Generative Economic Modeling*

Incomplete and preliminary

Hanno Kase
European Central Bank

Matthias Rottner
Bank for International Settlements
Deutsche Bundesbank

Fabio Stohler
University of Bonn

September 7, 2025

Abstract

We introduce a novel approach for solving quantitative economic models: generative economic modeling. Our method combines neural networks with conventional solution techniques. Specifically, we train neural networks on simplified versions of an economic model to generate approximations of the full model's dynamic behavior. By relying on these less complex satellite models, we circumvent the curse of dimensionality and are able to employ well-established numerical methods. We demonstrate our approach on models with nonlinear dynamics and heterogeneous agents. Finally, we apply generative economic modeling to solve a high-dimensional HANK model with financial frictions.

Keywords— Machine Learning, Neural networks, Nonlinearities, Heterogeneous Agents
JEL codes— C11, C45, D31, E32, E52

*hanno.kase@ecb.europa.eu, matthias.rottner@bis.org, fabio.stohler@uni-bonn.de. We would like to thank Christian Bayer, Jesús Fernández-Villaverde, and Jamie Lenney for their very helpful suggestions. Fabio Stohler gratefully acknowledges support by the German Research Foundation (DFG) through CRC TR 224 (Project C05). The views in this paper are those of the authors and should not be interpreted as reflecting the views of the Bank for International Settlements, the Deutsche Bundesbank, the European Central Bank, or any other person associated with the Eurosystem.

1 Introduction

The advances in artificial intelligence provide substantial opportunities for quantitative economics by shifting the production-possibility frontier of modeling. Deep learning has emerged as a powerful tool for solving dynamic economic models that were previously considered intractable. The curse of dimensionality - first articulated by Bellman (1957) - typically limits the complexity of economic models to only a few state variables. However, deep learning can help to tame this problem, as discussed in Fernández-Villaverde, Nuño and Perla (2024). Unfortunately, successfully employing deep learning in practice often necessitates meticulous and detailed adjustments tailored to the specific model at hand, which makes it challenging to employ. In contrast, established conventional solution methods, though constrained by the curse of dimensionality, are already specifically designed and optimized for particular types of economic models and feature well-understood emergence properties. To combine the strengths of both artificial intelligence and conventional solution methods, we propose a novel approach: generative economic modeling.

Our approach uses neural networks to approximate the full economic model. However, instead of using the complete model as input for the training process, the neural network learns from a collection of simplified models, which we refer to as "satellite models". Each satellite model includes only a subset of features and states, featuring only partially the dynamics of the full model.[1] By relying on these satellite models, we can solve these simpler models with conventional methods without running in the curse of dimensionality. We can then use the solved model to simulate it. The generated simulated data is used to train the neural network to learn the dynamics of the model. While each satellite model captures only some features of the underlying economic model, we ensure an overlap of features in the satellite models. The neural network is then trained on this data to learn to approximate the solution of a complete model that includes all features and states.

Our method belongs to the class of generative artificial intelligence because we employ the neural network to generate results for the complete model that includes all features and states, something we have not used for the training process. Generative artificial intelligence has achieved significant success, especially in the context of large language models, which are very large deep learning models. While this success also holds promise for our approach, the generative performance of neural networks in our context of economic modeling is less clear.

In this paper, we establish that we can use our method for quantitative economic mod-

[1] Most researchers are used to working with a satellite version of the model, as computational requirements often force them to limit the size of their model. Our approach formalizes how to use a set of satellite models and apply deep learning techniques to work with a more complex version of the model.

els that feature nonlinear dynamics or heterogeneous agents. We first focus on practical demonstrations of our method to solve a variety of models with increasing complexity. We benchmark our new methodology against analytical solutions or widely used alternative global solution methods. Our reference model is a real business cycle model, where we work with three versions: i) a simplified version that can be solved analytically, ii) a medium-sized version with state-dependent investment costs that result in distinct nonlinearities, and iii) a heterogeneous agents version with partially uninsurable income risk in line with Krusell and Smith (1998). We construct for all these models overlapping satellite models and then use simulated data to train our neural network. Thus, we can evaluate how well our approach is designed to capture nonlinearities and how applicable it is to the class of heterogeneous agent versions.

Our methodology provides a very good approximation in all three versions of the model. When evaluated against the true model, our approach consistently yields low prediction errors below one percent, highlighting its robustness and accuracy. Notably, the magnitude of these errors is comparable to that of the neural network trained on data generated by the full model. The similarity in performance suggests that the observed errors stem from imperfections in neural network training itself, rather than from differences in the training datasets. This holds for the aggregate dynamics in the analytical model, the heterogeneous agent version as well as the nonlinear dynamics in the medium-sized model with piecewise nonlinear capital adjustment costs, highlighting its flexibility. Importantly, the latter result underscores the potential of our method for addressing models with higher-order nonlinearities, which typically require computationally intensive global solution techniques. To summarize, our generative economic modeling approach effectively approximates the dynamics of the reference model with sufficient precision. This finding underscores the potential of our approach in efficiently learning and replicating complex model dynamics from simplified model versions.

When solving the heterogeneous agent model, we leverage an additional advantage of our methodology that further reduces the numerical burden. Heterogeneous agent models with multiple aggregate shocks face a threefold curse of dimensionality. First, expanding the state space significantly increases the computational complexity of solving the household problem. Second, incorporating additional shocks complicates the accurate computation of expected values. Third, introducing more states necessitates a more involved calculation of the perceived laws of motion to forecast the future evolution of payoff-relevant aggregate variables. Our methodology addresses these challenges by effectively reducing the state space by only solving satellite models, enabling researchers to solve more intricate models with greater efficiency. By accelerating the solution process for complicated models and enabling the analysis of yet unsolvable economic environments, our approach provides a powerful tool for advancing research on nonlinear and high-dimensional macroeconomic models, especially in the heterogeneous agent literature.

We apply generative economic modeling to solve a high-dimensional heterogeneous agent New-Keynesian (HANK) model with portfolio choice over four assets. Solving such a problem with either numerical approach is challenging since the size of the state space of the household scales with the size of the asset grid of each additional asset included.[2] We circumvent the curse of dimensionality by solving satellite models where households only choose between two different assets, fixing the portfolio choice of the other assets. We train the surrogate model on the simulated data from the set of satellite models and approximate the dynamics of an economy with four assets.

Another important feature of our method is that we can cross-check its fit using metrics such as the Euler equation errors. Such metrics evaluate how well the neural network can approximate the results. Therefore, the fit of the method can be evaluated also when we move away from the laboratory setting provided above. These metrics also help set up the satellite models and specify the neural network. Instead of providing a clear handbook or proof, our approach follows the data science literature. We evaluate which combination of satellite models and neural networks provides the best approximation and use this specification for our problem at hand.

As initial motivation for our approach, we highlighted the potential fragility of deep learning when used directly to solve the economic model. However, our approach does not feature this problem because of a difference in the type of training data. The data to train in our approach is not affected by the neural network training because it is precalculated and comes from the conventional solution method. When deep learning is directly used to solve economic equations, the inputs used depend on the model solution from the neural network. Therefore, the inputs are endogenous (instead of exogenous) because they depend on the neural network. This feedback complicates finding the solution substantially, an issue that is absent in our generative economic modeling approach.

Literature Review Our paper belongs to the fast-growing literature that uses deep learning to solve dynamic economic models. The areas of application have been HANK models (Fernández-Villaverde et al., 2024; Kase, Melosi and Rottner, 2022), heterogeneous agents (Azinovic, Gaegauf and Scheidegger, 2022; Fernández-Villaverde, Hurtado and Nuno, 2023; Gorodnichenko et al., 2021; Han, Yang and E, 2021; Kahou et al., 2021;

[2] To circumvent the curse of dimensionality Reiter (2009), Gornemann, Kuester and Nakajima (2016), Ahn et al. (2017), Bayer and Luetticke (2020), Bayer, Born and Luetticke (2024) provide techniques for dimensionality reduction. Even with these techniques, the problem remains numerically intractable.

Maliar and Maliar, 2022; Maliar, Maliar and Winant, 2021), overlapping generations and life-cycle (Azinovic and Zemlicka, 2024; Druedahl and Røpke, 2025; Pascal, 2024), finance (Chen, Didisheim and Scheidegger, 2023; Duarte, Duarte and Silva, 2024; Duarte and Fonseca, 2024), labor markets and search (Adenbaum, Babalievsky and Jungerman, 2024; Jungerman, 2024; Payne, Rebei and Yang, 2024), monetary policy (Chen et al., 2021; Nuño, Renner and Scheidegger, 2024), climate change (Fernández-Villaverde, Gillingham and Scheidegger, 2024; Friedl et al., 2023), and behavorial macroeconomics (Ashwin, Beaudry and Ellison, 2025; Kahou et al., 2024). However, our approach to using neural networks deviates strongly from these papers, as we are not interested in directly solving the economic equations. More closely related to our work is the usage of neural networks as surrogate models as in Kase, Melosi and Rottner (2022) and Chen, Didisheim and Scheidegger (2023).[3] Yet, these papers use the complete underlying model for the training, thereby excluding the generative part. Finally, we also differ by following a hybrid approach that exploits the advantages of conventional solution methods and deep learning. Generative economic modeling does not feature this limitation, but applies to portfolio problems of different liquidity, as well.

Our methodology allows us to speak to the literature featuring portfolio choice in HANK models. Prominent examples of quantitative HANK models with more than one asset are Bayer et al. (2019), Auclert, Rognlie and Straub (2020), McKay and Wieland (2021), Kekre and Lenel (2022), Bayer, Born and Luetticke (2022), Bhandari et al. (2023), and Bayer, Born and Luetticke (2024). We extend the research-possibility frontier in this area with our methodology by allowing researchers to go beyond two assets to solve models with even four assets. One important exception is Auclert et al. (2024), who solve a HANK model with potentially limitless assets. However, the authors only demonstrate their methodology to solve models with portfolios of identical liquidity.

Our work also builds on the broader literature on conventional solution methods that do not rely on deep learning. Given the vast array of contributions across different fields of computational economics, providing a comprehensive review would be infeasible. Instead, we refer to the influential books on numerical methods by Judd (1998), Miranda and Fackler (2004) and Heer and Maussner (2024), which provide a great overview of the methods available. Our approach builds upon these traditional methods while leveraging deep learning to enhance their capacity to handle higher levels of complexity. The combination of the methods makes it possible to tackle problems that were previously computationally intractable.

[3] A surrogate model is an approximate model that mimics the behavior of a more complex model.

2 Generative Economic Modeling

This section outlines our generative economic modeling approach, which is designed to solve a large class of dynamic general equilibrium models.[4]

2.1 Underlying Dynamic General Equilibrium Model

The dynamics of a dynamic general equilibrium model can be expressed as a transition equation:

$$\mathbb{S}_t = f\left(\mathbb{S}_{t-1}, \nu_t | \Theta\right) \tag{1}$$

where the state vector $\mathbb{S}_t \in \mathbb{R}^m$ describes the economy in period t. Note that such a representation can contain heterogeneous agents or behavioral models. The economy is also subject to exogenous shocks that follow a Markov process, which is captured by the vector $\nu_t \in \mathbb{R}^n$. There is also a vector of structural parameters $\Theta \in \mathbb{R}^d$, which affects the dynamics of the model. The function $f(\cdot)$ determines the mapping from the previous period state variables $\mathbb{S}_{t-1}$ and current period shocks ν_t to the current period state variables $\mathbb{S}_t$ conditional on the structural parameters. To solve the transition equation, it is needed to find the policy function (decision function) which maps the model state variables $\mathbb{S}_t$ to a set of choices ψ_t.

This mapping is usually unknown and needs to be solved with numerical methods. Luckily, there already exists a large set of solution methods - more general solution approaches, like value function iteration, policy function iteration, or the endogenous grid point method, and very tailored solution methods, such as the Krusell-Smith approach for heterogeneous agent economies with aggregate risk. In general, the idea is to find an equilibrium function that maps the state variables to a set of control variables, $\psi_t = \psi(\mathbb{S}_t|\Theta)$. These policy functions satisfy a set of equations derived from the model.

$$\mathbb{F}(\psi_t(\cdot)) = 0 \tag{2}$$

Once equipped with these policy functions, we can solve for the transition equation.

The advantage of our approach is that we do not attempt to innovate on this dimension, and we leave the conventional solution steps unchanged as they have been tailored for years or even decades for specific problems. However, such conventional global solution methods face the curse of dimensionality. The exponential growth of the state space as the number of states increases limits the complexity of the economic model that can be

[4] We focus on dynamic Markov economic models, where agents solve a Markov decision problem, as e.g. in Maliar, Maliar and Winant (2021) and Fernández-Villaverde, Nuño and Perla (2024).

considered. For instance, solution methods that use full grid-based approaches require N^{m+n} points per dimension, which results in exponential computing costs, as shown by the order of the integration error $O(N^{m+n})$.[5]

2.2 Satellite Models of the Complete Model

The curse of dimensionality often forces the modeler to reduce the complexity of the studied model by limiting the number of state variables and shocks. In other words, a simplified *satellite model* is derived from the underlying complete model. In practice, using a satellite model instead of the most comprehensive model available is mostly the norm when working with global solution methods.[6] In that regard, economists are well-trained to use satellite models, and it is likely the common approach.

We denote the variables in the satellite version with a $\sim$ and rewrite the transition equation of the satellite model as:

$$\tilde{\mathbb{S}}_t = \tilde{f}\left(\tilde{\mathbb{S}}_{t-1}, \tilde{\nu}_t | \tilde{\Theta}\right) \tag{3}$$

where the dimension of the states $\tilde{\mathbb{S}}_t \in \mathbb{R}^{\tilde{m}}$ and shocks $\tilde{\nu}_t \in \mathbb{R}^{\tilde{n}}$ is smaller than in the full model, that is $(\tilde{m} < m) \vee (\tilde{n} < n)$. Note that the set of structural parameters $\tilde{\Theta} \in \mathbb{R}^{\tilde{d}}$ is then also potentially smaller, that is $\tilde{d} \leq d$.

However, we can now specify not only one satellite model, but instead several satellite models that capture different elements of the underlying model, that is $\tilde{\mathbb{S}}^a_t, \tilde{\mathbb{S}}^b_t, \tilde{\mathbb{S}}^c_t, \ldots$, where the superscript indicates the satellite model.

$$\tilde{\mathbb{S}}^i_t = \tilde{f}^i\left(\tilde{\mathbb{S}}^i_{t-1}, \tilde{\nu}^i_t | \tilde{\Theta}^i\right), \text{ for } i = a, b, c, \ldots \tag{4}$$

All these satellite models together form a set $\bar{f}_t$:

$$\bar{f}\left(\tilde{\mathbb{S}}_{t-1}, \tilde{\nu}_t | \tilde{\Theta}\right) = \left\{\tilde{f}^a\left(\tilde{\mathbb{S}}^a_{t-1}, \tilde{\nu}^a_t | \tilde{\Theta}^a\right), \tilde{f}^b\left(\tilde{\mathbb{S}}^b_{t-1}, \tilde{\nu}^b_t | \tilde{\Theta}^b\right), \tilde{f}^c\left(\tilde{\mathbb{S}}^c_{t-1}, \tilde{\nu}^c_t | \tilde{\Theta}^c\right), \ldots\right\} \tag{5}$$

Importantly, the numerical costs to extend the set increase linearly instead of exponentially (conditional on keeping the same number of states and shocks for each satellite model). The satellite models can be specified to be complete and overlapping. We define completeness as each state $\mathbb{S}_t$ and shock ν_t is at least covered in one satellite model. Therefore,

[5] Note that refinements to the solution method can lower the computational costs, e.g. adaptive sparse grids (see Brumm and Scheidegger (2017)).

[6] Models that are solved with perturbation methods are usually much larger than models solved with global solution methods, however, there also exist limits on the size of the problems. In the HANK literature, the size of the household problem is such a limiting factor.

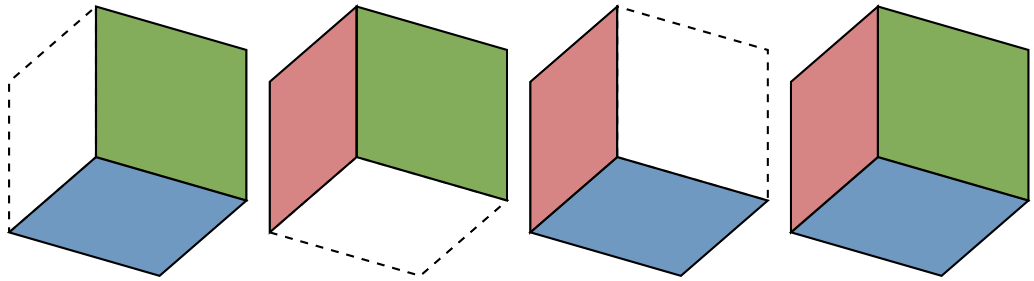


Figure 1 Illustration of Satellite Models

at least one satellite model captures one specific part of the underlying model. This requirement ensures that the set of satellite models and the true model have the same states, shocks, and parameters. We define overlapping as each state $\mathbb{S}_t$ and shock ν_t should be at least covered in two different satellite models, so that the different subsets overlap.

