Self-justified equilibria: Existence and computation*

Felix Kubler
DBF, University of Zurich
and Swiss Finance Institute
fkubler@gmail.com

Simon Scheidegger
Department of Finance, HEC Lausanne
simon.scheidegger@unil.ch

September 23, 2018

Abstract

In this paper, we introduce the concept of "self-justified equilibrium" as a tractable alternative to rational expectations equilibrium in stochastic general equilibrium models with a large number of heterogeneous agents. A self-justified equilibrium is a temporary equilibrium where agents trade in assets and commodities to maximize the sum of current utility and expected future utilities that are forecasted on the basis of current endogenous variables and the current exogenous shock. Agents' characteristics include a set of admissible forecasting functions as well as a rule that maps the long run behavior of equilibrium into this set.

We provide sufficient conditions for the existence of self-justified equilibria, and we develop a computational method to approximate them numerically. For this, we focus on a convenient special case where agents project current endogenous variables into a lower dimensional subspace and where the dimension of this subspace can be viewed as optimally trading off the accuracy of the forecast and its complexity. Using Gaussian process regression coupled with active subspaces, we can solve models with hundreds of heterogeneous agents.

Keywords: Dynamic General Equilibrium, Rational Expectations, Active Subspaces, Gaussian Process Regression.

JEL Classification: C63, C68 , D50, D52.

*We thank Michael Reiter, Victor Rios-Rull, Tom Sargent as well as seminar participants at various universities and conferences for useful conversations and comments, and gratefully acknowledge financial support from PASC and the SFI.

1 Introduction

The assumption of rational expectations and the use of recursive methods to analyze dynamic economic models has revolutionized financial economics, macroeconomics, and public finance (see, e.g., Ljungqvist and Sargent (2012)). Unfortunately, for stochastic general equilibrium models with a large number of heterogeneous agents, rational expectations equilibria are generally not tractable, computational methods to approximate these equilibria numerically are often ad hoc, and a rigorous error analysis seems impossible (see, e.g, Brumm et al. (2017b), and references therein). In this paper, we develop an alternative to rational expectations equilibria and consider temporary equilibria with forecasting functions that depend on the long-run properties of the equilibrium but that might lead to incorrect forecasts at any given time. We derive simple sufficient conditions that ensure the existence of these "self-justified" equilibria, and we show that by restricting the complexity of agents' forecasts one can numerically approximate them for models with very many agents.

The basic idea of the approach is as follows. In a temporary equilibrium, agents use current endogenous variables and the exogenous shock to forecast future marginal utilities for assets; prices for commodities and assets in the current period ensure that markets clear. Forecasting functions are assumed to lie in a pre-specified class[1] - the agent chooses a function according to a (possibly agent-specific) rule that maps the long-run behavior of the economy to forecasting functions. As an example, it is useful to consider the case where the forecasting-function minimize the long-run average of the squared difference between realizations of marginal utilities along the equilibrium path and the forecasts. We discuss a different example in detail below, but for the theoretical analysis, we want to keep the rule as general as possible.

In the temporary equilibrium, the expectations might be far from correct and agents might make significant mistakes. The concept does not require identical expectations or identical forecasts across agents. Different types of agents can have different expectations and different forecasting functions.

We introduce the concept in the context of an infinite horizon pure exchange economy with overlapping generations, a single perishable commodity, and aggregate uncertainty. This allows us to investigate the properties of a self-justified equilibrium with as little notation as possible. An extension to production economies with several commodities (e.g. along the lines of Brumm et al. (2017)) is conceptually straightforward.

To prove the existence of self-justified equilibrium we make the simplifying assumption that accounting is finite. That is to say, we assume that beginning-of-period portfolios across agents

[1] A very simple example of this is the set of all polynomial functions of fixed degree. In our application below, we consider the set of functions that are compositions of continuous functions with low-dimensional domains and linear functions that are defined on a much higher dimensional domain.

lie on some finite (arbitrarily fine) grid and that agents' portfolio-choices in the current period induce a probability distribution over this grid. This assumption can be viewed as a technical approximation to a continuous model, but one can also think of bounded rationality justifications. For example, one might want to assume that at the beginning of a period, an agent cannot measure his financial wealth with arbitrary precision and makes small errors in rounding. In any case, while the assumption is necessary for the technical argument, it does not affect the computed solutions since all computations are necessarily using finite precision arithmetics.

In order to develop a tractable computational method, we consider a specific form for the forecasting function in that we assume that each agent projects the current endogenous variables into a relatively low dimensional subspace and approximates forecasts over this subspace globally.[2] The agent chooses an optimal projection to minimize the unexplained variation in future marginal utilities as measured by the mean-squared gradient. The forecasting function is then defined on a domain that is much lower dimensional than the space of current endogenous variables, and it is assumed to minimize mean-squared deviations from actual marginal utilities. This example of the concept of a self-justified equilibrium (which we refer to as dimension-reduced self-justified equilibrium) turns out to be computationally tractable in models with very many agents.

Following Scheidegger and Bilionis (2017), we achieve this tractability this by combining Gaussian process regression (see, e.g., Rasmussen and Williams (2005)) with the exploitation of so-called active subspaces (see, e.g., Constantine et al. (2014)). Using this combination allows us to construct a method that determines an economically intuitive projection for a fixed dimension of the subspace. This combination directly gives rise to a simple algorithm that trades off complexity and simplicity of the forecasting function and allows us to approximate self-justified equilibria numerically.

We demonstrate that our computational method can be applied to large-scale heterogeneous agents models by applying it to an overlapping generations economy with segmented financial markets. We assume that agents live for 60 periods and that there are two types of agents per generation, resulting in 120 agents altogether. We first consider the simplest case where an agent only uses his own asset-holding (together with the exogenous shock) to forecast future utilities (i.e. the asset holdings across all agents are projected into own asset holdings). This turns out to work very well in standard calibrations of the model. However, once we assume sufficient heterogeneity in tastes across generations, this simple method leads to large forecasting errors. We then exploit active subspace methods (see Constantine et al. (2014)) to show that adding one additional explanatory variable, that consists of a weighted mean of asset holdings across agents, reduces forecasting errors substantially. This observation will allow us to use the methods developed

[2] We use the term "global solution" for a solution that is computed using equilibrium conditions at many points in the state space of a dynamic model—in contrast to a "local solution", which rests on a local approximation around a steady state of the model.

here to tackle models with hundreds to thousands of agents. It is subject to further research to explore models where the dimension of the active subspace is larger.

There is a large and diverse body of work exploring deviations from rational expectation (see, e.g., Sargent (1993), Kurz (1994), Woodford (2013), Gabaix (2014), Adam et al. (2016)). Much of this work is motivated by insights from behavioral economics about agents' behavior or by the search for simple economic mechanisms that enrich the observable implications of standard models. The motivation of this paper is rather different in that we want to develop a simple alternative to rational expectations that allows researchers to rigorously analyze stochastic dynamic models with a very large number of heterogeneous agents.[3]

As Sargent (1993) points out, "when implemented numerically ... rational expectations models impute more knowledge to the agent within the model ... than is possessed by an econometrician", and a sensible approach to relax rational expectations is "expelling rational agents from our model environment and replacing them with 'artificially intelligent' agents who behave like econometricians." This quote embodies the idea underlying self-justified equilibria—to construct a tractable model of the macro-economy that takes into account substantial heterogeneity across agents one needs to assume that the modeler can compute agents' expectations.

There is also a large body of literature on the numerical approximation of rational expectations equilibria in models with heterogeneous agents (see Maliar and Maliar (2014) for a comprehensive overview). In our numerical strategy, we use several ideas from this literature.

Applied dynamic general equilibrium modeling has been criticized for its failure to take into account the considerable heterogeneity in tastes and technologies across agents. Farmer and Foley (2009) make this point forcefully and strongly advocate the use of so-called agent-based models to understand macro-economy dynamics. As they point out, in agent-based models, the agents can be as diverse as needed, but behavioral rules are often arbitrary. Up to now—especially in the presence of aggregate and idiosyncratic shocks—it seemed too complicated to incorporate substantial heterogeneity into large-scale dynamic general equilibrium models because existing solution methods are not able to handle this amount of heterogeneity (see Brumm et al. (2017b), Scheidegger et al. (2018)). Using the concept of self-justified equilibria, one can incorporate large-scale heterogeneity into general equilibrium models, potentially improve their usefulness for applied work and bridge the gap between agent-based modeling and applied general equilibrium.

The rest of the paper is organized as follows. In Section 2, the general economy is introduced, and a self-justified equilibrium is defined. In Section 3, we prove existence. In Section 4 we consider a variation of the concept which has the attractive features that it is tractable and that forecasts can be viewed as a trade-off between complexity and accuracy. In Section 5 we describe

[3]The methods developed in Krusell and Smith (1997) and in Evans and Phillips (2014) can also be interpreted to arise from this motivation and there are some important similarities to our work.

our computational strategy. In Section 6 we give a simple example to illustrate both the concept of self-justified equilibria and our computational method.

2 A general dynamic Markovian economy

We consider a Bewley-style overlapping generations model (see Bewley (1984)) with incomplete financial markets and a continuum of agents. Time is indexed by $t \in \mathbb{N}_0$. Exogenous shocks z_t realize in a finite set $\mathbf{Z} = \{1, \dots, Z\}$, and follow a first-order Markov process with transition probability $\pi(z'|z)$. A history of shocks up to some date t is denoted by $z^t = (z_0, z_1, \dots, z_t)$ and called a date event. Whenever convenient, we use t instead of z^t.

At each date event, a continuum of ex-ante identical agents enter the economy, live for A periods, and differ ex-post by the realization of their idiosyncratic shocks. Each agent faces idiosyncratic shocks, $y_1, ..., y_A$, that have support in a finite set $\mathbf{Y}^A$. We denote by $\eta_{y^a}(y_{a+1})$ the (conditional) probability of idiosyncratic shock y_{a+1} for an agent with shock history y^a, $\eta_0(y_1)$ to denote the probability of idiosyncratic shock y_1 at the beginning of life, and, $\eta(y^a)$ to denote the probability of a history of idiosyncratic shocks. We assume that the idiosyncratic shocks are independent of the aggregate shock, that they are identically distributed across agents within each type and age and, as in the construction in Proposition 2 in Feldman and Gilles (1985), that they "cancel out" in the aggregate, that is, the joint distribution of idiosyncratic shocks within a type ensures that at each history of aggregate shocks, z^t, for any $y^a \in \mathbf{Y}^a$ the fraction of agents with history $y^a = (y_1, ..., y_a)$ is $\eta(y^a)$. This allows the focus on equilibria for which prices and aggregate quantities only depend on the history of aggregate shocks, z^t. We denote the set of all date events at time t by $\mathbf{Z}^t$ and, taking z_0 as fixed, we write $z^t \in \mathbf{Z}^t$ for any $t \in \mathbb{N}_0$ (including $t = 0$). At each z^t, there are finitely many different agents actively trading (distinguishing themselves by age and history of shocks), who are collected in a set $\mathbf{I} = \cup_{a=1}^{A} \mathbf{Y}^a$. A specific agent at a given node z^t is denoted by $y^a \in \mathbf{I}$.

At each date event, there is a single perishable commodity, the individual endowments are denoted by $e_{y^a}(z^t) \in \mathbb{R}_+$ and assumed to be time-invariant and measurable functions of the current aggregate shock.[4] Each agent who can be identified by his date-event of birth, z^t, has a time-separable expected utility function

$$U_{z^t}((x_{t+a})_{a=0}^{A-1}) = \sum_{a=1}^{A} \sum_{z^{t+a-1} \succeq z^t} \sum_{y^a} \eta(y^a) \pi(z^{t+a-1}|z^t) u_{y^a}\left(x_{y^a}(z^{t+a-1})\right),$$

where $x_{y^a}(z^{t+a-1}) \in \mathbb{R}_+$ denotes the agent y^a's (stochastic) consumption at date $t + a - 1$.

