Inequality, Business Cycles and Fiscal-Monetary Policy

Anmol Bhandari
bhandari@umn.edu

David Evans
devans@uoregon.edu

Mikhail Golosov
golosov@uchicago.edu

Thomas J. Sargent
thomas.sargent@nyu.edu

August 24, 2017

Abstract

We study fluctuations in macroeconomic aggregates and cross-section income and wealth distributions in a heterogeneous agent model with incomplete markets and sticky nominal prices. Optimal fiscal-monetary policy balances gains from "fiscal hedging" against benefits from "redistributional hedging" that responds to social concerns about inequality. A Ramsey planner uses inflation to offset inequality-increasing shocks to the cross-section distribution of labor earnings. A calibration that imitates how US recessions reshape that cross-section distribution in ways documented by Guvenen et al. (2014) indicates that the optimal response of tax rates, nominal interest rates and inflation are an order of magnitude larger in comparison to a representative agent benchmark.

Key words: Fiscal hedging, redistributional hedging, affine tax schedule, sticky prices, heterogeneity.

1 Introduction

The empirical labor literature has documented that dispersions of labor earnings, assets, and other measures of inequality comove with aggregate business cycle fluctuations. Meanwhile, the quantitative macroeconomic literature that studies the conduct of monetary and/or fiscal policy over the business cycle relies almost exclusively either on a representative agent assumption or on highly stylized and simplified models of heterogeneity. We want to know how their treatments of heterogeneity affect the reliability of the the quantitative prescriptions that arise from the existing optimal policy literature. The goal of this paper is to explore the conduct of the monetary and fiscal policy in a workhorse New Keynesian model augmented to fit rich heterogeneity across agents and other empirical facts regarding co-movements of aggregate variables and measures of inequality.

We study a New Keynesian economy populated by a continuum of heterogeneous agents who are subject to idiosyncratic wage risk. Agents differ in both the permanent and transitory components of shock processes that we calibrate to emulate dynamics of the distribution of U.S. labor earnings documented by Guvenen et al. (2014). Financial markets are incomplete and cannot provide perfect consumption smoothing. Agents differ in their equity holdings and their abilities to access financial markets. We study how the Ramsey planner optimally sets nominal interest rates, transfers and proportional labor taxes in response to aggregate shocks.

A key challenge is computational. Existing methods that study economies with heterogeneity, incomplete markets, and aggregate shocks mainly focus on approximating the ergodic state of a competitive equilibrium for a given set of policies and cannot easily be extended to answer normative questions. Ramsey problems entail additional challenges. First, because it contains Lagrange multipliers on individual and aggregate forward looking constraints, the part of state variables that summarizes the history dependence for the planner makes the state vector larger than what is needed to characterize ordinary competitive equilibria. Second, due to market incompleteness, some elements of the state space exhibit extremely slow rates of mean-reversion, implying that an approximation around the mean of the invariant distribution poorly approximates an optimal policy during a transition from a given distribution of state variables calibrated to data.

This paper contributes a new computational technique that allows us to optain good approximations to optimal government policies for economies with such large state spaces. Our numerical methods build on perturbation theory that uses small noise expansions with respect to a one-dimensional parameterization of uncertainty as in Fleming (1971), Fleming and Souganidis (1986) and its applications in economics as in Anderson et al. (2012)

among others. These are related to but differ from expansions in Judd and Guu (1993, 1997) and Judd (1996, 1998) that explore small noise expansions with respect to shocks and state variables about a deterministic steady state. We repeatedly take Taylor expansions of policy functions around a current state vector with respect to a parameter that scales both idiosyncratic and aggregate shocks. The current state vector can include a distribution of idiosyncratic states. We thus update the point around which local approximations are taken each period. This makes our approximations accurate even in settings where transition dynamics are slow.

To manage heterogeneity, we approximate the distribution of individual state variables using a discrete grid with a sufficiently large number of points. Our contribution here is to derive explicit formulas for coefficients occurring in the expansions of individual agents' and aggregate policy functions. We show that these formulas require matrix inversions of manageable dimensions, often equal to the number of *aggregate* variables, and that they can be efficiently computed. In this way, our procedure allows fast approximations even for a large number of agents and we can also construct impulse responses that account for how distributions across agents evolve after an aggregate shock. In Section 3.3 we elaborate on the specific steps in our algorithm and the how our method compares to other approaches found in the literature.

Applying our methods to a calibrated New Keynsian economy with heterogeneous agents, we find that attitudes about inequality induce the planner to use fiscal and monetary tools to redistribute towards agents who are especially adversely affected by a recession. Monetary policy ties nominal interest rates to real rates by keeping expected inflation near zero. While setting expected inflation close to zero, it is still desirable to have high unanticipated inflation in recessions because that redistributes from agents with high holdings of nominal bonds towards agents with low holding of nominal bonds. An optimal plan induces such surprise inflation by increasing the tax rate, which raise real wages and marginal costs for firms. Furthermore, as in data, recessions in our calibrated economy are accompanied by persistent increases inequality. This generates a motive to redistribute labor income from productive agents and to increase transfers. The planner achieves this by keeping marginal labor tax rates high long after output had recovered.

Quantitatively we find that in response to a productivity shock that lowers output growth by 3%, there is a nearly permanent increase in the labor tax rates of about 0.5 percentage points and a 0.15 percentage points jump in inflation for one period. As a point of comparison, the optimal tax rate and inflation in a economy without heterogeneity such as Siu (2004) and Schmitt-Grohe and Uribe (2004b) are an order of magnitude lower for similar shocks. The main driver for these differences is the Planner's desire to redistribute towards agents who are adversely affected by aggregate shocks. Lower nominal and real rates and higher transitory inflation transfers asset income from savers to borrowers and higher tax rates in part redistribute from productive agents to unproductive agents.

We begin by describing our model and some properties of the Ramsey allocation in Section 2. The numerical method and its comparison to alternative are discussed in Section 3. We use our method to obtain quantitative results in the calibrated economy in Section 4. Section 6 concludes.

2 Environment

A continuum of infinitely lived households face idiosyncratic shocks to their productivities. Individual i's preferences over stochastic processes for a final consumption good $\{c_{i,t}\}$ and labor supply $\{n_{i,t}\}$ are ordered by

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t U\left(c_{i,t}, n_{i,t}, \Theta_t\right)$$

where

$$U\left(c, n, \Theta\right) = \frac{c^{1-\nu}}{1-\nu} - \exp\left((1-\nu)\Theta\right) \frac{n^{1+\gamma}}{1+\gamma}, \tag{1}$$

$\mathbb{E}_t$ is a mathematical expectations operator conditioned on time t information and $\beta \in (0,1)$ is a time discount factor. With separable preferences, scaling the disutility of labor with $\exp\left((1-\nu)\Theta\right)$ keeps n stationary when the aggregate shocks have a stochastic trend.