Illustration of satellite models To illustrate the notion of a satellite model, consider an economic model with a state vector $\mathbb{S}$ that is too complex to solve in full. Instead, we solve satellite models, each based on a subset $\tilde{\mathbb{S}}$. Figure 1 illustrates three satellite models, each containing only a subset of the states from the full model. However, we solve satellite models with overlap, such that each satellite model features two of the subsets of states. For instance, satellite model A includes the blue and green subset of states, while others incorporate different combinations. In total, we have three satellite models (A, B, and C), capturing all possible subset combinations.[7]

2.3 Deep Learning Approach to Reconstruct the Full Model

The set of satellite models is a subset of the true underlying model, that is

$$\bar{f}\left(\tilde{\mathbb{S}}_{t-1}, \tilde{\nu}_t | \tilde{\Theta}\right) \subseteq f\left(\mathbb{S}_{t-1}, \nu_t | \Theta\right) \tag{6}$$

The idea of this paper is to evaluate whether a rich specified set of satellite models is sufficient to approximate the true underlying model. Although the set of satellite models is complete and overlapping, the satellite models are only partial representations of the full models. For this reason, we want to use deep learning to learn the underlying dynamics of the full model from the set of satellite models, leveraging deep learning's generative

[7] Let $n_{\mathbb{S}}$ and $n_{\tilde{\mathbb{S}}}$ denote the number of states in the full model and in the satellite model, respectively. The number of required satellite models to achieve completeness and overlapping as defined above is determined by the binomial coefficient: $\binom{n_{\mathbb{S}}}{n_{\tilde{\mathbb{S}}}} = \frac{n_{\mathbb{S}}!}{n_{\tilde{\mathbb{S}}}!(n_{\mathbb{S}}-n_{\tilde{\mathbb{S}}})!}$.

capacity. We approximate the transition equation of the satellite models using a *surrogate model* in the form of a deep neural network $\bar{f}_{DNN}$ such that

$$\bar{f}_{DNN}\left(\tilde{\mathbb{S}}_{t-1}, \tilde{\nu}_t|\tilde{\Theta}\right) \approx \bar{f}\left(\tilde{\mathbb{S}}_{t-1}, \tilde{\nu}_t|\tilde{\Theta}\right). \tag{7}$$

We then check whether the surrogate model approximates the dynamics of the true model. Hence, we test whether for the true states $\mathbb{S}_t$, shocks ν_t and parameters Θ it holds that

$$\bar{f}_{DNN}\left(\mathbb{S}_{t-1}, \nu_t|\Theta\right) \approx f\left(\mathbb{S}_{t-1}, \nu_t|\Theta\right). \tag{8}$$

Since training a neural network introduces errors in the approximation, besides checking the approximation (8), we also check the relative performance of our surrogate model trained on data from the satellite models compared to a surrogate model trained on the true data generated from the true full model. Below, we illustrate the individual steps of this analysis.

Analysis for the full model In the proof of concept, we solve the full versions of the models using conventional solution methods and hence know the transition equation for the underlying full model. To generate the true data, we draw a sequence of vectors of shocks $\{\nu_t\}_{t=1}^N$. Given these shocks and an initial condition, we update the states $\mathbb{S}_t$ and controls ψ_t period-by-period. This results in the following dataset:

$$\mathbb{D} = \{\mathbb{S}_t, \psi_t, \nu_t\}_{t=1}^N \tag{9}$$

We use this dataset for direct evaluation of the goodness of fit of our surrogate model. Moreover, using the simulated dataset, we train a neural network to learn the mapping from previous period state variables $\mathbb{S}_t$ and shocks ν_t to current period state variables $\mathbb{S}_t$, in other words the function $f(\cdot)$. In particular, we train the neural network to minimize the error between its predicted values and simulated values, summarized as loss $\mathbb{L}$. The neural network minimizes its weights using the provided data

$$f_{DNN} = \arg\min_W \mathbb{L}(W|\mathbb{D}) \tag{10}$$

Analysis for the satellite models In practice, due to the curse of dimensionality, it would be too costly to solve the full model, so we resort to our satellite model approach. Figure 2 illustrates the individual steps graphically. First, we need to solve the individual satellite models, ensuring completeness and overlap of the satellite models. The choice of solution

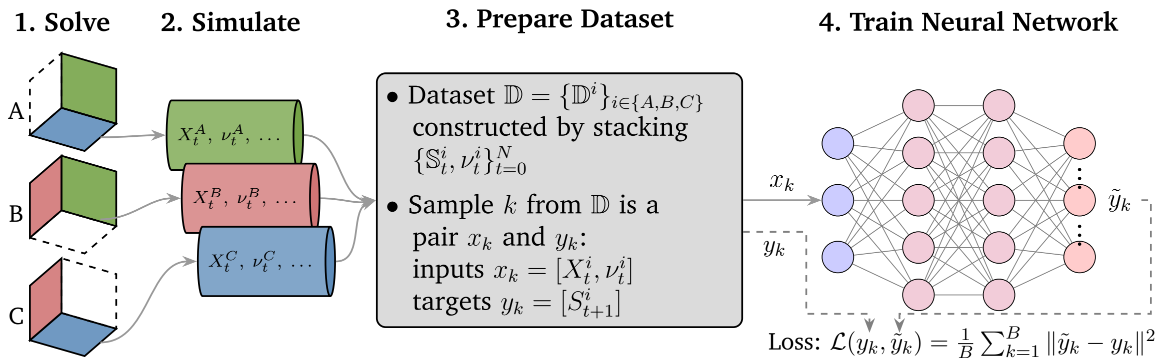


Figure 2 Flow chart of the generative economic modeling method

algorithm is left to the researcher, as our approach is compatible with any method that can solve the dynamics of the model.[8] Next, we simulate time series for each satellite model separately. We then merge all simulations, creating a long data series in which only a specific subset of states and shocks is active in different periods. The dataset that holds the simulation for all satellite models is:

$$\bar{\mathbb{D}} = \left\{ \left\{ \tilde{S}_t^a, \psi S_t^a, \tilde{\nu}_t^a \right\}_{t=1}^N, \left\{ \tilde{S}_t^b, \psi S_t^b, \tilde{\nu}_t^b \right\}_{t=1}^N, \left\{ \tilde{S}_t^c, \psi S_t^c, \tilde{\nu}_t^c \right\}_{t=1}^N, \ldots \right\} \tag{11}$$

Because our set of satellite models is complete, this training data covers all states and shocks, that is $\mathbb{S}_t$ and ν_t. We also divide the collected dataset into a training and validation sample.[9]

Finally, we train the neural network using the datasets from the satellite models by minimizing the mean squared errors between the predicted values from the neural network and the observed values from the satellite models:

$$\bar{f}_{DNN} = \arg\min_{\bar{W}} \mathbb{L}(\bar{W}|\bar{\mathbb{D}}) \tag{12}$$

By training the network on data from multiple satellite models, it learns the distinct transmission mechanisms of individual shocks, effectively generating an economic model that integrates the key features of the underlying data. This approach allows us to approximate the full model's dynamics by leveraging insights from its components.

[8] While any solution method can be used, the methodology yields optimal performance when applied to solution techniques that minimize approximation errors. The surrogate model's accuracy depends on how well the training data represents the true data generation process. In our applications later on, we employ global solution techniques to solve the model as accurately as possible.

[9] The simulated data for each satellite model is distributed between the training and validation sample to ensure that both samples contain the same share of each satellite model.

Validation For the proof of concept, we know the true model and can benchmark the performance of our methodology against the true data-generating process. Hence, we can check whether equation (8) holds approximately. When applying our methodology to a model that is otherwise unsolvable, this option is not available. While we cannot provide proof that a sufficiently rich set of satellite models approximates the true underlying model, we can directly employ standard methods to check the approximation error. In particular, we can use selected equilibrium conditions and calculate the associated Euler equation errors. Hence, we can check whether

$$\mathbb{F}(\bar{f}_{DNN}(\mathbb{S}_{t-1}, \nu_t|\Theta_t)) \approx 0. \tag{13}$$

The equilibrium conditions can be taken directly from the conventional solution step for the full model, which we never solve. However, we now use our mapping from the generative neural network to assess the fit. Note that we can also evaluate expectations over variables using methods such as Monte Carlo or Quadrature rules.

3 Proof of Concept

This section demonstrates generative economic modeling by applying our methodology to a range of models and comparing the results against those obtained with traditional global solution methods. We first describe the model environment, including its variations that entail nonlinear dynamics and heterogeneous agents. Afterwards, we showcase our new solution approach.

3.1 Model Environment: Variants of the Real Business Cycle Model

As a benchmark to evaluate our method, we choose three different versions of the real business cycle (RBC) model with distortionary taxes to showcase our method: i) a simplified version that can be solved analytically, ii) a medium-sized version with state-dependent investment costs that result in distinct nonlinearities, and iii) a heterogeneous agents version with partially uninsurable income risk in line with Krusell and Smith (1998).

We use an extended stochastic RBC model composed of a firm sector, a household sector, and a government sector to test the predictive power of our method. The firm sector comprises (i) final goods producers who bundle the intermediate goods, (ii) intermediate goods producers who rent out labor services and capital from perfectly competitive markets but face monopolistic competition in the goods market as they produce differentiated goods, and (iii) producers of capital goods who turn final goods into capital subject to adjustment costs. The adjustment costs are state-dependent, which introduces a non-linearity in the model. Prices remain flexible so monopolistic competition only has redistributive effects.

Households earn income from supplying labor n_{it} and capital k_{it}, and earn profits Π_{it} from owning the firm sector. Households spend their income for consumption c_{it} and capital investment k_{it+1}. Households can be heterogeneous in capital holdings k_{it}, in their idiosyncratic income component h_{it}, and in the allocation of profits Π_{it}.

Finally, the government levies distortionary labor- and capital-income taxes with tax rates τ^L and τ^K, besides a value-added tax on consumption τ^C. Raising taxes is purely distortionary since the government returns the tax revenues to the household via lump-sum transfers T_t.

Production sector: The production sector features final, intermediate, and capital goods producers. Final goods producers bundle varieties j of differentiated intermediate goods according to the Dixit-Stiglitz aggregator

$$Y_t = \left(\int y_{jt}^{\frac{1}{\mu_t}} dj \right)^{\mu_t}, \tag{14}$$

with elasticity of substitution $\frac{\mu_t - 1}{\mu_t}$. We allow the markup μ_t to evolve stochastically as an AR(1) in its log

$$\ln \mu_t = (1 - \rho_\mu) \left(\ln \mu - \frac{\sigma_\mu^2}{2} \right) + \rho_\mu \ln \mu_{t-1} + \epsilon_t^\mu \quad \text{with} \quad \epsilon_t^\mu \sim N(0, \sigma_\mu^2). \tag{15}$$

The shock ϵ_t^μ is normally distributed with mean zero and variance σ_μ^2. Firms can adjust prices in each period, hence markup shocks only redistribute between profits and the factor incomes.[10]

Final goods producers purchase a variety of goods from a continuous range of intermediate producers indexed by j. Production of intermediate goods occurs according to constant returns to scale Cobb-Douglas production technology which combines labor N_{jt} and capital services $u_{jt}K_{jt}$ taking into account capital utilization u_{jt} according to

$$Y_{jt} = A_t \left(u_{jt} K_{jt} \right)^\alpha \left(Z_t N_{jt} \right)^{1-\alpha}, \tag{16}$$

where α denotes the capital share in the Cobb-Douglas production function, A_t denotes

[10] Each differentiated good is offered at price p_{jt}, the aggregate price level is $P_t = \left(\int p_{jt}^{\frac{1}{1-\mu_t}} dj \right)^{1-\mu_t}$ and demand for each of the varieties is $y_{jt} = \left(\frac{p_{jt}}{P_t} \right)^{\frac{1-\mu_t}{\mu_t}} Y_t$. In a symmetric equilibrium, this boils down to $p_{jt} = P_t \, \forall j$ and $y_{jt} = Y_t \, \forall j$ and we do not need to keep track of prices hereafter.

aggregate productivity and Z_t denotes labor-augmenting technology. Firms can choose the intensity with which they use their capital stock K_{jt} by adjusting the capacity utilization u_{jt}. An intensity higher than normal results in increased depreciation of capital according to $\delta(u_{jt}) = \delta_{0t} + \delta_1(u_{jt} - 1) + \delta_2/2(u_{jt} - 1)^2$, which is an increasing an convex function of utilization if $\delta_1, \delta_2 > 0$.

The producer minimizes costs, $w_t N_{jt} - [r_t + q_t\delta(u_{jt})]K_{jt}$, where r_t and q_t are the rental rate and the (producer) price of capital goods and w_t is the real wage. Factor markets are perfectly competitive and all intermediate goods producers are symmetric. Therefore, we drop all indices j and only refer to the aggregate variables. We can characterize the first-order conditions for labor and effective capital as

$$r_t + q_t\delta(u_t) = \frac{\alpha}{\mu_t} A_t u_t \left(\frac{u_t K_t}{Z_t N_t}\right)^{\alpha-1} = \frac{\alpha}{\mu_t}\frac{Y_t}{K_t}, \tag{17}$$

$$\text{and} \quad w_t = \frac{1-\alpha}{\mu_t} A_t Z_t \left(\frac{u_t K_t}{Z_t N_t}\right)^{\alpha} = \frac{1-\alpha}{\mu_t}\frac{Y_t}{N_t}. \tag{18}$$

The optimal utilization choice is given by

$$q_t\left[\delta_1 + \delta_2(u_t - 1)\right] = \frac{\alpha}{\mu_t} A_t K_t \left(\frac{u_t K_t}{Z_t N_t}\right)^{\alpha-1} = \frac{\alpha}{\mu_t}\frac{Y_t}{u_t}. \tag{19}$$

As a result, aggregate profits are $\Pi_t = \mu_t Y_t$. The logarithm of productivities A_t and Z_t evolve stochastically according to AR(1) processes

$$\ln A_t = (1-\rho_A)\left(\ln A - \frac{\sigma_A^2}{2}\right) + \rho_A \ln A_{t-1} + \epsilon_t^A \quad \text{with} \quad \epsilon_t^A \sim N(0, \sigma_A^2), \tag{20}$$

$$\text{and} \quad \ln Z_t = (1-\rho_Z)\left(\ln Z - \frac{\sigma_Z^2}{2}\right) + \rho_Z \ln Z_{t-1} + \epsilon_t^Z \quad \text{with} \quad \epsilon_t^Z \sim N(0, \sigma_Z^2). \tag{21}$$

ρ_i and σ_i^2 with $i \in \{A, Z\}$ denote the autocorrelation of the log-technology shocks and the variance of their normally distributed innovations, while A and Z denote the unconditional means of the stochastic processes. Moreover, we allow time-varying depreciation rates δ_{0t}, which evolves according to

$$\delta_{0t} = \delta_0 + \epsilon_t^\delta \quad \text{with} \quad \epsilon_t^\delta \sim N(0, \sigma_\delta^2). \tag{22}$$

Finally, capital goods producers take the relative price of capital goods, q_t, as given when determining their output. They face capital adjustment costs as in Hayashi (1982) and maximize

$$\max_{\{I_t\}_{t=0}^{\infty}} \mathbb{E}_0 \sum_{t=0}^{\infty} \Lambda_{0,t} \left\{ q_t \left[I_t - \frac{\phi_t}{\kappa} \left(\frac{I_t}{K_t} - \delta_{0t} \right)^{\kappa} K_t \right] - I_t \right\}. \tag{23}$$

where $\kappa > 1$. To enhence the complexity of the mode, the adjustment costs feature a non-linear element. In particular, ϕ_t is state-dependent and depends on the level of aggregate capital:

$$\phi_t = \begin{cases} \overline{\phi} & \text{if } K_t > \overline{K} \\ \underline{\phi} & \text{if } K_t \geq \overline{K} \end{cases} \tag{24}$$

where $\overline{K} \geq \underline{K} \geq 0$. Note that capital good producers take the adjustment costs as given as they depend on aggregate capital.

Optimization yields the optimality condition

$$q_t = \left[1 - \phi_t \left(\frac{I_t}{K_t} - \delta_{0t} \right)^{\kappa - 1} \right]^{-1}. \tag{25}$$

Each capital goods producer will adjust its production, until (25) is satisfied. Since all capital goods producers are symmetric, we obtain a law of motion for aggregate capital

$$K_{t+1} = (1 - \delta(u_t))\, K_t + I_t - \frac{\phi_t}{\kappa} \left(\frac{I_t}{K_t} - \delta \right)^{\kappa} K_t. \tag{26}$$

Having specified the production sector, we now describe the households in the economy.

