CERGE
Center for Economic Research and Graduate Education
Charles University

Master Thesis

2021 **Jan Žemlička**

CERGE

Center for Economic Research and Graduate Education

Charles University

Macro-Epidemic Modelling: A Deep Learning Approach

Jan Žemlička

Master Thesis

Prague, July 2021

Author: Bc. Jan Žemlička
Supervisor: Ctirad Slavík, Ph.D.

Declaration of Autorship:
I hereby proclaim that I wrote my master thesis on my own under the leadership of my supervisor and that the reference list include all resources and literature I have used.

I grant a permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, ——————————-
July 26, 2021 Signature

Master Thesis Project

Author: Bc. Jan Žemlička

Supervisor: Ctirad Slavík, Ph.D.

Academic Year: 2020/2021.

Topic: Macro-Epidemic Modelling: A Deep Learning Approach

Motivation and Literature Review:

Conventional solution methods for macroeconomic models aren't able to tackle the challenge presented by macro-epidemic models: strong nonlinearity of epidemic process coupled with rich heterogeneity of economic agents. Due to this fact, macro-epidemic modeling is restricted to rather simplistic environments abstracting either from heterogeneity among households and firms (e.g., Gonzalez-Eiras and Niepelt, 2020) or/and an uncertainty about aggregate dynamics (e.g., Eichenbaum et. al., 2020; Kaplan et. al., 2020). Because of these limitations of macro-epidemic models, policymakers are forced to rely on "mechanistic" SIRD models. Those models preclude analysis how people react to changing macro-epidemic variables and expectations about future government policies.

This thesis aims to contribute to macro-epidemic literature in two main directions. Firstly, it aims to provide an accurate and scalable solution method for general recursive formulation of macro-epidemic models that would allow to analyze macro-epidemic dynamics in realistic environments featuring complex economic system together with aggregate uncertainty. Secondly, it aims to facilitate optimal policy computation by leveraging the "manifold-learning" approach of Duarte (2018) for simultaneously solving the model for all relevant combination of government policy parameters and hence reducing the optimal policy computation into a simple optimization problem. To the best of my knowledge, this thesis would be the first paper, which applies deep learning based projection method for solving for competitive equilibria, and optimal policies in macro-epidemic models.

Research Question and Contribution:

My thesis would contribute to the macro-epidemic literature in two directions. Firstly, I will develop an efficient deep learning solution method for recursive macro-epidemic models. Relative to the existing literature (e.g., Gonzalez-Eiras and Niepelt, 2020; Garibaldi and Pissarides, 2020) which rely on closed-form solutions or linear projection methods in a spirit of Judd (1998), my method would allow to analyze richer economies featuring more complex epidemic processes and economic structure. Secondly, I will utilize deep learning to accelerate computation of optimal government policies. Instead of re-solving the model many times for different parameterizations of government policies, I will use deep learning to approximate a whole equilibrium manifold indexed by policy parameters. This pre-computation in a spirit of Duarte (2018) should reduce the runtime of optimal policy computations by transforming the search for optimal policy parameters into a standard numerical optimization of readily-available function.

Methodology:
I will extend the deep learning solution method of Maliar et. al. (2019) to accommodate peculiarities of macro-epidemic model, specifically absence of non-degenerate stationary distribution of state variables, which precludes usage of the ergodic grid method of Maliar et. al. (2019). Besides that, I will show how the deep learning method could be used to speed-up optimal policy computations by treating policy parameters as additional "pseudo-states" included in the network.

To provide an use case of my method, I will extend the benchmark macro-epidemic model of Eichenbaum, Rebelo, and Trabandt (2020) by providing its recursive formulation. Besides solving the basic model, I will demonstrate broad applicability of my method by solving a stochastic version of the model featuring stochastic changes in the epidemic characteristics (mutations) and solving for the optimal containment policy using the manifold learning approach. Because of data availability, I will focus the empirical part of my thesis on the case of the Czech Republic. For calibration of model parameters I will use data provided by the Czech Ministry of Health, the Czech National Bank, and the Czech Statistical Office.

Outline:

1. Introduction
2. Literature
3. Stylized Facts
4. Model
    - Private Sector
    - Epidemic Dynamics
    - Government
    - Recursive Equilibrium
    - Policy Problem
5. Calibration
6. Numerical Solution
7. Results
8. Conclusion

Contents

Acknowledgments

I would like to express my gratitude to my supervisor Ctirad Slavík for his guidance and support during writing the thesis. Besides him, I am grateful to Marek Kapička for important suggestions, and to Filip Matějka for support. I also want to thank Lukáš Supík and Martin Kosík for being my companions during our MA studies. Finally, I want to thank my family for their patience and unconditional support.

Bibliographic note

ŽEMLIČKA, Jan. *Macro-Epidemic Modeling: A Deep Learning Approach.* Prague 2021. X pages. Master thesis. Center for Economic Research and Graduate Education - Economics Institute, A joint workplace of Charles University and the Economics Institute of the Czech Academy of Sciences. Thesis supervisor: PhDr. Mgr. Ctirad Slavík, Ph.D.

Abstract

I develop a novel method for computing globally accurate solutions to recursive macro-epidemic models featuring aggregate uncertainty and a potentially large number of state variables. Compared to the previous literature which either restricts attention to perfect-foresight economies amendable to sequence-space representation or focuses on highly simplified, low dimensional models that could can be analyzed using standard dynamic programming and linear projection techniques, I develop a deep learning-based algorithm that can handle rich environments featuring both aggregate uncertainty and large numbers of state variables. In addition to solving for particular model equilibria, I show how the deep learning method could be extended to solve for a whole set of models, indexed by the parameters of government policy. By pre-computing the whole equilibrium set, my deep learning method greatly simplifies computation of optimal policies, since it bypasses the need to re-solve the model for many different values of policy parameters.

Key Words Macro-Epidemic Model, Recursive Equilibrium, Aggregate Risk, Projection, Deep Learning

Abstrakt

V této práci prezentuji novou metodu pro výpočet globálně přesných řešení rekurzivních stochastických makro-epidemických modelů s potenciálně vysoko-dimenzionálním stavovým prostorem. V porovnání s existujícími studiemi, které budto studují deterministické ekonomiky pomocí sekvenčních metod, nebo analyzují stylizované modely řešitelné standardními metodami dynamického programování a lineární projekce, v této práci aplikuji algoritmus založený na hlubokém učení, který umožňuje analyzovat komplexní ekonomiky s agregátní nejistotou a velkým počtem stavových proměnných. Kromě řešení modelu vůči dané hodnotě parametrů prezentuji též rozšířený algoritmus, který umožňuje vyřešit celou množinu modelů indexovanou parametry reakční funkce vlády. Tento krok tak výrazně zjednodušuje výpočet optimální reakční funkce vlády, jelikož obchází nutnost opakovaného řešení modelu pro různé parametrizace vládní reakční funkce.

Klíčová slova: Makro-epidemický model, rekurzivní equilibrium, agregátní riziko, projekce, hluboké učení

Introduction [1]

With the outbreak of SARS-CoV-2 in December 2019, the issue of integrated macro-epidemic policy became a central topic of macroeconomic debate. Starting with the pioneering contribution of Eichenbaum, Rebelo and Trabandt (2020), the macroeconomics profession hurried to develop nascent literature on joint macro-epidemic dynamics. Important progress has been made both on the side of the positive analysis of interactions between macroeconomic variables, government policies, and the state of epidemics, as well as the normative side of determining welfare-maximizing government policies during an epidemic.

On the positive side, Kaplan, Moll and Violante (2020) studied the epidemic health-economy tradeoff in a complex New Keynesian economy model with heterogeneous agents (HANK) with wealth and income inequality, liquid and illiquid assets, and occupational and sectoral heterogeneity, and with limited capacity in intensive care units. Glover, Heathcote, Krueger and Rıos-Rull (2020) used an overlapping generations heterogeneous agents economy to study an inter-generational nexus of epidemic tradeoffs. Guerrieri, Lorenzoni, Straub and Werning (2020) used the multisectoral incomplete markets model to show, how epidemic-induced negative supply shocks might cause shortfalls in aggregate demand and generate a classical Keynesian recession.

On a normative site, Gonzalez-Eiras and Niepelt (2020) and Moser and Yared (2020) analyzed the problem of the social planner in economies during epidemics and derived corresponding optimality conditions characterizing the first-best allocations. By characterizing the first-best allocation, these papers quantify the welfare costs of externalities associated with consumption and labor activities of private households and provide an upper bound on achievable welfare gains using Ramsey-style distortionary policy instruments. The Ramsey problem in epidemics economy was firstly analyzed by Eichenbaum, Rebelo and Trabandt (2020) who employed numerical

[1]Previous versions of this paper were submitted to *Research Writing II*, *Labor Economics I*, *Public Finance*, *Combined Skills I* and *Macro Topics II* at CERGE-EI during the Fall and Spring of 2020/2021. This thesis also utilizes my GAUK submission material from the Fall of 2020.

optimization to solve the policy problem in sequence space. Finally, Moser and Yared (2020) explored the role of government commitment in a simple, three-period model with perfect foresight.

Although this literature has provided many important qualitative and quantitative insights about tradeoffs faced by a society confronted with an epidemic outbreak, the scope of their results remains limited, chiefly due to computational constraints. The dominant strand of the literature assumes away aggregate uncertainty in order to solve their models using tractable sequence-space techniques. While some of these papers analyze quite complex environments featuring many salient features of the real-world economy, the assumption of a deterministic aggregate could severely alter the dynamics of their model economies, and especially their welfare implications [2]. Because the behavior of welfare in the model is tightly linked to the issue of optimal policies, possible misinterpretation of welfare caused by the certainty-equivalence assumption could lead to incorrect policy advice.

To the best of my knowledge, the only paper that analyzes a macro-epidemic model with aggregate uncertainty using suitable recursive methods [3] is Gonzalez-Eiras and Niepelt (2020). Gonzalez-Eiras and Niepelt (2020) solve for the recursive competitive equilibria and the first-best allocation in a continuous-time economy featuring stochastic vaccine arrival and mutation risk. However, computational considerations forced them to simplify their economic environment a great deal, so they abstract from agent heterogeneity, constraints on the capacity of intensive care units, etc., so their study cannot serve as a reliable quantitative laboratory for performing policy experiments.

The aim of this thesis is to develop an efficient and highly scalable method for computing potentially high-dimensional recursive

[2]This fact is has been recognized in the business cycle literature at least since J. Kim and S.H. Kim (2003). Since the macro-epidemic environment features are significantly nonlinear, there are reasons to worry about a loss of accuracy caused by the certainty-equivalence approximation which is implicitly applied by the dominant strand of the macro-epidemics literature

[3]Recursive formulation greatly facilitates analysis of environments with aggregate uncertainty and government commitment issues. However, these methods tend to be much computationally expensive, complex to implement, and suffer from the curse of dimensionality.

equilibria, which would bypass the need to make a choice between rich, but deterministic economies and highly simplified stochastic "toy models". My contribution is twofold. Firstly, I develop a baseline algorithm for computing a globally accurate approximation to the model equilibria in a spirit of projection methods, à la Judd et al. (1991), designed to overcome both the curse of dimensionality and the numerical challenges presented by the highly nonlinear nature of macro-epidemic models. Although the existing quantitative macroeconomics literature proposes various approaches for solving high-dimensional recursive models, including the Smolyak interpolation (Krueger and Kubler 2004), adaptive sparse grid methods (Brumm and Scheidegger 2017,; Eftekhari and Scheidegger 2020), and deep learning projection (L. Maliar, S. Maliar and Winant 2019; Azinovic, Gaegauf and Scheidegger 2021), this literature has focused on business cycles and balanced growth path economies, and hence is not well suited to deal with highly nonlinear dynamics on non-hypercubic domains which are ubiquitous features in macro-epidemic models.

My solution method is based on the deep learning $\mathbb{L}^2$ projection method of L. Maliar, S. Maliar and Winant (2019). My contribution with respect to their paper is an adaptation of the algorithm to accommodate difficulties presented by macro-epidemic models. Specifically, I show how to improve the neural network approximator used in the deep learning method by selecting an efficient network structure and by constructing appropriate boundary conditions which exploit steady-state convergence of epidemic models to accelerate the process of computing the optimal approximation. Also, I discuss various alternatives for generation of the state-space grid on which the model should be solved.

Secondly, building on the baseline solution algorithm and the related approach of Duarte (2018a) [4], I develop a novel method for simplifying optimal policy computations using deep learning. Instead of solving the model for one particular value of the parameter vector, I include a subset of model parameters that index the government policy as an additional set of state variables, and

[4] Duarte (2018a), building on the previous contribution of Norets (2012), use related approach in context of model estimation. To the best of my knowledge, to this date, no paper used this algorithm for optimal policy computation.

train the neural network to solve the model on the expanded state space. This procedure simultaneously solves for a whole set of models indexed by the selected vector of policy rule parameters. The process of finding the optimal value of government policy parameters is then reduced to a simple numerical optimization of a function with closed-form expression [5], as opposed to a standard procedure in the spirit of Lippi, Ragni and Trachter (2015), in which the model is repeatedly solved for many different parameterizations of the policy rule [6]. Since neural networks are well known for their asymptotic efficiency (Barron 1993) in the representation of complex objects in high-dimensional spaces, my state-augmentation algorithm offers a potentially significant speed-up when solving for complex policy rules.

[5] The extended algorithm approximates all equilibrium objects, including value functions, as explicit functions of both natural state variables and policy rule parameters. The standard parametric Ramsey problem amounts to selecting policy parameterization which maximizes the total weighted welfare of households in the economy. In my framework, the Ramsey problem could be solved after the algorithm solves for the model set just by fixing the initial state and then using a numerical optimization algorithm to find a value of policy vector which maximize the total welfare, subject to some given set of Pareto weights. Because the algorithm solves the model for a whole set of possible parameterizations of the policy rule, this solution allows for computing the welfare as a closed-form function of policy parameters and possibly some given Pareto weights. However, neural networks are highly-nonlinear functions, hence the optimal value of policy vector had to be found using numerical routine.

