Approximating Transition Dynamics with Discrete Choice

Anmol Bhandari* David Evans† Ellen McGrattan‡

December 10, 2025

Abstract

This paper develops a method for analyzing policy reforms in general equilibrium settings with discrete choice. Computing transition paths in these settings is computationally challenging, particularly in models with substantial heterogeneity and many endogenous states. We extend perturbation methods to handle discrete choice by appropriately tracking both intensive-margin changes conditional on discrete choices that are relatively small and extensive-margin changes resulting from a switch in a discrete choice that are relatively large. The method is fast, scalable, and efficient, providing good initial estimates for global solution methods. We demonstrate our method by analyzing optimal business taxation in a model with occupational choice between entrepreneurship and paid employment.[1]

*University of Minnesota

†University of Oregon

‡University of Minnesota and Federal Reserve Bank of Minneapolis

[1]We thank Fernando Alvarez, Adrien Auclert, Adrien Bilal, Lars Peter Hansen, Greg Kaplan, Matt Rognlie, and participants of the 2024 Economic Dynamics, Uncertainty, and Computation conference for useful comments. This research has been supported by National Science Foundation grant SES-2214248. The replication code for the application studied in the paper is available on https://github.com/apb-umn/BEM. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

1 Introduction

The analysis of policy reforms in dynamic economies has long been a central focus of macroeconomics. Although substantial progress has been made in analyzing policies in representative agent frameworks, the incorporation of rich household and firm heterogeneity and discrete choices remains computationally challenging. This challenge is particularly acute when studying important policy questions impacting occupational choice, firm entry and exit, default decisions, and migration. In these settings, policy changes affect both continuous choices (for example, consumption and investment) and discrete choices (for example, whether to become an entrepreneur). The latter environments often feature non-convexities that complicate traditional solution methods. This paper extends the perturbation method of Bhandari et al. (2023) to handle these non-convexities and demonstrates the utility of the method for analyzing transition dynamics and optimal business taxation in a model with occupational choice between paid employment and entrepreneurship.

The standard approach to computing transition dynamics in these models involves solving for the equilibrium path between initial and final steady states by finding a path of prices and allocations that simultaneously satisfy all individual optimization conditions and market clearing requirements at all dates. This requires solving a massive system of nonlinear equations, typically by truncating the transition at some distant horizon and applying iterative or Newton-type algorithms. This global solution approach suffers from severe computational limitations if there are many endogenous aggregate variables, or if individual optimization problems are computationally intensive.

Our aim in this paper is to develop a complementary solution method that is fast and scalable, delivering an entire transition path at nearly the same cost as computing a steady state. To do that, we will use a perturbational approach in which we represent the equilibrium transition path using an appropriate perturbation of the distributional state around the post-reform steady state. This reformulation transforms the computationally hard global problem into a tractable local approximation around a known solution. To deal with equilibrium policy functions that have discontinuities due to discrete choice, we utilize taste shocks with positive variance. These shocks ensure tractability when computing extensive-margin changes resulting from agents switching between discrete alternatives. While intensive-margin responses are smooth and amenable to standard perturbation techniques, extensive-margin responses—for example, the mass of agents switching occupations—require special treatment. With taste shocks, the set of potentially switching agents spans the entire state space rather than concentrating on a lower-dimensional manifold of indifference. Our approach transforms the complex problem of tracking an endogenous density along an indifference boundary into a simple reweighting of the existing distribution using smooth probability functions.

Our approach offers several advantages. First, it is computationally efficient and scalable to high-dimensional state spaces, making it feasible to analyze models with substantial household and firm heterogeneity. Second, the approach provides a natural decomposition of policy effects into intensive- and extensive-margin responses, which facilitates economic interpretation. Third, even when global solution methods are feasible, our method generates good initial guesses that can accelerate their convergence. Finally, the approach readily handles general equilibrium effects and transition dynamics, which are often abstracted from in discrete choice settings.

We use a second-order approximation method to compute the optimal tax rate on private business income in a model with occupational choice between paid employment and entrepreneurship. This application introduces features of interest that make standard solution methods challenging: a discrete occupational choice, heterogeneity in assets and productivity, financial constraints that generate non-convexities, and general equilibrium price effects. Using parameters consistent with U.S. national accounts and government budgets, we compute transitional dynamics for a one-time rate change and for a change that is phased in over time. Since U.S. business owners have limited third-party reporting, the effective tax rate on their business profits is roughly 20 percent—about half of what is assessed on wages and salaries of paid employees. Computing the optimal tax and the associated transition path, we find an optimal tax rate for business income of 64.2 percent, which yields an average consumption-equivalent welfare gain of 0.8 percent. The main findings are quantitatively robust across a range of parameterizations of financial constraints, tax elasticities, and income risk.

1.1 Related Literature

Our work contributes to two strands of literature in macroeconomics: the analysis of discrete choice in dynamic economies and the development of computational methods to solve heterogeneous agent models. We bridge these literatures by developing techniques that can handle both rich household and firm heterogeneity and discrete choices while maintaining computational tractability.

In the main application of our method, we compute transitional dynamics in a general equilibrium model of occupational choice featuring borrowing and collateral constraints, which are now standard in theories of entrepreneurship (see, for example, Evans and Jovanovic (1989), Quadrini (2000), Cagetti and De Nardi (2006), Buera, Kaboski, and Shin (2011), and Midrigan and Xu (2014)). Our framework allows for analysis of how policy reforms interact with financial constraints in determining occupational choice and capital allocation, while properly accounting for transition dynamics and general equilibrium effects. Previous studies have focused primarily on steady-state analysis. Recent exceptions are Catherine (2022), who estimates the option value for a return to paid employment if an entrepreneur fails in business, and Brüggemann (2021), who estimates the top marginal income tax rate in an environment with occupational choice.

Our methodological approach builds on two recent advances in computational methods. First, the projection-perturbation approach of Reiter (2009) opened the door to tractably analyzing aggregate shocks in heterogeneous agent economies.[2] Second, the sequence-space approaches developed by Boppart, Krusell, and Mitman (2018) and refined by Auclert et al. (2021) enabled more efficient computation of first-order impulse responses to an unanticipated shock in discrete time.[3] Our work is closest to that of Bhandari et al. (2023), who show how to unify and extend the earlier methods by characterizing equilibrium dynamics through sequences of linear systems for directional derivatives. Our contribution here is to further extend the Bhandari et al. (2023) approach so that it accommodates environments in which discrete choices are central, and to do so by exploiting techniques from the theory of distributions and weak derivatives. The idea is to account properly for both the intensive-margin responses of continuous choices and the extensive-margin responses of discrete choices.[4]

The remainder of the paper is organized as follows. Section 2 presents the model environment,

[2]Reiter (2009) directly approximates the full first order law of motion for the distributional state. This can quickly become intractable as the size of the idiosyncratic state space grows, necessitating some kind of model reduction technique to reduce the size of those derivatives. See, for example, Ahn et al. (2017), Childers (2018), Winberry (2018), Gornemann, Kuester, and Nakajima (2021), Bayer, Born, and Luetticke (2022); and Reiter (2023).

[3]A parallel branch of the literature studies continuous time versions of these heterogeneous agent models. See, for example, Ahn et al. (2017), Kaplan and Violante (2018), Moll et al. (2022), Alvarez, Lippi, and Souganidis (2023), and Bilal (2023).

[4]See also Bardóczy (2021) for an adaptation of the toolkit based on Auclert et al. (2021) to a model with discrete labor supply and aggregate shocks.

which serves as a motivating application. Section 3 develops our methodology for approximating transition dynamics with discrete choice. Section 4 discusses the numerical implementation and empirical application to optimal business taxation. Section 5 concludes.

2 Motivating Application

We consider an incomplete-markets economy in discrete time populated by a continuum of infinitely lived individuals indexed by $i \in [0, 1]$. Individuals face idiosyncratic risk and choose between paid employment and entrepreneurship. A representative firm operates in the corporate sector. The government levies taxes on income, profits, and consumption and provides lump-sum transfers.

The discrete choice made by individual i at time t is summarized by the value of d_{it}, which is equal to one for paid employment and zero for entrepreneurship. The choice depends on the individual's productivity in paid work, θ^w_{it}, their idiosyncratic taste for paid work, η_{it}, their productivity operating a business, θ^b_{it}, and their beginning-of-period assets, a_{it-1}.

We include the taste shock η_{it} to capture non-pecuniary reasons for shaping occupational choice. A negative value of $\eta_{i,t}$ could reflect non-pecuniary benefits of self-employment as in Hamilton (2000). The shock is assumed to be independently and identically distributed (i.i.d.) with distribution Γ. If there were no such motivations, we could set the variance of the shock to zero and the occupational choice would be found by maximizing over the value functions for the two occupations. A positive variance of taste shocks makes the choice probabilistic. Economically, the size of the taste shock variance governs the elasticity of occupational choice with respect to changes in the value functions. The variance ensures that the ex-ante value functions are smooth in the individual states, which can have computational advantages.

It is convenient to log and stack the productivity shocks—which are assumed to be exogenous stochastic processes—into a vector $\theta_{it} = [\log \theta^w_{it}, \log \theta^b_{it}]$. Let $v^o_t(a_{i,t-1}, \theta_{it})$ be the value of an individual in occupation $o \in \{w, b\}$ with financial wealth $a_{i,t-1}$ and productivity θ_{it}, and let $v_t(a_{i,t-1}, \theta_{it}, \eta_{it})$ be the value of the same individual before choosing an occupation but after observing the taste shock η_{it}. These value functions satisfy Bellman equations,

$$v_t(a_{i,t-1}, \theta_{it}, \eta_{it}) = \max_{d_{it} \in \{0,1\}} \; d_{it}\{v^w_t(a_{i,t-1}, \theta_{it}) + \eta_{it}\} + (1 - d_{it}) v^b_t(a_{i,t-1}, \theta_{it}). \tag{1}$$

The value functions for occupation o are given by

$$v_t^o(a_{i,t-1}, \theta_{it}) = \max_{c_{it}, a_{it} \geq \underline{a}} U(c_{it}) + \beta \mathbb{E}_t[v_{t+1}(a_{it}, \theta_{i,t+1}, \eta_{i,t+1})] \tag{2}$$

$$\text{subject to} \quad (1 + \tau_{ct})c_{it} + a_{it} = R_t a_{i,t-1} + y_{it}^o, \tag{3}$$

where $\beta \in (0, 1)$ is the discount factor, $U(\cdot)$ is a strictly increasing, strictly concave utility function, τ_{ct} is the period-t tax rate on consumption, R_t is the return on assets, and y_{it}^o is the income after taxes and transfers for individuals choosing occupation o. Income and business profits are given by

$$y_{it}^w = (1 - \tau_{wt}) W_t \theta_{it}^w + T_t \tag{4}$$

$$y_{it}^b = (1 - \tau_{bt}) \pi_{it} + T_t \tag{5}$$

$$\pi_{it} = \max_{\substack{k_{it}^b \leq \chi a_{i,t-1}, \\ n_{it}^b}} \theta_{it}^b f^b(k_{it}^b, n_{it}^b) - (R_t - 1 + \delta) k_{it}^b - W_t n_{it}^b, \tag{6}$$

where W_t is the wage and labor earnings are taxed at rate τ_{wt}, π_{it} is the business profit taxed at rate τ_{bt}, T_t is government transfers, $\chi \geq 1$ is the maximum leverage ratio, k_{it}^b and n_{it}^b are inputs of capital and labor, respectively, for entrepreneurs that operate a decreasing-returns-to-scale technology $f^b(\cdot)$, and δ is the rate of capital depreciation. The decreasing-returns-to-scale captures the managerial factor that limits the size of the business. Without entrepreneurship, the individual's problem is similar to that analyzed by Aiyagari (1994). Without paid employment, the problem is similar to that analyzed by Lucas (1978) and its extensions in Quadrini (2000) and Cagetti and De Nardi (2006).

A representative firm operates in the corporate sector with a constant-returns-to-scale production technology:

$$Y_{ct} = \Theta f^c(K_{ct}, N_{ct}), \tag{7}$$

where Θ is aggregate productivity, K_{ct} is corporate capital, and N_{ct} is corporate labor. Profit maximization yields factor prices R_t and W_t equal to the marginal products of capital and labor, respectively. In equilibrium, these prices clear the asset and labor markets, and the government's budget balances:

$$K_{ct} + \int (1 - d_{it}) k^b_{it} \, \mathrm{di} + B_{t-1} = \int a_{i,t-1} \, \mathrm{di} \tag{8}$$

$$N_{ct} + \int (1 - d_{it}) n^b_{it} \, \mathrm{di} = \int d_{it} \theta^w_{it} \, \mathrm{di} \tag{9}$$

$$\begin{aligned} G_t + T_t + R_t B_{t-1} = B_t + \tau_{ct} \int c_{it} \, \mathrm{di} + \tau_{wt} \int d_{it} W_t \theta^w_{it} \, \mathrm{di} \\ + \tau_{bt} \int (1 - d_{it}) \pi_{it} \, \mathrm{di} + \tau_{pt} (Y_{ct} - W_t N_{ct} - \delta K_{ct}), \end{aligned} \tag{10}$$

where G_t is government spending on goods and services, $R_t B_{t-1}$ is gross repayments on the last period's debt, B_t denotes government debt held by households at the end of period t, and τ_{pt} is the tax rate on corporate profits.

Given initial conditions $\{a_{i,-1}, \theta_{i,0}\}_{i \in [0,1]}$, a competitive equilibrium consists of factor prices $\{R_t, W_t\}_{t \geq 0}$, individual choices $\{c_{it}, a_{it}, d_{it}, k^b_{it}, n^b_{it}\}_{t \geq 0, i \in [0,1]}$, corporate choices $\{K_{ct}, N_{ct}\}_{t \geq 0}$, and government fiscal variables $\{G_t, T_t, B_t, \tau_{ct}, \tau_{wt}, \tau_{bt}, \tau_{pt}\}_{t \geq 0}$ such that (i) given prices and fiscal policies, individual choices solve the worker and entrepreneur problems; (ii) corporate choices maximize profits; (iii) the government budget constraint holds; and (iv) markets clear.[5]

3 Approximation Methods

In this section, we describe approximation methods to compute the equilibrium transition path between an initial steady state and a post-reform steady state assuming individuals have perfect foresight about the trajectory of policy changes. Before presenting our approximation method, we establish notation for a class of economic environments to which our methods will be applicable. We then review the insights from Bhandari et al. (2023) that we build on and extend to environments with discrete choice.

3.1 Representation of Economies

We start by introducing notation for the aggregate states, individual choices, and aggregate variables that we need to keep track of. We then summarize conditions for an equilibrium.

[5] If there is growth in the economy—say, because of changes in population or technology—the variables would be detrended before an equilibrium is computed, and we would include a growth rate γ as a parameter.

Let $Z_t = [A_{t-1}, \Omega_t]'$ denote the aggregate state vector, where A_{t-1} is the vector of predetermined aggregate variables and Ω_t is the distribution over the individual states. In our motivating application of Section 2, the predetermined aggregate variable A_{t-1} is the aggregate beginning-of-period asset holdings $\int a_{i,t-1}\,\mathrm{di}$, and the distribution Ω_t is over the individual states $(a_{i,t-1}, \theta_{it})$.[6]

Let x^d_{it} denote the vector of individual choices, indexed with 'd' to indicate that the choices are conditional on the individual making one of D possible discrete choices. Let $d_{it} \in D$ denote the discrete choice made by individual i at time t. Let $x_{it} := \sum_{d \in D} \iota(d_{it} = d)\, x^d_{it}$ denote the value after occupational choice, where $\iota(\cdot)$ is an indicator function that is equal to 1 if the argument is true and 0 otherwise. In our motivating application, there are two values for the superscript d—namely, w for workers in paid employment and b for entrepreneurs. The elements of x_{it} include consumption, next period's assets, factors of production in business, and value functions. The occupational choice d_{it} is determined by solving the problem in equation (1). For a setting with $D > 2$ choices, we expect the researcher to introduce $D - 1$ taste shocks and modify equation (1) to be the maximum over all choices. In our $D = 2$ case, the choice variable d_{it} is equal to 0 or 1. The ex-ante probability of being a worker given states $(a_{i,t-1}, \theta_{it})$ is given by $\Gamma(v^b_{it} - v^w_{it})$, where Γ is the complementary cumulative distribution function (survival function) of the taste shock η_{it}.

Let X_t denote the vector of aggregate variables that are not distributions, with $A_t \subseteq X_t$ denoting the subset of predetermined aggregate variables. To represent the equilibrium conditions compactly, it is convenient to stack together all variables relevant for period t decision-making: $Y_t = [A_{t-1}, X_t, X_{t+1}]$.

3.1.1 Representation of equilibrium conditions

We can summarize the main conditions for an equilibrium as follows:

$$0 = F^d\Big(a_{i,t-1}, \theta_{it}, x^d_{it}, E_{it}x_{i,t+1}, Y_t\Big), \qquad \text{for all } i, t, d, \tag{11}$$

$$0 = G\Big(\textstyle\int x_{it}\,\mathrm{di}, Y_t\Big), \qquad \text{for all } t, \tag{12}$$

$$0 = \mathbb{E}_{it}\Big[H\Big(\{x^d_{i,t+1}\}_d\Big)\Big] - E_{it}x_{i,t+1}, \qquad \text{for all } i, t. \tag{13}$$

[6] Although aggregate beginning-of-period assets can be obtained as a moment of the distribution Ω_t, we keep it as a separate component to allow a cleaner representation of the class of models.

Here, F^d includes the individual optimality conditions for discrete choice d, G includes equilibrium conditions involving aggregates, and H summarizes how expectations over next period's discrete choices depend on the taste shocks. In our motivating application, F^w and F^b include the individual budget constraints, the first-order conditions, and the Bellman equations for the dynamic programs of workers in paid employment and entrepreneurs, respectively. Also included in F^b are additional conditions related to entrepreneurial production. The conditions in G include asset market clearing and government budget balance. The elements of H include expectations of the marginal values and the value functions with respect to the realized taste-shock process, that is,

$$H(\{x^d_{i,t+1}\}_d) = \int \sum_{d\in D} \left(x^d_{i,t+1}(a_{it}, \theta_{i,t+1}) + u^d_x\, \eta_{i,t+1} \right) \iota(d_{i,t+1}(a_{it}, \theta_{i,t+1}, \eta_{i,t+1}) = d)\; \mathrm{d}\Gamma(\eta_{i,t+1}),$$

where u^d_x is a selector vector that equals 1 in the component of $x^d_{i,t+1}$ corresponding to the value function v^d and 0 otherwise, so that the taste shock $\eta_{i,t+1}$ affects only the value-function component as in equation (1). Later, we impose that the taste shock η_{it} follows a Gumbel distribution, and the integral can then be solved in closed form.

There are two expectation operators appearing in equations (11)–(13). The variable $E_{it}x_{i,t+1}$ is the expectation of the individual's choices in the next period. This expectation—like the other arguments in (11)—only depends on the individual states $(a_{i,t-1}, \theta_{it})$. For example, we include the expected continuation value in the Bellman equation. The operator $\mathbb{E}_{it}$ in (13) is an expectation over next period's productivity shocks. The fact that it is not an expectation over the taste shock follows from the fact that we included the taste-shock expectations when defining the function H.

With this notation, we can describe the solution concept that we are interested in approximating. As before, given initial conditions and sequences of exogenous policy parameters, an *equilibrium path* is characterized by sequences $\{X_t,\, x^d_{it}(\theta^t),\, d_{it}(\theta^t, \eta^t)\}$ that satisfy F^d, G, and H.

3.2 Approximations

Consider a one-time permanent change in a policy variable, such as a tax rate, that occurs at date 0. Let $Z_{-1} = [A_{-1}, \Omega_0]'$ denote the initial aggregate state and $Z^* = [A^*, \Omega^*]'$ denote the final steady state following the policy reform. If we know Z_{-1} and have a good guess for the sequence $\{X_t\}$ over time—for example, for the sequence of prices and government policies—we can use this information together with F^d in (11) to construct the decision rules for individual and aggregate policies, and then check to see if market clearing conditions and other conditions in G are satisfied. Such an exercise is the same as solving for the fixed point of a very large system of equations—say,

$$\mathcal{F}(\{X_t\}; Z_{-1}) = 0 \tag{14}$$

for $\{X_t\}$, where $\mathcal{F}$ is the set of the equilibrium equations shown in (11)–(13).