Household sector: There exists a unit continuum of (potentially heterogeneous) households indexed by $i \in [0, 1]$ which maximize their lifetime utility discounted by the factor β. The households obtain utility from consumption c_{it} and disutility from supplying labor n_{it}. To smooth consumption, households accumulate capital k_{it+1}. The household's objective function is

$$U_{it} = \max_{c_{it}, n_{it}, k_{it+1}} \mathbb{E}_t \sum_{t=0}^{\infty} \beta^t \zeta_t u(c_{it}, n_{it}), \tag{27}$$

with $\mathbb{E}_t$ denoting the expectation operator over all stochastic processes given the information set as of time t and $u(c_{it}, n_{it})$ denotes the per period felicity function of the household over consumption c_{it} and labor n_{it}. ζ_t is a stochastic aggregate shock to the discount factor in period t. The logarithm of the discount factor shock ζ_t evolves stochastically according to an AR(1) process

$$\ln \zeta_t = -(1-\rho_\zeta)\frac{\sigma_\zeta^2}{2} + \rho_\zeta \ln \zeta_{t-1} + \epsilon_t^\zeta \quad \text{with} \quad \epsilon_t^\zeta \sim N(0, \sigma_\zeta^2). \tag{28}$$

ρ_ζ denotes the autocorrelation of the logarithmic discount factor shock, and the shock ϵ_t^ζ is normally distributed with mean zero and variance σ_ζ^2. Households optimize the objective function (27) subject to the budget constraint

$$(1+\tau^C)c_{it} + q_t k_{it+1} = \left(q_t + (1-\tau^K)r_t\right) k_{it} + (1-\tau^L)w_t h_{it} n_{it} + T_t + \Pi_{it}. \tag{29}$$

r_t and w_t denote the interest rate and wage rate as specified above and h_{it} denotes households' idiosyncratic income component. τ^C, τ^K, and τ^L denote the value-added-tax, the capital income tax, and the labor income tax, while T_t denotes the transfers the households obtain from the government. Π_{it} denotes the individual part in aggregate profits. In all applications, households face a borrowing constraint, such that they are prohibited from holding negative amounts of assets. Individual productivity h_{it} evolves according to

$$\log h_{it} = -(1-\rho_h)\frac{\sigma_h^2}{2} + \rho_h \log h_{it-1} + \epsilon_{it}^h \quad \text{with} \quad \epsilon_{it}^h \sim N(0, \sigma_h^2). \tag{30}$$

with ϵ_{it}^h as a normally distributed shock with variance σ_ϵ^2 and mean zero.

The solution of the household problem can be characterized by the Euler equation on capital and the optimal labor supply schedule below

$$q_t u_C(c_{it}, n_{it}) = \beta \mathbb{E}_t \left[\frac{\zeta_{t+1}}{\zeta_t}\left(q_{t+1} + (1-\tau^K)r_{t+1}\right) u_C(c_{it+1}, n_{it+1})\right] \tag{31}$$

$$-u_L(c_{it}, n_{it}) = (1-\tau^L)w_t h_{it}\frac{u_C(c_{it}, n_{it})}{1+\tau^C}. \tag{32}$$

$u_C(c_{it}, n_{it}) = \frac{\partial u}{\partial c_{it}}(c_{it}, n_{it})$ denotes the partial derivative of the felicity function with respect to consumption and $u_L(c_{it}, n_{it}) = \frac{\partial u}{\partial n_{it}}(c_{it}, n_{it})$ denotes the partial derivative of the felicity function with respect to labor.

Government sector: The government levies distortionary capital and labor income taxation at flat rates τ^L and τ^K, and claims a value-added-tax τ^C on consumption. It uses the tax revenues to finance lump-sum transfers T_t to the household. Therefore, the role of the government is purely to redistribute between factor incomes and consumption and leisure. The budget constraint is

$$T_t = \tau^C C_t + \tau^K r_t K_t + \tau^L w_t N_t. \tag{33}$$

Government transfers T_t adjust residually to make the government budget constraint hold.

Market clearing and equilibrium: The labor market, the capital market, and the goods market have to clear at all periods. Labor and capital market clearing requires

$$N_t = \int_0^1 n_{it} di \quad \text{and} \quad K_t = \int_0^1 k_{it} di. \tag{34}$$

Given these aggregate quantities, prices are determined by their marginal products on the factor inputs as denoted in equations (17) and (18). The goods market clears when

$$Y_t = C_t + I_t, \tag{35}$$

where $I_t = K_{t+1} - (1 - \delta(u_t))K_t + \frac{\phi}{\kappa}\left(\frac{I_t}{K_t} - \delta_0\right)^{\kappa} K_t$ denotes aggregate investment into next periods capital stock net of adjustment costs and $C_t = \int_0^1 c_{it} di$ is aggregate consumption. The goods market clears due to Walras-Law whenever the capital and the labor market clear.

Dynamic equilibrium: We define a dynamic equilibrium in this economy as follows. Firms and households take prices as given. Households behave optimally to maximize their lifetime utility (27) subject to the associated budget constraint (29) and the stochastic processes. Firms choose their factor inputs to maximize profits given their Cobb-Douglas production technology until the optimality conditions (17), (18), and (19). Lump sum taxes adjust such that the government budget constraint (33) holds, while the labor and asset markets (34), and the goods market (35) clear.

3.2 Generative Economic Modeling

We are now using our generative economic modeling approach to solve the different variants of the real business cycle model. For each model version, we execute the following steps: 1. Solve the set of satellite models, 2. simulate data from the satellite models, 3. train a surrogate neural network on the set of simulated data from the satellite models, 4. solve the true full model version, 5. simulate data from the true full model, 6. train a neural network on the simulated data from the true model, 7. evaluate the performance of the surrogate model on the data simulated from the full model, 8. compare the approximation of the true data generation process through the surrogate model with the neural network.

3.2.1 Generative Economic Modeling with an Analytical Solution

To begin with, we illustrate our methodology using a stochastic representative agent version of the model, which admits an analytical solution following the approach of Brock and Mirman (1972). To solve the model analytically, we abstract from some features. The following proposition summarizes the assumptions to obtain an analytical solution to the model.

Proposition 1. *If all households are ex-ante and ex-post identical, depreciation is deterministic and full, $\delta(u_t) = 1$, capacity utilization is fixed at $u_t = 1$, there are no capital adjustment costs $\phi = 0$, the discount factor shock is inactive $\sigma_\zeta^2 = 0$, and per period felicity is of King, Plosser and Rebelo (1988) (KPR)-form given by $u(C_t, N_t) = \ln C_t - \omega \frac{N_t^{1+\gamma}}{1+\gamma}$. Then the policy functions of the representative household are*

$$N_t(\mu_t) = \left[\frac{\mu(1-\tau^L)(1-\alpha)}{\mu_t \omega (1+\tau^C)(\mu - (1-\tau^K)\alpha\beta)}\right]^{\frac{1}{1+\gamma}}, \tag{36}$$

$$C_t(K_t, A_t, Z_t, \mu_t) = \left(1 - \frac{\alpha\beta}{\mu_t}\right) Y_t\left(K_t, A_t, Z_t, N_t(\mu_t)\right), \tag{37}$$

$$\text{and} \quad K_{t+1}(K_t, A_t, Z_t, \mu_t) = \frac{\alpha\beta}{\mu_t} Y_t\left(K_t, A_t, Z_t, N_t(\mu_t)\right), \tag{38}$$

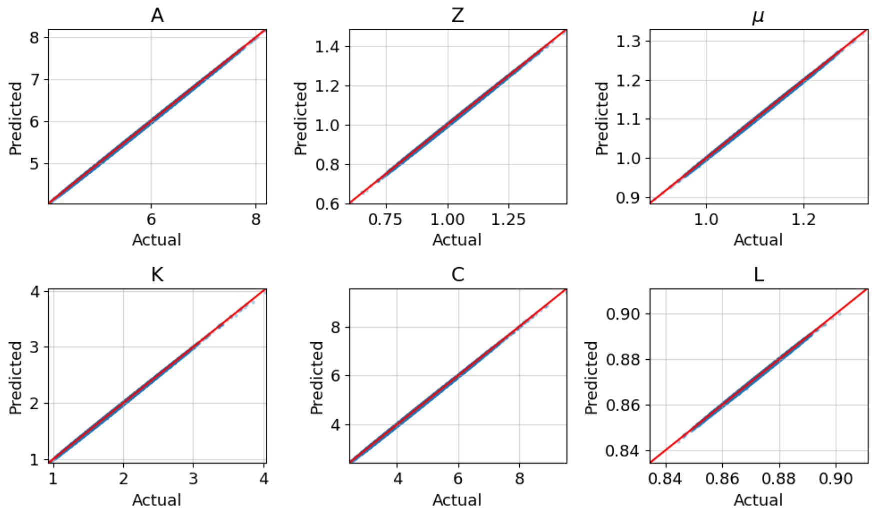
where $Y_t(K_t, A_t, Z_t, N_t(\mu_t))$ denotes the Cobb-Douglas production function. Given the policy functions of the household, the prices in the economy can be determined by equations (17) and (18). The transfers to the households are then determined by the government budget constraint (33).

Proof. See Appendix II. □

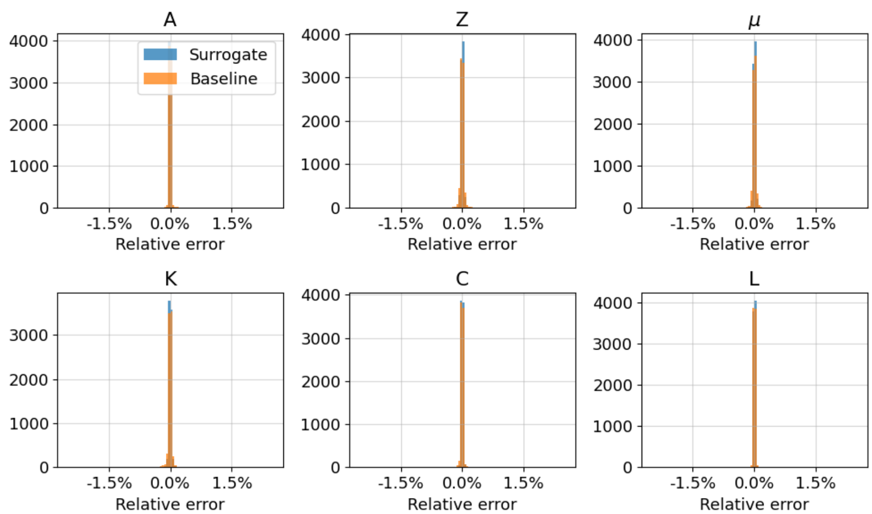
Proposition 1 presents the solution for the representative agent economy. Fluctuations in the markup μ_t drive changes in labor supply N_t, while shocks to productivity (A_t, Z_t) affect output Y_t. Consumption C_t and capital investment K_{t+1} are linear functions of output. As noted by Brock and Mirman (1972), this model is highly simplistic and omits key real-world dynamics. Despite these limitations, we use it as a benchmark, providing an exact reference to evaluate our methodology. We calibrate the model to conventional values, which are summarized in Appendix II. We use the analytical solution of proposition 1 to simulate long time series of four economies: We simulate three satellite economies with (I) TFP and labor-augmenting productivity shocks, (ii) TFP and markup shocks, (iii) labor-augmenting productivity and markup shocks. Finally, we simulate a time series of the full model (iv) with all three shocks at the same time. We train two neural networks, one on

Figure 3 Evaluation of the surrogate model with an analytical RBC model

(a) Prediction of surrogate against data



(b) Error distribution of surrogate model



the simulated data from the true model and one on the combined dataset of our satellite models. The neural networks consist of three hidden layers, each with 128 neurons, using the CELU activation function. The optimizer employed is AdamW, and training minimizes the mean squared error between the predicted and true values. The learning rate follows a cosine annealing schedule, starting at 10^{-3} and decaying to 10^{-7}. [color = red!40]Include the graphics containing the errors in the approximation

Figure 3 illustrates the performance of our methodology. Panel 3a) compares the surrogate model's predictions to the actual data generated by the full model, demonstrating the accuracy with which our approach captures the true dynamics. Panel 3b) presents the errors of the surrogate model relative to the true data generation process, alongside the errors of the deep neural network trained directly on that process. This comparison allows us to assess the performance loss incurred by restricting our neural network to training only on satellite model data.

When evaluating our generative economic modeling algorithm against the true data in Panel 3a), we find that the methodology effectively captures the dynamics of the exogenous stochastic processes A_t, Z_t, and μ_t. The surrogate model performs well overall, with minor imprecision occurring only in the case of large and infrequent shocks, where it tends to underpredict values. For the endogenous variables K_t, C_t, and L_t, the model maintains strong performance, albeit with slightly lower accuracy than for the exogenous processes. Specifically, it underpredicts capital K_t and consumption C_t at the upper end of the data range while slightly overpredicting labor supply L_t. This behavior likely stems from the limited number of observations in these regions, requiring the neural network to extrapolate. The fit can be further improved by expanding the dataset and optimizing the number of training epochs, enhancing the model's ability to capture extreme values more accurately.

We can assess the prediction errors of the neural network in relative terms in Panel 3b). The blue bins represent the relative errors of the surrogate model compared to the predicted variable values, while the orange bins show the corresponding errors of a deep neural network trained on the full dataset rather than on slices generated from satellite models. Since the surrogate model is trained on limited data, it is expected that its relative errors are larger than those of the fully trained deep neural network.

Both neural networks achieve high accuracy in predicting the exogenous processes A_t, Z_t, and μ_t, with relative errors well below one percent. The largest relative errors for the surrogate model occur in the prediction of capital K_t, though they remain under one percent. Notably, even the deep neural network trained on the full dataset struggles to predict capital stock with high precision, highlighting an inherent challenge in predicting this variable. In contrast, the surrogate model performs exceptionally well in predicting consumption C_t and labor supply L_t, achieving a level of accuracy comparable to the fully trained neural network.

While the surrogate model exhibits some errors, its overall performance is strong. It effectively captures the dynamics of the full model, with only minor difficulties in predicting the capital stock—a challenge that even the fully trained neural network faces. Importantly, all errors remain small in relative terms, such that the method of generative economic modeling provides a reliable and useful approximation of the full model's dynamics.

Besides visual inspection, table 1 illustrates the fit of generative economic modeling quantitatively. For that we use the Mean Squared Error (MSE) and the Mean Absolute

Table 1 Fit of the method for different model versions

Model	Mean Squared Error	Mean Absolute Error	Euler Eq. Error (%)
Analytical	2.88×10^{-6}	7.10×10^{-4}	-
Nonlinear	3.36×10^{-4}	6.57×10^{-3}	1.09
Krusell-Smith	9.50×10^{-8}	1.99×10^{-4}	0.76
HANK	-	-	3.23

NOTE - Mean Squared Error (MSE) and Mean Absolute Error (MSA) compare the predictions of the surrogate model against the true data. The Euler Equation Error (EEE) calculates the relative error in the Euler equation for the model. We cannot calculate the MSE or the MAE for the HANK model, as we do not know the true model solution. In the Krusell-Smith and the HANK model the measure is calculated as the weighted Euler equation error, where the weights are the measures in the distribution.

Error (MAE) of the predictions of the model approximated through our methodology compared to the true data generation process. Both, the MSE and the MAE are low considering that we use a standard specification of the neural network and do not fine tune the hyperparameters to the application at hand.

3.2.2 Generative Economic Modeling with a nonlinear model

In this section, we illustrate our methodology using a nonlinear medium-sized version of our RBC model. We abstract from shocks to labor-augmenting productivity Z_t and to the markup μ_t, hence shocks to TFP A_t are the only drivers of supply-side fluctuations. We assume that households remain ex-ante and ex-post identical, such that we can represent the household side with a representative agent. Moreover, households have KPR-preferences

$$u(C_t, N_t) = \frac{C_t^{1-\sigma} - 1}{1-\sigma} - \omega \frac{N_t^{1+\gamma}}{1+\gamma} \tag{39}$$

such that optimal household behavior can be described by the Euler equation and the optimal labor supply condition:

$$\begin{aligned} C_t^{-\sigma} &= \beta \mathbb{E}_t \left[(q_{t+1} + (1-\tau^K) r_{t+1}) C_{t+1}^{-\sigma} \right] \\ \omega N_t^{\gamma} &= \frac{1-\tau_t^L}{1+\tau^C} w_t C_t^{-\sigma}. \end{aligned}$$

Contrary to the former section, we allow for capacity utilization choice and allow for capital adjustment costs with a nonlinear specification as illustrated in equations (23) and (24). The model is solved with global methods, specifically policy function iteration, to account for all nonlinear features. Within the class of policy function iteration methods, we use time iteration with linear interpolation as in Richter, Throckmorton and Walker (2014) and Bianchi, Melosi and Rottner (2021). The parameter choices are conventional and chosen to ensure strong nonlinearities in the propagation of shocks, as outlined in Appendix II. Figure 8 illustrates the nonlinearities in response to shocks due to these features.

We generate time series data for three satellite models, each subject to two out of the three possible shocks. Hence, we solve and simulate satellite models with (I) TFP and discount factor shocks, (ii) TFP and depreciation shocks, and (iii) discount factor and depreciation shocks. Moreover, we simulate a fourth model in which all three shocks are active simultaneously. We train two neural networks, one on the simulated data from the true model and one on the combined dataset of our satellite models. The neural networks consist of five hidden layers, each with 128 neurons, using the CELU activation function. The optimizer employed is AdamW, and training minimizes the mean squared error between the predicted and true values. The learning rate follows a cosine annealing schedule, starting at 10^{-3} and decaying to 10^{-10}. [color = red!40]Include the graphics containing the errors in the approximation

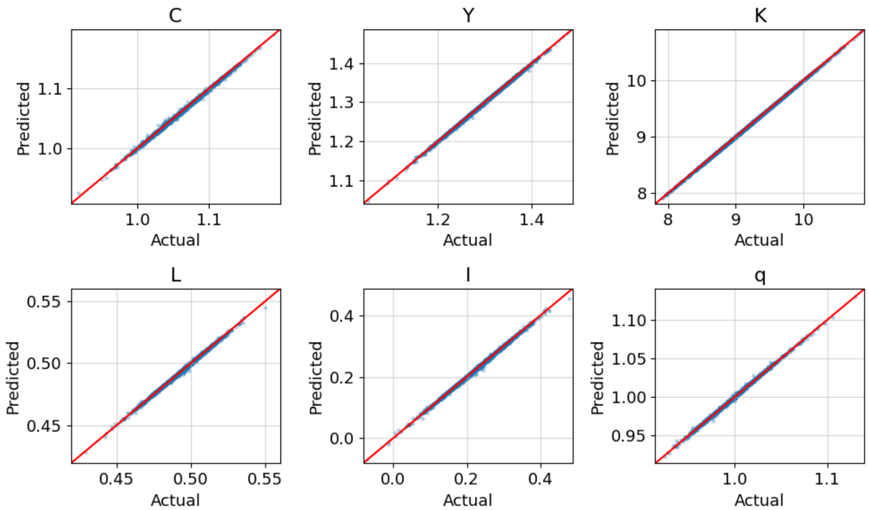
Figure 4 illustrates the performance of our algorithm for the medium-sized nonlinear DSGE model. Panel 4a) illustrates the fit of the surrogate model against the true data, while panel 4b) illustrates the relative errors of the surrogate model and of a deep neural network that is trained on the dataset generated by the full model.