[6] Availability of closed-form expression for welfare in terms of policy parameters became allow for utilization of both first and second-order optimization methods which are typically unfeasible if the repeated solution approach is utilized. The key difference is that with closed-form expression in form of a neural network, both gradient vector and Hessian matrix could be efficiently obtained using automatic differentiation (backpropagation), whereas finite-difference computation of gradient using the standard approach requires solving the economy $N + 1$ times, where N denotes the number of policy rule parameters. Because of this bottleneck, the optimal policy literature is biased towards derivative-free or at most first-order optimization methods, because gradient and especially hessian matrix computation are typically far too costly.

1 Related Literature

1.1 Macro-Epidemic Modeling

The literature on joint modeling of macroeconomic and epidemic dynamics is quite a recent one. Presumably, since full-blown epidemics of deadly pathogen did not happen in developed countries for nearly hundred of years, modeling of the interaction between epidemic dynamics, the behavior of modern industrialized economies and government policies was totally eclipsed by classical topics of macroeconomics research, such as business cycles, economic growth, or optimal fiscal and monetary policies, to name a few. To the best of my knowledge, the very first paper that merged standard optimizing general equilibrium macroeconomics framework with dynamic epidemic model à la Kermack and McKendrick (1927) is Eichenbaum, Rebelo and Trabandt (2020), which was written few months after the December 2019 coronavirus epidemics outbreak.

The model of Eichenbaum, Rebelo and Trabandt (2020) is admittedly a simple one. Yet, it captures main qualitative insights about the nature of epidemics externality of consumption and work behavior of households in competitive equilibrium and provides a basic case for government intervention in form of consumption and work repression aimed to slow down dynamics of epidemics. It contributed to the macro-epidemics literature both on the positive side by building a first competitive equilibrium model of economy-epidemics interactions, and to the normative side of the literature by analyzing an optimal policy problem in which government aims to control epidemic dynamics by imposing restrictive, lockdown-like measures. Compared to some further papers which used more advanced techniques to analyze social planner problems (Alvarez, Argente and Lippi 2020; Gonzalez-Eiras and Niepelt 2020), the value of Eichenbaum, Rebelo and Trabandt (2020) is that they solved a Ramsey problem in which the government had access to a limited set of potentially distortive instruments, whereas the social planner has access to all possible instruments [7]. Extended versions of their model include state-dependent mortality rate which aims to capture

[7] Even more than Mirleesian planner, since classical Pareto social planner does not have to deal with private information and associated incentive constraints.

health care capacity constraints, uncertain vaccination arrival [8] and re-infection possibility.

The key shortcoming of the analysis presented by Eichenbaum, Rebelo and Trabandt (2020), besides relative simplicity of their economic environment, is the sequence-space formulation which precludes analysis of environments featuring aggregate risk, such as stochastic mutations which could change virus transmission rate, or potentially more complicated optimal policy problems, such as problems featuring government complex time-consistency problems (eg. Clymo, Lanteri and Villa 2020). Moreover, computation of Ramsey-optimal policies in sequence-space is computationally extremely expensive, since it amounts to running numerical optimization algorithm over very highly-dimensional parameter vector [9]. For example, benchmark optimal policy exercise in Eichenbaum, Rebelo and Trabandt (2020) involves optimization over vector of consumption tax rates with 250 elements. Such an optimization problem is both highly computationally challenging [10] and also theoretically challenging, since optimization problems can be ill-conditioned and potentially non-convex. Recursive methods typically offer away, how to bypass a large number of aforementioned problems.

Alvarez, Argente and Lippi (2020) and Gonzalez-Eiras and Niepelt (2020) analyzed the epidemics externality by comparing social welfare implied by competitive equilibria allocations and Pareto social planner solutions. Both of these papers employed recursive continuous-time techniques à la Achdou et al. (2020) and Phelan and Eslami (2021) to formulate and solve social planner problems in relatively simple epidemic economies. While Alvarez, Argente and Lippi (2020) focuses on the deterministic environment, later paper by Gonzalez-Eiras and Niepelt (2020) analyzes social planner problems in the economy with aggregate mutation risk and stochastic vaccine arrival.

[8] Vaccination is modeled as a permanent, one-off event which is hence amenable to analysis using sequence-space methods employed by Eichenbaum, Rebelo and Trabandt (2020)

[9] Because government policies in sequence-space are sequences of per-period policies. Furthermore, most macroeconomic models, including epidemic ones, are set in infinite-horizon environments, and hence had to be approximated using long parameter sequences.

[10] Because efficient optimization methods typically require computation or at least approximation of hessian matrix with respect to policy parameters. Hessian computation has typically quadratic computational complexity, moreover, typically it had to be inverted in each iteration, and matrix-inversion has cubic complexity.

While these papers bring important methodological contributions and some valuable qualitative insights, the generality of their results is severely restricted by the simplicity of considered environments [11] and also by their focus on the analysis of Pareto planner problems, whose relevancy to real-world policymaking remains unclear.

Besides the standard Ramsey literature, Moser and Yared (2020) examined the issue of government commitment for efficiency of epidemic-containment measures. They found that the government trying to control the epidemic in a dynamic economy faces a time consistency problem which mirrors similar results appearing in optimal dynamic taxation literature (Chari and Kehoe 1990; Chang 1998). The key shortcoming of their analysis is the analytical, three-period framework which by definition precludes any meaningful quantitative analysis. Up to this date, a rich quantitative model of the time consistency problem remains one of the key blind spots of the normative macro-epidemics literature.

Kaplan, Moll and Violante (2020) and Glover, Heathcote, Krueger and Rıos-Rull (2020) studied macro-epidemic dynamics in heterogeneous agent economies without aggregate risk [12]. Glover, Heathcote, Krueger and Rıos-Rull (2020) focused on a continuous-time heterogeneous agent economy featuring overlapping generations structure which allows for analyzing the intergenerational aspects of the epidemic crisis. In the second part of their paper, Glover, Heathcote, Krueger and Rıos-Rull (2020) used numerical optimization to compute optimal containment policies coupled with intergenerational cash transfers. They found that welfare-maximizing government policy complements restrictive epidemic containment measures with large cash transfers from old to young households. The economic fundament behind these transfers is the fact that young households tend to be more hurt by restrictive measures, and thanks to their better health status, they do not have significant welfare gains from the improved epidemic situation. On the other hand, old households

[11] They consider even simpler economies than Eichenbaum, Rebelo and Trabandt (2020). The economy of Eichenbaum, Rebelo and Trabandt (2020) feature three continuous aggregate state variables, whereas Alvarez, Argente and Lippi (2020) and Gonzalez-Eiras and Niepelt (2020) consider economies with at most two continuous state variables

[12] This assumption allows them to solve for aggregate dynamics using sequence-space techniques

are vulnerable to the disease, and hence restrictive government policy improves their welfare quite significantly. The government could use these resource transfers to compensate young households for the restrictions which hurt the young generation disproportionately more and make old households to pay [13] for the benefit of improving the epidemic situation.

Kaplan, Moll and Violante (2020) analyzed macro-epidemic dynamics in a rich heterogeneous agents New Keynesian economy with wealth and income heterogeneity, nominal frictions, multiple sectors, and incomplete markets. Kaplan, Moll and Violante (2020) used their quantitative framework is a numerical characterization of the so-called pandemics possibility frontier which describes the tradeoff between lost economic welfare and the number of deceased households. In complementary work, Guerrieri, Lorenzoni, Straub and Werning (2020) studied effects of epidemics shocks to a complex multisectoral economy and showed that multisectoral structure coupled with incomplete market environment might under realistic calibration of model parameters lead to a situation in which negative supply shock counterintuitively cause also negative aggregate demand shock.

1.2 Solving Recursive Models

Since macroeconomic models typically do not possess closed-form solutions, the system of functional equations [14] which defines recursive equilibria of the economy of interest had to be approximately solved using functional approximation techniques. At least since Magill (1977) and Kydland and Prescott (1982), macroeconomics literature developed a large number of computational methods for solving recursive macroeconomic models. In principle, all of these methods could be used for approximating the solution of macro-epidemic models. These methods could be broadly classified into three categories; local perturbation techniques which use implicit function theorems to build a series approximation around some de-

[13] This result could be used as an argument for inflationary post-epidemics monetary policy, which tends to redistribute real resources from relatively wealthy old households who tends to hold a large amount of nominal financial assets towards younger households who in turn tends to have large nominal debts, such as mortgage balances.

[14] These equations include; bellman equations in discrete time, or HJB equations in continuous time, first-order conditions, no-arbitrage conditions, market-clearing conditions, etc.

generate solution, finite difference methods for solving differential equation models in continuous time and general projection methods which use functional analysis ideas to build computable global approximations to unknown functions.

Perturbation methods approximate the solution of a recursive model by a series expansion around some fixed point in a state-space. The key advantage of this approach is computational efficiency. Perturbation methods allow for quick solution of large-scale models with the minimal amount of computational resources (Aruoba, Fernandez-Villaverde and Rubio-Ramirez 2006). However, perturbation methods are limited by their local nature: they provide high accuracy around one point, but the accuracy may deteriorate quickly as the economy moves further away from approximation point Fernández-Villaverde and Levintal (2018). Since dynamics of macro-epidemic models include long transition paths instead of just small fluctuations around a steady-state, application of perturbation methods on macro-epidemic models leads to unacceptably large numerical errors.

Finite-difference methods constitute a separate class of solution methods, distinct from both perturbation methods which are based on local series expansions, and from projection methods that adopt functional analysis-inspired approach (Judd 1998, Chapter 11). Finite-difference methods discretize ordinary or partial differential equations into a finite set of equations by taking some version of grid-based finite difference approximation to differential operators present in the differential equation. Various finite difference methods differ in their choices of finite difference approximation, grid generation schemes, and algorithms for solving the resulting set of equations. The most influential finite-difference method used for solving recursive macroeconomic models is the implicit upwind scheme introduced into the macroeconomics literature by Achdou et al. (2020). Since this thesis deals with macro-epidemic models formulated in a discrete-time framework, finite-difference methods might seem less relevant compared to perturbation and projection methods. However, they could be a highly efficient tool for solving continuous-time versions of these models.

Projection methods constitute a traditional complement of perturbation methods in a toolkit of quantitative macroeconomics (Fernández-Villaverde, Rubio-Ramırez and Schorfheide 2016). Most often, they took form of pseudospectral (Boyd 2001) or finite element methods (Brenner and Scott 2007). These methods approach the problem of solving a functional equation by specifying a linear subspace of basis functions [15] and projecting the unknown function against the subspace by minimizing some loss function containing weighted residual of the equation [16]. The advantage of the projection approach is its ability to provide an arbitrarily high degree of accuracy on the arbitrarily large area of state-space. Also, projection methods are able to deal with challenging problems such as models with occasionally binding constraints or non-convexities that induce kinks and discontinuities into a model solution.

The main obstacle for the application of projection methods to the solution of macro-epidemic models is the curse of dimensionality. The computational complexity of standard tensor product-based projection methods grows exponentially with dimensionality of approximated function (Fernández-Villaverde, Rubio-Ramırez and Schorfheide 2016). This explosive growth of computational complexity precludes the application of standard projection methods to models whose state-space includes more than 3-4 state variables. Since a bare-bone SIRD model includes 3 state variables, just a "toy" version of the macro-epidemic model pushes the standard projection approach to its very limit.

During the last two decades, macroeconomic literature developed several approaches that extend classical projection methods to higher-dimensional spaces. Those methods are based on sparse grids (Krueger and Kubler 2004), adaptive sparse grids (Brumm and Scheidegger 2017) and ergodic set grids (L. Maliar and S. Maliar 2015). While methods developed in this literature constitute an important contribution to the toolkit of methods available for analysis

[15] While the subspace of basis functions has a linear structure, basis functions themselves are arbitrary nonlinear functions.

[16] Different choices of basis functions and weighting scheme leads to different projection methods. The most common approach in the macroeconomic literature is Chebyshev collocation that combines basis composed of Chebyshev orthogonal polynomials with Dirac weighting on Chebyshev zeros.

of recursive macroeconomic models, their applicability is still severely limited by a number of important bottlenecks. Sparse grid methods are limited to models with fewer than 30 continuous dimensions and smooth solutions. Ergodic set methods fail in models where dynamics of the economy is not bounded within a relatively small region of the state space (Fernández-Villaverde and Levintal 2018). Adaptive sparse grids allow for non-smooth functions and less regular dynamics, but they still face limitations when the dimensionality of the non-smooth model exceeds approximately 20 variables (Brumm and Scheidegger 2017).

In the last two years, a new strand of computational macroeconomic literature emerged. Building on recent groundbreaking advances in machine learning-based solution methods for highly dimensional partial differential equations in the field of computational physics (Sirignano and Spiliopoulos (2018), Raissi, Perdikaris and Karniadakis (2019)), starting with papers by Duarte (2018b), Villa and Valaitis (2019) and L. Maliar, S. Maliar and Winant (2019), macroeconomic literature began to develop neural networks based methods for solving recursive models.

Within the class of projection methods, the neural network-based "deep learning" constitutes an alternative to pseudo-spectral and finite element methods (Fernández-Villaverde, Nuño, Sorg-Langhans and Vogler 2020). Instead of approximating the target function as a linear combination of nonlinear basis functions, the deep neural network approach rely on a nonlinear combination of nonlinear basis functions. By building the structure of connected "hidden" layers of sigmoid, rectified linear (ReLu) or other suitable basis functions, these approximators gain unparalleled flexibility in approximation of complex and high dimensional objects. The empirical success of neural networks in approximating highly dimensional functions (Silver et al. 2016) is predicted by a number of theoretical results in the approximation theory. Hornik, Stinchcombe, White et al. (1989) proved, that neural networks can approximate any Borel measurable function mapping one finitely dimensional space into another up to an arbitrary degree of accuracy, as the size of the network grows. Poggio et al. (2017) survey the literature on how deep neural networks escape the curse of dimensionality for a broad class of approximation problems.

Within the nascent macroeconomic literature, Duarte (2018b) pioneered the application of machine learning techniques for solving recursive models. His approach has two main limitations. Firstly, he focuses on continuous-time problems exclusively. While continuous-time environment offers a large number of significant computational advantages (Achdou et al. 2020), the majority of macroeconomic research is conducted in discrete-time environments, and this fact might limit the applicability of Duarte's approach. Secondly, the application of his method requires dynamic programming representation of the problem of interest. Hence, its applicability outside the class of planner and partial equilibrium problems is quite limited.