The economy is subject to aggregate and idiosyncratic shocks. The aggregate shocks are labor productivity Θ_t and government expenditures G_t that follow stochastic processes given by

$$\Theta_t = \bar{\Theta} + \Theta_{t-1} + \mathcal{E}_{\Theta,t},$$

$$\frac{G_t}{\Theta_t} = \bar{G} + \mathcal{E}_{G,t},$$

where $\mathcal{E}_{\Theta,t}, \mathcal{E}_{G,t}$ are mean-zero, i.i.d. random variables. Aggregate and idiosyncratic shocks affect individual's labor productivity $\theta_{i,t}$ by the stochastic processes

$$\theta_{i,t} = \Theta_t + e_{i,t} + \varsigma_{i,t}, \tag{2}$$

$$e_{i,t} = \rho_e e_{i,t-1} + f_t\left(e_{i,t-1}\right) \mathcal{E}_{\Theta,t} + \eta_{i,t}, \tag{3}$$

where $\varsigma_{i,t}$, $\eta_{i,t}$ are also mean-zero i.i.d. random variables. This specification of idiosyncratic shocks builds closely on formulations used in labor literature, e.g. Storesletten et al. (2001), Low et al. (2010) where $\varsigma_{i,t}$ and $\eta_{i,t}$ correspond to transitory and persistent shocks to individual productivity. The function $f_t(e_{i,t-1})$ capture loading of aggregate shocks on individuals with different skills. In particular, it will allow us to match the cross-sectional implications of business cycle fluctuations as documented by Guvenen et al. (2014). We assume that all shocks take values on a compact set.

Agent i who works $n_{i,t}$ hours supplies $\exp(\theta_{i,t})n_{i,t}$ units of effective labor to a competitive labor market at nominal wage P_tW_t, where P_t is the nominal price of the final consumption good at time t. There is a common proportional labor tax rate τ_t and a common lump transfer T_tP_t. Agents trade a one-period risk-free nominal bond with each other and with the government. We use $P_tb_{i,t}$, P_tB_t to denote bond holdings of agent i and the government respectively, and ι_t,π_t to denote the nominal interest rate and inflation. Finally, D_t are dividends from intermediate goods producers measured in units of the final good. We assume in our baseline specification that they are distributed equally across households but will drop this assumption in Section 5. We take as given an initial price level $P_{-1} < \infty$ and set $\iota_{-1} = \beta^{-1} - 1$.

Agent i's budget constraint can be written as

$$c_{i,t} + b_{i,t} = (1-\tau_t)W_t\exp(\theta_{i,t})n_{i,t} + T_t + D_t + \left(\frac{1+\iota_{t-1}}{1+\pi_t}\right)b_{i,t-1}. \tag{4}$$

The government budget constraint at time t is

$$G_t + T_t + \left(\frac{1+\iota_{t-1}}{1+\pi_t}\right)B_{t-1} = \tau_t\int_i W_t\exp(\theta_{i,t})n_{i,t}di + B_t.$$

The final good Y_t is produced by competitive firms that use a continuum of intermediate goods $\{y_t(j)\}_{j\in[0,1]}$ in a production function

$$Y_t = \left[\int_0^1 y_t(j)^{\frac{\epsilon-1}{\epsilon}}dj\right]^{\frac{\epsilon}{\epsilon-1}}.$$

The final good producer takes the final good prices P_t and intermediate goods prices $\{p_t(j)\}_j$ as given and solves

$$\max_{\{y_t(j)\}_{j\in[0,1]}} P_t\left[\int_0^1 y_t(j)dj\right]^{\frac{\epsilon}{\epsilon-1}} - \int_0^1 p_t(j)y_t(j)dj. \tag{5}$$

Outcomes of optimization problem (5) are a demand function for intermediate goods

$$y_t(j) = \left(\frac{p_t(j)}{P_t}\right)^{-\epsilon} Y_t,$$

and a nominal price satisfying

$$P_t = \left(\int_0^1 p_t(j)^{1-\epsilon}\right)^{\frac{1}{1-\epsilon}}.$$

Intermediate goods $y_t(j)$ are produced by monopolists having linear technologies

$$y_t(j) = \tilde{n}_t(j),$$

where $\tilde{n}_t(j)$ is the amount of effective labor hired by firm j. These monopolists face downward sloping demand curves $\left(\frac{p_t(j)}{P_t}\right)^{-\epsilon} Y_t$ and choose prices $p_t(j)$ while bearing quadratic Rotemberg (1982) price adjustment costs $\frac{\exp(\Theta_t)\psi}{2}\left(\frac{p_t(j)}{p_{t-1}(j)} - 1\right)^2$ measured in units of the final consumption good. Firm j chooses prices $\{p_t(j)\}_t$ to solve

$$\max_{\{p_t(j)\}_t} \mathbb{E}_0 \sum_t \beta^t \left(\frac{C_t}{C_0}\right)^{-\nu} \left\{ \left[\frac{p_t(j)}{P_t} - W_t\right] \left(\frac{p_t(j)}{P_t}\right)^{-\epsilon} Y_t - \frac{\exp(\Theta_t)\psi}{2}\left(\frac{p_t(j)}{p_{t-1}(j)} - 1\right)^2 \right\}, \quad (6)$$

where for convenience we imposed that a firm values profit streams with a stochastic discount factor that is driven by aggregate consumption $C_t = \int c_{i,t} di$.[1]

In the symmetric equilibrium $p_t(j) = P_t$, $y_t(j) = Y_t$ for all j and market clearing conditions in labor, goods, and bond markets are:

$$y_t(j) = Y_t = \int \exp(\theta_{i,t}) n_{i,t} di \quad (7)$$

$$C_t + G_t = Y_t - \frac{\exp(\Theta_t)\psi}{2}\pi_t^2 \quad (8)$$

$$\int_i b_{i,t} di = B_t. \quad (9)$$

Definition 1. An *allocation* is a sequence $\{c_{i,t}, n_{i,t}\}_{i,t}$. A *bond profile* is a sequence $\left\{\{b_{i,t}\}_i, B_t\right\}_t$. A *price* system is a sequence $\{W_t, P_t\}_t$. A *fiscal-monetary policy* is a sequence $\{\tau_t, T_t, \iota_t\}_t$.