There are J assets, $j \in \mathbf{J} = \{1, \dots, J\}$ traded at each date event. Assets can be infinitely lived Lucas trees in unit net supply or one-period financial assets in zero net supply. The net supply of

[4] As opposed to the standard formulation where an agent's fundamentals are functions of his current idiosyncratic shock, y, we assume that they are functions of the history of all shocks - clearly these formulations are equivalent if one allows for a sufficiently rich set $\mathbf{Y}$.

an asset j is denoted by $\bar{\theta}_j \in \{0, 1\}$. Assets are traded at prices q, and their (non-negative) payoffs depend on the aggregate shock and possibly on the current prices of the assets $f_j : \mathbb{R}^J_+ \times \mathbf{Z} \to \mathbb{R}_+$. If asset j is a Lucas tree (i.e., an asset in positive net supply), then $f_j(q, z) = q_j + d_j(z)$ for some dividends $d_j : \mathbf{Z} \to \mathbb{R}_+$. Asset j could also be a collateralized loan whose payoff depends on the value of the underlying collateral, or an option, or simply a risk-free asset. The aggregate dividends of the trees are defined as $d(z_t) = \bar{\theta} \cdot f(q(z^t), z_t) - \bar{\theta} \cdot q(z^t)$. An agent y^a faces trading constraints $\theta \in \mathbf{\Theta}_{y^a} \subset \mathbb{R}^J$, where $\mathbf{\Theta}_{y^A} = \{0\}$ for all $y^A \in \mathbf{Y}^A$. To simplify notation we write $\vec{\theta} = (\theta_{y^a})_{y^a \in \mathbf{I}}$, $\vec{\theta}^- = (\theta^-_{y^a})_{y^a \in \mathbf{I}}$ and $\vec{x} = (x_{y^a})_{y^a \in \mathbf{I}}$.

It is useful to define the set of possible portfolio holdings with market-clearing built-in as

$$\mathbf{\Theta} = \{\vec{\theta} : \sum_{y^a \in \mathbf{I}} \eta(y^a)\theta_{y^a} = \bar{\theta}, \quad \theta_{y^a} \in \mathbf{\Theta}_{y^a} \text{ for all } y^{a-1} \in \mathbf{I}\}.$$

Similarly, let the set of all beginning-of-period portfolio holdings be

$$\mathbf{\Theta}^- = \{\vec{\theta}^- : \theta^-_{y^1} = 0, \sum_{y^{a-1} \in \mathbf{I}} \eta(y^{a-1})\theta^-_{y^a} = \bar{\theta} \text{ and } \theta^-_{y^a} \in \mathbf{\Theta}_{y^{a-1}} \text{ for all } y^a\}.$$

We define the state space to be $\mathbf{S} = \mathbf{Z} \times \mathbf{\Theta}^-$ with Borel σ-algebra $\mathcal{S}$. The law of motion of the exogenous shock, π, and current choices $\vec{\theta}$ determine a probability distribution over next period's state - we write $\mathbb{Q}(\cdot|z, \vec{\theta})$. We will make assumptions on this probability distribution below which turn out to simplify the analysis but which are not typical in this strand of literature.

2.1 Self justified equilibria

In a competitive environment, agents choose asset-holdings in the current period to maximize expected lifetime utility and current prices ensure that markets clear. To understand how today's asset choices affect future utilities the agent needs to form some expectations about future prices and compute his optimal life-cycle asset-holdings under these prices. As already mentioned, it turns out to be useful to model the forecasting of prices and the recursive solution of the agents' problem in one step and assume that the agent makes a current decision given expectations over the next period's marginal utility of asset holdings. These expectations are based on current endogenous variables and the exogenous shock. While in rational expectations these expectations are always correct, we allow them to be imperfect and heterogeneous across agents.

In a temporary equilibrium each agent, $y^a \in \mathbf{I}$, is characterized by a function

$$M_{y^a} : \mathbf{S} \times \mathbb{R}^I_+ \times \mathbf{\Theta} \times \mathbb{R}^J_+ \to \mathbb{R}^J_+,$$

that predicts marginal utilities of assets in the next period on the basis of the current state, current prices and current consumptions and portfolio-holdings across agents. In our formulation, the agent forecasts marginal utilities from asset holdings. It might seem more standard to assume that the agent forecasts prices and then solves his life-cycle optimization problem on the basis of forecasted prices. However, this turns out to be much more complicated because he has to forecast prices over his entire life-cycle and not just one-period ahead. Moreover, we illustrate in a simple example below that forecasting prices might be more complicated than forecasting marginal utilities from asset-holdings. Finally, one could argue that the agent might forecast his value function in the next period to solve the maximization problem. This turns out to be too complicated since he has to forecast an entire function.[5]

We denote by $\vec{M} = (M_{y^a})_{y^a \in \mathbf{I}}$ the forecasting functions across all agents. Throughout this paper, we assume that $M_{y^A}(\cdot) = 0$ for all $y^A \in \mathbf{Y}^A$, forecasts of agents of age A are irrelevant. Assuming concavity of utility, the first order conditions are necessary and sufficient for agents' optimality and, given prices q and beginning-of-period asset-holdings $\theta^-_{y^a}$, we can write an agent y^a's maximization problem as

$$\max_{x \in \mathbb{R}_+, \theta \in \Theta_{y^a}} u_{y^a}(x) + M_{y^a}(s, \vec{x}, \vec{\theta}, q) \cdot \theta \quad \text{s.t.} \qquad (1)$$
$$x + \theta \cdot q - e_{y^a}(z) - \theta^-_{y^a} \cdot f(q, z) \leq 0.$$

The agent takes as given current average portfolio- and consumption choices across all agents, $\vec{\theta}, \vec{x}$, and current prices q. For now, the function $M_{y^a}(\cdot)$ is given—we endogenize this for our definition of self-justified equilibrium below.

Given forecasting functions across agents, $\vec{M}$, we define the temporary equilibrium correspondence

$$\mathbf{N}_{\vec{M}} : \mathbf{S} \rightrightarrows \mathbb{R}^I_+ \times \mathbf{\Theta} \times \mathbb{R}^J$$

as a map from the current state to current prices and choices that clear markets and that are optimal for the agents, given their forecasting functions, i.e.,

$$\begin{aligned} \mathbf{N}_{\vec{M}}(s) \quad = \quad & \{(\vec{x}, \vec{\theta}, q) \in \mathbb{R}^I_+ \times \mathbf{\Theta} \times \mathbb{R}^J_+ : \qquad (2) \\ & (x_{y^a}, \theta_{y^a}) \in \arg \max_{x \in \mathbb{R}_+, \theta \in \Theta_{y^a}} u_{y^a}(x) + M_{y^a}(s, \vec{x}, \vec{\theta}, q) \cdot \theta \text{ s.t.} \\ & x + \theta \cdot q - e_{y^a}(z) - \theta^-_{y^a} \cdot f(q, z) \leq 0 \quad \text{for all } y^a \in \mathbf{I}\}. \end{aligned}$$

Assuming that for a given $\vec{M}$ the set $\mathbf{N}_{\vec{M}}(s)$ is non-empty for all $s \in \mathbf{S}$ and that there exists a single-valued selection $N(s)$, we write

$$N(s) = \left(N_{\vec{x}}(s), N_{\vec{\theta}}(s), N_q(s)\right).$$

It should be kept in mind that the function $N(s)$ also depends on $\vec{M}$. However, to simplify notation, we do not make this explicit.

[5]It is true that one could approximate the value function by a finitely parameterized family of functions and the agent simply forecasts the finite dimensional vector of parameters, but this would still be substantially more complicated than simply forecasting a number.

The crucial innovation of this paper is to allow for heterogeneous and possibly incorrect forecasts across agents while still allowing for the possibility that agents are rational. For this, we assume that the agents cannot evaluate (or store) arbitrarily complicated functions, but instead, approximate the equilibrium forecasts by "simple" functions. These functions could be relatively simple because they aggregate $\vec{\theta}$ into a lower dimensional vector (cf. Section 4 below), or because they belong to some convenient class of functions - a simple example would be polynomial functions of fixed degree. For the definition of a self-justified equilibrium, we, therefore, assume that agents' expectations are characterized by sets of admissible forecasting functions, $\mathbf{M}_{y^a}$, $y^a \in \mathbf{I}$ (as above, we write $\mathbf{M} = \times_{y^a \in \mathbf{I}} \mathbf{M}_{y^a}$) as well as a rule that maps the long-run behavior of the economy to forecasts. To formalize this, denote by $\mathbf{F}^N$ the set of measurable functions from $\mathbf{S}$ to $\mathbb{R}^I_+ \times \mathbf{\Theta} \times \mathbb{R}^J_+$, and denote by $\mathbf{F}^{\mathbb{Q}}$ the set of measures on $(\mathbf{S}, \mathcal{S})$. We assume that for each y^a there is a functional $R_{y^a} : \mathbf{F}^N \times \mathbf{F}^{\mathbb{Q}} \to \mathbf{M}_{y^a}$.

We have the following definition.

Definition 1 *A self-justified equilibrium consists of forecasts $\vec{M} \in \mathbf{M}$, a selection $N(\cdot)$ of the temporary equilibrium correspondence, $\mathbf{N}_{\vec{M}}(\cdot)$, and measure $\mathbb{Q}^*$ on $(\mathbf{S}, \mathcal{S})$, such that*

1. *$\mathbb{Q}^*$ is invariant given the law of motion induced by $N(\cdot)$ and by $\mathbb{Q}(\cdot,\cdot)$. That is to say for all $\mathbf{B} \in \mathcal{S}$*
$$\mathbb{Q}^*(\mathbf{B}) = \int_{s \in \mathbf{S}} \mathbb{Q}(\mathbf{B}|z, N_{\vec{\theta}}(s)) d\mathbb{Q}^*(s).$$

2. *For each y^a, $a < A$,*
$$M_{y^a} = R_{y^a}(N, \mathbb{Q}^*).$$

One important case arises if we assume that R_{y^a} minimizes the long-run average of the squared difference between the forecasted marginal utility and realized marginal utility, among all functions in $\mathbf{M}_{y^a}$. To make optimal current choices, agents need to know the marginal utility of their asset holdings in the next period. This is an equilibrium object since it depends on all future prices over the agent's life-cycle. Given a selection $N(s)$ of the equilibrium correspondence, it is given by

$$m_{y^a}(z, \vec{\theta}) = \int_{s' \in \mathbf{S}} f(N_q(s'), z') \sum_{y_{a+1} \in \mathbf{Y}} \eta_{y^a}(y_{a+1}) u'_{y^{a+1}}(N_{x_{y^{a+1}}}(s')) d\mathbb{Q}(s'|z, \vec{\theta}). \tag{3}$$

We then assume that R_{y^a} maps into the forecasting function that provides the best average approximation under the invariant measure, i.e.,

$$M_{y^a} \in \arg \min_{M \in \mathbf{M}_{y^a}} \int_{s \in \mathbf{S}} \|M(s, N(s)) - m(z, N_\theta(s))\|_2^2 d\mathbb{Q}^*(s).$$

Similarly to the concept of "self-confirming" equilibrium (see e.g. Fudenberg and Levine (1993) or Cho and Sargent (2009)), a self-justified equilibrium can be interpreted as a stationary point of a learning process which itself is not modeled in the theory. The crucial difference is that in a self-justified equilibrium, an agent's forecasts can be incorrect in every step.

Both rational expectations equilibria and self-justified equilibria are special cases of a temporary equilibrium in this model. For the special case where

$$m_{y^a}(z, N_{\vec{\theta}}(s)) = M_{y^a}(s, N(s)) \text{ for all } s \in \mathbf{S},$$

we obtain a standard rational expectations equilibrium. Under the stated examples, as the set of admissible functions, $\mathbf{M}_{y^a}$ becomes sufficiently rich (e.g. includes all measurable functions) a self-justified equilibrium converges to a rational expectations equilibrium. The main contribution of this paper is to explore what happens if the agent is unable to approximate m_{y^a} perfectly. In this case, self-justified equilibria can be arbitrarily far from a rational expectations equilibrium.

3 Existence

To prove the existence of simple equilibria in heterogeneous agents models with incomplete markets, one needs to impose strong assumptions on fundamentals. Brumm et al. (2017) present one possible set of strong assumptions and argue that without strong assumptions, simple equilibria might fail to exist (Kubler and Polemarchakis (2004) provide simple counterexamples). We show that under the (strong) assumption of finite accounting proving existence is relatively straightforward.