Fernández-Villaverde, Nuño, Sorg-Langhans and Vogler (2020) developed an interesting extension of the continuous-time dynamic programming approach. Relative to Duarte (2018b), they exploited information contained in first-order conditions to accelerate the convergence of their algorithm. Thank to this improvement, and application of Itô's lemma for the elimination of expectation operator, they were able to solve highly nonlinear dynamic programming problems up to 75 dimensions. In a related paper, Fernández-Villaverde, Hurtado and Nuno (2020) developed an extension of the celebrated method of Krusell and Smith (1998) for computation of Bewley models with aggregate risk. They replaced linear approximation of the perceived aggregate law of motion in the macro-finance version of Krusell-Smith economy with the neural network and demonstrated that neural network offers superior performance and accuracy in the highly nonlinear environment of their macro-finance economy.

Villa and Valaitis (2019) contributed to the macro-machine learning literature by extending the parameterized expectations method of Den Haan and Marcet (1990) by approximating the conditional expectations in dynamic equations of recursive macroeconomic models by shallow [17] neural networks. Their contribution is twofold. Firstly, they provide quite a general method for solving discrete-time models that do not rely on straightforward dynamic programming representation. Secondly, they showed that machine learning could be used to

[17]Neural network with one hidden layer.

solve equations characterizing Ramsey-optimal [18] policy in dynamic stochastic environments. The key shortcoming of their paper is a technological one. Instead of relying on industry-quality machine learning libraries, such as TensorFlow or Flux.jl, they build their own Matlab code that uses outdated shallow networks and trains them using Levenberg-Marquardt algorithm instead of standard stochastic gradient descent methods. Because of these technical bottlenecks, the method of Villa and Valaitis (2019) is not well suited to benefit from the possibility of massive parallelization on clusters of graphical processing units (GPUs) or machine learning accelerators such as Google's TPUs.

L. Maliar, S. Maliar and Winant (2019) and Azinovic, Gaegauf and Scheidegger (2021) presented two related discrete-time approaches that are closest to the methodology proposed by this project. Azinovic, Gaegauf and Scheidegger (2021) developed a highly parallel deep learning methodology for the solution of large-scale stochastic overlapping generations models. Their method approaches the problem of approximating the solution of a recursive model by directly replacing functions of interest with neural networks and training these networks on data iteratively simulated from the ergodic set of the model. They were able to solve a non-smooth stochastic OLG model up to 292 state variables, an unparalleled achievement in the computational macroeconomics literature. The potential downside of this paper is its strong focus on the utilization of supercomputers available to the authors. While their method was able to solve problems of impressive complexity, it is not clear, how these results could be replicated by researchers without access to supercomputing resources.

The key contribution of L. Maliar, S. Maliar and Winant (2019) is a very general methodology for formulating traditional [19] recursive macroeconomics models as objective functions for machine learning libraries, such as TensorFlow. It offers three distinct objectives: euler or bellman equation errors minimization and alternative approach in terms of lifetime discounted utility maximization. Their showcase is a 1000 agent finite version of Krusell and Smith (1998) economy

[18] Full commitment

[19] Models with finitely many agent types, as opposed to Bewley economies featuring a continuum of heterogeneous agents.

which leads to state space with 2001 dimensions. Because of the relatively low degree of nonlinearity of this model, authors were able to solve it on a standard laptop in a few hours. Moreover, they provide a user-friendly interface: the Dolo modeling language, that relieves the end-user from the necessity to deal with low-level programming and construction of objective function. While this paper constitutes a very useful contribution to the toolkit of computational macroeconomics, it doesn't provide the right tools necessary for an analysis of macro-epidemic models. Firstly, it utilizes the ergodic set approach for the construction of the training grid, and hence, it faces crucial difficulties when approximating the economy along transition trajectories characteristic for SIRD-based models. Secondly, this approach remains silent about optimal policy computations.

2 Solution Algorithm

2.1 Deep Learning

This thesis utilizes a nonlinear projection algorithm (deep learning) to develop a new method for solving recursive macro-epidemic models. The basic structure of deep learning algorithm broadly follows classical collocation and pseudo-spectral linear projection techniques à la Judd et al. (1991) and Boyd (2001). The key difference with respect to those linear projection methods is the choice of projection basis. Whereas linear projection methods employ a linear combination of nonlinear basis function, such as orthogonal polynomials (Chebyshev, Hermite, etc.), deep learning algorithms utilize a nonlinear combination of nonlinear functions, specifically the deep neural network architecture.

While neural networks came in many different flavors, I focus on simple fully connected feed-forward neural networks. The basic building block of a neural network is an artificial neuron, which is basically a function, which takes a linear combination of input values and transforms that linear combination using some function. Let θ_i denote parameters of linear combination and σ denote a nonlinear "activation" function. Then, an artificial neuron could be written as.

$$\mathcal{N}(x, \theta) = \sigma\left(\theta_0 + \sum_{i=1}^{N} \theta_i x_i\right) \tag{1}$$

The artificial neuron is a mapping from some $\mathbb{R}^N$ to some possibly restricted set of $\mathbb{R}$. Artificial neurons are grouped into layers, where layer represent a mapping from $\mathbb{R}^N$ to $\mathbb{R}^M$

$$\mathcal{H}(X, \Theta) = \big(\mathcal{N}(x, \theta^{(1)}), ..., \mathcal{N}(x, \theta^{(K)})\big) \tag{2}$$

The key idea behind deep neural networks is a composition of multiple (nonlinear) layers, so the output of one layer serves as the input of another, so the final output goes through multiple stages of parameterized nonlinear transformations. This structure gives deep neural networks immense capacity [20] for approximating complex

[20]Because the composition of nonlinear functions is a much richer operation than a simple addition (Fernández-Villaverde, Nuño, Sorg-Langhans and Vogler 2020)

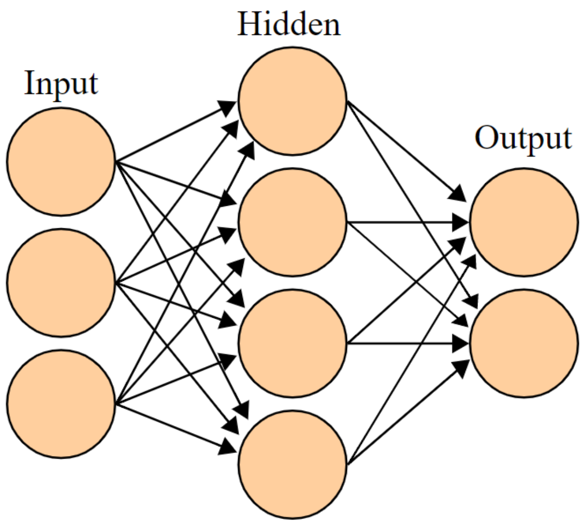

Figure 1: Simple FFN $\mathbb{R}^3 \to \mathbb{R}^2$ with one hidden layer

high-dimensional functions. In general, a deep neural network should be thought of simply as a parametric function indexed by a coefficient matrix $\mathbb{O}$.

$$\mathcal{NN}(X, \mathbb{O}) = \mathcal{H}(\mathcal{H}(\mathcal{H}(...\mathcal{H}(X, \Theta^{(1)})..., \Theta^{(L-2)}), \Theta^{(L-1)}), \Theta^{(L)}) \quad (3)$$

The general idea, how deep neural networks could be used to solve functional equations, such as recursive equilibria of macroeconomic models mirrors the linear projection approaches. Let $\mathcal{M}$ denote the set of model equations, $\mathbb{X}$ the state-space, $F(X\acute{u}$ set of unknown functions. Because the original problem is a general functional equation, it is typically an infinitely-dimensional object, which can not be tackled directly.

$$\mathcal{M}(F(x)) = 0 \qquad \forall X \in \mathbb{X} \quad (4)$$

To transform it into a computable finitely-dimensional problem, we had to approximate the unknown function $F(X)$ by some parametric approximator (ansatz), in this case a deep neural network $\mathcal{NN}(X, \mathbb{O})$. Plugging the neural network approximator into the set of model equations form an error function.

$$\mathcal{E}(X, \mathbb{O}) = \mathcal{M}(\mathcal{NN}(X, \mathbb{O})) \quad (5)$$

Intuitively, the solution to the functional equation should be a parameter matrix $\mathbb{O}^*$ for which $\mathcal{E} \approx 0 \quad \forall X \in \mathbb{X}$. However, $F(X)$, in general, has infinitely many degrees of freedom, and hence can not be exactly characterized by a finitely-dimensional parametric family. Instead, the projection approach builds a loss function, which is a weighted integral of the error function over the state-space.

$$\mathbb{L}(\mathbb{O}) = \int_{\mathbb{X}} \mathcal{W}\left(\mathcal{E}(X, \mathbb{O}), X\right) dX \tag{6}$$

The optimal value of the parameter matrix $\mathbb{O}$ is a minimizer of the loss function.

$$\mathbb{O}^* = argmin_{\mathbb{O}} \quad \mathbb{L}(\mathbb{O}) \tag{7}$$

While in theory many different loss function structures could be employed, the deep learning literature typically utilizes a simple mean squared error (MSE) approach, because the Galerkin type of scheme is infeasible due to the nonlinear structure of neural networks. The MSE loss function is simply an average squared residual of model equations over a large enough sample of states, denoted as $\tilde{X}$. This sample is typically obtained by means of Monte Carlo simulations.

$$\mathbb{L}^2(\tilde{X}, \mathbb{O}) = \frac{1}{N} \sum_{i=1}^{K} \mathcal{E}(\tilde{X}_i, \mathbb{O})^2 \tag{8}$$

2.2 Loss Function and Grid Sampling

My approach for the construction of the loss function broadly follows the euler equation method L. Maliar, S. Maliar and Winant (2019) and the algorithm of Azinovic, Gaegauf and Scheidegger (2021). The loss function composes of mean squared residuals in the first-order conditions, bellman equations, and constraint-violation error [21]. This form of loss function allows for directly attacking the model recursive equilibria without the necessity for engaging in some type of fixed-point iterative procedure.

L. Maliar, S. Maliar and Winant (2019) and Azinovic, Gaegauf and Scheidegger (2021) solve their model on random grids which are sampled from the ergodic distribution implied by approximate model solutions. L. Maliar, S. Maliar and Winant (2019), building on

[21] Especially non-negativity errors.

previous work of L. Maliar and S. Maliar (2015), discussed how those ergodic set methods help to mitigate the curse of dimensionality. Unfortunately, macro-epidemic economies in a spirit of Eichenbaum, Rebelo and Trabandt (2020) feature highly nonlinear dynamics that do not possess an ergodic steady state, hence the ergodic set method can not be used. There are two possible responses to this challenge.

Firstly, one might simulate a large number of transition paths of the economy, and use those transition paths as a projection grid. However, the wild nonlinearity of macro-epidemic models requires great care for preventing numerical instabilities during the training process. Secondly, one might use uniform hypercube sampling and try to limit the size to the necessary volume. In my algorithm, I use the second approach, specifically, my method aims to truncate the size of the hypercube in the direction of state variables which induce large nonlinearities into the model solution.

3 Benchmark Implementation

As a benchmark implementation of my method, I select the influential macro-epidemic model of Eichenbaum, Rebelo and Trabandt (2020). Although this economy might be considered as quite stylized, even its standard configuration includes 3 continuous state variables, and to the best of my knowledge, no paper up to this time solved fully recursive macro-epidemic model with more than two state variables [22]. Moreover, its relative simplicity is useful for a detailed exposition of inner details of the implementation of my method.

In this chapter, I use my algorithm to solve four different modifications of the model of Eichenbaum, Rebelo and Trabandt (2020). Firstly, I solve the simple laissez-faire version of their model for which a large part of the equilibrium could be solved in a closed form and use this result to check results provided by a general algorithm that numerically approximates whole recursive equilibria of the economy. After solving the baseline economy of Eichenbaum, Rebelo and Trabandt (2020), I use a slightly extended algorithm to solve a version of their model featuring aggregate stochastic virus mutations. Finally, I demonstrate the flexibility of my deep learning algorithm by solving a whole set of model equilibria indexed by parameters controlling government epidemic containment policies and show, how this algorithm could be used to simplify the computation of optimal government policies in macro-epidemic economies [23].

3.1 Economy of Eichenbaum, Rebelo and Trabandt (2020)

3.1.1 Model Economy

Time is discrete and infinite. The economy is populated by a unit measure of households who are divided into three groups; susceptible households which haven't been infected yet, infected households who

[22] Gonzalez-Eiras and Niepelt (2020) solved epidemic model with stochastic vaccine arrival and mutation risk, however, they model these variables as discrete Poisson processes. Their model feature only one continuous state variable. Alvarez, Argente and Lippi (2020) solved a recursive planner problem in a deterministic epidemics economy with two state variables by converting the continuous-time problem into a discrete-time and solving the resulting bellman equation using numerical value function iteration.

[23] I believe that this method could be also useful for optimal policy computations in the standard business cycle and growth models.

can further spread the disease and face the probability of dying to the disease, and recovered households who survived the disease and are no longer infectious. Susceptible households are denoted by s superscript, infected by i superscript and recovered by r superscript. The aggregate state of the economy consists of population shares of susceptible, infected, and recovered households [24]. Share of susceptible households at time t is denoted by S_t, share of infected households by I_t, and share of recovered households by R_t. The residual which represents deceased households is denoted by D_t.

The aggregate state evolves to the following system of "SIR" equations, where T_t denote share of newly infected households, C_t^s and C_t^i denote consumption of susceptible and infected households, N_t^s and N_t^i denote labor supply. Parameters $\{\pi_i\}_{i=1}^3$ determine relative shares of people infected while "shopping", infected at the workplace, and infected from reasons unrelated to economic activity. Also, the number of newly infected is increasing both in the share of susceptible and infected households. Intuitively, a higher number of susceptible people increase the probability that an infected person will transmit the disease because she meets more potential victims. A higher number of infected people obviously imply a higher number of infection spreaders and hence a higher number of newly infected.

$$T_t = \pi_1 C_t^s S_t C_t^i I_t + \pi_2 N_t^s S_t N_t^i I_t + \pi_3 S_t I_t \tag{1}$$

Because the model is calibrated to weekly frequency, the share of new born people is vanishingly small, hence the equation governing dynamics of the share of susceptible households collapses into a particularly simple form featuring only outflow of newly infected.

$$S_{t+1} = S_t - T_t \tag{2}$$

The dynamics of the share of infected households include two terms. Firstly, there is and outflow term which includes newly recovered and newly deceased households. Intensity of this outflow is governed by parameters π_r and π_d which represents probability of recovery and probability of death respectively. Inflow of newly infected is governed by equation (1).

$$I_{t+1} = (1 - \pi_r - \pi_d) I_t + T_t \tag{3}$$

[24]Number of deceased households is implied by knowledge of these three shares since the total population is normalized to unity.