[1] In the economies with heterogeneous agents and incomplete markets one has to take a stand on how firms are valued. Using aggregate consumption to drive a stochastic discount factor process has the advantage that it allows us to get a representative agent economy as a special case of our heterogeneous agent economy by appropriately setting some of our parameters. This choice aligns with Kaplan et al. (2016).

Definition 2. Given an initial bond distribution $\left(\{b_{i,-1}, e_{i,-1}\}_i, B_{-1}\right)$ and initial price levels $p_{-1}(j) = P_{-1}$ for all j, a *competitive equilibrium* is a fiscal-monetary policy $\{\tau_t, T_t, \iota_t\}_t$ and a sequence $\left\{\{c_{i,t}, n_{i,t}, b_{i,t}\}_i, B_t, W_t, P_t\right\}_t$ such that (i) $\{c_{i,t}, l_{i,t}, b_{i,t}\}_{i,t}$ maximize (1) subject to (4) and natural debt limits; (ii) final goods firms choose $\{y_t(j)\}_j$ to maximize (5); (iii) intermediate goods producers prices solve (6) and their prices satisfy $p_t(j) = P_t$; and (iv) market clearing conditions (7), (8) and (9) are satisfied.

A utilitarian Ramsey planner orders allocations by

$$\mathbb{E}_0 \int \sum_{t=0}^{\infty} \beta^t \left[\frac{c_{i,t}^{1-\nu}}{1-\nu} - \exp\left((1-\nu)\Theta_t\right) \frac{n_{i,t}^{1+\gamma}}{1+\gamma} \right] di. \tag{10}$$

In Section 5, we also check the robustness of our results to alternative choices of Pareto weights.

Definition 3. A *Ramsey problem* seeks a monetary-fiscal policy that supports a competitive equilibrium allocation that maximizes (10). A maximizing monetary-fiscal policy is called a *Ramsey plan*; an associated allocation is called a *Ramsey allocation*.

2.1 Ramsey plan

As in Kydland and Prescott (1980) and Farhi (2010) we use firms' and household's optimality conditions to derive the implementability constraints and express the Ramsey problem as solution to two Bellman equations: a continuation planning problem for $t \geq 1$ and a $t = 0$ problem. These steps are standard and we relegate them to the Appendix A and state the $t \geq 1$ recursive problem here.

We use hats to denote variables scaled by aggregate productivity. For example, $\hat{c}_{i,t} = \frac{c_{i,t}}{\exp(\Theta_t)}$, $\hat{b}_{i,t} = \frac{b_{i,t}}{\exp(\Theta_t)}$, $\hat{T}_t = \frac{T_t}{\exp(\Theta_t)}$ and so on for other idiosyncratic and aggregate variables. The period utility function satisfies

$$U(c, n, \Theta) = \exp\left((1-\nu)\Theta\right) \left(\frac{\hat{c}^{1-\nu}}{1-\nu} - \frac{n^{1+\gamma}}{1+\gamma} \right) = \exp\left((1-\nu)\Theta\right) U(\hat{c}, n, 1)$$

We use $U_{\hat{c}}$ and U_n to denote the derivatives of $U(\hat{c}, n, 1)$ with respect to the first and second argument. The state variables for the $t \geq 1$ recursive formulation are marginal utility adjusted assets, $\hat{a}_{it} = \hat{c}_{i,t}^{-\nu} \hat{b}_{i,t}$, inverse marginal utility $\hat{m}_{i,t} = \hat{c}_{i,t}^{\nu}$, the persistent component of idiosyncratic shocks $e_{i,t}$, and the individual's Pareto weight ω_i. Let $z_{i,t} = (\hat{a}_{i,t}, \hat{m}_{i,t}, e_{i,t})$ and Z_t denote the distribution of individual states $z_{i,t}$ and it also represents the aggregate state for the Planner.

In the optimum, aggregate allocations in period t are functions of the previous period aggregate state Z_{t-1} and the aggregate shocks $\mathcal{E}_t = (\mathcal{E}_{\Theta,t}, \mathcal{E}_{G,t})$. The individual variables are functions of $(Z_{t-1}, \mathcal{E}_t)$, individual state z_{t-1} and idiosyncratic shocks $\varepsilon_t = (\eta_t, \varsigma_t)$. Let Φ and ϕ denote distributions of $\mathcal{E}_t$ and ε_t. We use the notation $\hat{a}(z)$ to refer to the first component of z and similarly for the rest. It will also be convenient to define $g(z, \varepsilon, \mathcal{E}) \equiv \exp\left[e + \varsigma + \eta + f(e, Z, \mathcal{E}_\Theta)\mathcal{E}_\Theta\right]$ as the productivity of an agent scaled by the aggregate productivity in state $(z, \varepsilon, \mathcal{E})$. Finally, let $\mathbb{E}_z$ be the expectation of individual variables conditional on z, Z. The Planners' problem can be written as

$$\hat{V}(Z) = \max_{\substack{\hat{c}, n, \hat{a}' \\ \hat{C}, W, \hat{Y}, \hat{D}, \tau, \hat{T}, \hat{Y}, \pi, \alpha}} \int e^{(1-\nu)\left[\bar{\Theta} + \mathcal{E}_\Theta\right]} w(z) \left\{ U\left(\hat{c}(z,\varepsilon,\mathcal{E}), n(z,\varepsilon,\mathcal{E}), 1\right) d\phi dZ + \beta \hat{V}(Z') \right\} d\Phi \tag{11}$$

subject to for all $(z, \varepsilon, \mathcal{E})$,

$$\begin{aligned} &\frac{\exp\left\{-\left[\bar{\Theta} + \mathcal{E}_\Theta\right]\right\} \hat{a}(z) U_{\hat{c}}(z,\varepsilon,\mathcal{E})(1+\pi(\mathcal{E}))^{-1}}{\beta \mathbb{E}_z[\exp\left\{-\nu\left[\bar{\Theta} + \mathcal{E}_{\Theta,t}\right]\right\} U_{\hat{c}}(1+\pi)^{-1}]} \\ &= U_{\hat{c}}(z,\varepsilon,\mathcal{E})\left[\hat{c}(z,\varepsilon,\mathcal{E}) - \hat{D}(\mathcal{E}) - \hat{T}(\mathcal{E})\right] + U_n(z,\varepsilon,\mathcal{E}) n(z,\varepsilon,\mathcal{E}) + \hat{a}'(z,\varepsilon,\mathcal{E}), \end{aligned} \tag{12}$$