The standard approach is to truncate at some large T and solve the nonlinear fixed-point problem in (14) over $\{X_t\}$ using iterative Newton or quasi-Newton methods. This approach does not scale well in models with high-dimensional X_t or computationally intensive individual problems. Our goal is to develop an approximation to the transition path that is fast and scalable—one that delivers an entire transition path at nearly the same cost as computing a steady state. This is done by building on the perturbation methods proposed by Bhandari et al. (2023), modified to allow for extensive-margin changes that occur because of discrete choice. We show next that despite the additional complication, the problem boils down to solving a small linear system for a path of directional derivatives—just as in Bhandari et al. (2023).

The perturbation method we use replaces the system of equations in (14) with a Taylor expansion of $\mathcal{F}$ around the *post-reform steady state*. More specifically, we replace the initial state in (14) with

$$Z_0 = Z^* + \sigma(Z_{-1} - Z^*), \tag{15}$$

and construct a transition path that is a function of the parameter σ. A value of $\sigma = 0$ corresponds to the post-reform steady state, and a value of $\sigma = 1$ corresponds to the pre-reform initial economy. A Taylor expansion with respect to σ at $\sigma = 0$ approximates the transition path as an impulse response of the dynamic system with respect to aggregate state Z.

In Sections C and D of the appendix, we show that a Taylor expansion of

$$\mathcal{F}\left(\{X_t(\sigma)\}; Z^* + \sigma(Z_{-1} - Z^*)\right) = 0 \tag{16}$$

generates a linear systems of equations for the vectors of the sequences $\{\hat{X}_t\}$ and $\{\hat{X}_{t,t}\}$, so that

$$X_t = \overline{X} + \hat{X}_t + \frac{1}{2}\hat{X}_{t,t} + o(\|Z_{-1} - Z^*\|^2), \tag{17}$$

where $\overline{X}$ is the vector of aggregate variables in the final steady state, and $\hat{X}_t$ and $\hat{X}_{t,t}$ are first- and second-order adjustments arising from the policy change that occurs at date 0.

To derive the linear system, we express the equilibrium conditions in state-space form (see appendix equations (47)–(49)). In this formulation, individual and aggregate policies are represented by functions $\overline{x}(a, \theta, Z)$ and $\overline{X}(Z)$, with the aggregate law of motion $\overline{Z}(Z)$. The first- and second-order adjustments correspond to directional derivatives of these functions with respect to the aggregate state Z, evaluated along directions that trace the equilibrium transition path. Specifically, the first-order adjustment satisfies $\hat{X}_t = \bar{X}_Z \cdot \hat{Z}_t$, where $\bar{X}_Z$ is the Fréchet derivative of $\overline{X}$ with respect to Z, and the direction $\hat{Z}_t$ evolves recursively according to $\hat{Z}_t = \bar{Z}_Z \hat{Z}_{t-1}$ with the initial direction $\hat{Z}_0 = Z_{-1} - Z^*$. Representing these adjustments as directional derivatives is useful because it yields a linear system for the sequences $\{\hat{X}_t\}$ and $\{\hat{X}_{t,t}\}$ in terms of the derivatives of F^d, G, and H (see (11)–(13)), together with the individual policy functions and aggregate variables evaluated at the final steady state.

Two key challenges must be addressed when analyzing this system of equations. First, perturbing equilibrium conditions involves differentiating objects that have kinks (due to borrowing constraints) and jumps (due to discrete choices). Bhandari et al. (2023) explain in detail how to deal with kinks that arise from borrowing constraints using the machinary of generalized functions. To deal with this jumps to do discrete choice, we utilize taste shocks with strictly positive variance. While this assumption is primarily technical, it provides substantial computational advantages. To understand its importance, consider that computing changes in aggregates requires tracking how the set of agents who are indifferent between discrete choices evolves with aggregate states—a key determinant of extensive margin responses to small changes in aggregate conditions. In the absence of taste shocks, this indifference set forms a lower-dimensional manifold in the (a, θ) space, and the measure of agents along this manifold depends on an endogenous density that must be computed as part of the equilibrium solution. This endogeneity significantly complicates the computational problem. By contrast, when taste shocks are present, the set of potentially indifferent agents spans the entire support of Ω, and the relevant measure becomes a tractable reweighting of the original distribution. This tractability greatly simplifies tracking how the extensive margin responds to aggregate state changes.

Researchers interested in the limiting case without taste shocks can implement our method using arbitrarily small taste shock variance, with the appropriate tolerance determined by standard heuristics regarding the insensitivity of aggregates to small perturbations. One important caveat is that smaller taste shock variances require finer discretization grids to accurately capture discrete choice responses to changes in idiosyncratic states. In our application, we take the pragmatic approach of viewing taste shocks as determinants of aggregate elasticities and calibrate their variance to match empirical estimates of the relevant elasticities.

The second challenge involving the system of equations (11)–(13) is the inclusion of the cross-sectional distribution over assets and productivity types $\Omega(a, \theta)$—an infinite-dimensional object—in the state vector. Next, we describe how this issue is handled when computing the first- and second-order approximations (with full details in the appendix).

3.2.1 First-order approximation

To compute a first-order approximation to the transition path, we first differentiate the F^d and H mappings to express the changes in individual decisions conditioned on discrete choice d as:

$$\hat{x}_t^d(a, \theta) = \sum_{s=0}^{\infty} \mathsf{x}_s^d(a, \theta) \hat{Y}_{t+s}, \tag{18}$$

where $\hat{Y}_t = \left[\mathsf{P}\hat{X}_{t-1}, \hat{X}_t, \hat{X}_{t+1}\right]^{\mathrm{T}}$ with matrix P selecting adjustments for the state variables in A from vector $\hat{X}$. The loadings $\{\mathsf{x}_s^d(a, \theta)\}_s$ in (18) are derived in closed form in the appendix (see Lemma 2). Once we have the individual policies, we can derive the approximation to the law of motion of the distributional state, Ω_t, which is fully characterized by a collection of linear operators over functions of (a, θ). The operators are used to account for the changes in distribution due to both intensive- and extensive-margin choices. To handle the extensive margin, one needs to keep track of the change in the discrete choice thresholds for all individuals. Since the change in the discrete choice thresholds depends on changes in the value functions, which are part of the individual choices x_{it}^d, we can use the loadings $\{\mathsf{x}_s^d(a, \theta)\}_s$ to compute them (see Lemma 3 in the appendix for details).

The final step in computing the first-order approximation is to differentiate the mapping G in (12), which results in the following linear system:

$$\mathsf{G}_Y \hat{Y}_t + \mathsf{G}_x \sum_{s=0}^{\infty} \mathsf{J}_{t,s} \hat{Y}_s + \mathsf{G}_x \mathsf{J}_t^{TD} = 0, \tag{19}$$

where the G coefficients are derivatives of the mapping G with respect to its arguments and are evaluated at the final steady state, and matrices $\mathsf{J}_{t,s}$ and J_t^{TD} are defined recursively using loadings $\{\mathsf{x}_s^d(a,\theta)\}_s$ and the set of operators used to approximate the law of motion of Ω_t. Formulas are provided in Section C.3 of the appendix.

3.2.2 Second-order approximation

Following the same steps, we obtain the second-order approximation as a linear system of equations for the sequence $\{\hat{X}_{t,t}\}$, which is given by

$$\mathsf{G}_Y \hat{Y}_{t,t} + \mathsf{G}_x \sum_{s=0}^{\infty} \mathsf{J}_{t,s} \hat{Y}_{s,s} + \mathsf{G}_x \mathsf{H}_{t,t} + \hat{\mathsf{G}}_{t,t} = 0, \tag{20}$$

where $\hat{\mathsf{G}}_{t,t}$ is the second-order adjustment to the equilibrium conditions and $\mathsf{H}_{t,t}$ is the second-order adjustment to the law of motion of the distributional state. The derivation of these terms is provided in Section D of the appendix.

Thus far, we have described how to compute the transition path following a one-time change in a policy variable. The mathematical formulas described above and in the appendix do not change when we compute transitions in response to changes in the policy *path*. We simply need to include the policy variable in vector X_t and keep track of the path of reforms enacted at time zero as well as all changes in the endogenous aggregate variables. As an example, consider a one-time increase in a tax rate versus a gradual increase over a specified time period. In the case of a one-time change, we simply change one parameter when computing the final steady state and transition. In the case of a path change, we add the tax rate to the aggregate state vector and keep track of all enacted changes.

4 Empirical Application

In this section, we return to the motivating application from Section 2 and apply the second-order perturbation method described in Section 3.2. We start by providing details of the numerical implementation, including user-defined inputs for computing the initial equilibrium, the final steady state, and the transition paths. We then calibrate the model using U.S. national accounts data. With the calibrated model, we report statistics for the initial equilibrium intended to mimic certain features of U.S. data. We compute and describe the transition path following a policy change that raises the tax rate on business income—either immediately or gradually. Finally, we compute the optimal business tax rate and do a sensitivity analysis as we vary parameters governing the tightness of the collateral constraint, the tax elasticity of aggregate labor demand, and the amount of income risk.

4.1 Numerical Implementation

For numerical implementation, we need to specify functional forms for preferences and technologies, and processes for exogenous shocks.

The functional forms for preferences and technologies that we use are given by

$$U(c) = \frac{c^{1-\mu} - 1}{1-\mu} \tag{21}$$

$$f^b(k,n) = k^{\phi} n^{\nu} \tag{22}$$

$$f^c(K,N) = K^{\alpha} N^{1-\alpha}. \tag{23}$$

The period utility function U is of the constant relative risk aversion type, and μ is the relative risk aversion parameter. The technology used by business owners is f^b and is assumed to exhibit decreasing returns to scale; that is, $\phi + \nu < 1$. The technology of the corporate sector is assumed to be Cobb-Douglas with the capital share parameter α.

The exogenous shocks in our problem are the productivities for workers, θ^w, and business owners, θ^b, and the individual's i.i.d. taste shocks η. As noted earlier, we define a new productivity vector $\theta_{it} = [\log \theta^w_{it}, \log \theta^b_{it}]$, which is modeled here as a Markov chain with discrete states and transition matrix Π_θ. For the taste shock, we use a Gumbel distribution because of its analytical tractability and parametrize it with the scale parameter σ_η. More specifically, given values for business v^b and work v^w, the probability p^w that the individual chooses paid employment is:

$$p^w(a_-, \theta) = \left(1 + \exp\left(\frac{v^b(a_-, \theta) - v^w(a_-, \theta)}{\sigma_\eta}\right)\right)^{-1}. \tag{24}$$

4.1.1 Zeroth Order Steady State

The zeroth-order steady state is the stationary equilibrium for a given set of parameters and time-independent government policy. Off-the-shelf methods can be used to compute such equilibria, and our approximation does not depend on the specific choice.[7] From those off-the-shelf methods, we expect the user to extract the stationary policy functions and transition matrix that governs the stationary distribution of individuals over the state space. We find it useful to represent the policy functions as weighted sums of basis functions:

$$c(a, \theta_i) \approx \sum_j \mathsf{c}_j^i \, b_j(a),$$

where $b_j(a)$ is typically a spline and the transition matrix is stored in sparse form. When derivatives of policies are needed we use second-order splines; otherwise, linear splines suffice. The grids on which the policies are computed are non-uniform and are denser in regions where the policies are more non-linear. The grids on which the stationary distribution and transition matrix are stored are also non-uniform with more points than the policy grids. Using the policy functions, we also store the cutoffs for the borrowing and collateral constraints as functions of exogenous productivity θ.

4.1.2 Setup for the Perturbation Method

Next, we discuss the implementation of $\{F^d\}_{d\in\{b,w\}}$, G, and H in equations (11)–(13). For our problem, we need four sets of inputs for the constraints that are included in F^b. The first set are the states that include beginning-of-period assets a_- and productivity shocks θ. The second set is the vector of individual choices, which has the same length as the number of conditions in F^b.

[7] With discrete choices the maximization problems may be non-concave, requiring additional care. See Druedahl (2021) for efficient methods in such settings. In our specific application, we found that an adaptation of the endogenous grid method (EGM) with ex-post verification through brute-force grid search works well. See Section B in the appendix for details.

In our case, the length is 10 and the vector is given by $x = [a, \pi, y^b, k^b, n^b, \lambda, c, n, v, d]$—and thus includes the individual assets a, business profits π, business output y^b, business factor inputs k^b and n^b, the marginal value of wealth adjusted by the consumption tax $\lambda := (1+\tau_c)V_a^o$, consumption c, labor supply n, value v, and occupational choice d. The third set of inputs is the vector of aggregates, which in this case includes the interest rate and transfers: $X = [R, T]$. We also include the tax rate on business, τ_b, in vector X when analyzing a gradual change in the tax rate. The fourth set of inputs is the vector of expectations, including the marginal values of wealth adjusted by the consumption tax and the value function. Here, we adopt a convenient shorthand notation, namely, $x^e = [\lambda^e, v^e]$.[8]

4.1.3 Mapping F^b

Corresponding to the input vector x are the constraints in F^b, which include the following 10 conditions: the budget constraint; the definition of π; the definition of y^b; the capital first-order condition; the labor first-order condition; the marginal value of wealth; the Euler equation; the condition equating labor supply and demand; the Bellman equation; and the occupational identity. Mathematically, these conditions are as follows:

$$0 = Ra_- + (1-\tau_b)\pi + T - (1+\tau_c)c - (1+\gamma)a \tag{25}$$

$$0 = \pi - y^b + Wn^b + (R - 1 + \delta)k^b \tag{26}$$

$$0 = y^b - \theta^b (k^b)^\phi (n^b)^\nu \tag{27}$$

$$0 = \begin{cases} k^b - \chi a_- & \text{if } a_- \leq \text{borrowing or collateral constraint cutoffs} \\ \phi y^b / k^b - R + 1 - \delta & \text{otherwise} \end{cases} \tag{28}$$

$$0 = \nu y^b / n^b - W \tag{29}$$

$$0 = \lambda - RU'(c) - \chi U'(c)(\phi y^b / k^b - R + 1 - \delta)(1 - \tau_b) \tag{30}$$

$$0 = \begin{cases} \underline{a} - a & \text{if } a_- \leq \text{borrowing constraint cutoff} \\ \beta \lambda^e - U'(c) & \text{otherwise} \end{cases} \tag{31}$$

$$0 = n^b - n \tag{32}$$

[8]These expectations appear more formally as the input $E_{it}x_{i,t+1}$ in functions F^d and H. See equations (11)–(13).

$$0 = U(c) + \beta v^e - v \tag{33}$$

$$0 = d. \tag{34}$$

If the borrowing constraint binds, we replace the capital first-order condition and the Euler equation as noted above. If the collateral constraint binds, we replace only the capital first-order condition. Although the conditions are not listed above, we set all business activity to zero if $a_- = 0$ in order to be consistent with the collateral constraint.

4.1.4 Mapping F^w

The corresponding set of conditions are included in F^w, which includes residuals for the paid-employed workers. The states and elements of x, X, and x^e are the same for workers but the variables related to business are set to zero.

$$0 = Ra_- + (1-\tau_w)W\theta^w + T - (1+\tau_c)c - (1+\gamma)a \tag{35}$$

$$0 = \pi = y^b = k^b = n^b \tag{36}$$

$$0 = \lambda - RU'(c) \tag{37}$$

$$0 = \begin{cases} \underline{a} - a & \text{if } a_- \leq \text{borrowing constraint cutoff} \\ \beta\lambda^e - U'(c) & \text{otherwise} \end{cases} \tag{38}$$

$$0 = \theta^w + n \tag{39}$$

$$0 = U(c) - v + \beta v^e \tag{40}$$

$$0 = d - 1. \tag{41}$$

Note that the labor supply condition is written so that $\int n_i \, \mathrm{d}i$ is equal to the labor demand of business owners less the labor supply of workers, which is convenient for imposing labor market clearing.

4.1.5 Mapping G

Next, we construct the mapping G in (12). There are two types of arguments: the integrals of individual choices and the aggregate variables relevant for period-t decision-making. For convenience, we write the integrals of individual choices compactly. For example, in the case of consumption, we write $\int c$ instead of $\int c_i \,\mathrm{di}$. For our problem, we create the two vector inputs that we need—namely, the integrals of individual choices $\int x$=[$\int c$, $\int n$, $\int n^b$, $\int k^b$, $\int y^b$] and the aggregate variables $X = [R, T]$. With the inputs in $\int x$ and X, we start with some intermediary calculations. First, we set $N_c = -\int n$, which is consistent with the labor market clearing, and we set the marginal product of capital (MPK) equal to user cost of capital, $\frac{R-1+\delta}{1-\tau_p}$. The latter relation is consistent with firm maximization in a corporate sector that faces a tax τ_p on taxable corporate profits. With N_c and MPK, we can compute $K_c = (MPK/(\alpha\Theta))^{1/(\alpha-1)} N_c$, $Y_c = \Theta K_c^{\alpha} N_c^{1-\alpha}$, and $W = (1-\alpha)Y_c/N_c$. Next, we sum up tax revenues for consumption, $\tau_c \int c$; labor earnings, $\tau_w W(N_c + \int n^b)$; corporate profits, $\tau_p(Y_c - WN_c - \delta K_c)$; and profits of entrepreneurs, $\tau_b(\int y^b - (R-1+\delta)\int k^b - W\int n^b)$. With these intermediate calculations complete, we construct G in (12):

$$0 = 1 - (K_c + \textstyle\int k^b + B)/A_{-} \tag{42}$$

$$0 = 1 - (\mathrm{Tax} - (R-1-\gamma)B - T)/G, \tag{43}$$

where "Tax" is the sum of the tax revenues.

4.1.6 Mapping H

The Euler and Bellman equations appearing in F^b and F^w—specifically, equations (31), (33), (38), and (40)—have expectations of marginal values of wealth adjusted by the consumption tax and values that depend on next period choices if in paid employment or entrepreneurship. This relation necessitates having the additional H mapping in (13) when solving dynamic models with discrete choice—something that would not be included in the Bhandari et al. (2023) implementation. The conditions in (13) are as follows:

$$\lambda^e = p^w \lambda^w + (1-p^b)\lambda^b \tag{44}$$

$$v^e = v^w + \sigma_\eta \log(1 + \exp((v^b - v^w)/\sigma_\eta)), \tag{45}$$

where p^w is defined in (24) and $p^b = 1 - p^w$.

Once we have the mappings F^d, G, and H, the next step is to compute their derivatives with respect to arguments in x, X, and x^e. We use the automatic differentiation package, `ForwardDiff.jl`, in Julia to compute the derivatives. As we noted above, the coefficients of the linear system in (20) can be computed directly using the formulas in the appendix. The numerical recipes for efficient computation of those coefficients are in Bhandari et al. (2023). To solve for the paths $\{\hat{X}_t, \hat{X}_{t,t}\}$, we impose a large horizon and terminal conditions: $\hat{X}_T = 0$ and $\hat{X}_{T,T} = 0$. The code to implement this approach in our motivating application is available at https://github.com/apb-umn/BEM. We turn next to describe the results for this application.