3.1 Assumptions

We first make a number of fairly standard assumptions on fundamentals:

Assumption 1

1. *For each $y^a \in \mathbf{I}$ the Bernoulli-utility function $u_{y^a}(\cdot)$ is continuously differentiable, strictly increasing, strictly concave, and satisfies an Inada condition*

$$u'_{y^a}(x) \to \infty \text{ as } x \to 0.$$

   Individual endowments are positive, i.e.,

$$e_{y^a}(z) > 0 \text{ for all } z \in \mathbf{Z}.$$

2. *The set $\mathbf{\Theta}$ is compact, and for each $y^a \in \mathbf{I}$, the set $\mathbf{\Theta}_{y^a}$ is a closed convex cone containing $\mathbb{R}^J_+$.*

3. *The payoff functions, $f : \mathbb{R}^J_+ \times \mathbf{Z} \to \mathbb{R}^J$, are non-negative valued and continuous. Moreover, for any $i = 1, \dots, J$ and $j = 1, \dots, J$ the payoff $f_j(q, z)$ only depends on q_i if $\bar{\theta}_i > 0$.*

4. *For all $y^a \in \mathbf{I}$ and all $\theta^-_{y^a} \in \mathbf{\Theta}_{y^a}$*

$$\theta^-_{y^a} \cdot f(q, z) \geq 0 \text{ for all } q \in \mathbb{R}^J_+, z \in \mathbf{Z}.$$

Assumptions 1.1-1.3 are standard (see, e.g., Kubler and Schmedders (2003)). Assumption 1.4 is motivated by collateral and default. These constraints ensure that agents cannot borrow against future endowments. In our formulation, this is true independently of prices and could be justified if we allow for default (see Kubler and Schmedders (2003) for a detailed motivation) or if agents face appropriate borrowing constraints.

The crucial and non-standard assumption of the paper is that accounting is finite, i.e., that beginning of period portfolios lie in a finite set (or at least that agents perceive them to lie in a finite set). This simplifies the analysis dramatically, and we will argue below that it has few practical disadvantages. Formally, we make the following assumptions:

Assumption 2

1. *There is a finite set* $\widehat{\mathbf{S}} \subset \mathbf{S}$ *such that the support of the transition function* $\mathbb{Q}(\cdot|z, \vec{\theta})$ *is a subset of* $\widehat{\mathbf{S}}$ *for all* $z \in \mathbf{Z}$ *and all* $\vec{\theta} \in \mathbf{\Theta}$.

2. *The measure* $\mathbb{Q}(\cdot|z, \vec{\theta})$ *is continuous in* $\vec{\theta}$ *for all* $z \in \mathbf{Z}$, $\vec{\theta} \in \mathbf{\Theta}$.

Assuming that $\widehat{\mathbf{S}}$ contains ZG elements, we then can take $\mathbb{Q}(\cdot|z, \vec{\theta})$ to be a vector in the $ZG-1$ dimensional unit simplex, $\mathbf{\Delta}^{ZG-1}$. Assumption 2.2 then simply states that this vector changes continuously in $\vec{\theta}$.

From a practical point of view, the assumption seems innocuous. Because of finite precision arithmetic in scientific computations, almost any numerical method will lead to $\vec{\theta}^-$ lying on a (possibly very fine) grid. Assumption 2.2 then states that there is some randomness in the rounding error. However, from a technical point, the assumption turns out to be crucial. It is not clear which of our results hold true in the limit as the grid becomes dense in $\mathbf{\Theta}^-$. The assumption will allow us to obtain simple existence results below, but it is certainly not a standard assumption in this literature and it is not compatible with full rationality of individuals.

Assuming finite accounting has several economic justifications. One interpretation is that actual portfolios lie in $\mathbf{\Theta}^-$ but that agent cannot measure portfolios arbitrarily finely and make their decisions based on rounded values, exhibiting some degree of bounded rationality. Our preferred interpretation is that agents take the fact that beginning-of-period portfolios always lie on a finite grid as a technological constraint. This viewpoint seems natural when one thinks of the grid to be extremely fine. For this interpretation, let $\widehat{\mathbf{\Theta}}^- \subset \mathbf{\Theta}^-$ be a finite set, and assume that given $\vec{\theta}(z^t)$, we have

$$\vec{\theta}^-(z^{t+1}) \in \arg \min_{\vec{\theta}^- \in \widehat{\mathbf{\Theta}}^-} \|\bar{\theta} + \epsilon_{t+1} - \vec{\theta}^-\|_2,$$

with $\bar{\theta}_{y^a} = \theta_{y^{a-1}}$ for all $a = 2, \ldots A$, $y^a \in \mathbf{Y}^a$ and $\bar{\theta}_{y^1} = 0$ for all $y^1 \in \mathbf{Y}$. In this formulation, ϵ_t should be interpreted as a (small) rounding error, and it is assumed that the support of $\epsilon(\cdot)$ is centered around zero, convex, and sufficiently small. We assume that ϵ_t is i.i.d. and that it only affects the current rounding error. In this formulation, it is easy to verify that Assumption 2.2 holds whenever ϵ_t has a continuous density function. Of course, the formulation of the agent's problem in (1) now potentially (depending on the set of admissible forecasting functions, $\mathbf{M}_{y^a}$) builds in another layer of bounded rationality, since the correct dynamic programming problem of an agent is no longer a standard convex program.

Since we assumed $\widehat{\mathbf{S}}$ to be finite and to contain GZ elements, for fixed $\vec{M} \in \mathbf{M}$, a selection of the temporary equilibrium correspondence can be viewed as a vector $N \in \left(\mathbb{R}_+^I \times \mathbf{\Theta} \times \mathbb{R}_+^J\right)^{GZ}$. We write $M_{y^a}(\cdot; N, \mathbb{Q}) = R_{y^a}(N, \mathbb{Q})$, and it is useful to note that for each z the function $M_{y^a}(z, \cdot)$ is defined on a subset of the Euclidean space. To make this more explicit, we write

$$M_{y^a}(z, \vec{\theta}^-, \vec{x}, \vec{\theta}, q, \nu, \mu),$$

where $\nu \in \left(\mathbb{R}_+^I \times \mathbf{\Theta} \times \mathbb{R}_+^J\right)^{GZ}$, and $\mu \in \Delta^{ZG-1}$.

We make the following reduced-form assumption on forecasting-functions:

Assumption 3

1. *For all $\mu \in \Delta^{ZG-1}$, all $\nu \in \left(\mathbb{R}_+^I \times \mathbf{\Theta} \times \mathbb{R}_+^J\right)^{GZ}$, all $s \in \widehat{\mathbf{S}}$ and all $\vec{\theta} \in \mathbf{\Theta}$, $\vec{x} \in \mathbb{R}_{++}^I, q \in \mathbb{R}_{++}^J$ the function $M_{y^a}(s, \vec{x}, \vec{\theta}, q; \nu, \mu)$ is jointly continuous in $\vec{x}$, $\vec{\theta}$, q, ν and μ.*

2. *For each agent $y^a \in \mathbf{I}$, all functions in $\mathbf{M}_{y^a}$ are uniformly bounded above, i.e., there is some $\bar{m}$ such that*

$$M_{y^a,j}(z, \vec{\theta}^-, q, \vec{\theta}, \vec{x}) < \bar{m} \text{ for all } z \in \mathbf{Z}, \vec{\theta}^-, q, \vec{\theta}, \vec{x}, j \in \mathbf{J} \text{ and all } M \in \mathbf{M}_{y^a}.$$

Assumption 3.1 is relatively standard and very likely to be satisfied in applied settings. Assumption 3.2 might appear to be rather strong. However, with enough structure on the sets $\mathbf{M}_{y^a}$, and with a more concrete description of the economy, one can typically find these bounds in an overlapping-generations setting. Clearly, with strictly positive endowments and borrowing constraints, all functions in $\mathbf{M}_{y^{A-1}}$ are bounded. A backward induction argument can then be used to justify Assumption 3.2. It is clear that in a framework with infinitely lived agents, this becomes much more difficult.

3.2 The main theoretical result

With these assumptions, the existence of a self-justified equilibrium reduces to the existence of a finite-dimensional fixed point. The main result of this section thus reads as follows:

Theorem 1 *Under Assumptions 1-3 there exists a self-justified equilibrium.*

Proof. We decompose the economy into sub-economies for each $s \in \mathbf{S}$ and construct a map from a compact and convex set of all agents' choices, prices, probabilities, μ, and forecasts, M_s, into itself.

Using Kakutani's theorem (see Border (1989)), we can show that this map has a fixed point, and we finish the proof by demonstrating that this is a self-justified equilibrium.

First, we need to find a suitable, convex and compact domain for the map. Assumption 1.3 implies that there exist l, r such that whenever $\vec{\theta} \in \mathbf{\Theta}$,

$$l \leq \theta_{y^a,j} \leq r \text{ for all } y^a \in \mathbf{I}, j \in \mathbf{J}.$$

Let the set of admissible asset holdings be $\mathbf{T} = [l, r]^J$, and let the set of admissible consumptions be

$$\mathbf{X} = \left[0, \max_{z \in \mathbf{Z} y^a \in \mathbf{I}} \frac{e_{y^a}(z) + d(z)}{\eta(y^a)}\right].$$

We construct a upper-hemi-continuous, non-empty and convex-valued correspondence, $\mathbf{\Phi}$, mapping choices and prices at each element in $\widehat{\mathbf{S}}$ as well as a probability measure over $\widehat{\mathbf{S}}$, to itself, which has a fixed point,

$$\mathbf{\Phi} : (\mathbf{X}^I \times \mathbf{T}^I \times \Delta^J)^{GZ} \times \mathbf{\Delta}^{GZ} \rightrightarrows (\mathbf{X}^I \times \mathbf{T}^I \times \Delta^J)^{GZ} \times \mathbf{\Delta}^{GZ}.$$

For all $y^a \in \mathbf{I}$ and all $s \in \widehat{\mathbf{S}}$, let

$$\begin{aligned} \mathbf{\Phi}_{y^a,s}((x_t, p_t, q_t)_{t \in \widehat{\mathbf{S}}}) \quad = \quad & \arg \max_{x \in \mathbf{X}, \theta \in \mathbf{\Theta}_{y^a} \cap \mathbf{T}} u_{y^a}(x) + \widetilde{M}_{y^a}(z, \vec{\theta}_s^-, \vec{x}_s, \vec{\theta}_s, \vec{q}_s) \cdot \theta \\ & \text{s.t.} \\ & (x - e_{y^a}(z)) + \theta . \frac{1}{p_s} q_s - \theta^- \cdot f(\frac{1}{p_s} q_s, z) \leq 0, \end{aligned}$$

where

$$\widetilde{M}_{y^a} = R_{y^a}(\nu, \mu), \quad \nu = (x_s, \theta_s, \frac{1}{p_s} q_s)_{s \in \widehat{\mathbf{S}}} \quad . \tag{4}$$

Define the price-players best response as

$$\mathbf{\Phi}_{0,s}(\vec{\theta}_s, \vec{x}_s) = \arg \max_{(p,q) \in \mathbf{\Delta}^J} p(\sum_{y^a \in \mathbf{I}} \eta(y^a)(x_{y^a,s} - e_{y^a}(z) - d(z))) + q \cdot (\sum_{y^a \in \mathbf{I}} \eta(y^a)(\theta_{y^a,s} - \bar{\theta})),$$

and let

$$\mathbf{\Phi}_\mu((\vec{\theta}_s)_{s \in \mathbf{S}}, \mu) = (\mu(s) \sum_{s' \in \mathbf{S}} \mathbb{Q}(s' | z, \vec{\theta}_s)(s'))_{s \in \mathbf{S}}.$$

Assumptions 1 - 3 guarantee that the mapping

$$\mathbf{\Phi} = \times_{s \in \mathbf{S}} \left((\times_{y^a \in \mathbf{I}} \mathbf{\Phi}_{y^a,s}) \times \mathbf{\Phi}_{0,s} \right) \times \mathbf{\Phi}_\mu$$

is non-empty, convex valued, and upper hemi-continuous. By Kakutani's fixed point theorem, there exists a fixed point. Assumption 1 guarantees that the additional constraints imposed by forcing choices to lie in $\mathbf{T} \times \mathbf{X}$ are not binding, and hence the forecasting functions defined by (4) at the fixed point, together with $\mathbb{Q}^* = \mu$ and the equilibrium values constitute a self-justified equilibrium.

□

The discretization of the state-space enables us to prove a very simple result. Without this, strong assumptions would be needed to ensure the existence of a recursive rational expectations equilibrium (see Brumm et al. (2017)), and the existence of a self-justified equilibrium thus would remain an open problem.

4 A tractable version of the model

To make the concept of self-justified equilibrium tractable, it is essential to find a simple domain for agents' forecasts. So far, we allowed forecasts to depend on all current endogenous variables which is clearly too general to be useful in applications. In particular the agents will face a *curse of dimensionality* (Bellman (1961)) when trying to approximate and evaluate functions on very high-dimensional domains.

