Sequence-Space Jacobians of Life Cycle Models

Bence Bardóczy *[1] and Mateo Velásquez-Giraldo †[1]

[1]Board of Governors of the Federal Reserve

May 29, 2025

Abstract

We provide an efficient algorithm to calculate the sequence-space Jacobians of overlapping generations models. The algorithm exploits agents' finite planning horizons and predictable transitions between ages. It only needs a function mapping life cycle parameters to policy functions and transition matrices—a usual implementation for partial equilibrium analyses. This makes the algorithm easy to apply to existing models and facilitates their analysis in general equilibrium. The algorithm also produces cohort-specific Jacobians that track the dynamic responses of different cohorts to shocks.

JEL Codes: C61, C63, D15.

*bence.a.bardoczy@frb.gov

†mateo.velasquezgiraldo@frb.gov

The views expressed in this paper are those of the authors and do not necessarily represent the views or policies of the Board of Governors of the Federal Reserve System or its staff.

1 Introduction

Advances in perturbation methods have accelerated the adoption of heterogeneous-agent models.[1] Using these advances, a growing literature has demonstrated that heterogeneity shapes macroeconomic processes like the transmission of monetary and fiscal policy.[2] This literature has emphasized several characteristics of households that matter for macro dynamics—for example, their wealth, portfolios, and marginal propensities to consume. Most of these characteristics change significantly as households age.[3] Yet, most existing work relies on infinite-horizon models, where the age of an agent does not change the problem he faces.

A contributing factor to this modeling decision is computational complexity. While most perturbation methods could readily accommodate meaningful life cycle dynamics by treating age as just an additional state variable in the agent's problem, this approach would significantly increase their computational burden. Naively, augmenting a model with A distinct ages without making any changes to its representation or solution algorithm would multiply the size of its state space by a factor of A and its number of potential state-to-state transitions by a factor of A^2. Therefore, one could roughly expect the life cycle version of a model to be at least A times as costly to use as its infinite horizon counterpart.

However, treating age as a generic state variable is inefficient. Age has various properties that model representations and solution algorithms can exploit. First, its possible transitions are limited. A 20-year-old cannot turn 38; he either turns 21 or dies. This implies that the majority of transitions in the naive A^2 augmentation can be ignored. It also means that all the future information relevant to a 20-year-old at time t is contained in the value function of a 21-year-old at time $t+1$, $v_{t+1}(21)$; he is unconcerned with $v_{t+1}(38)$. Second, in models without explicit dynastic considerations, agents care only about events that will occur within their lifetime. Therefore, shocks that happen too far in the future have no effect on their actions. These ideas underpin the backward solution and forward simulation methods used in the life cycle literature.

This paper formalizes these special properties of age as a state variable and provides a practical guide to exploit them in the sequence-space Jacobian (SSJ) method of Auclert et al. (2021a). The result is a method that can significantly outperform the naive estimate of an A-fold increase in computational costs. We focus on computing Jacobians for life cycle household models in isolation; these Jacobians can later be combined with other blocks to form larger macroeconomic models following the original SSJ framework. In Table 1, we compare the costs of calculating Jacobians for a simple infinite horizon household model with those of applying our method to a 75-year life cycle analogue.[4] While the number of

[1] For perturbation methods, see, for example, Reiter (2009), Boppart, Krusell, and Mitman (2018), Auclert et al. (2021a), and Bhandari et al. (2023).

[2] See, for example, Kaplan, Moll, and Violante (2018), Auclert, Rognlie, and Straub (2018), Auclert (2019), and Luetticke (2021).

[3] Studies like Doepke and Schneider (2006), Auclert et al. (2021b), Peterman and Sager (2022), Bardóczy and Velásquez-Giraldo (2024), Beaudry, Cavallino, and Willems (2024), and Gruss et al. (2025) have demonstrated the significance of these life cycle dynamics for several macroeconomic debates.

[4] The life cycle model is the one described in Section 6 and the infinite horizon model is based on the parametrization of the first period of the life cycle model. Both models use the exact same within-period

Table 1: Costs of Augmenting a Household Model with Life Cycle Dynamics

Step	Inf. Horizon	Life Cycle (75 Ages)	Ratio (LC/IH)
No. of idiosyncratic states	357	$26,775$	75
Computing time (Average across 100 runs), seconds			
Steady state	0.0022 s	0.0007 s	0.32
Single in.-out. Jacobian ($H = 300$)	0.1214 s	0.8884 s	7.32
Peak memory use, megabytes			
Single in.-out. Jacobian ($H = 300$)	50.06 MB	124.04 MB	2.48

Notes: The Life Cycle model that the table refers to is the one described in Section 6. The infinite horizon model simply takes the parameters (income process, preferences) of the first period of the life cycle model; we set its survival probability to 0.96 per year to make the model stationary. We use exactly the same single-period solver for both models to make computing times comparable. The steady state calculation of the life cycle model is faster because we use a non-iterative procedure to find it (see Lemma 1). All the tests were performed on a laptop with an Intel(R) Core(TM) Ultra 7 155H processor at 3.80 GHz. For Jacobians, the input is the interest rate and the output is consumption.

possible idiosyncratic state configurations that must be solved for grows by a factor of 75, the time it takes to find a Jacobian of the model grows only by a factor of around 7. Peak memory use, which can become a consideration for more complex models, grows only by a factor of around 2.5.

We construct our algorithm using objects and methods that most implementations of life cycle models already have, and we hope this will facilitate its adoption. The algorithm requires only a function that, given a sequence of the aggregate inputs (for example, taxes, interest rates, or wages) that an agent will face throughout his life $\{\mathbf{X}_a\}_{a=0}^{A-1}$, returns the sequence of relevant policy functions evaluated on age-specific grids $\{\mathbf{y}_a\}_{a=0}^{A-1}$ and the transition matrices that describe the stochastic movement of agents through those grids conditional on their survival, $\{\mathcal{L}_a\}_{a=0}^{A-1}$.[5]

We see two main groups of applications for the method we develop. First, the method will facilitate the inclusion of life cycle dynamics in heterogeneous-agent macroeconomic models. We believe this is valuable given the large body of studies that highlight age as a driver of heterogeneity in key quantities like marginal propensities to consume and interest rate exposures.[6] Furthermore, as demonstrated in Auclert et al. (2021a,b), Boehl (2023), and Gruss et al. (2025), sequence-space Jacobians can accelerate the calculation of nonlinear

solution and simulation code.

[5]These transition matrices will be readily available in studies that take a non-stochastic approach to forward simulation. Young (2010) and Ocampo and Robinson (2023) compare this method with Monte Carlo simulation.

[6]See, for example, Auclert (2019), Fagereng, Holm, and Natvik (2021), and Andersen et al. (2023).

transition paths in heterogeneous-agent economies. Thus, we expect our method to facilitate explorations of regime changes in contexts where generational differences and life cycle considerations matter. These contexts include studies of fiscal reforms and demographic transitions.

Second, the method will benefit studies using life cycle models to answer counterfactual questions involving partial equilibrium effects. For concreteness, imagine a microeconomic study using an estimated model to evaluate the effect of counterfactual income taxes on housing decisions over the life cycle. To account for the effects of endogenous movements in house prices, the household block could be combined with an auxiliary housing supply sector and a market-clearing condition. In this type of model, the Jacobians that we deliver can be used to give a first-order approximation of the effect of transitory changes, or to accelerate the numerical search of the transition path to new permanent regimes. Additionally, our method produces "age-specific" Jacobians that can instantly deliver the response of different cohorts of agents to changes in their environments, which can be of interest in this class of applications.

The rest of this paper is organized as follows. Section 2 defines the class of life cycle problems that we consider and represents it in the SSJ framework. Section 3 defines age-specific sequence-space Jacobians and discusses their relationship with aggregate Jacobians. Section 4 defines and describes age-specific "fake news" matrices and shows they are sufficient to construct age-specific Jacobians. Section 5 provides our algorithm for computing age-specific fake news matrices, relating it to a typical life cycle model implementation. Section 6 demonstrates our method, calculating age-specific and aggregate Jacobians for a life cycle consumption-saving problem. Section 7 concludes.

2 Life Cycle Problems in Sequence Space

Background. The SSJ framework represents macroeconomic models as collections of blocks. A block can represent, for example, all the households in the economy, firms, or a fiscal authority. Blocks interact through aggregate variables that they either take as given (an "input") or that they produce (an "output"). For example, the fiscal authority can set the tax rate, give it to the household block as an input, and the household block outputs aggregate consumption and savings. Other blocks include equilibrium restrictions that aggregate variables must satisfy; for example, market clearing and no-arbitrage conditions.

Since we are concerned with life cycle problems, we restrict our attention to heterogeneous-agent blocks: those that represent collections of agents that can be in different idiosyncratic states. The state space of agents in these blocks is discretized and represented with a grid. Policy functions, value functions and distributions are vectors. The ith entry of a vector is the value of the relevant function at the ith possible configuration of the idiosyncratic state, or grid point. The representation of a block with inputs $\mathbf{X}$ and outputs $\mathbf{Y}$ is

$$\begin{aligned}\mathbf{v}_t &= v(\mathbf{v}_{t+1}, \mathbf{X}_t), \quad \mathbf{\Lambda}_t = \Lambda(\mathbf{v}_{t+1}, \mathbf{X}_t), \quad \mathbf{y}_t = y(\mathbf{v}_{t+1}, \mathbf{X}_t) \\ \mathbf{D}_{t+1} &= \mathbf{\Lambda}_t' \mathbf{D}_t, \qquad\qquad \mathbf{Y}_t = \mathbf{y}_t' \mathbf{D}_t,\end{aligned} \tag{1}$$

where $\mathbf{v}_t$ is the value function, $\mathbf{D}_t$ is the distribution of agents over states, and $\mathbf{y}_t$ is the individual-level outcome of interest. Aggregate outputs are averages of individual outcomes. The matrix $\mathbf{\Lambda}_t$ is the transition matrix describing the stochastic movement of agents through states. Agents are forward-looking. The functions $v(\cdot,\cdot)$, $\Lambda(\cdot,\cdot)$, and $y(\cdot,\cdot)$ capture the dependence on the continuation value $\mathbf{v}_{t+1}$ as well as on contemporaneous inputs $\mathbf{X}_t$, embedding Bellman's principle of optimality.

Zooming out, the equations in Equation 1 define a function that, given an initial distribution $\mathbf{D}_0$, maps sequences of $\mathbf{X}$ into sequences of $\mathbf{Y}$,

$$f(\{\mathbf{X}_t\}_{t=0}^{\infty}) = \{\mathbf{Y}_t\}_{t=0}^{\infty}. \tag{2}$$

Assumptions and representation of life cycle problems. We now characterize the class of problems that we confront in this paper. A life cycle problem is a dynamic optimization problem that includes a specific state variable we call "age" and denote with a. We enumerate the assumptions that we make about the problem and the age variable and derive some implications after each assumption.

1. Age has a finite domain. We denote its number of distinct possible values with A and index them starting with 0, $a \in \{0, 1, 2, ..., A-2, A-1\}$. We partition the vectors in equation 1 into age-specific sub-vectors as

$$\begin{aligned}
\mathbf{v}_t' &= \left[\mathbf{v}_t'(0)\ , ..., \mathbf{v}_t'(A-1)\right], \\
\mathbf{D}_t' &= \left[\mathbf{D}_t'(0)\ , ..., \mathbf{D}_t'(A-1)\right], \\
\mathbf{y}_t' &= \left[\mathbf{y}_t'(0)\ , ..., \mathbf{y}_t'(A-1)\right].
\end{aligned} \tag{3}$$

This partition allows for every aggregate output to be decomposed additively into contributions of each living cohort,

$$\mathbf{Y}_t = \sum_{a=0}^{A-1} \mathbf{y}_t(a)'\mathbf{D}_t(a).$$

2. From age a, an agent either survives and advances to age $a+1$ or dies and is replaced by a newborn agent. The probability of these events can depend on age but must be independent of the other state variables and of aggregate inputs. At age a, δ_a denotes the probability of death and $\not\delta_a \equiv 1 - \delta_a$ the probability of survival. Agents die with certainty at age $A-1$, $\delta_{A-1} = 1$.

3. Agents who die are replaced by newborns ($a = 0$). Newborns draw their initial states (other than age) from a distribution that is fixed over time and independent of all other variables in the model. We denote this distribution with η.