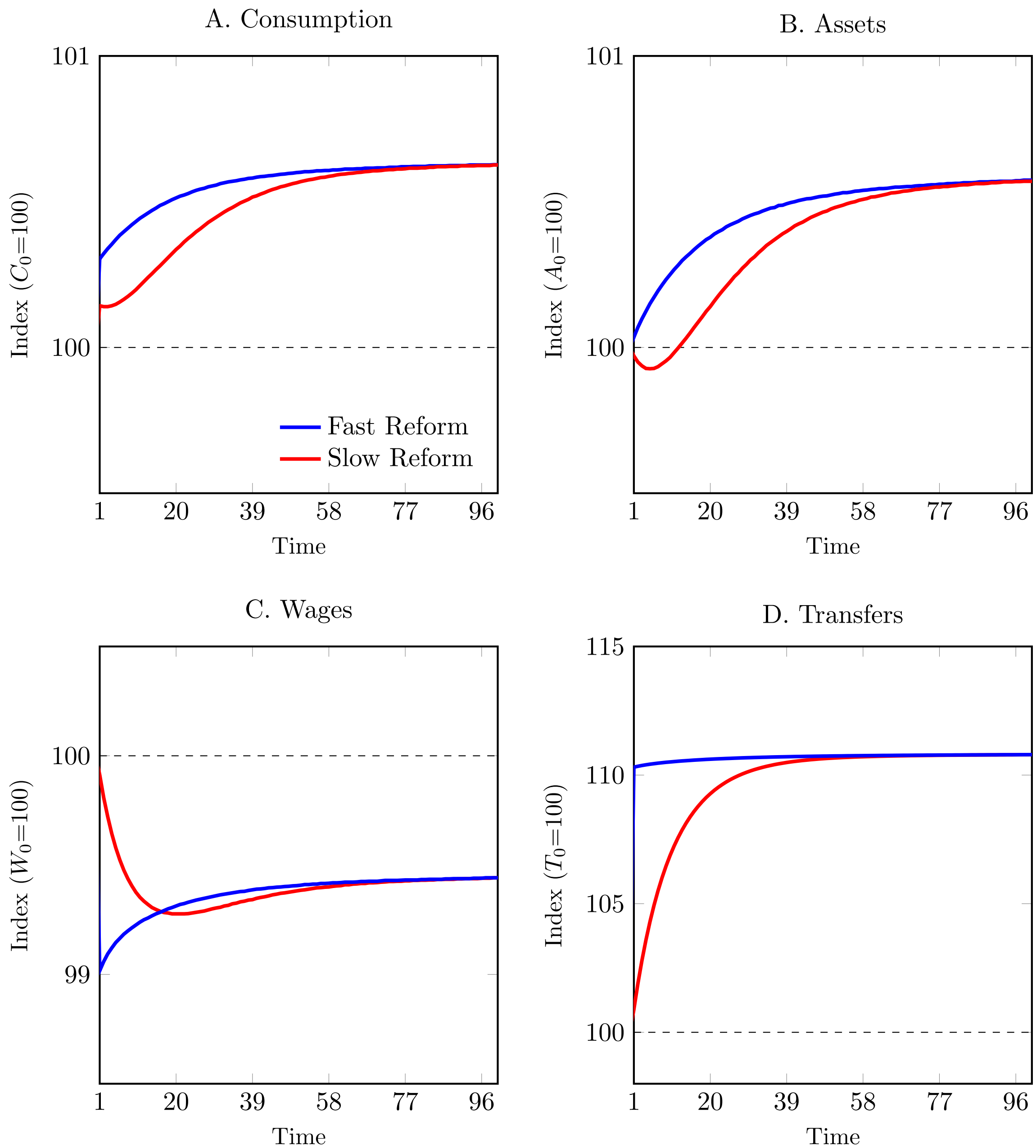
4.2 Results

In this section, we report on results for the motivating application in Section 2.

4.3 Calibration

We summarize here the main calibration choices and empirical targets, relegating detailed parameter values and data sources to Section F in the appendix. Our calibration aligns the model with major aggregates in the U.S. national accounts and government budgets. Output in the model corresponds to GDP, decomposed into production by privately-held businesses—sole proprietorships, partnerships, and S corporations—and a residual "corporate" sector that includes publicly-traded firms and nonbusiness entities. Preference parameters are chosen to match standard macroeconomic values for risk aversion, discounting, and trend growth. Production parameters imply roughly equal factor shares for private businesses and standard capital shares for corporate firms, with depreciation rates consistent with U.S. tangible capital estimates. Financing parameters are selected to match observed leverage and borrowing constraints, with the collateral limit calibrated to reproduce the empirical ratio of business loans to GDP. Government spending, taxation, and debt levels are set to replicate U.S. fiscal ratios. A key behavioral parameter—the Gumbel scale σ_η governing occupational sorting and labor demand elasticity—is disciplined by the estimated semi-elasticity of employment with respect to business tax rates from Giroud and Rauh (2019). Finally, idiosyncratic productivity processes for entrepreneurs and workers follow Markov chains consistent with estimates in Bhandari and McGrattan (2021). The detailed parameter values appear in Table 2, and the corresponding steady-state moments that align with U.S. data are reported in Table 3 in the appendix Section F.

Figure 1: Transition Paths for Aggregates



Notes: The transition paths follow an increase in the tax on business τ_b from the baseline 20 percent to 40 percent. The fast reform is implemented immediately, while the slow reform reaches the target with a persistence parameter of $\rho_\tau = 0.9$.

4.3.1 Transition Paths

We next compute transition paths following a change in business taxes, and we then compute the optimal rate. These results are compared for several recalibrations of the model with key parameters varied.

Our first reform increases the tax rate on business, τ_b, from 20 percent to 40 percent. Such a reform is in the spirit of enhanced enforcement: business owners in the United States are paying roughly half of what they owe to the Internal Revenue Service because of limited third-party reporting. Two experiments are considered, which differ in the timing. The first ("fast reform") assumes that the tax rate is changed immediately and the second ("slow reform") assumes that the higher rate is phased in according to:

$$\tau_{bt+1} = \rho_\tau \tau_{bt} + (1 - \rho_\tau)\tau_b^*, \tag{46}$$

where τ_b^* is the higher rate in the final steady state. In each case, we keep the public debt fixed and adjust the lump-sum transfers T_t every period to ensure the government budget constraint in (43) holds with equality.

We begin by assessing the accuracy of our second-order approximation. Specifically, we take the approximated path of X_t and compute the residuals for the aggregate equilibrium conditions in (42) and (43). To do so, we globally solve the individual problems backward in time given the path of X_t and then propagate the distribution forward from the initial steady state. Using the resulting policy functions and distributions, we recompute aggregates and evaluate the residuals of the two equilibrium conditions for each date. For ease of interpretation, we scale the asset market condition by the asset-to-GDP ratio (A/Y) and the government budget constraint by the government-expenditure share (G/Y), and compute the residuals as percentages of output. If we consider the asset market clearing condition over the first 100 years following an immediate tax change, we find an average residual equal to 0.19 percent of output, with the maximum residual in absolute value equal to 0.29 percent of output. If we consider the government budget constraint, we find an average residual equal to 0.01 percent of output, with the maximum residual in absolute value equal to 0.04 percent of output.

In Figure 1, we plot asset holdings, consumption, wages, and transfers, as a percent of their pre-reform values. For the slow reform, we set ρ_τ equal to 0.9. We see from the upper left panel of the figure that consumption is 0.6 percent higher. In the upper right panel, we see a similar rise in assets, which are near the new steady state in roughly 60 periods. We see from the lower panels that a big part of the story is a large increase in transfers. Wages do fall as the tax rate on private business rises and demand for labor in the sector falls. But the fact that transfers rise so sharply—even in the case of the slower reform—implies that the individuals are able to ensure a smooth path of consumption during the transition.

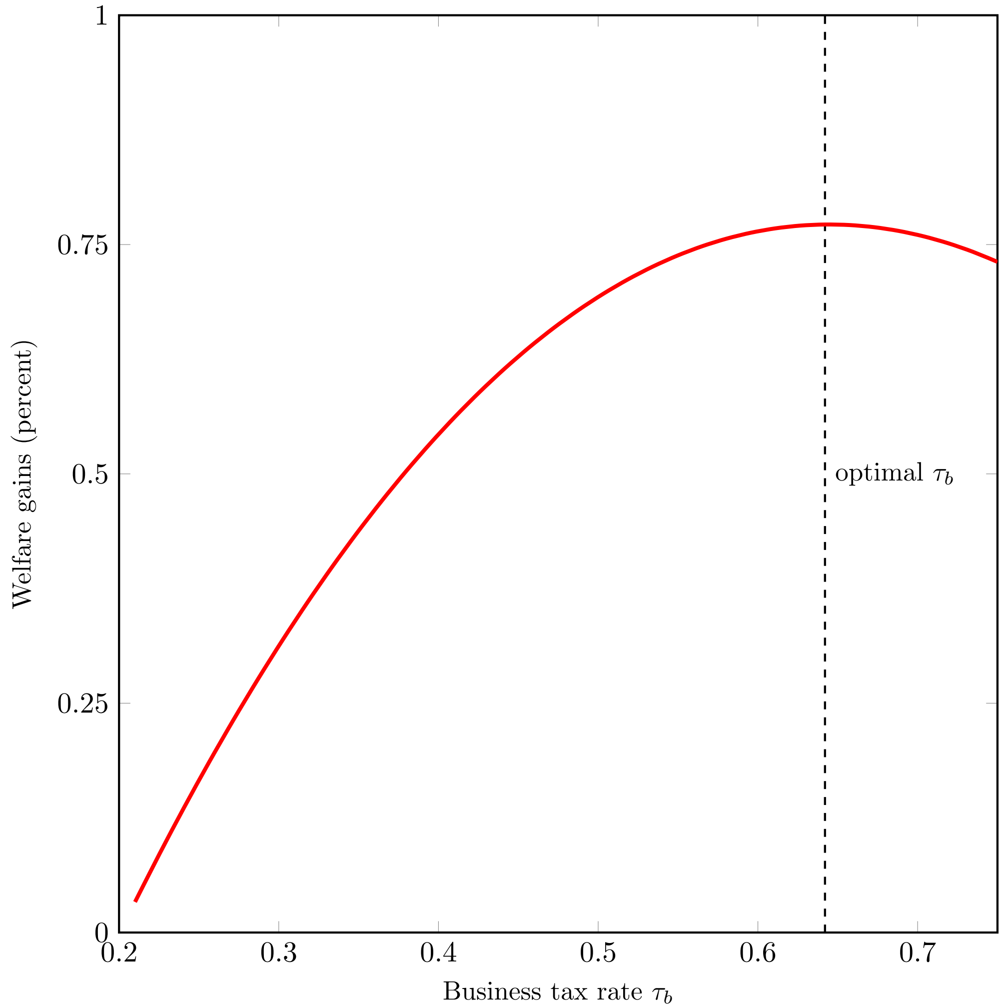
For private business, the plots of economic activity are the mirror image of the path of transfers: we find a sharp decline in profits and the share of individuals choosing to run a business with similar timing to the rise in transfers. Profits in the new steady state fall 11.5 percent and the fraction of individuals running businesses falls -5.4 percent. Along the transition, the average productivity of the private business sector improves because the individuals with lower θ_t^b do not enter when τ_{bt} is increased. These efficiency gains make it possible to finance higher average consumption. It is important to note, however, that we have restricted our analysis in key ways—for example, by assuming inelastic labor supplies and by assuming that owners do not make investments in business capital. We restrict the environment in this way to highlight the impact of higher taxation on the discrete occupational choice.

4.3.2 Optimal Taxation

In this section, we compute the optimal one-time permanent reform for a utilitarian social planner who maximizes the average welfare of individuals in the initial steady state. As in the exercises in the previous sections, pursuant to the reform, transfers are adjusted each period to ensure that the government budget constraint holds with equality and the public debt is unchanged. It is fairly straightforward to extend the analysis to consider reforms richer than one-time changes in τ_b, but we leave this to future work.

We find an optimal tax on business owners at 64.2 percent. In Figure 2, we plot the consumption-equivalent welfare gain, computed as the per-period percentage change in consumption relative to the intitial allocation required for individuals to be indifferent between the status quo and higher taxes. At the optimum, the gain is 0.8 percent. Our optimal value for τ_b is significantly higher than the effective 20 percent tax rate currently paid by business owners in the United States. This

Figure 2: Consumption-Equivalent Welfare Gains



Notes: The welfare gains are computed as the per-period percentage change in consumption relative to the initial allocation required for individuals to be indifferent between having the policy reform and not having it. The dashed vertical line marks the optimal business tax rate.

reflects the presence of several, potentially conflicting, motives for taxing business owners in our model. We begin by reviewing the main forces identified in related environments and then explain how these forces manifest in our setting.

In Aiyagari-type models focused on paid employees and calibrated to U.S. data, the typical finding is that observed transfer levels are too low from a utilitarian planner's perspective. Optimal reforms in such models involve increasing transfers financed by more progressive taxes, which improve welfare through better insurance and redistribution, albeit at the cost of lower output.

See, for example, Aiyagari and McGrattan (1998), Heathcote, Storesletten, and Violante (2017), Ferrière et al. (2023), and Bhandari et al. (2025). In contrast, models where income stems from entrepreneurial profits sometimes recommend even negative profit taxes. For example, Guvenen et al. (2023) show that profits can reflect distortions such as borrowing constraints or wedges induced by size-dependent markups, as in Boar and Midrigan (forthcoming). In both cases, productive firms are inefficiently small, and a subsidy on profits can improve efficiency by reallocating resources to more productive firms.

Our model includes both types of income—labor and business—and is calibrated to U.S. data in terms of risk, income shares, and borrowing behavior. The finding that the optimal tax on business owners is high is driven primarily by efficiency considerations. First, weak enforcement lowers effective tax rates on business income, allowing low-productivity entrepreneurs to remain in the private sector; taxing them more heavily induces reallocation toward the more productive corporate sector. Second, taxing highly productive owners is efficient because their extensive-margin responses are relatively inelastic and their income derives from a fixed factor. Finally, while taxing business income may worsen capital misallocation, such losses are minor given the relatively small income share in GDP and limited borrowing of private business owners. Taken together, these factors imply that a higher tax on business owners is optimal.[9]

If we compare the transition dynamics in the case that τ_b is set optimally, the figure has the same patterns as Figure 1, although not surprisingly, we have larger responses. Consumption, assets, and transfers are all higher in the end, rising by 1.6, 1.8, and 21.2 percent, respectively, and wages are 0.9 percent lower. For business owners, the decline in profits and the share of owners is 25.8 percent and 12.2 percent, respectively.

4.3.3 Sensitivity Analysis

In this section, we compare the baseline results to those of four alternative economies with tighter collateral constraints (χ =1.05), looser collateral constraints (χ =2), a higher semi-elasticity of aggregate labor demand to an increase in τ_b, and higher income risk for both wage earners and business owners. In case of higher income risk, we set the standard deviations of log productivities

[9] Our results are related to those of Brüggemann (2021) and Imrohoroglu et al. (2023), who study optimal taxation of high earners—both in entrepreneurship and paid employment—and also find high optimal rates, around 60 percent.

Table 1: Sensitivity Analysis

Economy	Percent Change in: C	A	W	T	$\#_b$	π	Optimal τ_b%	Welfare Gain%
Baseline	1.6	1.8	−0.9	21.2	−12.2	−25.8	64.2	0.77
Tighter collateral	1.8	2.4	−0.5	22.9	−12.3	−25.6	64.4	0.87
Looser collateral	1.3	1.4	−1.3	19.8	−11.7	−24.6	61.6	0.68
Higher tax elasticity	2.9	4.9	−0.6	14.9	−49.6	−58.5	53.2	1.51
Higher income risk	0.3	−0.1	−1.1	12.9	−6.0	−17.1	56.0	0.39

Notes: C = consumption; A= assets; W= wage; T= transfers; $\#_b$= fraction of business owners; π= profit of business owners; τ_b= tax rate on business. Parameters for the baseline economy are described in Section F. For the "tighter collateral" economy, we set χ =1.05. For the "looser collateral" economy, we set χ =2. For the "higher tax elasticity" economy, we set σ_η =0.075. For the "higher income risk" economy, we multiply elements of $\ln\theta_t^w$ and $\ln\theta_t^b$ by a factor of 1.5. All economies were re-calibrated so that they match the national income and product accounts and government budgets shown in Table 3.

equal to 1.5 times the baseline estimates. In all cases, we re-calibrate the parameters of the model to ensure that the national income and product accounts and government budget constraint are in line with U.S. accounts before we recompute the transition dynamics and welfare gains.

The results of these exercises are shown in Table 1. The first six columns of the table report the percent changes in the variables of interest when comparing the final steady state and the initial values. The final columns are the optimum tax rate on business profits and the consumption-equivalent welfare gain. Results for the baseline economy are reported in the first row of the table. We can compare this to the next two rows that vary the tightness of the collateral constraint. As the table shows, varying the tightness of the constraint does little to change our main results—qualitatively or quantitatively. This is perhaps surprising given the attention paid to the financing constraints of business owners. However, this result can be understood if we take into account that business profits are a much smaller share of GDI than wage earnings, and thus changing conditions on their financing cannot affect aggregates that much.

For the next experiment (labelled "higher tax elasticity") in Table 1, we reduced the Gumbel scale parameter σ_η from 0.40 to 0.075 to generate a tax semi-elasticity of labor demand equal to −1. For this case, we find larger responses in consumption and assets and a smaller decline in wages—which implies a larger welfare gain. Not surprisingly, we find a much larger drop in the fraction of business owners and a lower optimal tax rate when compared to the baseline that has a lower tax elasticity (in absolute value).

Results for the experiment with higher income risk are reported in the last row of Table 1. For this case, we find smaller changes in aggregate variables when compared to the baseline case, with the exception of wages. The tax rate achieving the optimum in this case is 56.0—lower than the baseline but still well above the tax rate levied on wages and salaries. The welfare gain in this case is 0.39. With higher income risk, there are more owners in the right tail of the productivity distribution with higher desired capital stock. The collateral constraint binds more frequently for these owners, and the planner lowers the optimal tax rate so that they can retain more of their profits to finance investment.

Overall, the results provide similar guidance with respect to business taxation, namely, that raising the tax rate on business owners well above effective U.S. levels is welfare-enhancing for the utilitarian planner.

5 Conclusion

In this paper, we showed how to extend perturbation methods to approximate transition dynamics in environments with discrete choice. We applied the methods to a general equilibrium model with occupational choice, heterogeneity in assets and productivity, and financial constraints. While we restricted attention to policy impacts on occupational choice, the methods are well suited for exploring much richer environments than the specific application studied here.

References

Ahn, SeHyoun, Greg Kaplan, Benjamin Moll, Thomas Winberry, and Christian Wolf. 2017. "When Inequality Matters for Macro and Macro Matters for Inequality." *NBER Macroeconomics Annual* 32:1–75.

Aiyagari, S Rao. 1994. "Uninsured Idiosyncratic Risk and Aggregate Saving." *Quarterly Journal of Economics* 109 (3):659–684.

Aiyagari, S. Rao and Ellen R. McGrattan. 1998. "The Optimum Quantity of Debt." *Journal of Monetary Economics* 42 (3):447–467.

Alvarez, Fernando, Francesco Lippi, and Panagiotis Souganidis. 2023. "Price Setting with Strategic Complementarities as a Mean Field Game." *Econometrica* 91 (6):2005–2039.

Auclert, Adrian, Bence Bardóczy, Matthew Rognlie, and Ludwig Straub. 2021. "Using the Sequence-Space Jacobian to Solve and Estimate Heterogeneous-agent Models." *Econometrica* 89 (5):2375–2408.

Bardóczy, Bence. 2021. "Spousal Insurance and the Amplification of Business Cycles." URL https://www.bencebardoczy.com/files/cv_bardoczy.pdf. Working paper, Northwestern University.

Bayer, Christian, Benjamin Born, and Ralph Luetticke. 2022. "Shocks, Frictions, and Inequality in US Business Cycles." Discussion Paper 14364, Centre for Economic Policy Research.

Bhandari, Anmol, Thomas Bourany, David Evans, and Mikhail Golosov. 2023. "A Perturbational Approach for Approximating Heterogeneous Agent Models." Working Paper 31744, National Bureau of Economic Research.

Bhandari, Anmol, David Evans, Mikhail Golosov, and Thomas Sargent. 2025. "Efficiency, Insurance, and Redistribution Effects of Government Policies." Tech. rep., University of Minnesota, University of Oregon, University of Chicago, NYU.

Bhandari, Anmol and Ellen R. McGrattan. 2021. "Sweat Equity in U.S. Private Business." *Quarterly Journal of Economics* 136 (2):727–781.

Bilal, Adrien. 2023. "Solving Heterogeneous Agent Models with the Master Equation." Working Paper 31103, National Bureau of Economic Research.

Boar, Corina and Virgiliu Midrigan. forthcoming. "Markups and Inequality." *Review of Economic Studies* .

Boppart, Timo, Per Krusell, and Kurt Mitman. 2018. "Exploiting MIT Shocks in Heterogeneous-agent Economies: The Impulse Response as a Numerical Derivative." *Journal of Economic Dynamics and Control* 89:68–92.

Brüggemann, Bettina. 2021. "Higher Taxes at the Top: The Role of Entrepreneurs." *American Economic Journal: Macroeconomics* 13 (3):1–36.

