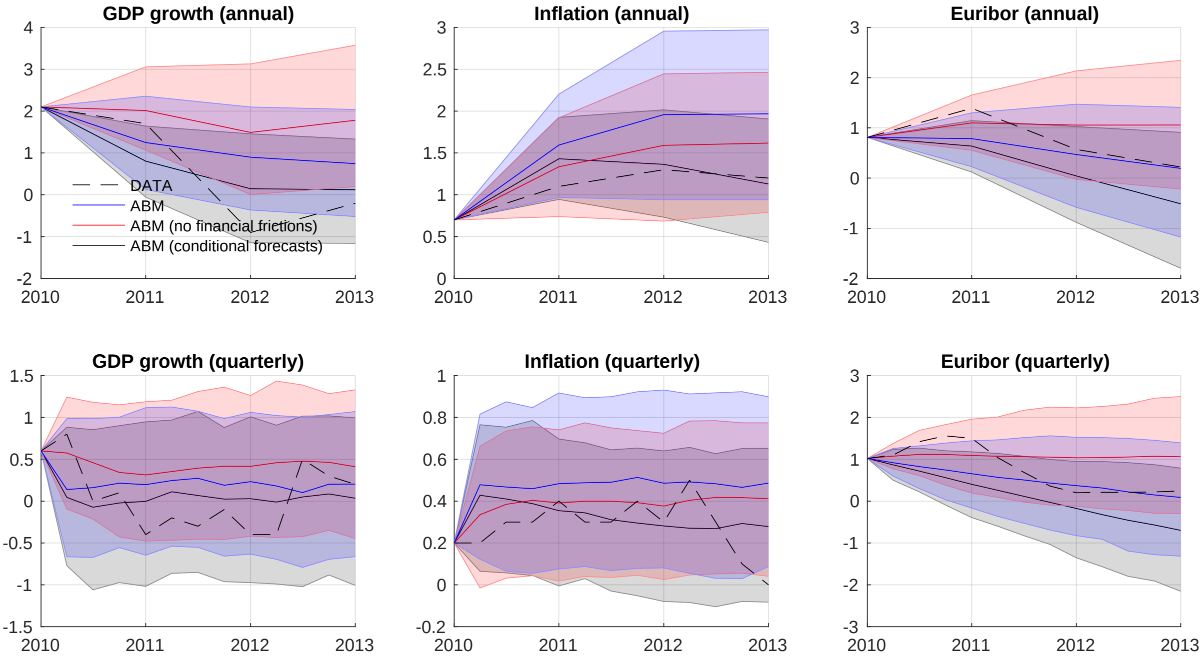

Figure 15. Forecasts estimated on data up to the fourth quarter of 2010 with and without financial frictions and government austerity. Conditional forecasts are conditional on real government consumption (government austerity). Top figures show GDP growth, inflation and the 3-month Euribor on an annualized basis, bottom figures depict quarterly growth, inflation and interest rates for ABM (black line), DSGE (red line), AR(1) (blue line), and observed Eurostat data for the euro area (dashed line). A 90 per cent confidence interval is plotted around the mean trajectory. Model results are obtained as an average over 500 Monte Carlo simulations. The DSGE model is estimated using Bayesian methods where a sample of 250,000 draws was created (neglecting the first 50,000 draws).

assume no government austerity in the euro area. Thus, government spending follows a trend from exogenous projections, and government austerity does not occur (Figures 16 and 17 (red and blue lines)). Clearly, with this setup, persistently low growth during the height of the European debt crisis can be observed in the ABM, as shown in (Figure 15 (blue lines)). In the third setup, we additionally include government austerity in the euro area (black lines). In this setup, forecasts are conditional on the observed data for real government consumption of the euro area, and agents have (forward-looking) expectations for government austerity (Figures 16 and 17 (black lines)). With this setup, forecasts for GDP and the main components (household consumption, government consumption, investment, exports and imports) are accurate for annual levels Figure 16 (black line) and very close for quarterly levels Figure 17 (black line). In particular, annual (Figures 16 and 17 (black line)) and quarterly (Figures 16 and 17 (black line)) GDP is more accurate in comparison to the setup without government austerity (blue lines). However, the overall differences in comparison to the

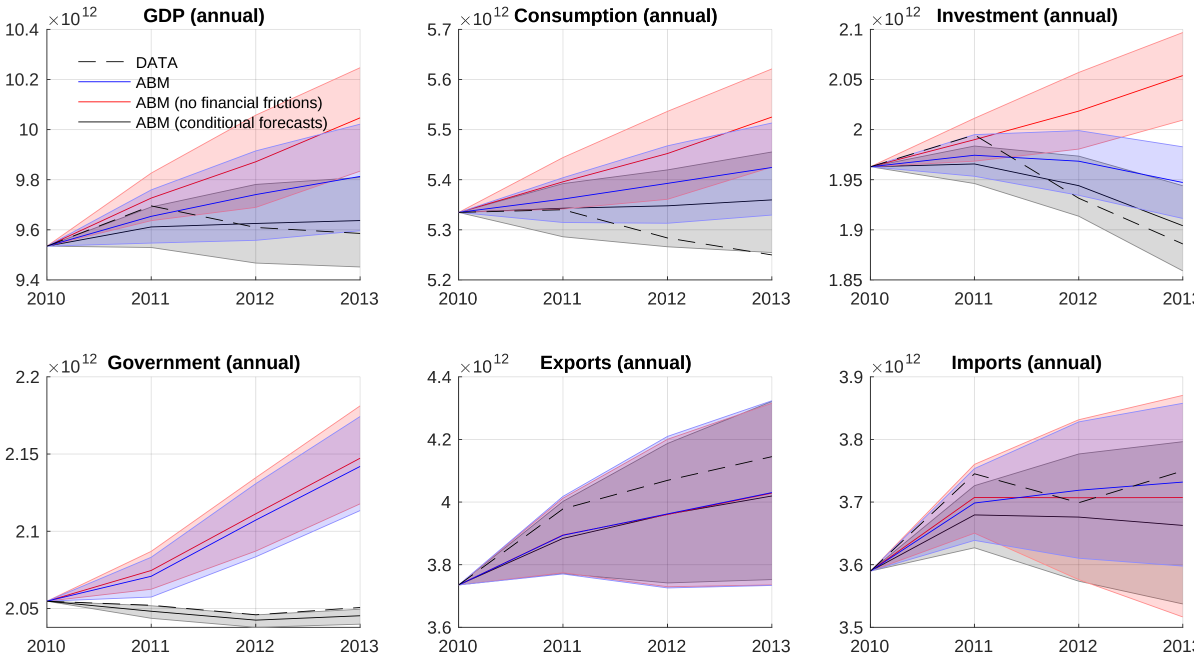

FIGURE 16. Forecasts estimated on data up to the fourth quarter of 2010 with and without financial frictions and government austerity. Conditional forecasts are conditional on real government consumption (government austerity). Figures show GDP (annually, in Euro and in real terms with base year 2010), household consumption (annually, in Euro and in real terms with base year 2010), fixed investment (annually, in Euro and in real terms with base year 2010), government consumption (annually, in Euro and in real terms with base year 2010), exports (annually, in Euro and in real terms with base year 2010), and imports (annually, in Euro and in real terms with base year 2010) for ABM (black line), DSGE (red line), AR(1) (blue line), and observed Eurostat data for the euro area (dashed line). A 90 per cent confidence interval is plotted around the mean trajectory. Model results are obtained as an average over 500 Monte Carlo simulations. The DSGE model is estimated using Bayesian methods where a sample of 250,000 draws was created (neglecting the first 50,000 draws).

setup without government austerity are relatively small. Thus, the European debt crisis seems to be driven, to a considerable extent, by the financial conditions of firms in the euro area rather than by government austerity.