Comparing the surrogate model's predictions with the true data in panel 4a) demonstrates that our methodology provides a strong fit, even for a medium-sized DSGE model with nonlinearities. The predicted values closely align with the true values, addressing the issue of under- and overprediction observed in the analytical version from the previous section. The model performs well in predicting exogenous processes (not illustrated) and provides a reasonable fit for endogenous variables. Among these, capital K_t exhibits the best fit, while investment I_t shows the largest deviations due to its high volatility, driven by nonlinear adjustment costs. Notably, the model's ability to predict the price of capital q_t remains unaffected by these nonlinearities.

The relative prediction errors in panel 4b) indicate that the surrogate model effectively captures the true dynamics of the approximated model. For all endogenous variables except investment I_t, relative errors remain well below one percent and are comparable to those generated by the deep neural network trained on the full dataset. In contrast, the relative errors for I_t reach up to 2%, highlighting a weaker performance in predicting investment. Furthermore, the deep neural network trained on the full dataset significantly outperforms the surrogate model. A possible explanation for this discrepancy is the nonlinear dynamics underlying investment: certain shocks have minimal impact on I_t, limiting the information available in the satellite datasets. As a result, the surrogate model extracts less relevant information during training, reducing its predictive accuracy for investment.

Figure 4 Evaluation of the surrogate model with a nonlinear medium-sized RBC model

(a) Prediction of surrogate against data



(b) Error distribution of surrogate model

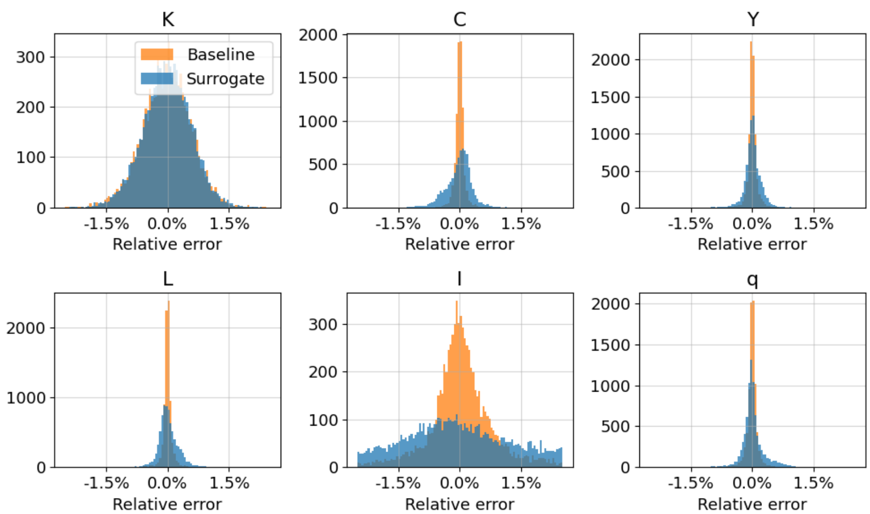


Table 1 illustrates quantitatively that our approximation performs reasonable well, but that the nonlinear dynamics reduce our accuracy compared to the analytical model as both MSE and MAE increase. For the nonlinear model we also compute the Euler equation Error (EEE).[11] With 1.09% the Euler equation error is large, however not exessive given that the economy features strong nonlinearities.

Overall, the surrogate model effectively captures the dynamics of the underlying model

[11] To calculate the EEE, we use the predictions of the surrogate model for consumption and capital today together with exogenous processes for the shocks. We then calculate the expectations on the right-hand-side of the Euler equation by Monte Carlo simulation.

and even improves its forecasting performance for endogenous variables compared to the previous section. Although forecast errors are higher for variables exhibiting nonlinearities, such as investment I_t, the relative errors remain small.

3.2.3 Generative Economic Modeling with Heterogeneous Agents

This section applies our methodology to a model with heterogeneous households. In contrast to earlier sections, we now incorporate both ex-ante and ex-post heterogeneity, following the framework of Krusell and Smith (1998). As a result, the joint distribution of wealth and income becomes a state variable of the model.

To simplify the solution of the model, we abstract from endogenous capacity utilization ($u_t = 1$) and set capital adjustment costs to zero ($\phi_t = \kappa = 0$). Households maximize a standard CRRA utility function:

$$u(c_{it}) = \frac{c_{it}^{1-\sigma} - 1}{1 - \sigma}$$

subject to the budget constraint in equation (29). Since labor supply entails no disutility, households supply one unit of labor inelastically.

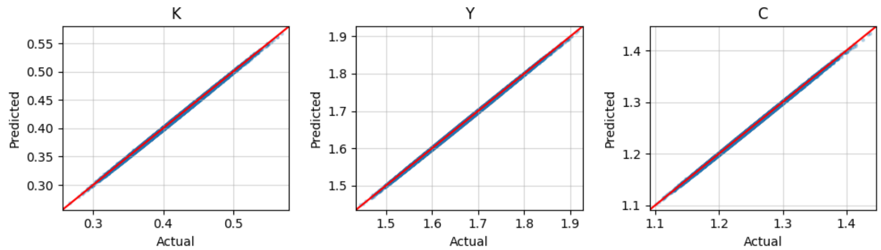
We shut down government activity and keep three aggregate shocks: productivity (A_t), discount factor (ζ_t), and depreciation (δ_t). To solve the model, we discretize both aggregate and idiosyncratic state processes using the method of Tauchen (1986). Each aggregate shock is approximated by a four-state Markov chain. Similarly, idiosyncratic income shocks h_{it} are modeled using a four-state Markov chain. Unlike Krusell and Smith (1997) and Krusell and Smith (1998), we abstract from any dependence of idiosyncratic income risk on aggregate TFP.[12]

A detailed description of the solution algorithm for the heterogeneous agent model with multiple aggregate shocks is provided in Appendix II. Here, we briefly summarize the approach. We solve the household problem using the endogenous grid point method of Carroll (2006), and simulate the economy using the non-stochastic simulation method from Young (2010). Since the model includes aggregate shocks, households require a perceived aggregate law of motion (ALM) for aggregate capital. Following Krusell and Smith (1998), we assume a state-dependent log-linear ALM and update it iteratively until convergence, ensuring that the gap between the true and perceived laws is minimal, in line with the concerns raised by Den Haan (2010). Figure 9 (in the appendix) shows that

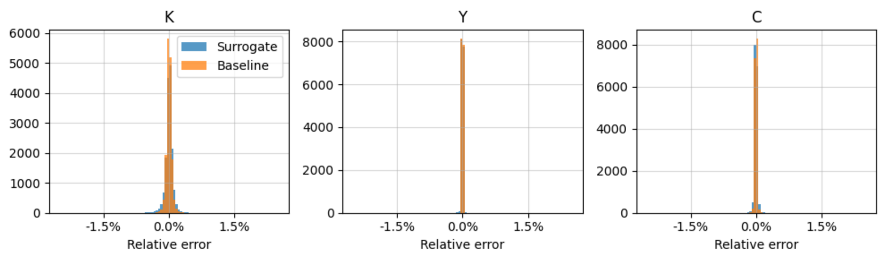
[12] Although we can allow for a correlation between idiosyncratic and aggregate risk, with more than two aggregate states there exist various possibilities to allow for cyclical variation in idiosyncratic income risk. We leave the exploration of the impact of cyclical idiosyncratic income risk on the performance of our method for future research.

Figure 5 Evaluation of the surrogate model with a heterogeneous agent model

(a) Prediction of surrogate against data



(b) Error distribution of surrogate model



the model-generated capital series closely match the ALM, with a maximum deviation of less than 0.01. We solve the model using four idiosyncratic income states, four aggregate states per shock, 100 asset grid points (with an exponential spacing to capture the lower wealth range), and 20 grid points for aggregate capital. The full model therefore spans $4^4 \times 100 \times 20 = 512{,}000$ state variables.

For our generative economic modeling approach, we solve three distinct satellite models, each featuring two out of the three aggregate shocks. Specifically, we solve satellite models with (i) TFP and discount factor shocks, (ii) TFP and depreciation shocks, and (iii) discount factor and depreciation shocks. Finally, we solve the full model featuring all three aggregate shocks. We use the policy functions of the solved economies to simulate a long time series of each economy. We train two neural networks, one on the simulated data from the true model and one on the combined dataset of our satellite models. The neural networks consist of five hidden layers, each with 128 neurons, using the CELU activation function. The optimizer employed is AdamW, and training minimizes the mean squared error between the predicted and true values. The learning rate follows a cosine annealing schedule, starting at 10^{-3} and decaying to 10^{-10}. [color = red!40]Include the graphics containing the errors in the approximation

Figure 5 evaluates the performance of our algorithm for the heterogeneous agent model. Panel 5a) shows the fit of the surrogate model relative to the true simulated data, while Panel 5b) compares the relative errors of the surrogate model and a deep neural network trained on the full model's dataset. Panel 5a) demonstrates that the surrogate model provides an excellent fit in the heterogeneous agent setting. The predicted values closely track the true data, and the model performs well in capturing the dynamics of the endogenous variables. Panel 5b) shows that the relative prediction errors of the surrogate model remain consistently below one percent across all endogenous variables, comparable to the performance of the deep neural network trained on the full dataset. While the neural network marginally outperforms the surrogate model, this difference may stem from the increased complexity of the heterogeneous agent environment—specifically, the interaction of multiple shocks affecting households near the borrowing constraint. These nonlinearities can reduce the amount of relevant structure the surrogate model can extract during training, particularly for variables such as investment.

Despite this, the surrogate model captures the core dynamics of the economy remarkably well. Compared to the results in the representative agent case, it even shows improved predictive performance for some endogenous variables, while maintaining low overall error levels. To calculate the Euler Equation Error of the model, we also applied our methodology to predict the policy functions, as well as the dynamics of the distribution of households at each individual point of the discretized individual state space point. In both applications, our method works well and approximates the true dynamics almost exactly using generative economic modeling from the dynamics of the satelite models. Table 1 supplements the graphical analysis and shows that in terms of MSE, MAE, and EEE the model approximation is satisfying. The average Euler equation error of 0.76% indicates that over the entire distribution of households the error in the Euler equation is below one percent.

4 Global HANK model with financial frictions

In this section, we use our method of generative economic modeling to globally solve a heterogeneous agent New Keynesian model with financial frictions. It is challenging to solve the model for three reasons. First, the solution requires forecasting the forward-looking part of the Phillips-curve using a perceived law of motion, hence facing the same issue as our heterogeneous agent model in section 3.2.3. Second, in the simulation stage, we need to find market-clearing levels of today's inflation as input for our estimation of the perceived law of motion. This requires a root-finding step for each simulation period, which is numerically costly. Finally, introducing the cash-in-advance constraint introduces asymmetric responses to the economy when financial constraint. When introducing a large number of aggregate shocks to the economy, all numerical issues are multiplied. Consequently, it provides a natural application for our methodology.

4.1 Description of the Model

The model is a HANK model that includes both idiosyncratic and aggregate risk. Households insure against both types of risks by saving in liquid assets subject to a borrowing limit. Intermediate goods are produced using labor under monopolistic competition, where firms face Rotemberg price adjustment costs and a cash-in-advance constraint limiting production in times of financial distress. A final goods bundler bundles intermediate goods into a final good. The government raises taxes to issue bonds and for government consumption, while the central bank sets the nominal interest rate as a function of price inflation. The model features shocks to the discount factor, shocks to aggregate productivity, monetary policy shocks, as well as shocks to the ability of firms to borrow through the financial sector.

Households There exists a continuum $i \in [0,1]$ of households which choose to obtain utility from consumption c_{it} and leisure, and save in liquid assets b_{it+1} such as to insure against idiosyncratic income fluctuations in labor productivity h_{it}. Labor productivity follows an AR(1) process in logs as in equation (30). Households maximize the following utility:

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \exp(\zeta_t) u(c_{it}, n_{it}), \tag{40}$$

where n_{it} denotes their labor supply and ζ_t denotes the shock to the discount factor. We assume that the discount factor shock follows an AR(1) process as in equation (28). We assume that household felicity is of Greenwood, Hercowitz and Huffman (1988) (GHH) form[13] together with a CRRA specification:

$$u(c_{it}, n_{it}) = \frac{\left(c_{it} - \omega h_{it} \frac{n_{it}^{1+\gamma}}{1+\gamma}\right)^{1-\sigma} - 1}{1-\sigma} \tag{41}$$

ω is a scalar for multiplying the disutility of supplying labor. Households maximize utility subject to the budget and borrowing constraint

$$c_{it} + b_{it+1} = (1-\tau_t) w_t n_{it} h_{it} + R_t b_{it} + (1-\tau_t)\Pi_{it}, \tag{42}$$

$$b_{it+1} \geq \bar{B} \tag{43}$$

[13] We follow Bayer, Born and Luetticke (2022) with this utility specification for its numerical advantages. It allows us to compute labor supply only as a function of the wage rate. Bayer, Born and Luetticke (2024) provides a detailed description in favor of the use of GHH preferences contrary to King, Plosser and Rebelo (1988) preferences.

where b_{it+1} denotes savings of the household, τ_t the income tax, $R_t = 1 + r_t$ the real interest rate, and Π_{it} profits from owning the firm sector. We distribute profits proportional to the idiosyncratic productivity h_{it}. $\bar{B}$ denotes the exogenous borrowing limit of households.

Firms A final goods producer bundles a continuum of differentiated varieties $j \in [0, 1]$ into a final good according to a Dixit-Stiglitz aggregator

$$Y_t = \left(\int_0^1 y_{jt}^{\frac{\eta-1}{\eta}} dj \right)^{\frac{\eta}{\eta-1}}, \tag{44}$$

with elasticity of substitution η. This yields an optimal demand for each variety j of

$$y_{jt} = \left(\frac{p_{jt}}{P_t} \right)^{-\eta} Y_t, \tag{45}$$

where $P_t = (\int_0^1 p_{jt}^{1-\eta} dj)^{\frac{1}{1-\eta}}$ denotes the price level. Each differentiated variety is produced by an intermediate goods producer with index j using labor as input. Production follows the linear production function

$$Y_{jt} = A_t N_{jt}, \tag{46}$$

where A_t denotes aggregate productivity that follows an AR(1) process in logs as in equation (20). Intermediate goods producers are subject to quadratic price adjustment costs in logarithmic price changes. Hence, for price-setting, the firm maximizes

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t Y_t \left\{ \left(\frac{p_{jt}}{P_t} - MC_t \right) \left(\frac{p_{jt}}{P_t} \right)^{-\eta} - \frac{\eta}{2\kappa} \left(\log \frac{p_{jt}}{p_{jt-1}} \right)^2 \right\}, \tag{47}$$

with a time-constant discount factor.[14] The producer's first-order condition gives rise to a New Keynesian Phillips curve in goods price inflation

$$\log(\pi_t) = \beta \mathbb{E}_t \left[\log(\pi_{t+1}) \frac{Y_{t+1}}{Y_t} \right] + \kappa \left(MC_t - \frac{\eta - 1}{1} \right), \tag{48}$$

where Π_t is the gross inflation rate $\Pi_t \equiv \frac{P_t}{P_{t-1}}$, and MC_t is the real marginal costs. The price adjustment then creates real costs $\frac{\eta}{2\kappa} Y_t \log(\Pi_t)^2$.

Finally, intermediate goods producers are subject to a financing constraint when paying their labor bill. Firms need to borrow their wage bill from a perfectly competitive financial intermediary at a zero intratemporal interest rate; however, due to agency costs are not

[14] There exist multiple possible discount factors for the price-setting problem. We use the standard constant discount factor from the DSGE literature based on the discussion of Bayer et al. (2019).

able to do so up to the full level of their revenue. Hence firms face the following borrowing constraint:

$$w_t N_{jt} \leq \lambda_t y_{jt}, \tag{49}$$

where λ_t denotes the fraction of output that firms are allowed to borrow. We assume λ_t to follow an AR(1) process in logs:

$$\ln \lambda_t = -(1-\rho_\lambda)\frac{\sigma_\lambda^2}{2} + \rho_\lambda \ln \lambda_{t-1} + \epsilon_t^\lambda \quad \text{with} \quad \epsilon_t^\lambda \sim N(0, \sigma_\lambda^2). \tag{50}$$

This implies that if $\lambda_t < MC_t$, the household is constrained in its labor bill and firms can only demand labor up to the wage rate $w_t = \lambda_t A_t$. Hence, if firms are financially constraint, they cannot produce up to their capacity, because they are limited in the wages they can pay. This introduces a nonlinearity in the economy, which makes the solution of the model numerically more demanding.