$$\alpha = \hat{m}(z)\mathbb{E}_z[\exp\left\{-\nu\left[\bar{\Theta} + \mathcal{E}_\Theta\right]\right\} U_{\hat{c}}(1+\pi)^{-1}], \tag{13}$$

$$U_n(z,\varepsilon,\mathcal{E}) = -(1-\tau(\mathcal{E}))W(\mathcal{E})U_{\hat{c}}(z,\varepsilon,\mathcal{E})g(z,\varepsilon,\mathcal{E}), \tag{14}$$

and, for all $\mathcal{E}$

$$\hat{C}(\mathcal{E}) = \int \hat{c}(z,\varepsilon,\mathcal{E}) d\phi dZ, \tag{15}$$

$$\hat{Y}(\mathcal{E}) = \int n(z,\varepsilon,\mathcal{E}) g(z,\varepsilon,\mathcal{E}) d\phi dZ, \tag{16}$$

$$\hat{C}(\mathcal{E}) + \hat{G}(\mathcal{E}) = \hat{Y}(\mathcal{E}) - \frac{\psi}{2}\pi(\mathcal{E})^2, \tag{17}$$

$$\hat{D}(\mathcal{E}) = (1 - W(\mathcal{E}))\hat{Y}(\mathcal{E}) - \frac{\psi}{2}\pi(\mathcal{E})^2. \tag{18}$$

The distribution Z' is generated by $\hat{a}'$,$\hat{m}'$ and shocks $\varepsilon, \mathcal{E}$.

Equation (12) is obtained by dividing the budget constraint (4) by $\exp(\Theta_t)$, then multiplying with $U_{\hat{c}_i,t}$ and substituting for after tax income from (28). Equations (13) and (14) ensure that the intra and inter-temporal choices of all agents are equalized to aggregate prices. Equations (15) - (18) are aggregate market clearing conditions.

The Bellman equation (11) closely resembles the one obtained in the representative agent economies, except now the state variable is a distribution of individual state variables. With any realistic degree heterogeneity it is a highly dimensional object, which would make it impossible to solve this Bellman equation directly. Our approach is to approximate the solution to this problem using the first order conditions and not compute the Bellman equation.

We will use tilde to denote policy function to the Bellman equation (11). Let $\tilde{x}\,(z, Z, \varepsilon, \mathcal{E})$ and $\tilde{X}\,(Z, \mathcal{E})$ denote the vectors of individual and aggregate policy functions respectively, let $\tilde{z}(z, Z, \varepsilon, \mathcal{E})$ and $\tilde{Z}(Z, \mathcal{E})$ be the laws of motion for the individual and aggregate states respectively, and we let $\tilde{z}$ be a component of $\tilde{x}$. We use $\mathbb{E}_z\tilde{x}$ to denote the expectation of $\tilde{x}\,(z, Z, \cdot, \cdot)$ conditional on (z, Z) and $\mathbb{E}_{\tilde{z}}\tilde{\tilde{x}}$ the expectation of $\tilde{x}\left(\tilde{z}\,(z, Z, \epsilon, \mathcal{E})\,, \tilde{Z}(\Psi, \mathcal{E}), \cdot, \cdot\right)$ conditional on $\left(\tilde{z}, \tilde{Z}\right)$.

From problem (11), $\tilde{x}$ includes $\{\hat{c}, n, \hat{a}'\}$ and Lagrange multipliers on constraints (12) - (14), $\tilde{X}$ contains $\left\{\hat{C}, W, \hat{Y}, \hat{D}, \tau, \hat{T}, \hat{Y}, \pi, \alpha\right\}$and the Lagrange multipliers on (16) - (18). We use N_z, N_x, and N_X be numbers of elements in z, x, and X respectively. In our implementation we substitute (15) into (17), (18) into (26), and follow Marcet and Marimon (2011) to use the co-state variable in place of $\hat{a}(z)$ in z to end up with $N_z = 3$, $N_x = 6$ and $N_X = 4$. More details are in Appendix B.

We start with the set of individual constraints and first order conditions to the problem defined by (11) with respect to x. They can be written, for a certain vector-valued function F, as

$$F\left(z, \mathbb{E}_z\tilde{x}, \tilde{x}(z, Z, \varepsilon, \mathcal{E}), \mathbb{E}_{\tilde{z}}\tilde{\tilde{x}}, \tilde{X}(Z, \mathcal{E}), \varepsilon, \mathcal{E}\right) = \mathbf{0} \quad \text{for all } z, \varepsilon, \mathcal{E}. \tag{19}$$

Similarly, first order conditions with respect to $\tilde{X}$ along with market clearing constraints can be written compactly as

$$\int R(z, \tilde{x}(z, Z, \varepsilon, \mathcal{E}), \tilde{X}(Z, \mathcal{E}), \varepsilon, \mathcal{E}) d\phi dZ = \mathbf{0} \quad \text{for all } \mathcal{E}. \tag{20}$$

for some vector-valued function R. The explicit functional form of F and R are given in Appendix B. Our goal is to compute approximations to $\tilde{x}\,(z, Z, \varepsilon, \mathcal{E})$ and $\tilde{X}\,(Z, \mathcal{E})$ that satisfy (19) and (20) for an arbitrary Z.

3 Numerical Method

The starting point for our approach is perturbation theory using small noise expansions as in Fleming (1971), Fleming and Souganidis (1986). Consider a family of stochastic processes parameterized by a positive scalar σ, where all shocks $(\varepsilon, \mathcal{E})$ are scaled by σ. Let $\tilde{x}\,(z, Z, \sigma \cdot \varepsilon, \sigma \cdot \mathcal{E}; \sigma)$ and $\tilde{X}(Z, \sigma \cdot \mathcal{E}; \sigma)$ denote the policy functions when the scaling parameter is equal to σ. For our application we will also assume that auto-correlation of the persistent component of idiosyncratic shock is parametrized as $\rho_e = 1 - \hat{\rho}_e \sigma$. This is not essential but will simplify the exposition of our method.