7. The COVID-19 recession

In this section, we use the model to analyse the COVID-19 recession in the euro area. The COVID-19 pandemic caused the second-largest global recession in history and had far-reaching economic consequences, from supply-side manufacturing issues to decreased consumer activity due to lockdown measures, as well as higher unemployment and rising debt levels. In the euro area, the pandemic

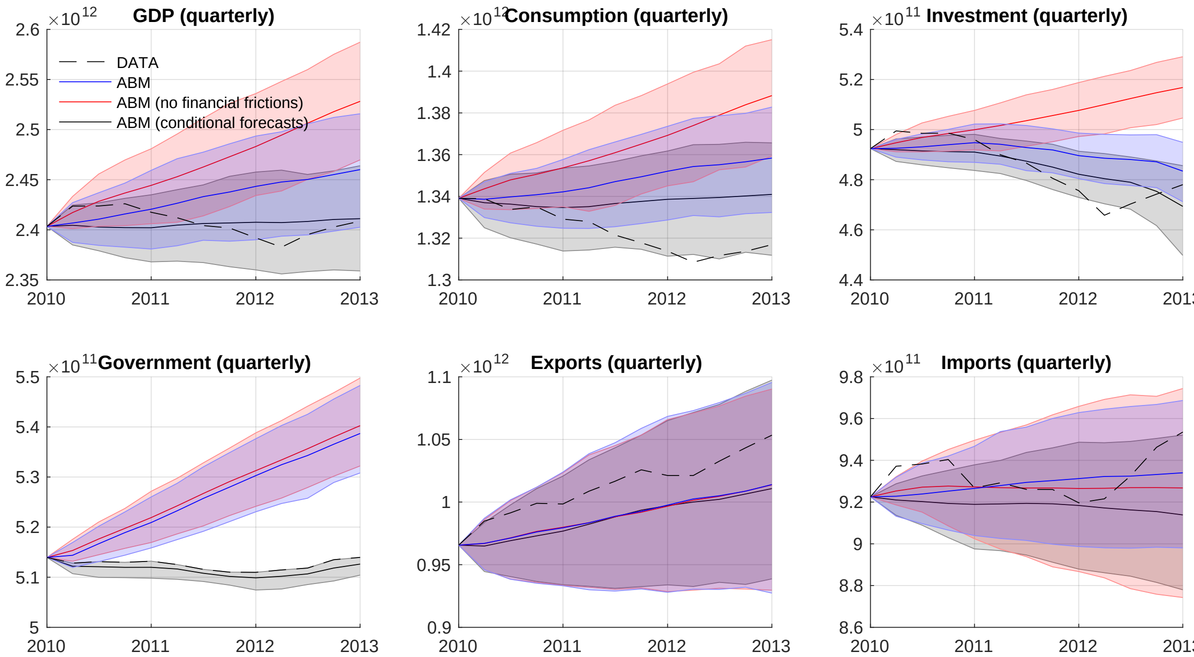

FIGURE 17. Forecasts estimated on data up to the fourth quarter of 2010 with and without financial frictions and government austerity. Conditional forecasts are conditional on real government consumption (government austerity). Figure show GDP (quarterly, in Euro and in real terms with base year 2010), household consumption (quarterly, in Euro and in real terms with base year 2010), fixed investment (quarterly, in Euro and in real terms with base year 2010), government consumption (quarterly, in Euro and in real terms with base year 2010), exports (quarterly, in Euro and in real terms with base year 2010), and imports (quarterly, in Euro and in real terms with base year 2010) for ABM (black line), DSGE (red line), AR(1) forecasts (blue line), and observed Eurostat data for the euro area (dashed line). A 90 per cent confidence interval is plotted around the mean trajectory. Model results are obtained as an average over 500 Monte Carlo simulations. The DSGE model is estimated using Bayesian methods where a sample of 250,000 draws was created (neglecting the first 50,000 draws).

began in February 2020, when most countries introduced lockdown measures that decreased economic activity in the service sector. Additionally, the euro area faced supply-side manufacturing issues from the global supply chain crisis. GDP in the second quarter of 2020 fell by more than 11 per cent compared to the previous quarter, which is the biggest drop in a single quarter on record. To limit the impact of the recession, euro area countries increased government spending and introduced a number of policy instruments, such as cash transfers, subsidies, and loan guarantees. Notably, Germany and most other euro area countries introduced government-subsidized short-time working schemes known as Kurzarbeit, which entailed a reduction in working time with the government making up for all or part of the lost wage. This policy represented one of the main instruments against the COVID-19 recession and enabled companies to avoid layoffs or bankruptcies.

7.1. Sceanrio analysis of the COVID-19 pandemic in the euro area

We start by performing a scenario analysis of the COVID-19 pandemic in the euro area. In the context of the ABM, we model the COVID-19 pandemic in the euro area with three shocks. These shocks include: (1) an initial shock caused by the restrictions of economic activities due to the lockdown measures in the first two quarters of 2020; (2) an export demand and import supply shock from the global supply chain crisis; (3) a fiscal shock from increased government spending. Additionally, we implement a short-time work policy instrument where companies can keep their employees at a salary of up to 90 per cent of the net remuneration received before short-time work, which is refunded by the government to the companies in full. To model loan guarantees by the government, we further do not allow credit rationing in the financial sector. Thus, sufficient funds are available to firms in the model.

For the analysis, we consider three different scenarios with the ABM. First, a baseline scenario where we assume business as usual without the COVID-19-related shocks. Second, a scenario with only the initial shock caused by the restrictions of economic activities due to the lockdown measures in the first two quarters of 2020. Third, a scenario with the initial shock from the lockdown measures and export demand and import supply shocks from the global supply chain crisis, as well as fiscal shocks throughout the simulation. The setup of the model for this scenario is similar to the conditional forecasts in Section 4.3. This scenario is conditional on the observed data for exports, imports, and government consumption of the euro area, as well as the respective deflators, which are exogenous in this setup of the ABM. For the implementation of the COVID-19-related shocks and the three scenarios in the model presented in Section 2, see Appendix A.1.

In Figures 18 to 20, we show the main macroeconomic aggregates of the different scenarios considered with the model. For this exercise, the model is calibrated with data up to the fourth quarter of 2019 before the COVID-19 pandemic hit the euro area. Macroeconomic aggregates are projected for eight quarters, which includes the peak of the COVID-19 recession and the recovery. In Figure 18, we show GDP growth, the inflation rate, and the 3-month Euribor—annually (top) and quarterly (bottom). In the baseline scenario (blue line), without the COVID-19-related shocks, we see, as expected, continuous low growth and inflation. However, with the initial shock caused by the restrictions of economic activities due to the lockdown measures in the first two quarters of 2020 (initial shock scenario shown in red), we see a sharp decline in the GDP growth rate in the first and second quarter of 2020 followed by a fast recovery in the following quarters. While the initial fall