The previous assumptions imply that we can partition the transition matrix $\mathbf{\Lambda}_t$ in Equation 1 as

$$\mathbf{\Lambda}_t = \begin{bmatrix} \delta_0 \times \mathbf{1}\eta' & \not\delta_0 \mathcal{L}_t(0) & \mathbf{0} & \dots & \mathbf{0} \\ \delta_1 \times \mathbf{1}\eta' & \mathbf{0} & \not\delta_1 \mathcal{L}_t(1) & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \delta_{A-2} \times \mathbf{1}\eta' & \mathbf{0} & \mathbf{0} & \dots & \not\delta_{A-2}\mathcal{L}_t(A-2) \\ \mathbf{1}\eta' & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \end{bmatrix}. \tag{4}$$

This equation implicitly defines $\mathcal{L}_t(a)$ as the transition matrix that describes the movement of agents from idiosyncratic states in age a to idiosyncratic states in age $a+1$ conditional on their survival. We will henceforth refer to $\{\mathcal{L}_t(a)\}_{a=0}^{A-1}$ as conditional transition matrices. No conditional transition matrix exists for the terminal age, but defining a placeholder $\mathcal{L}_t(A-1) = \emptyset$ lightens our notation. The first block column of Equation 4 accounts for dying agents, who draw their states from η.[7]

A more succinct way to express the law of motion for age-specific distributions implied by equation 4 is

$$\mathbf{D}_{t+1}(a) = \begin{cases} \left(\sum_{j=0}^{A-1} \delta^j \times \mathbf{1}'\mathbf{D}_t(j)\right)\eta, & a = 0 \\ \not\delta^{a-1}\mathcal{L}_t(a-1)'\mathbf{D}_t(a-1), & 0 < a \leq A-1. \end{cases} \tag{5}$$

4. The dynamic optimization problem that agents solve is such that there are age-specific functions $\{v[a](\cdot), y[a](\cdot), \mathcal{L}[a](\cdot)\}_{a=0}^{A-1}$ that satisfy

$$\begin{aligned} \mathbf{v}_t(a) &= v[a]\left(\mathbf{v}_{t+1}\left(a+1\right), \mathbf{X}_t\right), \\ \mathbf{y}_t(a) &= y[a]\left(\mathbf{v}_{t+1}\left(a+1\right), \mathbf{X}_t\right), \\ \mathcal{L}_t(a) &= \mathcal{L}[a]\left(\mathbf{v}_{t+1}\left(a+1\right), \mathbf{X}_t\right) \end{aligned} \tag{6}$$

for $0 \leq a < A-1$, and $\mathbf{v}_t(A-1) = v[A-1]\left(\mathbf{X}_t\right)$ and $\mathbf{y}_t(A-1) = y[A-1]\left(\mathbf{X}_t\right)$.

This assumption characterizes the type of dynamic problem that our method allows. Its main requirement is that the present actions, transitions, and utility of agents of age a at time t are affected by information about the future only through the value function of agents of age $a+1$ at time $t+1$. Most dynamically-consistent life cycle models that can be written in Bellman form satisfy this restriction. Examples of models that do not satisfy this restriction are those with time-inconsistent preferences (for example, Laibson 1997) or those where agents care about the path of aggregate inputs $\mathbf{X}$ after their death.[8]

As in SSJ, we focus on shocks around a steady state: a setting where aggregate inputs $\mathbf{X}$ have been constant for an indefinitely long time and are expected to remain so, and

[7] An extension of our method that allows for endogenous death probabilities and newborn distributions that depend on the microeconomic state of the agents they replace is possible. The first block column of Equation 4 becomes a series of time and age-varying matrices that describe the state of newborns as functions of the state of the agents they replace. The effects of this generalization are discussed in Footnote 11.

[8] The last case could arise, for example, in dynastic models where a parent thinks about the macroeconomic environment that will affect their children.

where the distribution of agents over states has become time-invariant. The following lemma characterizes the steady state distribution.

Lemma 1. *(Steady State Distribution) The only possible steady state distribution of agents across states in the life cycle model is recursively defined as*

$$\mathbf{D}_{ss}(a) = \begin{cases} \left(\sum_{j=0}^{A-1} \not\Delta^{j}\right)^{-1} \times \eta, & a = 0 \\ \not\delta^{a-1} \mathcal{L}_{ss}(a-1)' \mathbf{D}_{ss}(a-1), & 0 < a \leq A-1, \end{cases}$$

where $\not\Delta^{0} = 1$ *and* $\not\Delta^{a} \equiv \prod_{k=0}^{a-1} \not\delta^{k}$ *denote the probability that an agent will live at least to age* a*. Proof in Appendix A.1.*

A convenient implication of Lemma 1 is that the exact steady state distribution can be computed with a finite number of operations: finding the newborn distribution and iterating it to age $A-1$.

Shocks to life cycle blocks. We now examine the effect of shocks to an input $\mathbf{X}$ on age-specific value functions, distributions, policy functions, and transition matrices. The system starts from a steady state with constant inputs $\mathbf{X}_{ss}$ and, unexpectedly at time $t=0$, there is a change in the future sequence of expected and realized inputs $\{\mathbf{X}_t\}_{t=0}^{\infty}$.

To represent these shocks and their effects, we extend the SSJ notation to the age partitions. Steady state values of variables and vectors are marked with underscript ss. For a given input $\mathbf{X}$ and shock size dx, we use $\mathbf{X}^s$ to denote a sequence of inputs that has the value $\mathbf{X}_{ss} + dx$ in its s-th entry and $\mathbf{X}_{ss}$ in every other. Indexed vectors $\{\mathbf{v}_t^s, \mathbf{D}_t^s, \mathbf{y}_t^s, \mathbf{\Lambda}_t^s\}_{t=0}^{\infty}$ represent the solution to the system in Equation 1 when the sequence of inputs is $\mathbf{X}^s$. We extend this notation to the partition of these objects into its age-specific constituents $\{\mathbf{v}_t^s(a), \mathbf{D}_t^s(a), \mathbf{y}_t^s(a), \mathcal{L}_t^s(a)\}_{t=0}^{\infty}$ for $0 \leq a \leq A-1$.[9] Finally, for any vector or variable Z, we use dZ to denote its deviation from its steady state value, $dZ \equiv Z - Z_{ss}$ and $dZ(a) \equiv Z(a) - Z_{ss}(a)$ for any $0 \leq a \leq A-1$.

The effects of single-period shocks on policy functions and transition matrices of life cycle problems satisfy the three following properties.

Lemma 2 (Invariance to past shocks)**.** *For any* $0 \leq a \leq A-1$*,* $t \geq 0$ *and* $s \geq 0$ *such that* $s < t$*, it is the case that*

$$\mathbf{y}_t^s(a) = \mathbf{y}_{ss} \quad \text{and} \quad \mathcal{L}_t^s(a) = \mathcal{L}_{ss}.$$

Lemma 3 (Time-shift symmetry)**.** *For any* $0 \leq a \leq A-1$*,* $t \geq 0$ *and* $s \geq 0$ *it is the case that*

$$\mathbf{y}_t^s(a) = \mathbf{y}_{t+k}^{s+k}(a) \quad \text{and} \quad \mathcal{L}_t^s(a) = \mathcal{L}_{t+k}^{s+k}(a) \quad \forall k \geq 0.$$

Particularly, for $t = 0$*,*

$$\mathbf{y}_0^s(a) = \mathbf{y}_k^{s+k}(a) \quad \text{and} \quad \mathcal{L}_0^s(a) = \mathcal{L}_k^{s+k}(a) \quad \forall k \geq 0.$$

[9] For example, $\mathbf{v}_t^{s\,\prime} = [\mathbf{v}_t^{s\,\prime}(0), ..., \mathbf{v}_t^{s\,\prime}(A-1)]$.

Lemma 4 (Truncated horizon)**.** *For any* $0 \leq a \leq A-1$ *and* (s,t) *such that* $s-t > (A-1)-a$*,*

$$y_t^s(a) = y^{ss}(a) \quad \text{and} \quad \mathcal{L}_t^s(a) = \mathcal{L}^{ss}(a).$$

Particularly, for $t = 0$*, if* $s > (A-1)-a$*,*

$$y_0^s(a) = y_{ss}(a) \quad \text{and} \quad \mathcal{L}_0^s(a) = \mathcal{L}_{ss}(a).$$

The proof of Lemmas 2, 3 and 4 is in Appendix A.2 and relies on writing out the recursion of value functions, policy vectors, and transition matrices as

$$\begin{aligned}
\mathbf{v}_t^s(a) &= v[a](\mathbf{X}_t^s) \circ v[a+1](\mathbf{X}_{t+1}^s) \circ ... \circ v[A-1](\mathbf{X}_{t+A-1-a}^s) \\
\mathbf{y}_t^s(a) &= y[a](\mathbf{X}_t^s) \circ v[a+1](\mathbf{X}_{t+1}^s) \circ ... \circ v[A-1](\mathbf{X}_{t+A-1-a}^s) \\
\mathcal{L}_t^s(a) &= \mathcal{L}[a](\mathbf{X}_t^s) \circ v[a+1](\mathbf{X}_{t+1}^s) \circ ... \circ v[A-1](\mathbf{X}_{t+A-1-a}^s).
\end{aligned}$$

The results follow from replacing $\mathbf{X}^s$ and its shifts into these expressions.

Lemma 2 says that policy functions and transition matrices are not affected by past shocks conditional on agent's current state. Lemma 3 says that, conditional on an agent's age, what matters for its policy function and transition matrix is the time remaining until a shock arrives $(s-t)$, not the specific calendar dates of s and t. Finally, lemma 4 says that agents do not react to the announcement of shocks that will occur after their death. The infinite horizon framework of Auclert et al. (2021a) shares the properties in Lemmas 2 and 3. The symmetries and invariances implied by these three properties underpin the efficient calculation of sequence-space Jacobians.

3 Age-Specific Sequence-Space Jacobians

Consider the functional representation of a model block, $f(\{\mathbf{X}_t\}_{t=0}^{\infty}) = \{\mathbf{Y}_t\}_{t=0}^{\infty}$. The Jacobian $\mathcal{J}$ is a matrix describing the effect of a perturbation to $\mathbf{X}_s$ on $\mathbf{Y}_t$,

$$\mathcal{J}_{t,s} = \frac{df_t(\{\mathbf{X}_j\}_{j=0}^{\infty})}{d\mathbf{X}_s} = \frac{d\mathbf{Y}_t}{d\mathbf{X}_s}. \tag{7}$$

These Jacobians are extremely useful objects. Not surprisingly, they are sufficient to solve for impulse responses to first order. Auclert et al. (2021a) also demonstrate that the Jacobians can be used to test local determinacy, evaluate the likelihood, and give good guesses for an iterative solution of nonlinear transition paths.

The SSJ method is operational thanks to the "fake news algorithm" that exploits the structure of Equation 1 to efficiently compute the Jacobians of heterogeneous agent blocks. Our contribution is a version of this algorithm that exploits the special features of age as a state variable to economize on the processing time and memory requirements in the case of OLG models.

Our algorithm decomposes the aggregate Jacobians of Equation 7 into the contributions of different living cohorts to changes in aggregate outputs. Formally, this decomposition is

$$\begin{aligned}\mathcal{J}_{t,s}dx = d\mathbf{Y}_t^s &= d\left(\sum_{a=0}^{A-1}\mathbf{y}_t^s(a)'\mathbf{D}_t^s(a)\right)\\ &= \sum_{a=0}^{A-1}\{d\mathbf{y}_t^s(a)'\mathbf{D}_{ss}(a)+\mathbf{y}_{ss}(a)'d\mathbf{D}_t^s(a)\} \equiv \sum_{a=0}^{A-1}\mathcal{J}_{t,s}(a)dx.\end{aligned} \tag{8}$$

Age-specific Jacobians are the elements of this decomposition. $\mathcal{J}_{t,s}$ measures the change in output $\mathbf{Y}$ at time t caused by a shock to input $\mathbf{X}$ that is announced at time 0 and occurs at time s. $\mathcal{J}_{t,s}(a)$ is the part of that change that is due to the cohort of agents that has age a at time t.

Definition 1 (Age-Specific Sequence-Space Jacobians)**.** *For an output* $\mathbf{Y}$ *and input* $\mathbf{X}$ *of interest, and for every age* $a \in \{0,1,...,A-1\}$*, the age-specific sequence-space Jacobians* $\{\mathcal{J}(a)\}_{a=0}^{A-1}$ *are matrices with entries such that*

$$\mathcal{J}_{t,s}(a)dx = d\mathbf{y}_t^s(a)'\mathbf{D}_{ss}(a)+\mathbf{y}_{ss}(a)'d\mathbf{D}_t^s(a).$$

for $t \geq 0$ *and* $s \geq 0$*.*

Age-specific sequence-space Jacobians have two immediate uses. First, since $\mathcal{J} = \sum_{a=0}^{A-1}\mathcal{J}(a)$, they can be used to construct the aggregate Jacobians of the heterogeneous-agent life cycle block. Second, age-specific Jacobians can be used to trace the response of particular cohorts to a shock over time. For example, the response of the cohort that had age a at time 0 to a shock that occurs at time s and is announced at time 0 is $\{\mathcal{J}_{t,s}(a+t)\}_{t=0}^{A-1-a}$. The t-th element in the set is the cohort's response at time t.

The rest of the paper describes an efficient algorithm to compute $\{\mathcal{J}(a)\}_{a=0}^{A-1}$.

4 Age-Specific Fake News Matrices

Our method to find age-specific sequence-space Jacobians is an adaptation of the "fake news algorithm" of Auclert et al. (2021a). As such, it requires age-specific analogues of many of the objects defined in that paper. Among these objects, the most important are fake news matrices, which measure the difference between the time t response to a time s shock, and the time $t-1$ response to a time $s-1$ shock.

Definition 2 (Age-Specific Fake News Matrices)**.** *For an output* $\mathbf{Y}$ *and input* $\mathbf{X}$ *of interest, and for every age* $a \in \{0,1,...,A-1\}$*, the age-specific fake news matrices* $\{\mathcal{F}(a)\}_{a=0}^{A-1}$ *have their entries defined as*

$$\mathcal{F}_{t,s}(a) = \begin{cases}\mathcal{J}_{t,s}(a), & \textit{If } t=0 \textit{ or } s=0,\\ \mathcal{J}_{t,s}(a)-\mathcal{J}_{t-1,s-1}(a), & \textit{Otherwise},\end{cases}$$

for $t \geq 0$ *and* $s \geq 0$*.*

Two properties of age-specific fake news matrices follow from Definition 2. First, age-specific Jacobians are partial sums of the diagonals of age-specific fake news matrices. For any $t \geq 0$, $s \geq 0$, and $a \in \{0, 1, ..., A-1\}$,

$$\mathcal{J}_{t,s}(a) = \sum_{k=0}^{\min\{t,s\}} \mathcal{F}_{t-k,s-k}(a). \tag{9}$$

Second, because age-specific Jacobians add up to aggregate Jacobians, age-specific fake news matrices also add up to aggregate fake news matrices:

$$\mathcal{F} = \sum_{a=0}^{A-1} \mathcal{F}(a), \tag{10}$$

where $\mathcal{F}$ is defined as $\mathcal{F}_{t,s} = \mathcal{J}_{t,s} - \mathcal{J}_{t-1,s-1}$.