Government The government operates a monetary and a fiscal authority. The monetary authority controls the nominal interest rate on liquid assets, while the fiscal authority issues government bonds to finance deficits and adjusts expenditures to stabilize debt in the long run and output in the short run.

We assume that monetary policy sets the nominal interest rate i_t on bonds following a Taylor-type rule:

$$(1 + i_{t+1}) = \Pi_t^{\phi_\pi} \exp(\iota_t), \tag{51}$$

where ϕ_π governs the extend to which the central bank attempts to stabilize inflation. The larger ϕ_π, the stronger the reaction of the central bank to changes in the inflation rate. ι_t is an exogenous monetary policy shock that follows an AR(1) process in logs:

$$\ln \iota_t = -(1-\rho_\iota)\frac{\sigma_\iota^2}{2} + \rho_\iota \ln \iota_{t-1} + \epsilon_t^\iota \quad \text{with} \quad \epsilon_t^\iota \sim N(0, \sigma_\iota^2) \tag{52}$$

The real interest rate is then determined using a Fisher relation $R_t = 1 + r_t = \frac{1+i_t}{\Pi_t}$

Moreover, we assume that the government issues bonds according to the rule

$$\frac{B_{t+1}}{\bar{B}} = \left(\frac{R_t B_t}{\bar{R}\bar{B}}\right)^{\rho_B} \left(\frac{\Pi_t}{\bar{\Pi}}\right)^{-\gamma_\pi} \left(\frac{\mathcal{T}_t}{\bar{\mathcal{T}}}\right)^{-\gamma_\mathcal{T}}, \tag{53}$$

using tax revenues, $\mathcal{T}_t = \tau Y_t$, to finance government consumption, G_t, and interest on outstanding debt. The coefficients $\bar{B}$, $\bar{\Pi}$, and $\bar{\mathcal{T}}$ are normalization constants. ρ_B captures whether and how fast the government seeks to repay its outstanding obligations, $B_t R_t$.

For $\rho_B < 1$, the government actively stabilizes real government debt, and for $\rho_B = 1$, the government rolls over all outstanding debt, including interest. The coefficients γ_π band $\gamma_\mathcal{T}$ capture the cyclicality of debt issuance: for $\gamma_\pi = \gamma_\mathcal{T} = 0$, new debt does not actively react to tax revenues and inflation, but only to the value of outstanding debt; for $\gamma_\pi > 0 > \gamma_\mathcal{T}$, debt is countercyclical; for $\gamma_\pi < 0 < \gamma_\mathcal{T}$, debt is procyclical. We assume that government expenditure G_t adjusts such that the government budget constraint is satisfied $G_t + R_t B_t = B_{t+1} + \mathcal{T}_t$

Market clearing Market clearing requires that the labor market, the bond market, as well as the goods market, clear. GHH preferences imply that households supply labor according to $n_{it} = (\frac{(1-\tau)w_t}{\omega})^{\frac{1}{\gamma}} = N_t$ where the last equality follows from $\int_0^1 h_{it} di = 1$. Hence, labor market clearing is achieved if

$$N_t = \begin{cases} \left(\frac{(1-\tau)A_t MC_t}{\omega}\right)^{\frac{1}{\gamma}} & \text{if unconstrained} \\ \left(\frac{(1-\tau)A_t \lambda_t}{\omega}\right)^{\frac{1}{\gamma}} & \text{if constrained} \end{cases}. \tag{54}$$

Bonds market clearing is achieved if

$$B_{t+1} = \int_0^1 b_{it+1} di, \tag{55}$$

and goods market clearing is achieved if

$$(1 - \frac{\eta}{2\kappa}(\ln \Pi_t)^2) Y_t = C_t + G_t, \tag{56}$$

where the left-hand side indicates production adjusted for price-adjustment costs.

Computational challenges The problem faces three computational challenges. First, to solve the model, we need to employ the algorithm of Krusell and Smith (1997) to forecast the forward-looking part $\mathbb{E}_t\left[\log(\pi_{t+1})\frac{Y_{t+1}}{Y_t}\right]$ in the Philips curve, as well as today's inflation Π_t using a perceived law-of-motion. This requires solving and simulating the model multiple times to update the law of motion until convergence. This issue is identical to the computation issue illustrated in section 3.2.3. Second, we need to calculate the market-clearing inflation rate in the simulation step to update the prediction of the nowcast of inflation, which is necessary for the solution of the household problem. Hence, after solving the household problem globally, we need to introduce a root-finding step. Concretely, we guess an inflation rate $\tilde{\Pi}_t$ that then determines the nominal interest rate i_{t+1}, relevant for the savings choice of the households. Given the aggregate states $(B_t, \lambda_t, A_t, \zeta_t, \iota_t)$, we impose labor market clearing by using equation (54) and then update the guess for inflation $\tilde{\Pi}_t$ until bond market clearing is achieved. This additional step requires us to solve the household problem (although only for one backward iteration) multiple times. This adds additional computational time. Third, adding the cash-in-advance constraint implies that the economy features nonlinear dynamics if intermediate good firms are financially constrained. This requires additional runtime to solve for accurate perceived laws of motion.

4.2 Generative Economic Modeling Solution

The global solution of this model with conventional methods remains numerically intractable. Consequently, we solve the model using our methodology of generative economic modeling. For that, we generate time series data for three satellite models, each subject to two out of the four possible shocks. Hence, we solve and simulate satellite models with (i) financial, and TFP shocks, (ii) financial and discount factor shocks, and (iii) financial and monetary shocks. We also solve a model version without financial shocks. For the approximation of the true solution, we solve three satellite models, each including one of the three non-financial shocks. Hence we solve and simulate models with (i) TFP shocks, (ii) discount factor shocks, and (iii) monetary shocks. We use this solution of the model to understand the effect of introducing financial shocks into a HANK model. We train two neural networks, each on the combined datasets of our satellite models. The neural networks consist of five hidden layers, each with 128 neurons, using the CELU activation function. The optimizer employed is AdamW, and the training minimizes the mean squared error between the predicted and true values. The learning rate follows a cosine annealing schedule, starting at 10^{-3} and decaying to 10^{-10}. Table 1 presents the Euler equation error of 3.23% for our approximation.

The error is the highest for all our applications, as we only use satellite models with two shocks each for the approximation of a model with a total of four shocks. Consequently, we only have a limited overlap of the model features in the sense illustrated in figure 1. Moreover, the model features nonlinearities that increase the Euler equation error, especially as we take the average over all agents in the economy. Incorporating more shocks simultaneously in the satellite models surely improves the fit of the method in this case. We consider the Euler equation error not prohibitively high to inspect nonlinearities of the model, as well as the effect of multiple shocks being present in the model at the same time.

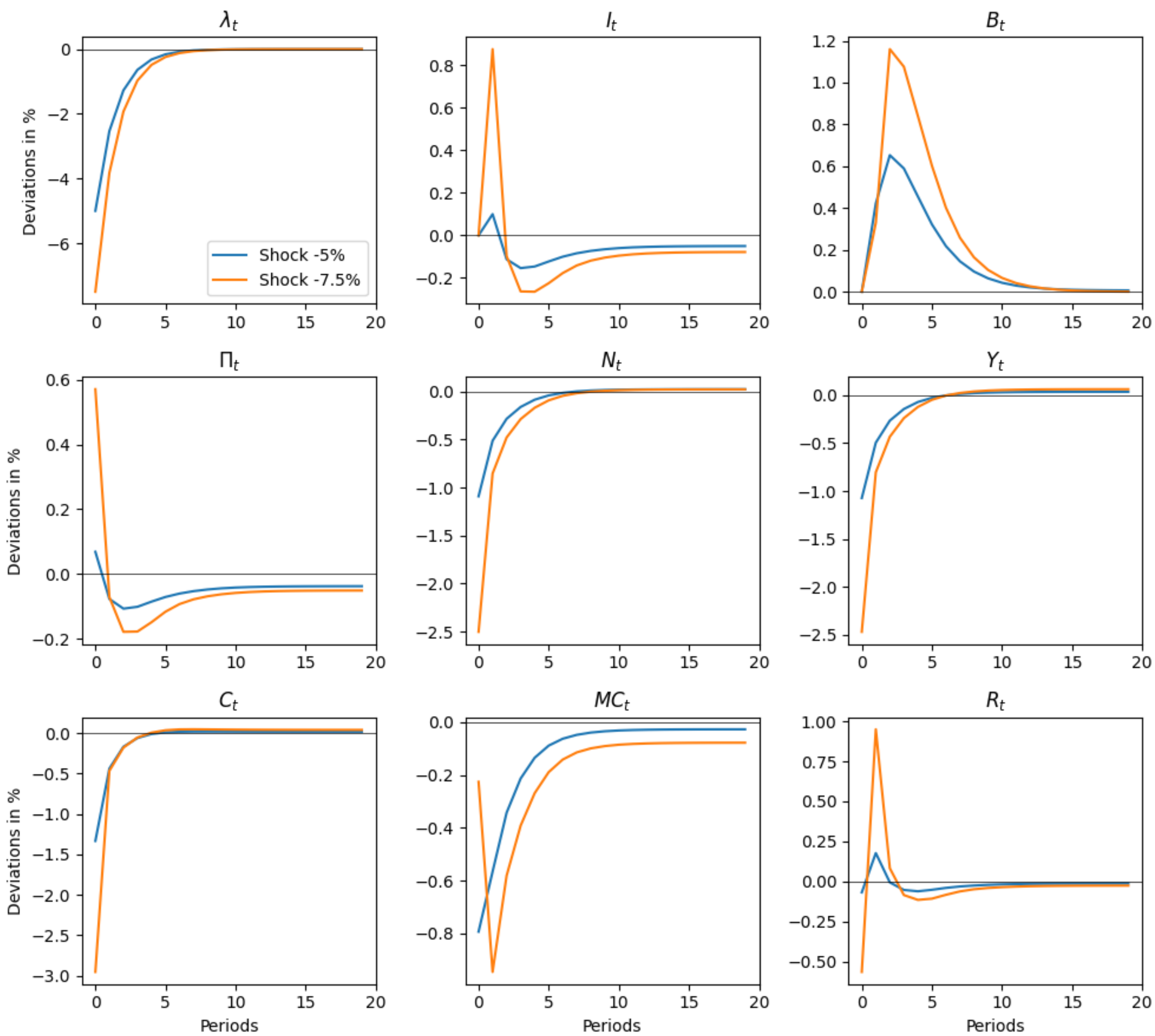


Figure 6 Generalized impulse response function of HANK economy with financial friction

4.2.1 Impact of financial friction shock

Having solved a HANK model with financial shocks, we are interested in the transmission of financial shocks in our model. To analyze the impact of a financial friction shock on the economy, we investigate the generalized impulse responses of the model to financial shocks of varying sizes. Concretely, we illustrate response of the economy to a shock in λ_t of -5% and -7.5% relative to its mean. Figure (6) illustrates the generalized impulse response function of the solved model to these shocks.

A drop in λ_t, the variable determining the space of the financial sector for intratemporal lending induces a recession for both shock sizes. As result of the drop, firms are constraint and hire less labor N_t, which reduces production Y_t. As result of this negative supply shock, consumption C_t drops, while inflation Π_t increases. The central bank increases the nominal interest rate I_t going forward in response to the hike in inflation, thereby increasing the real interest rate R_t. Bond supply B_t increases by the government, triggering countercyclical government expenditure.

Besides having these qualitative responses in common, the economy features nonlinearity in response to different sizes of shocks. The 7.5% decrease in lambda triggers a substantially larger recession than the 5% decrease in lambda. This manifests itself in a larger reduction in labor N_t, a larger decrease in output Y_t, and a larger decrease in consumption C_t. As the shock to the supply side is more severe, inflation Π_t increases more, triggering a larger increase in the nominal interest rate I_t. The shock is so contractionary that the marginal costs for production does even increases in response to the shock before dropping, indicating that in the first period the costs of production increase due to the financial friction.

To summarize, shocks to λ_t that serves as a financial shock here triggers nonlinear dynamics in response to different shock sizes. Such dynamics can be especially important when trying to understand financial crisis. Generative economic modeling serves as a useful tool to solve models with nonlinear dynamics, while keeping it numerically tractable.

4.2.2 Implications of more aggregate shocks

Second, we are interested in the interaction of different sources of risk with each other. For that, we compare the dynamics of the economy with four shocks with an economy that only features the financial shock and a TFP shock. Through this analysis, we aim to understand how the presence of more shocks in the model shapes the response of the model. Figure (7) illustrates impulse responses of the economy with four shocks to a TFP shock compared to the impulse response of the satelite economy with only the TFP shock and the lambda shock.

The comparison between the two impulse responses shows how incorporating more shocks in a model alters the dynamics of the model. For all variables illustrated, the responses of the variables after the financial shock is attenuated in the model with all four shocks compared to only featuring a TFP shock. Hence, for the model environment we solve here, introducing more shocks beyond the financial shock and the TFP shock reduces the response of endogenous variables to the financial shock. One economic explanation for this is that in the presence of more aggregate shocks, households have a larger precautionary incentive. Hence, knowing that they will face larger aggregate volatility with more shocks, they insure themselves better through precautionary savings, for example. In response to one of these shocks hitting, households are then better insured.

In the context of our model here, integrating more shocks dampens the response of

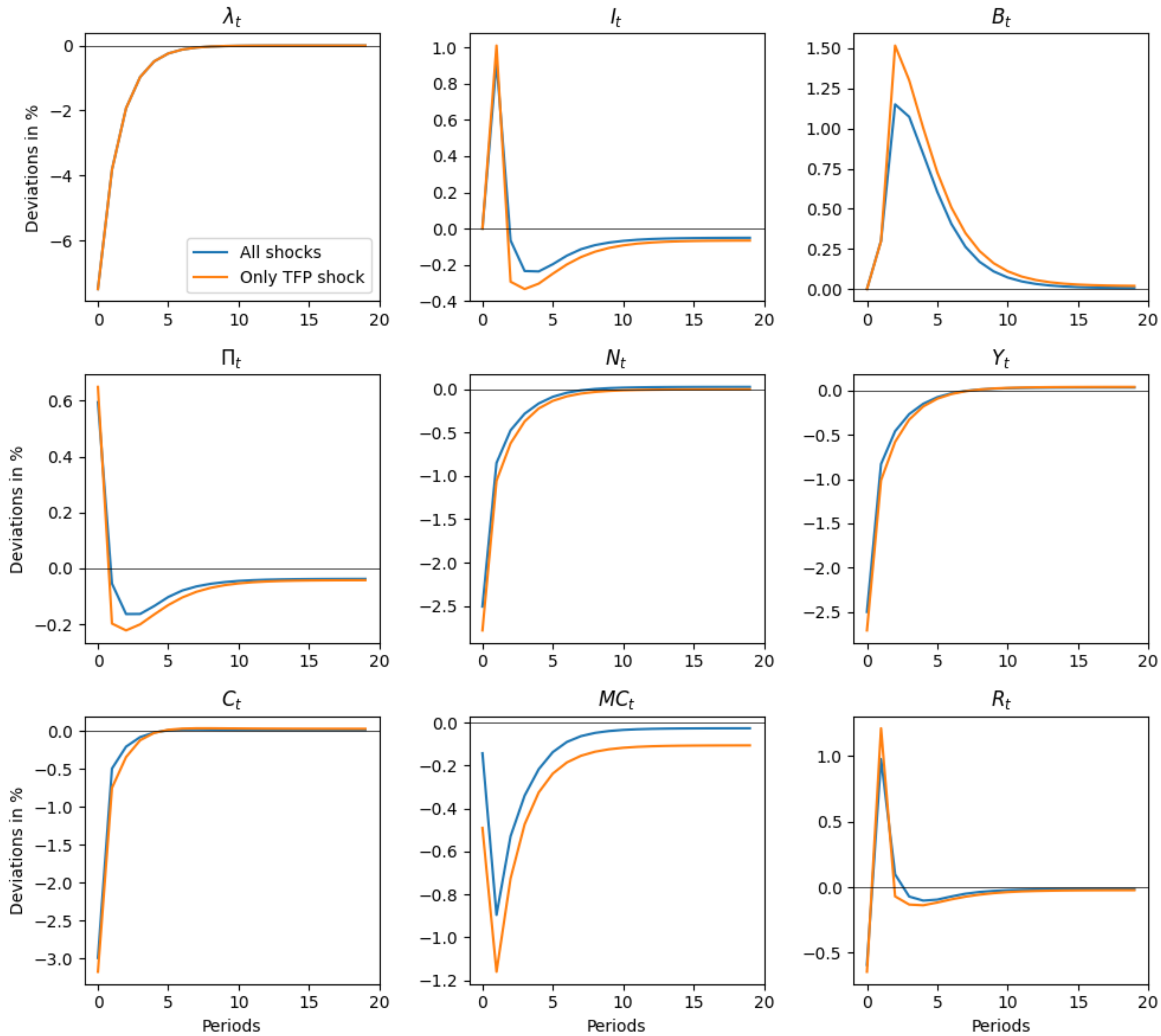


Figure 7 Comparison of generalized impulse responses to a lambda shock for model with two and four shocks

the economy to the nonlinear financial shock. The methodology of generative economic modeling allows users to integrate more shocks to study the implications that integrating more shocks has for their model.