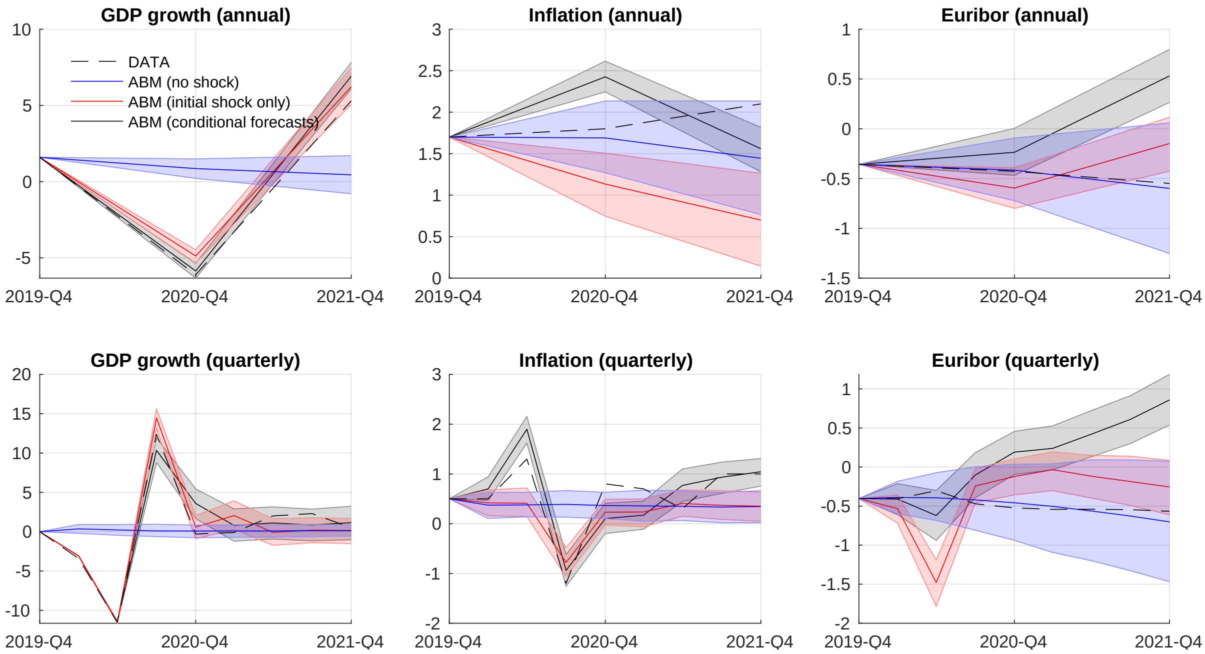

FIGURE 18. Forecasts estimated on data up to the fourth quarter of 2019 with and without financial frictions and exogenous export demand shocks. Conditional forecasts are conditional on real exports from the euro area (export demand shock). Top figures show GDP growth, inflation and the 3-month Euribor on an annualized basis, bottom figures depict quarterly growth, inflation and interest rates for ABM with financial frictions and exogenous export demand shocks (black line), ABM without financial frictions (red line), ABM with financial frictions (blue line), and observed Eurostat data for the euro area (dashed line). A 90 per cent confidence interval is plotted around the mean trajectory. Model results are obtained as an average over 500 Monte Carlo simulations. The DSGE model is estimated using Bayesian methods where a sample of 250,000 draws was created (neglecting the first 50,000 draws).

and recovery of the GDP growth rate are almost accurately captured by the model, with the initial shock only (red line), the initial and subsequent rise of inflation, as in the observed data (dashed line), is not captured. Only with conditional forecasts (black line), including the initial shock and, additionally, export demand and import supply shocks from the global supply chain crisis, as well as fiscal shocks, we observe the sharp initial rise and fall, as well as the subsequent continuous rising of the inflation rate in line with the observed data (dashed line). In this scenario, the sharp initial rise in the inflation rate is caused by the fiscal shock, and the continuous rising of inflation after the lifting of the lockdown measures is due to the global supply chain crisis. Both with only the initial shock (red line) and with the global supply chain crisis (black line), the Taylor rule of the ABM suggests that an inflation-targeting central bank would raise interest rates, which does not match the observed data of the 3-month Euribor.

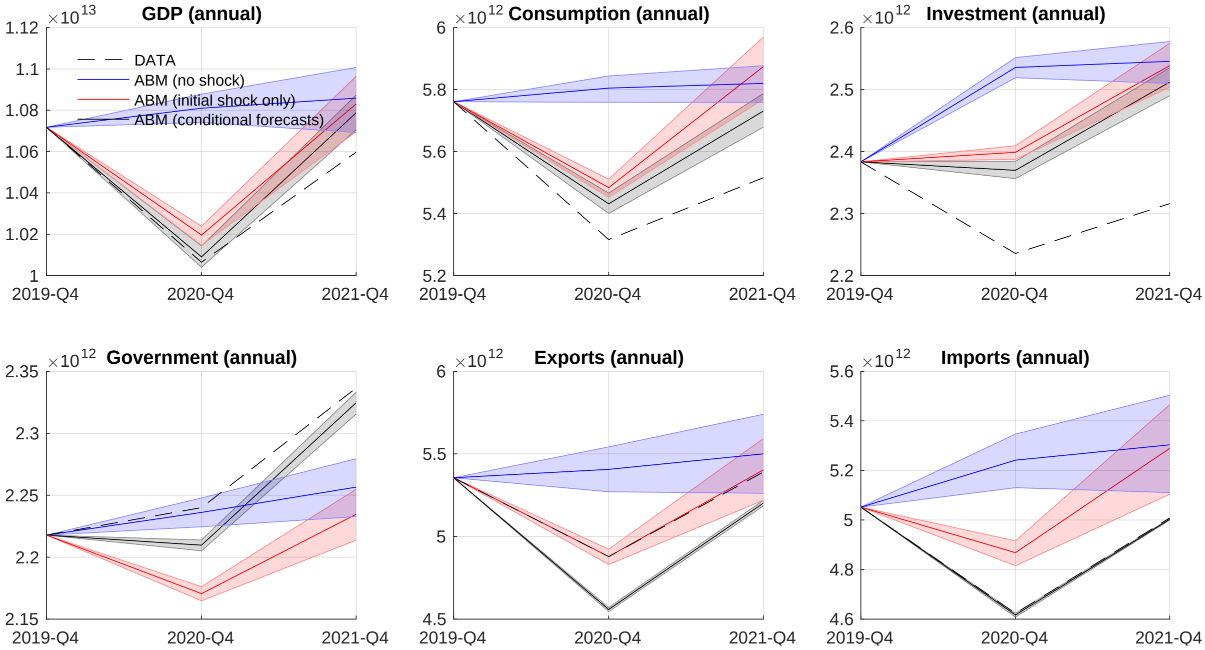

FIGURE 19. Forecasts estimated on data up to the fourth quarter of 2019 with and without financial frictions and export demand shocks. Conditional forecasts are conditional on real exports from the euro area (export demand shock). Figures show GDP (annually, in Euro and in real terms with base year 2010), household consumption (annually, in Euro and in real terms with base year 2010), fixed investment (annually, in Euro and in real terms with base year 2010), government consumption (annually, in Euro and in real terms with base year 2010), exports (annually, in Euro and in real terms with base year 2010), and imports (annually, in Euro and in real terms with base year 2010) for ABM with financial frictions and exogenous export demand shocks (black line), ABM without financial frictions (red line), ABM with financial frictions (blue line), and observed Eurostat data for the euro area (dashed line). A 90 per cent confidence interval is plotted around the mean trajectory. Model results are obtained as an average over 500 Monte Carlo simulations. The DSGE model is estimated using Bayesian methods where a sample of 250,000 draws was created (neglecting the first 50,000 draws).