It is clear, therefore, that if we had age-specific fake-news matrices $\{\mathcal{F}(a)\}_{a=0}^{A-1}$, we would be able to construct age-specific and aggregate Jacobians $\{\mathcal{J}(a)\}_{a=0}^{A-1}$ and $\mathcal{J}$. Before presenting our method for calculating these age-specific fake news matrices, we characterize them in terms of simpler objects.

4.1 Building Blocks of Age-Specific Fake News Matrices

The first objects that we define are age-specific expectation vectors $\mathcal{E}_t(a)$, another specialization of an infinite-horizon object from Auclert et al. (2021a).

Definition 3 (Age-Specific Expectation Vectors)**.** *For $0 \leq a$, $0 \leq t$, and $0 \leq a+t \leq A-1$, the age-specific expectation vector $\mathcal{E}_t(a)$ is*

$$\mathcal{E}_t(a) = \begin{cases} \mathbf{y}_{ss}(a), & \textit{If } t = 0, \\ \left(\prod_{k=a}^{a+t-1} \not{\beta}^k \mathcal{L}_{ss}(k)\right) \mathbf{y}_{ss}(a+t) & \textit{otherwise.} \end{cases}$$

Expectation vectors trace the expected path of agents' outcomes over time. The n-th entry of $\mathcal{E}_t(a)$ is the expected value that the output of an agent who is on the n-th gridpoint of the age-a state space at time 0 will have in period t, when he reaches age $a+t$. These expectations account for the probability that the agent will die and not make it to period t.

The other two types of object that we will use in building age-specific fake news matrices capture the initial effect of the time 0 announcement of a shock scheduled to occur at time $s \geq 0$.[10] Age-specific "policy shifts" capture the effect of the shock on agents' policy functions at time 0. Age-specific "distributional shifts" capture the effect of the shock on the distribution of agents over states at time 1. These objects already fit our notation as $d\mathbf{y}_0^s(a)$ and $d\mathbf{D}_1^s(a)$, but they directly appear in fake news matrices and thus deserve a name and definition.

[10] If $s \leq 1$ these objects contain the effect of both the announcement and realization of the shock.

Definition 4 (Age-Specific Policy Shifts and Distributional Shift)**.** *For age $0 \leq a \leq A-1$ and shock date $s \geq 0$, the age-specific policy shift $d\mathbf{y}_0^s(a)$ and age-specific distributional shift $dD_1^s(a)$ are*

$$d\mathbf{y}_0^s(a) = \mathbf{y}_0^s(a) - \mathbf{y}_{ss}(a) \quad \text{and} \quad d\mathbf{D}_1^s(a) = \mathbf{D}_1^s(a) - \mathbf{D}_{ss}(a).$$

The following lemma characterizes age-specific distributional shifts, showing that they are null in many cases.

Lemma 5. *For $s \geq 0$ and $0 \leq a \leq A-1$,*

$$d\mathbf{D}_1^s(a) = \begin{cases} \not{\beta}^{a-1}\mathcal{L}_0^s(a-1)'\mathbf{D}_{ss}(a-1) - \mathbf{D}_{ss}(a), & \text{If } 0 \leq a-1 \leq A-1-s. \\ 0, & \text{otherwise.} \end{cases} \tag{11}$$

Proof in Appendix A.4.

Agents that are represented in $d\mathbf{D}_1^s(a)$ are of age $a-1$ when the relevant shock is announced ($t = 0$). If $s > (A-1) - (a-1)$, they will be dead by the time the shock occurs. In that case, they do not react to the shock and we have $d\mathbf{D}_1^s(a) = 0$.

4.2 Structure of Fake News Matrices

The following theorem characterizes every entry of age-specific fake news matrices in terms of expectation vectors, policy shifts and distributional shifts.

Theorem 1 (Structure of Age-Specific Fake News Matrices)**.** *For any $0 \leq a \leq A-1$, age-specific Fake-News matrices $\mathcal{F}(a)$ from Definition 2 have the following structure:*

- *First row, $\mathbf{t = 0}$:*

$$\mathcal{F}_{0,s}(a) = \begin{cases} d\mathbf{y}_0^s(a)'\mathbf{D}_{ss}(a), & \text{If } a \leq A-1-s, \\ 0, & \text{Otherwise.} \end{cases}$$

- *Upper-left block outside the first row, $\mathbf{1 \leq t \leq A-1}$ and $\mathbf{0 \leq s \leq A-1}$:*

$$\mathcal{F}_{t,s}(a) = \begin{cases} \mathcal{E}_{t-1}(a-t+1)'d\mathbf{D}_1^s(a-t+1), & \text{If } 0 \leq a-t \leq A-1-s \\ 0, & \text{Otherwise.} \end{cases}$$

- *Outside of upper-left block, $\mathbf{t > A-1}$ or $\mathbf{s > A-1}$:*

$$\mathcal{F}_{t,s}(a) = 0.$$

  Proof in Appendix A.5.

Each case of Theorem 1 has an intuitive interpretation. To that end, remember that $\mathcal{F}_{t,s}(a)$ is linked to the incremental time t response from the cohort of age a to a shock scheduled to occur at time s and announced at time 0. That cohort is of age $a-t$ when the shock is announced and will be of age $a-t+s$ when the shock happens. The main condition for nonzero elements in Theorem 1 is

$$0 \leq a-t \leq A-1-s.$$

This condition requires that the cohort has been born when the shock is announced ($0 \leq a-t$) and that it is also alive when the shock arrives ($a-t \leq A-1-s$).[11]

The first row of $\mathcal{F}(a)$ captures the effect at time 0 of announcing shocks scheduled for time s. Theorem 1 says that, since time-0 distributions are predetermined, contemporaneous responses happen only through time-0 changes in policy functions $d\mathbf{y}_0^s(a)$. Additionally, there is no reaction to shocks expected to arrive after agents' terminal age ($a+s > A-1$).

The second case of Theorem 1 pertains to changes in aggregate outcomes after shocks are announced ($t > 1$). The theorem says that, to a first order approximation, the difference between response at time t to a shock that happens at time s, and the response at time $t-1$ to a shock that happens at time $s-1$, $\mathcal{J}_{t,s}(a) - \mathcal{J}_{t-1,s-1}(a)$, is that the (t,s) case has one additional period of anticipatory effects. This additional period is time 0, when the relevant cohort has age $a-t$. The announcement produces a distributional shift $d\mathbf{D}_1^s(a-t+1)$ and the effect of that shift on the output at time t is $\mathcal{E}_{t-1}(a-t+1)' d\mathbf{D}_1^s(a-t+1)$. There is no additional anticipation if agents had not been born at the time of the announcement ($a-t < 0$), or if the shock arrives after their terminal age ($a-t \leq A-1-s$).

The last case of Theorem 1 says that all nonzero elements of age-specific fake news matrices must be in their first A rows and first A columns. Entries after the Ath row pertain responses at time t to shocks that were announced when none of the cohorts alive at time t had been born. Hence, increasing t and s brings no additional anticipatory effects to aggregate responses. Entries after the Ath column pertain responses to shocks scheduled to occur after the death of all the cohorts that are alive at the time of the announcement. These cohorts do not react to the shocks and thus generate no additional anticipatory effects.

5 Computing Age-Specific Fake News Matrices

This section explains how to compute age-specific fake news matrices $\{\mathcal{F}(a)\}_{a=0}^{A-1}$ using methods and routines that are commonly available in implementations of life cycle models. We start by describing our assumptions about these common implementations and defining additional objects that use them.

[11] The main effect of adopting a more flexible representation of death (one in which death probabilities and the distribution of newborns are not exogenous and constant) is that nonzero terms start appearing in the $t > a$ regions of Fake News matrices. The reason is that "Fake news" shocks can affect agents that were not alive at their announcement through endogenous movements in the distribution of newborns. Terms with the form $d\mathbf{D}_t^s(0)$, which currently disappear, start to matter. However, these terms require no additional solutions of the life cycle model: they can be obtained using steady-state transitions to propagate elements that our method already calculates.

Implementation of life cycle models. Our starting point is a computational implementation of a life cycle model that is solved by backward induction. Aggregate variables are taken as given. We build up from a microeconomic implementation that represents the life of a single cohort and, therefore, has no meaningful distinction between time and age: the cohort has age 0 at time 0, age 1 at time 1, and so on. This type of implementation commonly has the following structure:

- There are age-specific grids that discretize the state-space of every age $\{\mathcal{G}_a\}_{a=0}^{A-1}$.[12]
- There are age-varying parameters that affect the agent's problem, for example, an age-profile of productivity or a sequence of tax rates. We separate the input with respect to which we want to compute Jacobians, $\mathbf{X}$, and write the sequence of age-varying parameters as $\{\vartheta_a, \mathbf{X}_a\}_{a=0}^{A-1}$, where ϑ collects the rest of parameters.
- There are age-specific solvers $\{\texttt{solve}_a(\cdot)\}_{a=0}^{A-1}$ which, given the solution of age a, find the solution of age $a-1$. To lighten the notation, we incorporate the auxiliary parameters $\{\vartheta_a\}$ into the definition of these solvers and stop writing them. The "solution" of an age comprises its policy function $\mathbf{y}(a)$, value function $\mathbf{v}(a)$, and transition matrix to next period, conditional on survival $\mathcal{L}(a)$.

$$\mathbf{y}_a(a),\, \mathbf{v}_a(a),\, \mathcal{L}_a(a) = \texttt{solve}_a\left(\mathbf{v}_{a+1}\left(a+1\right), \mathbf{X}_a\right).$$

  The vectors $\mathbf{y}_a(a)$ and $\mathbf{v}_a(a)$ are defined on grid $\mathcal{G}_a$, and transition matrix $\mathcal{L}_a(a)$ has dimensions $|\mathcal{G}_a| \times |\mathcal{G}_{a+1}|$.

With this representation, we define two additional objects:

- Auxiliary functions that solve the dynamic problem up to an age a. These functions, which we will denote with $\overline{\texttt{solve}}$, take as input the continuation value function $\mathbf{v}_{a+1}\left(a+1\right)$ and parameters up to age a:

$$\{\mathbf{y}_j(j),\, \mathbf{v}_j(j),\, \mathcal{L}_j(j)\}_{j=0}^{a} = \overline{\texttt{solve}}_a\left(\mathbf{v}_{a+1}\left(a+1\right), \{\mathbf{X}_j\}_{j=0}^{a}\right).$$

  Each function $\overline{\texttt{solve}}_a$ chains backward applications of single-period solvers $\{\texttt{solve}_j\}_{j=0}^{a}$.
- The steady state solution is that obtained when the aggregate input is at its steady state value $\mathbf{X} = \mathbf{X}_{ss}$ at every age

$$\{\mathbf{y}_{ss}(j),\, \mathbf{v}_{ss}(j),\, \mathcal{L}_{ss}(j)\}_{j=0}^{A-1} = \overline{\texttt{solve}}_{A-1}\left(\{\mathbf{X}_{ss}\}_{j=0}^{A-1}\right).$$

  Since $A-1$ is the terminal age, $\overline{\texttt{solve}}_{A-1}$ does not have a continuation value function as input. The steady state distribution $\{\mathbf{D}_{ss}(a)\}_{a=0}^{A-1}$ is defined in Lemma 1.

These objects are sufficient to calculate age-specific fake news matrices and sequence space Jacobians.

[12] A reason why grids might be age-varying is that certain states might disappear after certain ages. For example, income shocks to a household might be turned off after retirement.

5.1 From life cycle solutions to building blocks

Theorem 1 expresses age-specific fake news matrices in terms of expectation vectors, distributional shifts, and policy shifts. We now explain how to calculate these building blocks for a given life cycle model using the steady state solution $\{\mathbf{y}_{ss}(j),\, \mathbf{v}_{ss}(j),\, \mathcal{L}_{ss}(j)\}_{j=0}^{A-1}$ and partial solvers $\{\overline{\texttt{solve}}_a(\cdot)\}_{a=0}^{A-1}$ that we just introduced.