Consider the second order Taylor expansion with respect to σ around $\sigma = 0$ at *a given state* Z:

$$\tilde{X}\left(Z, \sigma\mathcal{E}; \sigma\right) = \bar{X} + \sigma\left(X_{\mathcal{E}}\mathcal{E} + X_{\sigma}\right) + \frac{\sigma^2}{2}\left(\mathcal{E}^{\mathsf{T}} X_{\mathcal{E}\mathcal{E}}\mathcal{E} + 2X_{\mathcal{E}\sigma}\mathcal{E} + X_{\sigma\sigma}\right) + \mathcal{O}(\sigma^3) \tag{21}$$

and

$$\begin{aligned}\tilde{x}\left(z, Z, \sigma\varepsilon, \sigma\mathcal{E}; \sigma\right) = {}& \bar{x} + \sigma\left(x_{\mathcal{E}}\mathcal{E} + x_{\varepsilon}\varepsilon + x_{\sigma}\right) \\ & + \frac{\sigma^2}{2}\left(\mathcal{E}^{\mathsf{T}} x_{\mathcal{E}\mathcal{E}}\mathcal{E} + \varepsilon^{\mathsf{T}} x_{\varepsilon\varepsilon}\varepsilon + 2\mathcal{E}^{\mathsf{T}} x_{\mathcal{E}\varepsilon}\varepsilon + 2x_{\mathcal{E}\sigma}\mathcal{E} + 2x_{\varepsilon\sigma}\varepsilon + x_{\sigma\sigma}\right) + \mathcal{O}(\sigma^3).\end{aligned} \tag{22}$$

where $\bar{X} \equiv \tilde{X}\left(Z, 0; 0\right)$, $\bar{x} \equiv \tilde{x}\left(z, Z, 0, 0; 0\right)$, $X_{\mathcal{E}}, X_{\sigma}$ denote the derivatives of $\tilde{X}\left(Z, \sigma\mathcal{E}; \sigma\right)$ with respect to the second and third arguments evaluated at $\sigma = 0$, and the derivatives of $\tilde{x}$ and higher order derivatives defined analogously.

Our main contribution is to show that the right hand side of these expressions can be computed fast and efficiently even if dimensionality of the underline state Z is very large. In particular we show that (a) the "no uncertainty" terms $\bar{X}$, $\bar{x}$ are a solution to a relatively simple system of non-linear equations corresponding to a certain static economy, and (b) we derive explicit formulas for all the higher order terms $X_{\mathcal{E}}$,x_{ε},... that can be computed using linear algebra. Moreover, the computation of these derivatives involve inversion of matrices that depend only on the dimensionality of aggregate variables but not on the dimensionality of Z. Thus, even if Z is very large (in our numerical application we discretize Z with 20,000 elements) these derivatives can be computed very fast.

3.1 Step 1: computing points of expansion

Our next proposition simplifies the calculation of the 'zeroth-order' terms $\bar{X} \equiv \tilde{X}\left(Z, 0; 0\right)$, $\bar{x} \equiv \tilde{x}\left(z, Z, 0, 0; 0\right)$.

Proposition 1. *The individual and aggregate states are stationary in the non-stochastic limit, i.e.,*

$$\tilde{z}(z, Z, 0, 0; 0) = z \text{ and } \tilde{Z}(Z, 0; 0) = Z.$$

Therefore $\left(\bar{x}, \bar{X}\right)$ *is the solution to*

$$F\left(z, \bar{x}, \bar{x}, \bar{x}, \bar{X}, 0, 0\right) = \mathbf{0} \quad \text{for all } z, \tag{23}$$

$$\int R(z, \bar{x}, \bar{X}, 0, 0) dZ = \mathbf{0}. \tag{24}$$

The intuition for this result is straightforward. In absence of shocks, markets are complete and the pair $(\bar{x}, \bar{X})$ corresponds to a stationary economy where all households smooth their consumption, which implies that the aggregate state Z stays unchanged too. In our case, we can invert *(23)* to express $\bar{x}$ in terms of $\bar{X}$ and then use a standard numerical root solver to solve *(24)* as system of equations in N_X unknowns.[2] To calculate the integral in equation (24) we discretize state space Z into K bins, with $z_1, ..., z_K$ corresponding to the values that z takes on that grid, so that the distribution Z assigns equal measure to each bin, where K is chosen to be sufficiently large to approximate Z well.[3]

Once $(\bar{x}, \bar{X})$ is obtained, we evaluate first-, second- and higher-order derivatives of functions F and R at these allocations. These derivatives can be found efficiently using automatic differentiation routines. Let $F_0^k, F_{x-}^k, F_x^k, F_{x+}^k, F_X^k$ be the derivatives of F with respect to its first five arguments evaluated at $(z_k, \bar{x}, \bar{x}, \bar{x}, \bar{X}, 0, 0)$ for $k = 1, .., K$. Similarly, R_0^k, R_x^k, R_X^k are the derivatives of R with respect to its first three arguments evaluated at $(z_k, \bar{x}, \bar{X}, 0, 0)$. The higher order terms are analogously defined.

3.2 Step 2: Finding derivatives of policy functions

We now describe how to compute the derivatives of the policy functions $\tilde{x}$ and $\tilde{X}$. Given our discretization procedure, the state variable for $\tilde{X}$ is a $N_z \times K$-dimensional variable object, and the state variable for $\tilde{x}$ is a $N_z \times (K + 1)$ dimensional object. We use X_k, x_k^l to denote the derivatives of $\tilde{X}(Z, \cdot; \cdot)$ and $\tilde{x}(z_l, Z, \cdot, \cdot; \cdot)$ with respect to the k^{th} element on Z evaluated at $\sigma = 0$ and $z = z_l$; and x_0^l for $l \in \{1, ..., K\}$ to denote the derivative of $\tilde{x}$ with respect to the individual state z evaluated at $\sigma = 0$ and $z = z_l$. It will be convenient to define a matrix Q that selects state variables $\tilde{z}$ from the vector $\tilde{x}$, i.e., $\tilde{z} = Q\tilde{x}$ and we use z_k^l, z_0^l in a similar way to define the derivatives of $\tilde{z}$. The higher order derivatives of $\tilde{x}$ and $\tilde{X}$ are defined analogously.

Although only the derivatives of policy functions with respect to shocks $(\varepsilon, \mathcal{E})$ appear in the Taylor expansion (21) - (22), computing those derivatives requires also computing the derivatives of $\tilde{x}$ and $\tilde{X}$ with respect to the state variables (z, Z). This raises an immediate hurdle. Consider just the first order terms. A total differentiation of (19) and (20) will generate a linear system that determines $\{X_k, x_k^l, x_0^l\}_{k,l}$. This is a system with $KN_XN_z +$

[2] More generally one would need to use a shooting algorithm to compute the deterministic path for the $\sigma = 0$ allocation.

[3] In our application we set $K = 20,000$.

$K^2 N_x N_z + K N_x N_z$ unknowns that requires inverting a $K N_X N_z + K^2 N_x N_z + K N_x N_z$ size Jacobian matrix. Since the terms grow with K^2, such inversions are computationally infeasible for large K. Fortunately, the next lemma simplifies things.