In Figures 19 and 20, we compare the scenarios for the levels of GDP and the main components (household consumption, government consumption, investment, exports and imports). While Figure 19 shows annual levels, in Figure 20, we show the respective quarterly levels for GDP and the main components. Again, in the baseline scenario, Figures 19 and 20 show projections of the trend for GDP and the main components. With the initial shock caused by the restrictions of economic activities due to the lockdown measures in the first two quarters of 2020 (initial shock scenario shown in red), we see a sharp decline in GDP and the main components in the first and second quarters of 2020. This sharp decline in the initial shock scenario (red line) is followed by a fast recovery in the following quarters, where the levels of GDP and the main components recover to the respective levels of the baseline scenario (blue line). With conditional forecasts (black line), the

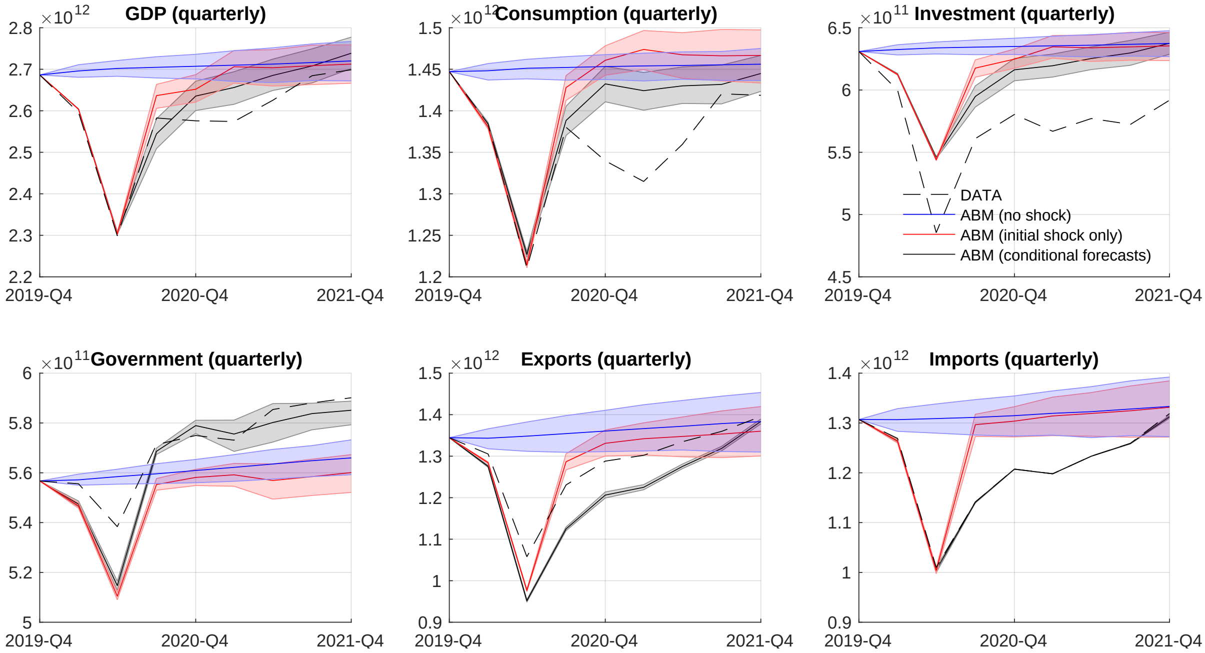

FIGURE 20. Forecasts estimated on data up to the fourth quarter of 2019 with and without financial frictions and export demand shocks. Conditional forecasts are conditional on real exports from the euro area (export demand shock). Figure show GDP (quarterly, in Euro and in real terms with base year 2010), household consumption (quarterly, in Euro and in real terms with base year 2010), fixed investment (quarterly, in Euro and in real terms with base year 2010), government consumption (quarterly, in Euro and in real terms with base year 2010), exports (quarterly, in Euro and in real terms with base year 2010), and imports (quarterly, in Euro and in real terms with base year 2010) for ABM with financial frictions and exogenous export demand shocks (black line), ABM without financial frictions (red line), ABM with financial frictions (blue line), and observed Eurostat data for the euro area (dashed line). A 90 per cent confidence interval is plotted around the mean trajectory. Model results are obtained as an average over 500 Monte Carlo simulations. The DSGE model is estimated using Bayesian methods where a sample of 250,000 draws was created (neglecting the first 50,000 draws).

levels of GDP and the main exogenous components show a dynamic in line with the data (dashed line).

7.2. Impact of the COVID-19 pandemic differentiated by industries and components of GDP

Next, we break down the scenarios from Section 7.1 for different sectors by economic activities. Figure 21 shows the growth rates for sectoral gross value added (GVA) of the three scenarios. GVA is disaggregated for ten economic activities (NACE*10) according to the statistical classification of economic activities in the European Community (NACE Rev. 2). See Table B.7 in the appendix

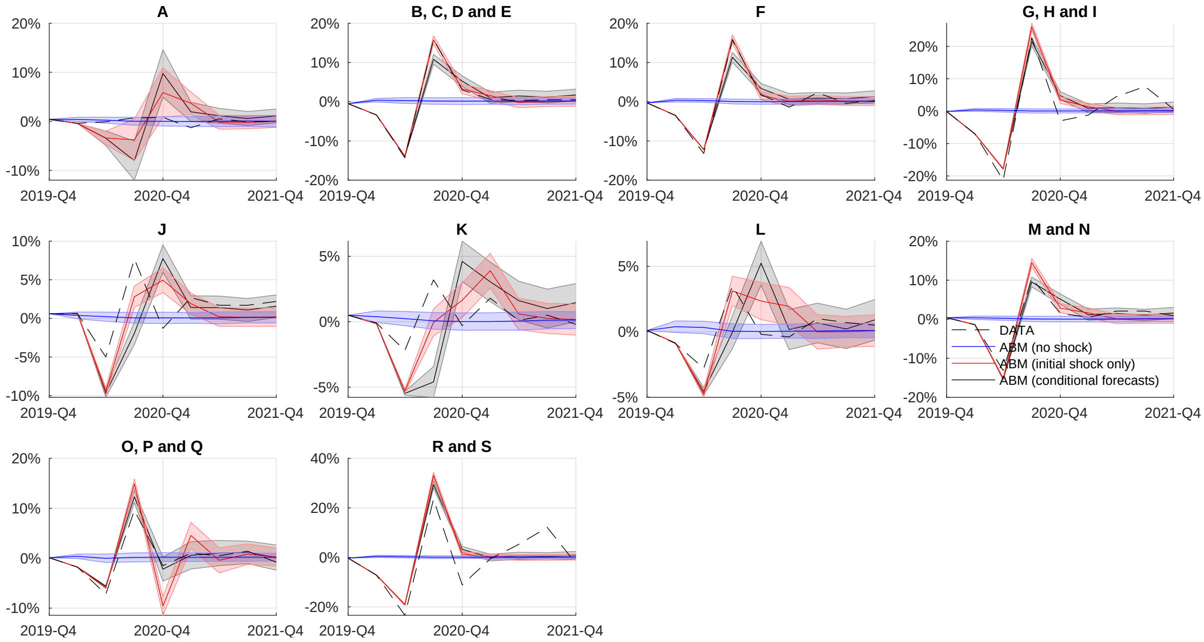

FIGURE 21. Sectoral conditional forecasts estimated on data up to the fourth quarter of 2019. Figures show sectoral gross value added (GVA) (annually, in Euro) for ABM (solid line, a 90 per cent confidence interval is plotted around the mean trajectory) and observed Eurostat data for the euro area (dashed line). GVA is disaggregated for 10 economic activities (NACE*10) according to the statistical classification of economic activities in the European Community (NACE Rev. 2), see Table B.7 in the appendix for details. Model results are obtained as an average over 500 Monte Carlo simulations.

for details.[21] Again, the baseline scenario shows a projection of the trend for the sectorial growth rates without the COVID-19 recession. With the domestic demand shock caused by the restrictions of economic activities due to the lockdown measures in the first two quarters of 2020 (initial shock scenario shown in red), we see a sharp decline of the sectoral growth rates in the first and second quarter of 2020 followed by a fast recovery in the following quarters in most sectors. Overall sectoral growth rates are in line or close to observed data for the euro area in almost all sectors (with the exception of the Agricultural sector). Notably, with only the initial shock (red line), both the sharp decline and the relatively swift recovery of the sectorial growth rates match the observed data for the euro area (dashed line) for the trade, transport and hospitality sector (G, H and I) and the arts and entertainment sector (R and S). These sectors were particularly affected by the lockdowns and are mostly driven by domestic demand.