The structure of the equilibrium suggests that it suffices to base forecasts only on the current shock and on (new) portfolio-choices across agents. As we will argue in the examples below, this often yields excellent results and is well suited for computational purposes. For the remainder of the paper, we assume that agents forecasts do not depend on the current endogenous state, on prices, or on consumption choices, and we write

$$M_{y^a} : \mathbf{Z} \times \mathbf{\Theta} \to \mathbb{R}^J_+.$$

From an economic point of view, it might make more sense to assume that households base their forecasts on current prices and possibly lagged shocks, since these are easily observable. However, it is clear that current portfolio-choices determine the (natural) endogenous state in the next period, and it is therefore a good choice from the perspective of the computational modeler.

In many applications, the set of current asset holdings, $\mathbf{\Theta}$, will be very high dimensional. Both as a matter of realism and for tractability, it seems advantageous to assume that the agents only take a low dimensional subspace of the actual state-space and use this for their forecasts. In our tractable version of the model, we assume that agents project $\vec{\theta}$ into a lower dimensional subspace and use the latter for the forecasts. That is to say, M_{y^a} is actually not defined on $\mathbf{\Theta}$, but instead on a subset of $\mathbb{R}^d_+$, with d typically being much smaller than IJ.

4.1 What are good projections

Given a $d \times IJ$ projection matrix $W_{y^a,z,j}$ for a given agent y^a, shock z and asset j, we define the set of admissible forecasting-functions to be

$$\mathbf{M}_{y^a,z,j}(d) = \{f : \mathbf{\Theta}^W_{y^a,z} \to \mathbb{R}\},$$

where

$$\mathbf{\Theta}^W_{y^a,z} = \{\phi \in \mathbb{R}^d : \phi = W^T_{y^a,z,j}\theta, \theta \in \mathbf{\Theta}_{y^a}\}.$$

For each shock $\bar{z} \in \mathbf{Z}$ and each asset $j \in \mathbf{J}$ the agent's forecasting function solves

$$\min_{M \in \mathbf{M}_{y^a,\bar{z},j}(d)} \int_{\vec{\theta} \in \mathbf{\Theta}} (M(\bar{z}, W^T_{y^a,\bar{z},j} N_{\vec{\theta}}(\bar{z}, \vec{\theta})) - m_{y^a,j}(\bar{z}, N_{\vec{\theta}}(\bar{z}, \vec{\theta})))^2 d\mathbb{Q}^*(\vec{\theta}|\bar{z}). \tag{5}$$

At this point, we take the sets $\mathbf{M}_{y^a}(d)$ as given. In our applications below, we take it to be the set of all continuous functions.

4.1.1 Choosing the projection matrix

We first discuss an optimal choice of $W_{y^a,z,j}$ given a fixed dimension d. In choosing $W_{y^a,z,j}$, two extremes are conceivable. First, one could view the projection matrices, $W_{y^a,z,j}$, $y^a \in \mathbf{I}$, $z \in \mathbf{Z}$, $j \in \mathbf{J}$ as fundamentals–agents have certain technologies that allow them to observe projections of the state into lower dimensional subspaces (for example, they observe parts of the wealth distribution). Second, one could take d as given and require that the matrices $W_{y^a,z,j}$ are optimal in the sense that they minimize the mean squared error in Equation (5). This would fit our definition of a self-justified equilibrium, but unfortunately, it turns out to amount to a non-convex optimization problem that is generally not tractable.[6] The problem is so complicated that it is not consistent with the whole idea of boundedly rational agents. Moreover, we wish to develop a theory of optimal projections that are independent of the sets $\mathbf{M}_{y^a}$ and only depend on the function that needs to be approximated. This allows us to disentangle the methods used to approximate a d-dimensional function from the method used to find an optimal projection of the IJ-dimensional vector $\vec{\theta}$ into $\mathbb{R}^d$.

In the following, we take the approach that lies between the two extremes, and we believe that it has an elegant micro-foundation and it turns out to be very tractable. In that approach, agents choose a matrix $W_{y^a,z,j}$ to minimize the unexplained part of the variations in $\hat{m}_{y^a,z,j}$ as measured by the mean squared derivative of $\hat{m}$ with respect to the orthogonal complement of the variables used for forecasting.

To this end, we assume that each agent y^a uses his own portfolio as the primary factor that influences next period's marginal utilities. This is a natural assumption, and if asset prices would only depend on the current and lagged shock, this would yield an optimal solution. However, in our model asset prices vary with the distribution of assets in the economy. We therefore write θ_{-y^a} to denote the portfolio of all other agents in the economy besides agent y^a, and we write $\vec{\theta} = (\theta_{y^a}, \theta_{-y^a})$. Clearly, θ_{-y^a} influences the agent's marginal utility for assets because it influences all future prices. We assume that the agent assesses the variability of his future marginal utility that is caused by this variability of future prices by the mean squared gradient, and chooses an "active subspace" (see Constantine et al. (2014)) to ensure that the unexplained part of fluctuations is minimized.

[6] In Appendix 7, we describe some of the difficulties that arise if one requires the matrix to be chosen optimally.

To formalize this idea, let $D = IJ - J$ be the dimension of other agents' asset holdings and write forecasts as $M_{y^a,j}(z, \theta_{y^a}, W_{y^a,z,j}\theta_{-y^a})$. Without loss of generality we assume that $W_{y^a,z,j}$ is an element of the d-dimensional Stiefel-manifold in $\mathbb{R}^D$, i.e.,

$$W_{y^a,z,j} \in \mathbf{V}_d(\mathbb{R}^D) = \left\{ A \in \mathbb{R}^{D\times d} : A^T A = I_{d\times d} \right\},$$

where $I_{d\times d}$ is the $d \times d$ identity matrix. Given a candidate $d \times D$ matrix $V_1 \in \mathbf{V}_d(\mathbb{R}^D)$, there is a $V_2 \in \mathbf{V}_{D-d}(\mathbb{R}^D)$ such that

$$[V_1, V_2] \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} = I_{D\times D},$$

and we can write

$$m_{y^a,j}(z, \vec{\theta}) = m_{y^a,j}\left(z, \left(\theta_{y^a}, [V_1 V_2] \begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} \theta_{-y^a}\right)\right) = m_{y^a,j}\left(z, \theta_{y^a}, V_1 V_1^T \theta_{-y^a} + V_2 V_2^T \theta_{-y^a}\right).$$

Defining $\phi_1 = V_1^T \theta_{-y^a}$ and $\phi_2 = V_2^T \theta_{-y^a}$ we obtain a function

$$\widehat{m}_{y^a,j}(z, \theta_{y^a}, \phi_1, \phi_2) = m_{y^a,j}\left(z, \theta_{y^a}, V_1\phi_1 + V_2\phi_2\right).$$

Strengthening Assumption 2.2, we assume that $\widehat{m}_{y^a}$ is continuously differentiable in θ_{-y^a} ($\mathbb{Q}^*$-a.e.). Given our justification for finite accounting, this simply amounts to assuming that the transition probability $\mathbb{Q}(\cdot|z, \vec{\theta})$ is continuously differentiable in θ and therefore does not seem substantially stronger than the original assumption.

We assume that the agent approximates the function $\widehat{m}_{y^a,j}$ using only (θ_{y^a}, ϕ_1). For simplicity assume for now that $\mathbf{M}_{y^a,j}$ consists of all (Borel-measurable) functions. For his case, we obtain

$$M_{y^a,j}(z, \theta_{y^a}, \phi_1) = \int_{\phi_2} \hat{m}_{y^a,j}(z, \phi_1, \phi_2) d\mathbb{Q}^*(\phi_2 | z, \theta_{y^a}, \phi_1),$$

where $\widehat{\mathbb{Q}}^*(z, (\theta_{y^a}, \phi_1, \phi_2))$ denotes the invariant distribution over

$$(z, (\theta_{y^a}\phi_1, \phi_2)) = (z, N_{\theta_{y^a}}(s), V_1 N_{\theta_{-y^a}}(s), V_2 N_{\theta_{-y^a}}(s)),$$

which is induced by $\mathbb{Q}^*$, and $\widehat{\mathbb{Q}}^*(\phi_2|\theta_{y^a}, \phi_1, z)$ denotes the invariant distribution of ϕ_2 conditional on z, θ_{y^a}, and ϕ_1.

This approximation is justified if the impact of ϕ_2 on the function $\widehat{m}_{y^a,j}$ is relatively small. How do agents decide that the effect of ϕ_2 on next period's marginal utility is small? We assume in this paper that they use the squared derivative with respect to ϕ_2, averaged along the stationary distribution, to measure the variability[7] with respect to ϕ_2 and define

$$\xi_{y^a,z,j}(V_1, V_2) = \int_{(\theta_{y^a}, \phi_1, \phi_2)} \left(\nabla_{\phi_2} \widehat{m}_{y^a,j}\left(z, \theta_{y^a}, \phi_1, \phi_2\right)\right)^T \left(\nabla_{\phi_2} \widehat{m}_{y^a,j}\left(z, \theta_{y^a}, \phi_1, \phi_2\right)\right) d\widehat{\mathbb{Q}}^*(\theta_{y^a}, \phi_1, \phi_2 | z),$$

[7] Sobol and Kucherenko (2009) discuss several different approaches to estimate the influence of individual factors and groups of factors and show that many of them can be effectively bounded by the average squared gradient of the function.

where for $x \in \mathbb{R}^D$,

$$\nabla_x f(x,y) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_D} \end{pmatrix},$$

and the partial derivatives are taken to be one-sided derivatives at the boundary of the domain.

We assume that $W_{y^a,z,j} = V_1$ solves

$$\min_{(V_1 V_2) \in \mathbf{V}_{IJ-J}(\mathbb{R}^{IJ-J})} \xi_{y^a,z,j}(V_1, V_2) \tag{6}$$

It turns out that there is a very simple characterization of an optimal solution to this minimization problem - in stark contrast to the case where the projection is chosen to minimize the mean squared average forecasting error. In computational sciences, this is known as the "classical" active subspace approach (see Constantine et al. (2014)).

Note that

$$\nabla_{\phi_2} \widehat{m}_{y^a,j}(z, \theta_{y^a}, \phi_1, \phi_2) = \nabla_{\phi_2} m_{y^a,j}\left(z, \theta_{y^a}, V_1\phi_1 + V_2\phi_2\right) = V_2^T \nabla_{\theta_{-y^a}} m_{y^a,j}(z, \theta_{y^a}, \theta_{-y^a}),$$

and that

$$\begin{aligned}
&\int_{(\theta_{y^a},\phi_1,\phi_2)} \left(\nabla_{\phi_2} \widehat{m}_{y^a,j}\left(z, \theta_{y^a}, \phi_1, \phi_2\right)\right)^T \left(\nabla_{\phi_2} \widehat{m}_{y^a,j}\left(z, \theta_{y^a}, \phi_1, \phi_2\right)\right) d\widehat{\mathbb{Q}}^*(\theta_{y^a}, \phi_1, \phi_2 | z) = \\
&\int_{(\theta_{y^a},\phi_1,\phi_2)} tr\left(\left(\nabla_{\phi_2} \widehat{m}_{y^a,j}\left(z, \theta_{y^a}, \phi_1, \phi_2\right)\right)^T \left(\nabla_{\phi_2} \widehat{m}_{y^a,j}\left(z, \theta_{y^a}, \phi_1, \phi_2\right)\right)\right) d\widehat{\mathbb{Q}}^*(\theta_{y^a}, \phi_1, \phi_2 | z) = \\
&\int_{(\theta_{y^a},\phi_1,\phi_2)} tr\left(\left(\nabla_{\phi_2} \widehat{m}_{y^a,j}\left(z, \theta_{y^a}, \phi_1, \phi_2\right)\right) \left(\nabla_{\phi_2} \widehat{m}_{y^a,j}\left(z, \theta_{y^a}, \phi_1, \phi_2\right)\right)^T\right) d\widehat{\mathbb{Q}}^*(\theta_{y^a}, \phi_1, \phi_2 | z).
\end{aligned}$$

Therefore, solving (6) amounts to solving

$$\min_{V_2 \in \mathbf{V}_{D-d}(\mathbb{R}^D)} tr\left(V_2^T C_{y^a,z,j} V_2\right),$$

where

$$C_{y^a,z,j} = \int_{(\theta_{y^a},\theta_{-y^a})} \left(\nabla_{\theta_{-y^a}} m_{y^a,j}\left(z, \theta_{y^a}, \theta_{-y^a}\right)\right) \left(\nabla_{\theta_{-y^a}} m_{y^a,j}\left(z, \theta_{y^a}, \theta_{-y^a}\right)\right)^T d\widehat{\mathbb{Q}}^*(\vec{\theta}|z). \tag{7}$$

Denoting $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_D$ the eigenvalues of $C_{y^a,z,j}$, it follows from the Courant-Fischer Theorem (see, e.g., Horn and Johnson (1985), Theorem 4.2.11) that since $C_{y^a,z,j}$ is a symmetric matrix the minimum is given by $\sum_{i=d+1}^{D} \lambda_i$ and one (not unique) minimizer is given by the matrix of associated eigenvectors.