Buera, Francisco, Joseph Kaboski, and Yongseok Shin. 2011. "Finance and Development: A Tale of Two Sectors." *American Economic Review* 101 (5):1964–2002.

Cagetti, Marco and Mariacristina De Nardi. 2006. "Entrepreneurship, Frictions, and Wealth." *Journal of Political Economy* 114 (5):835–870.

Catherine, Sylvain. 2022. "Keeping Options Open: What Motivates Entrepreneurs?" *Journal of Financial Economics* 144 (1):1–21.

Childers, David. 2018. "On the Solution and Application of Rational Expectations Models with Function Valued States." Working paper, Carnegie Mellon. URL https://donskerclass.github.io/files/pdf/Function_Valued_States_1-15-2018.pdf.

Druedahl, Jeppe. 2021. "A Guide on Solving Non-convex Consumption-Saving Models." *Computational Economics* 58 (3):747–775.

Evans, David S. and Boyan Jovanovic. 1989. "An Estimated Model of Entrepreneurial Choice Under Liquidity Constraints." *Journal of Political Economy* 97 (4):808–827.

Ferrière, Axelle, Philipp Grübener, Gaston Navarro, and Oliko Vardishvili. 2023. "On the Optimal Design of Transfers and Income Tax Progressivity." *Journal of Political Economy Macroeconomics* 1 (2):276–333.

Giroud, Xavier and Joshua Rauh. 2019. "State Taxation and the Reallocation of Business Activity: Evidence from Establishment-Level Data." *Journal of Political Economy* 127 (3):1262–1316.

Gornemann, Nils, Keith Kuester, and Makoto Nakajima. 2021. "Doves for the Rich, Hawks for the Poor? Distributional Consequences of Systematic Monetary Policy." Discussion Paper 89, ECONtribute.

Guvenen, Fatih, Gueorgui Kambourov, Burhan Kuruscu, Sergio Ocampo, and Daphne Chen. 2023. "Use It or Lose It: Efficiency and Redistributional Effects of Wealth Taxation." *Quarterly Journal of Economics* 138 (2):835–894.

Hamilton, Barton H. 2000. "Does Entrepreneurship Pay? An Empirical Analysis of the Returns to Self-Employment." *Journal of Political Economy* 108 (3):604–631.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante. 2017. "Optimal Tax Progressivity: An Analytical Framework." *Quarterly Journal of Economics* 132 (4):1693–1754.

Imrohoroglu, Ayse, Cagri Kumru, Jiu Lian, and Arm Nakornthab. 2023. "Revisiting Taxes on High Incomes." *Review of Economic Dynamics* 51:1159–1184.

Kaplan, Greg and Giovanni L Violante. 2018. "Microeconomic Heterogeneity and Macroeconomic Shocks." *Journal of Economic Perspectives* 32 (3):167–194.

Lucas, Robert E. 1978. "On the Size Distribution of Business Firms." *Bell Journal of Economics* 9 (2):508–523.

Midrigan, Virgiliu and Daniel Yi Xu. 2014. "Finance and Misallocation: Evidence from Plant-level Data." *American Economic Review* 104 (2):422–458.

Moll, Benjamin, Yves Achdou, Jiequn Han, Jean-Michel Lasry, and Pierre-Louis Lions. 2022. "Income and Wealth Distribution in Macroeconomics." *Review of Economic Studies* 89 (1):227–262.

Quadrini, Vincenzo. 2000. "Entrepreneurship, Saving, and Social Mobility." *Review of Economic Dynamics* 3 (1):1–40.

Reiter, Michael. 2009. "Solving Heterogeneous-agent Models by Projection and Perturbation." *Journal of Economic Dynamics and Control* 33 (3):649–665.

———. 2023. "State Reduction and Second-Order Perturbations of Heterogeneous Agent Models." Working Paper 49, Institute for Advanced Studies.

Winberry, Thomas. 2018. "A Method for Solving and Estimating Heterogeneous Agent Macro Models." *Quantitative Economics* 9 (3):1123–1151.

A State Space Representation

Written in state space representation with aggregate state $Z = [A, \Omega]$, the equilibrium conditions can be expressed as

$$0 = F^d\Big(a, \theta, \bar{x}^d(a, \theta, Z), \mathbb{E}^d[\bar{x}|a, \theta, Z], \bar{Y}(Z)\Big), \text{ for all } a, \theta, Z \tag{47}$$

$$0 = G\Big(\int \sum_d \bar{x}^d(a, \theta, Z)\bar{p}^d(a, \theta, Z)\, \mathrm{d}\Omega(\mathrm{a}, \theta), \bar{Y}(Z)\Big), \text{ for all } Z, \tag{48}$$

$$\mathbb{E}^d[\bar{x}|a, \theta, Z] = \int H\Big(\{\bar{x}^{d'}\left(a^d(a, \theta, Z), \rho_\theta \theta + \epsilon, \bar{Z}(Z)\right)\}_{d'}\Big)\mu(\epsilon)d\epsilon, \text{ for all } a, \theta, Z, \tag{49}$$

where F^d includes the individual optimality conditions for discrete choice d; $\mu(\epsilon)$ is the transition density for the idiosyncratic shock process; and $\bar{Z}(Z) = [\mathsf{P}\bar{X}(Z), \bar{\Omega}(Z)]$ is the law of motion for the aggregate state that depends on $\bar{A}(Z) = \mathsf{P}\bar{X}(Z) \in \bar{X}(Z)$ and the law of motion for the distribution.

$$\bar{\Omega}(Z)\langle a', \theta' \rangle = \iiint \sum_d \iota(\bar{a}^d(a, \theta, Z) \leq a')\iota(\rho_\theta \theta + \epsilon \leq \theta')\bar{p}^d\,(a, \theta, Z))\,\mu(\epsilon)d\epsilon d\Omega(a, \theta). \tag{50}$$

The variable $\bar{p}^d(a, \theta, Z)$ captures the probability that an agent in idiosyncratic state (a, θ) chooses the occupational choice o. Without loss of generality, we will assume there are two possible occupational choices, w and b. The probability that the agent chooses one of these occupations will depend on the cutoff $\bar{\kappa}(a, \theta, Z) = \bar{v}^b(a, \theta, Z) - \bar{v}^w(a, \theta, Z)$ and the complement of the cumulative distribution function of the taste shock, Γ,

$$p^d(a, \theta, Z) = \begin{cases} \Gamma(\bar{\kappa}(a, \theta, Z)) & \text{if } d = w \\ 1 - \Gamma(\bar{\kappa}(a, \theta, Z)) & \text{otherwise.} \end{cases}$$

B Zeroth Order Approximation

We approximate equilibrium responses around the steady state of the deterministic economy after the reform. We let Ω^* be the invariant distribution in such economy. We denote this steady state by $Z^* = [A^*, \Omega^*]^{\mathrm{T}}$. Let $\overline{\Lambda}^d(a', \theta', a, \theta)$ be the transition kernel from (a, θ) to (a', θ') in the steady state for agents with occupational choice d. Our perturbation is only with respect to the initial state $Z_0 = Z^* + \sigma(Z_{-1} - Z^*)$. Thus, the policy rules will not explicitly depend on σ since all of the dependence is captured through the aggregate state.

We drop explicit dependence of policy functions on Z when $Z = Z^*$. Thus, for example, $\overline{x}\,(a,\theta)$ will be understood as $\overline{x}\,(a,\theta,Z^*)$. We use $\overline{X}_Z$, $\overline{X}_{ZZ}$, $\overline{Z}_Z$, $\overline{x}_Z(a,\theta)$, and so on, to denote Frechet derivatives of policy functions of various orders. In heterogeneous-agent settings, these are infinite dimensional linear operators rather than the finite dimensional matrices that would be analyzed in a standard representative-agent setting, but their analytical properties used in the proofs are largely unaffected by this distinction. For further details on the necessary assumptions and how we handle kinked policy rules, see Bhandari et al. (2023).

Finally we will use $\mathsf{F}^d_x(a,\theta)$ (and similar notation for other derivatives) to denote derivatives of the (47) evaluated at the long-run steady state. We will use $\mathsf{H}_{x^d}(a,\theta)$ (and similar notation for other derivatives) to denote derivatives of the (49) with respect to the third argument evaluated at the long-run steady state. Finally we will use G_x and G_Y to denote the derivatives of the (48) with respect to the first and second arguments evaluated at the long-run steady state.

C First Order Approximation

Let $X_t(\sigma)$ represent the path of aggregate variables as a function of σ, which scales the deviation of the initial non-stochastic steady state from the long-run steady state:

$$Z_0 = Z^* + \sigma(Z_{-1} - Z^*).$$

Setting $\sigma = 0$ corresponds to the economy being initially at the long-run steady state, while $\sigma = 1$ corresponds to the pre-reform economy.

We define a set of directions using the Frechet derivatives of X and the law of motion of the aggregate state. We let $\hat{Z}_0 = Z_{-1} - Z^*$ and define $\hat{Z}_t$ recursively as $\hat{Z}_t = \bar{Z}_Z \hat{Z}_{t-1}$. We then let $\hat{X}_t = \bar{X}_Z \cdot \hat{Z}_t$, where $\bar{X}_Z$ is the Frechet derivative of X with respect to the aggregate state. Our first result characterizes the first-order response of aggregates given these directional derivatives.

Lemma 1. *To first order, X_t is approximated by*

$$X_t = \bar{X} + \hat{X}_t + o(\|Z_{-1} - Z^*\|).$$

Proof. By definition, X_t solves the following recursive system:

$$
\begin{aligned}
X_t(\sigma) &= \bar{X}(Z_t) \\
Z_t(\sigma) &= \bar{Z}(Z_{t-1}) \text{ for } t > 0,
\end{aligned} \tag{51}
$$

with $Z_0(\sigma) = Z^* + \sigma(Z_{-1} - Z^*)$. Differentiating the both equations with respect to σ yields

$$
\begin{aligned}
\bar{X}_{t,\sigma} &= \bar{X}_Z \cdot \bar{Z}_{t,\sigma} \\
\bar{Z}_{t,\sigma} &= \bar{Z}_Z \cdot \bar{Z}_{t-1,\sigma} \text{for } t > 0,
\end{aligned}
$$

with $\bar{Z}_{0,\sigma} = (Z^*_{-1} - Z^*) = \hat{Z}_0$. Using the definition of $\hat{Z}_t$ and $\hat{X}_t$, we have that $\bar{Z}_{t,\sigma} = \hat{Z}_t$ which implies that $\bar{X}_{t,\sigma} = \bar{X}_Z \cdot \hat{Z}_t = \hat{X}_t$, which yields the desired result. □

In order to construct the first-order approximation we must find these directional derivatives $\hat{X}_t$. Our main result is that these derivatives solve the linear system (19), which is summarized in Proposition 1 below. To derive that result, we proceed by differentiating the system of equations (47)–(50) to characterize how individual policies and, as a direct result, the law of motion of the distribution responds to the changes in the aggregate state.

C.1 The F^d equation

We begin by differentiating equations (47) and (49) in direction $\hat{Z}_t$. Differentiating equation (47) yields

$$
\mathsf{F}^d_x(a,\theta)\hat{x}_t(a,\theta) + \mathsf{F}^d_{x'}(a,\theta)\left(\mathbb{E}^d[\bar{x}|a,\theta]\right)_Z \cdot \hat{Z}_t + \mathsf{F}^d_Y(a,\theta)\hat{Y}_t = 0,
$$

where $\hat{x}_t(a,\theta) = \bar{x}_Z(a,\theta)\cdot\hat{Z}_t$. The change in expectations $\left(\mathbb{E}^d[\bar{x}|a,\theta]\right)_Z\cdot\hat{Z}_t$ is found by differentiating (49) in direction $\hat{Z}_t$,

$$
\begin{aligned}
\left(\mathbb{E}^d[\bar{x}|a,\theta]\right)_Z \cdot \hat{Z}_t = &\int \sum_{d'} \mathsf{H}_{x^{d'}}\left(\bar{a}^d(a,\theta), \rho_\theta\theta + \epsilon\right) \\
&\cdot \left(\hat{x}^{d'}_{t+1}(\bar{a}^d(a,\theta), \rho_\theta\theta + \epsilon) + \bar{x}^{d'}_a(\bar{a}^d(a,\theta), \rho_\theta\theta + \epsilon)\mathsf{p}\hat{x}^d_t(a,\theta)\right)\mu(\epsilon)d\epsilon \\
= &\bar{\mathbb{E}}^d\left[\sum_{d'} \mathsf{H}_{x^{d'}}(\cdot,\cdot)\hat{x}^{d'}_{t+1}(\cdot,\cdot) + \mathsf{H}_{x^{d'}}(\cdot,\cdot)\bar{x}^{d'}_a(\cdot,\cdot)\mathsf{p}\hat{x}^d_t(a,\theta)\big|a,\theta\right]
\end{aligned}
$$

$$= \bar{\mathbb{E}}^d \left[\sum_{d'} \mathsf{H}_{x^{d'}}(\cdot,\cdot)\hat{x}_{t+1}^{d'}(\cdot,\cdot) \Big| a,\theta \right] + \bar{x}_a^{d,e}(a,\theta)\mathsf{p}\hat{x}_t^d(a,\theta),$$

where

$$\bar{\mathbb{E}}^d[y(\cdot,\cdot)|a,\theta] \equiv \int y(\bar{a}^d(a,\theta), \rho_\theta\theta + \epsilon)\mu(\epsilon)d\epsilon$$

and

$$\bar{x}_a^{d,e}(a,\theta) \equiv \bar{\mathbb{E}}^d \left[\sum_{d'} \mathsf{H}_{x^{d'}}(\cdot,\cdot)\bar{x}_a^{d'}(\cdot,\cdot) \big| a,\theta \right]$$

represents taking the expectation over the steady-state policy rules and idiosyncratic shocks.

Combining these two derivatives yields the following lemma:

Lemma 2. *For any t,*

$$\hat{x}_t^d(a,\theta) = \sum_{s=0}^{\infty} \mathsf{x}_s^d(a,\theta)\hat{Y}_{t+s}, \tag{52}$$

where $\mathsf{x}_s^d(a,\theta)$ satisfies the following recursive equations

$$\mathsf{x}_0^d(a,\theta) = -\left(\mathsf{F}_x^d(a,\theta) + \mathsf{F}_{x'}^d(a,\theta)\bar{x}_a^{d,e}(a,\theta)\mathsf{p}\right)^{-1} \mathsf{F}_Y^d(a,\theta) \tag{53}$$

$$\mathsf{x}_{s+1}^d(a,\theta) = -\left(\mathsf{F}_x^d(a,\theta) + \mathsf{F}_{x'}^d(a,\theta)\bar{x}_a^{d,e}(a,\theta)\mathsf{p}\right)^{-1} \mathsf{F}_{x'}^d(a,\theta)\mathsf{x}_s^{d,e}(a,\theta) \tag{54}$$

$$\mathsf{x}_s^{d,e}(a,\theta) = \bar{\mathbb{E}}^d \left[\sum_{d'} \mathsf{H}_{x^{d'}}(\cdot,\cdot)\mathsf{x}_s^{d'}(\cdot,\cdot)|a,\theta \right]. \tag{55}$$

Proof. We will show that (52) solves the recursive system above. Substituting for $\hat{x}_{t+1}^d$, we have that

$$\begin{aligned}
\mathsf{F}_{x'}^d(a,\theta)\left(\mathbb{E}^d[\bar{x}|a,\theta]\right)_Z \cdot \hat{Z}_t &= \sum_{s=0}^{\infty} \mathsf{F}_{x'}^d(a,\theta)\bar{\mathbb{E}}^d \left[\sum_{d'} \mathsf{H}_{x^{d'}}(\cdot,\cdot)\mathsf{x}_{s+1}^{d'}(\cdot,\cdot)|a,\theta \right] \hat{Y}_{t+s+1} \\
&\qquad + \mathsf{F}_{x'}^d(a,\theta)\bar{x}_a^{d,e}(a,\theta)\mathsf{p}\hat{x}_t^d(a,\theta) \\
&= \sum_{s=0}^{\infty} \mathsf{F}_{x'}^d(a,\theta)\mathsf{x}_s^{d,e}(a,\theta)\hat{Y}_{t+s+1} \\
&\qquad + \mathsf{F}_{x'}^d(a,\theta)\bar{x}_a^{d,e}(a,\theta)\mathsf{p}\hat{x}_t^d(a,\theta) \\
&= -\left(\mathsf{F}_x^d(a,\theta) + \mathsf{F}_{x'}^d(a,\theta)\bar{x}_a^{d,e}(a,\theta)\mathsf{p}\right) \sum_{s=0}^{\infty} \mathsf{x}_{s+1}^d(a,\theta)\hat{Y}_{t+s+1} \\
&\qquad + \mathsf{F}_{x'}^d(a,\theta)\bar{x}_a^{d,e}(a,\theta)\mathsf{p}\hat{x}_t^d(a,\theta)
\end{aligned}$$

$$
\begin{aligned}
&= -\mathsf{F}_x^d(a,\theta)\hat{x}_t^d(a,\theta) + \left(\mathsf{F}_x^d(a,\theta) + \mathsf{F}_{x'}^d(a,\theta)\bar{x}_a^{d,e}(a,\theta)\mathsf{p}\right)\mathsf{x}_0^d(a,\theta)\hat{Y}_t \\
&= -\mathsf{F}_x^d(a,\theta)\hat{x}_t^d(a,\theta) - \mathsf{F}_Y^d(a,\theta)\hat{Y}_t,
\end{aligned}
$$

where the third equality follows from $\hat{x}_t^d(a,\theta) = \mathsf{x}_0^d(a,\theta)\hat{Y}_t + \sum_{s=0}^{\infty}\mathsf{x}_{s+1}^d(a,\theta)\hat{Y}_{t+s+1}$. This implies

$$
\mathsf{F}_x^d(a,\theta)\hat{x}_t(a,\theta) + \mathsf{F}_{x'}^d(a,\theta)\left(\mathbb{E}^d[\bar{x}|a,\theta]\right)_Z \cdot \hat{Z}_t + \mathsf{F}_Y^d(a,\theta)\hat{Y}_t = 0,
$$

which completes the proof. □

This lemma provides a simple way to construct a response of individual policy rules to changes in aggregates. The coefficients $\mathsf{x}_s^d(a,\theta)$ have the economic interpretation that they are $\frac{\partial x_t}{\partial Y_{t+s}}$, for example, the change in individual policy rules today to the expected change in aggregates s periods in the future.

C.2 The Law of Motion of Ω

Our next step is to determine how individual policies aggregate into changes in the distribution. To achieve this, we differentiate equation (50) in direction $\hat{Z}_t$ to obtain an expression for $\hat{\Omega}_{t+1}\langle a,\theta\rangle = \bar{\Omega}_Z \cdot \hat{Z}_t\langle a,\theta\rangle$:

$$
\begin{aligned}
\frac{\partial^2}{\partial a'\partial\theta'}\hat{\Omega}_{t+1}\langle a',\theta'\rangle = &\int\sum_d \delta\left(a' - \bar{a}^d(a,\theta)\right)\mu\left(\theta' - \rho_\theta\theta\right)\bar{p}^d(a,\theta)d\hat{\Omega}_t(a,\theta) \\
&+ \int\sum_d \delta\left(a' - \bar{a}^d(a,\theta)\right)\mu\left(\theta' - \rho_\theta\theta\right)\hat{p}_t^d(a,\theta)d\Omega^*(a,\theta) \\
&- \int\sum_d \frac{\partial}{\partial a'}\left(\delta\left(a' - \bar{a}^d(a,\theta)\right)\right)\mu\left(\theta' - \rho_\theta\theta\right)\bar{p}^d(a,\theta)\hat{a}_t^d(a,\theta)d\Omega^*(a,\theta),
\end{aligned} \tag{56}
$$

where

$$
\begin{aligned}
\hat{p}_t^d(a,\theta) = \bar{p}_Z^d(a,\theta)\cdot\hat{Z}_t = &\begin{cases} \Gamma'(\bar{\kappa}(a,\theta))\hat{\kappa}_t(a,\theta) & \text{if } d = w \\ -\Gamma'(\bar{\kappa}(a,\theta))\hat{\kappa}_t(a,\theta) & \text{if } d = b \end{cases} \\
\equiv &\ \bar{p}_\kappa^d(a,\theta)\hat{\kappa}_t(a,\theta)
\end{aligned}
$$

and $\hat{\kappa}_t(a,\theta) = \mathsf{p}_v \hat{x}_t^b(a,\theta) - \mathsf{p}_v \hat{x}_t^w(a,\theta)$, with p_v being the selector matrix such that $\bar{v}(a,\theta,Z) = \mathsf{p}_v \bar{x}(a,\theta,Z)$. Also, $\hat{a}_t^d(a,\theta) = \mathsf{p}\hat{x}_t^d(a,\theta)$, where p is the selector matrix such that $\bar{a}^d(a,\theta,Z) = \mathsf{p}\bar{x}^d(a,\theta,Z)$.