Expectation vectors only depend on the steady state solution. From Definition 3, it follows that we can calculate expectation vectors recursively as

$$\mathcal{E}_t(a) = \begin{cases} \mathbf{y}_{ss}(a), & \text{If } t = 0 \\ \not\beta^a \mathcal{L}_{ss}(a)\mathcal{E}_{t-1}(a+1), & \text{If } 0 < t \leq A-1-a. \end{cases} \tag{12}$$

Now consider the distributional shifts and policy shifts that appear in Theorem 1 for any given $0 \leq s \leq A-1$. For $s = 0$, they are $\{\mathbf{y}_0^0(j)\}_{j=0}^{A-1}$ and $\{d\mathbf{D}_1^0(j)\}_{j=1}^{A-1}$. For $s \geq 1$, they are $\{\mathbf{y}_s^0(j)\}_{j=0}^{A-1-s}$ and $\{d\mathbf{D}_1^s(j)\}_{j=1}^{A-s}$. What policy vectors $\mathbf{y}$ and conditional transition matrices $\mathcal{L}$ would we need to produce these shifts? Visualizing them in tables where s increases with rows and the counter j increases with columns, and crossing out the elements that we do not need, we have[13]

$$\begin{array}{ccccc} \mathbf{y}_0^0(0) & \mathbf{y}_0^0(1) & \cdots & \mathbf{y}_0^0(A-2) & \mathbf{y}_0^0(A-1) \\ \mathbf{y}_0^1(0) & \mathbf{y}_0^1(1) & \cdots & \mathbf{y}_0^1(A-2) & \text{✗} \\ \mathbf{y}_0^2(0) & \mathbf{y}_0^2(1) & \cdots & \text{✗} & \text{✗} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ \mathbf{y}_0^{A-2}(0) & \mathbf{y}_0^{A-2}(1) & \cdots & \text{✗} & \text{✗} \\ \mathbf{y}_0^{A-1}(0) & \text{✗} & \cdots & \text{✗} & \text{✗} \end{array} \quad \text{and} \quad \begin{array}{cccc} \mathcal{L}_0^0(0) & \mathcal{L}_0^0(1) & \cdots & \mathcal{L}_0^0(A-2) \\ \mathcal{L}_0^1(0) & \mathcal{L}_0^1(1) & \cdots & \mathcal{L}_0^1(A-2) \\ \mathcal{L}_0^2(0) & \mathcal{L}_0^2(1) & \cdots & \text{✗} \\ \vdots & \vdots & \cdots & \vdots \\ \mathcal{L}_0^{A-2}(0) & \mathcal{L}_0^{A-2}(1) & \cdots & \text{✗} \\ \mathcal{L}_0^{A-1}(0) & \text{✗} & \cdots & \text{✗} \end{array}\,. \tag{13}$$

The s-th row of these tables contains the time-0 response to the announcement of a shock that takes place at time $s \geq 0$. The j-th column contains the responses of agents aged j at the announcement. Agents of age $j > A-1-s$ will be past the terminal age when the shock arrives and, thus, they do not respond.

Now consider sequences of aggregate inputs of length a that have the steady state value $\mathbf{X}_{ss}$ every period except the last (the a-th), where they have $\mathbf{X}_{ss} + dx$. These would be the first a entries of $\mathbf{X}^a$ in the notation of Auclert et al. (2021a). Solving the life cycle model up to age a with such sequences of aggregate inputs and assuming that aggregates revert to steady state afterwards would yield

$$\left\{\mathbf{y}_j^a(j),\, \mathbf{v}_j^a(j),\, \mathcal{L}_j^a(j)\right\}_{k=0}^{a} = \overline{\texttt{solve}}_a\left(\mathbf{v}_{ss}\left(a+1\right), \{\mathbf{X}_k^a\}_{k=0}^{a}\right). \tag{14}$$

These are the policy functions, value functions, and transition matrices of agents that expect a shock to $\mathbf{X}$ at age a that was announced at the time of their birth. We have used our partial solver and the steady state solution to avoid solving for ages greater than a.

Concentrating on policy functions and transition matrices, Lemma 3 implies

$$\left\{\mathbf{y}_j^a(j),\, \mathcal{L}_j^a(j)\right\}_{j=0}^{a} = \left\{\mathbf{y}_0^{a-j}(j),\, \mathcal{L}_0^{a-j}(j)\right\}_{j=0}^{a}.$$

[13] Here we have used Lemma 5, which implies that to get $d\mathbf{D}_1^s(j)$ we need $\mathcal{L}_0^s(j-1)$.

These are precisely the elements of the a-th "anti-diagonal" (bottom-left to top-right diagonals) of the tables in Equation 13. Therefore, given the steady state solution and expectation vectors, we only have to evaluate each of the partial solvers $\{\overline{\texttt{solve}}_a(\cdot)\}_{a=0}^{A-1}$ once and to apply Lemma 5 to get the remaining building blocks that we need for $\{\mathcal{F}(a)\}_{a=0}^{A-1}$. Since partial solver $\overline{\texttt{solve}}_a(\cdot)$ nests the evaluation of single-period solvers $\{\texttt{solve}_j(\cdot)\}_{j=0}^{a}$, this entails evaluating a total of $A \times (A+1)/2$ single-period solvers.

The reduced number of single period solutions explains the performance of our method. Return to the experiment in Table 1, which finds a 300-period Jacobian of an infinite horizon household and compares its cost with those of finding the same Jacobian in an analogue life cycle model with 75 possible ages. Applying the traditional SSJ method to the infinite horizon model requires $H = 300$ solutions of the single-period problem. Applying our method to the life cycle model requires $75 \times 76/2 = 2{,}850$ single period solutions. Therefore, the number of single-period problems that must be solved in the life cycle version of the model grows only by a factor of $2{,}850/300 = 9.5$, which is close to the speed ratio we report in Table 1.

5.2 Algorithm

With the results from the previous subsection, one could first apply partial solvers to obtain all the objects in Equation 13, then use Lemma 5 to get distributional shifts, and then use Theorem 1 to construct age-specific fake news matrices. This approach could become impractical for models of moderate size. The number of points in age-specific state spaces can easily be in the thousands. This means that each of the policy vectors ($\mathbf{y}$) and transition matrices ($\mathcal{L}$) in Equation 13 can have thousands and millions of elements, respectively.[14] Therefore, storing and wrangling the $A \times (A+1)/2$ policy vectors and transition matrices simultaneously can become cumbersome for moderate values of A and sizes of the age-specific state space. For such cases, this section presents an algorithm that constructs age-specific fake news matrices progressively, evaluating partial solvers and updating all the relevant matrix entries before calling the next solver, so that all outputs do not need to be stored in memory simultaneously.

Algorithm 1 presents the routine, which takes the steady state solution and expectation vectors (calculated using equation 12) as its inputs and produces age-specific fake news matrices for every age as its outputs. The algorithm initializes all age-specific fake news matrices with zeros. After that, for every age $0 \leqq k \leq A-1$, it solves the model up to age k with inputs $\mathbf{X}^k$ in line 3, obtaining $\{\mathcal{L}_0^{k-l}(l), \mathbf{y}_0^{k-l}(l)\}_{l=0}^{k}$. Lines 4 to 6 update the first row elements of fake news matrices that use the policy shifts obtained in the partial solution. Lines 7 to 12 update elements outside the first row of fake news matrices that use distributional shifts generated by the transition matrices of the partial solution. The policy vectors and transition matrices can be deleted from memory each time k is updated.[15] After

[14] In practice, transition matrices are often sparse and could be stored using specialized representations.

[15] The algorithm could be even more efficient if $\mathcal{L}_0^{k-l}(l)$ and $\mathbf{y}_0^{k-l}(l)$ were calculated one backward step at a time. All the relevant entries of fake news matrices could be updated after each backward step, and then $\mathcal{L}_0^{k-l}(l)$ and $\mathbf{y}_0^{k-l}(l)$ could be deleted. We believe that the current exposition facilitates the use of the

Algorithm 1 Computing age-specific Fake-News matrices for a single input and output

Input: S.S. Solution: $\{\boldsymbol{\mathcal{L}}_{ss}(a), \mathbf{y}_{ss}(a), \mathbf{D}_{ss}(a)\}_{a=0}^{A-1}$, Expectation Vectors $\{\{\mathcal{E}_t(a)\}_{t=0}^{a}\}_{a=0}^{A-1}$

Initialize A matrices $\{\mathcal{F}(a)\}_{a=0}^{A-1}$ of dimension $A \times A$ with zeros
for $0 \le k \le A-1$ **do**
    $\{\boldsymbol{\mathcal{L}}_0^{k-l}(l), \mathbf{y}_0^{k-l}(l)\}_{l=0}^{k} \leftarrow \overline{\texttt{solve}}_k \left(\mathbf{v}_{ss}\left(k+1\right), \{\mathbf{X}_{ss} + \mathbf{1}_{j=k} \times dx\}_{j=0}^{k}\right)$
    for $0 \le l \le k$ **do**
        $\mathcal{F}_{0,k-l}(l) \leftarrow d\mathbf{y}_0^{k-l}(l)'\mathbf{D}_{ss}(l)$
    end for
    for $0 \le l \le \min\{k, A-2\}$ **do**
        $d\mathbf{D}_1^{k-l}(l+1) \leftarrow \not{\delta}^l \boldsymbol{\mathcal{L}}_0^{k-l}(l)'\mathbf{D}_{ss}(l) - \mathbf{D}_{ss}(l+1)$
        for $1 \le m \le (A-1) - l$ **do**
            $\mathcal{F}_{m,k-l}(l+m) \leftarrow \mathcal{E}_{m-1}(l+1)'d\mathbf{D}_1^{k-l}(l+1)$
        end for
    end for
end for

Output: Fake news matrices $\{\mathcal{F}(a)\}_{a=0}^{A-1}$

completing the previous steps, all the nonzero entries of all age-specific fake news matrices $\{\mathcal{F}(a)\}_{a=0}^{A-1}$ will be filled; the following Lemma formalizes this claim.

Lemma 6. *All the possible nonzero elements of matrices $\{\mathcal{F}(a)\}_{a=0}^{A-1}$ according to Theorem 1 are calculated by Algorithm 1. Also, all the entries of $\{\mathcal{F}(a)\}_{a=0}^{A-1}$ that are calculated in Algorithm 1 are possible nonzero entries according to Theorem 1.*

Proof in Appendix A.6

6 Demonstration

This section applies our method to a simple life cycle version of the household model with idiosyncratic shocks and borrowing constraints that is standard in heterogeneous agent new Keynesian models. The sole goal of the model is to provide concrete depictions of objects such as age-specific fake news matrices and Jacobians, and to illustrate their interpretation.

6.1 Model Description

In our illustrative model, time advances in one-year increments and, since the microeconomic definition of the problem does not distinguish between "time" and "age" we denote both

algorithm for models with existing implementations of partial solvers $\overline{\texttt{solve}}$.

with a. Households are born at the age of 26 ($a = 0$) and live up to a maximum age of 100 ($a = 100 - 26 = 74$); there are $A = 75$ ages in total.

Each period, households receive income either from their work if they are 65 or younger, or from a pension plan if they are older than 65. We use "age fixed-effects" $\{f_a\}_{a=0}^{74}$ that shape the average age profile of earnings. During their working lives, indexing households with i, their income is

$$y_{i,a} = \exp\{f_a + z_{i,a}\} \times w_a, \tag{15}$$

where $z_{i,a}$ is an idiosyncratic shock and w_a is the economy-wide wage rate that prevails when the household reaches age a. From age 66 onward, their income is

$$y_{i,a} = \exp\{f_a + z_{i,a}\} \times d, \tag{16}$$

where d is a parameter that controls the economy-wide level of pensions. We follow the common practice of calculating individual pensions based on the value of the idiosyncratic shock in the household's last working period, $z_{i,(65-26)}$.[16] This amounts to setting $z_{i,a} = z_{i,a-1}$ for $a \geq 40$. Both labor earnings and pensions are taxed at a rate τ.

Households get utility from their consumption only, through an isoelastic function $u(c) = c^{1-\rho}/(1-\rho)$ and discounting the future with an annual factor β. They decide how much of their accumulated assets to consume each period. They can save to protect their consumption against fluctuations in their earnings as they age and receive shocks. We denote their level of savings at the end of age a with $b_{i,a}$. Their savings earn a return factor R_a and they can never borrow ($b_{i,a} \geq 0$).

In sum, the recursive representation of the problem faced by households before the terminal age ($a \leq 73$) is

$$\begin{aligned} V_a(z_{i,a}, b_{i,a-1}) = &\max_{c_{i,a}} \frac{c_{i,a}^{1-\rho}}{1-\rho} + \beta \not{\delta}^a \mathbb{E}\left[V_{a+1}\left(z_{i,a+1}, b_{i,a}\right)\right] \\ &\text{Subject to:} \\ &\quad b_{i,a} = R_a \times b_{i,a-1} + (1-\tau) y_{i,a} - c_{i,a}, \\ &\quad b_{i,a} \geq 0. \end{aligned}$$

where $y_{i,a}$ is defined in Equations 15 and 16, and $z_{i,a+1}$ follows a first order discrete Markov process $z_{i,a+1} \sim \Pi_{a+1}(z_{i,a})$.[17] We represent these Markov processes with age-specific vectors of possible values and transition matrices. After age 65, since $z_{i,a+1} = z_{i,a}$, grids become constant and transitions become identity matrices for $a \geq (65 - 26)$.

In the notation introduced in Section 5, the various elements of our model would be:

- Grids for each age $\{\mathcal{G}_a\}_{a=0}^{A-1}$ would comprise every possible combination of the (discretized) values of $b_{i,a-1}$ and $z_{i,a}$.

[16] See, for example, Carroll (1997) and Kaplan and Violante (2014).

[17] In the terminal age, households consume all their resources and receive no utility from future events $V_{74}(z_{i,74}, b_{i,73}) = \left(R_{74} \times b_{i,73} + (1-\tau)\, y_{i,74}\right)^{1-\rho} / (1-\rho)$.