This suggests the following construction of projection-matrices. Since $C_{y^a,z,j}$ is symmetric positive definite, it admits the form

$$C_{y^a,z,j} = V \Lambda V^T, \tag{8}$$

where $\Lambda = \text{diag}(\lambda_1, \cdots, \lambda_D)$ is a diagonal matrix containing the eigenvalues of C in decreasing order, $\lambda_1 \geq \cdots \geq \lambda_D \geq 0$, and $V \in \mathbf{V}_D(\mathbb{R}^D)$ is an orthonormal matrix whose columns correspond to the eigenvectors of C. separating the d largest eigenvalues from the rest,

$$\Lambda = \begin{bmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{bmatrix}, \quad V = \begin{bmatrix} V_1 & V_2 \end{bmatrix}, \tag{9}$$

(here $\Lambda_1 = \text{diag}(\lambda_1, \ldots, \lambda_d), V_1 = [v_{11} \ldots v_{1d}]$, and Λ_2, V_2 are defined analogously), and setting the projection matrix to

$$W_{y^a,z,j} = V_1. \tag{10}$$

Intuitively, $W_{y^a,z,j}$ rotates the input space in such a manner that the directions associated with the largest eigenvalues correspond to directions of maximal function variability (see Constantine (2015)).

Our above discussion gives directly rise to the following proposition (which is Lemma 2.2 in Constantine et al. (2014)), which makes the active subspace method very attractive for our model.

PROPOSITION 1 *The mean squared gradients of $\widehat{m}$ with respect to ϕ_1 and ϕ_2 satisfy*

$$\int_{(\theta_{y^a},\phi_1,\phi_2)} \left(\nabla_{\phi_1} \widehat{m}_{y^a,j}\left(z, \theta_{y^a}, \phi_1, \phi_2\right)\right)^T \left(\nabla_{\phi_1} \widehat{m}_{y^a,j}\left(z, \theta_{y^a}, \phi_1, \phi_2\right)\right) d\widehat{\mathbb{Q}}^*(\theta_{y^a}, \phi_1, \phi_2 | z) \leq \lambda_1 + \ldots + \lambda_d$$

and

$$\int_{(\theta_{y^a},\phi_1,\phi_2)} \left(\nabla_{\phi_2} \widehat{m}_{y^a,j}\left(z, \theta_{y^a}, \phi_1, \phi_2\right)\right)^T \left(\nabla_{\phi_2} \widehat{m}_{y^a,j}\left(z, \theta_{y^a}, \phi_1, \phi_2\right)\right) d\widehat{\mathbb{Q}}^*(\theta_{y^a}, \phi_1, \phi_2 | z) \geq \lambda_{d+1} + \ldots + \lambda_{IJ-J}.$$

Both inequalities hold with equality if (V_1, V_2) are chosen according to (8) and (9).

As mentioned, the matrix consisting of the eigenvectors associated with the d largest eigenvalues is not the unique solution to (6) - nevertheless we will assume in the following that each projection matrix $W_{y^a,z,j}$, $y^a \in \mathbf{I}$, $z \in \mathbf{Z}$, is determined by (8) and (9).

4.1.2 Determining d

We assume[8] that the dimension of the projective space of agent y^a is independent of the shock z, and the asset j, and is determined by two parameters, $\epsilon^1_{y^a}, \epsilon^2_{y^a}$. First we say that a dimension d is acceptable for the agent y^a if for all $z \in \mathbf{Z}$ and all asset $j \in \mathbf{J}$

$$\int_{\vec{\theta} \in \Theta} \left| \frac{M_{y^a,j}(z, W^T_{y^a,z,j} N_{\vec{\theta}}(z, \vec{\theta})) - m_{y^a,j}(z, N_{\vec{\theta}}(z, \vec{\theta}))}{m_{y^a,j}(z, N_{\vec{\theta}}(z, \vec{\theta}))} \right| d\mathbb{Q}^*(\vec{\theta} | \bar{z}) < \epsilon^1_{y^a}, \tag{11}$$

where $M_{y^a,j}$ solves (5). For sufficiently small $\epsilon^1_{y^a}$, this criterion is comparable to what is used in Krusell and Smith (1997) and much of the literature that follows it.

[8]This simplifies the exposition. There is no reason why one cannot have the sensitivity levels depend on shocks and/or assets.

However, in some models, this might require too much rationality and agents might simply not be able to achieve very accurate predictions. Therefore we assume in addition that the agent is satisfied with a $(IJ - J) \times d$ matrix V_1 if adding one additional dimension (i.e. instead of using V_1 using a $D \times d + 1$ matrix $\widetilde{V}_1$) the variability of the unexplained part is not reduced significantly,

$$\frac{\min_{(\widetilde{V}_1, \widetilde{V}_2) \in \mathbf{V}_D(\mathbb{R}^D)} \xi(\widetilde{V}_1, \widetilde{V}_2)}{\min_{(V_1 V_2) \in \mathbf{V}_D(\mathbb{R}^D)} \xi(V_1, V_2)} < \epsilon_{y^a}^2. \tag{12}$$

Proposition 1 above implies that this can be written equivalently as

$$1 + \frac{\lambda_d}{\sum_{i=d+1}^{D} \lambda_i} > \frac{1}{\epsilon_{y^a}^2}.$$

These are the criteria used in our numerical examples below.

4.2 Self-justified equilibrium in a tractable economy

With this, an economy is described by assets, trading constraints, preferences, admissable forecasting functions and endowments, but also $\epsilon_{y^a}^1, \epsilon_{y^a}^2$ for all (active) agents $y^a \in \mathbf{I}$. We allow ϵ to depend on the agent to incorporate heterogeneity in forecasts into the model.

A "dimension-reduced self-justified equilibrium" then consists of dimensions d_{y^a}, $D \times d_{y^a}$ matrices, $W_{y^a,z,j}$, $z \in \mathbf{Z}, j \in \mathbf{J}$, and forecasts $M_{y^a,z,j} \in \mathbf{M}_{y^a,j}(d_{y^a})$ for all $y^a \in \mathbf{I}$ as well as a selection $N(\cdot)$ of the temporary equilibrium correspondence, $\mathbf{N}_{\vec{M}}(\cdot)$, and measure $\mathbb{Q}^*$ on $(\mathbf{S}, \mathcal{S})$, such that

- for each agent, each shock, and each asset $W_{y^a,z,j} = V_1$ where V_1 satisfies (8), (9).
- for each agent, d_{y^a} satisfies (11) or (12).
- $\mathbb{Q}^*$ is invariant given the law of motion induced by $N(\cdot)$ and by $\mathbb{Q}(\cdot, \cdot)$. That is to say, for all $\mathbf{B} \in \mathcal{S}$
$$\mathbb{Q}^*(\mathbf{B}) = \int_{s \in \mathbf{S}} \mathbb{Q}(\mathbf{B}|z, N_{\vec{\theta}}(s)) d\mathbb{Q}^*(s).$$
- For each y^a $a < A$ and each $z \in \mathbf{Z}$, $j \in \mathbf{J}$,
$$M_{y^a,z,j} \in \arg \min_{M \in \mathbf{M}_{y^a,\bar{z},j}(d)} \int_{\vec{\theta} \in \mathbf{\Theta}} (M(\bar{z}, W_{y^a,\bar{z},j}^T N_{\vec{\theta}}(\bar{z}, \vec{\theta})) - m_{y^a,j}(\bar{z}, N_{\vec{\theta}}(\bar{z}, \vec{\theta})))^2 d\mathbb{Q}^*(\vec{\theta}|\bar{z}). \tag{13}$$

It is clear that this is a special case of our general definition. However, it should be noted that since d_{y^a} are integers, Assumption 3 is not likely to hold as stated.

To prove existence for fixed d_{y^a}, a version of the proof of Theorem 1 can be used, as long as the continuity assumptions are strengthened to differentiability. An important point to note is that the eigenvectors of C will change continuously as elements of C change continuously (keeping the matrix symmetric and definite) - see Horn and Johnson (1985).

In the application below, we will assume that the sets $\mathbf{M}_{y^a}(d)$ consist of all continuous functions—it is easy to see that in this case $M_{y^a,z}(z, W_{y^a,z}\vec{\theta})$ is equal to the conditional expectation of $\hat{m}$, given ϕ_2 (if this is continuous), i.e.,

$$M_{y^a}(z, \theta_{y^a}, \phi_1) = \int_{\phi_2} \hat{m}(z, \phi_1, \phi_2) d\widehat{\mathbb{Q}}^*(\phi_2 | z, \theta_{y^a}, \phi_1).$$

In the next part of the paper, we will describe computational methods to solve for this self-justified equilibrium efficiently. For this, is it important to first note that we impose "too much" rationality on the agent to be able to solve his problem exactly. The fact that forecasts minimize the least-squared error under the (a priori unknown) invariant distribution makes it impossible to compute the forecast accurately. Instead, we will have to resort to Monte-Carlo simulations and approximate the invariant distribution by finitely many draws. At the same time, we are also unable to compute the conditional expectation exactly. Hence, we will need to approximate it using a numerical method. To this end, we choose Gaussian process regressions. The latter has been proven to be very useful in other contexts (see Scheidegger and Bilionis (2017), and references therein).

It is computationally much easier to assume that $\mathbf{M}_{y^a}$ consists of all linear functions. This often yields acceptable approximations but in our examples below errors would be very high. Restricting agents' forecasts to be linear in today's aggregated asset-choices across agents seems overly strong.

5 Computation

To numerically approximate a self-justified equilibrium in a model where agents use optimal projections to form their forecasts, the main computational issues are (i) how to find projections that minimize the mean squared gradient of the unexplained variation, (ii) how to approximate the (low) dimensional forecasting functions well, and (iii) how to solve for them.

5.1 Finding $W_{y^a,z}$

As explained in detail above, to compute optimal projections, we use so-called active subspace methods developed by Constantine et al. (2014) (see also Scheidegger and Bilionis (2017)). This version of dimension-reduction turns out to fit well our economic model and produces very good results in our examples below. Reiter (2010) considers an alternative approach which is better suited for models with 100,000 agents which differ only in their asset holdings but it does not fit well into our framework where we target models with 100 - 1000 heterogeneous agents.

It is impossible to evaluate (7) exactly. Instead, the usual procedure is to approximate the integral (7) via Monte Carlo, that is, assuming that the observed inputs are drawn from a simulated path of the economy, one approximates C using the observed gradients by

$$C_N = \frac{1}{N} \sum_{i=1}^{N} g^{(i)} \left(g^{(i)} \right)^T . \tag{14}$$

In practice, the eigenvalues and eigenvectors of C_N are found using the singular value decomposition of C_N. Clearly in our framework, the gradient, g^i cannot be evaluated analytically (in fact they are not guaranteed to exist), so we generally approximate (14) by finite differences.[9]

Active subspace methods are attractive in practice because it turns out that for many multivariate functions that occur for example in engineering models and the natural sciences, one observes sharp drops in the spectrum of C at relatively small values of d (see Constantine (2015) and the references therein). In our iterative computational strategy described below, we start with $W_{y^a,z} = 0$, i.e., the agents only use their own asset holdings to forecasts future marginal utilities. If the resulting forecasting errors are large, we update the matrix along the iterations.

To make the algorithm operable beyond the case of linear functions, we first need to understand how to conveniently approximate functions on arbitrary domains. For this, we use so-called Gaussian process (GP) regression, which is a method from Bayesian statistics (see, e.g., Barry (1986)) and now often used in supervised machine learning (see, e.g., Rasmussen and Williams (2005)). There are many examples in the literature where the combination of GP-regression and active subspaces proves very fruitful (see, e.g., Tripathy et al. (2016), or Scheidegger and Bilionis (2017)).