Here and below, we define operators for the law of motion of the distribution. The first set of operators translate the distribution over time:

$$\mathcal{L} \cdot \mathsf{y}\langle a',\theta'\rangle = \iint \sum_d \overline{\Lambda}^d(a',\theta',a,\theta)\mathsf{y}(a,\theta)dad\theta \tag{57}$$

$$\mathcal{L}^{(a)} \cdot \mathsf{y}\langle a',\theta'\rangle = \iint \sum_d \overline{\Lambda}^d(a',\theta',a,\theta)\bar{a}_a^d(a,\theta)\mathsf{y}(a,\theta)dad\theta \tag{58}$$

$$\mathcal{L}_{(a)} \cdot \mathsf{y}\langle a',\theta'\rangle = \iint \sum_d \overline{\Lambda}^d(a',\theta',a,\theta)\frac{\bar{p}_a^d(a,\theta)}{\bar{p}^d(a,\theta)}\mathsf{y}(a,\theta)dad\theta, \tag{59}$$

where $\overline{\Lambda}^d(a',\theta',a,\theta) = \delta(a' - \bar{a}^d(a,\theta))\mu(\theta' - \rho_\theta\theta)\bar{p}^d(a,\theta)$ is the steady state transition density for occupational choice d. We can define another set of operators that translate changes in policy rules into changes in the distribution

$$\mathcal{M}^{(a)} \cdot \mathsf{x}\langle a',\theta'\rangle = \iint \sum_d \overline{\Lambda}^d(a',\theta',a,\theta)\mathsf{p}\mathsf{x}^d(a,\theta)\omega^*(a,\theta)dad\theta$$

$$\mathcal{M} \cdot \mathsf{x}\langle a',\theta'\rangle = \iint \sum_d \overline{\Lambda}^d(a',\theta',a,\theta)\frac{\bar{p}_\kappa^d(a,\theta)}{\bar{p}^d(a,\theta)}\left(\mathsf{p}_v\mathsf{x}^b(a,\theta) - \mathsf{p}_v\mathsf{x}^w(a,\theta)\right)\omega^*(a,\theta)dad\theta.$$

With these operators in hand, it's possible to use (56) to prove the following Lemma.

Lemma 3. *The change in the distribution, $\hat{\Omega}_t$, satisfies*

$$\frac{d^2}{dad\theta}\hat{\Omega}_t(a,\theta) = \hat{\omega}_t(a,\theta) - \frac{d}{da}\hat{\omega}_t^{(a)}(a,\theta), \tag{60}$$

where $\hat{\omega}_t$ and $\hat{\omega}_t^{(a)}$ satisfy the following recursion:

$$\hat{\omega}_{t+1} = \mathcal{L}\cdot\hat{\omega}_t + \mathcal{L}_{(a)}\cdot\hat{\omega}_t^{(a)} + \mathcal{M}\cdot\hat{x}_t$$
$$\hat{\omega}_{t+1}^{(a)} = \mathcal{L}^{(a)}\cdot\hat{\omega}_t^{(a)} + \mathcal{M}^{(a)}\cdot\hat{x}_t,$$

with $\hat{\omega}_0 = \omega_{-1}^* - \omega^*$ and $\hat{\omega}_0^{(a)} = 0$.

Proof. Starting with the initial conditions, we have that $\frac{\partial^2}{\partial a \partial \theta}\hat{\Omega}_0(a,\theta)$ satisfies (60) with $\hat{\omega}_0 = \omega^*_{-1} - \omega^*$ and $\hat{\omega}_0^{(a)} = 0$. We next proceed with an induction argument. Using $\hat{a}_t^d(a,\theta) = \mathsf{p}\hat{x}_t^d(a,\theta)$ and

$$\hat{p}_t^d(a,\theta) = \bar{p}_\kappa^d(a,\theta)\left(\mathsf{p}_v \hat{x}_t^b(a,\theta) - \mathsf{p}_v \hat{x}_t^w(a,\theta)\right),$$

we have that equation (56) implies that

$$\begin{aligned}\frac{\partial^2}{\partial a \partial \theta}\hat{\Omega}_{t+1}\langle a,\theta\rangle = &\int \sum_d \delta\left(a' - \bar{a}^d(a,\theta)\right)\mu\left(\theta' - \rho_\theta \theta\right)\bar{p}^d(a,\theta) d\hat{\Omega}_t(a,\theta) \\ &+ \mathcal{M}\cdot \hat{x}_t\langle a,\theta\rangle - \frac{\partial}{\partial a'}\mathcal{M}^{(a)}\cdot \hat{x}_t\langle a,\theta\rangle\end{aligned}$$

Substituting for $\hat{\Omega}_t$ using (60), we have that

$$\begin{aligned}&\int \sum_d \delta\left(a' - \bar{a}^d(a,\theta)\right)\mu\left(\theta' - \rho_\theta \theta\right)\bar{p}^d(a,\theta) d\hat{\Omega}_t(a,\theta) \\ &\quad = \iint \sum_d \delta\left(a' - \bar{a}^d(a,\theta)\right)\mu\left(\theta' - \rho_\theta \theta\right)\bar{p}^d(a,\theta)\left(\hat{\omega}_t(a,\theta) - \frac{\partial}{\partial a}\hat{\omega}_t^{(a)}(a,\theta)\right) d\Omega^*(a,\theta) da d\theta \\ &\quad = \mathcal{L}\cdot\hat{\omega}_t\langle a',\theta'\rangle + \iint \sum_d \delta\left(a' - \bar{a}^d(a,\theta)\right)\mu\left(\theta' - \rho_\theta \theta\right)\bar{p}_a^d(a,\theta)\hat{\omega}_t^{(a)}(a,\theta) da d\theta \\ &\qquad - \iint \sum_d \frac{\partial}{\partial a'}\left(\delta\left(a' - \bar{a}^d(a,\theta)\right)\right)\mu\left(\theta' - \rho_\theta \theta\right)\bar{p}^d(a,\theta)\bar{a}_a^d(a,\theta)\hat{\omega}_t^{(a)}(a,\theta) da d\theta \\ &\quad = \mathcal{L}\cdot\hat{\omega}_t\langle a',\theta'\rangle + \mathcal{L}_{(a)}\cdot\hat{\omega}_t^{(a)}\langle a',\theta'\rangle - \frac{\partial}{\partial a'}\left(\mathcal{L}^{(a)}\cdot\hat{\omega}_t^{(a)}\langle a',\theta'\rangle\right),\end{aligned}$$

which completes the proof. □

These operators have intuitive interpretations. The operator $\mathcal{M}$ captures the large changes in the distribution resulting from occupational choice. The operator $\mathcal{L}$ takes those changes in the distribution and pushes them forward using the steady-state transition density. The operator $\mathcal{M}^{(a)}$ captures changes in the distribution induced by small changes in the savings decisions. It does that by weighting those changes by the steady state density and then integrating over the steady state transition density. The operator $\mathcal{L}^{(a)}$ pushes those changes in the distribution forward through time, while the operator $\mathcal{L}_{(a)}$ captures the changes in occupation choice induced by those small savings decisions.

We can use Lemma 3 to construct the following representation of the distribution.

Corollary 1. *$\hat{\omega}_t(a,\theta)$ and $\hat{\omega}_t^{(a)}(a,\theta)$ can be expressed as*

$$\hat{\omega}_t(a,\theta) = \sum_{s=0}^{\infty} \mathsf{A}_{t,s}\hat{Y}_s + \mathcal{L}^t \cdot \hat{\omega}_0 \tag{61}$$

$$\hat{\omega}_t^{(a)}(a,\theta) = \sum_{s=0}^{\infty} \mathsf{A}_{t,s}^{(a)}\hat{Y}_s, \tag{62}$$

where $\mathsf{A}_{0,s} = \mathsf{A}_{0,s}^{(a)} = 0$ for all s and $\mathsf{A}_{t,s}$ and $\mathsf{A}_{t,s}^{(a)}$ satisfy the following recursive system:

$$\mathsf{A}_{t+1,s}^{(a)} = \mathcal{L}^{(a)} \cdot \mathsf{A}_{t,s}^{(a)} + \mathcal{M}^{(a)} \cdot \mathsf{x}_{s-t}$$
$$\mathsf{A}_{t+1,s} = \mathcal{L} \cdot \mathsf{A}_{t,s} + \mathcal{L}_{(a)} \cdot \mathsf{A}_{t,s}^{(a)} + \mathcal{M} \cdot \mathsf{x}_{s-t}.$$

Proof. We can use the recursive system in Lemma 3 to prove this via induction. First of all, the initial conditions imply that $\mathsf{A}_{0,s} = \mathsf{A}_{0,s}^{(a)} = 0$ for all s. We then proceed by induction:

$$\begin{aligned}\hat{\omega}_{t+1}^{(a)} &= \mathcal{L}^{(a)} \cdot \hat{\omega}_t^{(a)} + \mathcal{M}^{(a)} \cdot \left(\sum_{j=0}^{\infty} \mathsf{x}_j \hat{Y}_{t+j}\right) \\ &= \mathcal{L}^{(a)} \cdot \left(\sum_{s=0}^{\infty} \mathsf{A}_{t,s}^{(a)}\hat{Y}_s\right) + \mathcal{M}^{(a)} \cdot \left(\sum_{s=0}^{\infty} \mathsf{x}_{s-t}\hat{Y}_s\right) \\ &= \sum_{s=0}^{\infty}\left(\mathcal{L}^{(a)} \cdot \mathsf{A}_{t,s}^{(a)} + \mathcal{M}^{(a)} \cdot \mathsf{x}_{s-t}\right)\hat{Y}_s,\end{aligned}$$

as desired. The second equation is proved in a similar manner. □

C.3 The Market Clearing Conditions: G

Finally, we can differentiate the market clearing conditions to get

$$\mathsf{G}_x \left(\int \bar{x} d\Omega\right)_Z \cdot \hat{Z}_t + \mathsf{G}_Y \hat{Y}_t = 0,$$

where $\left(\int \bar{x} d\Omega\right)_Z \cdot \hat{Z}_t$ is the derivative of the aggregate individual decisions in direction $\hat{Z}_t$. Those derivatives equal

$$\left(\int \bar{x} d\Omega\right)_Z \cdot \hat{Z}_t = \int \sum_d \hat{x}_t^d(a,\theta)\bar{p}^d(a,\theta)d\Omega^*(a,\theta) + \int \sum_d \bar{x}^d(a,\theta)\bar{p}_\kappa^d(a,\theta)\hat{\kappa}_t(a,\theta)d\Omega^*(a,\theta) + \int \sum_d \bar{x}^d(a,\theta)\bar{p}^d(a,\theta)d\hat{\Omega}_t(a,\theta). \tag{63}$$

The first two terms can be expressed in terms of $\hat{Y}_{t+s}$ directly from Lemma 2. If we define the operators

$$\mathcal{I} \cdot \mathsf{y} = \int \sum_d \bar{x}^d(a,\theta)\bar{p}^d(a,\theta)\mathsf{y}(a,\theta)dad\theta$$
$$\mathcal{I}^{(a)} \cdot \mathsf{y} = \int \sum_d \left(\bar{x}_a^d(a,\theta) + \bar{x}^d(a,\theta)\frac{\bar{p}_a^d(a,\theta}{\bar{p}^d(a,\theta)}\right)\bar{p}^d(a,\theta)\mathsf{y}(a,\theta)dad\theta,$$

then integration by parts gives us that

$$\int \sum_d \bar{x}^d(a,\theta)\bar{p}^d(a,\theta)d\hat{\Omega}_t(a,\theta) = \mathcal{I}\cdot\hat{\omega}_t + \mathcal{I}^{(a)}\cdot\hat{\omega}_t^{(a)}.$$

All of these results can be combined to obtain the following corollary.

Corollary 2. *The aggregated decision rules satisfy*

$$\left(\int \bar{x} d\Omega\right)_Z \cdot \hat{Z}_t = \sum_{s=0}^{\infty} \mathsf{J}_{t,s}\hat{Y}_s + \mathsf{J}_t^{TD}, \tag{64}$$

where $\mathsf{J}_t^{TD} = \mathcal{I}\cdot\mathcal{L}^t\cdot\hat{\omega}_0$ *and* $\mathsf{J}_{t,s}$ *solves the following recursive system:*

$$\mathsf{J}_{t,s} = \mathsf{J}_{t-1,s-1} + \mathcal{I}\cdot\mathcal{L}^{t-1}\cdot\mathcal{M}\cdot\mathsf{x}_s + \mathcal{I}\mathcal{L}_t^{(a)}\cdot\mathcal{M}^{(a)}\cdot\mathsf{x}_s$$
$$\mathcal{I}\mathcal{L}_t^{(a)} = \mathcal{I}\mathcal{L}_{t-1}^{(a)}\cdot\mathcal{L}^{(a)} + \mathcal{I}\cdot\mathcal{L}^{t-2}\cdot\mathcal{L}_{(a)},$$

with the initial conditions $\mathcal{I}\mathcal{L}_1^{(a)} = \mathcal{I}^{(a)}$ *and*

$$\mathsf{J}_{0,s} = \int \sum_d \mathsf{x}_s^d(a,\theta)\bar{p}^d(a,\theta)d\Omega^*(a,\theta) + \int \sum_d \bar{x}^d(a,\theta)\bar{p}_\kappa^d(a,\theta)\mathsf{p}_v\left(\mathsf{x}_s^b(a,\theta) - \mathsf{x}_s^w(a,\theta)\right)d\Omega^*(a,\theta).$$

Proof. We begin by noting that (63) and Corollary 1 imply that

$$\mathsf{J}_{t,s} = \int \sum_d \mathsf{x}^d_{s-t}(a,\theta)\bar{p}^d(a,\theta)d\Omega^*(a,\theta) + \int \sum_d \bar{x}^d(a,\theta)\bar{p}^d_\kappa(a,\theta)\mathsf{p}_v\left(\mathsf{x}^b_{s-t}(a,\theta) - \mathsf{x}^w_{s-t}(a,\theta)\right) d\Omega^*(a,\theta)$$

$$+ \mathcal{I}\cdot\mathsf{A}_{t,s} + \mathcal{I}^{(a)}\cdot\mathsf{A}^{(a)}_{t,s}.$$

We note that $\mathsf{J}_{0,s}$ satisfies the initial conditions for all s. Next, we note that

$$\mathsf{J}_{t,s} - \mathsf{J}_{t-1,s-1} = \mathcal{I}\cdot(\mathsf{A}_{t,s} - \mathsf{A}_{t-1,s-1}) + \mathcal{I}^{(a)}\cdot\left(\mathsf{A}^{(a)}_{t,s} - \mathsf{A}^{(a)}_{t-1,s-1}\right).$$

We can then use the recursive system for $\mathsf{A}_{t,s}$ and $\mathsf{A}^{(a)}_{t,s}$ to obtain the recursive system for $\mathsf{J}_{t,s}$.

Let $B^{(a)}_t = \left(\mathcal{L}^{(a)}\right)^{t-1}\cdot\mathcal{M}^{(a)}$, our first step will be to show that $A^{(a)}_{t,s} - A^{(a)}_{t-1,s-1} = B^{(a)}_t\cdot\mathsf{x}_s$. We observe that $A^{(a)}_{1,s} - A^{(a)}_{0,s-1} = \mathcal{M}^{(a)}\cdot\mathsf{x}_s$ and then proceed by induction using the recursive system for $\mathsf{A}^{(a)}_{t,s}$ to find:

$$\begin{aligned}
A^{(a)}_{t,s} - A^{(a)}_{t-1,s-1} &= \mathcal{L}^{(a)}\cdot A^{(a)}_{t-1,s} + \mathcal{M}^{(a)}\cdot\mathsf{x}_{s-t+1} - \left(\mathcal{L}^{(a)}\cdot\mathsf{A}^{(a)}_{t-2,s-1} + \mathcal{M}^{(a)}\cdot\mathsf{x}_{s-t+1}\right)\\
&= \mathcal{L}^{(a)}\cdot(\mathsf{A}_{t-1,s} - \mathsf{A}_{t-2,s-1}) = \mathcal{L}^{(a)}\cdot\left(\mathcal{L}^{(a)}\right)^{t-2}\cdot\mathcal{M}^{(a)}\cdot\mathsf{x}_s\\
&= B^{(a)}_t\cdot\mathsf{x}_s
\end{aligned}$$

Next, we show that $A_{t,s} - A_{t-1,s-1} = (\mathcal{L})^{t-1}\cdot\mathcal{M}\cdot\mathsf{x}_s + \mathsf{B}_t\cdot\mathcal{M}^{(a)}\cdot\mathsf{x}_s$, where $\mathsf{B}_1 = 0$ and $\mathsf{B}_t = \mathcal{L}\cdot\mathsf{B}_{t-1} + \mathcal{L}_{(a)}\cdot\mathsf{B}^{(a)}_{t-1}$. It holds true for $t=1$ since $A_{1,s} - A_{0,s-1} = \mathcal{M}\cdot\mathsf{x}_s$. We then proceed by induction to show that it holds for all t. Specifically,

$$\begin{aligned}
A_{t,s} - A_{t-1,s-1} &= \mathcal{L}\cdot A_{t-1,s} + \mathcal{L}_{(a)}\cdot A^{(a)}_{t-1,s} + \mathcal{M}\cdot\mathsf{x}_{s-t} - \left(\mathcal{L}\cdot A_{t-2,s-1} + \mathcal{L}_{(a)}\cdot A^{(a)}_{t-2,s-1} + \mathcal{M}\cdot\mathsf{x}_{s-t}\right)\\
&= \mathcal{L}\cdot(A_{t-1,s} - A_{t-2,s-1}) + \mathcal{L}_{(a)}\cdot\left(A^{(a)}_{t-1,s} - A^{(a)}_{t-2,s-1}\right)\\
&= \mathcal{L}\cdot\left(\mathcal{L}^{t-2}\cdot\mathcal{M}\cdot\mathsf{x}_s + \mathsf{B}_{t-1}\cdot\mathcal{M}^{(a)}\cdot\mathsf{x}_s\right) + \mathcal{L}_{(a)}\cdot\left(\mathcal{L}^{(a)}\right)^{t-2}\cdot\mathcal{M}^{(a)}\cdot\mathsf{x}_s\\
&= (\mathcal{L})^{t-1}\cdot\mathcal{M}\cdot\mathsf{x}_s + \mathsf{B}_t\cdot\mathcal{M}^{(a)}\cdot\mathsf{x}_s
\end{aligned}$$

as desired.