21. Note the varying scales for the sectors of different sizes.

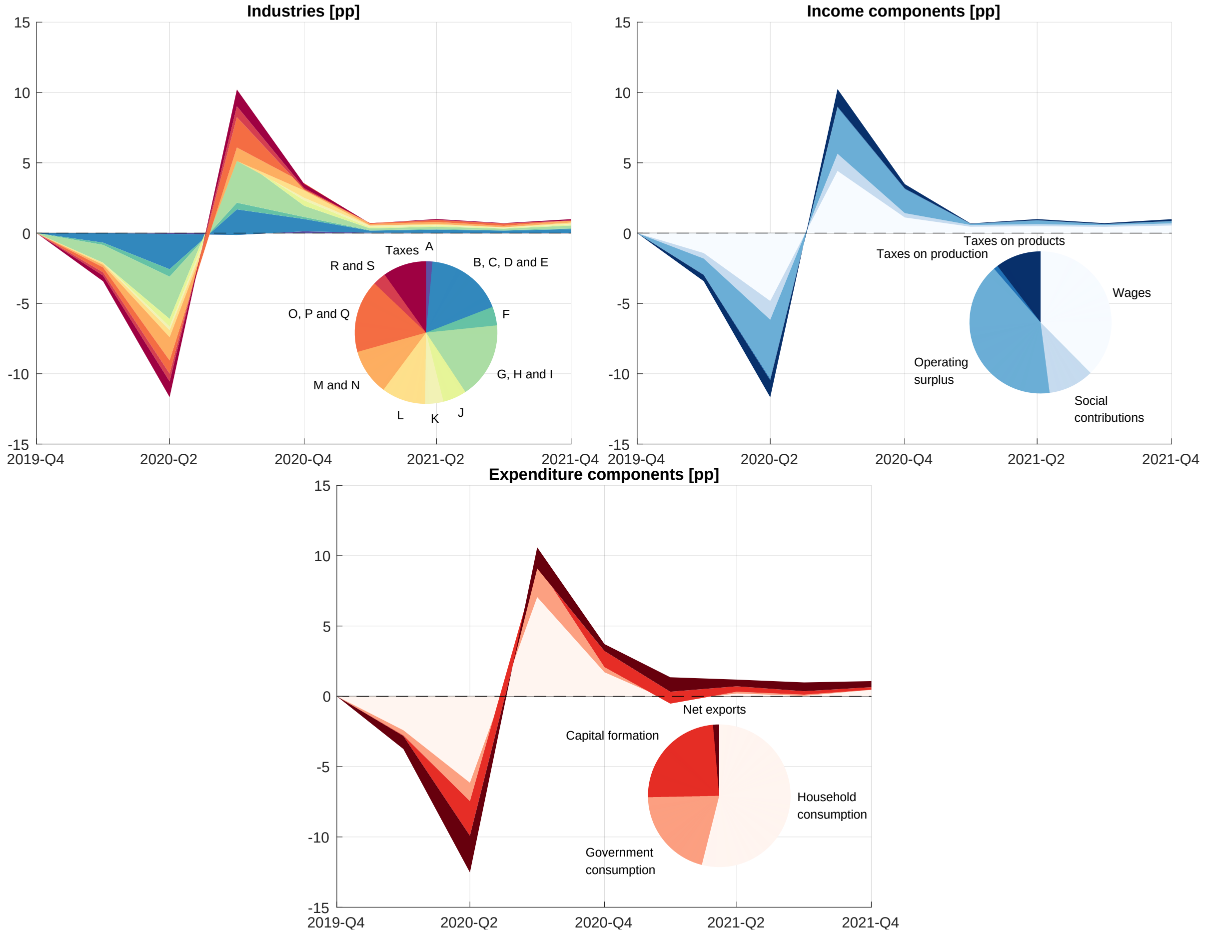

FIGURE 22. Composition of GDP according to production, income and expenditure approaches. Colored areas show conditional forecasts of the ABM estimated on data up to the fourth quarter of 2019, again as an average over 500 Monte Carlo simulations. The dashed line shows the corresponding values from observed data for the euro area.

Finally, we look into the contribution of individual sectors and industries to the total changes in GDP due to the COVID-19 pandemic. These contributions can be seen in Figure 22, which shows the contribution of individual sectors to the total changes in GDP due to the COVID-19 pandemic (the stacked area under the line diagrams) and relates them to the total shares of these sectors in the economy (the pie diagram in the right lower corner). In Figure 22 (top left), we show the contribution of individual industries to the total changes in GDP of the conditional forecasts. GDP is disaggregated for ten economic activities (NACE*10) according to the statistical classification of economic activities in the European Community (NACE Rev. 2). See Table B.7 in the appendix for details. Again, the disproportionate contribution of the trade, transport and hospitality sectors (G, H and I) and the arts and entertainment sector (R and S) is clearly visible. Additionally, as can be seen in Figure 22 (top left), the government and healthcare sector has a disproportionate contribution to the recovery in the second half of 2020 in the euro area.

In Figure 22, we show the contributions of the income (top right) and the expenditure components (bottom centre) of GDP of the conditional forecasts. The decomposition into expenditure components, Figure 22 (bottom centre), shows that household consumption has by far the largest contribution of all expenditure components to the sharp decline and the relatively swift recovery of GDP growth. On the other hand, it is interesting to note that according to the breakdown in GDP income components, Figure 22 (top right), this reduction in household consumption translates into a loss of operating surplus and lower wage income in the first half of 2020, which is then followed by a recovery in the second half of 2020 and 2021. Clearly, from the decomposition into expenditure components, Figure 22 (bottom centre), it is visible that government consumption made a disproportionate contribution to the decline and recovery of GDP growth. In particular, in the second half of 2020, government consumption made a substantial contribution to the recovery.

8. Conclusion

In this paper, we developed an ABM for analyzing and forecasting economic crises. The model has many features required of next-generation macroeconomic models. It relaxes rational expectations to bounded rationality by approximating them with adaptive learning. It further incorporates financial frictions and allows for non-linear responses. We demonstrated the model on the three recent major economic crises of the euro area: the Financial crisis of 2007-2008 and the subsequent Great Recession, the European sovereign debt crisis, and the COVID-19 recession.

In Poledna et al. (2023), we showed that an ABM can be competitive with benchmark VAR and DSGE models in out-of-sample forecasting of macro variables of a small open economy. In this paper, we demonstrated that the model can be adopted for a large open economy. This result is more significant for the following reasons. First, the usefulness of the model has been demonstrated for the euro area, the third largest economy after the U.S. and China. Second, unlike the results in Poledna et al. (2023), the out-of-sample forecasts are, in general, significantly better (or close to the ten per cent significance level) than the forecasts of the benchmark models. Third, the sample period includes the Financial crisis of 2007-2008.

By incorporating a financial accelerator mechanism, we further demonstrated that the ABM can be calibrated to analyze and potentially forecast financial crises without an exogenous shock. In the ABM, the Financial crisis of 2007-2008 and the subsequent Great Recession unfolds in the following way. Before the financial crisis, we observed a period of relatively high growth and investment, as well as high-interest rates and low profitability, leading to high debt levels. When in the model, the debt-to-asset ratio of firms crosses a threshold due to a fluctuation in asset prices, firms experience a credit crunch and cannot invest anymore. This non-linear response is then reinforced through the financial accelerator and causes an economic crisis without an exogenous shock.

We demonstrated the model with the financial accelerator during the Financial crisis of 2007-2008 and the subsequent Great Recession, as well as the European sovereign debt crisis—two crises that DSGE models struggled to predict and have difficulties explaining. By making use of out-of-sample forecasting exercises, we have shown that the model predicts an endogenous crisis around the most intense phase of the Great Recession in the euro area, albeit with lower severity without a global downturn (which is exogenous to the model). With conditional forecasts, which include an exogenous shock on exports from the global downturn, we have demonstrated that the model explains both the severity and slow recovery of the Great Recession.