- Age-varying parameters that are not aggregate inputs $\{\vartheta_a\}_{a=0}^{A-1}$ include the age fixed effects of earnings f_a, and the transition matrices and possible values of income shocks $\Pi_{a+1}(\cdot)$. While constant in our example, elements like the discount factor β could be made age-dependent and would, in that case, be included in ϑ.
- The aggregate inputs that we will use in our example are the interest rate and wage rate, $\{\mathbf{X}_a\}_{a=0}^{A-1} = \{R_a, w_a\}_{a=0}^{A-1}$. Other clear candidates would be the level of pensions and the tax rate on earnings.
- $\mathbf{v}(a)$ are vectors with all the properties of the value function that are necessary for the solution of the model, evaluated on every point of $\mathcal{G}_a$. This particular model can be solved with Carroll's (2006) method of endogenous gridpoints (EGM), which requires only the derivative of the value function with respect to assets,

$$\frac{\partial}{\partial\, b_{i,a-1}} V_a(z_{i,a}, b_{i,a-1}) = R_a \times \left[c_a^*\left(z_{i,a}, b_{i,a-1}\right)\right]^{-\rho}.$$

  where $c_a^*(\cdot)$ is the optimal consumption function for age a. More complex models can require tracking more properties of the value function.[18]
- $\mathbf{y}(a)$ are vectors evaluating outcomes of interest in every point of $\mathcal{G}_a$. Consumption $c_a^*\left(z_{i,a}, b_{i,a-1}\right)$, "cash-in-hand" $m(z_{i,a}, b_{i,a-1}) \equiv R_a \times b_{i,a-1} + (1-\tau)y_{i,a}$, and savings $m(z_{i,a}, b_{i,a-1}) - c_a^*\left(z_{i,a}, b_{i,a-1}\right)$ are examples.
- Transition matrices $\mathcal{L}(a)$ contain the probability that a household will transit from state $(z_{i,a}, b_{i,a-1})$ to state $(z_{i,a+1}, b_{i,a})$ conditional on its survival, for every pair in $\mathcal{G}_a \times \mathcal{G}_{a+1}$.

The solution method for this model, incorporating auxiliary parameters into the function definition, is $\mathbf{y}(a)$, $\mathbf{v}(a)$, $\mathcal{L}(a) = \texttt{solve}_a\left(\mathbf{v}\left(a+1\right), \{R_a, w_a\}\right)$. Practically, the function has four steps. First, it applies EGM to find optimal consumption c_a^* on a grid that does not match $\mathcal{G}_a$. Second, it interpolates consumption onto $\mathcal{G}_a$. Third, it uses on-grid consumption to evaluate the marginal value function and any relevant outcome points needed to construct $\mathbf{v}(a)$ and $\mathbf{y}(a)$. Fourth, it uses on-grid consumption to find on-grid savings and uses the lottery/histogram method of Young (2010) and the income process to find the transition probabilities in $\mathcal{L}(a)$.

Table 2 presents the model's calibration. We normalize the model by the average deterministic component of income, $\bar{y} \equiv (A-1)^{-1} \times \sum_{a=0}^{A-1} e^{f_a}$. For assets $b_{i,a}$, we use a doubly-nested exponential grid with 50 points that goes from 10^{-4} to 500 times $\bar{y}$. We add $b_{i,a} = 0$ as the 51st gridpoint and use the same grid for every age. Productivity $z_{i,a}$ follows an autoregressive process in working years, discretized using 7 points. In total, our state

[18] For example solution methods for discrete-continuous problems can require both the level and derivative of the value function (see, for example, Iskhakov et al. 2017; Dobrescu and Shanker 2024)

[19] We use cross-sectional probabilities for the year 2004.

[20] We regress income on a 5th degree polynomial of age and use the fitted values. Income is the sum of wages and earnings, and pension and Social Security payments.

[21] We use the Rouwenhorst method. Annual persistence is 0.95 and the shocks' standard deviation is 0.2. We use 7 points.

Table 2: Calibration of Illustrative Model

Object	Notation	Value / Source / Description
Relative risk aversion	ρ	2.0
Time discount factor	β	0.98
Death probabilities	$\{\delta^a\}_{a=0}^{A-1}$	SSA Life tables[19]
Life cycle income intercepts	$\{f_a\}_{a=0}^{A-1}$	Survey of Consumer Finances, 2019[20]
Income shock distribution	$\Pi(\cdot)$	Discretized AR(1)[21]
Interest factor	R_{ss}	1.02
Tax rate	τ	0.3
Wage rate	w_{ss}	Normalized to 1.0
Retirement benefits	d	Normalized to 1.0

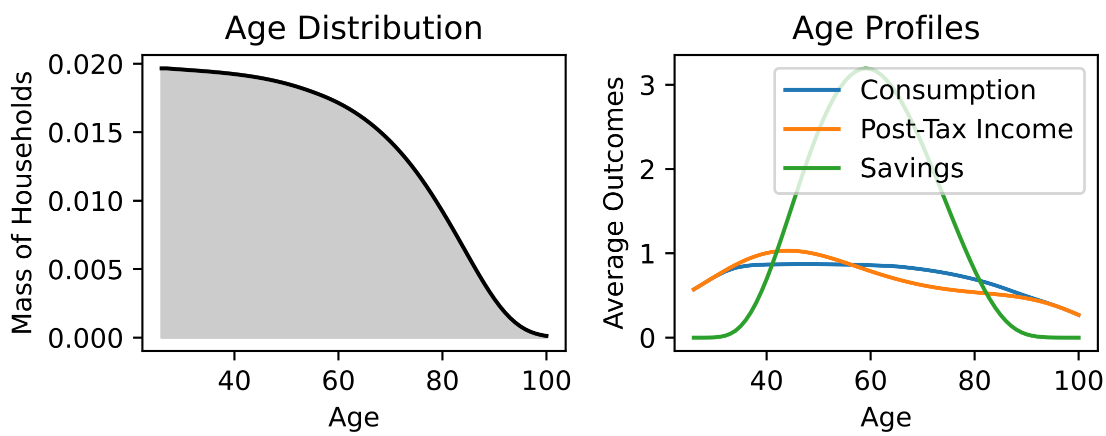
space grids $\mathcal{G}_a$ have $51 \times 7 = 357$ points for every age. For the distribution of newborns over states, η, we assume households enter the model without assets and draw their income shocks from the stationary distribution of the shock process.

6.2 Solution, Fake News Matrices, and Jacobians

Figure 1 depicts the steady state distribution of the model. The left panel shows its demographic structure. The mass of agents declines monotonically with age. The right panel shows the average consumption, post-tax income, and assets among households of every given age. The age profiles reproduce the main features of Modigliani's (1986) life cycle hypothesis of saving. Households save in their peak-earning years anticipating the decline in their income, and they deplete their savings during retirement. They achieve a consumption age profile that is smoother than that of income. In reality, many old households do not deplete their savings (De Nardi, French, and Jones 2010), but our illustrative model lacks the ingredients necessary to reproduce this fact.

We apply Algorithm 1 to our model and obtain age-specific fake news matrices $\{\mathcal{F}(a)\}_{a=0}^{A-1}$ for the interest rate and wage rate as inputs and consumption as the sole output. The time and memory costs are reported in Table 1. Given the steady state solution, it takes on average 0.89 seconds to find the expectation vectors, fake news matrices, and Jacobians of a single input-output pair and a horizon of $T = 300$ periods.[22] Runtimes depend, among other things, on details of the implementation of single-period solvers ($\texttt{solve}_a$). Therefore, we compare these runtimes with the standard fake news algorithm applied to an infinite-horizon model that uses the exact same one-period solver, grids, and stochastic processes as our life cycle model—it just removes age variation. Table 1 also reports the costs of the

[22] The time was averaged over 100 test runs. Tests were performed in a laptop with an Intel(R) Core(TM) Ultra 7 155H processor at 3.80 GHz. The input is the interest rate and the output is consumption.



Notes: the left panel depicts the total mass of agents of each age in the steady state distribution of the model. The total mass of agents is 1.0. The right panel depicts the mean of various outcomes conditional on agents' age.

Figure 1: Steady State Age Structure and Life Cycle Profiles

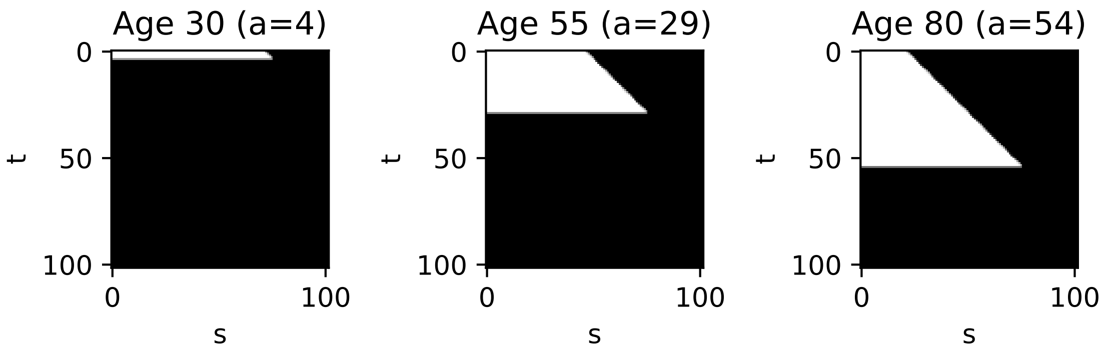
infinite horizon model. Given the steady state solution, it takes on average 0.12 seconds to obtain the aggregate 300-period sequence-space Jacobian for this model. In this simple test, our method performs much better than a naive benchmark of multiplying the infinite-horizon runtime by the $A = 75$ possible ages we add: runtimes increase by a factor closer to 7.3.

We first depict age-specific fake news matrices. The characterization of these matrices in Theorem 1 specifies $0 \leq a - t \leq A - 1 - s$ as a necessary condition for $\mathcal{F}_{t,s}(a) \neq 0$. Figure 2 depicts this condition by plotting the nonzero elements of age-specific fake news matrices for the response of consumption to the interest factor R. The condition configures trapezoidal nonzero regions that become larger as a increases. Zeros may exist in these regions, as the condition is necessary but not sufficient. This may happen, for example, with changes in parameters that do not affect households of all ages; for example, a working household does not react to a change in wages scheduled to happen after he has retired.

We use Equation 9 to obtain age-specific Jacobians and then Equation 8 to obtain aggregate Jacobians. Figure 3 depicts examples of their entries. Panel a) presents the dynamic responses of consumption to a 1 percentage point increase in the interest rate by 30-year-olds, 55-year-olds, and in the aggregate. Panel b) presents the dynamic responses of consumption to a 10 percent increase in the wage rate by 55-year-olds, 80-year-olds, and in the aggregate.

Aggregate Jacobians have their intuitive usual shapes. When households expect a future increase in the interest rate, they reduce their consumption to take advantage of greater returns and, after the shock happens, they slowly consume their excess savings. The aggregate response to wage increases combines the protracted reaction of households with savings that smooth the shock intertemporally, with that of hand-to-mouth agents who simply consume the whole increase when it arrives.

Age-specific Jacobians can deviate considerably from these known shapes, differing qualitatively from aggregate Jacobians, as Figure 3 shows. The differences in their shape are



Notes: each panel depicts the zero and nonzero elements of $\mathcal{F}(a)$ for different values of a. Nonzero elements are white and zero elements are black. The depicted matrices correspond to the response of consumption ($c_a^*(z_{i,a}, b_{i,a-1})$) to the interest factor R.

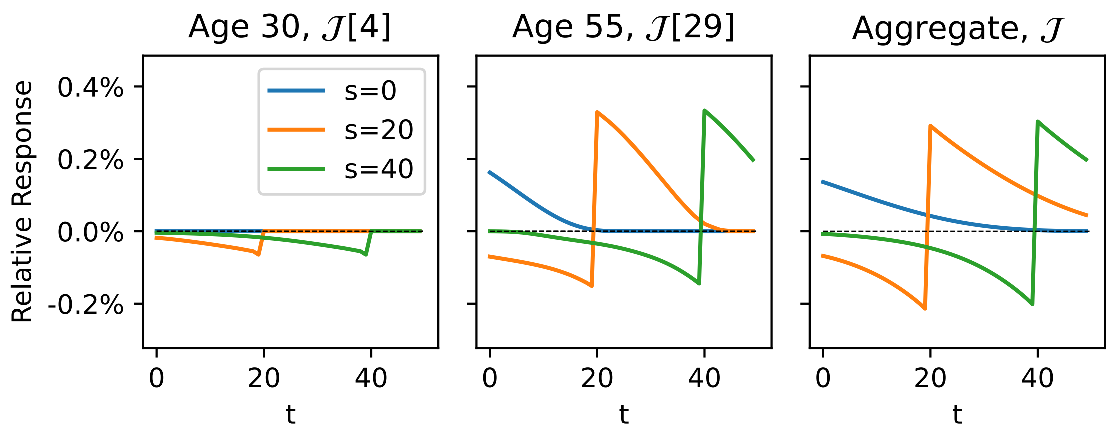
Figure 2: Nonzero Elements of Age-Specific Fake-News Matrices $\mathcal{F}_{t,s}(a)$

illustrative of their interpretation and we explore them in turn.