5 Conclusion

Our study introduces generative economic modeling, a novel approach that combines conventional solution methods with artificial intelligence to overcome computational barriers in solving complex dynamic economic models. By using neural networks trained on data generated from satellite models, we provide an alternative to standard deep learning-based approaches, which often require extensive fine-tuning and can suffer from instability due to their endogenous feedback loops. In contrast, our methodology ensures stability and scalability by using precomputed solutions from conventional methods, allowing for efficient training and accurate approximations of the full economic model.

The results demonstrate that this approach successfully captures model dynamics with high precision, yielding prediction errors comparable to those of deep neural networks trained on full model data. Importantly, our method extends the applicability of conventional global solution methods by using recent advances in artificial intelligence. This is particularly valuable for models featuring higher-order nonlinearities and heterogeneous agents, where the curse of dimensionality poses significant computational challenges.

Our general approach offers several promising avenues in the future. First, it can be applied to more complex environments, such as solving nonlinear HANK models by training on simplified RANK and linearized HANK models. Second, it has the potential to enhance model estimation techniques, where fast and reliable solutions are critical. Lastly, we are developing evaluation metrics to assess neural network performance systematically, which will guide the optimal design of satellite models and network specifications.

References

Adenbaum, J., Babalievsky, F. and Jungerman, W. (2024). Learning on the Job. Working Paper.

Ahn, S., Kaplan, G., Moll, B., Winberry, T. and Wolf, C. (2017). When Inequality Matters for Macro and Macro Matters for Inequality. Tech. rep.

Ashwin, J., Beaudry, P. and Ellison, M. (2025). Neural network learning for nonlinear economies. *Journal of Monetary Economics* **149**, 103723.

Auclert, A., Rognlie, M. and Straub, L. (2020). Micro Jumps, Macro Humps: Monetary Policy and Business Cycles in an Estimated HANK Model. Tech. rep.

Auclert, A., Rognlie, M., Straub, L. and Tapak, T. (2024). When do Endogenous Portfolios Matter for HANK? Tech. rep.

Azinovic, M., Gaegauf, L. and Scheidegger, S. (2022). Deep equilibrium nets. *International Economic Review* **63** (4), 1471–1525.

Azinovic, M. and Zemlicka, J. (2024). Intergenerational consequences of rare disasters.

Barillas, F. and Fernández-Villaverde, J. (2007). A generalization of the endogenous grid method. *Journal of Economic Dynamics and Control* **31** (8), 2698–2712.

Bayer, C., Born, B. and Luetticke, R. (2022). The Liquidity Channel of Fiscal Policy. *Journal of Monetary Economics*.

— (2024). Shocks, Frictions, and Inequality in US Business Cycles. *American Economic Review* **114** (5), 1211–1247.

Bayer, C. and Luetticke, R. (2020). Solving Heterogeneous Agent Models in Discrete Time with Many Idiosyncratic States by Perturbation Methods. *Quantitative Economics* **11** (4), 1253–1288.

Bayer, C., Luetticke, R., Pham-Dao, L. and Tjaden, V. (2019). Precautionary Savings, Illiquid Assets, and the Aggregate Consequences of Shocks to Household Income Risk. *Econometrica* **87** (1), 255–290.

Bellman, R. (1957). *Dynamic Programming*. Princeton, NJ: Princeton University Press.

Bhandari, A., Bourany, T., Evans, D. and Golosov, M. (2023). A Perturbational Approach for Approximating Heterogeneous Agent Models. Tech. rep.

Bianchi, F., Melosi, L. and Rottner, M. (2021). Hitting the elusive inflation target. *Journal of Monetary Economics* **124**, 107–122.

Brock, W. A. and Mirman, L. J. (1972). Optimal economic growth and uncertainty: The discounted case. *Journal of Economic Theory* **4** (3), 479–513.

Brumm, J. and Scheidegger, S. (2017). Using adaptive sparse grids to solve high-dimensional dynamic models. *Econometrica* **85** (5), 1575–1612.

Carroll, C. D. (2006). The method of endogenous gridpoints for solving dynamic stochastic optimization problems. *Economics Letters* **91** (3), 312–320.

Chen, H., Didisheim, A. and Scheidegger, S. (2023). Deep surrogates for finance: With an application to option pricing. *Available at SSRN 3782722*.

Chen, M., Joseph, A., Kumhof, M., Pan, X. and Zhou, X. (2021). Deep reinforcement learning in a monetary model.

Den Haan, W. J. (2010). Assessing the accuracy of the aggregate law of motion in models with heterogeneous agents. *Journal of Economic Dynamics and Control* **34** (1), 79–99.

Druedahl, J. and Røpke, J. (2025). Deep Learning Algorithms for Finite-Horizon Models. *To be added*.

Duarte, V., Duarte, D. and Silva, D. H. (2024). Machine learning for continuous-time finance. *The Review of Financial Studies* **37** (11), 3217–3271.

Duarte, V. and Fonseca, J. (2024). Global identification with gradient-based structural estimation.

Fernández-Villaverde, J., Gillingham, K. and Scheidegger, S. (2024). Climate Change Through the Lens of Macroeconomic Modeling. NBER Working Paper 32963. 1050 Massachusetts Avenue, Cambridge, MA 02138: National Bureau of Economic Research.

Fernández-Villaverde, J., Hurtado, S. and Nuno, G. (2023). Financial frictions and the wealth distribution. *Econometrica* **91** (3), 869–901.

Fernández-Villaverde, J., Marbet, J., Nuño, G. and Rachedi, O. (2024). Inequality and the zero lower bound. *Journal of Econometrics*, 105819.

Fernández-Villaverde, J., Nuño, G. and Perla, J. (2024). Taming the curse of dimensionality: quantitative economics with deep learning. Tech. rep. National Bureau of Economic Research.

Friedl, A., Kübler, F., Scheidegger, S. and Usui, T. (2023). Deep uncertainty quantification: with an application to integrated assessment models. Tech. rep. Working Paper University of Lausanne.

Gornemann, N., Kuester, K. and Nakajima, M. (2016). Doves for the Rich, Hawks for the Poor? Distributional Consequences of Monetary Policy. *International Finance Discussion Paper* **2016** (1167), 1–40.

Gorodnichenko, Y., Maliar, L., Maliar, S. and Naubert, C. (2021). Household savings and monetary policy under individual and aggregate stochastic volatility.

Greenwood, J., Hercowitz, Z. and Huffman, G. (1988). Investment, Capacity Utilization, and the Real Business Cycle. *American Economic Review* **78** (3), 402–417.

Han, J., Yang, Y. and E, W. (2021). DeepHAM: A global solution method for heterogeneous agent models with aggregate shocks.

Hayashi, F. (1982). Tobin's Marginal q and Average q: A Neoclassical Interpretation. *Econometrica* **50** (1), 213.

Heer, B. and Maussner, A. (2024). *Dynamic general equilibrium modeling*. Springer.

Hintermaier, T. and Koeniger, W. (2010). The method of endogenous gridpoints with occasionally binding constraints among endogenous variables. *Journal of Economic Dynamics and Control* **34** (10), 2074–2088.

Judd, K. L. (1998). *Numerical Methods in Economics*.

Jungerman, W. (2024). Dynamic Monopsony and Human Capital. Working Paper.

Kahou, M. E., Fernández-Villaverde, J., Gomez-Cardona, S., Perla, J. and Rosa, J. (2024). Spooky Boundaries at a Distance: Inductive Bias, Dynamic Models, and Behavioral Macro. Tech. rep. National Bureau of Economic Research.

Kahou, M. E., Fernández-Villaverde, J., Perla, J. and Sood, A. (2021). Exploiting symmetry in high-dimensional dynamic programming. Tech. rep. National Bureau of Economic Research.

Kase, H., Melosi, L. and Rottner, M. (2022). Estimating Nonlinear Heterogeneous Agents Models with Neural Networks. Tech. rep. DP17391. Paris & London: CEPR Discussion Paper No. 17391.

Kekre, R. and Lenel, M. (2022). Monetary Policy, Redistribution, and Risk Premia. *Econometrica* **90** (5), 2249–2282.

King, R. G., Plosser, C. I. and Rebelo, S. T. (1988). Production, growth and business cycles. *Journal of Monetary Economics* **21** (2-3), 195–232.

Krusell, P. and Smith, A. A. (1997). Income and Wealth Heterogeneity, Portfolio Choice, and Equilibrium Asset Returns. *Macroeconomic Dynamics* **1** (02).

— (1998). Income and Wealth Heterogeneity in the Macroeconomy. *Journal of Political Economy* **106** (5), 867–896.

Maliar, L. and Maliar, S. (2022). Deep learning classification: Modeling discrete labor choice. *Journal of Economic Dynamics and Control* **135**, 104295.

Maliar, L., Maliar, S. and Winant, P. (2021). Deep learning for solving dynamic economic models. *Journal of Monetary Economics* **122**, 76–101.

McKay, A. and Wieland, J. F. (2021). Lumpy Durable Consumption Demand and the Limited Ammunition of Monetary Policy. *Econometrica* **89** (6), 2717–2749.

Miranda, M. J. and Fackler, P. L. (2004). *Applied computational economics and finance*. MIT press.

Nuño, G., Renner, P. and Scheidegger, S. (2024). Monetary policy with persistent supply shocks.

Pascal, J. (2024). Artificial neural networks to solve dynamic programming problems: A bias-corrected Monte Carlo operator. *Journal of Economic Dynamics and Control* **162**, 104853.

Payne, J., Rebei, A. and Yang, Y. (2024). Deep learning for search and matching models.

Reiter, M. (2009). Solving heterogeneous-agent models by projection and perturbation. *Journal of Economic Dynamics and Control* **33** (3), 649–665.

Richter, A. W., Throckmorton, N. A. and Walker, T. B. (2014). Accuracy, speed and robustness of policy function iteration. *Computational Economics* **44**, 445–476.

Tauchen, G. (1986). Finite state markov-chain approximations to univariate and vector autoregressions. *Economics Letters* **20** (2), 177–181.

Young, E. R. (2010). Solving the incomplete markets model with aggregate uncertainty using the Krusell–Smith algorithm and non-stochastic simulations. *Journal of Economic Dynamics and Control* **34** (1), 36–41.

Appendix

I Appendix: Solution of the Asset Pricing Model

This section illustrates our methodology using an analytical asset-pricing model and develops intuition for why the methodology works. In all our applications, the solution of a model is a set of policy functions that express controls as a function of state variables. In this first analytical example, the control variable we are looking for is the price of a nominal bond q_t. Let's assume that the price satisfies the Euler equation:

$$q_t = \beta \mathbb{E}_t \left[\frac{1}{\Pi_{t+1}} \frac{C_{t+1}^{-\gamma}}{C_t^{-\gamma}} \right] = \beta \mathbb{E}_t \Big[\exp\big(- \pi_{t+1} - \gamma \Delta c_{t+1} \big) \Big]$$

with $\pi_t \equiv \ln \Pi_t$, $\Delta c_t = \ln C_t - \ln C_{t-1}$, where Π_t is gross inflation between period t and $t-1$, C_t denotes consumption, as well as β, and γ are the discount factor and risk-aversion, respectively. We can write the equation more compact as

$$q_t = \beta \mathbb{E}_t \exp(\vec{\psi}' y_{t+1}), \tag{57}$$

where $y_t = [\pi_t, \Delta c_t]'$ and $\vec{\psi} = [-1, -\gamma]'$. Hence, the control variable q_t is a forward-looking variable, which depends nonlinearly on the expectations over the dynamics of the state variables y_{t+1}. In quantitative models, the policy functions we aim to solve for have a similar form as (57). For example, households make consumption-savings decisions and firms make capital accumulation decision by forming expectations about the future. To solve for the exact policy function, we need to introduce a law of motion for the states. We assume that the dynamics of the states y_t are described by a VAR(1) without intercept

$$y_t = A_1 y_{t-1} + e_t, \tag{58}$$

with $e_t = \eta \epsilon_t$, where ϵ_t is distributed $N(0, I)$. η is a $n_y \times n_\epsilon$ matrix, where $n_y = 2$ is the number of states, and n_ϵ is the number of shocks. This implies that e_t is distributed according to a $N(0, \Sigma)$, with variance-covariance matrix $\Sigma = \eta\eta'$. We give the shocks an economic interpretation by assuming $\epsilon_t = [\epsilon_t^a, \epsilon_t^\zeta, \epsilon_t^\mu]'$. Hence, the first shock denotes a TFP-shock, the second shock denotes a discount factor shock, and the last shock denotes a markup shock. The η matrix then denotes the loadings of the shocks onto the state variables. With these assumptions, the solution for the asset price q_t can be expressed as[15]

[15] The VAR(1) specifies that the vector of variables y_t is distributed according to a multivariate normal distribution. Together with the identity $\mathbb{E}_t \exp(x_{t+1}) = \exp(\mu_x + \frac{1}{2}\Sigma_x)$ for a normally distributed vector

$$q_t = \beta \exp\left(\vec{\psi}'\mu_t + \frac{1}{2}\vec{\psi}'V_t\vec{\psi}\right) \tag{59}$$

where $\mu_t = A_1 y_t$, denotes the conditional mean forecast, and $V_t = \Sigma$ the conditional variance of the variables in y_t. Let η_{ij}, and a_{ij} denote the entries in the i'th row and j'th column of the shock impact matrix η and the matrix of the VAR A_1 respectively. Then the full solution for the price of the asset can be written as

$$\begin{aligned} q_t = \beta \exp\Bigg(& -\left(a_{11} + \gamma a_{21}\right)\pi_t \; - \; \left(a_{12} + \gamma a_{22}\right)\Delta c_t \\ & + \tfrac{1}{2}\left[\left(\eta_{11} + \gamma\eta_{21}\right)^2 + \left(\eta_{12} + \gamma\eta_{22}\right)^2 + \left(\eta_{13} + \gamma\eta_{23}\right)^2\right]\Bigg). \end{aligned} \tag{60}$$

Equation (60) expresses how the bond price depends on the inclusion of different shocks to our dynamic system of equations. While the first line components remain unchanged, the second line changes depending on the shocks that we include in the model. Consequently, this model serves as a natural illustration for the functionality of our approach.

When solving a simplified version of the above model that contains only two of the three shocks, the entries $\eta_{1\cdot}$ and $\eta_{2\cdot}$ are equal to zero, where the $\cdot$ is a placeholder for the shock which is not included, anymore. Let us denote the resulting equilibrium price without a shock i by $q_t^{\setminus i}$. Consequently, simplified model versions feature the prices

$$\begin{aligned} q_t^{\setminus a} = \beta \exp\Bigg(& -\left(a_{11} + \gamma a_{21}\right)\pi_t \; - \; \left(a_{12} + \gamma a_{22}\right)\Delta c_t \\ & + \tfrac{1}{2}\left[\left(0 + \gamma \cdot 0\right)^2 + \left(\eta_{12} + \gamma\eta_{22}\right)^2 + \left(\eta_{13} + \gamma\eta_{23}\right)^2\right]\Bigg) \end{aligned} \tag{61}$$

$$\begin{aligned} q_t^{\setminus \zeta} = \beta \exp\Bigg(& -\left(a_{11} + \gamma a_{21}\right)\pi_t \; - \; \left(a_{12} + \gamma a_{22}\right)\Delta c_t \\ & + \tfrac{1}{2}\left[\left(\eta_{11} + \gamma\eta_{21}\right)^2 + \left(0 + \gamma \cdot 0\right)^2 + \left(\eta_{13} + \gamma\eta_{23}\right)^2\right]\Bigg) \end{aligned} \tag{62}$$

$$\begin{aligned} q_t^{\setminus \mu} = \beta \exp\Bigg(& -\left(a_{11} + \gamma a_{21}\right)\pi_t \; - \; \left(a_{12} + \gamma a_{22}\right)\Delta c_t \\ & + \tfrac{1}{2}\left[\left(\eta_{11} + \gamma\eta_{21}\right)^2 + \left(\eta_{12} + \gamma\eta_{22}\right)^2 + \left(0 + \gamma \cdot 0\right)^2\right]\Bigg) \end{aligned} \tag{63}$$

$x_{t+1} \sim N(0, \Sigma_x)$ we obtain the closed form expression for the asset price.

Equations (61) - (63) illustrate that shutting down individual shocks reduces the price of the risk-free bond compared to the specification with all shocks by a constant.

II Appendix: Real Business Cycle Model

II.1 Appendix: Equilibrium and steady-state of the model

This section illustrates the steady state of the economy. We first illustrate all equations that characterize the equilibrium before characterizing the steady state.