By analysing the COVID-19 recession, we further demonstrated the potential of the model for scenario analysis with exogenous shocks. By implementing an industry-specific shock caused by the restrictions of economic activities due to the lockdown measures, we showed that the model reproduces the observed deep recession followed by a swift recovery. By additionally adding an export demand and import supply shock from the global supply chain crisis, as well as a fiscal shock from increased government spending, we demonstrated that the model also captures the persistent rise of inflation following the COVID-19 recession.

The calibration of parameters related to the financial side of the economy and the financial accelerator mechanism, in particular, remains a difficult task. Unlike non-financial variables, financial variables such as assets, debt, and interest remain unavailable at a sufficiently disaggregated level in official statistics, i.e., these variables are not available at the two-digit NACE level. Thus, assumptions on the distribution of assets and debt over sectors need to be made to calibrate the model. This lack of data impedes the potential to forecast future financial crises and may also contribute to an, in general, lower forecast performance of the ABM with the financial accelerator similar to DSGE models with financial frictions that also tend to have a lower forecasting performance than DSGE models without financial frictions.

These limitations could be overcome by using additional microdata and increased computing power. Computing power and registry-based firm and individual-level data are increasingly available and could be used to calibrate parameters on the agent level. Thus, the grand challenge and long-term objective remain to create a "Big Data ABM" research program to develop ABMs based on micro and macro data to monitor the economy in real-time using supercomputers.

References

Assenza, Tiziana, Domenico Delli Gatti, and Jakob Grazzini (2015). "Emergent dynamics of a macroeconomic agent based model with capital and credit." *Journal of Economic Dynamics and Control*, 50, 5–28.

Axtell, Robert L. (2001). "Zipf distribution of US firm sizes." *Science*, 293(5536), 1818–1820.

Basurto, Alessandro, Herbert Dawid, Philipp Harting, Jasper Hepp, and Dirk Kohlweyer (2022). "How to design virus containment policies? A joint analysis of economic and epidemic dynamics under the COVID-19 pandemic." *Journal of Economic Interaction and Coordination.*

Blanchard, Olivier (2016). "Do DSGE Models have a Future?" *PIIE Policy Brief*, PB 16-11.

Blattner, Tobias S. and Emil Margaritov (2010). "Towards a Robust Monetary Policy Rule for the Euro Area." *ECB Working Paper No. 1210.*

Brayton, Flint, Eileen Mauskopf, David Reifschneider, Peter Tinsley, John Williams, Brian Doyle, and Steven Sumner (1997). "The Role of Expectations in the FRB/US Macroeconomic Model." *Federal Reserve Bulletin*, April 1997.

Brzoza-Brzezina, Michał and Marcin Kolasa (2013). "Bayesian Evaluation of DSGE Models with Financial Frictions." *Journal of Money, Credit and Banking*, 45(8), 1451–76.

Canova, Fabio and Luca Sala (2009). "Back to square one: Identification Issues in DSGE Models." *Journal of Monetary Economics*, 56, 431 – 449.

Chatterjee, Satyajit, Dean Corbae, Makoto Nakajima, and José-Víctor Ríos-Rull (2007). "A quantitative theory of unsecured consumer credit with risk of default." *Econometrica*, 75(6), 1525–1589.

Christiano, Lawrence J., Martin S. Eichenbaum, and Mathias Trabandt (2018). "On DSGE models." *Journal of Economic Perspectives*, 32(3), 113–40.

Colander, David, Michael Goldberg, Armin Haas, Katarina Juselius, Alan Kirman, Thomas Lux, and Brigitte Sloth (2009). "The financial crisis and the systemic failure of the economics profession." *Critical Review*, 21:2-3, 249 – 267.

Dawid, Herbert and Domenico Delli Gatti (2018). "Agent-Based Macroeconomics." In *Handbook of Computational Economics*, *Handbook of Computational Economics*, vol. 4, edited by Cars Hommes and Blake LeBaron, pp. 63 – 156. Elsevier.

De Grauwe, Paul and Yuemei Ji (2020). "Structural reforms, animal spirits, and monetary policies." *European Economic Review*, 124, 103395.

Del Negro, Marco, Raiden B. Hasegawa, and Frank Schorfheide (2016). "Dynamic prediction pools: An investigation of financial frictions and forecasting performance." *Journal of Econometrics*, 192(2), 391–405. Innovations in Multiple Time Series Analysis.

Del Negro, Marco and Frank Schorfheide (2013). "DSGE model-based forecasting." In *Handbook of economic forecasting*, vol. 2, pp. 57–140. Elsevier.

Delli Gatti, Domenico and Severin Reissl (2022). "Agent-Based Covid economics (ABC): Assessing non-pharmaceutical interventions and macro-stabilization policies." *Industrial and Corporate Change*, 31(2), 410–447.

Diebold, Francis X and Roberto S Mariano (1995). "Comparing Predictive Accuracy." *Journal of Business & Economic Statistics*, 13(3).

Edge, Rochelle M. and Refet S. Gurkaynak (2010). ""How Useful Are Estimated DSGE Model Forecasts for Central Bankers?" *Brookings Papers on Economic Activity*, 41(2), 209–259.

Evans, George W. and Seppo Honkapohja (2001). *Learning and expectations in macroeconomics*. Princeton University Press.

Haldane, Andrew G. and Arthur E. Turrell (2018). "An interdisciplinary model for macroeconomics." *Oxford Review of Economic Policy*, 34(1-2), 219–251.

Harvey, David, Stephen Leybourne, and Paul Newbold (1997). "Testing the Equality of Prediction Mean Squared Errors." *International Journal of Forecasting*, 13(2), 281–291.

Hendry, David F and John N J Muellbauer (2018). "The future of macroeconomics: macro theory and models at the Bank of England." *Oxford Review of Economic Policy*, 34(1-2), 287–328.

Hommes, Cars (2021). “Behavioral and Experimental Macroeconomics and Policy Analysis: A Complex Systems Approach.” *Journal of Economic Literature*, 59(1), 149–219.

Hommes, Cars, Mario He, Melissa Siqueira, Sebastian Poledna, and Yang Zhang (2022a). “CANVAS: A Canadian Behavioral Agent-Based Model.” *Bank of Canada Staff Working Paper*, 2022-51.

Hommes, Cars, Kostas Mavromatis, Tolga Özden, and Mei Zhu (2022b). “Behavioral learning equilibria in the New Keynesian model.” *Bank of Canada Staff Working Paper*, 2022-42.

Hommes, Cars and Mei Zhu (2014). “Behavioral learning equilibria.” *Journal of Economic Theory*, 150, 778–814.

Ijiri, Yuji and Herbert A. Simon (1977). *Skew distributions and the sizes of business firms*. North-Holland.

Kaplan, Greg, Benjamin Moll, and Giovanni L Violante (2018). “Monetary policy according to HANK.” *American Economic Review*, 108(3), 697–743.

Kaplan, Greg and Giovanni L Violante (2014). “A model of the consumption response to fiscal stimulus payments.” *Econometrica*, 82(4), 1199–1239.

Kaplan, Greg and Giovanni L Violante (2018). “Microeconomic heterogeneity and macroeconomic shocks.” *Journal of Economic Perspectives*, 32(3), 167–94.

Khan, Aubhik and Julia K Thomas (2008). “Idiosyncratic shocks and the role of nonconvexities in plant and aggregate investment dynamics.” *Econometrica*, 76(2), 395–436.

Kirman, Alan (2010). “The Economic Crisis is a Crisis for Economic Theory.” *CESifo Economic Studies*, 56 (4), 498–535.

Krugman, Paul (2011). “The Profession and the Crisis.” *Eastern Economic Journal*, 37, 307 – 312.

Krugman, Paul (2018). “Good enough for government work? Macroeconomics since the crisis.” *Oxford Review of Economic Policy*, 34(1-2), 156–168.

Lindé, Jesper (2018). “DSGE models: still useful in policy analysis?” *Oxford Review of Economic Policy*, 34(1-2), 269–286.