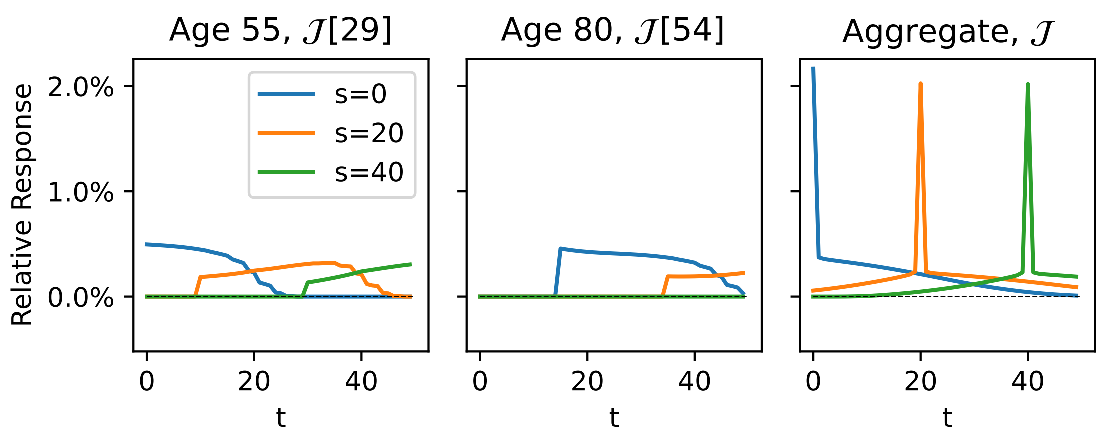
First, we highlight the responses of 30-year-olds to interest rate increases, depicted in the leftmost panel of Figure 3, Panel a). These responses feature increased savings in anticipation to interest rate increments, but no noticeable increase in consumption after their occurrence. The reason behind this apparent mismatch is that the group of agents that configure the response captured in $\{\mathcal{J}_{t,s}[a]\}_{t=0}^{\infty}$ is different for every t—it is the age and not the cohort that is kept constant as time advances. Therefore, it would be wrong to conclude that the agents represented in the $s = 20$ line have accumulated an unusual level of savings by $t \geq 20$ just because the line is negative in $t < 20$. Each point on that line represents a different group of agents and, in fact, for t large enough, the relevant agents had not been born at the time of the shock or its announcement, so their response has to be zero.

Another set of illustrative responses are those of the consumption of 80-year-olds to wage increases, which are depicted in the middle panel of Figure 3, Panel b). The responses to shocks occurring at $s = 0$ and $s = 20$ are initially null, but as t increases they have a discontinuous jump that makes them positive. This shows it would be wrong to conclude that, since 80-year-olds are retired and receive no wages, these Jacobians must be zero. Instead, these Jacobians capture the responses of agents who are 80 years old at time t and, for t sufficiently larger than s, these agents would have been in the job market at the time of the wage increase and benefited from it. Therefore, the discontinuity represents the point at which $80 - (t - s)$—the age of these agents at the time of the shock—starts falling below the age of retirement. After this point, 80-year-olds have a higher consumption due to their higher working-age earnings.

These characteristics of age-specific Jacobians highlight a compromise that we strike in capturing the response of age groups instead of cohorts. There is always the same set of age groups in the economy, while cohorts instead exit the model completely after their last member dies. Splitting responses by ages and not cohorts ensures that one can, for example, calculate the aggregate Jacobian by summing up age-specific Jacobians, with a fixed set *a) Consumption Response to 1 p.p. increase in R.*

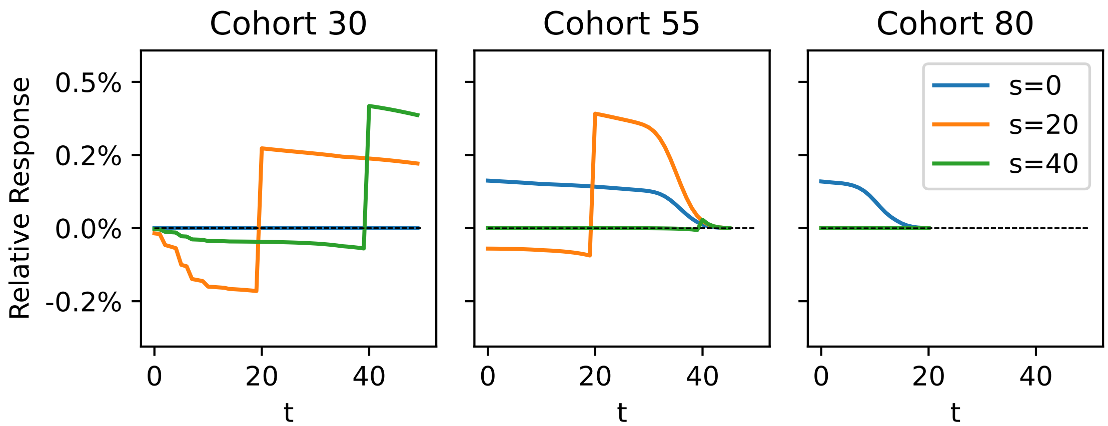


b) Consumption Response to 10% increase in w.



Notes: t denotes the time since the shock was announced and s denotes the time when the shock occurs. Aggregate responses are normalized by steady state aggregate consumption, and age-specific responses are normalized by the total steady state consumption from agents of the given age.

Figure 3: Age-Specific and Aggregate Jacobians of Consumption



Notes: t denotes the time since the shock was announced and s denotes the time when the shock occurs. Responses are normalized by the total steady state consumption from agents of the given age. For example, points in $t = 20$ of the "Cohort 30" panel are normalized by the steady state consumption of 50-year-olds.

Figure 4: Consumption Response of Different Cohorts to a 1 p.p. Increase in R.

of ages. The downside of this decision is that the interpretation and visualization of age Jacobians is muddled by the fact that they capture a different group of agents every period.

However, we can easily construct the dynamic response of any given cohort using age-specific Jacobians. As mentioned in Section 3, the response to a time-s shock announced at time 0 from the cohort that has age a at time 0 is $\{\mathcal{J}_{t,s}(a+t)\}_{t=0}^{A-1-a}$. Figure 4 presents the dynamic response of different cohorts to increases in the interest rate R. Cohorts are labeled using their age at time 0. Tracking consistent groups of agents over time, these responses are easier to interpret and more similar to their aggregate counterparts. Anticipated interest rate increases lead agents to save more until the shock happens, and to consume their extra savings thereafter. Agents that do not anticipate being alive when the shock happens are the exception. For example, the figure shows that "Cohort 80"—those who are 80 years old at time 0—do not respond to shocks that will happen at $s = 40$, because by that time they will have passed the terminal age (100).

7 Concluding Remarks

We have two main goals with the development of this paper. First, we want the derivations and algorithm that we put forth to be an asset for the growing literature that uses models with overlapping generations of heterogeneous agents to tackle what we believe to be fundamental macroeconomic questions (see, for example, Auclert et al. 2021b; Platzer and Peruffo 2022; Gruss et al. 2025). We hope this paper will accelerate the development of existing models and lower the technical barriers for the creation of new ones. Second, we want this paper to serve as a bridge between areas of the microeconomic literature that use life cycle models (for example, labor economics and household finance) and the expanding sequence-space framework for macroeconomic modeling. Beyond the dynamic equilibrium analyses that this framework facilitates, there have been further advances on, for example, incorporating imperfect expectations (Auclert, Rognlie, and Straub 2020; Bardóczy and Guerreiro 2024) and studying optimal policy (Davila and Schaab 2025; Auclert et al. 2024). The literature around this framework has prioritized modularity and interoperability, often taking the sequenc-space Jacobians of a “block” as a starting point. Hence, we hope that researchers will be able to use our method to calculate Jacobians of their life cycle models and use them as blocks to leverage this growing set of tools.

Many technical advances not mentioned so far are available to lower the computational burden of the class of models that we consider in this paper. In particular, most of the techniques that accelerate the solution of the microeconomic life cycle problem or make its representation more efficient directly apply to our setup. A few examples are: using specialized sparse representations for transition matrices, using sparse grids to represent age-specific state spaces (as in Brumm and Scheidegger 2017), or modeling decisions as sequences of stages (as in Bardóczy 2022; Sun 2023). The results developed in this paper can also be adapted to different or broader classes of models by altering our assumptions. We have provided explicit statements of these assumptions and the proofs behind our results to facilitate this process.

References

Andersen, Asger Lau, Niels Johannesen, Mia Jørgensen, and José-Luis Peydró (2023). "Monetary Policy and Inequality". In: *The Journal of Finance* 78.5, pp. 2945–2989.

Auclert, Adrien (2019). "Monetary Policy and the Redistribution Channel". In: *American Economic Review* 109.6, pp. 2333–2367.

Auclert, Adrien, Bence Bardóczy, Matthew Rognlie, and Ludwig Straub (2021a). "Using the Sequence-Space Jacobian to Solve and Estimate Heterogeneous-Agent Models". In: *Econometrica* 89.5, pp. 2375–2408.

Auclert, Adrien, Michael Cai, Matthew Rognlie, and Ludwig Straub (2024). *Optimal Long-Run Fiscal Policy with Heterogeneous Agents*. Pre-published.

Auclert, Adrien, Hannes Malmberg, Frederic Martenet, and Matthew Rognlie (2021b). *Demographics, Wealth, and Global Imbalances in the Twenty-First Century*. URL: https://www.nber.org/papers/w29161 (visited on 04/22/2025). Pre-published.

Auclert, Adrien, Matthew Rognlie, and Ludwig Straub (2018). *The Intertemporal Keynesian Cross*. URL: https://www.nber.org/papers/w25020 (visited on 03/07/2024). Pre-published.

— (2020). *Micro Jumps, Macro Humps: Monetary Policy and Business Cycles in an Estimated HANK Model*. URL: https://www.nber.org/papers/w26647 (visited on 03/04/2024). Pre-published.

Bardóczy, Bence (2022). *Spousal Insurance and the Amplification of Business Cycles*.

Bardóczy, Bence and Joao Guerreiro (2024). *Unemployment Insurance in Macroeconomic Stabilization with Imperfect Expectations*. Pre-published.

Bardóczy, Bence and Mateo Velásquez-Giraldo (2024). *HANK Comes of Age*. URL: https://www.federalreserve.gov/econres/feds/hank-comes-of-age.htm (visited on 09/18/2024). Pre-published.

Beaudry, Paul, Paolo Cavallino, and Tim Willems (2024). *Life-Cycle Forces Make Monetary Policy Transmission Wealth-centric*. URL: https://www.nber.org/papers/w32511 (visited on 06/03/2024). Pre-published.

Bhandari, Anmol, Thomas Bourany, David Evans, and Mikhail Golosov (2023). *A Perturbational Approach for Approximating Heterogeneous Agent Models*. URL: https://www.nber.org/papers/w31744 (visited on 01/15/2025). Pre-published.

Boehl, Gregor (2023). *HANK on Speed: Robust Nonlinear Solutions Using Automatic Differentiation*. URL: https://papers.ssrn.com/abstract=4433585 (visited on 01/21/2025). Pre-published.

Boppart, Timo, Per Krusell, and Kurt Mitman (2018). "Exploiting MIT Shocks in Heterogeneous-Agent Economies: The Impulse Response as a Numerical Derivative". In: *Journal of Economic Dynamics and Control*. Fed St. Louis-JEDC-SCG-SNB-UniBern Conference, Titled: "Fiscal and Monetary Policies". 89, pp. 68–92.

Brumm, Johannes and Simon Scheidegger (2017). "Using Adaptive Sparse Grids to Solve High-Dimensional Dynamic Models". In: *Econometrica* 85.5, pp. 1575–1612.

Carroll, Christopher D. (1997). "Buffer-Stock Saving and the Life Cycle/Permanent Income Hypothesis*". In: *The Quarterly Journal of Economics* 112.1, pp. 1–55.

Carroll, Christopher D. (2006). "The Method of Endogenous Gridpoints for Solving Dynamic Stochastic Optimization Problems". In: *Economics Letters* 91.3, pp. 312–320.

Davila, Eduardo and Andreas Schaab (2025). "Welfare Assessments with Heterogeneous Individuals". In: *Journal of Political Economy*.

De Nardi, Mariacristina, Eric French, and John B. Jones (2010). "Why Do the Elderly Save? The Role of Medical Expenses". In: *Journal of Political Economy* 118.1, pp. 39–75.

Dobrescu, Loretti and Akshay Shanker (2024). *Using Inverse Euler Equations to Solve Multidimensional Discrete-Continuous Dynamic Models: A General Method*. URL: https://papers.ssrn.com/abstract=4850746 (visited on 05/28/2025). Pre-published.

Doepke, Matthias and Martin Schneider (2006). "Inflation and the Redistribution of Nominal Wealth". In: *Journal of Political Economy* 114.6, pp. 1069–1097.

Fagereng, Andreas, Martin B. Holm, and Gisle J. Natvik (2021). "MPC Heterogeneity and Household Balance Sheets". In: *American Economic Journal: Macroeconomics* 13.4, pp. 1–54.

Gruss, Bertrand, Eric Huang, Andresa Lagerborg, Diaa Noureldin, and Galip Kemal Ozhan (2025). "The Rise of the Silver Economy: Global Implications of Population Aging". In: *IMF World Economic Outlook, April 2025: A Critical Juncture Amid Policy Shifts*. World Economic Outlook. International Monetary Fund.

Iskhakov, Fedor, Thomas H. Jørgensen, John Rust, and Bertel Schjerning (2017). "The Endogenous Grid Method for Discrete-Continuous Dynamic Choice Models with (or without) Taste Shocks". In: *Quantitative Economics* 8.2, pp. 317–365.

Kaplan, Greg, Benjamin Moll, and Giovanni L. Violante (2018). "Monetary Policy According to HANK". In: *American Economic Review* 108.3, pp. 697–743.

Kaplan, Greg and Giovanni L. Violante (2014). "A Model of the Consumption Response to Fiscal Stimulus Payments". In: *Econometrica* 82.4, pp. 1199–1239.

Laibson, David (1997). "Golden Eggs and Hyperbolic Discounting*". In: *The Quarterly Journal of Economics* 112.2, pp. 443–478.

Luetticke, Ralph (2021). "Transmission of Monetary Policy with Heterogeneity in Household Portfolios". In: *American Economic Journal: Macroeconomics* 13.2, pp. 1–25.