Equilibrium characterization: In a symmetric equilibrium, the production side is characterized by

$$Y_t: \quad Y_t = A_t(u_t K_t)^\alpha (Z_t N_t)^{1-\alpha} \tag{64}$$

$$r_t: \quad r_t + q_t\delta(u_t) = \frac{\alpha}{\mu_t}\frac{Y_t}{K_t} \tag{65}$$

$$w_t: \quad w_t = \frac{1-\alpha}{\mu_t}\frac{Y_t}{N_t} \tag{66}$$

$$u_t: \quad q_t[\delta_1 + \delta_2(u_t - 1)] = \frac{\alpha}{\mu_t}\frac{Y_t}{u_t} \tag{67}$$

$$\Pi_t: \quad \Pi_t = \mu_t Y_t \tag{68}$$

$$q_t: \quad q_t = \left[1 - \phi\left(\frac{I_t}{K_t} - \delta_{0t}\right)^{\kappa-1}\right]^{-1} \tag{69}$$

$$I_t: \quad K_{t+1} = (1-\delta(u_t))\,K_t + I_t - \frac{\phi}{\kappa}\left(\frac{I_t}{K_t} - \delta_{0t}\right)^{\kappa} K_t \tag{70}$$

With four exogenous processes:

$$\mu_t: \quad \ln\mu_t = (1-\rho_\mu)\left(\ln\mu - \frac{\sigma_\mu^2}{2}\right) + \rho_\mu \ln\mu_{t-1} + \epsilon_t^\mu \quad \text{with} \quad \epsilon_t^\mu \sim N(0,\sigma_\mu^2) \tag{71}$$

$$A_t: \quad \ln A_t = (1-\rho_A)\left(\ln A - \frac{\sigma_A^2}{2}\right) + \rho_A \ln A_{t-1} + \epsilon_t^A \quad \text{with} \quad \epsilon_t^A \sim N(0,\sigma_A^2), \tag{72}$$

$$Z_t: \quad \ln Z_t = (1-\rho_Z)\left(\ln Z - \frac{\sigma_Z^2}{2}\right) + \rho_Z \ln Z_{t-1} + \epsilon_t^Z \quad \text{with} \quad \epsilon_t^Z \sim N(0,\sigma_Z^2), \tag{73}$$

$$\delta_{0t}: \quad \delta_{0t} = \delta_0 + \epsilon_t^\delta, \tag{74}$$

The household side is characterized by the aggregate functions $\mathcal{C}_t$, $\mathcal{N}_t$ and $\mathcal{K}_{t+1}$ that represent the aggregate consumption, labor, and capital function of the (potentially heterogenous) household side of the model. We can characterize the aggregate functions as

$$\mathcal{C}_t: \qquad \mathcal{C}_t(\{r_t, w_t, q_t, \zeta_t, \tau^k, \tau^L, \tau^C\}) = \int c^*_{it} d\Theta_t, \tag{75}$$

$$\mathcal{N}_t: \qquad \mathcal{N}_t(\{r_t, w_t, q_t, \zeta_t, \tau^k, \tau^L, \tau^C\}) = \int n^*_{it} d\Theta_t, \tag{76}$$

$$\mathcal{K}_{t+1}: \qquad \mathcal{K}_{t+1}(\{r_t, w_t, q_t, \zeta_t, \tau^k, \tau^L, \tau^C\}) = \int k^*_{it} d\Theta_t, \tag{77}$$

with the integral integrating over the distribution of households Θ_t and lowercase letters with asterisks denoting the policy functions of a household in period t that are determined by the first-order conditions

$$c_{it}: \qquad q_t u_C(c_{it}, n_{it}) = \beta \mathbb{E}_t \left[\frac{\zeta_{t+1}}{\zeta_t} \left(q_{t+1} + (1-\tau^K) r_{t+1} \right) u_C(c_{it+1}, n_{it+1}) \right] \tag{78}$$

$$n_{it}: \qquad -u_L(c_{it}, n_{it}) = (1-\tau^L) w_t h_{it} \frac{u_C(c_{it}, n_{it})}{1+\tau^C}, \tag{79}$$

and residually via the budget constraint:

$$k_{it+1}: (1+\tau^C) c_{it} + q_t k_{it+1} = \left(q_t + (1-\tau^K) r_t \right) k_{it} + (1-\tau^L) w_t h_{it} n_{it} + T_t + \Pi_{it} \tag{80}$$

Generally, the distribution Θ_t of agents develops according to a law of motion

$$\Theta_{t+1}: \quad \Theta_{t+1}(h_{it+1}, k_{it+1}) = \int_{k_{it+1}=k^*_{it}} \Phi(h_{it}, h_{it+1}) d\Theta_t(h_{it}, k_{it}), \tag{81}$$

with Φ denoting the transition of idiosyncratic states. In the representative agent case, the aggregate functions just boil down to the first-order conditions (78), (79), and (80). On the household side, we then have the exogenous processes:

$$\zeta_t: \qquad \ln \zeta_t = -(1-\rho_\zeta)\frac{\sigma_\zeta^2}{2} + \rho_\zeta \ln \zeta_{t-1} + \epsilon^\zeta_t \quad \text{with} \quad \epsilon^\zeta_t \sim N(0, \sigma^2_\zeta). \tag{82}$$

$$h_{it}: \qquad \log h_{it} = -(1-\rho_h)\frac{\sigma_h^2}{2} + \rho_h \log h_{it-1} + \epsilon_{it}. \tag{83}$$

Finally, the government sector is characterized by

$$T_t: \quad T_t = \tau^C C_t + \tau^K r_t K_t + \tau^L w_t N_t, \tag{84}$$

while market clearing is characterized by

$$N_t = \mathcal{N}_t, \tag{85}$$

$$K_t = \mathcal{K}_t, \tag{86}$$

$$C_t = \mathcal{C}_t \tag{87}$$

$$Y_t = C_t + I_t, \tag{88}$$

Hence, in equilibrium, the 25 equations above determine the 25 variables $\{Y_t, r_t, w_t, u_t, \Pi_t, q_t, I_t, \mu_t, A_t, Z_t, \delta_{0t}, \{c^*_{it}\}_i, C_t, \mathcal{C}_t, \{n^*_{it}\}_i, N_t, \mathcal{N}_t, K_t, \{k^*_{it}\}_i, K_{t+1}, \mathcal{K}_{t+1}, \zeta_t, h_{it}, T_t, \Theta_t\}$.

Steady-State: Based on the description of the equilibrium, we now can turn to describe the steady state. First, except for idiosyncratic income risk h_{it} all exogenous processes revert to their long-run averages such that:

$$\mu^{SS}: \qquad \ln\mu^{SS} = \ln\mu - \frac{\sigma^2_\mu}{2} \tag{89}$$

$$A^{SS}: \qquad \ln A^{SS} = \ln A - \frac{\sigma^2_A}{2} \tag{90}$$

$$Z^{SS}: \qquad \ln Z^{SS} = \ln Z - \frac{\sigma^2_Z}{2} \tag{91}$$

$$\delta_0^{SS}: \qquad \delta_0^{SS} = \delta_0 \tag{92}$$

$$\zeta^{SS}: \qquad \ln\zeta^{SS} = -\frac{\sigma^2_\zeta}{2}. \tag{93}$$

In the heterogeneous agent version of the model in the steady state, there still exists heterogeneity on the household side in a steady state. The steady state on the household side is then characterized by constant *aggregates*, while still featuring fluctuations at the individual household level. Hence

$$\mathcal{C}^{SS}: \qquad \mathcal{C}^{SS} = \mathcal{C}_t = \mathcal{C}_{t+1} = \int c^*_i d\Theta^{SS}, \tag{94}$$

$$\mathcal{N}^{SS}: \qquad \mathcal{N}^{SS} = \mathcal{N}_t = \mathcal{N}_{t+1} = \int n^*_i d\Theta^{SS}, \tag{95}$$

$$\mathcal{K}^{SS}: \qquad \mathcal{K}^{SS} = \mathcal{K}_{t+1} = \mathcal{K}_{t+2} = \int k^*_i d\Theta^{SS}, \tag{96}$$

with Θ^{SS} denoting the time-invariant stationary distribution of households that solves

$$\Theta^{SS}: \quad \Theta^{SS}(h_i, k_i) = \int_{k_i = k^*_i} \Phi(h_{it}, h_{it+1}) d\Theta^{SS}(h_i, k_i). \tag{97}$$

The lower-case letters with asterisks subscript i denote the policy functions given constant aggregates. In the representative version of the model, this boils down to

$$r^{SS}: \quad q^{SS} = \beta\left[\left(q^{SS} + (1-\tau^K)r^{SS}\right)\right] \tag{98}$$

$$\mathcal{N}^{SS}: \quad -u_L(\mathcal{C}^{SS}, \mathcal{N}^{SS}) = \frac{1-\tau^L}{1+\tau^C} w^{SS} u_C(\mathcal{C}^{SS}, \mathcal{N}^{SS}), \tag{99}$$

$$\mathcal{C}^{SS}: \quad (1+\tau^C)\mathcal{C}^{SS} = (1-\tau^L)w^{SS}\mathcal{N}^{SS} + (1-\tau^K)r^{SS}\mathcal{K}^{SS} + T^{SS} + \Pi^{SS}. \tag{100}$$

Using the market clearing conditions determines aggregate consumption and aggregate labor

$$\mathcal{C}^{SS} = C^{SS} \quad \text{and} \quad \mathcal{N}^{SS} = N^{SS} \tag{101}$$

We use steady-state labor supply N^{SS} and the steady-state real interest rate r^{SS} to pin down the production side:

$$Y^{SS}: \quad Y^{SS} = A^{SS}(u^{SS}K^{SS})^{\alpha}(Z^{SS}N^{SS})^{1-\alpha} \tag{102}$$

$$K^{SS}: \quad r^{SS} + q^{SS}\delta(u^{SS}) = \frac{\alpha}{\mu^{SS}}\frac{Y^{SS}}{K^{SS}} \tag{103}$$

$$w^{SS}: \quad w^{SS} = \frac{1-\alpha}{\mu^{SS}}\frac{Y^{SS}}{N^{SS}} \tag{104}$$

$$u^{SS}: \quad q_t[\delta_1 + \delta_2(u^{SS}-1)] = \frac{\alpha}{\mu^{SS}}\frac{Y^{SS}}{u^{SS}} \tag{105}$$

$$\Pi^{SS}: \quad \Pi_t = \mu_t Y_t \tag{106}$$

$$q^{SS}: \quad q^{SS} = \left[1 - \phi\left(\frac{I^{SS}}{K^{SS}} - \delta_0^{SS}\right)^{\kappa-1}\right]^{-1} \tag{107}$$

$$I^{SS}: \quad I^{SS} = \left(\delta(u^{SS}) + \frac{\phi}{\kappa}\left(\frac{I^{SS}}{K^{SS}} - \delta_0^{SS}\right)^{\kappa}\right)K^{SS} \tag{108}$$

Government transfers are then determined by

$$T^{SS}: \quad T^{SS} = \tau^C C^{SS} + \tau^K r^{SS} K^{SS} + \tau^L w^{SS} N^{SS}, \tag{109}$$

Finally, the capital and the goods market have to clear

$$K^{SS} = \mathcal{K}^{SS}, \tag{110}$$

$$\text{and} \quad Y^{SS} = C^{SS} + I^{SS}. \tag{111}$$

Having defined all equations that have to hold in steady state, we can now characterize the solution for a specific steady state.

Steady-State determination: It is easiest to start by imposing some normalization on the production side. We want to normalize capacity utilization u^{SS} and the price for capital to unity in the steady state q^{SS}. From equations (105) and (107) this derives the parameter restrictions

$$\frac{I^{SS}}{K^{SS}} = \delta_0^{SS} \quad \text{and} \quad \delta_1 = \frac{\alpha}{\mu^{SS}} Y^{SS}. \tag{112}$$

which pins down the depreciation values besides investment in steady state. The rest of the steady state can then be characterized by a joint solution of the household side and the firm side that satisfies market clearing conditions. In the heterogenous agent case, we solve the household problem for each household along the state space given a guess for aggregate capital K_t and aggregate labor N_t, obtain the aggregate labor function (95) and capital function (96) and update the guesses for capital and labor until labor market clearing (101) and capital market clearing (110) are satisfied.

In the representative agent case, equation (98) determines the steady state real interest rate as

$$r^{SS}: \quad r^{SS} = \frac{1-\beta}{\beta(1-\tau^K)}. \tag{113}$$

In both cases, an interest rate also determines the steady-state capital-to-labor ratio and wage rate

$$K^{SS}: \qquad \frac{K^{SS}}{N^{SS}} = \left(\frac{\alpha}{r^{SS}+\delta_0^{SS}} \frac{A^{SS}(Z^{SS})^{1-\alpha}}{\mu^{SS}}\right)^{\frac{1}{1-\alpha}} \tag{114}$$

$$w^{SS}: \qquad w^{SS} = (1-\alpha)\frac{A^{SS}(Z^{SS})^{1-\alpha}}{\mu^{SS}} \left(\frac{K^{SS}}{N^{SS}}\right)^{\alpha}. \tag{115}$$

II.2 The analytical model

Derivation of the analytical model The part below derives the proof related to the analytical model in section 3.2.1. To solve the model analytically, we abstract from shocks to the discount factor (ζ_t) and depreciation rate (δ_t). We keep the shocks to technology (A_t), productivity (Z_t), shocks to the markup (μ_t). Finally, we abstract from capital income taxation ($\tau^K = 0$) and assume full depreciation ($\delta_{0t} = 1$).[16] Moreover, we abstract from household heterogeneity and let households be ex-ante identical by assuming away

[16] The combination of the assumptions renders the model unrealistic, as already noted by Brock and Mirman (1972) themselves. We do not employ the model for realistic reasons, but because it provides us with an analytical benchmark we can use. This also motivates the choice of our shocks.

differences in idiosyncratic income $h_{i0} = 1$ and $\Pi_{it} = \Pi_t \,\forall\, i$ and initial capital holdings are identical $k_{i0} = K_0 \,\forall\, i$. We make households ex-post identical by disregarding idiosyncratic income risk $\sigma_h^2 = 0$. The absence of ex-ante or ex-post heterogeneity enables us to represent the household side through a representative agent. Therefore, we drop the individual index i to describe the variables of interest.

Proof. The proof employs a guess-and-verify approach. Guess that the policy function for savings is given by $K_{t+1} = \Gamma Y_t$. Substituting the guess into the goods market clearing condition (35) while imposing the parameter restriction $\delta = 1$ yields

$$C_t = (1 - \Gamma) Y_t.$$

We use the two guesses and substitute into the Euler equation (31)

$$\frac{1}{C_t} = \beta \mathbb{E}_t \left[\frac{\frac{\alpha}{\mu_{t+1}} \frac{Y_{t+1}}{K_{t+1}}}{C_{t+1}} \right] \Leftrightarrow \frac{1}{(1-\Gamma)Y_t} = \beta \mathbb{E}_t \left[\frac{\alpha}{\mu_{t+1}(1-\Gamma)K_{t+1}} \right],$$

from which it is straightforward to see that $\Gamma = \frac{\alpha\beta}{\mu}$ given that $\mathbb{E}_t \frac{1}{\mu_{t+1}} = \mu^{-1}$ with $\rho_\mu = 0$. Note that the value-added-tax $(1 + \tau^C)$ drops from the Euler equation, since it is constant over time. To obtain the policy function (36) we substitute the guesses with specified Γ into the labor-supply condition (32)

$$\frac{(1-\tau_t^N) w_t}{(1+\tau^C) C_t} = \omega N_t^\gamma \Leftrightarrow \frac{(1-\tau_t^N)(1-\alpha)}{\mu_t (1+\tau^C)(1-\Gamma) N_t} = \omega N_t^\gamma,$$

from which we obtain expression (36) when solving for N_t. □

With KPR-preferences with log-felicity over consumption the income and substitution effect of wage changes cancel out. Therefore, only shocks to the wage tax rate τ_t^L and the markup μ_t impact the level of equilibrium labor supply.

Calibration Table 2 illustrates the parameter values that we use for solving the model. We largely use standard values from the literature, but some variables require further explanation. First, we have a steady-state markup μ of 1.5, which is very high. We introduce such a high markup value such as to amplify the effects of markup shocks on the economy. Moreover, we shut down the government by setting all tax rates equal to zero. Finally, we do not only simulate the economy with fixed volatilities of the shocks but allow for different volatility levels. While the model only features three shocks, the shocks can have different volatilities, such that we generate data sets for all three shock combinations with

Table 2 Parameter values of the analytical Brock and Mirman (1972) model

Parameter	Value	Description	Parameter	Value	Description
Households			**Exogenous processes**		
β	0.96	Discount factor	A	1.0	Steady state TFP
γ	5	Inverse Frisch	ρ_a	0.9	TFP persistence
ω	1.0	Scale labor disutility	σ_a	{0.0, 0.01, 0.05}	TFP std.
			Z	1.0	Steady state labor prod.
Firms			ρ_z	0.9	Labor prod. persistence
α	0.33	Capital share	σ_z	{0.0, 0.01, 0.05}	Labor prod. std.
δ	1.0	Depreciation rate	μ	1.1	Steady State Markup
			ρ_μ	0.0	Markup persistence
Government			σ_μ	{0.0, 0.01, 0.05}	Markup std.
τ^L	0.0%	Labor tax rate level			
τ^R	0.0%	Capital tax rate level			
τ^C	0.0%	Value-added tax rate level			

NOTE - All parameters in the table are calibrated to a yearly frequency.