5.2 Gaussian process regression

Given a data set $\{(x^{(i)}, y^{(i)}) \,|i = 1, ..., n\}$ consisting of n vectors $x^{(i)} \in \mathbb{R}^d$ and corresponding, potentially noisy, observations,

$$y^{(i)} = f(x^{(i)}) + \epsilon_i, \tag{15}$$

we want to construct a function $\hat{f}$ that trades of smoothness and approximation in an optimal way. Given a reproducing kernel Hilbert space, $\mathbf{H}$ with a positive definite kernel $K(x, y)$, classical regularization theory (see, e.g., Evgeniou et al. (2000), and there references therein) solves the following problem:

$$\min_{f \in \mathbf{H}} \frac{1}{n} \sum_{i=1}^{n} (y_i - f(x^{(i)}))^2 + \lambda \|f\|_K^2, \tag{16}$$

where $\|.\|_K$ is the norm defined by $K(\cdot)$. It can be shown that the solution to (16) can be written as

$$\hat{f}(x) = \sum_{i=1}^{n} \alpha_i K(x, x_i), \tag{17}$$

where α solves

$$(K + \lambda I)\alpha = y, \quad (K)_{ij} = K(x_i, x_j), \; y = (y^{(1)}, \dots, y^{(n)})^T.$$

As Rasmussen and Williams (2005) point out, the representation of f can also be obtained as the posterior mean of a Gaussian process. The advantages of that formulation are that it naturally

[9] Alternatively, one may use the Bayesian information criterion to discover the active subspace. For the latter, see Tripathy et al. (2016).

leads to systematic ways for choosing $K(\cdot)$ and λ and that the standard deviation of the Gaussian process can be used as an indication of goodness of fit. We provide a very brief introduction to Gaussian process regression based on Rasmussen and Williams (2005) (see also Scheidegger and Bilionis (2017) for a more detailed introduction).

A Gaussian process is a collection of random variables, any finite number of which have a joint Gaussian distribution. We say that $f(\cdot)$ is a GP with *mean function* $m(\cdot)$ and *covariance function* $k(\cdot,\cdot)$, and write

$$f(\cdot) \sim \mathrm{GP}\left(m(\cdot), k(\cdot,\cdot)\right) \tag{18}$$

The covariance function can be chosen, but must be positive semi-definite and symmetric. Throughout our work, we either use the so-called *square exponential* (SE)

$$k_{\mathrm{SE}}(x,x') = \sigma^2 \exp\left\{-\frac{1}{2}\sum_{i=1}^{r}\frac{(x_i - x_i')^2}{\ell_i^2}\right\}, \tag{19}$$

or the Matern-3/2 covariance kernel:

$$k_{\mathrm{mat}}(x,x') = \sigma^2\left(1+\sqrt{3}\sum_{i=1}^{l}\frac{(x_i - x_i')^2}{\ell_i^2}\right)\exp\left(-\sqrt{3}\sum_{i=1}^{l}\frac{(x_i - x_i')^2}{\ell_i^2}\right), \tag{20}$$

where $\ell_i > 0$ and $\sigma > 0$ in both kernels denotes the characteristic length-scale of the i-th input, and the signal strength. The "hyper-parameters" of the covariance function are typically estimated by maximum-likelihood (see Scheidegger and Bilionis (2017)). In our implementation, we use a self-customized version of the software package Limbo (see Cully et al. (2018)), which provides several options for this step.

The specification of the mean function $m(\cdot)$ is similar to the specification of a prior in Bayesian statistics. In our numerical examples below, we set $m(\cdot) = 0$. Note that this does not imply the posterior mean (which we use as our approximating function) is zero. Rasmussen and Williams (2005, Chapter 2.7) discuss several ways to model a mean function.

Let us define the matrix of so-called "training inputs" as

$$X = \left\{x^{(1)}, \ldots, x^{(n)}\right\}. \tag{21}$$

Given X, we have a Gaussian prior on the corresponding response outputs,

$$\vec{f} = \left\{f\left(x^{(1)}\right), \ldots, f\left(x^{(n)}\right)\right\}.$$

In particular,

$$\vec{f}|X \sim \mathcal{N}\left(m, K\right), \tag{22}$$

where $m := m(X) \in \mathbb{R}^n$ being the mean function evaluated at all points in X, and $K \in \mathbb{R}^{n\times n}$ is the covariance matrix with

$$K_{ij} = k(x^{(i)}, x^{(j)}), \tag{23}$$

and $k(x^{(i)}, x^{(j)})$ given by (19) or (20).

In order to derive an explicit expression for the likelihood, we assume that the noise-terms ϵ_i in (15) are i.i.d. normal with mean zero and variance s^2. This assumption is not going to be satisfied in our application. However, it turns out that the method works well even if the noise is not i.i.d. normal. Using the independence of the observations, we obtain

$$y|\vec{f}, s \sim \mathcal{N}\left(y\middle|\vec{f}, s^2 I_n\right). \tag{24}$$

The *likelihood*-function of the observations is then given by

$$y|X, s \sim \mathcal{N}\left(y|m, K + s^2 I_n\right). \tag{25}$$

Bayes' rule combines the prior GP (see (18) with the likelihood (see (25) and yields the *posterior* GP

$$f(\cdot)|X, y, s \sim \mathcal{GP}\left(f(\cdot)\middle|\tilde{m}(\cdot), \tilde{k}(\cdot, \cdot)\right), \tag{26}$$

where the *posterior* mean and covariance functions are given by

$$\tilde{m}(x) = m(x) + K(x, X)\left(K + s^2 I_n\right)^{-1}(y - m) \tag{27}$$

and

$$\begin{aligned} \tilde{k}(x, x') &:= \tilde{k}(x, x';, s) \\ &= k(x, x') - K(x, X)\left(K + s^2 I_n\right)^{-1} K(X, x), \end{aligned} \tag{28}$$

respectively.

To carry out interpolation tasks when iterating on policies, one has to work with the predictive (marginal) distribution of the function value $f(x^*)$ for a single test input x^*. That is, given our posterior for the GP $f(\cdot)$, we can derive the marginal distribution of $f(\cdot)$ at any point. We obtain,

$$f(x^*)|X, y, s \sim \mathcal{N}\left(\tilde{m}(x^*), \tilde{\sigma}(x^*)\right), \tag{29}$$

where $\tilde{m}(x^*) = \tilde{m}(x^*)$ is the *predictive mean* given by (27), and $\tilde{\sigma}^2(x^*) := \tilde{k}(x^*, x^*; s)$ is the *predictive variance.*

Throughout our computations, we use the predictive mean as the value of the unknown function. Hence, we derive the same formula as in (17). The advantage of this procedure is that we can use maximum likelihood to estimate the hyper-parameters and s^2 from our training data. In principle, it would be useful also to make use of the variance-covariance term that indicates how accurate the forecast is at that point. Incorporating this into our economic model is subject to further research.

Standard GPs are not able to deal with very high input dimensions because they rely on the Euclidean distance to define input-space correlations. Since the Euclidean distance becomes uninformative as the dimensionality of the input space increases, the number of observations required to learn the function grows exponentially. To this end, following Scheidegger and Bilionis (2017), we couple GPs to active subspaces, which is consistent with our economic modeling.

5.3 The basic computational strategy

In our setup, the computation of self-justified equilibria is straightforward and reduces to Gaussian regression and the repeated solution of non-linear systems of equations. In particular, we employ an iterative scheme to solve for the optimal forecasting functions. In many respects, our method is very close to standard stimulation based projection-methods pioneered by den Haan and Marcet (1990). The basic details of the algorithm are then as follows:

1. Fix a stopping criterium, η, the size of the training sample, n, as well as the number of samples used for estimating C_N, N.
   Initial guess for each agent's forecasting:
   Initially, we assume that agents only use own asset holdings to forecast, i.e., $d = J$ and each $IJ \times d$ projection matrix $W_{y^a,z}$ project on agent y^a's asset holdings. Next, either use linear regression or construct the Gaussian processes whose posterior means approximate
   $$M^0_{y^a,z'} : \mathbf{Z} \times \mathbb{R}^d \to \mathbb{R}_+.$$
   Then, choose an approximation accuracy ξ and choose an initial condition $z_0, \vec{\theta}(z^{-1})$.

2. Iteration step:
   Simulate a temporary equilibrium path for given forecasts $\vec{M}^0$.
   For $i = 1, n$
   (a) Solve numerically for a temporary equilibrium, set $\vec{x}_i, \vec{\theta}_i, q_i$ to the equilibrium values and set $z_i = z$.
   (b) Using pseudo random numbers draw a new z' and set $\theta^-_{y^a} = \theta_{y^{a-1}}$ for all agents y^a.

3. For each y^a regress the equilibrium values of $f(q_i, z_i)u'(x_{y^{a+1},i})$ on $W_{y^a,z_{i-1}}\vec{\theta}_{i-1}$ and z_{i-1} to obtain a new Gaussian process whose posterior mean gives a new forecasting function $M^1_{y^a}$

4. If
   $$\|M^1 - M^0\| < \eta$$
   then set $M^* = M^1$. Else set $M^0 = M^1$ and repeat time iteration step 2.

5. Compute C_N as defined in Equation 14 and its eigenvalues, λ. If all agents' criteria (11) or (12) are satisfied, terminate. Else include one more eigenvector of C_N into the projective matrix W_{y^a}, make a new initial guess for Gaussian processes and go to time iteration step 2.

The computation of the temporary equilibrium is done using a simple Newton-method, the derivatives needed for the computation of C_N (cf. (14)) are approximated using one-sided finite differences. The package Limbo of Cully et al. (2018) for the Gaussian process regressions.

6 A simple example

To illustrate the concept of self-justified equilibria our general computational strategy, it is useful to focus on a specific simple example. Concretely, we assume that agents live for A periods and that there are two types of agents per generation and there are no idiosyncratic shocks. An agent is then characterized by (y, a), where $y = 1, 2$ denotes the initial shock. The agents distinguish themselves by trading constraints, endowment risk and possibly preferences. Type 1 agents can trade in a single Lucas-tree and Arrow securities. In our framework, it is useful to assume that the Arrow-securities pay in the Lucas-tree (as in Gottardi and Kubler (2015) or Chien and Lustig (2011)). Type 2 agents can only trade in the Lucas tree. Both agents face borrowing constraints. The model is a simplified OLG-version of Chien et al. (2011).

For concreteness, it is useful to define the temporary equilibrium system of inequalities as the system of all agents' KKT-conditions together with the market clearing conditions, i.e.,

$$
\begin{aligned}
-u'_{1,a}(e_{1,a}(z) + \theta^-_{(1,a-1),z}(\sum_{z' \in \mathbf{Z}} q_{z'} + d(z)) - q \cdot \theta_{1,a}) + \beta M_{1,a}(z, z', W_{1,a}\vec{\theta}) + \kappa_{1,a}, &\quad \text{for all } a, z' \\
\kappa_{1,a} \cdot \theta_{1,a} & \\
-u'_{2,a}(e_{2,a}(z) + \theta^-_{2,a-1}(\sum_{z' \in \mathbf{Z}} q_{z'} + d(z)) - \sum_{z' \in \mathbf{Z}} q_{z'}\theta_{2,a}) + \beta M_{2,a}(z, z', W_{1,a}\vec{\theta}) + \kappa_{2,a} &\quad \text{for all } a, z' \\
\kappa_{2,a}\theta_{2,a} & \\
\sum_a (\theta_{(1,a),z} + \theta_{2,a}) - 1, &\quad \text{for all } z \in \mathbf{Z}.
\end{aligned}
$$

We can combine $\kappa_{i,a}$ and $\theta_{i,a}$ into one variable and obtain a system with $(A-1)Z + (A-1) + Z$ equations and unknowns. This system has to be solved at every simulation step 2(a) in our algorithm (see. Sec. 5.3) and is the most time-consuming part of the computation.