Modigliani, Franco (1986). "Life Cycle, Individual Thrift, and the Wealth of Nations". In: *Science* 234.4777, pp. 704–712.

Ocampo, Sergio and Baxter Robinson (2023). "Computing Longitudinal Moments for Heterogeneous Agent Models". In: *Computational Economics*.

Peterman, William B. and Erick Sager (2022). "Optimal Public Debt with Life Cycle Motives". In: *American Economic Journal: Macroeconomics* 14.4, pp. 404–437.

Platzer, Josef and Marcel Peruffo (2022). *Secular Drivers of the Natural Rate of Interest in the United States: A Quantitative Evaluation*. URL: https://www.imf.org/en/Publications/WP/Issues/2022/02/11/Secular-Drivers-of-the-Natural-Rate-of-Interest-in-the-United-States-A-Quantitative-512755 (visited on 09/26/2024). Pre-published.

Reiter, Michael (2009). “Solving Heterogeneous-Agent Models by Projection and Perturbation”. In: *Journal of Economic Dynamics and Control* 33.3, pp. 649–665.
Sun, Jeffrey E (2023). *Continuation Value Is All You Need: A Deep Learning Method for Solving Heterogeneous-Agent Models*. Pre-published.
Young, Eric R. (2010). “Solving the Incomplete Markets Model with Aggregate Uncertainty Using the Krusell–Smith Algorithm and Non-Stochastic Simulations”. In: *Journal of Economic Dynamics and Control*. Computational Suite of Models with Heterogeneous Agents: Incomplete Markets and Aggregate Uncertainty 34.1, pp. 36–41.

A Proofs

A.1 Proof of Lemma 1

Proof. Let $\mu_t(a) \equiv \mathbf{1}'\mathbf{D}_t(a)$, denote the total mass of agents of a given age at time t. Since $\mathbf{D}_t$ is a probability distribution, it must always be the case that

$$1 = \left(\sum_{a=0}^{A-1} \mu_t(a) \right). \tag{17}$$

Also, pre-multiplying Equation 5 by $\mathbf{1}'$, we get

$$\mu_t(a) = \begin{cases} \sum_{a=0}^{A-1} \delta^a \times \mu_{t-1}(a), & a = 0 \\ \not\delta^{a-1} \mu_{t-1}(a-1), & 0 < a \leq A-1. \end{cases} \tag{18}$$

This equation implies that $\mu_t(a) = \mu_{t-a}(0) \times \not\Delta^a$. Replacing this into Equation 17, we have

$$1 = \left(\sum_{a=0}^{A-1} \mu_{t-a}(0) \times \not\Delta^a \right).$$

Therefore, any steady state distribution of the model must satisfy $\mu_{ss}(0) = \left(\sum_{a=0}^{A-1} \not\Delta^a \right)^{-1}$.

Now consider a steady state version of Equation 5 for $a = 0$,

$$\begin{aligned} \mathbf{D}_{ss}(a) &= \left(\sum_{a=0}^{A-1} \delta^a \times \mu_{ss}(a) \right) \eta = \left(\sum_{a=0}^{A-1} \delta^a \times \mu_{ss}(0) \times \not\Delta^a \right) \eta \\ &= \mu_{ss}(0) \times \underbrace{\left(\sum_{a=0}^{A-1} \delta^a \times \not\Delta^a \right)}_{=1} \eta = \mu_{ss}(0) \times \eta = \left(\sum_{a=0}^{A-1} \not\Delta^a \right)^{-1} \times \eta. \end{aligned}$$

Plugging this initial condition into the law of motion in Equation 5 gives the result. □

A.2 Proof of Lemmas 2, 3 and 4

Proof. Apply the recursive definitions in Equation 6 until the terminal age to arrive at

$$\begin{aligned} \mathbf{v}_t^s(a) &= v[a] \left(\mathbf{v}_{t+1}^s (a+1), \mathbf{X}_t^s \right) \\ &= v[a] \left(v[a+1] \left(\mathbf{v}_{t+2}^s (a+2), \mathbf{X}_{t+1}^s \right), \mathbf{X}_t^s \right) \\ &\equiv v[a](\mathbf{X}_t^s) \circ v[a+1](\mathbf{X}_{t+1}^s) \circ \mathbf{v}_{t+2}^s (a+2) \\ &\dots \\ &= v[a](\mathbf{X}_t^s) \circ v[a+1](\mathbf{X}_{t+1}^s) \circ \dots \circ v[A-1](\mathbf{X}_{t+A-1-a}^s) \end{aligned} \tag{19}$$

and, likewise,

$$\begin{aligned}\mathbf{y}_t^s(a) &= y[a](\mathbf{X}_t^s) \circ v[a+1](\mathbf{X}_{t+1}^s) \circ ... \circ v[A-1](\mathbf{X}_{t+A-1-a}^s) \\ \mathcal{L}_t^s(a) &= \mathcal{L}[a](\mathbf{X}_t^s) \circ v[a+1](\mathbf{X}_{t+1}^s) \circ ... \circ v[A-1](\mathbf{X}_{t+A-1-a}^s).\end{aligned} \tag{20}$$

Now remember that $\mathbf{X}^s$ is a series with the value of $\mathbf{X}_{ss} + dx$ in its s-th entry and $\mathbf{X}_{ss}$ in every other. That structure has the following implications:

- If $s < t$ or $s - t > A - 1 - a$, all the entries of $\mathbf{X}^s$ that appear in Equation 20 are X_{ss}. Therefore,

$$\begin{aligned}\mathbf{y}_t^s(a) &= y[a](\mathbf{X}_{ss}) \circ v[a+1](\mathbf{X}_{ss}) \circ ... \circ v[A-1](\mathbf{X}_{ss}) = \mathbf{y}_{ss}(a) \\ \mathcal{L}_t^s(a) &= \mathcal{L}[a](\mathbf{X}_{ss}) \circ v[a+1](\mathbf{X}_{ss}) \circ ... \circ v[A-1](\mathbf{X}_{ss}) = \mathcal{L}_{ss}(a).\end{aligned} \tag{21}$$

  This proves Lemmas 2 and 4.

- For any $t \geq 0$, $s \geq 0$ and $k \geq 0$, $\mathbf{X}_t^s = \mathbf{X}_{t+k}^{s+k}$ (its value depends only on whether the subscript is equal to the superscript). Therefore,

$$\begin{aligned}\mathbf{y}_{t+k}^{s+k}(a) &= y[a](\mathbf{X}_{t+k}^{s+k}) \circ v[a+1](\mathbf{X}_{t+1+k}^{s+k}) \circ ... \circ v[A-1](\mathbf{X}_{t+A-1-a+k}^{s+k}) \\ &= y[a](\mathbf{X}_t^s) \circ v[a+1](\mathbf{X}_{t+1}^s) \circ ... \circ v[A-1](\mathbf{X}_{t+A-1-a}^s) = \mathbf{y}_t^s.\end{aligned}$$

  and

$$\begin{aligned}\mathcal{L}_{t+k}^{s+k}(a) &= \mathcal{L}[a](\mathbf{X}_{t+k}^{s+k}) \circ v[a+1](\mathbf{X}_{t+1+k}^{s+k}) \circ ... \circ v[A-1](\mathbf{X}_{t+A-1-a+k}^{s+k}) \\ &= \mathcal{L}[a](\mathbf{X}_t^s) \circ v[a+1](\mathbf{X}_{t+1}^s) \circ ... \circ v[A-1](\mathbf{X}_{t+A-1-a}^s) = \mathcal{L}_t^s.\end{aligned}$$

  This proves Lemma 3.

□

A.3 Additional properties of distributions

The following two properties will be used in the proofs that follow.

Lemma 7 (Shocks do not change initial distributions)**.** *For any $s \geq 0$ and $0 \leq a \leq A-1$, $\mathbf{D}_0^s(a) = \mathbf{D}_{ss}(a)$.*

Proof. Applying Equation 5 to period $t = 0$ and noting that we are assuming that the system starts from steady state so that $\mathbf{D}_{-1}(a) = \mathbf{D}_{ss}(a)$ and $\mathcal{L}_{-1}(a) = \mathcal{L}_{ss}(a)$ for any a, we have

$$\mathbf{D}_0(a) = \begin{cases} \left(\sum_{j=0}^{A-1} \delta^j \times \mathbf{1}'\mathbf{D}_{ss}(j)\right)\eta = \mathbf{D}_{ss}(0), & a = 0 \\ \not{\delta}^{a-1}\mathcal{L}_{ss}(a-1)'\mathbf{D}_{ss}(a-1) = \mathbf{D}_{ss}(a), & 0 < a \leq A-1. \end{cases}$$

□

Lemma 8 (Shocks do not change newborn distribution)**.** *For any $t \geq 0$ and $s \geq 0$, $\mathbf{D}_t^s(0) = \mathbf{D}_{ss}(0)$.*

Proof. The rows of any valid transition matrix must sum to 1. Therefore, $\mathcal{L} \times \mathbf{1} = \mathbf{1}$ and this implies that Equation 18 holds even when there are shocks that change transition matrices.

Lemma 7 implies that $\mu_0^s(a) = \mu_{ss}(a)$ for any a. Then, applying Equation 18 to time $t = 1$, we have

$$\mu_1(a) = \begin{cases} \sum_{j=0}^{A-1} \delta^j \times \mu_{ss}(j) = \mu_{ss}(0), & a = 0 \\ \not\delta^{a-1} \mu_{ss}(a-1) = \mu_{ss}(a), & 0 < a \leq A-1. \end{cases}$$

Thus $\mu_1^s(a) = \mu_{ss}(a)$.

We can keep going period by period to conclude that

$$\mu_t^s(a) = \mu_{ss}(a) \qquad \forall t \geq 0,\ s \geq 0,\ 0 \leq a \leq A-1.$$

Therefore, replacing this result into the law of motion for the newborn distribution (Equation 5),

$$\mathbf{D}_{j+1}^s(0) = \left(\sum_{a=0}^{A-1} \delta^a \mu_j^s(a) \right) \eta = \left(\sum_{a=0}^{A-1} \delta^a \mu_{ss}(a) \right) \eta = \mathbf{D}_{ss}(0) \qquad \forall j.$$

□

A.4 Proof of Lemma 5

Proof. We know that

$$\mathbf{D}_1^s(a) = \begin{cases} \mathbf{D}_{ss}(0), & \text{If } a = 0, \\ \not\delta^{a-1} \mathcal{L}_0^s(a-1)' \mathbf{D}_{ss}(a-1), & \text{If } 1 \leq a \leq A-1. \end{cases}$$

The distribution of newborns never changes (Lemma 8). For $a \geq 1$, the distribution of agents at time 1 is the steady state distribution at time 0 (when the shock was announced) rolled forward by the $\mathcal{L}_0^s(a-1)$ transition matrix that accounts for the response to the shock, and adjusted for survival.

Subtracting $\mathbf{D}_{ss}(a)$,

$$d\mathbf{D}_1^s(a) = \begin{cases} 0, & \text{If } a = 0, \\ \not\delta^{a-1} \left(\mathcal{L}_0^s(a-1) - \mathcal{L}_{ss}(a-1) \right)' \mathbf{D}_{ss}(a-1), & \text{If } 1 \leq a \leq A-1. \end{cases} \tag{22}$$

Finally, we know from Lemma 4 that $\mathcal{L}_0^s(a-1) = \mathcal{L}_{ss}(a-1)$ if $(a-1)+s > A-1$, which simplifies to $a > A-s$. Replacing these elements and ranges into Equation 22 yields Lemma 5.

□

A.5 Proof of Theorem 1

Theorem 1 is a combination of results describing different sections of age-specific fake news matrices. We first establish these results.

First row.

Lemma 9. *For $t = 0$, $s \geq 0$, and $0 \leq a \leq A - 1$,*

$$\mathcal{F}_{0,s}(a) = \mathcal{J}_{0,s}(a)\, dx = \begin{cases} d\mathbf{y}_0^s(a)'\mathbf{D}_{ss}(a), & \textit{If } s \leq A - 1 - a, \\ 0, & \textit{Otherwise.} \end{cases}$$

Proof. Note that, for $0 \leq s$,

$$\mathcal{J}_{0,s}(a)dx = (\mathbf{y}_0^s(a)'\mathbf{D}_0^s(a)) = d\mathbf{y}_0^s(a)'\mathbf{D}_{ss}(a)$$

because $d\mathbf{D}_0^s(a) = 0$ given that $\mathbf{D}_0$ is fixed at the time of the shock. Furthermore, Lemma 4 implies $d\mathbf{y}_0^s(a) = 0$ for $s > A - 1 - a$. Therefore,

$$\mathcal{J}_{0,s}(a) = \begin{cases} d\mathbf{y}_0^s(a)'\mathbf{D}_{ss}(a), & \text{If } s \leq A - 1 - a, \\ 0, & \text{Otherwise.} \end{cases}$$

□

First column.

Lemma 10. *For $t \geq 1$, $s = 0$, and $0 \leq a \leq A - 1$*

$$\mathcal{F}_{t,0}(a) = \mathcal{J}_{t,0}(a)\, dx = \begin{cases} \mathcal{E}_{t-1}(a - t + 1)'d\mathbf{D}_1^0(a - t + 1) & \textit{If } 0 \leq a - t, \\ 0 & \textit{Otherwise.} \end{cases}$$

Proof. Consider

$$\begin{aligned} \mathcal{J}_{t,0}(a)dx &= d(\mathbf{y}_t^0(a)'\mathbf{D}_t^0(a)) \\ &= d\mathbf{y}_t^0(a)'\mathbf{D}_{ss}(a) + \mathbf{y}_{ss}(a)'d\mathbf{D}_t^0(a) \end{aligned}$$

For $t > a$ both $d\mathbf{y}_t^0(a)$ and $d\mathbf{D}_t^0(a)$ are zero because the relevant cohort had not been born at the time of the shock's occurrence.