The dynamics of the share of recovered include only an inflow term, since recovered households are assumed to receive perfect immunity and hence there are no re-infections.

$$R_{t+1} = R_t + \pi_r I_t \tag{4}$$

Government in this economy sets path of consumption tax $\{\mu_t\}_{t=\tau}^{\infty}$ and lump-sum transfers $\{\Gamma_t\}_{t=\tau}^{\infty}$. Consumption tax serves as a proxy for lockdowns and other epidemics-containment measures. Government is assumed to obey a balanced budget constraint which links consumption tax revenues to lump-sum transfers.

$$\mu_t \left(C_t^s \times S_t + C_t^i \times I_t + C_t^r \times R_t \right) = \Gamma_t \left(S_t + I_t + R_t \right) \tag{5}$$

Production sector of the economy is competitive in both input and output markets and composes of unit measure of identical firms. Without loss of generality, production sector can by summarized by means of representative firm, which operates a linear production technology that transforms labor supplied by households into final output.

$$Y_t = \alpha \times N_t = \alpha \times \left(N_t^s S_t + N_t^i I_t + N_t^r R_t \right) \tag{6}$$

This structure of production sector implies constant wage rates [25].

$$\omega = \alpha \tag{7}$$

All households in the economy have the same period utility function (henceforth felicity function) over consumption and labor supply. Felicity function is assumed to take log-quadratic form with labor dis-utility parameter θ. The log-quadratic form of felicity function allows for closed-form characterization of steady-state

$$u(c_t, n_t) = \log c_t - \frac{\theta}{2} n_t^2 \tag{8}$$

Following Eichenbaum, Rebelo and Trabandt (2020), I wrote decision problems of households in terms of recursively described lifetime utility. However, it should be stressed that Eichenbaum, Rebelo and Trabandt (2020) formulated and solved the model in the sequence-space. Although the utility is written in terms of recursion, value

[25] Hence, wages are independent of aggregate state. This fact greatly simplifies model solution. In laissez-faire equilibrium it generates block-recursive structure which allows to analyticaly characterize a large part of the equilibrium dynamics.

functions are modeled as time-dependent. To made the problem recursive, a sufficient set of state variables had to be determined, and all equilibrium objects had to be written such that they solely depend on those state variables.

The decision problem of a susceptible household is described by the following bellman equation. $\mathcal{V}_t$ denotes the value function of susceptible households. Following Lucas Jr and Prescott (1971), lower case variables (c_t^s, n_t^s) denote individual decision variables of households. Susceptible households maximize the sum of its felicity flow and discounted continuation value. Compared to the canonical household problem of dynamics macroeconomics household doesn't engage in the consumption-savings decision. There are no savings in this economy, neither in form of productive capital, neither bonds. Households are indexed only by their health status. A susceptible household faces classical static consumption-leisure modified by health concern summarized by the difference between continuation value of remaining susceptible (healthy) denoted as $\mathcal{V}_{t+1}$, continuation value of getting infected $\mathcal{X}_{t+1}$ and infection probability τ_t. Infection probability mirrors the structure of the aggregate infection equation. The key difference is the presence of individual decision variables and omission of the share of susceptible households[26].

The possibility of getting infected, and the difference between the continuation value of susceptible versus infected household enters as a wedge into the otherwise static consumption-leisure problem and makes it dynamic[27]. Because of this forward-looking element, the model can't be solved as a simple system of equations and had to be tackled by means of sequence-space or even functional analysis tools.

$$\mathcal{V}_t = max_{c_t^s, n_t^s} \{u(c_t^s, n_t^s) + \beta(1-\tau_t)\mathcal{V}_{t+1} + \beta\tau_t\mathcal{X}_{t+1}\}$$

$$(1+\mu_t)c_t^s \leq n_t^s w + \Gamma_t \tag{9}$$

$$\tau_t = \pi_1 c_t^s C_t^i I_t + \pi_2 n_t^s N_t^i I_t + \pi_3 I_t$$

[26]Share of susceptible households doesn't influence the probability that individual susceptible household get infected. This probability depends only on consumption and labor choices of susceptible household together with the share and behavior of infected households

[27]Because current choices became influenced by forward-looking continuation values, together with current state variables and current decisions.

In the case of infected households, forward-looking terms enter only continuation value, but continuation value itself is independent of current decisions, hence the problem faced by infected households is the static consumption-leisure problem. Infected households face twofold penalization. Firstly, they are less productive, and hence their wage is scaled-down by a factor $\phi \in (0, 1)$. Secondly, they face the risk of succumbing to disease and dying. The period death probability is denoted by π_d. The continuation value of a deceased person is parameterized by $\mathcal{D}$, it could be interpreted as the magnitude of "fear of death" [28]. If an infected household survives in period t, it either remains infected with probability $(1 - \pi_r - \pi_d)$ and receives continuation value of infected household or recovers from disease with probability π_r and receives continuation value of the recovered household.

$$\begin{gathered}\mathcal{X}_t = u(c_t^i, n_t^i) + \beta(1 - \pi_r - \pi_d)\mathcal{X}_{t+1} + \beta\pi_r\mathcal{W}_{t+1} + \beta\pi_d\mathcal{D} \\ (1 + \mu_t)c_t^i \leq n_t^i w\phi + \Gamma_t\end{gathered} \tag{10}$$

Recovered households also face a very standard static consumption-leisure problem, since they face only a static budget constraint. With the log-quadratic utility, policy functions of both infected and recovered households could be obtained in closed-form. However, value functions are more complex objects, since they are inherently forward-looking. In laissez-faire equilibrium, they could be also analyzed in closed-form, since both consumption taxes and government transfers disappear from budget constraint and wages are fixed. Outside laissez-faire equilibrium, value functions could not be solved in closed-form, because they depend on future dynamics of government policies together with future behavior of susceptible households which co-determine the size of future government transfers to recovered and infected households. Hence in the case of active government, value functions had to be approximated by a projection to a suitable

[28] This parameter is not present in the original model of Eichenbaum, Rebelo and Trabandt (2020). Instead, they set it implicitly as zero. My version hence nests their model as a special case for $\mathcal{D} = 0$. While Eichenbaum, Rebelo and Trabandt (2020) calibrated the model to match weekly wages in the United States economy in dollar terms, I use normalized consumption units, where pre-epidemics consumption is calibrated to 1. In contrast with their calibration, which delivers high nominal consumption and positive continuation values, my normalized calibration induces a negative continuation value. Hence, setting the continuation value of the death household to zero would imply counterfactual "suicidal" behavior of susceptible households. To prevent this problem, I introduced the $\mathcal{D}$ parameter, which could be calibrated to be negative enough to prevent suicidal behavior of households.

functional basis [29].

$$\mathcal{W}_t = u(c_t^r, n_t^r) + \beta \mathcal{W}_{t+1}$$
$$(1+\mu_t)c_t^r \leq n_t^r w + \Gamma_t \tag{11}$$

3.1.2 Recursive Equilibrium

To solve the model using functional approximation tools and facilitate further analysis of the model with aggregate uncertainty, I re-formulate the sequence-space problem into a functional-space problem by defining a recursive competitive equilibrium (RCE) of the economy.

The recursive competitive equilibrium of the economy is a:

1. Set of functions $\mathcal{V}, \mathcal{X}, \mathcal{W}, c^s, c^i, c^r, n^s, n^i, n^r, C^s, C^i, C^r, N^s, N^i, N^r, \mu, \Gamma$
2. Such that households value and policy functions solve their respective optimization problems, taking the rest of the economy as given

$$\mathcal{V}(\Omega) = max_{c^s,n^s}\{u(c^s,n^s) + \beta(1-\tau)\mathcal{V}(\Omega') + \beta\tau\mathcal{X}(\Omega')\}$$
$$(1+\mu(\Omega))c^s \leq n^s w(\Omega) + \Gamma(\Omega)$$
$$\tau = (\pi_1 c^s C^i(\Omega) I + \pi_2 n^s N^i(\Omega) I + \pi_3 I)$$
$$\mathcal{X}(\Omega) = max_{c^i,n^i}\{u(c^i,n^i) + \beta((1-\pi_r-\pi_d)\mathcal{X}(\Omega') + \pi_r\mathcal{W}(\Omega') + \pi_d\mathcal{D})\} \tag{12}$$
$$(1+\mu(\Omega))c^i \leq n^i w(\Omega)\phi + \Gamma(\Omega)$$
$$\mathcal{W}(\Omega) = max_{c^r,n^r}\{u(c^r,n^r) + \beta\mathcal{W}(\Omega')\}$$
$$(1+\mu(\Omega))c^r \leq n^r w(\Omega) + \Gamma(\Omega)$$

[29] Moreover, outside laissez-faire equilibrium, the policy is characterized in closed-form only as a function of government transfer, which in turn depends on the behavior of susceptible household. Hence, if policies of infected and recovered households should be characterized as functions of state variables, they had to be also numerically approximated.

3. Government policies μ, Γ satisly government budget constraint
$$\mu(\Omega)(C^s(\Omega)S + C^i(\Omega)I + C^r(\Omega)R) = \Gamma(\Omega)(S + I + R) \quad (13)$$
4. State $\Omega = \begin{pmatrix} S & I & R \end{pmatrix}$ evolves according to is law of motion, described by equations $(1-4)$. Those equations are summarized by operator $\mathcal{H}$.
$$\begin{gathered} \Omega' = \mathcal{H}(\Omega) \\ T = \pi_1 C^s(\Omega) S C^i(\Omega) I + \pi_2 N^s(\Omega) S N^i(\Omega) I + \pi_3 S I \\ S' = S - T \\ I' = (1 - \pi_r - \pi_d) I + T \\ R' = R + \pi_r I \end{gathered} \quad (14)$$
5. And rational expectations fixed-point condition holds
$$\begin{matrix} c^s = C^s & c^i = C^i & c^r = C^r \\ n^s = C^s & n^i = N^i & n^r = N^r \end{matrix} \quad (15)$$

Besides lifting description of household problems from the space of sequences to space of functions I transformed the set of government policies from the space of all feasible [30] tax and transfer sequences to a space of feasible tax and transfer functions. Those tax and transfer functions $\mu : \Omega \to \mathbb{R}$ and $\Gamma : \Omega \to \mathbb{R}$ describes possibly nonlinear mappings from the aggregate state space to actual tax and transfer quantities. With sufficiently rich functions μ and Γ this transformation doesn't lose expressivity given by possibility of choosing from all feasible policy sequences.

Finally, fixed-point conditions formally encode the structure of competitive equilibria under rational expectations. Household chooses their individual policy (functions), denoted by lower-case letters while taking choices of all other agents in the economy, denoted by upper-case letters, as given. Since all households in each group are identical, all of them had to choose the same value of control variables. Hence their policy functions are identical and equal to the aggregate policy functions [31]. Yet, it is crucial to distinguish

[30] In a sence that they satisfy the government balanced budget constraint.

[31] Intuitively this concept could be related to the iterative procedure in a spirit of Krusell and Smith (1998), abstracting from distribution approximation issues. Firstly, households guess how the rest of the economy behaves, and given that they compute their optimal response to the behavior of the rest of the economy. In a second step, households update their guess using the solution of their optimization problem, re-compute their best response and iterate on this procedure until a fixed-point is reached.

between individual and aggregate policy functions. Households do not control aggregate policy functions, and hence they do not take into account how their behavior will influence aggregate epidemic dynamics, which gave rise to the negative consumption and labor externality of susceptible and infected households.

3.1.3 First-Order Conditions and Euler Equation

Since optimization problems of all households are convex [32], their optimal behavior could be characterized by a set of first-order conditions coupled with bellman recursions determining their value functions. By deriving first-order conditions which determine optimal household behavior, and applying rational expectations fixed-point condition the recursive competitive equilibria of the economy get reduced into a relatively standard [33] set of functional equations that could be tackled directly using functional approximation tools, such as linear projection or deep learning.

The key part of the system is the intertemporal euler equation which describes the consumption-leisure-health decision problem of susceptible households. The striking difference with respect to the classical euler equation appearing in the canonical consumption-savings model in a spirit of Ramsey (1928) is the presence of value functions in the equation. The reason for this difference is the discrete nature of the individual state in this economy. Whereas households in classical consumption-savings problem à la Ramsey (1928) are indexed by their assets, which is a continuous variable [34], households in the epidemics economy of Eichenbaum, Rebelo and Trabandt (2020) are indexed only by their health status; susceptible, infected, and recovered. Because of this discrete nature of idiosyncratic household states, there is no envelope formula, which could be used to substitute-out the value function from the euler equation expression.

[32] Household utility is strictly concave in all their arguments, and constraint sets are linear in their choice variables. Also, the government imposes just linear taxes and lump-sum transfers, neither of them could spoil the convexity of household optimization

[33] Those equations no longer feature maximum operators. Instead, the maximization behavior of households is encoded in the structure of both static consumption-leisure conditions and in the intertemporal euler equation. Those equations feature then only standard mathematical expressions which could be calculated or approximated directly without relying on fixed-point iteration procedures which are typically slow and relatively unstable.

[34] At least in theoretical formulation, before the model is discretized and solved on a computer.

Hence, value functions had to be solved simultaneously with policy function as a solution to simultaneous functional equations [35].