Lemma 1. *For $z = z_l$, let x^l_X, x^l_0 be matrices of dimension $N_x \times N_X$ and $N_x \times N_z$ defined as*

$$x^l_X = -\left[F^l_{x-} + F^l_x + F^l_{x+}\right]^{-1} F^l_X,$$

$$.$$

The derivatives of the policy rules with respect to states $\left\{x^l_0, X_k, x^l_k\right\}_{k,l}$ are given by

$$\begin{aligned}
x^l_0 &= -\left[F^l_{x-} + F^l_x + F^l_{x+}\right]^{-1} F^l_0 \\
X_k &= \left(\sum_l \left(R^l_x x^l_X + R^l_X\right)\right)^{-1} \left[R^k_0 + R^k_x x^k_0\right] \\
x^l_k &= x^l_X X_k.
\end{aligned}$$

Lemma 1 is a critical step that makes our approach computationally feasible. It shows that a $\left\{x^l_k\right\}_{l,k}$ that has $N_z \times K^2$ elements can be decomposed into a product of two sets of terms, $\left\{x^l_X\right\}_l$ and $\{X_k\}_k$, each of which has $N_z \times K$ elements. Therefore the computational complexity grows linearly in K rather than K^2. Critically, the matrix inversions in the fomulas, for example x^l_X or X_k require inversions of $N_X \times N_X$ size matrices, which can be done quickly since the number of aggregate variables is small.

The economic intuition that yields this simplification comes from Evans (2015). The term x^l_k captures the effect on group of agents with state variable $z = z_l$ by a small change in the state variable of the group $z = z_k$. In our economy, the interaction between groups occurs only through aggregates such as prices and taxes. The means that the pairwise effects x^l_k can be broken down into how group k affects the aggregates X and how a change in aggregates X affect individuals in group l and ultimately, the number of computations scale linearly in the number of points required to approximate Z.

Using this lemma, we can compute the coefficients of the first order expansion of the policy functions.

Proposition 2. *The coefficients for the first order expansion of the policy function $\tilde{x}$ are given by: $x^l_\sigma = 0$,*

$$x^l_\varepsilon = -\left[F^l_x + F^l_{x+} x^l_0 Q\right]^{-1} F^l_\varepsilon,$$

$$x^l_{\mathcal{E}} = \left(x^l_{\mathcal{E},1} + x^l_{\mathcal{E},3}\left(I - \sum_k X_k Q x^k_{\mathcal{E},3}\right)^{-1}\left(\sum_k X_k Q x^k_{\mathcal{E},1}\right)\right)$$
$$+ \left(x^l_{\mathcal{E},2} + x^l_{\mathcal{E},3}\left(I - \sum_k X_k Q x^k_{\mathcal{E},3}\right)^{-1}\left(\sum_k X_k Q x^k_{\mathcal{E},2}\right)\right) X_{\mathcal{E}},$$

$$x^l_{\sigma} = \left(x^l_{\sigma,1} + x^l_{\sigma,3}\left(I - \sum_k X_k Q x^k_{\sigma,3}\right)^{-1}\left(\sum_k X_k Q x^k_{\sigma,1}\right)\right)$$
$$+ \left(x^l_{\sigma,2} + x^l_{\sigma,3}\left(I - \sum_k X_k Q x^k_{\sigma,3}\right)^{-1}\left(\sum_k X_k Q x^k_{\sigma,2}\right)\right) X_{\sigma}.$$

where

$$x^l_{\mathcal{E},1} = -\left(F^l_x + F^l_{x+} x^l_0 Q\right)^{-1} F^l_{\mathcal{E}}, \quad x^l_{\mathcal{E},2} = -\left(F^l_x + F^l_{x+} x^l_0 Q\right)^{-1} F^l_X, \quad x^l_{\mathcal{E},3} = -\left(F^l_x + F^l_{x+} x^l_0 Q\right)^{-1} F^l_{x+} x^l_X$$

$$x^l_{\sigma,1} = -\left(F^l_{x-} + F^l_x + F^l_{x+} + F^l_{x+} x^l_0 Q\right)^{-1} F^l_{\sigma}, \quad x^l_{\sigma,2} = -\left(F^l_{x-} + F^l_x + F^l_{x+} + F^l_{x+} x^l_0 Q\right)^{-1} F^l_X,$$
$$x^l_{\sigma,3} = -\left(F^l_{x-} + F^l_x + F^l_{x+} + F^l_{x+} x^l_0 Q\right)^{-1} F^l_{x+} x^l_X$$

and policy functions $\tilde{X}$ *are given by*

$$X_{\mathcal{E}} = -\left(\sum_k \left[R^k_X + R^k_x x^k_{\mathcal{E},2} + R^k_x x^k_{\mathcal{E},3}\left(I - \sum_l X_l Q x^l_{\mathcal{E},3}\right)^{-1}\left(\sum_l X_l Q x^l_{\mathcal{E},2}\right)\right]\right)^{-1}$$
$$\times \left(\sum_k \left[R^k_{\varepsilon} + R^k_x x^k_{\mathcal{E},1} + R^k_x x^k_{\mathcal{E},3}\left(I - \sum_l X_l Q x^l_{\mathcal{E},3}\right)^{-1}\left(\sum_l X_l Q x^l_{\mathcal{E},1}\right)\right]\right),$$

and

$$X_{\sigma} = -\left(\sum_k \left[R^k_X + R^k_x x^k_{\sigma,2} + R^k_x x^k_{\sigma,3}\left(I - \sum_l X_l Q x^l_{\sigma,3}\right)^{-1}\left(\sum_l X_l Q x^l_{\sigma,2}\right)\right]\right)^{-1}$$
$$\times \left(\sum_k \left[R^k_x x^k_{\sigma,1} + R^k_x x^k_{\sigma,3}\left(I - \sum_l X_l Q x^l_{\sigma,3}\right)^{-1}\left(\sum_l X_l Q x^l_{\sigma,1}\right)\right]\right).$$

Another message from Proposition 2 is that the formulas for the derivatives necessary for the approximations can be expressed in terms of matrices that are $N_x \times N_x$ or $N_X \times N_X$ and do not scale with the number of agents K. This allows us to handle very large state spaces.