Using the initial conditions that $\mathsf{B}_1 = 0$, the recursive system implies that $\mathsf{B}_t = \sum_{j=0}^{t-2}\mathcal{L}^{t-2-j}\cdot\mathcal{L}_{(a)}\cdot\left(\mathcal{L}^{(a)}\right)^j$. Therefore, B_t can also be constructed recursively via $\mathsf{B}_1 = 0$ and $\mathsf{B}_t = \mathsf{B}_{t-1}\cdot\mathcal{L}^{(a)} +$

$\mathcal{L}^{t-2} \cdot \mathcal{L}_{(a)}$. This recursive structure implies that $\mathcal{I} \cdot \mathsf{B}_t + \mathcal{I}^{(a)} \cdot \left(\mathcal{L}^{(a)}\right)^{t-1} = \mathcal{I}\mathcal{L}_t^{(a)}$ and thus

$$\begin{aligned}
\mathsf{J}_{t,s} - \mathsf{J}_{t-1,s-1} &= \mathcal{I} \cdot \mathsf{A}_{t,s} + \mathcal{I}^{(a)} \cdot \mathsf{A}_{t,s}^{(a)} \\
&= \mathcal{I} \cdot \left(\mathcal{L}^{t-1} \cdot \mathcal{M} \cdot \mathsf{x}_s + \mathsf{B}_t \cdot \mathcal{M}^{(a)} \cdot \mathsf{x}_s\right) + \mathcal{I}^{(a)} \cdot \left(\mathcal{L}^{(a)}\right)^{t-1} \cdot \mathcal{M}^{(a)} \cdot \mathsf{x}_s \\
&= \mathcal{I} \cdot \mathcal{L}^{t-1} \cdot \mathcal{M} \cdot \mathsf{x}_s + \mathcal{I} \cdot \mathsf{B}_t \cdot \mathcal{M}^{(a)} \cdot \mathsf{x}_s + \mathcal{I}^{(a)} \cdot \left(\mathcal{L}^{(a)}\right)^{t-1} \cdot \mathcal{M}^{(a)} \cdot \mathsf{x}_s \\
&= \mathcal{I} \cdot \mathcal{L}^{t-1} \cdot \mathcal{M} \cdot \mathsf{x}_s + \mathcal{I}\mathcal{L}_t^{(a)} \cdot \mathcal{M}^{(a)} \cdot \mathsf{x}_s
\end{aligned}$$

as desired. □

We can now use this corollary to show that the sequence $\{\hat{X}_t\}_t$ solves a linear system.

Proposition 1. *The sequence $\{\hat{X}_t\}_t$ solves the linear system*

$$\mathsf{G}_Y \hat{Y}_t + \mathsf{G}_x \sum_{s=0}^{\infty} \mathsf{J}_{t,s} \hat{Y}_s + \mathsf{G}_x \mathsf{J}_t^{TD} = 0, \tag{65}$$

where $\hat{Y}_t = [\mathsf{P}\hat{X}_t, \hat{X}_t, \hat{X}_{t+1}]$, $\mathsf{P}\hat{X}_{-1} = X^*_{-1} - X^*$ and $\lim_{t\to\infty} \hat{X}_t = 0$.

Proof. This is a direct consequence of substituting Corollary 2 into the market clearing conditions (48) along with the initial conditions $\hat{Z}_0 = Z^*_{-1} - Z^*$. □

D Second Order Approximation

To perform the second order approximation, we define a set of directions via the Frechét derivatives of $\bar{Z}(Z)$. We define $\hat{Z}_{t,t}$ recursively via $\hat{Z}_{0,0} = 0$ and

$$\hat{Z}_{t+1,t+1} = \bar{Z}_Z \cdot \hat{Z}_{t,t} + \bar{Z}_{ZZ} \cdot (\hat{Z}_t, \hat{Z}_t),$$

where $\bar{Z}_{ZZ} \cdot (\cdot, \cdot)$ represents the bi-linear form associated with the second Frechét derivative of Z. We then define the direction $\hat{X}_{t,t}$ via

$$\hat{X}_{t,t} = \bar{X}_Z \cdot \hat{Z}_{t,t} + \bar{X}_{ZZ} \cdot (\hat{Z}_t, \hat{Z}_t).$$

We now show that the second order approximation to the transition path can be approximated using these directions.

Lemma 4. *To second order, X_t is approximated by*

$$X_t = \bar{X} + \hat{X}_t + \frac{1}{2}\hat{X}_{t,t} + o(\|Z_{-1} - Z^*\|^2).$$

Proof. We begin by taking the second derivative of the recursive system (51) with respect to σ. Doing so yields

$$\bar{X}_{t,\sigma\sigma} = \bar{X}_{ZZ} \cdot \left(\bar{Z}_{t,\sigma}, \bar{Z}_{t,\sigma}\right) + \bar{X}_Z \cdot \bar{Z}_{t,\sigma\sigma}$$
$$\bar{Z}_{t,\sigma\sigma} = \bar{Z}_{ZZ} \cdot \left(\bar{Z}_{t-1,\sigma}, \bar{Z}_{t-1,\sigma}\right) + \bar{Z}_Z \cdot \bar{Z}_{t-1,\sigma\sigma} \text{ for } t > 0,$$

with $\bar{Z}_{0,\sigma\sigma} = 0$.[10] From the proof of Lemma 1, we know that $\bar{Z}_{t,\sigma} = \hat{Z}_t$. We can therefore conclude $\bar{Z}_{t,\sigma\sigma}$ satisfies the law of motion for $\hat{Z}_{t,t}$ with the same initial condition. Thus $\bar{Z}_{t,\sigma\sigma} = \hat{Z}_{t,t}$ and

$$\bar{X}_{t,\sigma\sigma} = \bar{X}_{ZZ} \cdot \left(\hat{Z}_t, \hat{Z}_t\right) + \bar{X}_Z \cdot \hat{Z}_{t,t} = \hat{X}_{t,t}$$

as desired. □

We proceed in the same manner as in the first-order approximation. To obtain the directions $\hat{X}_{t,t}$, we first show how individual policy rules depend on these aggregates and then use that dependence for approximating the law of motion of the distribution. This procedure allows us to express the derivative of (48) in terms of the directional derivatives $\hat{X}_{t,t}$ and terms already known to the first order.

D.1 The F^d Equation

We begin by defining some functions. We let

$$\hat{x}_t^{d,e}(a,\theta) = \bar{\mathbb{E}}^d\left[\sum_{d'} \mathsf{H}_{x^{d'}}(\cdot,\cdot)\hat{x}_{t+1}^{d'}(\cdot,\cdot)\middle| a,\theta\right]$$
$$\left(\mathbb{E}^d[\bar{x}|a,\theta]\right)_Z \cdot \hat{Z}_t = \hat{x}_t^{d,e}(a,\theta) + \bar{x}_a^{d,e}(a,\theta)\mathsf{p}\hat{x}_t^d(a,\theta).$$

[10]Since all of the effect of the perturbation is transferred through the aggregate state, the derivatives $\bar{X}_{\sigma\sigma}$ and $\bar{X}_{\sigma Z}$ are zero, and similarly for Z. Thus, we drop these terms.

Both can be constructed from the first order terms. Differentiating equation (49) twice in direction $\hat{Z}_t$ and then adding to it the derivative in direction $\hat{Z}_{t,t}$, we get

$$\begin{aligned}
&\left(\mathbb{E}^d[\bar{x}|a,\theta]\right)_Z \cdot \hat{Z}_{t,t} + \left(\mathbb{E}^d[\bar{x}|a,\theta]\right)_{ZZ} \cdot (\hat{Z}_t, \hat{Z}_t) \\
&\qquad = \bar{\mathbb{E}}^d\left[\sum_{d'} \mathsf{H}_{x^{d'}}(\cdot,\cdot)\hat{x}^{d'}_{t+1,t+1}(\cdot,\cdot)\middle| a,\theta\right] + \bar{x}^{d,e}_a(a,\theta)\mathsf{p}\hat{x}^d_{t,t}(a,\theta) \\
&\qquad\quad + \bar{\mathbb{E}}^d\left[\sum_{d_1',d_2'} \mathsf{H}_{x^{d_1'},x^{d_2'}}(\cdot,\cdot)\cdot\left(\hat{x}^{d_1'}_{t+1}(\cdot,\cdot),\hat{x}^{d_2'}_{t+1}(\cdot,\cdot)\right)\middle| a,\theta\right] + 2\hat{x}^{d,e}_{t,a}(a,\theta)\mathsf{p}\hat{x}^d_t(a,\theta) \\
&\qquad\quad + \bar{x}^{d,e}_{aa}(a,\theta)\cdot\left(\mathsf{p}\hat{x}^d_t(a,\theta),\mathsf{p}\hat{x}^d_t(a,\theta)\right),
\end{aligned}$$

where $\hat{x}^{d,e}_{t,a}(a,\theta)$ represents the deriative of $\hat{x}^{d,e}_t(a,\theta)$ with respect to a.

We can then differentiate equation (47) twice in direction $\hat{Z}_t$ and then add to it the derivative in direction $\hat{Z}_{t,t}$ to get

$$\begin{aligned}
0 =& \mathsf{F}^d_x(a,\theta)\hat{x}^d_{t,t}(a,\theta) + \mathsf{F}^d_Y(a,\theta)\hat{Y}_{t,t} + \mathsf{F}^d_{t,t}(a,\theta) \\
&+ \mathsf{F}^d_{x'}(a,\theta)\left(\bar{\mathbb{E}}^d\left[\sum_{d'}\mathsf{H}_{x^{d'}}(\cdot,\cdot)\hat{x}^{d'}_{t+1,t+1}(\cdot,\cdot)\middle| a,\theta\right] + \bar{x}^{d,e}_a(a,\theta)\mathsf{p}\hat{x}^d_{t,t}(a,\theta)\right),
\end{aligned} \tag{66}$$

where $\hat{x}^d_{t,t}(a,\theta) = \bar{x}^d_Z(a,\theta)\cdot\hat{Z}_{t,t} + \bar{x}^d_{ZZ}(a,\theta)\cdot(\hat{Z}_t,\hat{Z}_t)$. The term

$$\begin{aligned}
\mathsf{F}^d_{t,t}(a,\theta) =& \mathsf{F}^d_{xx}(a,\theta)\cdot(\hat{x}^d_t(a,\theta),\hat{x}^d_t(a,\theta)) + 2\mathsf{F}^d_{xY}(a,\theta)\cdot(\hat{x}^d_t(a,\theta),\hat{Y}_t) \\
&+ 2\mathsf{F}^d_{xx'}(a,\theta)\cdot\left(\hat{x}^d_t(a,\theta),\left(\mathbb{E}^d[\bar{x}|a,\theta]\right)_Z\cdot\hat{Z}_t\right) + \mathsf{F}^d_{YY}\cdot(\hat{Y}_t,\hat{Y}_t) \\
&+ 2\mathsf{F}^d_{Yx'}(a,\theta)\cdot\left(\hat{Y}_t,\left(\mathbb{E}^d[\bar{x}|a,\theta]\right)_Z\cdot\hat{Z}_t\right) \\
&+ \mathsf{F}^d_{x'x'}(a,\theta)\cdot\left(\left(\mathbb{E}^d[\bar{x}|a,\theta]\right)_Z\cdot\hat{Z}_t,\left(\mathbb{E}^d[\bar{x}|a,\theta]\right)_Z\cdot\hat{Z}_t\right) \\
&+ \mathsf{F}^d_{x'}(a,\theta)\Bigg(\bar{\mathbb{E}}^d\left[\sum_{d_1',d_2'}\mathsf{H}_{x^{d_1'},x^{d_2'}}(\cdot,\cdot)\cdot\left(\hat{x}^{d_1'}_{t+1}(\cdot,\cdot),\hat{x}^{d_2'}_{t+1}(\cdot,\cdot)\right)\middle| a,\theta\right] + 2\hat{x}^{d,e}_{t,a}(a,\theta)\mathsf{p}\hat{x}^d_t(a,\theta) \\
&\quad + \bar{x}^{d,e}_{aa}(a,\theta)\cdot\left(\mathsf{p}\hat{x}^d_t(a,\theta),\mathsf{p}\hat{x}^d_t(a,\theta)\right)\Bigg).
\end{aligned}$$

contains all of the second order cross interactions. The change in expectations $\left(\mathbb{E}^d[\bar{x}|a,\theta]\right)_Z\cdot\hat{Z}_t$ is found by differentiating (49) twice in direction $\hat{Z}_t$ and adding to it the derivative in direction $\hat{Z}_{t,t}$.

In the next lemma, we follow the same steps as in Lemma 2.

Lemma 5. *For any* t

$$\hat{x}^d_{t,t}(a,\theta) = \sum_{s=0}^{\infty} \mathsf{x}^d_s(a,\theta)\hat{Y}_{t+s,t+s} + \mathsf{x}^d_{t,t}(a,\theta), \tag{67}$$

where x^d_s *is the same as in Lemma 2 and* $\mathsf{x}^d_{t,t}$ *satisfies*

$$\mathsf{x}^d_{t,t}(a,\theta) = -\left(\mathsf{F}^d_x(a,\theta) + \mathsf{F}^d_{x'}(a,\theta)\bar{x}^{d,e}_a(a,\theta)\mathsf{p}\right)^{-1}\left(\mathsf{F}^d_{x'}(a,\theta)\mathsf{x}^{d,e}_{t,t}(a,\theta) + \mathsf{F}^d_{t,t}(a,\theta)\right)$$
$$\mathsf{x}^{d,e}_{t,t}(a,\theta) = \bar{\mathbb{E}}\left[\sum_{d'} \mathsf{H}_{x^{d'}}(\cdot,\cdot)\mathsf{x}^{d'}_{t+1,t+1}(\cdot,\cdot)|a,\theta\right].$$

Proof. Using the same steps as the proof Lemma 2, we can show that

$$\begin{aligned}
&\mathsf{F}^d_{x'}(a,\theta)\left(\bar{\mathbb{E}}\left[\sum_{d'} \mathsf{H}_{x^{d'}}(\cdot,\cdot)\hat{x}^{d'}_{t+1,t+1}(\cdot,\cdot)\middle| a,\theta\right] + \bar{x}^{d,e}_a(a,\theta)\mathsf{p}\hat{x}^d_{t,t}(a,\theta)\right) \\
&\quad = \sum_{s=0}^{\infty} \mathsf{F}^d_{x'}(a,\theta)\mathsf{x}^{d,e}_s(a,\theta)\hat{Y}_{t+s+1,t+s=1} + \mathsf{F}^d_{x'}(a,\theta)\mathsf{x}^{d,e}_{t,t}(a,\theta) + \mathsf{F}^d_{x'}(a,\theta)\bar{x}^{d,e}_a(a,\theta)\mathsf{p}\hat{x}^d_{t,t}(a,\theta) \\
&\quad = -\mathsf{F}^d_x(a,\theta)\hat{x}^d_{t,t}(a,\theta) - \mathsf{F}^d_{t,t}(a,\theta) - \mathsf{F}^d_Y(a,\theta)\hat{Y}_{t,t}.
\end{aligned}$$

We can then conclude that (66) holds with equality. □

D.2 The Law of Motion of Ω

Our next step is to determine how individual policies aggregate into changes in the distribution. To achieve this, we first note that the first derivative of the law of motion evaluated at an arbitrary point in the state space is

$$\begin{aligned}
\frac{\partial^2}{\partial a'\partial\theta'}\bar{\Omega}_Z(Z)\cdot\hat{Z}_t\langle a',\theta'\rangle &= \int\sum_d \delta\left(a' - \bar{a}^d(a,\theta,Z)\right)\mu\left(\theta' - \rho_\theta\theta\right)\bar{p}^d(a,\theta,Z)d\hat{\Omega}_t(a,\theta) \\
&\quad + \int\sum_d \delta\left(a' - \bar{a}^d(a,\theta,Z)\right)\mu\left(\theta' - \rho_\theta\theta\right)\hat{p}^d_t(a,\theta,Z)d\Omega(a,\theta) \\
&\quad - \int\sum_d \delta'\left(a' - \bar{a}^d(a,\theta,Z)\right)\mu\left(\theta' - \rho_\theta\theta\right)\bar{p}^d(a,\theta,Z)\hat{a}^d_t(a,\theta,Z)d\Omega(a,\theta).
\end{aligned} \tag{68}$$

Differentiating this in direction $\hat{Z}_t$, we obtain

$$
\begin{aligned}
&\frac{\partial^2}{\partial a' \partial \theta'} \bar{\Omega}_{ZZ} \cdot (\hat{Z}_t, \hat{Z}_t) \langle a', \theta' \rangle \qquad (69) \\
&\quad = -2 \int \sum_d \delta' \left(a' - \bar{a}^d(a,\theta) \right) \mu \left(\theta' - \rho_\theta \theta \right) \bar{p}^d(a,\theta) \hat{a}_t(a,\theta) d\hat{\Omega}_t(a,\theta) \\
&\quad + 2 \int \sum_d \delta \left(a' - \bar{a}^d(a,\theta) \right) \mu \left(\theta' - \rho_\theta \theta \right) \hat{p}^d_t(a,\theta) d\hat{\Omega}_t(a,\theta) \\
&\quad + \int \sum_d \delta \left(a' - \bar{a}^d(a,\theta) \right) \mu \left(\theta' - \rho_\theta \theta \right) \bar{p}^d_{ZZ}(a,\theta) \cdot (\hat{Z}_t, \hat{Z}_t) d\Omega^*(a,\theta) \qquad (70) \\
&\quad - 2 \int \sum_d \delta' \left(a' - \bar{a}^d(a,\theta) \right) \mu \left(\theta' - \rho_\theta \theta \right) \hat{p}^d_t(a,\theta) \hat{a}^d_t(a,\theta) d\Omega^*(a,\theta) \\
&\quad - \int \sum_d \delta' \left(a' - \bar{a}^d(a,\theta) \right) \mu \left(\theta' - \rho_\theta \theta \right) \bar{p}^d(a,\theta) \bar{a}^d_{ZZ}(a,\theta) \cdot (\hat{Z}_t, \hat{Z}_t) d\Omega^*(a,\theta) \\
&\quad + \int \sum_d \delta'' \left(a' - \bar{a}^d(a,\theta) \right) \mu \left(\theta' - \rho_\theta \theta \right) \bar{p}^d(a,\theta) \hat{a}^d_t(a,\theta) \hat{a}^d_t(a,\theta) d\Omega^*(a,\theta).
\end{aligned}
$$

We can then add to this the derivative of the law of motion in the direction of $\hat{Z}_{t,t}$,

$$
\begin{aligned}
\frac{\partial^2}{\partial a' \partial \theta'} \bar{\Omega} \cdot \hat{Z}_{t,t} \langle a', \theta' \rangle &= \int \sum_d \delta \left(a' - \bar{a}^d(a,\theta,Z) \right) \mu \left(\theta' - \rho_\theta \theta \right) \bar{p}^d(a,\theta) d\hat{\Omega}_{t,t}(a,\theta) \\
&+ \int \sum_d \delta \left(a' - \bar{a}^d(a,\theta) \right) \mu \left(\theta' - \rho_\theta \theta \right) \bar{p}^d_Z(a,\theta) \cdot \hat{Z}_{t,t} d\Omega^*(a,\theta) \qquad (71) \\
&- \int \sum_d \delta' \left(a' - \bar{a}^d(a,\theta) \right) \mu \left(\theta' - \rho_\theta \theta \right) \bar{p}^d(a,\theta,Z) \bar{a}^d_Z(a,\theta) \cdot \hat{Z}_{t,t} d\Omega^*(a,\theta),
\end{aligned}
$$

to obtain the law of motion for $\hat{\Omega}_{t+1,t+1} = \bar{\Omega}_{ZZ} \cdot (\hat{Z}_t, \hat{Z}_t) + \bar{\Omega}_Z \cdot \hat{Z}_{t,t}$. Specifically, if we define the operators

$$
\begin{aligned}
\mathcal{L}^{(aa)} \cdot \mathsf{y} \langle a', \theta' \rangle &= \iint \sum_d \bar{\Lambda}^d(a',\theta',a,\theta) \bar{a}^d_a(a,\theta) \bar{a}^d_a(a,\theta) \mathsf{y}(a,\theta) da d\theta \\
\mathcal{L}^{(a)}_{(aa)} \cdot \mathsf{y} \langle a', \theta' \rangle &= \iint \sum_d \bar{\Lambda}^d(a',\theta',a,\theta) \left(\bar{a}^d_{aa}(a,\theta) + 2\bar{a}^d_a(a,\theta) \frac{\bar{p}^d_a(a,\theta)}{\bar{p}^d(a,\theta)} \right) \mathsf{y}(a,\theta) da d\theta \\
\mathcal{L}_{(aa)} \cdot \mathsf{y} \langle a', \theta' \rangle &= \iint \sum_d \bar{\Lambda}^d(a',\theta',a,\theta) \frac{\bar{p}^d_{aa}(a,\theta)}{\bar{p}^d(a,\theta)} \mathsf{y}(a,\theta) da d\theta,
\end{aligned}
$$

then we can show the following result.