The intertemporal euler equation of susceptible households in this economy features two distinct elements. Firstly, it is the standard consumption-leisure first-order condition, denoted in red. In the static problem, the optimality principle would require the household to equalize marginal utility of consumption with the lost marginal utility of leisure, and hence, the red expression should sum up to zero. However, susceptible households face a dynamic problem, because their consumption and work behavior influence the probability of getting infected, and hence it has a direct impact on continuation value because the continuation value of remaining healthy is different from the continuation value of getting infected. This health consideration manifests itself in the second term present in the intertemporal euler equation, denoted in blue. It acts efficiently as a wedge in the static first-order condition. In absence of this health consideration [36], the wedge term is equal to zero, which implies that the intertemporal euler equation collapses into a standard static first-order condition.

$$\begin{gathered}\beta \times (\mathcal{V}(\Omega') - \mathcal{X}(\Omega')) \times (w(\Omega)\pi_1 C^i I + \pi_2 N^i I(1+\mu(\Omega)) = \\ u_c(C^s, N^s)w(\Omega) + u_n(C^s, N^s)(1+\mu(\Omega)) \\ \mathcal{V}(\Omega) = u(C^s, N^s) + \beta \times (1-\tau)\mathcal{V}(\Omega') + \tau\mathcal{X}(\Omega')\end{gathered} \tag{16}$$

Optimal behavior of recovered households is characterized by a very simple consumption-leisure first-order conditions, and their value function solves traditional bellman equation.

$$\begin{gathered}u_c(C^s, N^s)w(\Omega) + u_n(C^s, N^s)(1+\mu(\Omega)) = 0 \\ \mathcal{W}(\Omega) = u(C^r, N^r) + \beta \times \mathcal{W}(\Omega')\end{gathered} \tag{17}$$

[35] In principle, this is a much more complicated problem relative to a standard set-up, for which envelope conditions allow for characterizing first-order conditions solely in terms of policy functions. If the only unknown in first-order conditions is the set of policy functions, a block-recursive approach could be used. One might firstly solve for optimal policies without having to simultaneously approximate value functions. After computing a reasonable approximation to the policy function, one might recover value functions relatively easily by simply solving the set of bellman equations while taking the policy functions computed in the previous step as given.

[36] For example in a steady-state, when the share of infected households is zero and hence the probability of getting infected is zero, regardless of the individual behavior of the susceptible household.

The first-order condition of the infected household mirrors the structure of the first-order condition which describes the optimal behavior of the recovered household. Value function of infected household solves following bellman recursion, maximum operator is omitted, since the optimizing behavior of household is encoded in the structure of the first-order condition. Because optimization behavior is already encoded in the first-order condition, which was derived under the assumption that individual households can not control aggregate policy functions, in this step, the rational expectations fixed-point condition is invoked, and individual policy functions are replaced by aggregate policy functions.

$$\begin{aligned} u_c(C^s, N^s)w(\Omega)\phi + u_n(C^s, N^s)(1 + \mu(\Omega)) = 0 \\ \mathcal{X}(\Omega) = u(C^r, N^r) + \beta \times (1 - \pi_r - \pi_d) \times \mathcal{X}(\Omega')+ \end{aligned} \tag{18}$$

The set of equations $(16 - 18)$, together with aggregate law of motion $(1 - 4)$ and government budget constraint (5), given some policy function μ or [37] Γ fully characterize the recursive equilibria of this economy. Since it this system no longer features maximum operators, it could be solved using linear projection or deep learning tools without the necessity to resolve the maximum operators in each and every iteration step of some nonlinear solver [38].

3.1.4 Steady-State

Before solving for the full recursive competitive equilibria it is useful to solve firstly for the model steady-state behavior. The steady-state of the epidemic economy is simply a state in which the share of infected households reaches zero. Compared to a standard business cycle economies with typically a unique ergodic steady-state to which the economy converges regardless of initial conditions, macro-epidemic

[37] Transfer function is implied by consumption tax function through government balanced budget constraint and vice-versa. Hence only one of those functions should be explicitly specified. Otherwise, great care is required to ensure that government budget constraint will hold with equality over the whole state-space.

[38] Presence of maximum operator and the need for iterative re-optimizations becomes especially costly when first or second-order methods are employed for minimization of the loss function, since these methods require computation of gradient or even hessian matrix. When the maximum operator had to be approximated using numerical routine, gradients or hessian matrices had to be approximated using finite-difference techniques which require many functional evaluations for computing gradients. This problem is especially severe in deep learning applications, where neural networks are parameterized by a very large number of parameters, and are optimized using gradient descent algorithms.

models feature much complex steady-state behavior. While all those steady-state values feature a zero share of infected households [39], different paths of epidemic lead to different final shares of susceptible, infected, recovered, and deceased households. However, the block-recursive structure of the epidemics economy of Eichenbaum, Rebelo and Trabandt (2020) imply that the steady-state behavior of households is the same, regardless final distribution of households across the health state [40].

Because in the probability of infection vanishes with a share of the infected household going to zero, in steady-state decision problems of susceptible household converges to the decision problem of the recovered household. Household behavior is hence characterized by a simple static first-order condition.

$$u_c(C,N)\alpha\phi + u_n(C,N) = 0 \tag{19}$$

The first-order condition is coupled with simple budget constraint.

$$c_{\leq}\alpha n \tag{20}$$

After substituting in the log-utility function, those two equations could be solved in closed-form for steady state consumption and labor supply of susceptible and recovered households.

$$\begin{aligned} c_{ss} &= \sqrt{\tfrac{1}{\theta}} \\ n_{ss} &= \alpha \times \sqrt{\tfrac{1}{\theta}} \end{aligned} \tag{21}$$

Those steady state consumption and labor choice imply also steady state level of utility, which could be then used to solve for steady state value functions of susceptible and recovered households.

$$\begin{aligned} u_{ss} &= \log c_{ss} - \tfrac{\theta}{2} n_{ss}^2 \\ \mathcal{V}_{ss} &= \tfrac{u_{ss}}{1-\beta} \end{aligned} \tag{22}$$

[39] Since epidemics stops only when the share of infected households drops to zero because then there is no possibility that remaining susceptible households get infected.

[40] To guarantee this property I assume that steady-state consumption tax is zero. This assumption is in line with the underlying economic logic of the economy, since abstracting from epidemics externality, which is absent in post-epidemic steady-state, this economy is Pareto efficient, and hence distortive government interventions can not improve welfare. A weaker assumption about steady-state policy could be used, however, I chose a stronger assumption for sake of exposition simplicity

Although the steady-state share of infected households, continuum definition of agent population allows for a zero-measure set of remaining infected households in the steady-state economy [41]. In the steady-state, the decision problem of households, including their steady-state value functions, could be also characterized in a closed-form, since they also face a constant aggregate environment. Because of the log-quadratic specification of the utility function, both infected, susceptible, and recovered households choose the same steady-state labor supply. However, steady-state consumption of infected households is lower, because of their productivity penalty ϕ.

$$\begin{gathered} c_{iss} = \sqrt{\frac{1}{\theta}} \\ n_{iss} = \phi \times \alpha \times \sqrt{\frac{1}{\theta}} \\ u_{iss} = \log c_{iss} - \frac{\theta}{2} n_{iss}^2 \\ \mathcal{X}_{iss} = \frac{u_{iss} + \beta\pi_r \times \mathcal{V}_{ss} + \beta\pi_d \times \mathcal{D}}{1 - \beta \times (1 - \pi_r - \pi_d)} \end{gathered} \tag{23}$$

Since the economy converges over time to the steady-state, I exploit knowledge of the steady-state behavior of the economy to construct a Dirichlet type of boundary conditions that restrict functions used to approximate the full dynamic equilibrium functions to be consistent with the long-run steady-state behavior of the economy.

3.1.5 Calibration

My calibration broadly follows the original approach employed by Eichenbaum, Rebelo and Trabandt (2020). Model is calibrated under the assumption that one time period lasts for one week. For calibration of π_r and π_d parameters, I directly use values suggested by Eichenbaum, Rebelo and Trabandt (2020). My calibration differs from Eichenbaum, Rebelo and Trabandt (2020) along two dimensions. Firstly, instead of calibrating the steady steady-state consumption to match nominal weekly wage in the economy of interest, I normalize steady-state consumption to one. Similarly, instead of calibrating worked hours to match nominal worked hours, I normalized labor supply to $\frac{1}{3}$, following the broad consensus in macroeconomics literature. Relative to nominal calibration of Eichenbaum, Rebelo and Trabandt (2020), my calibration does not guarantee positive utility

[41] Steady state problem of infected households allows for the convenient derivation of calibration equation for death disutility parameters D.

and continuation value of households, which implies that setting the continuation value of deceased households to zero might induce them to engage in counterfactual suicidal behavior, where households would try to get infected.

This issue leads to the second difference of my calibration with respect to Eichenbaum, Rebelo and Trabandt (2020). To prevent the suicidal paradox, I introduce an additional parameter $\mathcal{D}$ which parameterize the continuation value of deceased households. By setting it to a sufficiently negative value, households get really afraid of getting infected and possibly dying. For calibration of this parameter, I exploited the steady-state problem of infected households, which allows computing a consumption equivalent of the $\mathcal{D}$ parameter. In my calibration strategy [42] I assume, that an infected household can choose between living in a stationary [43] economy in which it faces the possibility of succumbing to the disease, and receiving one-off compensating payment, denoted as consumption multiplier λ, versus living in a deathless economy but not receiving this compensating lump-sum transfer. The consumption equivalent is a value of λ which solves the following equation.

$$\mathcal{V}^{ss}_{mortal}(\lambda, \Theta) = \mathcal{V}^{ss}_{deathless}(\Theta) \tag{24}$$

Given the rest of the parameters [44] which influence steady-state objects and policy functions, each value of $\mathcal{D}$ implies distinct value of λ. Conversely, for a given consumption equivalent, equation (24) could be solved for an implied value of $\mathcal{D}$. In order to exploit this relationship, I use data provided by Vocelka et al. (2017) and the Czech Statistical Office to estimate the consumption equivalent of lost life of the average Czech individual. They estimated the average valuation of one quality-adjusted year of life (QALY) to 22712 EUR. Assuming net average monthly wage of 1094 EUR, the average life expectancy of 79.48 years, the average current age of 42.5 years these

[42] This thesis focus on the solution method and particular calibration details are of minor interest. Hence my description of the calibration strategy is heuristic. This calibration should be understood only as a procedure for obtaining roughly plausible parameter values for which the model could be solved. For a normal quantitative exercise, a more careful calibration procedure should be used.

[43] Steady state economy

[44] Denoted by Θ. This set includes $\mathcal{D}$

data implies $\lambda = 3336$ [45].

3.1.6 Approximating Function and Boundary Conditions

To solve for the recursive competitive equilibria of the macro-epidemic economy of Eichenbaum, Rebelo and Trabandt (2020), my algorithm had to approximate six unknown functions. Those are value functions of susceptible, infected, and recovered households, denoted by $\mathcal{V}(\Omega)$, $\mathcal{X}(\Omega)$ and $\mathcal{W}(\Omega)$ respectively, and consumption functions denoted as $C_s(\Omega)$, $C_i(\Omega)$, and $C_r(\Omega)$ [46]. All of those unknown functions are represented using ansatz which combines fully connected feed-forward neural networks with Dirichlet-type boundary conditions constructed using steady-state values of value and policy functions.

Resulting approximating functions hence could be written as [47]

$$
\begin{aligned}
\hat{\mathcal{V}}(\Omega|\mathcal{P}_1) &= \mathcal{V}_{ss} - (\Omega[2]) \times \mathcal{NN}(\Omega|\mathcal{P}_1) \\
\hat{\mathcal{X}}(\Omega|\mathcal{P}_2) &= \mathcal{X}_{iss} - (\Omega[2]) \times \mathcal{NN}(\Omega|\mathcal{P}_2) \\
\hat{\mathcal{W}}(\Omega|\mathcal{P}_3) &= \mathcal{W}_{ss} - (\Omega[2]) \times \mathcal{NN}(\Omega|\mathcal{P}_3) \\
\hat{C}_s(\Omega|\mathcal{P}_4) &= C_{ss} - (\Omega[2]) \times \mathcal{NN}(\Omega|\mathcal{P}_4) \\
\hat{C}_i(\Omega|\mathcal{P}_5) &= C_{iss} - (\Omega[2]) \times \mathcal{NN}(\Omega|\mathcal{P}_5) \\
\hat{C}_r(\Omega|\mathcal{P}_6) &= C_{ss} - (\Omega[2]) \times \mathcal{NN}(\Omega|\mathcal{P}_6)
\end{aligned} \tag{25}
$$

All these functions are functions $\mathbb{R}^3 \to \mathbb{R}$. $\{\mathcal{P}_i\}_{i=1}^6$ are parameter vectors that index approximating neural networks. I use $\mathcal{P} = \{\mathcal{P}_i\}_{i=1}^6$ to denote a matrix which stacks together parameters of all networks used for approximation of the recursive equilibria.

After some experimentation, I settled on feed-forward neural networks with four hidden layers and 32 neurons per hidden layer, using swish (Ramachandran, Zoph and Le 2017) function as activation which performs nonlinear transformations in the hidden part of the

[45] Average current age and life expectancy implies, that in case of epidemic-induced death, average household lose 36.68 years. Using average QALY valuation estimated by Vocelka et al. (2017), this implies a monetary loss of 839889 EUR. Given an average monthly wage of 1094 this monetary loss is an equivalent of 63.98 net yearly wages or 3336 weekly wages.

[46] Labor supply functions does not have to be approximated, since given consumption function, labor supply could be obtained in closed-form by simply inverting budget constraint.

[47] Where $\Omega = (S, I, R)$. Hence $\Omega[2]$ denotes share of infected, and $\Omega[2]_{ss} = 0$, hence boundary condition encoded in the equation (25) constraints approximating functions to be consistent with steady state objects.

network. Since consumption had to be always non-negative [48], the usual practice in macroeconomic deep learning methods, such as Azinovic, Gaegauf and Scheidegger (2021), is to endowed output layers of neural networks which approximate consumption policy functions with softplus activation functions because those functions guarantee positive output and hence prevent possible numerical accidents during the training procedure. However, in this application, the structure of boundary conditions precludes this approach, since the neural network is subtracted from the steady-state value [49], which implies that consumption non-negativity constraints can not be directly encoded into the approximating functions. For all neural networks used in approximation functions, I employed linear activation [50], since it provides the flexibility required for universal approximation results in a spirit of Hornik, Stinchcombe, White et al. (1989).

$$\begin{aligned} Swish(x) = x \times \sigma(x) = \tfrac{x}{1-\exp -x} \\ Softplus(x) = \log(1+\exp x) \end{aligned} \tag{26}$$

Besides the swish function, which is a linearly scaled cumulative distribution function of the logistic distribution, the gelu [51] or the tanh [52] function could be also used very efficiently.

$$\begin{aligned} Gelu = x \times \Phi(x) \\ Tanh = \tfrac{\exp x - \exp -x}{\exp x + \exp -x} \end{aligned} \tag{27}$$

While some influential papers on deep learning solution methods in macroeconomics (eg.L. Maliar, S. Maliar and Winant 2019; Azinovic, Gaegauf and Scheidegger 2021) recommend using simple rectified-linear (ReLU) activation functions, I found their performance to be inferior relative to tanh or scaled cumulative distribution activations, even in some applications beyond problems analyzed in this thesis.