With similar reasoning, we find higher-order derivatives with respect to $(z, Z, \varepsilon, \mathcal{E})$ and derivatives of the policy function with respect to σ. Conceptually, the number of terms required for higher order derivatives grows exponentially. For example, the second order derivatives with respect to the state variables $\left\{x^l_{jk}, X_{jk}, x^l_{0,k}\right\}_{k,l,j}$ would require solving for $K^3 N_x N_z + K^2 N_X N_z + K^2 N_x N_z$ terms, which grow at the rate K^3. We show that a counterpart of Lemma 1 can be established and computations of all the second order derivatives can be broken down into simpler terms that scale at most linearly with K and furthermore requires inversion of small-dimensional matrices. This logic extends and preserves the computational advantages of our approach for all higher-order expansions. The next proposition summarizes these findings and we list the detailed expressions for all the second-order terms in the Appendix C.3.

Proposition 3. *The terms* $\left\{x^l_{\varepsilon\varepsilon}, x^l_{\mathcal{E}\varepsilon}, x^l_{\varepsilon\mathcal{E}}, x^l_{\mathcal{E}\mathcal{E}}, x^l_{\sigma\sigma}, X_{\mathcal{E}\mathcal{E}}, X_{\sigma\sigma}\right\}$ *in the second order expansion can be expressed in terms of matrices of dimensions at most max* $\{N_x, N_X\}$ *by max* $\{N_x, N_X\}$ *and sums that involve at most* K *elements.*

The second order terms capture interesting economic behavior in addition to providing more accurate approximations. For instance, included in the derivative $x_{\sigma\sigma}$ is the fact that agents with different asset holdings save more or less in anticipation of future uninsurable shocks. In our economy, these self insurance motives not only affect individual savings behavior, but through market clearing conditions also change aggregate prices and optimal responses of government policies. Such aggregate responses are encoded in the $X_{\sigma\sigma}$ terms. Our formulas for second order derivatives combine these partial and general equilibrium effects of future risk on individual behavior. From a computational standpoint, these terms are important in capturing the evolution of the distributional state variable Z. In our quantitative Section 4, we show that ignoring second order terms significantly affects the responses of policies to aggregate shocks.

3.3 Comparison to other methods

Our method is related to approaches taken by Campbell (1998), Reiter (2009). These papers extend perturbation techniques of Judd and Guu (1993, 1997) and Judd (1996, 1998) to environments with heterogeneity. In those approaches, responses to aggregate shocks are approximated using first-order expansions of policy rules around a steady state Z^{SS} obtained by shutting down aggregate shocks.[4] We differ from these in two ways.

[4] Mertens and Judd (2013) approximate around a point of no heterogeneity. Ahn et al. (2017) extend the Reiter (2009) method to settings with continuous time, while Winberry (2016) uses a variant in which

First, we follow small noise expansions with respect to a one-dimensional parameterization of uncertainty, σ as in Fleming (1971), Fleming and Souganidis (1986) and Anderson et al. (2012). This means that our point of approximation is history dependent, Z_{t-1}, and we repeatedly perform Taylor expansions with respect to uncertainty as the economy moves along with aggregate shocks. There are many reasons why approximating around Z^{SS} is not desirable in our case. First, computing Z^{SS} in a Ramsey setting is difficult. One needs to jointly solve for agents behavior, which is endogenous to government policies, and optimal (deterministic) path for such policies using time consuming non-linear solution methods. Furthermore, in many settings with incomplete markets, the speed of mean-reversion to Z^{SS} is extremely low. The state variables are driven by martingale like dynamics and drift a lot over extremely low frequencies.[5] From a computational point of view, using perturbation around the fixed point provides a poor approximation for the optimal policy in many states.

Secondly, along the lines of Evans (2015), in Proposition 2 and 3 we are able to analytical characterize all the derivatives used in the small noise expansion in terms of matrices which are of small (typically $N_x \times N_x$) dimension. This allows us to perform higher order expansions in a fast and efficient way. As mentioned in the previous section, higher order terms are required to compute transition paths and accurate responses of aggregate variables to shocks as the underlying distribution of state variables drifts due to idiosyncratic risk.

Alternative to perturbation methods the literature has also used *projection methods* like Krusell and Smith (1998), Den Haan (1997), Algan et al. (2010). Projection methods summarize the infinite dimensional state variable using a subset of moments and approximate value functions and policy functions by using functional approximations and simulations for the aggregate laws of motion that describe the ergodic behavior of moments. These methods work well when individual state variables are low dimensional and economies exhibit what Krusell and Smith termed an "approximate aggregation" property in which a function of the first moment of the distribution of idiosyncratic states predicts next period's prices accurately.

In our application, we also use several groups of agents. These groups differ in the stochastic process of idiosyncratic skills and in their ability to participate in asset markets. For the optimal allocation, we need to keep track of keep track of a much more complicated distributional state variable that includes individual assets, past shocks, and Lagrange mul-

parametric forms capture the steady state distributions rather than the histograms used by Reiter (2009). Legrand and Ragot (2017) study an optimal fiscal policy problem with idiosyncratic risk and aggregate shocks after truncating individual histories.

[5] To give an extreme example, debt follows a random walk in a canonical incomplete market model of Barro (1979), so that the speed of the mean reversion is 0. Aiyagari et al. (2002) showed that a slow margtingale-like component is generally present in incomplete market economies. In Bhandari et al. (2017) we compute analytically the speed of mean reversion for several incomplete markets economies.

tipliers on all individuals' Euler equations for each group separately along with additional aggregate state variables. The state space for a problem like ours is too large to use projection methods. Furthermore, projection methods, like the perturbation methods cited before only approximate around the long run ergodic distribution and cannot compute transition paths.

4 Quantitative Application

We apply our equilibrium approximation algorithm to economies initially calibrated to recent U.S. data, assess quantitatively the properties of Ramsey policies, and contrast them with those of benchmark representative agent settings.

4.1 Calibration

We set $\nu = 1$ and $\gamma = 2$ to attain an intertemporal elasticity of substitution of 1 and a Frisch elasticity of labor supply equal to 0.5. We set the time discount factor $\beta = 0.96$, which implies that an annual net return in an economy would be 4% per year. Following Schmitt-Grohé and Uribe (2004a), we then set $\psi = 15$ to match the slope of the New Keynsian Phillip's curve as estimated by Sbordone (2002) and $\epsilon = 6$ to target a value-added markup of 0.2. We calibrate the mean and the standard deviation of the aggregate productivity growth rate process to 2% per year.

The parameters for the individual productivity process (2) - (3) are calibrated to match several moments of labor earnings. Since labor supply is endogenous, we simulate earnings from a competitive equilibrium of our model with constant labor tax, $\tau = 24\%$ and flexible prices. The variance of the transitory component, the auto-correlation and the variance of the persistent component are calibrated to match the variance of earnings and earnings growth at several horizons. We construct the process f_t to match some aspects of evidence in Guvenen et al. (2014) that expected earning drops in recessions vary with pre-recession earnings via a loadings $f_t(e_i)\mathcal{E}_\Theta$.