Lemma 6. *The change in the distribution, $\hat{\Omega}_t$, satisfies*

$$\frac{\partial^2}{\partial a \partial \theta}\hat{\Omega}_{t,t}(a,\theta) = \hat{\omega}_{t,t}(a,\theta) - \frac{\partial}{\partial a}\hat{\omega}^{(a)}_{t,t}(a,\theta) + \frac{\partial^2}{\partial a^2}\hat{\omega}^{(aa)}_{t,t}(a,\theta), \tag{72}$$

where $\hat{\omega}_{t,t}$, $\hat{\omega}^{(a)}_{t,t}$, and $\hat{\omega}^{(aa)}_{t,t}$ satisfy the following recursion

$$\begin{aligned}
\hat{\omega}_{t+1,t+1} &= \mathcal{L}\cdot\hat{\omega}_{t,t} + \mathcal{L}_{(a)}\cdot\hat{\omega}^{(a)}_{t,t} + \mathcal{L}_{(aa)}\cdot\hat{\omega}^{(aa)}_{t,t} + \mathcal{M}\cdot\hat{x}_{t,t} + \mathsf{c}_{t,t}\\
\hat{\omega}^{(a)}_{t+1,t+1} &= \mathcal{L}^{(a)}\cdot\hat{\omega}^{(a)}_{t,t} + \mathcal{L}^{(a)}_{(aa)}\cdot\hat{\omega}^{(aa)}_{t,t} + \mathcal{M}^{(a)}\cdot\hat{x}_{t,t} + \mathsf{c}^{(a)}_{t,t}\\
\hat{\omega}^{(aa)}_{t+1,t+1} &= \mathcal{L}^{(aa)}\cdot\hat{\omega}^{(aa)}_{t,t} + \mathsf{c}^{(aa)}_{t,t}
\end{aligned}$$

with $\hat{\omega}_{0,0} = \hat{\omega}^{(a)}_{0,0} = \hat{\omega}^{(aa)}_{0,0} = 0$ and $\mathsf{c}_{t,t}$, $\mathsf{c}^{(a)}_{t,t}$, and $\mathsf{c}^{(aa)}_{t,t}$ are all constructed from interactions of terms known to first order.

Proof. Adding together (70) and (71), we obtain

$$\begin{aligned}
\hat{\Omega}_{t+1,t+1}\langle a',\theta'\rangle = &\iint\sum_d \bar{\Lambda}^d(a',\theta',a,\theta)d\hat{\Omega}_{t,t}(a,\theta)\\
&- 2\frac{\partial}{\partial a'}\int\sum_d \bar{\Lambda}^d(a',\theta',a,\theta)\hat{a}^d_t(a,\theta)d\hat{\Omega}_t(a,\theta)\\
&+ 2\int\sum_d \delta\left(a' - \bar{a}^d(a,\theta)\right)\mu\left(\theta' - \rho_\theta\theta\right)\hat{p}^d_t(a,\theta)d\hat{\Omega}_t(a,\theta)\\
&+ \int\sum_d \delta\left(a' - \bar{a}^d(a,\theta)\right)\mu\left(\theta' - \rho_\theta\theta\right)\hat{p}^d_{t,t}(a,\theta)d\Omega^*(a,\theta)\\
&- 2\frac{\partial}{\partial a'}\int\sum_d \delta\left(a' - \bar{a}^d(a,\theta)\right)\mu\left(\theta' - \rho_\theta\theta\right)\hat{p}^d_t(a,\theta)\hat{a}^d_t(a,\theta)d\Omega^*(a,\theta)\\
&- \frac{\partial}{\partial a'}\int\sum_d \bar{\Lambda}^d(a',\theta',a,\theta)\hat{a}^d_{t,t}(a,\theta)d\Omega^*(a,\theta)\\
&+ \frac{\partial^2}{\partial a'^2}\int\sum_d \delta\bar{\Lambda}^d(a',\theta',a,\theta)\hat{a}^d_t(a,\theta)\hat{a}^d_t(a,\theta)d\Omega^*(a,\theta),
\end{aligned}$$

where $\hat{p}^d_t(a,\theta) = \bar{p}^d_\kappa(a,\theta)\hat{\kappa}_t(a,\theta)$ and

$$\hat{p}^d_{t,t}(a,\theta) = \bar{p}^d_\kappa(a,\theta)\hat{\kappa}_{t,t}(a,\theta) + \bar{p}^d_{\kappa\kappa}\hat{\kappa}_t(a,\theta)\hat{\kappa}_t(a,\theta).$$

From this we can conclude that $\hat{\Omega}_{t,t}$ has the structure of (72). To derive the recursions, we substitute

$\frac{\partial^2}{\partial a \partial \theta}\hat{\Omega}_{t,t}(a,\theta)$ and $\frac{\partial^2}{\partial a \partial \theta}\hat{\Omega}_{t}(a,\theta)$ into the above equation and apply integration by parts to find that $\hat{\omega}_{t,t}$, $\hat{\omega}_{t,t}^{(a)}$, and $\hat{\omega}_{t,t}^{(aa)}$ satisfy the above recursion with

$$\begin{aligned}
\mathsf{c}_{t,t}^{(aa)}(a',\theta') &= \iint \sum_d \bar{\Lambda}^d(a',\theta',a,\theta)\left(2\hat{a}_t^d(a,\theta)\bar{a}_a^d(a,\theta)\hat{\omega}_t^{(a)}(a,\theta) + \hat{a}_t^d(a,\theta)^2\omega^*(a,\theta)\right) da d\theta \\
\mathsf{c}_{t,t}^{(a)}(a',\theta') &= \iint \sum_d \bar{\Lambda}^d(a',\theta',a,\theta)\left[2\hat{a}_t^d(a,\theta)\frac{\hat{p}_t^d(a,\theta)}{\bar{p}^d(a,\theta)}\omega^*(a,\theta) + 2\hat{a}_t^d(a,\theta)\hat{\omega}_t(a,\theta)\right. \\
&\quad \left. + 2\left(\hat{a}_t^d(a,\theta)\frac{\bar{p}_a^d(a,\theta)}{\bar{p}^d(a,\theta)} + \bar{a}_a^d(a,\theta)\frac{\hat{p}_t^d(a,\theta)}{\bar{p}^d(a,\theta)} + \hat{a}_{t,a}^d(a,\theta)\right)\hat{\omega}_t^{(a)}(a,\theta)\right] da d\theta \\
\mathsf{c}_{t,t}(a',\theta') &= \iint \sum_d \bar{\Lambda}^d(a',\theta',a,\theta) \\
&\quad \cdot \left[\frac{\bar{p}_{\kappa\kappa}^d(a,\theta)\hat{\kappa}_t(a,\theta)^2}{\bar{p}^d(a,\theta)}\omega^*(a,\theta) + 2\frac{\hat{p}_t^d(a,\theta)}{\bar{p}^d(a,\theta)}\hat{\omega}_t(a,\theta) + 2\frac{\hat{p}_{ta}^d(a,\theta)}{\bar{p}^d(a,\theta)}\hat{\omega}_t^{(a)}(a,\theta)\right] da d\theta.
\end{aligned}$$

□

Our next step is to use the law of motion of the distribution to derive how the distribution depends on aggregate variables to second order. The next corollary shows that the coefficients of the second order terms are the same as the first order terms. The additional terms are all interactions of first order terms.

Corollary 3. *$\hat{\omega}_{t,t}$, $\hat{\omega}_{t,t}^{(a)}$, and $\hat{\omega}_{t,t}^{(aa)}$ can be expressed as*

$$\begin{aligned}
\hat{\omega}_{t,t} &= \sum_{s=0}^{\infty} \mathsf{A}_{t,s}\hat{Y}_{s,s} + \mathsf{C}_{t,t} \\
\hat{\omega}_{t,t}^{(a)} &= \sum_{s=0}^{\infty} \mathsf{A}_{t,s}^{(a)}\hat{Y}_{s,s} + \mathsf{C}_{t,t}^{(a)} \\
\hat{\omega}_{t,t}^{(aa)} &= \mathsf{C}_{t,t}^{(aa)},
\end{aligned} \tag{73}$$

where $\mathsf{A}_{t,s}$, $\mathsf{A}_{t,s}^{(a)}$ are the same as Corollary 1 and $\mathsf{C}_{t,t}$, $\mathsf{C}_{t,t}^{(a)}$, and $\mathsf{C}_{t,t}^{(aa)}$ are constructed from interactions of first order terms and satisfy

$$\begin{aligned}
\mathsf{C}_{t+1,t+1} &= \mathcal{L}\cdot\mathsf{C}_{t,t} + \mathcal{L}_{(a)}\cdot\mathsf{C}_{t,t}^{(a)} + \mathcal{L}_{(aa)}\cdot\mathsf{C}_{t,t}^{(aa)} + \mathcal{M}\cdot\mathsf{x}_{t,t} + \mathsf{c}_{t,t} \\
\mathsf{C}_{t+1,t+1}^{(a)} &= \mathcal{L}^{(a)}\cdot\mathsf{C}_{t,t} + \mathcal{L}_{(aa)}^{(a)}\cdot\mathsf{C}_{t,t}^{(aa)} + \mathcal{M}^{(a)}\cdot\mathsf{x}_{t,t} + \mathsf{c}_{t,t}^{(a)} \\
\mathsf{C}_{t+1,t+1}^{(aa)} &= \mathcal{L}^{(aa)}\cdot\mathsf{C}_{t,t} + \mathsf{c}_{t,t}^{(aa)},
\end{aligned}$$

with $\mathsf{C}_{0,0} = \mathsf{C}_{0,0}^{(a)} = \mathsf{C}_{0,0}^{(aa)} = 0$.

Proof. We begin by noting that Lemma 5 can be used to write

$$\hat{x}_{t,t} = \sum_{s=0}^{\infty} \mathsf{x}_{s-t} \hat{Y}_{s,s} + \mathsf{x}_{t,t}.$$

We also note that (73) holds with $t = 0$ because all the terms are zero. Proceeding by induction, we have that

$$\begin{aligned}
\hat{\omega}_{t+1,t+1} &= \mathcal{L} \cdot \hat{\omega}_{t,t} + \mathcal{L}_{(a)} \cdot \hat{\omega}_{t,t}^{(a)} + \mathcal{L}_{(aa)} \cdot \hat{\omega}_{t,t}^{(aa)} + \mathcal{M} \cdot \hat{x}_{t,t} + \mathsf{c}_{t,t} \\
&= \mathcal{L} \cdot \left(\sum_{s=0}^{\infty} \mathsf{A}_{t,s} \hat{Y}_{s,s} + \mathsf{C}_{t,t} \right) + \mathcal{L}_{(a)} \cdot \left(\sum_{s=0}^{\infty} \mathsf{A}_{t,s}^{(a)} \hat{Y}_{s,s} + \mathsf{C}_{t,t}^{(a)} \right) + \mathcal{L}_{(aa)} \cdot \left(\mathsf{C}_{t,t}^{(aa)} \right) \\
&\quad + \mathcal{M} \cdot \left(\sum_{s=0}^{\infty} \mathsf{x}_{s-t} \hat{Y}_{s,s} + \mathsf{x}_{t,t} \right) + \mathsf{c}_{t,t} \\
&= \sum_{s=0}^{\infty} \left(\mathcal{L} \cdot A_{t,s} + \mathcal{L}_{(a)} \cdot A_{t,s}^{(a)} + \mathcal{M} \cdot \mathsf{x}_{s-t} \right) \hat{Y}_{s,s} + \mathcal{L} \cdot C_{t,t} + \mathcal{L}_{(a)} \cdot C_{t,t}^{(a)} + \mathcal{L}_{(aa)} \cdot C_{t,t}^{(aa)} \\
&\quad + \mathcal{M} \cdot \mathsf{x}_{t,t} + \mathsf{c}_{t,t} \\
&= \sum_{s=0}^{\infty} \mathsf{A}_{t+1,s} \hat{Y}_{s,s} + \mathsf{C}_{t+1,t+1}.
\end{aligned}$$

The other recursions follow from the same argument. □

A key property of the recursions in Corollary 3 is that the coefficients of the second order terms are the same as the first order terms.

D.3 Market Clearing

Finally, the path of aggregates $\hat{X}_{t,t}$ is pinned down by taking the second derivative of (48) in direction $\hat{Z}_t$ and adding to it the derivative in direction $\hat{Z}_{t,t}$. Doing so yields

$$\mathsf{G}_Y \hat{Y}_{t,t} + \mathsf{G}_x \left[\left(\int \bar{x} d\Omega \right)_{ZZ} \cdot (\hat{Z}_t, \hat{Z}_t) + \left(\int \bar{x} d\Omega \right)_{Z} \cdot \hat{Z}_{t,t} \right] + \hat{\mathsf{G}}_{t,t} = 0, \tag{74}$$

where

$$\hat{\mathsf{G}}_{t,t} = \mathsf{G}_{YY} \cdot (\hat{Y}_t, \hat{Y}_t) + 2\mathsf{G}_{Yx} \cdot (\hat{Y}_t, \left(\int \bar{x} d\Omega\right)_Z \cdot \hat{Z}_t)$$
$$+ \mathsf{G}_{xx} \cdot \left(\left(\int \bar{x} d\Omega\right)_Z \cdot \hat{Z}_t, \left(\int \bar{x} d\Omega\right)_Z \cdot \hat{Z}_t\right) \tag{75}$$

contains all of the interactions of the second order terms. Applying Lemmas 5 and 6, we can show the following corollary

Corollary 4. *he second derivative of the aggregate individual decisions satisfies*

$$\left(\int \bar{x} d\Omega\right)_{ZZ} \cdot (\hat{Z}_t, \hat{Z}_t) + \left(\int \bar{x} d\Omega\right)_Z \cdot \hat{Z}_{t,t} = \sum_{s=0}^{\infty} \mathsf{J}_{t,s} \hat{Y}_{s,s} + \mathsf{H}_{t,t}, \tag{76}$$

where $\mathsf{H}_{t,t}$ is composed of interaction terms known to first order.

Proof. Taking derivatives, we have

$$\left(\int \bar{x} d\Omega\right)_{ZZ} \cdot (\hat{Z}_t, \hat{Z}_t) + \left(\int \bar{x} d\Omega\right)_Z \cdot \hat{Z}_{t,t}$$
$$= \int \sum_d \hat{x}^d_{t,t}(a,\theta) \bar{p}^d(a,\theta) d\Omega^*(a,\theta) + \int \sum_d \bar{x}^d(a,\theta) \hat{p}^d_{t,t}(a,\theta) d\Omega^*(a,\theta)$$
$$+ \int \sum_d \bar{x}^d(a,\theta) \bar{p}^d(a,\theta) d\hat{\Omega}_{t,t}(a,\theta) + 2\int \sum_d \hat{x}^d_t(a,\theta) \hat{p}^d_t(a,\theta) d\Omega^*(a,\theta)$$
$$+ 2\int \sum_d \hat{x}^d_t(a,\theta) \bar{p}^d(a,\theta) d\hat{\Omega}_t(a,\theta) + 2\int \sum_d \bar{x}^d(a,\theta) \hat{p}^d_t(a,\theta) d\hat{\Omega}_t(a,\theta).$$

Substituting for $\hat{\Omega}_{t,t}$ and applying integration by parts, we have

$$\int \sum_d \bar{x}^d(a,\theta) \bar{p}^d(a,\theta) d\hat{\Omega}_{t,t}(a,\theta)$$
$$= \int \sum_d \bar{x}^d(a,\theta) \bar{p}^d(a,\theta) \left(\hat{\omega}_{t,t}(a,\theta) - \frac{d}{da}\hat{\omega}^{(a)}_{t,t}(a,\theta) + \frac{d^2}{da^2}\hat{\omega}^{(aa)}_{t,t}(a,\theta)\right) da d\theta$$
$$\int \sum_d \bar{x}^d(a,\theta) \bar{p}^d(a,\theta) d\hat{\Omega}_{t,t}(a,\theta)$$
$$= \int \sum_d \bar{x}^d(a,\theta) \bar{p}^d(a,\theta) \left(\hat{\omega}_{t,t}(a,\theta) - \frac{\partial}{\partial a}\hat{\omega}^{(a)}_{t,t}(a,\theta) + \frac{\partial^2}{\partial a^2}\hat{\omega}^{(aa)}_{t,t}(a,\theta)\right) da d\theta$$
$$= \mathcal{I} \cdot \hat{\omega}_{t,t} + \mathcal{I}^{(a)} \cdot \hat{\omega}^{(a)}_{t,t} + \iint \sum_d \Big(\bar{x}^d_{aa}(a,\theta) \bar{p}^d(a,\theta)$$

$$+ 2\bar{x}^d_a(a,\theta)\bar{p}^d_a(a,\theta) + \bar{x}^d(a,\theta)\bar{p}^d_{aa}(a,\theta)\Big)\hat{\omega}^{(aa)}_{t,t}(a,\theta)dad\theta$$

$$= \mathcal{I} \cdot \left(\sum_{s=0}^{\infty} \mathsf{A}_{t,s}\hat{Y}_{s,s} + \mathsf{C}_{t,t}\right) + \mathcal{I}^{(a)} \cdot \left(\sum_{s=0}^{\infty} \mathsf{A}^{(a)}_{t,s}\hat{Y}_{s,s} + \mathsf{C}^{(a)}_{t,t}\right) + \mathcal{I}^{(aa)} \cdot \left(\mathsf{C}^{(aa)}_{t,t}\right),$$

where

$$\mathcal{I}^{(aa)} \cdot \mathsf{y} = \iint \sum_d \left(\bar{x}^d_{aa}(a,\theta)\bar{p}^d(a,\theta) + 2\bar{x}^d_a(a,\theta)\bar{p}^d_a(a,\theta) + \bar{x}^d(a,\theta)\bar{p}^d_{aa}(a,\theta)\right)\mathsf{y}(a,\theta)dad\theta.$$

In a similar manner, we have that

$$\int \sum_d \hat{x}^d_{t,t}(a,\theta)\bar{p}^d(a,\theta)d\Omega^*(a,\theta) = \sum_{s=0}^{\infty}\int\sum_d \mathsf{x}_{s-t}\bar{p}^d(a,\theta)d\Omega^*(a,\theta)\hat{Y}_{s,s} + \int\sum_d \mathsf{x}_{t,t}(a,\theta)\bar{p}^d(a,\theta)d\Omega^*(a,\theta)$$

and

$$\int \bar{x}^d(a,\theta)\hat{p}^d_{t,t}(a,\theta)d\Omega^*(a,\theta) = \sum_{s=0}^{\infty}\int\sum_d \bar{x}^d(a,\theta)\bar{p}^d_\kappa(a,\theta)\mathsf{p}_v(\mathsf{x}^b_{s-t}(a,\theta) - \mathsf{x}^w_{s-t}(a,\theta))d\Omega^*(a,\theta)\hat{Y}_{s,s} + \int\sum_d \bar{x}^d(a,\theta)\bar{p}^d_\kappa(a,\theta)\mathsf{p}_v(\mathsf{x}^b_{t,t}(a,\theta) - \mathsf{x}^w_{t,t}(a,\theta))d\Omega^*(a,\theta).$$

Putting these together implies (76) with

$$\begin{aligned}
\mathsf{H}_{t,t} &= \mathcal{I} \cdot \mathsf{C}_{t,t} + \mathcal{I}^{(a)} \cdot \mathsf{C}^{(a)}_{t,t} + \mathcal{I}^{(aa)} \cdot \mathsf{C}^{(aa)}_{t,t} + \int\sum_d \mathsf{x}_{t,t}(a,\theta)\bar{p}^d(a,\theta)d\Omega^*(a,\theta) \\
&+ \int\sum_d \bar{x}^d(a,\theta)\bar{p}^d_\kappa(a,\theta)\mathsf{p}_v(\mathsf{x}^b_{t,t}(a,\theta) - \mathsf{x}^w_{t,t}(a,\theta))d\Omega^*(a,\theta) + 2\int\sum_d \hat{x}^d_t(a,\theta)\hat{p}^d_t(a,\theta)d\Omega^*(a,\theta) \\
&+ 2\iint\sum_d \left(\hat{x}^d_t(a,\theta)\bar{p}^d(a,\theta) + \bar{x}^d(a,\theta)\hat{p}^d_t(a,\theta)\right)\hat{\omega}_t(a,\theta)dad\theta \\
&+ 2\iint\sum_d \left(\hat{x}^d_{ta}(a,\theta)\bar{p}^d(a,\theta) + \hat{x}^d_t(a,\theta)\bar{p}^d_a(a,\theta) + \bar{x}^d_a(a,\theta)\hat{p}^d_t(a,\theta) + \bar{x}^d(a,\theta)\hat{p}^d_{ta}(a,\theta)\right) \\
&\qquad \hat{\omega}^{(a)}_t(a,\theta)dad\theta.
\end{aligned}$$

□

This allows us to state our second order proposition.