[48] As implied by logarithmic utility of consumption.

[49] Because epidemics consumption, labor supply, as well as continuation values, are lower than their steady-state counterparts

[50] Also known as linear activation

[51] Gelu is a linearly scaled cumulative distribution function of the normal distribution. In general, I found scaled cumulative distributions as particularly efficient activation functions, especially for unbounded problems.

[52] Tanh activation is mentioned by Fernández-Villaverde, Nuño, Sorg-Langhans and Vogler (2020) and Gopalakrishna (2021) as particularly efficient activation function for challenging macroeconomic applications. However, I found scaled cumulative distributions, and especially the swish function as more convenient.

However, relative to L. Maliar, S. Maliar and Winant (2019) and especially Azinovic, Gaegauf and Scheidegger (2021) who analyzed very large scale problems using high-performance computation units with large clusters of graphical processing units (GPUs), where ReLU activations could be efficient due to their simplicity, which allows for efficient utilization of massively parallel GPU processing cores [53].

$$ReLU = \max\{0, x\} \tag{28}$$

3.1.7 Loss Function

The loss function defines the least-squares projection condition. Using notation from Chapter 2, the loss function is defined as a weighted sum of model residuals indexed by the set of parameters which characterize approximating neural networks [54]. For sake of tractable notation, I directly restrict attention to least-squares loss function where $\mathcal{M}$ denotes the set of model equations, $\mathcal{G}$ set of approximating neural networks, $\mathcal{P}$ matrix of parameters that index neural networks of interest and $\{x_i\}_{i=1}^{N}$ denotes a sample of states used as a possibly randomized projection grid [55].

$$\mathcal{L}(\mathcal{P}) = \sum_{i=1}^{N} \mathcal{M}\left(\mathcal{G}(x_i, \mathcal{P})\right) \mathcal{M}\left(\mathcal{G}(x_i, \mathcal{P})\right)^T \tag{29}$$

In the first exercise, I solve the recursive equilibrium of the Eichenbaum, Rebelo and Trabandt (2020) economy. Firstly, I solve the laissez-faire economy, in which the absence of government intervention generates a block-recursive structure that could be used for analytical characterization of a large part of the recursive equilibria.

[53] Those ReLU-related efficiency gains could be even more pronounced when using tensor processing units (TPUs), which use even more specialized architecture which is aggressively optimized to perform massively parallel linear algebra operations at cost of sacrificing some accuracy and ability to efficiently perform more complex computational tasks, which are needed for example for computation of more complicated activation functions.

[54] Following Fernández-Villaverde, Hurtado and Nuno (2020) this set of neural networks could be for thought off as one network with multiple outputs.

[55] In my implementation, I employ a uniform hypercubic sampling scheme which rule-of-thumb policy simulations to bound the relevant region of share of an infected household. This restriction is extremely important for numerical stability of the training process since the high nonlinearity of the macro-epidemics model tends to generate extremely challenging behavior of the model equilibrium outside economically reasonable part of the state space. In this application, the simple rule-of-thumb method was sufficient to ensure fast convergence of the training algorithm. However, in more challenging environments, an iterative procedure might be necessary. One might firstly guess an upper bound, solve the model, and then check whether the simulated dynamics of the economy stays with this bound.

The only component of the laissez-faire economy which had to be approximated using deep learning is the value and policy function of susceptible household, which face genuinely dynamic decision problem, due to the interaction of their forward-looking continuation value with their consumption-leisure margin. In this very simple economy, the $\mathcal{M}$ equation set composes only from the euler and bellman equation described in the equation (16) [56]. The rest of equilibrium objects is available in closed form, and hence does not have to been approximated.

Secondly, I solve the general economy in which government could set non-zero consumption tax. For exposition simplicity, I assume that the government policy rule takes a form of a simple linear function of the current share of infected households parameterized by a coefficient $\varphi \geq 0$.

$$\mu(\Omega) = \varphi \times \Omega[2] = \varphi \times I \tag{30}$$

In case of $\varphi > 0$ the block-recursive structure of the laissez-faire equilibria is lost, because non-zero consumption taxes implies non-zero transfers, and since transfers depend jointly on both on the state-dependent consumption tax rate and state-dependent behavior of susceptible behavior, value and policy functions of infected and recovered households became quite complex dynamic objects which had to be also approximated. The $\mathcal{M}$ equation set in this case includes 3 bellman equations and 3 euler equations described in equations $(16-18)$. However, for $\varphi = 0$ results of both deep learning problems should coincide, up to some numerical errors.

As mentioned in the Chapter 2 optimal value of parameter matrix $\mathcal{P}$ which indexes the set of approximating neural network is pinned-down by minimizing the loss function using adaptive stochastic gradient descent algorithm known as ADAM (Kingma and Ba 2014). While the machine learning literature developed a plethora of competing network optimization algorithms, ADAM is widely considered to be a good default optimizer that is simple to implement, computationally efficient, and offers fast and relatively stable convergence properties. In all presented problems, ADAM delivered a reason-

[56] Subject to the aggregate law of motion defined by the SIR model described by equations $(1-4)$

able performance, so there was no reason for using some alternative method for network optimization.

3.1.8 Results

Since the general model economy nests, the laissez-faire economy as a special case for $\phi = 0$, the comparison of results obtained by numerical approximation of the laissez-faire economy with the general economy under $\varphi = 0$ could serve as a basic test of the accuracy of my solution method. Figures 1 and Figure 2 show, that both algorithms generated indistinguishable economic dynamics, which fact speaks in favor of accuracy of my method [57].

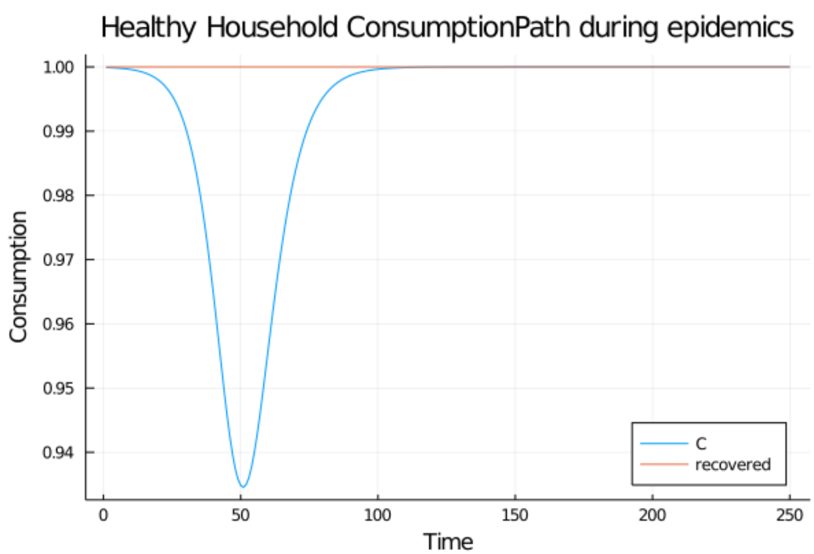

Figure 1: Consumption path of susceptible households in the laissez-faire economy

[57]Both algorithms also converged to the required loss function threshold.

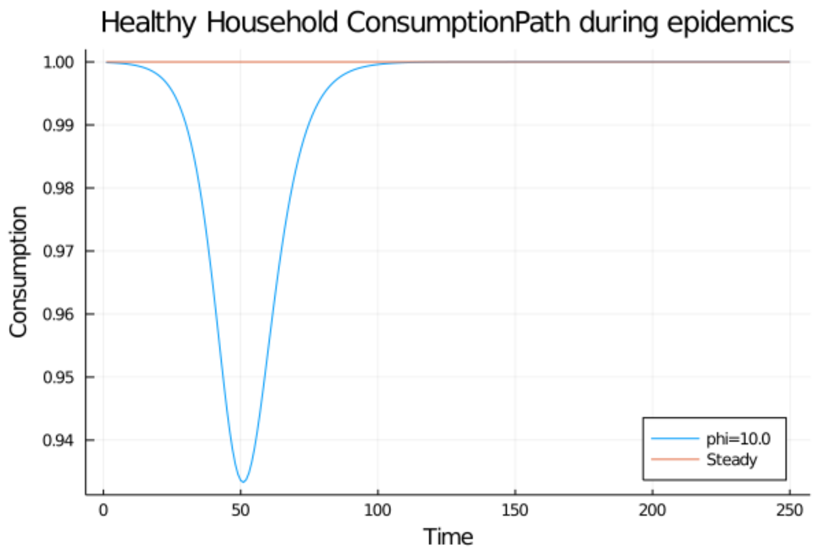

Figure 2: Consumption path of susceptible households in the general economy with $\varphi = 0$

3.2 Stochastic Economy

In this exercise, I focus on competitive equilibria of the Eichenbaum, Rebelo and Trabandt (2020) economy with aggregate shocks. Compared to the vanilla Eichenbaum, Rebelo and Trabandt (2020), who work in an environment with deterministic aggregates, I extend the set of epidemic SIR equations with mutation risk that changes virulence of the epidemics.

$$T_t = \pi_1 C_t^s S_t C_t^i I_t + \pi_2 N_t^s S_t N_t^i I_t + \pi_3 S_t I_t \tag{1}$$

$$S_{t+1} = S_t - T_t \times \mathcal{E}_t \tag{2}$$

$$I_{t+1} = (1 - \pi_r - \pi_d) I_t + T_t \times \mathcal{E}_t \tag{3}$$

$$R_{t+1} = R_t + \pi_r I_t \tag{4}$$

For the stochastic economy, I use the same calibration procedure as for the baseline economy [58]. Compared to the original Eichenbaum,

[58] Parameters governing the mutation Markov chain governing the virulency of epidemics,

Rebelo and Trabandt (2020) paper, I assume that number of infected doesn't depend just on the original set of state variables $\Omega_t = (S_t, I_t, R_t)$ and household choices, I introduce a $\mathcal{E}_t$ element which represents mutation of the virus. I assume that there is a low-virulency, standard, and high-virulency mutation [59].

$$\mathcal{E} \in \{0.5, 1.0, 1.5\} \tag{5}$$

I assume that these mutations follow an exogeneous markov process with transition matrix

$$\mathbb{M} = \begin{bmatrix} 0.9 & 0.1 & 0 \\ 0.05 & 0.9 & 0.05 \\ 0 & 0.1 & 0.9 \end{bmatrix} \tag{6}$$

The problem of susceptible household could be written as [60]

$$\mathcal{V}_t = max_{\{c_t^s, c_t^n\}} \left\{ u(c_t^s, n_t^s) + \beta(1-\tau_t)E_t\left[V_{t+1}\right] + \beta\tau_t\mathbb{E}_t\left[\mathcal{X}_{t+1}\right] \right\}$$

$$(1+\mu_t)c_t^s \leq n_t^s w + \Gamma_t \tag{7}$$

$$\tau_t = (\pi_1 c_t^s C_t^i I_t + \pi_2 n_t^s N_t^i I_t + \pi_3 I_t) \times \mathcal{E}_t$$

Infected household

$$\mathcal{X}_t = max_{\{c_t^i, c_t^i\}} \left\{ u(c_t^i, n_t^i) + \beta(1-\pi_r-\pi_d)\mathbb{E}_t\left[\mathcal{X}_{t+1}\right] + \beta\pi_r\mathbb{E}_t\left[\mathcal{W}_{t+1}\right] + \beta\pi_d\mathcal{D} \right\}$$

$$(1+\mu_t)c_t^i \leq n_t^i w\phi + \Gamma_t \tag{8}$$

Recovered household

$$\mathcal{W}_t = max_{\{c_t^r, c_t^r\}} \left\{ u(c_t^r, n_t^r) + \beta\mathbb{E}_t\left[\mathcal{W}_{t+1}\right] \right\}$$

$$(1+\mu_t)c_t^r \leq n_t^r w + \Gamma_t \tag{9}$$

The recursive equilibria in the stochastic economy is defined as follows

which were not present in the baseline economy are calibrated to "plausible" values which could be used for testing of the algorithm.

[59] While I didn't calibrate those parameters formally, they serve as a proof of concept for a "non-zero" aggregate risk case.

[60] $\mathbb{E}_t$ denotes a conditional expectations operator formed using information set available at period t.