Lindé, Jesper, Frank Smets, and Rafael Wouters (2016). “Challenges for Central Banks’ macro models.” In *Handbook of macroeconomics*, vol. 2, pp. 2185–2262. Elsevier.

McKay, Alisdair and Ricardo Reis (2016). “The role of automatic stabilizers in the US business cycle.” *Econometrica*, 84(1), 141–194.

Milani, Fabio (2007). “Expectations, learning and macroeconomic persistence.” *Journal of Monetary Economics*, 54(7), 2065–2082.

Milani, Fabio (2012). "The Modeling of Expectations in Empirical DSGE Models: A Survey." In *DSGE Models in Macroeconomics: Estimation, Evaluation, and New Developments (Advances in Econometric, Vol. 28)*, edited by Nathan Balke, Fabio Canova, Fabio Milani, and Mark A. Wynne, pp. 3–38. Emerald Group Publishing Limited.

Pangallo, Marco, Alberto Aleta, R Chanona, Anton Pichler, David Martín-Corral, Matteo Chinazzi, François Lafond, Marco Ajelli, Esteban Moro, Yamir Moreno, et al. (2022). "The unequal effects of the health-economy tradeoff during the COVID-19 pandemic." *arXiv preprint arXiv:2212.03567*.

Poledna, Sebastian, Michael Gregor Miess, Cars Hommes, and Katrin Rabitsch (2023). "Economic forecasting with an agent-based model." *European Economic Review*, 151, 104306.

Poledna, Sebastian, Elena Rovenskaya, Jesus Crespo Cuaresma, Serguei Kaniovski, and Michael Miess (2020). "Recovery of the Austrian Economy Following the COVID-19 Crisis Can Take up to Three Years." *IIASA Policy Brief #26*.

Romer, Paul (2016). "The trouble with macroeconomics." *The American Economist*, 20, 1–20.

Slobodyan, Sergey and Rafael Wouters (2012). "Learning in a medium-scale DSGE model with expectations based on small forecasting models." *American Economic Journal: Macroeconomics*, 4(2), 65–101.

Smets, Frank and Rafael Wouters (2007). "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach." *American Economic Review*, Vol. 97, No. 3, 586 – 606.

Stiglitz, Joseph E. (2011). "Rethinking Macroeconomics: What failed, and how to repair it." *Journal of the European Economic Association*, 9(4), 591–645.

Stiglitz, Joseph E (2018). "Where modern macroeconomics went wrong." *Oxford Review of Economic Policy*, 34(1-2), 70–106.

Vines, David and Samuel Wills (2018). "The rebuilding macroeconomic theory project: an analytical assessment." *Oxford Review of Economic Policy*, 34(1-2), 1–42.

Wren-Lewis, Simon (2018). "Ending the microfoundations hegemony." *Oxford Review of Economic Policy*, 34(1-2), 55–69.

Xiao, Wei and Junyi Xu (2014). "Expectations and optimal monetary policy: A stability problem revisited." *Economics Letters*, 124(2), 296–299.

Appendix A: Additional details on the agent-based model

A.1. Modifications to the agent-based model to assess the COVID-19 recession

For the analysis of the COVID-19 recession, we made a number of modifications with respect to the model presented in Section 2. Implementation of the initial domestic shock caused by the restrictions of economic activities due to the lockdown measures in the first two quarters of 2020 involves modifying Equation (3), which is replaced by

$$Q_i^{\mathrm{s}}(t) = Q_i^{\mathrm{d}}(t-1)(1+\gamma_i^{\mathrm{X}}(t))(1+\gamma^{\mathrm{e}}(t)) \,, \tag{A.1}$$

where $\gamma_i^{\mathrm{X}}(t)$ is the domestic shock to firm i at time t caused by the restrictions of economic activities due to the lockdown measures. Thus, the supply choice of firm i is reduced by a factor that represents the duration and extent of the lockdown to the firm. To calibrate the domestic shock, we use industry-specific growth rates at the one-digit NACE level for the first two quarters of 2020.

To implement the initial shock caused by the lockdown measures in the rest of the world, we replace Equations (11), (14) and (17) and set imports, exports and government consumption according to observed data for the first two quarters of 2020. After the initial shock, we assume the supply of imports to be exogenously given and following the pre-pandemic trend.[22] Thus, the total supply of imports $Y^{\mathrm{I}}(t)$ is assumed to follow an autoregressive process of lag order one (AR(1)):

$$\log\left(Y^{\mathrm{I}}(t)\right) = \alpha^{\mathrm{I}} \log\left(Y^{\mathrm{I}}(t-1)\right) + \beta^{\mathrm{I}} + \epsilon^{\mathrm{I}}(t-1) \,, \tag{A.2}$$

where $\epsilon^I(t)$ is normally distributed with standard deviation σ^I.

To implement the short-time work policy instrument, Equations (A.41) and (A.60) from Poledna et al. (2023), as well as the search-and-matching mechanism in the labour market are modified. Equation (A.41) is replaced by

$$w_h(t) = \begin{cases} \max(0.9\bar{w}_i(t), w_i(t)) & \text{if on short-time work at firm } i \\ w_i(t) & \text{if employed by firm } i \\ w_h(t-1) & \text{otherwise, i.e. if unemployed.} \end{cases} \tag{A.3}$$

22. This implies that imports to the domestic economy are subject to a supply constraint. However, demand for imports is endogenous.

Thus if firm i makes use of the short-time work policy instrument, employees always receive at least 90 per cent of the average wage ($\bar{w}_i$) equivalent to a full-time position at firm i. To account for short-time work in the government expenditures, Equation (A.60) is replaced by

$$\begin{aligned}\Pi^{\mathrm{G}}(t) = & \underbrace{\sum_{h\in H^{\mathrm{inact}}} \bar{P}^{\mathrm{HH}}(t) sb^{\mathrm{inact}}(t) + \sum_{h\in H^{\mathrm{U}}(t)} \bar{P}^{\mathrm{HH}}(t)\theta^{\mathrm{UB}} w_h(t) + \sum_{h} \bar{P}^{\mathrm{HH}}(t) sb^{\mathrm{other}}(t)}_{\text{Social benefits and transfers}} \\ & + \underbrace{\sum_j C_j(t)}_{\text{Government consumption}} + \underbrace{r^{\mathrm{G}} L^{\mathrm{G}}(t-1)}_{\text{Interest payments}} \\ & + \underbrace{\sum_i (1+\tau^{\mathrm{SIF}}) \max(0, 0.9\bar{w}_i(t) - w_i(t)) N_i(t) \bar{P}^{\mathrm{HH}}(t)}_{\text{Short-time work}} - \underbrace{Y^{\mathrm{G}}(t)}_{\text{Government revenues}} .\end{aligned} \tag{A.4}$$

Therefore the government pays the difference between the reduced salary ($w_i(t)$) and the equivalent of 90 per cent of the wage of a full-time position ($\bar{w}_i$). The search-and-matching mechanism in the labour market is modified to give firms the option to use short-time work instead of laying off employees, for which we assume a probability of two-thirds. Moreover, we assume the short-time work policy instrument can be used by firms until the end of the second quarter of 2021, i.e. for up to 6 simulation periods.

For the scenario with the initial shock from the lockdown measures and export demand and import supply shocks from the global supply chain crisis, as well as a fiscal shock, we additionally modify Equation (15) and set imports, exports, and government consumption, as well as the respective deflators, according to observed data throughout the simulations. Thus, to model the global supply chain crisis, we assume euro area countries are price takers and subject to an import supply constraint.

Appendix B: Additional tables

Table B.1. Optimized log-likelihood of VAR models of different lag orders

Order of the VAR	Log-likelihood
VAR(1)	1886.46
VAR(2)	1846.95
VAR(3)	1808.87

Note: All models are estimated using the period 1996:Q1 to 2016:Q4.