6.1 A simple self-justified equilibrium with accurate forecasts

For the simplest example, assume that $A = 60$, $Z = 3$. In our first example all, agents have identical CRRA utility functions with $u_{y,a}(c) = \beta_y^a \frac{c^{1-\gamma_y}}{1-\gamma_y}$. We take $\beta_y = 0.97$ for $y = 1, 2$, and $\gamma_1 = \gamma_2 = 2.5$. Individual endowments are

$$e_{1,a}(1) = 0.95(0.4 + a/500),\ e_{1,a}(2) = 0.4 + a/500, e_{1,a}(3) = 0.95 * (0.4 + a/500) \text{ for } a < 50,$$

$$e_{1,a}(z) = 0.3 \text{ for } a \geq 50, z = 1, 2, 3,$$

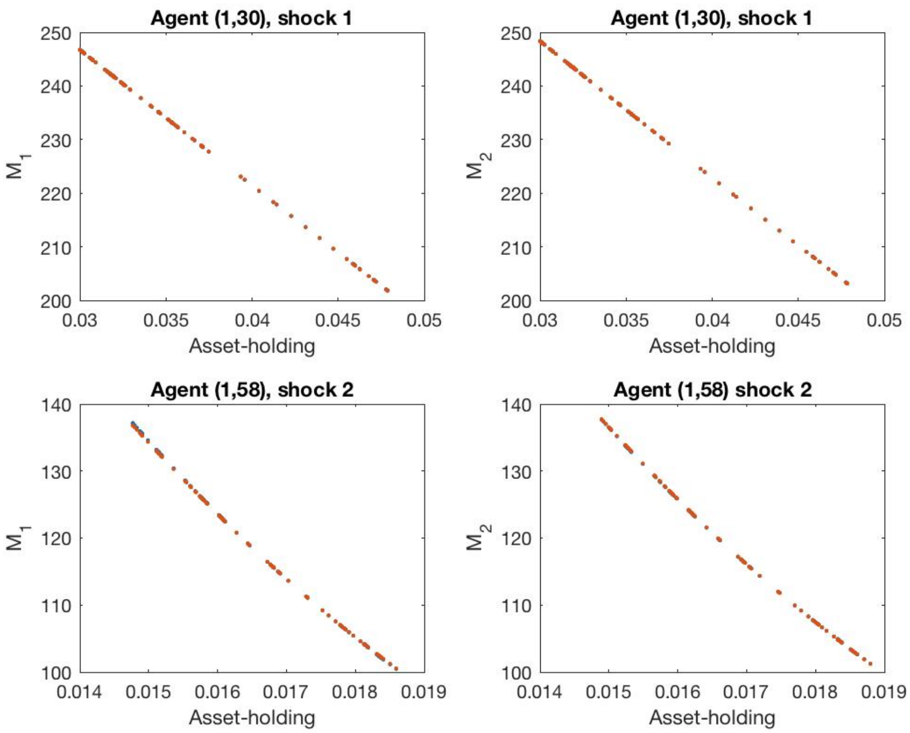
$$e_{2,a}(z) = e_{1,a}(z) \text{ for } z = 1, 2, \quad e_{2,a}(3) = 0.85 e_{1,a}(3).$$

Moreover, we also assume that $d(z) = 2$ for all $z = 1, 2, 3$, and that

$$\pi = \begin{pmatrix} 0.6 & 0.3 & 0.1 \\ 0.3 & 0.6 & 0.1 \\ 0.1 & 0.1 & 0.8 \end{pmatrix}.$$

As mentioned above we start off by assuming that agents only use their own asset holdings to forecast future marginal utilities. It is natural to assume that agent 1 (who can trade in three assets) assumes that his holdings in asset z (that pays if shock z realizes) only affects marginal utility in shock z for each $z = 1, 2, 3$. In the computed self-justified equilibrium, forecasting errors,

Figure 1: Simple forecasts.

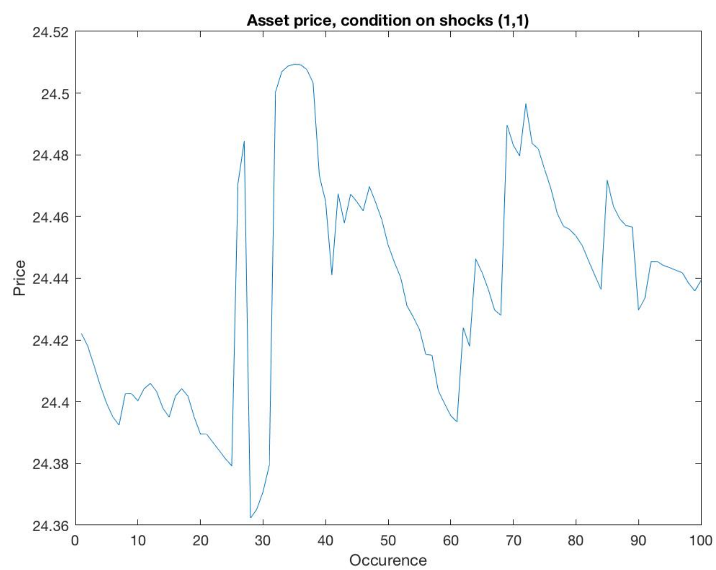


as measured by the average relative deviation between forecasted marginal utilities and realized marginal utilities, are tiny and uniformly (for all agents, shocks and assets) below 10^{-4}. That is to say that we computed a dimension-reduced self-justified equilibrium for $\epsilon^1_{y^a} = 10^{-4}$ for all y^a.

Moreover, many agents' forecasts are almost linear. In Figure 1, we show the forecast of an agent of age 29 and type 1 for the next period as a function of his asset holding, plotted against the realized marginal utility of that agent when he is of age 30. The plots almost perfectly overlap, and forecasting functions appear to be linear. In the same figure (second row) we repeat the same exercise for forecasts of agents of age 57 and realized marginal utilities of 58-year-olds. Again, the plots almost perfectly overlap, but this time a non-linearity in optimal forecasts is apparently visible.

The fact that agents forecasts are so accurate even though they do not take any aspect of the wealth distribution into account seems surprising at first. It turns out to hold for a wide variety of calibrations of the model (differences in risk-aversion, idiosyncratic shocks after the first period, etc.). This result is, of course, consistent with many examples in the literature, where one finds pseudo aggregation (most notably Krusell and Smith (1996)) and Chien and Lustig (2011), Chien et al. (2011), but also Storesletten et al. (2007)). The main reason why the simple forecasts are well in this example is that there is almost no variation in asset holdings and that asset prices are mostly a function of the current and past exogenous shock. In Figure 2, we show the asset prices conditional on the current shock being shock 1 and the previous shock being shock 1 and confirm that there is indeed minimal variation. Asset prices depend on the wealth distribution, but in equilibrium, this changes relatively little and effects on prices are very small. For other histories of shocks, the graphs look very similar.

Figure 2: Realized prices for shocks (1,1).



The linearity of forecasts is an artifact of this particular example. As it turns out, the result that linear forecasts are quite accurate holds true for a wide variety of parameter specifications but as this example already shows many forecasts are not exactly linear.

To go beyond this simple first example, we now construct an example where forecasts that do not take into account the wealth distribution across agents do not do a very good job.

6.2 Moving away from the simple example

One particular case where the simplicity of forecasts breaks down can be obtained by assuming that agents across generations have different, shock-dependent discounting. While this does not completely fit our model, it is a straightforward variation, and it gives us a modeling testbed to show the advantages of our algorithm.

In the concrete case, we modify above's example along 2 aspects. First, we assume that agents have heterogeneous risk-aversion with $\gamma_1 = 2.5$ and $\gamma_2 = 0.5$. Second, we assume that

$$u_{y,a}(c; z^{t+a}) = \Pi_{i=1}^{a} \beta_{y,a}(z_{t+i}) u_y(c),$$

that is, discounting across agents is heterogneous and might depend on the current shock. In particular, we choose for this example

$$\beta_{2,a}(z) = 0.98 \text{ for all } a, z,$$

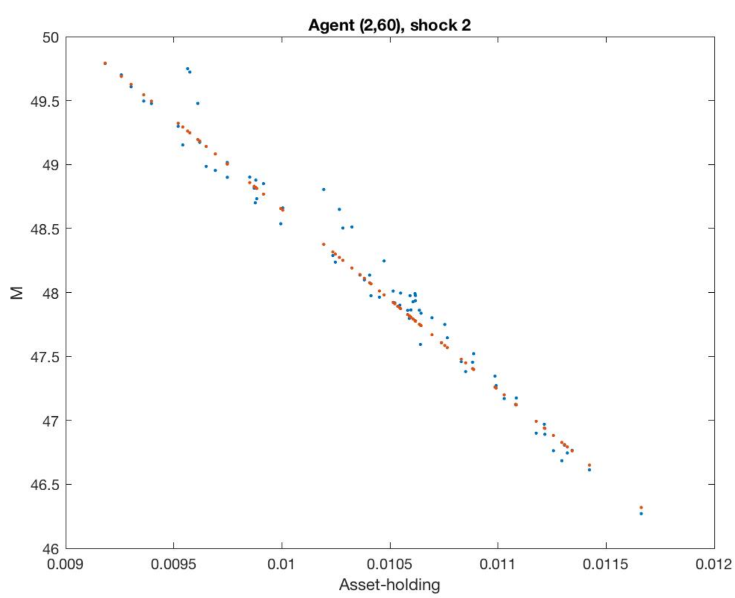
$$\beta_{1,a}(z) = 0.97 \text{ for } a = 1, \ldots, 40, 50 \ldots, 60, \text{ all } z,$$

and for $a = 51, \ldots, 60$,

$$\beta_{1,a} = \begin{cases} 1.06 & z = 1 \\ 0.88 & z = 2 \\ 0.97 & z = 3. \end{cases}$$

With this specification, forecasts are systematically misspecified—future marginal utilities for asset holding do not only depend on own choices. Figure 3 depicts forecasts of a 59-year-old agent of type 2 plotted against the (average) realized marginal utilities of the 60-year-old agent but for this specification with heterogeneous beliefs. Clearly, there are variables in addition to own asset holdings that have significant effects on marginal utilities.

Figure 3: Own asset-choice is not enough.

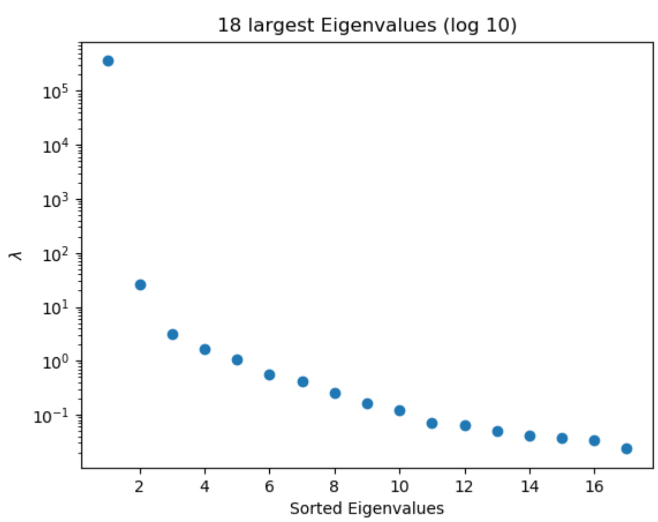


It turns out that for this specification the active subspace is two-dimensional. In addition to an agent's own asset holding, a single one-dimensional variable is needed to obtain accurate forecasts. The additional variable turns out to be a weighted sum of asset holdings across all agents, (roughly) weighted by the agents' marginal propensity to consume. Employing a higher-dimensional space to forecast future marginal utilities turns out to add very little. We compute the matrix C_N (cf. 14) by Monte-Carlo draws and finite differences and find that one single eigenvalue (in addition to the ones associated with own asset holdings) dominates all others. As explained above, we project all other agents asset holdings into a one-dimensional vector which we add to an agent owns asset holding. We solve for the new self-justified equilibrium iteratively, updating the projection a few times. At convergence, the average errors are uniformly below 3×10^{-3}. Adding the additional variable hence turns out to reduce forecasting errors substantially, comparable to the case in the Section above. Increasing the dimension of the active subspace further from 2 to 3 thus has almost no effect on the forecasting power.

In Figure 4, we plot the 18 largest eigenvalues on a log-scale (for the agent (2,60) whose realized marginal utilities are plotted in Figure 3 above). The figure confirms that all other eigenvalues are negligibly small compared to the one that corresponds the weighted sum of asset holdings across agents. In the computed self-justified equilibrium (using two variables to forecast future marginal utilities) both average forecasting errors are relatively small (in the order of 10^{-3}), and the reduction of variation by adding an additional explanatory variable is small. The jump from the largest to next largest eigenvalue is in the order of 10,000.