Proposition 2. *The second order response $\hat{X}_{t,t}$ solves*

$$\mathsf{G}_Y \hat{Y}_{t,t} + \mathsf{G}_x \sum_{s=0}^{\infty} \mathsf{J}_{t,s} \hat{Y}_{s,s} + \mathsf{G}_x \mathsf{H}_{t,t} + \hat{\mathsf{G}}_{t,t} = 0,$$

where $\mathsf{P}\hat{X}_{-1,-1} = 0$ *and* $\lim_{t\to\infty} \hat{X}_{t,t} = 0$*.*

Proof. This is a direct consequence of Corollary 4 and equation (74). □

For details on how to construct numerical versions of the operators in the above proposition, see Bhandari et al. (2023).

E Adaptation of EGM Algorithm

We describe our adaptation of the Carroll's (2006) endogenous grid method (EGM) to compute the zeroth-order steady state in an occupation choice setting. We proceed in two steps. First we obtain a candidate solution using the algorithm we describe below. Then we check the candidate solution using brute-force grid search. If the candidate solution is not sufficiently accurate, we use the candidate solution as an initial guess for a more refined solution using grid search. In practice, we find that the candidate solution is very accurate and requires no further refinement.

The algorithm for the candidate solution proceeds as follows:

1. *Initialize marginal value of wealth.* Set an initial guess for the marginal value of wealth $\lambda(a,\theta) = \frac{\partial v}{\partial a}$, where $v(a,\theta)$ is the value function before the discrete choice.

2. *Solve the worker problem using EGM.* Apply the standard EGM algorithm to the worker's optimization problem:

   (a) *Forward step:* For each point a' on the asset grid, use the Euler equation to compute the implied consumption explicitly, then apply the budget constraint to determine the corresponding asset level today. Let a_0 denote the minimum implied asset level.

   (b) *Handle the borrowing constraint:* If $a_0 < \underline{a}$ (the borrowing limit), impose the borrowing constraint over the region $[a_0, \underline{a}]$ and solve for the constrained consumption and asset policies.

(c) *Store policy functions:* Interpolate the consumption policy and store the consumption function as well as the marginal value of wealth for workers, $\lambda^w(a, \theta)$, which can be obtained from the envelope condition.

3. *Solve the owner problem using modified EGM.* Repeat steps 2(a)–2(c) for the business owner, accounting for the additional production-related conditions. The mapping from consumption implied by the Euler equation to assets becomes implicit when the collateral constraint binds. Handle this efficiently by pre-computing the mapping and using fast lookup tables during EGM iterations. Store the consumption policy and the marginal value of wealth for owners, $\lambda^b(a, \theta)$, which includes the shadow value of relaxing the collateral constraint.

4. *Update value functions.* Using the consumption policies and marginal values of wealth from both occupations, compute the corresponding value functions and update the guess for the marginal value of wealth using the Bellman equation combined with the logsumexp formula and the derivative of the logsumexp for the discrete choice.

5. *Iterate to convergence.* Repeat steps 2-4 until the marginal value of wealth converges to the desired tolerance level.

In Figures 3 and 4, we plot the policy functions from the candidate solution and the brute-force grid search solution for our benchmark calibration. The two solutions are visually indistinguishable, confirming the accuracy of our EGM-based candidate solution.

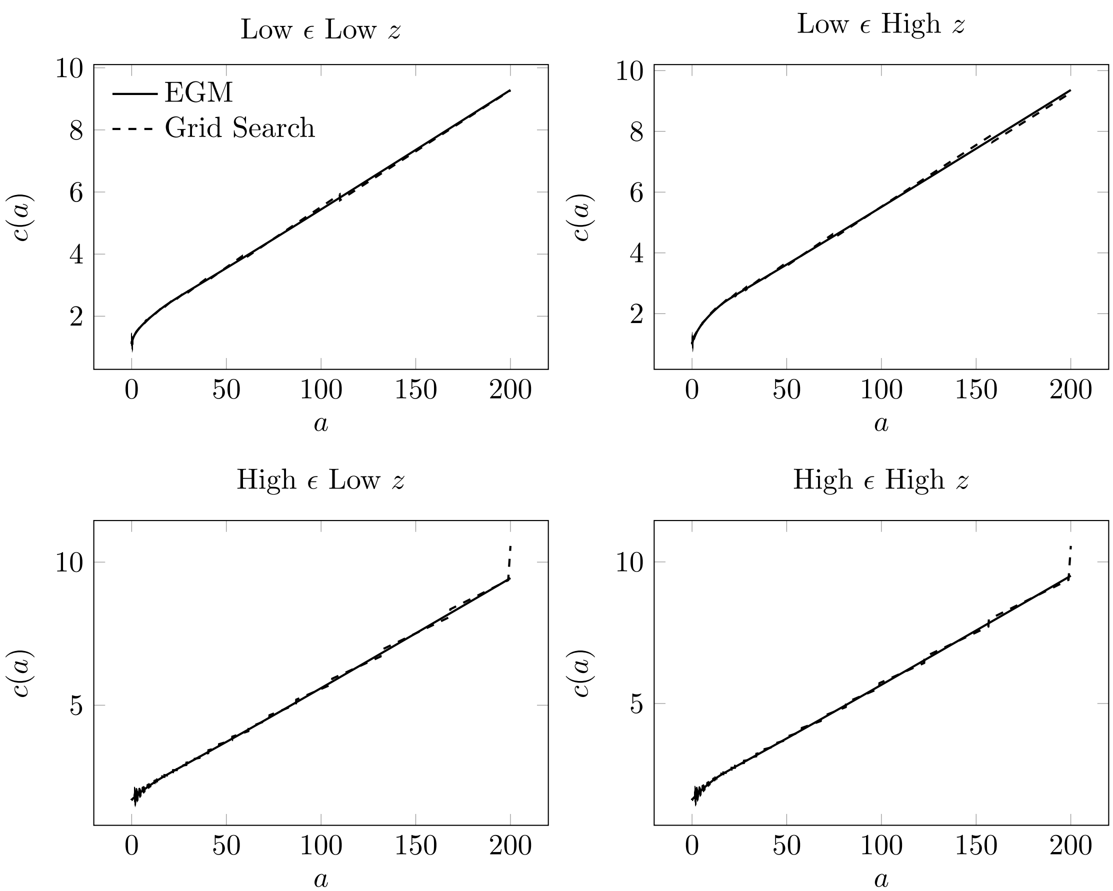


Figure 3: Consumption policies (workers): EGM versus grid search.

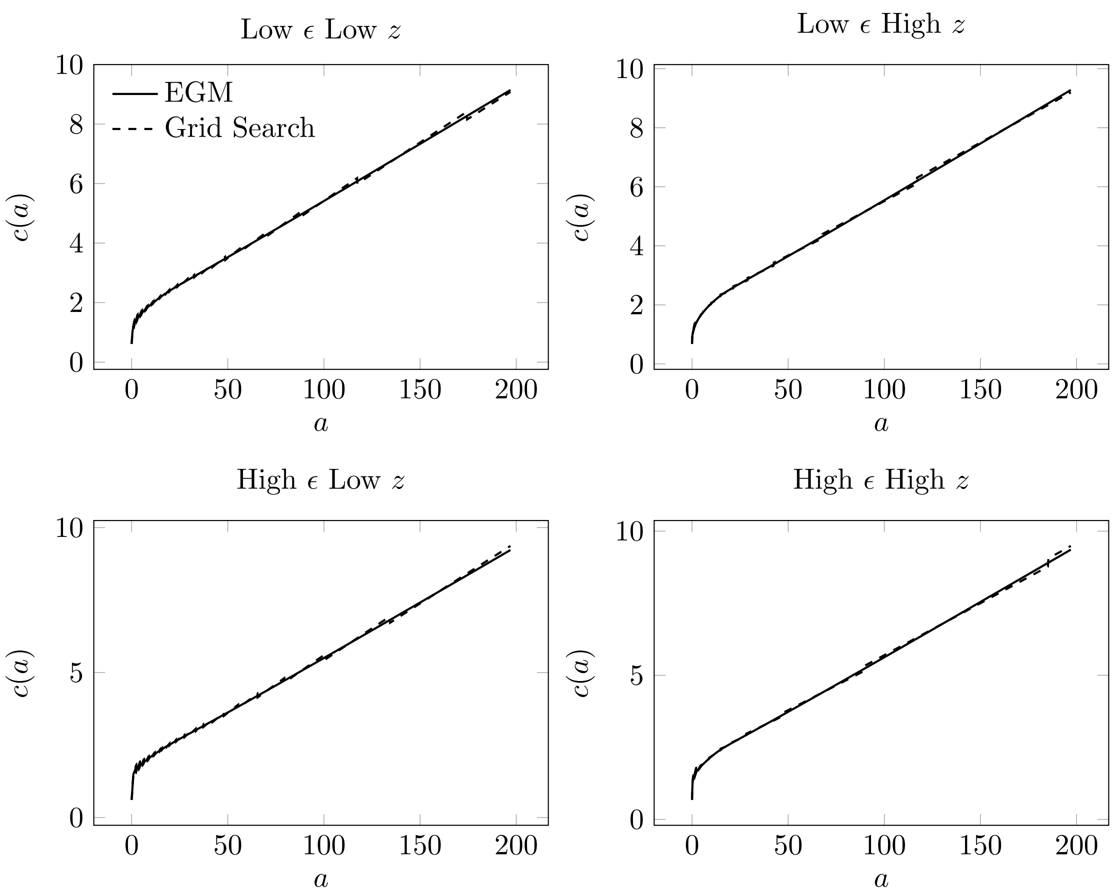


Figure 4: Consumption policies (business owners): EGM versus grid search.

Table 2: Model Parameters
A. Preferences, Technologies, Private and Government Financing

Parameter	Expression	Value
Preferences		
Relative risk aversion	μ	1.5
Discount factor	β	0.97
Growth	γ	0.02
Gumbel scale parameter	σ_η	0.40
Technologies		
Private business fixed asset share	ϕ	0.33
Private business labor share	ν	0.33
Corporate TFP	Θ	1.15
Corporate fixed asset share	α	0.45
Fixed asset depreciation	δ	0.041
Business financing		
Asset holding lower bound	$\underline{a}$	0
Maximum leverage	$\chi - 1$	0.25
Government financing		
Defense spending (level)	G	0.11
Debt (level)	B	3.00
Tax rate on labor income	τ_w	0.37
Tax rate on business income	τ_b	0.20
Tax rate on corporate income	τ_p	0.20
Tax rate on consumption	τ_c	0.06

F Calibration

In this section, we discuss choices of model parameters and the implied moments for the data generated by our model. These moments can be directly compared to U.S. national accounts and government budgets.

Table 2: Model Parameters (cont.)
B. Productivity Shocks

Productivity in t	Productivity in $t+1$				
Business owner, θ_t^b	0.432	0.657	1.000	1.522	2.317
0.432	0.612	0.170	0.098	0.065	0.055
0.657	0.172	0.551	0.187	0.064	0.025
1.000	0.099	0.191	0.475	0.190	0.045
1.522	0.060	0.055	0.164	0.558	0.164
2.317	0.046	0.009	0.034	0.135	0.776
Worker, θ_t^w	0.509	0.713	1.000	1.402	1.965
0.509	0.424	0.549	0.027	0	0
0.713	0.046	0.621	0.327	0.005	0
1.000	0.001	0.145	0.709	0.145	0.001
1.402	0	0.005	0.327	0.621	0.046
1.965	0	0	0.027	0.549	0.424

Notes: See Section F for details of the parameterization of the model.

F.1 Model Parameters

We turn now to parameterizing the model to be consistent with U.S. national accounts. We assume that the empirical analogue of total output in our model economy is U.S. GDP—which is split into output $\int y_{it}^b$ di of private businesses run by entrepreneurs, and all other output Y_c. In the United States, the private business sector includes sole proprietorships, partnerships, S corporations, and C corporations that are not publicly listed. Not included in the private business sector are publicly-traded corporate businesses and nonbusiness entities (specifically, government, non-profits, and households). We will continue to refer to the latter as "corporate" since publicly-traded companies are a large part of the output, but it should be noted that the Y_c excludes S corporate output and includes the nonbusiness output. On the product side of the national accounts, we assume that consumption C is private consumption expenditures plus non-defense spending on goods and services, and G is defense spending.

In Table 2, we list the parameters that we use in calibrating the model. These include parameters in the utility and production functions shown in equations (21)–(23) as well as the discount factor, growth rate, depreciation rate, financing parameters, government policy parameters, and processes governing productivities in entrepreneurship and paid employment. For preferences, we use a relative risk aversion parameter of 1.5 and a discount factor of 0.97 before detrending variables to account for growth at a rate of 2 percent.

A key target in our calibration is the semi-elasticity of aggregate labor demand N_b with respect to changes in the business tax rate τ_b. The mechanism is straightforward: when τ_b rises, the effective marginal cost of labor for firms increases, and labor demand contracts. In our model, the magnitude of this response is governed by the dispersion of idiosyncratic shocks, captured by the Gumbel scale parameter σ_η. A smaller σ_η implies sharper sorting across alternatives and hence a more elastic labor demand response, whereas a larger σ_η dampens the adjustment.

Empirical guidance comes from Giroud and Rauh (2019), who exploit variation in state corporate tax rates across the United States to estimate how firms adjust employment in response to taxation. Their baseline estimate implies that a one percentage point increase in the state corporate tax rate reduces employment by about 0.4 percent. Calibrating our model to this target implies a value of 0.40 for the Gumbel scale parameter σ_η, which delivers a model-implied semi-elasticity consistent with the estimates of Giroud and Rauh (2019).

For the private business sector production f^b, we assume equal shares for fixed assets, external labor, and owner income, which implies $\phi = \nu = 0.33$ and is roughly consistent with income shares for private businesses. For the corporate sector production f^c, we use a capital share of 0.45 and calibrate the corporate sector productivity shifter Θ to ensure the model generates an appropriate split between public and private business income. The depreciation rate for fixed assets is 4.1 percent, consistent with U.S. estimates for tangible capital.

The next set of parameters in Table 2 are related to private and government financing. The private financing depends on two parameters—namely, $\underline{a}$ and χ—that appear in the borrowing constraint in equation (2) and the collateral constraint in (6). All individuals in the model are subject to a non-negativity constraint on net worth, with $\underline{a} = 0$. For business owners, external financing is additionally limited by a collateral constraint, which requires that the capital they rent cannot exceed a multiple χ of their own net worth. This condition imposes an upper bound on leverage, with maximum leverage equal to $\chi - 1$. We set the baseline value of χ to 1.25 so that the implied steady-state ratio of business loans to GDP is in line with U.S. data.

For the government financing, we choose values for the level of defense spending G, debt B, and tax rates that ensure model budget shares and ratios to GDP are consistent with current combined budgets for federal, state, and local governments in the United States. We should note that the tax rate on corporate income is lower than estimates for the United States because within our "corporate" sector, we have included government enterprises, households, and non-profits that serve households.

The final set of parameters in Table 2 are the Markov chains governing the productivity processes reported in Bhandari and McGrattan (2021). The top panel shows the values and transition matrix for θ_t^b, and the bottom panel shows values and transition matrix for θ_t^w. To construct Π_θ, we compute a Kronecker product of the two transition matrices.

F.2 Steady State Moments

In Table 3, we report the model national accounts, government budget and variables of interest for the initial economy based on U.S. data. In the table, we use capital letters to denote aggregates of private business inputs and outputs, $K_b = \int k_i^b \,\mathrm{di}$, $N_b = \int n_i^b \,\mathrm{di}$, and $Y_b = \int y_i^b \,\mathrm{di}$; aggregate output, $Y = Y_c + Y_b$; aggregate capital, $K = K_c + K_b$; aggregate labor, $N = N_c + N_b$; total assets, $A = \int a_i \,\mathrm{di}$; total assets by occupation, $A_d = \int \mathbf{1}_{i \in d} a_i \mathrm{di}$, where $\mathbf{1}_{i \in d}$ is the occupational indicator; and business loans, $L_b = \int \mathbf{1}_{i \in b}(k_i^b - a_i) \,\mathrm{di}$, where $\mathbf{1}_{i \in b}$ indicates i runs a business. For convenience in writing expressions, we also replace $R-1$ with r. In Table 3, we include statistics related to the collateral constraints of business owners. We let $\#_{fc}/\#_b$ denote the share of business owners with $k_i^b = \chi a_i$. We let K_{fc} be the amount capital used by those constrained.

We report the national accounts for the model in the top panel of Table 3, starting with components of gross domestic income (GDI) and then moving on to gross domestic product.[11] Values in the third column are shares, which sum to 100 percent. Compensation to workers accounts for 49 percent of GDI but does not include labor earnings of business owners. We refer to the latter as "sweat" income but should note that it does not include any compensation for investing in business assets in the version of the model we analyze here. Sweat income in our baseline model is roughly 11 percent of GDI. The net operating surplus includes the taxable profits of corporations

[11] In the U.S. accounts, retail sales estimates used to construct personal consumption expenditures are recorded at purchaser prices and thus include sales tax. On the income side, the sales taxes are included with taxes on imports and production. For the model, we work with adjusted measures of GDI and GDP that do not include the sales taxes $\tau_c C$.

Table 3: Steady State Moments of the Model

Variable	Expression	Value
National Income and Product (% GDP)		
Gross domestic income	Y	100.0
Compensation	WN	48.5
Corporate	WN_c	38.8
Private business	WN_b	9.7
Sweat income	π	10.7
Net operating surplus	$Y - WN - \delta K - \pi$	25.8
Corporate	$Y_c - WN_c - \delta K_c$	20.5
Private business	$Y_b - WN_b - \delta K_b - \pi$	5.3
Consumption of fixed assets	δK	15.0
Gross domestic product	Y	100.0
Consumption	C	74.1
Government defense	G	3.6
Investment	$(\gamma + \delta)K$	22.3
Government Budget (% GDP)		
Revenues	Tax	28.6
Taxes on wages	$\tau_w WN$	18.0
Taxes on sweat income	$\tau_b(Y_b - (r+\delta)K_b - WN_b)$	2.1
Taxes on corporate income	$\tau_p(Y_c - WN_c - \delta K_c)$	4.1
Taxes on consumption	$\tau_c C$	4.4
Expenditures	$G + T + (r-\gamma)B$	28.6
Defense	G	3.6
Transfers	T	21.1
Net interest on debt	$(r-\gamma)B$	3.9
Variables of Interest		
Wealth-to-GDP ratio	A/Y	4.6
Business owners	A_b/Y	1.7
Workers	A_w/Y	2.9
Business loans-to-GDP ratio	L_b/Y	7.2
Share of owners constrained	$\#_{fc}/\#_b\%$	22.2
Share of capital constrained	$K_{fc}/K_b\%$	39.7

Notes: See Section 2 for details of the model and Section 4.1 for details of the algorithm to compute the steady state of the model.

and private business owners and accounts for 26 percent of GDI. The remaining income accounting for 15 percent of GDI is consumption of fixed assets. On the product side of the accounts, we have total consumption C—including non-defense public consumption—at 74 percent of GDP, defense spending at 4 percent, and total investment at 22 percent.

The second set of statistics in Table 3 include revenues and expenditures in the government accounts. The tax revenues total 29 percent of GDP and fund spending on defense, transfers—including non-defense spending—and net interest on the debt. The main revenue source in both the data and the model is taxation of employee wages and salaries, which are predicted here to be 18 percent of GDP. The tax revenue from business owners is much lower at 2.1 percent of GDP, in part because these owners are a much smaller fraction of the population and because higher rates of non-compliance in business implies lower effective tax rates. In Section 4.3.2, we consider policy reforms that increase the tax rate on business owners.

In the bottom panel of Table 3, we report several key aggregates for the initial economy. First, the wealth-to-GDP ratio is 4.6, with separate values shown for owners and workers. Finally, the choice of χ affects the extent to which owners are borrowing-constrained. In our baseline economy, about 22 percent of owners face binding collateral constraints, accounting for 40 percent of business capital.