Table B.2. Out-of-sample forecast performance of VAR models of different lag orders

	GDP	Inflation	Euribor	Government consumption	Exports
VAR(1)	*RMSE-statistic for different forecast horizons*				
1q	0.74	0.21	0.09	0.31	2.1
2q	1.63	0.21	0.18	0.48	4.88
4q	3.59	0.23	0.39	0.88	10.73
8q	6.98	0.25	0.71	1.75	20.46
12q	7.72	0.22	0.7	2.61	22.47
VAR(2)	*Percentage improvements (+) or losses (-) relative to VAR(1) model*				
1q	-1 (0.64)	5.4 (0.27)	-10 (0.87)	-11.7 (0.94)	-0.9 (0.74)
2q	0.6 (0.35)	6.1 (0.30)	-12 (0.84)	-15.4 (0.95)	0.9 (0.21)
4q	3.3 (0.03)	6.5 (0.31)	-18.3 (0.84)	-11.6 (0.92)	5.1 (0.02)
8q	8.6 (0.09)	-0.4 (0.53)	-11.5 (0.82)	-7.4 (0.76)	12 (0.04)
12q	12.3 (0.10)	-28.1 (0.89)	4.1 (0.20)	-4.2 (0.72)	17.6 (0.01)
VAR(3)	*Percentage improvements (+) or losses (-) relative to VAR(1) model*				
1q	-6 (0.83)	-7.3 (0.96)	-13.7 (0.94)	-34.6 (0.99)	-6.7 (0.79)
2q	-3.1 (0.65)	-5 (0.71)	-15.6 (0.85)	-29.3 (0.91)	-5.7 (0.75)
4q	-2.5 (0.67)	0.8 (0.48)	-27.4 (0.86)	-15.2 (0.86)	-3.2 (0.69)
8q	7.1 (0.12)	1.3 (0.45)	-21.8 (0.82)	-14.1 (0.76)	11.5 (0.02)
12q	23.7 (0.13)	-43.5 (0.90)	15.2 (0.04)	-8.2 (0.73)	34.7 (0.03)

Note: The forecast period is 2005:Q2 to 2019:Q4. The VAR models are estimated starting in 1996:Q1 and are re-estimated each quarter. In parentheses, we show p-values of (modified) Diebold-Mariano tests (Harvey et al., 1997), where we test whether forecasts are significantly different in accuracy than the VAR(1) (the null hypothesis of the test is that the VAR(2) and the VAR(3) are less accurate than the VAR(1)).

Table B.3. Optimized log-likelihood of AR models of different lag orders

Order of the AR	GDP	Inflation	Euribor	Household consumption	Investment
AR(1)	327.51	389.82	469.18	356.09	224.99
AR(2)	315.10	377.63	468.83	352.18	224.23
AR(3)	304.16	372.06	450.68	341.58	217.32

Note: All models are estimated using the period 1996:Q1 to 2016:Q4.

Table B.4. Out-of-sample forecast performance of AR models of different lag order

	GDP	Inflation	Euribor	Household consumption	Investment
AR(1)	*RMSE-statistic for different forecast horizons*				
1q	0.71	0.2	0.1	0.36	2.15
2q	1.54	0.21	0.18	0.68	3.12
4q	3.35	0.21	0.28	1.42	5.35
8q	7.5	0.21	0.37	2.82	8.81
12q	12.36	0.2	0.4	4.11	11.44
AR(2)	*Percentage improvements (+) or losses (-) relative to AR(1) model*				
1q	-2.9 (0.83)	4.7 (0.06)	15.4 (0.06)	13.3 (0.00)	4.5 (0.20)
2q	-5.5 (0.76)	5.2 (0.02)	5.4 (0.25)	18.5 (0.02)	1.1 (0.45)
4q	-16.6 (0.84)	5.3 (0.00)	-6.8 (0.75)	14.2 (0.04)	-7.1 (0.82)
8q	-51.3 (0.85)	2.6 (0.01)	-27.8 (0.91)	8.1 (0.06)	-17.5 (0.87)
12q	-111 (0.87)	3.2 (0.00)	-43.8 (0.96)	5.7 (0.16)	-19.8 (0.89)
AR(3)	*Percentage improvements (+) or losses (-) relative to AR(1) model*				
1q	-4 (0.83)	1.4 (0.34)	12.7 (0.07)	13.4 (0.02)	4.9 (0.21)
2q	-5.6 (0.77)	2.1 (0.12)	4.5 (0.28)	22.2 (0.04)	-0.7 (0.52)
4q	-12.8 (0.84)	3.7 (0.03)	-6.6 (0.75)	17.8 (0.07)	-14.4 (0.82)
8q	-30.5 (0.85)	2.2 (0.00)	-27.1 (0.90)	10.6 (0.13)	-40.5 (0.86)
12q	-56.1 (0.87)	2 (0.00)	-41.6 (0.96)	9.2 (0.25)	-60.7 (0.88)

Note: The forecast period is 2005:Q2 to 2019:Q4. The AR models are estimated starting in 1996:Q1 and are re-estimated each quarter. In parentheses, we show p-values of (modified) Diebold-Mariano tests (Harvey et al., 1997), where we test whether forecasts are significantly different in accuracy than the AR(1) (the null hypothesis of the test is that the AR(2) and the AR(3) are less accurate than the AR(1)).

Table B.5. Optimized log-likelihood of VARX models of different lag orders

Order of the VARX	Log-likelihood
VARX(1)	1245.80
VARX(2)	1229.09
VARX(3)	1191.57

Note: All models are estimated using the period 1996:Q1 to 2016:Q4.

Table B.6. Conditional forecast performance of VARX models of different lag orders

	GDP	Inflation	Euribor
VARX(1)	*RMSE-statistic for different forecast horizons*		
1q	0.63	0.22	0.1
2q	1.07	0.27	0.17
4q	1.65	0.37	0.3
8q	2.21	0.35	0.59
12q	2.62	0.26	0.85
VARX(2)	*Percentage improvements (+) or losses (-) relative to VARX(1) model*		
1q	4.4 (0.09)	-0.5 (0.55)	1.5 (0.44)
2q	5.6 (0.06)	10.1 (0.03)	-7.8 (0.87)
4q	7.6 (0.03)	11.7 (0.14)	-4.5 (0.99)
8q	8.1 (0.04)	11.5 (0.11)	4.2 (0.27)
12q	6.4 (0.02)	10.7 (0.07)	8.6 (0.15)
VARX(3)	*Percentage improvements (+) or losses (-) relative to VARX(1) model*		
1q	7.7 (0.05)	-10.8 (0.85)	3 (0.37)
2q	8.5 (0.14)	-2.9 (0.58)	1.3 (0.41)
4q	15.7 (0.01)	-12.1 (0.73)	11.2 (0.19)
8q	12.7 (0.02)	-40 (0.83)	16.9 (0.16)
12q	10.6 (0.00)	-61.1 (0.85)	12.8 (0.21)

Note: The forecast period is 2005:Q2 to 2019:Q4. The VARX models are estimated starting in 1996:Q1 and are re-estimated each quarter. In parentheses, we show p-values of (modified) Diebold-Mariano tests (Harvey et al., 1997), where we test whether forecasts are significantly different in accuracy than the VARX(1) (the null hypothesis of the test is that the VARX(2) and the VARX(3) are less accurate than the VARX(1)).

Table B.7. Statistical classification of economic activities in the European Community (NACE Rev. 2)

NACE	Description
A	Agriculture, forestry and fishing
B, C, D and E	Industry (except construction)
F	Construction
G, H and I	Wholesale and retail trade, transport, accomodation and food service activities
J	Information and communication
K	Financial and insurance activities
L	Real estate activities
M and N	Professional, scientific and technical activities; administrative and support service activities
O, P and Q	Public administration, defence, education, human health and social work activities
R and S	Arts, entertainment and recreation; other service activities; activities of household and extra-territorial organizations and bodies