While in the discussion we focused on the forecasts of the agent (2,59), it turns out that all other agents' forecasts are actually more accurate, to begin with. It turns out that agents of type 1 can make very accurate forecasts using only their own asset position. The large variation in prices has little effects on the variation of marginal utilities. Again, we computed a self-justified equilibrium. Note that in practice, ordinary GP-regressions easily scale up about 10 dimensions. The simple

Figure 4: Largest eigenvalues for agent (2,59), shock 2



example in this section illustrates that this is likely to be enough to obtain very accurate forecasts even in much more complicated models, where active subspaces might be larger than one or two dimensions.

7 Conclusion

This paper makes three contributions. First, we define the concept of self-justified equilibrium as a natural generalization of rational expectations equilibrium, and we provide sufficient conditions for existence. Second, we argue that active subspace methods provide a natural way to formalize bounded rationality in very high dimensional models. Third, we provide an implementation to approximate self-justified equilibria numerically. In a relatively small model with 120 agents, we show that the method can potentially be used for large-scale applications.

We allow for the possibility of idiosyncratic shocks and a continuum of agents. However, in our current implementation, when solving for the temporary equilibrium, we compute the optimal demand for each agent in the economy. If there is a continuum of agents (that differ ex post by the realization of an idiosyncratic shock), one needs to aggregate groups of agents with similar wealth levels into one type of agent to make this step feasible. This adds another layer of approximation to our method, but is very simple in practice.

Future research includes production economies as well as economies with several consumption goods. While this is conceptually straightforward, it is not clear if the dimension of the active subspace is as low as in our examples.

It is also future research to allow for the possibility of non-convexities in agents' preferences or budget-sets. This is more complicated since one can no longer use the first order conditions to characterize optimal decisions.

Appendix A: Optimal ridge approximation and active subspaces

In our economic model, agents do not search for the optimal projection but are satisfied with finding an active subspace that reduces most of the "noise" from the forecasts. It turns out that the problem of finding an optimal projection is a difficult non-convex problem, but that the active subspace methods our agents use often provide reasonable approximations to an optimal projection.

Constantine et al. (2014) have the following theoretical result which makes concrete how well active subspace methods lead to a good approximation. Let $\tilde{\rho}(y, z) = \rho(V_1 y + V_2 z)$ and define the conditional expectation of the function value, given y as

$$G(y) = \int_z f(V_1 y + V_2 z)\tilde{\rho}(z|y)dz.$$

Theorem 3.1 in Constantine et al. (2014) now states

$$\int_x (f(x) - G(V_1^T x))^2 \rho(x) dx \leq C(\lambda_{d+1} + \ldots + \lambda_D),$$

where C is the Poincaré constant that depends on the pdf ρ.

Unfortunately, in this framework, Poincaré bounds are known to be far away from tight upper bounds (the exception being the standard normal distribution). Therefore, Theorem 3.1 in Constantine et al. (2014) does not tell us much about how far we are from an optimal projection.

The situation is slightly different if ρ is standard normal. In this case, the Poincaré constant is known to be 1, and it is easy to see that it can be obtained in a worse case scenario. As Zahm et al. (2018) point out, this can be extended to non-standard normal densities. Assuming that the normal density has covariance matrix Σ, they show that If one takes as projection matrix

$$P = (\sum_i v_i v_i^T) \Sigma^{-1},$$

where (λ_i, v_i) solves

$$C v_i = \lambda_i \Sigma^{-1} v_i,$$

one can obtain to following upper bound:

$$\int_x (f(x) - G(P^T x))^2 \rho(x) dx \leq (\lambda_{d+1} + \ldots + \lambda_D).$$

While our ergodic distributions are unlikely to be normal, the result is useful, since mixture of normal distributions typically can describe the distributions in our model. It is subject to further research to explore this in more detail. In any case, even this is not the optimal projection.

An optimal projection can easily be defined, but hardly ever computed in higher dimensions. Suppose that for a given function $f : \mathbb{R}^d \to \mathbb{R}$ and a given $n \ll d$, one wants to find a a $n \times d$ matrix $V_1 \in \mathbf{V}_n(\mathbb{R}^d)$ that allows for an "optimal "approximation of $f(\cdot)$ by a function $g : \mathbb{R}^n \to \mathbb{R}$, setting

$$f(x) \simeq g(V_1 x).$$

We want to define optimality as minimizing the L^2 norm with respect to a probability density over $\mathbb{R}^d$, $\rho(x)$. For given V_1, we can define $V_2 = I - V_1 V_1^T$ and write $x = V_1^T y + V_2^T z$ for $y = V_1^T x$, $z = V_2^T y$. We can define $\widetilde{\rho}(y, z) = \rho(V_1 x + V_2 y)$ and marginal and conditional densities by the standard equations. The conditional expectation is

$$\mathbb{E}\left(f(x)|y\right) = \int f(V_1 y + V_2 z) \widetilde{\rho}(z|y) dz.$$

The optimal V_1 solves the following optimization problem:

$$\min_{V_1 \in \mathbf{V}_n(\mathbb{R}^d)} \int_x (f(x) - \mathbb{E}(f(x)|V_1^T x))^2 \rho(x) dx. \tag{30}$$

Unfortunately, this is a very complicated, non-convex optimization problem, and even the search for a stationary point turns out to be very costly in high dimensions (see e.g. Cohen et al. (2012)). Constantine et al. (2017) propose to use active subspace methods to obtain an approximation for a stationary point. Since the problem is non-convex, there is, unfortunately, no guarantee that the stationary point is, in fact, a minimum. However, Constantine et al. (2017) also provide various examples to illustrate that one can sometimes expect to obtain a good approximation from active subspaces.

References

[1] Adam, K, A. Marcet and J.P. Nicolini, (2016), "Stock Market Volatility and Learning", *Journal of Finance*, 71, 33–82.

[2] Barry, D., (1986), "Nonparametric Bayesian Regression". *Annals of Statistics*, 14, 934–953.

[3] Bellman, R., (1961), *Adaptive Control Processes: A Guided Tour*, Princeton University Press.

[4] Bengio, Y., O. Delalleau and N.Le Roux, (2005), "The curse of highly variable functions for local kernel machines", *Neural Information Processing Systems.*

[5] Bewley, T. (1984), "Fiscal and Monetary Policy in a General Equilibrium Model", Cowles Foundation Discussion Paper 690.

[6] Blei, D. and M Jordan,(2005), "Variational inference for Dirichlet process mixtures" *Bayesian Analysis* 1, 121–144.

[7] Border, K. C., (1989), *Fixed Point Theorems with Applications to Economics and Game Theory.* Cambridge University Press.

[8] Brumm, J. , D. Kryczka and F. Kubler, (2017), "Recursive Equilibria in Dynamic Economies with Stochastic Production", (2017), *Econometrica*, 85, 1467–1499.

[9] Brumm, J. , F. Kubler and S. Scheidegger, (2017b), "Computing equilibria in dynamic stochastic macro-models with heterogeneous agents", *Advances in Economics and Econometrics, Eleventh World Congress*, 185–230.

[10] Chien, Y. and H. Lustig (2011), "The Market Price of Aggregate Risk and the Wealth Distribution," *The Review of Financial Studies* 23, 1596–1650.

[11] Chien, Y, H. Cole and H. Lustig, (2011), "A Multiplier Approach to Understanding the Macro Implications of Household Finance", *Review of Economic Studies* 78, 199–234.

[12] Cho, I.-K. and T. Sargent (2009), “Self-confirming equilibria”in *The New Palgrave Dictionary of Economics, 2nd edition*

[13] Citanna, A. and P. Siconolfi (2012), “Recursive equilibria in stochastic OLG economies: Incomplete markets,” *Journal of Mathematical Economics*, 48, 322–337.

[14] Cohen, A., I. Daubechies, R. DeVore, G. Kerkyacharian and D. Picard, (2012), “Capturing ridge functions in high dimensions from point queries”, *Constructive Approximation*, 35, 225 –243.

[15] Constantine, P., (2015), *Active Subspaces: Emerging Ideas for Dimension Reduction in Parameter Studies* SIAM, Philadelphia.

[16] Constantine, P. E. Dow and Q. Wang, (2014), “Active Subspace Methods in Theory and Practice: Applications to Kriging Surfaces”*SIAM Journal on Scientific Computing*, 36, A1500-A1524.

[17] Constantine, P., A. Eftekhari, J. Hokanson and R. Ward, (2017), “A near-stationary subspace for ridge approximation”, *Computer Methods in Applied Mechanics and Engineering*, 326, 402–421.

[18] Cully, A., K. Chatzilygeroudis, F. Allocati and J-B Mouret, (2018), “Limbo: A Flexible High-performance Library for Gaussian Processes modeling and Data-Efficient Optimization”, preprint.

[19] Den Haan, W. J. and A. Marcet, (1990), “Solving the Stochastic Growth Model by Parameterizing Expectations,” *Journal of Business and Economic Statistics*, 8, 31–34.

[20] Evans, R. and K.L. Phillips, (2014), “OLG Life Cycle Model Transition Paths: Alternate Model Forecast Method”, *Computational Economics*, 43, 1095–1110.

[21] Evgeniou, T., M. Pontil and T. Poggio, (2010), “Regularization Networks and Support Vector Machines”, *Advances in Computational Mathematics*, 13 : 1.

[22] Farmer, D.F and D. Foley, (2009), “The economy needs agent-based modeling”, Nature Vol 460, 485–486

[23] Fudenberg, D. and D. Levine (1993), “Self-confirming equilibira”, *Econometrica*, 61, 523–545.

[24] Gabaix, X, (2014), “A sparsity based model of bounded rationality”, *The Quarterly Journal of Economics*, 129, 1661–1710.

[25] Gottardi, P. and F. Kubler (2015) “Dynamic Competitive Economies with Complete Markets and Collateral Constraints”, *Review of Economic Studies*, 82, 1119–1153.

[26] Horn, R.A. and C.R. Johnson, (1985), *Matrix Analysis*, Cambridge University Press.

[27] Krueger, D. and F. Kubler, "Computing equilibrium in OLG models with stochastic production", (2004), *Journal of Economic Dynamics and Control*, 28, 1411-1436.

[28] Krusell, P. and A. Smith (1998), "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 106, 867–896.

[29] Kubler, F. and K. Schmedders (2003), "Stationary Equilibria in Asset-Pricing Models with Incomplete Markets and Collateral," *Econometrica*, 71, 1767–1795.

[30] Kubler, F. and H.M. Polemarchakis (2004), "Stationary Markov Equilibria for Overlapping Generations," *Economic Theory*, 24, 623–643.

[31] Kurz, (1994), "On Rational Belief Equilibria, " *Economic Theory*, 4, 859–876.

[32] Ljungqvist, L. and T.J. Sargent, (2012), *Recursive Macroeconomic Theory.* MIT Press.

[33] Maliar, L. and S. Maliar (2014), "Numerical Methods for Large Scale Dynamic Economic Models", *Handbook of Computational Economics,* Volume 3, 325 – 477.

[34] Rasmussen, C. and C. Williams, Christopher, (2005), *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*, MIT Press.

[35] Reiter, M., (2010), "Approximate and Almost-Exact Aggregation in Dynamic Stochastic Heterogeneous-Agent Models", IHS Economics Series 258.

[36] Sargent, T.J., (1993), *Bounded Rationality in Macroeconomics*, Oxford University Press.

[37] Scheidegger, S. and I. Bilionis (2017). "Machine learning for high-dimensional dynamic stochastic economies", working paper, University of Zurich.

[38] Scheidegger, J. D. Mikushin, F. Kubler and O. Schenk, "Rethinking large-scale economic modeling for efficiency: optimizations for GPU and Xeon Phi clusters", (2018), *IPDPS 2018.*

[39] Sobol, I.M. and S. Kucherenko, (2009), "Derivative based local sensitivity measures and their link with global sensitivity indices". *Mathematics and Computers in Simulation*, 79, 3009–3017.

[40] Storesletten, K., C. I. Telmer, and A. Yaron (2007) "Asset pricing with idiosyncratic risk and overlapping generations." *Review of Economic Dynamics*, 10, 519 548.

[41] Tripathy, R., and I. Bilionis, and M. Gonzalez (2016) "Gaussian processes with built-in dimensionality reduction: Applications to high-dimensional uncertainty propagation ". *Journal of Computational Physics*, 321, 191–223.

[42] Woodford, M., (2013), "Macroeconomic Analysis without the Rational Expectations Hypothesis", *Annual Review of Economics*, 5, 303–346.

[43] Zahm, O., P. Constantine, C. Prieur and Y. Marzouk (2018), "Gradient-based dimension reduction of multivaraite vector-valued functions", preprint.