Table 3 Parameter values of the nonlinear medium-sized RBC model

Parameter	Value	Description	Parameter	Value	Description
Households			**Firms**		
σ	1	Risk aversion	α	0.33	Capital share
β	0.99	Discount factor	δ_0	0.025	Depreciation rate
γ	1	Inverse Frisch	δ_1	αY^{SS}	Depreciation rate
ω	0.5	Scale labor disutility	δ_2	$5\delta_1$	Depreciation rate
			κ	2	Cap. adj. cost curvature
			$\bar{\phi}$	2.5	High slope of cap. adj. cost
Exogenous processes			$\underline{\phi}$	0.025	Low slope of cap. adj. cost
A	1.0	Steady state TFP			
ρ_a	0.95	TFP persistence	**Government**		
σ_a	0.01	TFP std.	τ^L	0.0%	Labor tax rate level
ρ_ζ	0.95	Discount factor persistence	τ^R	0.0%	Capital tax rate level
σ_ζ	0.05	Discount factor std.	τ^C	0.0%	Value-added tax rate level
σ_δ	0.004	Depreciation std.			

NOTE - All parameters in the table are calibrated to a quarterly frequency.

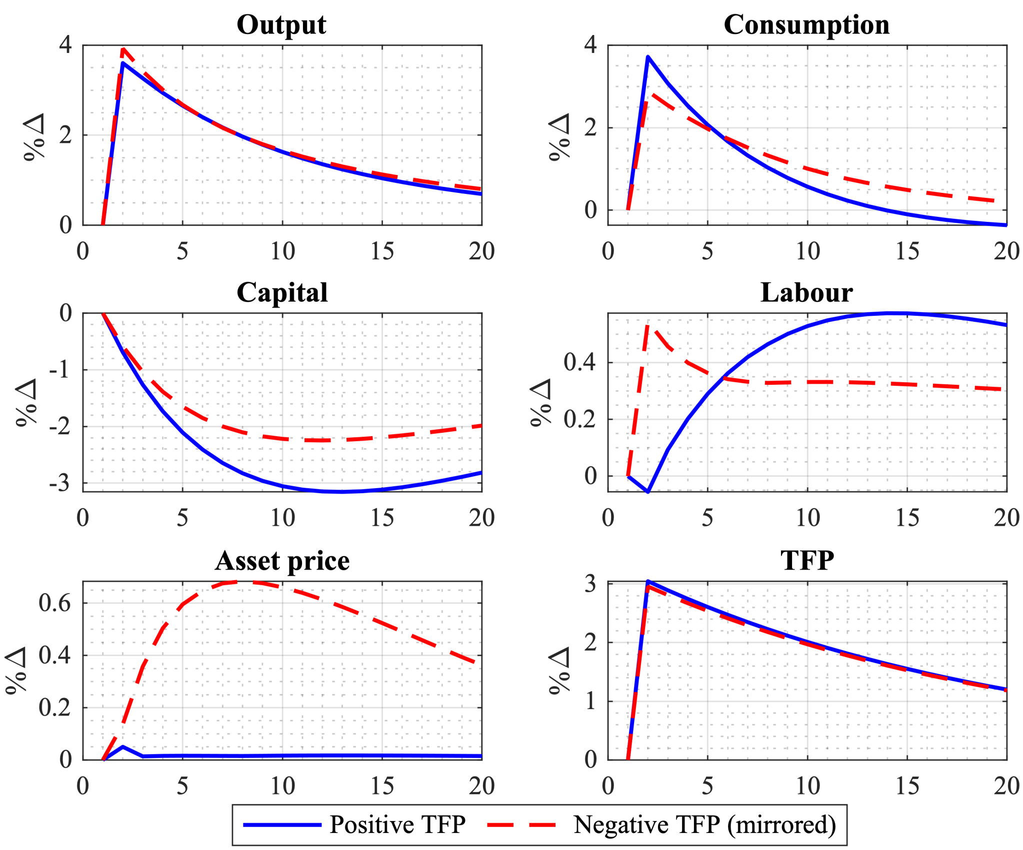
different volatilities. This also challenges the surrogate network since it needs to learn the model dynamics for different shock volatilities.

II.3 Nonlinear version

Table 3 presents the parameter choices for solving the model. Most parameters align with standard values in the literature but are calibrated at a quarterly frequency. Compared to the model in Section 3.1, we introduce partial depreciation, a capacity utilization choice, a lower Frisch elasticity of labor supply, and capital adjustment costs. Additionally, we incorporate nonlinear capital adjustment costs by allowing ϕ_t to take values $\bar{\phi}$ and $\underline{\phi}$ depending on the capital stock. The other parameters are standard.

The model features strong nonlinearities as the impulse response functions (IRFs) in Figure 8 highlight. In this simulation, we compare the impact of an expansionary and contractionary three-standard deviation TFP shock. We display the percentage deviation from

Figure 8 IRFs and Nonlinear Propagation of the TFP Shock



the stochastic steady state and mirror the IRFs of the negative TFP for easier comparison. The state-dependent investment costs result in strong differences between positive and negative shocks. We observe a similar behavior when evaluating the preference and discount rate shock. This is an important precondition for our analysis as we want to evaluate the performance of our approach in a highly nonlinear environment.

II.4 Heterogeneous agent version

Solution approach This subsection presents the solution to the heterogeneous agent model following the methodology of Krusell and Smith (1997) and Krusell and Smith (1998).[17] In the consumption-savings problem, households require a prediction of next period's capital given today's state space. With heterogeneous agents, this would typically require households to track the entire distribution of households over the state space, Θ_t as an additional state variable, which renders the problem numerically intractable. To address this, Krusell and Smith demonstrate that households do not need to keep track of the

[17] The code is available at `https://github.com/Fabio-Stohler/KS1998`.

full distribution Θ_t, but a few moments of the distribution suffice to forecast future capital. Specifically, they approximate the law of motion for capital using its mean, allowing households to form expectations based on a simplified perceived law of motion. Let $\vec{A}$, $\vec{\zeta}$, and $\vec{\delta}$ denote the discretized grid of aggregate productivity, discount factor, and depreciation shock with N_A, N_ζ, and N_δ shocks. We generalize on the original paper and assume that the law of motion takes the following state-dependent functional form:

$$\begin{aligned}\ln K_{t+1} = \beta_0 + \sum_{i=1}^{N_A} \beta_{A,i}\, \mathbf{1}_{\{A_t=A^i\}} + \sum_{j=1}^{N_\zeta} \beta_{\zeta,j}\, \mathbf{1}_{\{\zeta_t=\zeta^j\}} + \sum_{k=1}^{N_\delta} \beta_{\delta,k}\, \mathbf{1}_{\{\delta_t=\delta^k\}} \\ + \beta_K \ln K_t + \sum_{i=1}^{N_A} \gamma_{A,i}\, \mathbf{1}_{\{A_t=A^i\}} \ln K_t + \sum_{j=1}^{N_\zeta} \gamma_{\zeta,j}\, \mathbf{1}_{\{\zeta_t=\zeta^j\}} \ln K_t + \sum_{k=1}^{N_\delta} \gamma_{\delta,k}\, \mathbf{1}_{\{\delta_t=\delta^k\}} \ln K_t\end{aligned} \tag{116}$$

Our extension allows for an arbitrary many realizations of the discretized shocks and allows for both the slope and the intercept to vary with each (discretized) aggregate state value. The solution algorithm consists of an inner and an outer loop. The outer loop iterates until the coefficients in the regression equation (116) converge. Let

$$\boldsymbol{\beta}^n = \left(\beta_0^n,\, \left\{\beta_{A,i}^n\right\}_{i=1}^{N_A},\, \left\{\beta_{\zeta,j}^n\right\}_{j=1}^{N_\zeta},\, \left\{\beta_{\delta,k}^n\right\}_{k=1}^{N_\delta},\, \beta_K^n,\, \left\{\gamma_{A,i}^n\right\}_{i=1}^{N_A},\, \left\{\gamma_{\zeta,j}^n\right\}_{j=1}^{N_\zeta},\, \left\{\gamma_{\delta,k}^n\right\}_{k=1}^{N_\delta}\right)'$$

denote the vector of regression coefficients for the perceived law of motion of iteration n of the algorithm. Convergence is determined by checking whether the coefficients remain unchanged across iterations. If the coefficients changed by less then a small ϵ, the algorithm terminates.[18]

The inner loop iterates until the household problem is globally solved for a given perceived law of motion for capital. To solve the household side, we discretize the space $(k_{it}, h_{it}, K_t, A_t, \zeta_t, \delta_t)$ and use the endogenous grid-point method (EGM) of Carroll (2006) to solve the household problem given the stochastic processes and the perceived law of motion. Household policies are updated iteratively until the (inverse) marginal values of consumption converge. Once the household problem is solved globally, we aggregate and simulate the economy for T periods using the stochastic simulation method of Young (2010). Finally, using the simulated time series of capital and the aggregate states, we estimate the regression equation (116) to update the law of motion.

We illustrate the algorithm as follows. Let $\vec{k}$, $\vec{h}$, denote the discretized vectors of individual capital holdings, individual productivity, respectively.

[18] We also verify that the true law of motion for capital closely matches the perceived law of motion, which is generally the case.

1. For each realization of aggregate states $\{K_t, A_t, \zeta_t, \delta_t\}$ compute labor L_t (which depends on the aggregate state), as well as the interest rate r_t and the wage rate w_t as the marginal products of capital and labor. For each realization of the individual state space $\{k_{it}, h_{it}, K_t, A_t, \zeta_t, \delta_t\}$ compute household incomes.

2. Initialize the coefficients for the law of motion (116). Typically, the intercepts (β_0, β_1, β_2, β_3) are set to zero, and the slopes (β_4, β_5, β_6, β_7) to one.

3. Given the coefficients for the law of motion, solve the household problem using EGM[19]:

   (a) Initialize guesses for the policy functions c_{it}^0, k_{it}^0 defined on the state space $(k_{it}, h_{it}, K_t, A_t, \zeta_t, \delta_t)$. Create an initial guess for the marginal value function $\frac{\partial V_{it}^0}{\partial k_{it}} = (1+r)\frac{\partial u(c_{it}^0)}{\partial c_{it}}$. The superscript denotes the iteration step mm of the EGM algorithm.

   (b) For each realization of the aggregate state today $\{K_t, A_t, \zeta_t, \delta_t\}$ forecast next period's capital stock using the perceived law of motion. Let $\tilde{K}_{t+1}$ denote the forecasted capital stock according to the perceived law of motion. Interpolate the marginal value $\frac{\partial V_{it}^{n-1}}{\partial k_{it}}$ from the exogenous grid $\vec{K}$ onto the perceived value in the next period $\tilde{K}_{t+1}$. Finally, compute the expected marginal value by integrating over the realizations of the aggregate $\{A_t, \zeta_t, \delta_t\}$ and idiosyncratic states $\{h_{it}\}$ and discount the expected value with the discount factor.[20]

   (c) Apply the inverse of the marginal utility function to the interpolated expected marginal value to find the policy function of consumption $\hat{c}_{it}$ on the endogenous grid $\tilde{k}_{it}$.

   (d) Compute the endogenous grid points $\tilde{k}_{it}$ from the budget constraint given the policy $\hat{c}_{it}$.

   (e) Interpolate the consumption policy function $\hat{c}_{it}$ from the endogenous grid $\tilde{k}_{it}$ onto the exogenous grid $\vec{k}$ to obtain an updated policy function c_{it}^m.

   (f) Enforce the borrowing constraint.

   (g) Check convergence by verifying for a small ϵ, whether the condition $|u'^{-1}\left(\frac{\partial V_{it}^m}{\partial k_{it}}\right) - u'^{-1}\left(\frac{\partial V_{it}^{m-1}}{\partial k_{it}}\right)| < \epsilon$ is true. $|\cdot|$ denotes the Euclidean norm. If the condition is not satisfied, repeat steps (b)–(g).

[19] There exist numerous resources that go into detail in the illustration of the method. We only highlight the differences that occur due to the presence of aggregate risk. The interested reader might consult Carroll (2006), Barillas and Fernández-Villaverde (2007), Hintermaier and Koeniger (2010), and the appendix of Bayer et al. (2019).

[20] Note that the idiosyncratic risk depends on the realization of the aggregate risk.

4. With the converged global policy functions, we aggregate and simulate the economy using stochastic simulation of Young (2010) for T periods:

   (a) We set the initial capital stock to the value from the deterministic steady state of a representative agent economy and denote the capital stock as K_1.[21]

   (b) In period t, we have capital stock K_t, which is generally off-grid. To evaluate the policy functions of the household, we evaluate the individual policy function at K_t by interpolating from the exogenous grid $\vec{K}$ on the current capital stock K_t. Evaluate the policy functions at the current aggregate state $\{A_t, \zeta_t, \delta_t\}$.

   (c) Given household policies evaluated at the state realizations today, we update the household distribution using the stochastic simulation method of Young (2010).

   (d) Repeat steps (b) and (c) for T periods.

5. Discard the first 1000 periods as a burn-in sample. Use the remaining time series to update the perceived law of motion by regressing the logarithm of the capital stock on the aggregate states and the lagged logarithm of the capital stock as in equation (116). Denote the resulting regression coefficients as $\tilde{\boldsymbol{\beta}}^m$

6. Check whether $|\tilde{\boldsymbol{\beta}}^n - \boldsymbol{\beta}^{n-1}| < \epsilon$ for a small ϵ. If the condition is met, stop; otherwise, update the coefficients as $\boldsymbol{\beta}^n = \varphi\tilde{\boldsymbol{\beta}}^n + (1-\varphi)\boldsymbol{\beta}^{n-1}$ with $\varphi \in (0, 0.5)$ and repeat steps (3) - (6).

Note that the description above accounts for all aggregate shocks but also accommodates cases with fewer aggregate shocks by keeping some of them fixed. We apply steps (1) to (6) to each satellite model, generating a dataset that is then used to train the surrogate model.

Calibration Table 4 reports the parameter values used to solve the satellite models and simulate the corresponding data. For household and firm behavior, we adopt the parameterization from Krusell and Smith (1998), calibrated at an annual frequency. The exogenous shock processes are also specified using standard annual values commonly found in the literature.

Moreover, figure 9 illustrates the perceived law of motion of the households in comparison to the true law of motion for capital in the economy. As the plot illustrates, the perceived law of motion and the true law of motion closely align, with the error between the two lines generally being below one percent of the capital stock.

[21] We also find a distribution that has the mean of K_1 and initialize the simulation with this distribution.

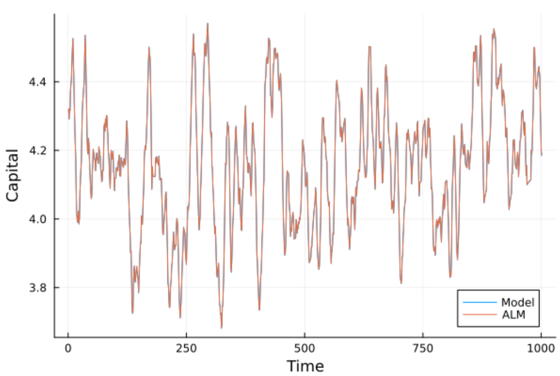
Table 4 Parameter values of heterogenous agent model

Parameter	Value	Description	Parameter	Value	Description
Households			**Exogenous processes**		
β	0.95	Discount factor	ρ_h	0.9	Idiosy. risk persistence
σ	1.0	Risk aversion	σ_h	0.15	Idiosy. risk std.
			A	1	Steady State TFP
Firms			ρ_a	0.75	TFP persistence
α	0.36	Capital share	σ_a	0.02	TFP std.
δ	0.1	Steady State depreciation	ζ	1	Steady State Discount fact.
			ρ_ζ	0.75	Discount fact. persistence
Government			σ_ζ	0.02	Discount Fact. std.
τ^L	0.0%	Labor tax rate level	σ_δ	0.01	Depreciation std.
τ^R	0.0%	Capital tax rate level			
τ^C	0.0%	VAT rate level			

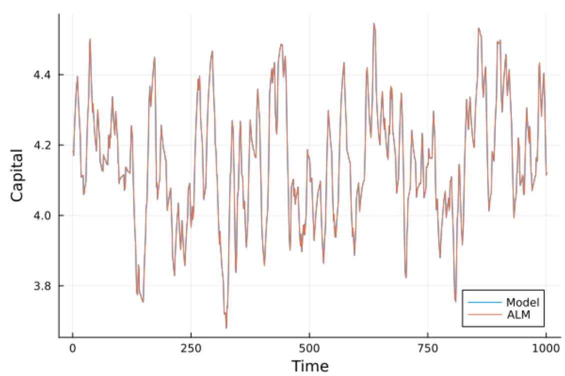
NOTE - All parameters in the table are calibrated to a yearly frequency.

Figure 9 Model implied series of capital and perceived aggregate law of motion (ALM)

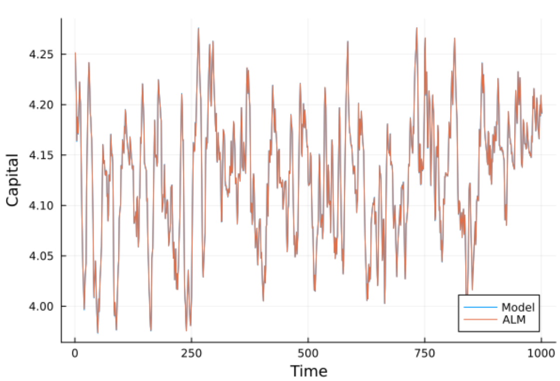
(a) Model with TFP and zeta shocks



(b) Model with TFP and delta shocks



(c) Model with zeta and delta shocks



(d) Model with TFP, zeta, and delta shocks