Now, for any $0 < t \leq a$ and $0 \leq a \leq A - 1$, Lemma 2 implies $d\mathbf{y}_t^0(a) = 0$ and, therefore

$$\mathcal{J}_{t,0}(a)dx = \mathbf{y}_{ss}(a)'d\mathbf{D}_t^0(a).$$

For $t = 1$, the last line is $\mathcal{E}_0(a)'\mathbf{D}_1^0(a)$ and we are done. For $t > 1$, we continue expanding,

$$
\begin{aligned}
\mathcal{J}_{t,0}(a)dx &= \mathbf{y}_{ss}(a)'d\mathbf{D}_t^0(a) \\
&= \mathbf{y}_{ss}(a)'d(\not\beta^{a-1}\mathcal{L}_{t-1}^0(a-1)'\mathbf{D}_{t-1}^0(a-1)) \\
&= \mathbf{y}_{ss}(a)'d(\not\beta^{a-1}\mathcal{L}_{ss}(a-1)'\mathbf{D}_{t-1}^0(a-1)) \\
&= \mathbf{y}_{ss}(a)'\not\beta^{a-1}\mathcal{L}_{ss}(a-1)'d\mathbf{D}_{t-1}^0(a-1) \\
&= \mathbf{y}_{ss}(a)'\not\beta^{a-1}\mathcal{L}_{ss}(a-1)'\not\beta^{a-2}\mathcal{L}_{ss}(a-2)'d\mathbf{D}_{t-2}^0(a-2) \\
&\dots \\
&= \mathbf{y}_{ss}(a)'\not\beta^{a-1}\mathcal{L}_{ss}(a-1)'\dots\not\beta^{a-(t-1)}\mathcal{L}_{ss}(a-(t-1))'d\mathbf{D}_1^0(a-(t-1)) \\
&= \mathcal{E}_{t-1}(a-t+1)'d\mathbf{D}_1^0(a-t+1).
\end{aligned}
$$

□

Elements outside the first row and column.

Lemma 11. *For $t \geq 1$, $s \geq 1$, $0 \leq a \leq A-1$*

$$
\mathcal{F}_{t,s}(a) \equiv \mathcal{J}_{t,s}(a) - \mathcal{J}_{t-1,s-1}(a) = \begin{cases} \mathcal{E}_{t-1}(a-t+1)'d\mathbf{D}_1^s(a-t+1) & \textit{If } 0 \leq a-t, \\ 0, & \textit{Otherwise.} \end{cases} \tag{23}
$$

Proof. Note that, for $s \geq 1$ and $t \geq 1$,

$$
\begin{aligned}
\mathcal{F}_{t,s}(a)dx &= d\mathbf{Y}_t^s(a) - d\mathbf{Y}_{t-1}^{s-1}(a) \\
&= \underbrace{(d\mathbf{y}_t^s(a) - d\mathbf{y}_{t-1}^{s-1}(a))}_{0}{}'\mathbf{D}_{ss}(a) + \mathbf{y}_{ss}(a)'(d\mathbf{D}_t^s(a) - d\mathbf{D}_{t-1}^{s-1}(a)) \\
&= \mathbf{y}_{ss}(a)'(d\mathbf{D}_t^s(a) - d\mathbf{D}_{t-1}^{s-1}(a)),
\end{aligned} \tag{24}
$$

where the second equality is due to Lemma 3.

Now consider the following cases for $d\mathbf{D}_t^s(a) - d\mathbf{D}_{t-1}^{s-1}(a) = \mathbf{D}_t^s(a) - \mathbf{D}_{t-1}^{s-1}(a)$:

- If $a = 0$: $\mathbf{D}_t^s(a) - \mathbf{D}_{t-1}^{s-1}(a) = 0$ due to Lemma 8.
- If $a > 0$: note that

$$
\begin{aligned}
d\mathbf{D}_t^s(a) =& d\left(\not\beta^{a-1}\mathcal{L}_{t-1}^s(a-1)'\mathbf{D}_{t-1}^s(a-1)\right) \\
=& \not\beta^{a-1}d\mathcal{L}_{t-1}^s(a-1)'\mathbf{D}_{ss}(a-1) + \\
& \not\beta^{a-1}\mathcal{L}_{ss}(a-1)'d\mathbf{D}_{t-1}^s(a-1).
\end{aligned}
$$

So that

$$
\begin{aligned}
d\mathbf{D}_t^s(a) - d\mathbf{D}_{t-1}^{s-1}(a) =& \not\beta^{a-1}(d\underbrace{\mathcal{L}_{t-1}^s(a-1) - d\mathcal{L}_{t-2}^{s-1}(a-1)}_{=0})'\mathbf{D}_{ss}(a-1) + \\
& \not\beta^{a-1}\mathcal{L}_{ss}(a-1)'(d\mathbf{D}_{t-1}^s(a-1) - d\mathbf{D}_{t-2}^{s-1}(a-1)) \\
=& \not\beta^{a-1}\mathcal{L}_{ss}(a-1)'(d\mathbf{D}_{t-1}^s(a-1) - d\mathbf{D}_{t-2}^{s-1}(a-1)).
\end{aligned} \tag{25}
$$

We can apply Equation 25 to itself recursively. The end of the recursion will depend on the relationship between a and t.

– If $t > a$, the recursion proceeds as

$$\begin{aligned}
d\mathbf{D}_t^s(a) - d\mathbf{D}_{t-1}^{s-1}(a) =& \beta^{a-1}\mathcal{L}_{ss}(a-1)'(d\mathbf{D}_{t-1}^s(a-1) - d\mathbf{D}_{t-2}^{s-1}(a-1)) \\
=& \beta^{a-1}\mathcal{L}_{ss}(a-1)'\beta^{a-2}\mathcal{L}_{ss}(a-2)' \\
& (d\mathbf{D}_{t-2}^s(a-2) - d\mathbf{D}_{t-3}^{s-1}(a-2)) \\
\ldots \\
=& \beta^{a-1}\mathcal{L}_{ss}(a-1)' \ldots \beta^{0}\mathcal{L}_{ss}(0)' \\
& \underbrace{(d\mathbf{D}_{t-a}^s(0) - d\mathbf{D}_{t-a-1}^{s-1}(0))}_{=0} = 0.
\end{aligned}$$

The last equality follows from Lemma 8.

– If $t \geq a$, the recursion proceeds as

$$\begin{aligned}
d\mathbf{D}_t^s(a) - d\mathbf{D}_{t-1}^{s-1}(a) =& \beta^{a-1}\mathcal{L}_{ss}(a-1)'(d\mathbf{D}_{t-1}^s(a-1) - d\mathbf{D}_{t-2}^{s-1}(a-1)) \\
=& \beta^{a-1}\mathcal{L}_{ss}(a-1)'\beta^{a-2}\mathcal{L}_{ss}(a-2)' \\
& (d\mathbf{D}_{t-2}^s(a-2) - d\mathbf{D}_{t-3}^{s-1}(a-2)) \\
\ldots \\
=& \beta^{a-1}\mathcal{L}_{ss}(a-1)' \ldots \beta^{a-(t-1)}\mathcal{L}_{ss}(a-(t-1))' \\
& (d\mathbf{D}_1^s(a-(t-1)) - \underbrace{d\mathbf{D}_0^{s-1}(a-(t-1))}_{=0}) \\
=& \beta^{a-1}\mathcal{L}_{ss}(a-1)' \ldots \beta^{a-(t-1)}\mathcal{L}_{ss}(a-(t-1))' \\
& d\mathbf{D}_1^s(a-(t-1)) \\
=& \left(\prod_{k=1}^{t-1} \beta^{a-k}\mathcal{L}_{ss}(a-k)'\right) d\mathbf{D}_1^s(a-(t-1)) \\
=& \mathcal{E}_{t-1}(a-(t-1))' d\mathbf{D}_1^s(a-(t-1)).
\end{aligned}$$

In the derivation, $d\mathbf{D}_0^{s-1}(a-(t-1)) = 0$ due to Lemma 7.

Plugging the different cases for $d\mathbf{D}_t^s(a) - d\mathbf{D}_{t-1}^{s-1}(a)$ into Equation 24 yields Lemma 11.

□

Proof of Theorem 1

Proof. The first case in the Lemma, $t = 0$, is a re-statement of Lemma 9.

The second case in the Lemma, $1 \leq t \leq A-1$ and $0 \leq s \leq A-1$, combines Lemmas 10, 11, and 5.

- If $s = 0$, then $a - t \leq A - 1 - s$ will always be true and, eliminating this restriction, the case becomes a re-statement of Lemma 10.

- If $s \geq 1$ instead, then Lemma 11 applies. From Lemma 5 we know that $d\mathbf{D}_1^s(a-t+1) \neq 0$ only if $0 \leq (a-t+1)-1 \leq A-1-s$. Adding this condition to Lemma 11, we have that $\mathcal{F}_{t,s}(a) = \mathcal{E}_{t-1}(a-t+1)' d\mathbf{D}_1^s(a-t+1)$ can differ from zero only if

$$0 \leq a-t \wedge 0 \leq (a-t+1)-1 \wedge (a-t+1)-1 \leq A-1-s.$$

  This expression reduces to $0 \leq a-t \leq A-1-s$, the condition in the case.

The third case of the Lemma has two parts. First, Lemmas 10 and 11 imply $\mathcal{F}_{t,s}(a) = 0$ for $a < t$ and any s. This implies $\mathcal{F}_{t,s}(a) = 0$ for $t \geq A$ and any (a, s) pair. Second, Lemma 5 implies that $d\mathbf{D}_1^s(a) = 0$ for any $s \geq A$ and we also know $d\mathbf{y}_0^s(a)$ for any $s > A-1-a$ from Lemma 4. Applying these facts to Lemmas 9 and 11, we know that $\mathcal{F}_{t,s}(a) = 0$ for $s \geq A$ and any (a, t) pair.

□

A.6 Proof of Lemma 6

A.6.1 Proof that all nonzero elements are computed by the algorithm

Proof. Given $0 \leq a \leq A-1$, $0 \leq t \leq A-1$ and $0 \leq s \leq A-1$, consider the possible cases for a nonzero $\mathcal{F}_{t,s}(a)$ according to Theorem 1:

- If $t = 0$, then it must be the case that $a \leq A-1-s$, which is $s+a \leq A-1$. Algorithm 1 finds $\mathcal{F}_{0,s}(a)$ on line 5, when $l = a$ and $k = s+a$. Line 5 is reached under these conditions because
  - The condition on line 4 is satisfied since

$$0 \leq l = a \leq k = s+a.$$

  - The condition on line 2 is satisfied since $0 \leq k = s+a \leq A-1$.
- If $t \geq 1$, then it must be the case that $0 \leq a-t \leq A-1-s$. Algorithm 1 finds $\mathcal{F}_{t,s}(a)$ on line 10, when $m = t$, $l = a-t$, and $k = s+a-t$. Line 10 is reached under these conditions because
  - The condition on line 9 is satisfied since $1 \leq t = m$ and

$$m = t \leq (A-1)-(a-t) = (A-1)-l$$

    since $t \leq (A-1)-(a-t) \leftrightarrow a \leq (A-1)$ which is true by the definition of a.
  - The condition in line 7 is satisfied. First, $l = a-t \leq (A-1)-1 = A-2$ because $a \leq A-1$ and $t \geq 1$. It is also the case that $l = a-t \leq s+a-t = k$ because we assumed $s \geq 0$. Finally, $0 \leq a-t = l$ by assumption of the current case.
  - The condition in line 2 is satisfied. First, $0 \leq s+(a-t) = k$ because $s \geq 0$ and $a-t \geq 0$. Second, $k = s+(a-t) \leq A-1 \leftrightarrow a-t \leq A-1-s$, which is an assumption of the current case.

□

A.6.2 Proof that all the elements that the algorithm computes are possible nonzeros

Proof. Algorithm 1 assigns entries to fake news matrices in lines 5 and 10. Both cases correspond to elements that Theorem 1 flags as potential nonzero terms.

- Line 5 assigns $\mathcal{F}_{0,k-l}(l)$ for some $0 \leq k \leq A-1$ and $0 \leq l \leq k$. Letting $s \equiv k-l \geq 0$ and $a \equiv l$ we are assigning $\mathcal{F}_{0,s}(a)$. We know that $a+s=k \leq A-1$ and thus $a \leq A-1-s$. This is the condition for nonzero elements in the first row of $\mathcal{F}(a)$ in Theorem 1.
- Line 10 assigns $F_{m,k-l}(l+m])$ for some $0 \leq k \leq A-1$, $0 \leq l \leq \min\{k, A-2\}$ and $1 \leq m \leq (A-1)-l$. Letting $t \equiv m$, $s \equiv k-l$, and $a \equiv l+m$, we are assigning $\mathcal{F}_{t,s}(a)$. Since $t=m \geq 1$, this is not an element in the first row of the matrix. First, $a-t=l \geq 0$ as a condition of line 7. Also, since $k \leq A-1$, $A-1-k+l \geq l$. Therefore,

$$l \leq A-1-(k-l) \to a-t \leq A-1-s.$$

  So that, in sum, we have, $t \geq 1$ and $0 \leq a-t \leq A-1-s$ which are the conditions for a nonzero element below the first row of $\mathcal{F}(a)$ in Theorem 1.

□