1. Set of functions $\mathcal{V}, \mathcal{X}, \mathcal{W}, c^s, c^i, c^r, n^s, n^i, n^r, C^s, C^i, C^r, N^s, N^i, N^r, \mu, \Gamma$

2. Such that households value and policy functions solve their respective optimization problems, taking the rest of the economy as given

$$
\begin{gathered}
\mathcal{V}(\Omega) = max_{c^s,n^s}\{u(c^s, n^s) + \beta(1-\tau)\mathbb{E}\left[\mathcal{V}(\Omega')\right] + \beta\tau\mathbb{E}\left[\mathcal{X}(\Omega')\}\right] \\
(1+\mu(\Omega))c^s \leq n^s w(\Omega) + \Gamma(\Omega) \\
\tau = (\pi_1 c^s C^i(\Omega)I + \pi_2 n^s N^i(\Omega)I + \pi_3 I) \times \mathcal{E}_t \\
\mathcal{X}(\Omega) = max_{c^i,n^i}\{u(c^i, n^i) + \beta((1-\pi_r-\pi_d)\mathbb{E}\left[\mathcal{X}(\Omega')\right] \\
+\pi_r\mathbb{E}\left[\mathcal{W}(\Omega')\right] + \pi_d\mathcal{D})\} \\
(1+\mu(\Omega))c^i \leq n^i w(\Omega)\phi + \Gamma(\Omega) \\
\mathcal{W}(\Omega) = max_{c^r,n^r}\{u(c^r, n^r) + \beta\mathbb{E}\left[\mathcal{W}(\Omega')\right]\} \\
(1+\mu(\Omega))c^r \leq n^r w(\Omega) + \Gamma(\Omega)
\end{gathered} \tag{10}
$$

3. Government policies μ, Γ satisly government budget constraint

$$\mu(\Omega)(C^s(\Omega)S + C^i(\Omega)I + C^r(\Omega)R) = \Gamma(\Omega)(S+I+R) \tag{11}$$

4. State $\Omega = \begin{pmatrix} S & I & R & \mathcal{E} \end{pmatrix}$ evolves according to is law of motion, where $\mathcal{F}$ denotes the Markov chain distribution described by equations (5) and (6)

$$
\begin{gathered}
\Omega' = \mathcal{H}(\Omega) \\
T = \pi_1 C^s(\Omega)SC^i(\Omega)I + \pi_2 N^s(\Omega)SN^i(\Omega)I + \pi_3 SI \\
S' = S - T \times \mathcal{E}_t \\
I' = (1-\pi_r-\pi_d)I + T \times \mathcal{E} \\
R' = R + \pi_r I \\
\mathcal{E}' = \mathcal{F}(\mathcal{E})
\end{gathered} \tag{12}
$$

5. And rational expectations fixed-point condition holds

$$
\begin{array}{ccc}
c^s = C^s & c^i = C^i & c^r = C^r \\
n^s = C^s & n^i = N^i & n^r = N^r
\end{array} \tag{13}
$$

As in the deterministic economy, optimization problems of households are fully characterized by a set of necessary and sufficient first-order conditions. The core of the equilibrium system is the following inter-temporal euler equation and associated bellman recursion.

$$\begin{gathered}\beta \times \mathbb{E}_{\Omega}[(\mathcal{V}(\Omega') - \mathcal{X}(\Omega'))] \times (w(\Omega)\pi_1 C^i I + \pi_2 N^i I(1+\mu(\Omega)) = \\ u_c(C^s, N^s)w(\Omega) + u_n(C^s, N^s)(1+\mu(\Omega)) \\ \\ \mathcal{V}(\Omega) = u(C^s, N^s) + \beta \times \mathbb{E}_{\Omega}[(1-\tau)\mathcal{V}(\Omega') + \tau\mathcal{X}(\Omega')]\end{gathered} \tag{14}$$

The rest of FOCs is trivial [61]. They are also coupled by bellman recursions for value functions of infected and recovered households [62], as well as the law of motion of the state and government budget constraint [63].

To solve for the recursive equilibrium of this economy, I used just a slight modification of the original algorithm. Since the mutation shock is represented using the N-state Markov chain, I approximate all equilibrium objects using N-output networks, where the i-th output of the network represents the function of interest in the i-th mutation state. This approach requires just a relatively limited change of the codebase when going from the deterministic to the deterministic economy. Also, networks with the same number of hidden layers as in a deterministic economy are typically sufficiently rich for representing shock-dependent functions in the stochastic economy.

Because in this application I assumed just a three-state Markov chain, I used a tensor-product approach for generating a training grid. I simply took the random sample from the Ω domain and evaluated all those points for all possible values of the Markov chain. This approach is efficient with such a simple exogenous state, however, this tensor-product approach will become infeasible when the dimensionality of the shock process increase. In that case, typical

[61] Static first-order conditions are identical to their counterparts in the deterministic economy.

[62] The key difference with respect to their structure in the deterministic economy is the presence of conditional expectations operator in continuation value expressions. Because uncertainty takes for of a discrete-state Markov chain, those conditional expectations could be computed as simple weighted sums without need for integral approximation techniques required in case of continuous-state uncertainty.

[63] While my code includes government budget constraint and simple linear policy, in this numerical experiment, I switched government off, so these are competitive equilibrium results. However, the same code could be used to solve for policy-distorted competitive equilibria.

set simulations in a style of (L. Maliar, S. Maliar and Winant 2019) might be employed to break the curse of dimensionality.

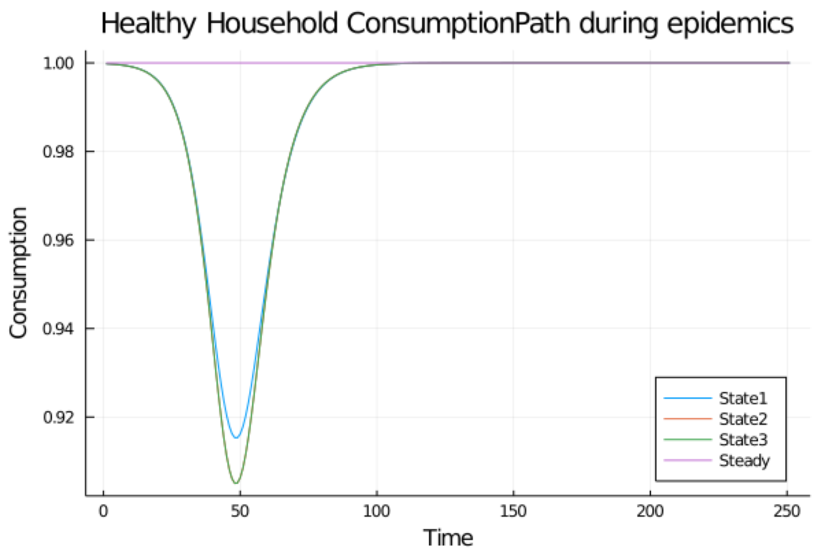

Figure 3: Consumption path of susceptible households in the stochastic economy. When the epidemics switch to the less aggressive mutation denoted by State1, susceptible household tends to work and consume more, because the risk of getting infected is lower compared to the situation under some more aggressive mutation.

3.3 Parametric Set Method

Finally, I present the extended algorithm which aims to approximate a whole set of model equilibria [64] indexed by a vector of policy rule parameters in order to simplify parametric Ramsey-style optimal policy computations which were historically plagued by the necessity for repeated solving of the model equilibria for a huge number of possible parameterizations of the policy rule. The only change with respect to the baseline algorithm which solves for the model equilibrium given one particular value of the parameter vector is that this extended algorithm treats the selected subset of model parameters as additional state variables. While those pseudo-states follow an identity law of motion, otherwise they are treated in the same way

[64] In this exercise, I solve the deterministic version of Eichenbaum, Rebelo and Trabandt (2020) economy.

as natural state variables of the problem.

The code used for solving the parameterized set of models differs from the code that implements the baseline economy only in a few minor details. Firstly, when there are N natural state variables and M parameters of interest, the model had to be solved over a $N + M$ dimensional state space. Hence neural networks used in the approximation had to have $N + M$ inputs. Secondly, the grid generating scheme had to be updated to sample from the whole extended state space. Naturally, training of networks that approximate the whole parametric model set is more computationally demanding than solving a single model, however, I did not experience any particular problems with the convergence of ADAM gradient descent when solving the model set problem. Convergence took required a bigger number of iterations but it was quite stable.

To check the accuracy of the model set solution algorithm, I created a plot (Figure 3)mirroring figures 1 and 2. It depicts the consumption path of susceptible households over the epidemic path generated by the very same initial conditions as in the case of figures 1 and 2. To generate it, I iterated the dynamic system generated by the model law of motion and computed policy functions. Figures 1 and 2 are created using simple policy functions. Figure 4 was constructed using generalized policy rule, which depends both on natural state variables and parameter pseudo-state whose value was fixed during the transition path to $\varphi = 0$. For this pseudo-state value, the transition path generated by generalized policy functions should be indistinguishable from transition paths generated by solving the laissez-faire and standard tax-distorted equilibria. Figure 4 shows, that the model set algorithm converged to an approximately correct solution since the transition path is indistinguishable from transition paths generated by previously discussed solution algorithms.

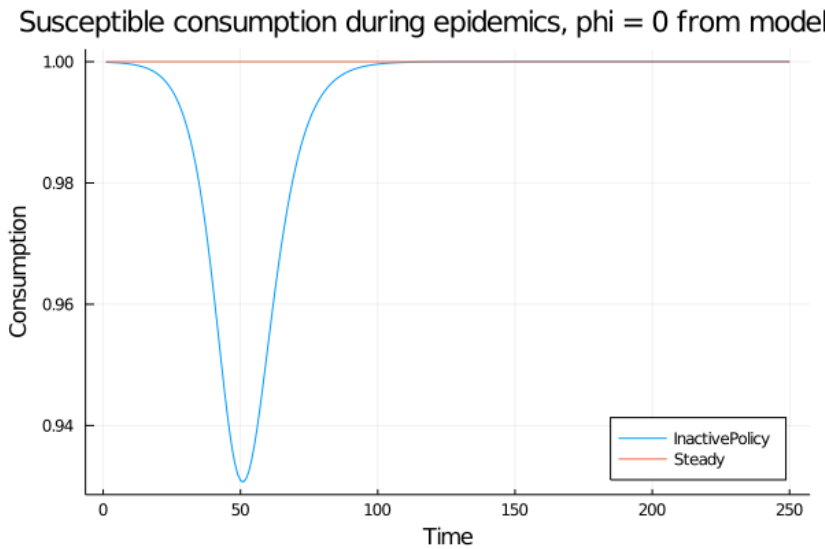

Figure 4: Consumption path of susceptible households in the model set economy with $\varphi = 0$

As a second check, I plotted transition paths of susceptible households for different values of the policy rule parameter φ one can see that economies with more aggressive government containment policy feature a sharper decline in consumption of susceptible households because their consumption is during the epidemics repressed by the increase of consumption tax implied by the aggressive policy rule.

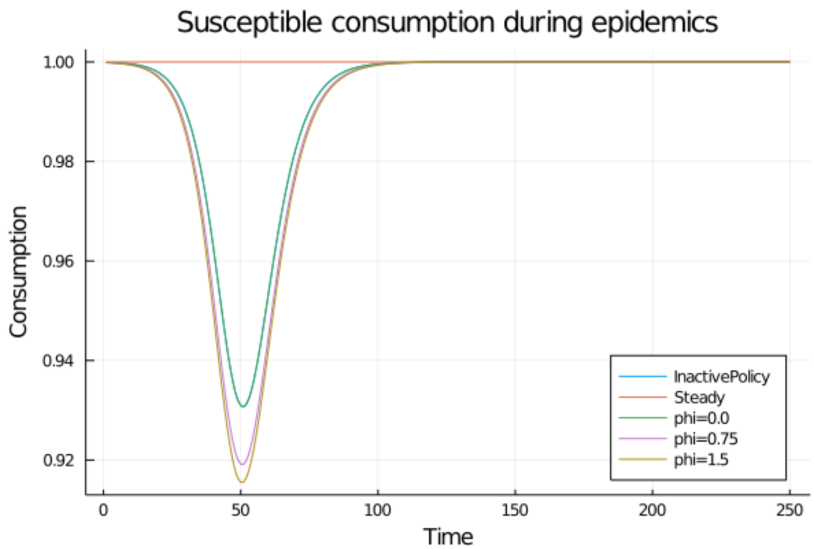

Figure 5: Consumption path of susceptible households in the model set economy with $\varphi = 0$

Finally, I used the approximated parametric model set to find the optimal value of the policy response parameter φ. To do so, I assume that government is utilitarian [65] and constructed a simple social welfare function which aggregates welfare of susceptible, infected, and recovered households by their relative population shares [66].

$$\begin{aligned} \mathcal{SWF}(S, I, R, \varphi) = S \times \mathcal{V}\,(S, I, R|\varphi) + \\ + I \times \mathcal{X}\,(S, I, R|\varphi) + R \times \mathcal{W}\,(S, I, R|\varphi) \end{aligned} \tag{15}$$

Since the model set algorithm solves for all of those value functions over both natural states and the policy parameter, the social welfare function could be also written as a function of state variables and φ. To solve for the optimal [67] value of φ, I maximized the social welfare function with respect to φ for initial condition of $\Omega = (0.9999, 0.0001, 0)$. This amounts to the standard approach in the Ramsey policy literature, where optimal policy is typically found

[65] Hence it aggregates utilities of all agents using constant weights. For convenience, I normalized those weights to one. Since utility and hence also social welfare is an ordinal concept, I could perform this normalization without loss of generality.

[66] Since utilitarian government weights all households exactly the same, effective Pareto weights in the social welfare function are equal to their respective population shares.

[67] Social welfare-maximizing.

by numerical maximization of the social welfare function for some given initial condition [68].

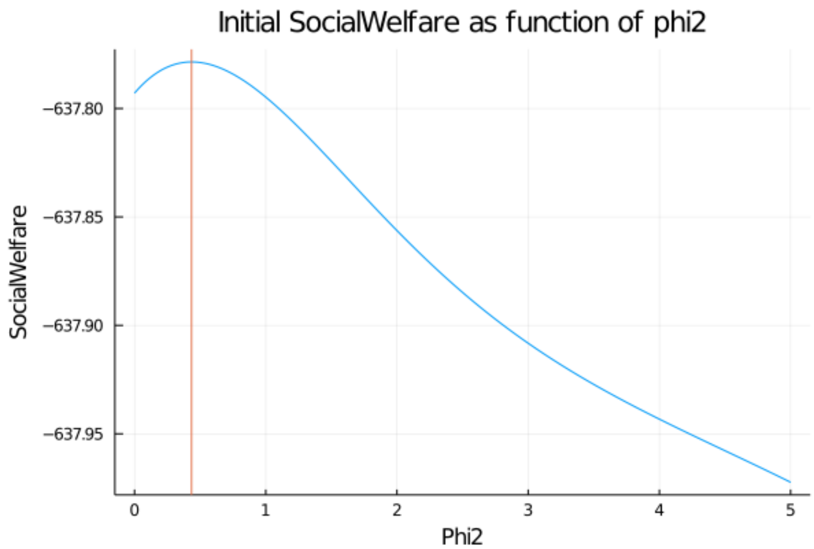

Figure 6: Social welfare for the initial state $\Omega = (0.9999, 0.0001, 0)$ as a function of φ While the social welfare function features a peak, it looks very sensitive to numerical errors. This problem is expected to disappear in case of more efficient policy rules

The key difference relative to standard policy optimization methods à la Dyrda and Pedroni (2018) is that the social welfare function generated by the model set method is a relatively simple function with closed-form expression as opposed to function which is implicitly given by the unknown solution to the model equilibrium. Since the policy rule in my version of the Eichenbaum, Rebelo and Trabandt (2020) economy involves only one parameter, the resulting optimization is just one dimensional, and hence I used the derivative-free Brent optimization algorithm Press, Flannery, Teukolsky, Vetterling et al. (1989). However, in the case of multidimensional problems, Newton or quasi-Newton methods are commonly utilized, and in that case gradients or even hessian matrices of the social welfare, function had to be computed. Thanks to machine learning libraries, such as Flux.jl (Innes et al. 2018), PyTorch (Paszke et al. 2019), or TensorFlow (Abadi et al. 2016), which are used for construction and

[68] The resulting optimal policy system crucially depends on the initial condition, because of ubiquitous time-consistency problems.

training of neural networks, those gradients, and Hessian matrices could be obtained in nearly automatic fashion at low computational costs.

3.3.1 Computer Implementation

I implemented all presented problems in Julia programming language [69] (Bezanson, Edelman, Karpinski and Shah 2017) using Flux.jl machine learning framework (Innes et al. 2018). All computations were run on a standard laptop computer [70] with runtime not exceeding one hour for any of the presented algorithms. For this application, I did not employ GPU acceleration, because from my experience possible speed gains implied by massively parallel computations facilitated by GPU are overshadowed by CPU-GPU communication overheads for neural networks featuring layers with less than approximately 200 neurons. However, for large-scale networks, GPUs or TPUs might provide immense speed gains relative to CPUs because those machine learning accelerators excel in handling massive linear algebra operations involved in neural network computations.

[69] All codes utilized to obtain results presented in this thesis could are available at the following GitHub repository: `https://github.com/Honza9723/ThesisCode` or at request.

[70] Acer Aspire 7 (A715-72G) station with Intel Core i7-8750H (2.2GHz, TB 4.1GHz) and 32 GB of RAM memory

4 Conclusion

In this thesis , I develop a new method for solving recursive macro-epidemic models and computing optimal government policies in epidemic economies using a deep learning algorithm. My contribution is twofold. Firstly, I provide an efficient and easy-to-use method for solving potentially high-dimensional recursive macro-epidemic models. While the existing quantitative macroeconomics literature proposes multiple different approaches for solving high-dimensional recursive models, such as the Smolyak interpolation (Krueger and Kubler 2004), adaptive sparse grid methods (Brumm and Scheidegger 2017,; Eftekhari and Scheidegger 2020), and deep learning projection (L. Maliar, S. Maliar and Winant 2019; Azinovic, Gaegauf and Scheidegger 2021), this literature has focused on business cycle and balanced growth path economies, and therefore is not well suited to deal with highly nonlinear dynamics on non-hypercubic domains, which are ubiquitous features in macro-epidemic models. Compared to the existing literature on solving recursive macroeconomic models using deep learning, I show how to tackle these peculiarities of macro-epidemic models using customization of training grid sampling and construction of appropriate boundary conditions.

Secondly, I extend the baseline solution algorithm with the parametric model set approach in the spirit of Duarte (2018a). Instead of solving the model for one particular parametric vector, I include a subset of model parameters that parameterize government policy as an additional state variable and train the neural network to solve the model on the expanded state space. This procedure simultaneously solves for a whole set of models indexed by the selected vector of policy rule parameters. The process of finding optimal values of government policy parameters is then reduced to the simple numerical optimization of a function with closed-form expression as opposed to a standard procedure in the spirit of Lippi, Ragni and Trachter (2015), in which the model is repeatedly solved for many different parameterizations of the policy rule. Since neural networks are well known for their asymptotic efficiency in the representation of complex objects in very high- dimensional spaces, my state-augmentation algorithm offers potentially significant speed-up when solving for complex policy rules.

To provide an example of my method, I applied it to the benchmark macro-epidemic model of Eichenbaum, Rebelo and Trabandt (2020), calibrated to replicate relevant properties of the Czech economy. After solving the benchmark model and checking the accuracy of the approximation, I used a modified version of my algorithm to solve the model with aggregate uncertainty in the form of virus mutation risk. Finally, I used my state-augmentation algorithm to solve for a set of model equilibria under different parameterizations of the government epidemic -containment policy rule. While the baseline model of Eichenbaum, Rebelo and Trabandt (2020) is quite stylized for capturing real-world epidemic dynamics, it presents an ideal benchmark for building confidence in my algorithm, since it features key numerical difficulties presented by macro-epidemic models.

Bibliography

References

Abadi, Martın, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard et al. (2016). 'Tensorflow: A system for large-scale machine learning'. In: *12th {USENIX} symposium on operating systems design and implementation ({OSDI} 16)*, pp. 265–283.

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions and Benjamin Moll (2020). *Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach*. Tech. rep.

Alvarez, Fernando E, David Argente and Francesco Lippi (2020). *A simple planning problem for covid-19 lockdown*. Tech. rep. National Bureau of Economic Research.

Aruoba, S Borağan, Jesus Fernandez-Villaverde and Juan F Rubio-Ramirez (2006). 'Comparing solution methods for dynamic equilibrium economies'. In: *Journal of Economic dynamics and Control* 30.12, pp. 2477–2508.

Azinovic, Marlon, Luca Gaegauf and Simon Scheidegger (2021). 'Deep equilibrium nets'. In: *Available at SSRN 3393482*.

Barron, Andrew R (1993). 'Universal approximation bounds for superpositions of a sigmoidal function'. In: *IEEE Transactions on Information theory* 39.3, pp. 930–945.

Bezanson, Jeff, Alan Edelman, Stefan Karpinski and Viral B. Shah (2017). *Julia: A Fresh Approach to Numerical Computing*. julialang.org/publications/julia-fresh-approach-BEKS.pdf. DOI: 10.1137/141000671.

Boyd, John P (2001). *Chebyshev and Fourier spectral methods*. Courier Corporation.

Brenner, Susanne and Ridgway Scott (2007). *The mathematical theory of finite element methods*. Vol. 15. Springer Science & Business Media.

Brumm, Johannes and Simon Scheidegger (2017). 'Using adaptive sparse grids to solve high-dimensional dynamic models'. In: *Econometrica* 85.5, pp. 1575–1612.

Chang, Roberto (1998). 'Credible monetary policy in an infinite horizon model: Recursive approaches'. In: *journal of economic theory* 81.2, pp. 431–461.

Chari, Varadarajan V and Patrick J Kehoe (1990). 'Sustainable plans'. In: *Journal of political economy* 98.4, pp. 783–802.

Clymo, Alex, Andrea Lanteri and Alessandro T Villa (2020). 'Capital and Labor Taxes with Costly State Contingency'. In:

Den Haan, Wouter J and Albert Marcet (1990). 'Solving the stochastic growth model by parameterizing expectations'. In: *Journal of Business & Economic Statistics* 8.1, pp. 31–34.

Duarte, Victor (2018a). 'Gradient-Based Structural Estimation'. In:

— (2018b). 'Machine Learning for Continuous-Time Economics'. In: *Available at SSRN 3012602*.

Dyrda, Sebastian and Marcelo Pedroni (2018). 'Optimal fiscal policy in a model with uninsurable idiosyncratic shocks'. In: *Available at SSRN 3289306*.

Eftekhari, Aryan and Simon Scheidegger (2020). 'High-Dimensional Dynamic Stochastic Model Representation'. In: *Available at SSRN 3603294*.

Eichenbaum, Martin, Sergio Rebelo and Mathias Trabandt (2020). *The macroeconomics of epidemics*. Tech. rep. National Bureau of Economic Research.

Fernández-Villaverde, Jesús, Samuel Hurtado and Galo Nuno (2020). *Financial Frictions and the Wealth Distribution*. Tech. rep.

Fernández-Villaverde, Jesús and Oren Levintal (2018). 'Solution methods for models with rare disasters'. In: *Quantitative Economics* 9.2, pp. 903–944.

Fernández-Villaverde, Jesús, Galo Nuño, George Sorg-Langhans and Maximilian Vogler (2020). *Solving High-Dimensional Dynamic Programming Problems using Deep Learning*. Tech. rep.

Fernández-Villaverde, Jesús, Juan Francisco Rubio-Ramırez and Frank Schorfheide (2016). 'Solution and estimation methods for DSGE models'. In: *Handbook of macroeconomics*. Vol. 2. Elsevier, pp. 527–724.

Glover, Andrew, Jonathan Heathcote, Dirk Krueger and José-Vıctor Rıos-Rull (2020). *Health versus wealth: On the distributional effects of controlling a pandemic*. Tech. rep. National Bureau of Economic Research.

Gonzalez-Eiras, Martın and Dirk Niepelt (2020). 'Optimally controlling an epidemic'. In:

Gopalakrishna, Goutham (2021). 'ALIENs and Continuous Time Economies'. In: *Swiss Finance Institute Research Paper* 21-34.

Guerrieri, Veronica, Guido Lorenzoni, Ludwig Straub and Iván Werning (2020). *Macroeconomic implications of COVID-19: Can negative supply shocks cause demand shortages?* Tech. rep. National Bureau of Economic Research.

Hornik, Kurt, Maxwell Stinchcombe, Halbert White et al. (1989). 'Multilayer feedforward networks are universal approximators.' In: *Neural networks* 2.5, pp. 359–366.

Innes, Michael, Elliot Saba, Keno Fischer, Dhairya Gandhi, Marco Concetto Rudilosso, Neethu Mariya Joy, Tejan Karmali, Avik Pal and Viral Shah (2018). 'Fashionable Modelling with Flux'. In: *CoRR* abs/1811.01457. arXiv: 1811.01457. URL: https://arxiv.org/abs/1811.01457.

Judd, Kenneth et al. (1991). *Minimum weighted residual methods for solving aggregate growth models*. Tech. rep. Federal Reserve Bank of Minneapolis.

Judd, Kenneth (1998). *Numerical methods in economics*. MIT press.

Kaplan, Greg, Benjamin Moll and Giovanni L Violante (2020). *The Great Lockdown and the Big Stimulus: Tracing the Pandemic Possibility Frontier for the US*. Tech. rep. National Bureau of Economic Research.

Kermack, William Ogilvy and Anderson G McKendrick (1927). 'A contribution to the mathematical theory of epidemics'. In: *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character* 115.772, pp. 700–721.

Kim, Jinill and Sunghyun Henry Kim (2003). 'Spurious welfare reversals in international business cycle models'. In: *journal of International Economics* 60.2, pp. 471–500.

Kingma, Diederik P and Jimmy Ba (2014). 'Adam: A method for stochastic optimization'. In: *arXiv preprint arXiv:1412.6980*.

Krueger, Dirk and Felix Kubler (2004). 'Computing equilibrium in OLG models with stochastic production'. In: *Journal of Economic Dynamics and Control* 28.7, pp. 1411–1436.

Krusell, Per and Anthony A Smith Jr (1998). 'Income and wealth heterogeneity in the macroeconomy'. In: *Journal of political Economy* 106.5, pp. 867–896.

Kydland, Finn E and Edward C Prescott (1982). 'Time to build and aggregate fluctuations'. In: *Econometrica: Journal of the Econometric Society*, pp. 1345–1370.

Lippi, Francesco, Stefania Ragni and Nicholas Trachter (2015). 'Optimal monetary policy with heterogeneous money holdings'. In: *Journal of Economic Theory* 159, pp. 339–368.

Lucas Jr, Robert E and Edward C Prescott (1971). 'Investment under uncertainty'. In: *Econometrica: Journal of the Econometric Society*, pp. 659–681.

Magill, Michael JP (1977). 'A local analysis of N-sector capital accumulation under uncertainty'. In: *Journal of Economic Theory* 15.1, pp. 211–219.

Maliar, Lilia and Serguei Maliar (2015). 'Merging simulation and projection approaches to solve high-dimensional problems with an application to a new Keynesian model'. In: *Quantitative Economics* 6.1, pp. 1–47.

Maliar, Lilia, Serguei Maliar and Pablo Winant (2019). *Will Artificial Intelligence Replace Computational Economists Any Time Soon?* Tech. rep.

Moser, Christian A and Pierre Yared (2020). *Pandemic lockdown: The role of government commitment*. Tech. rep. National Bureau of Economic Research.

Norets, Andriy (2012). 'Estimation of dynamic discrete choice models using artificial neural network approximations'. In: *Econometric Reviews* 31.1, pp. 84–106.

Paszke, Adam, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga et al. (2019). 'Pytorch: An imperative style, high-performance deep learning library'. In: *Advances in neural information processing systems* 32, pp. 8026–8037.

Phelan, Thomas and Keyvan Eslami (2021). 'Applications of Markov Chain Approximation Methods to Optimal Control Problems in Economics'. In:

Poggio, Tomaso, Hrushikesh Mhaskar, Lorenzo Rosasco, Brando Miranda and Qianli Liao (2017). 'Why and when can deep-but not shallow-networks avoid the curse of dimensionality: a review'. In: *International Journal of Automation and Computing* 14.5, pp. 503–519.

Press, William H, Brian P Flannery, Saul A Teukolsky, William T Vetterling et al. (1989). *Numerical recipes*.

Raissi, Maziar, Paris Perdikaris and George E Karniadakis (2019). 'Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations'. In: *Journal of Computational Physics* 378, pp. 686–707.

Ramachandran, Prajit, Barret Zoph and Quoc V Le (2017). 'Searching for activation functions'. In: *arXiv preprint arXiv:1710.05941*.

Ramsey, Frank Plumpton (1928). 'A mathematical theory of saving'. In: *The economic journal* 38.152, pp. 543–559.

Silver, David, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot et al. (2016). 'Mastering the game of Go with deep neural networks and tree search'. In: *Nature* 529.7587, p. 484.

Sirignano, Justin and Konstantinos Spiliopoulos (2018). 'DGM: A deep learning algorithm for solving partial differential equations'. In: *Journal of Computational Physics* 375, pp. 1339–1364.

Villa, Alessandro T and Vytautas Valaitis (2019). 'Machine Learning Projection Methods for Macro-Finance Models'. In: *Economic Research Initiatives at Duke (ERID) Working Paper Forthcoming*.

Vocelka, M, M Haluska, T Mazel, A Lukacisinova and I Stefancikova (2017). 'Willingness To Pay For Qaly In The Czech Republic Between 2013 And 2017: A Review'. In: *Value in Health* 20.9, A671–A672.