Figures I exhibit the fit of our model labor earnings. The right panel plots the change in variance of log earnings at horizon $t+h$ from a reference period t. As discussed in Storesletten et al. (2004), the slope and the curvature of the this variance curve is informative about drivers of the persistent component of log earnings. In addition, we match the standard deviation of change in log earnings that pins down the variance of the transitory component of workers productivity. In the left panel we simulate a recession that yields an aggregate output loss of 3% for one year. Following the procedure in Guvenen et al. (2014), we bin

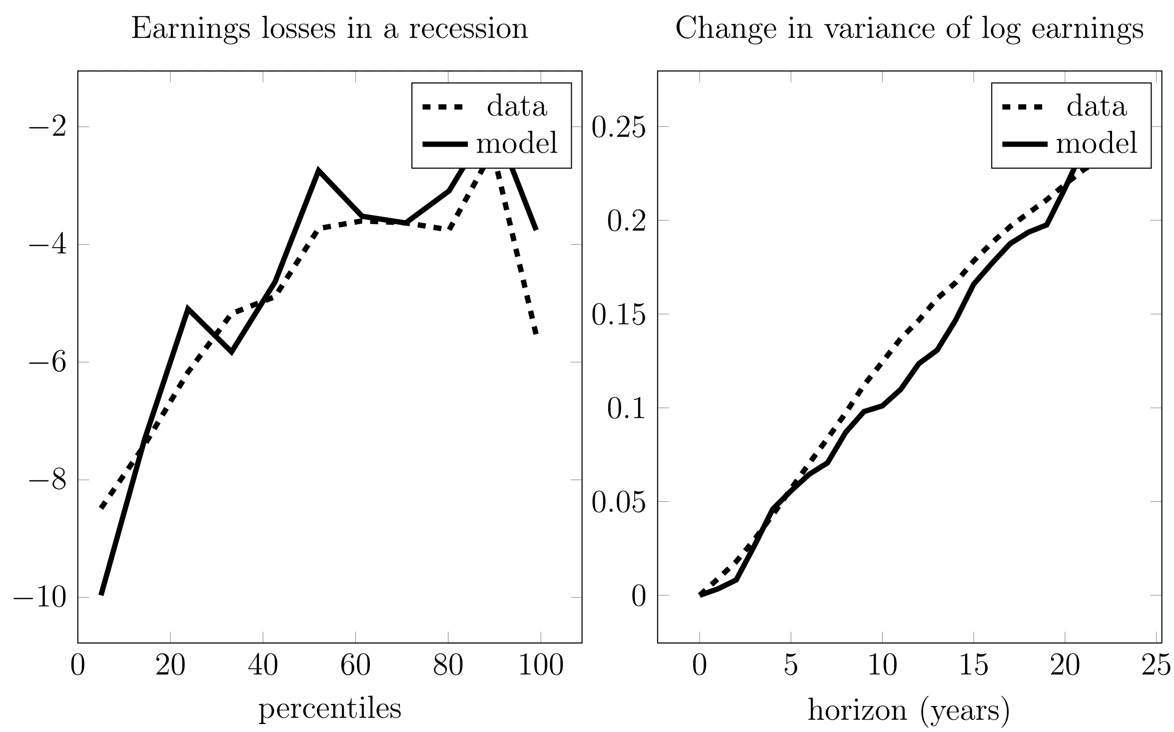

Figure I: The left panel plots annual earnings losses using simulated earnings from the model and data in Guvenen et al. (2014). The right panel plots change in variance of log earnings for several horizons using simulated earnings from the model and data from Guvenen et al. (2012)

workers into quantiles using their log earnings 5 years before the recession and measure the income loss for each bin. In both the panels the solid and the dashed lines compare the model simulations to data, respectively.

For initial debt and initial values for the persistent component of skills we use use data from the 2013 wave of the Survey of Consumer Finance (SCF) with the sample restricted to married households. The SCF provides information on households' total labor earnings, as well as hours worked by the primary and secondary earners. This lets us construct average household wages. We then measure households' holdings of government debt[6] and target wages and average bond holdings by wage quintiles. Parameters of our baseline model are in Table I

As a point of comparison, we take a Ramsey allocation in a representative agent economy with no transfers and a linear labor income tax, obtained as a special case of our economy by imposing $b_{1,-1} = B_{-1}$ and $T_t = 0$, setting $f_{1,t} = 0$, and to $\theta_{i,t} = \Theta_t$. This recovers what is

[6]We sum direct holdings plus indirect holdings through government bond mutual funds (taxable and nontaxable), saving bonds, money market accounts, and components of retirement accounts that are invested in government bonds.

Parameters	Values
Inter temporal elasticity, ν^{-1}	1
Frisch elasticity, γ^{-1}	0.5
Subjective discount factor, β	0.96
Std. deviation of $\mathcal{E}_\Theta$	2%
Mean growth rate $\bar{\Theta}$	2%
Mean expenditure to output ratio	12%
Std. deviation of expenditure to output ratio	3.5%
Std. of innovation to the AR(1) component, η	10
Autocorr. of AR(1) component, e_i	0.999
Std. i.i.d component ς	25%
Elasticity of substitution, ϵ	6
Menu cost, ψ	15
Government spending, G	12%

Table I: Parameters

almost a Schmitt-Grohé and Uribe (2004a) economy.[7]

4.2 Results

We first report results using our baseline calibration and then analyze our main findings by constructing special cases that turn off one feature at a time.

Baseline

Figure II depicts responses to a one-time, two-standard-deviation negative impulse to aggregate productivity $\mathcal{E}_\Theta$ occurring in period $t = 1$. This shock induces a drop in growth rate of output of about 3 percentage points on impact. Optimal policy in the baseline case (solid black line) responds by reducing the tax rate by 0.5 percentage point when the shock hits and then raising it by about 0.5 percentage points for a duration that lasts much longer than does the effect of the the shock on output growth. The effect on inflation is transitory: nominal prices rise by 0.15 percentage points on impact, then inflation returns to zero after one period. In terms of magnitudes, the movements in the tax rate and inflation are quite large relative to those in the representative agent model (dotted lines). The most significant

[7]This representative agent economy is also close to Siu's (2004) economy that uses a consumption tax instead of a proportional labor tax and and Correia et al.'s (2008) economy that uses both consumption and proportional labor taxes.

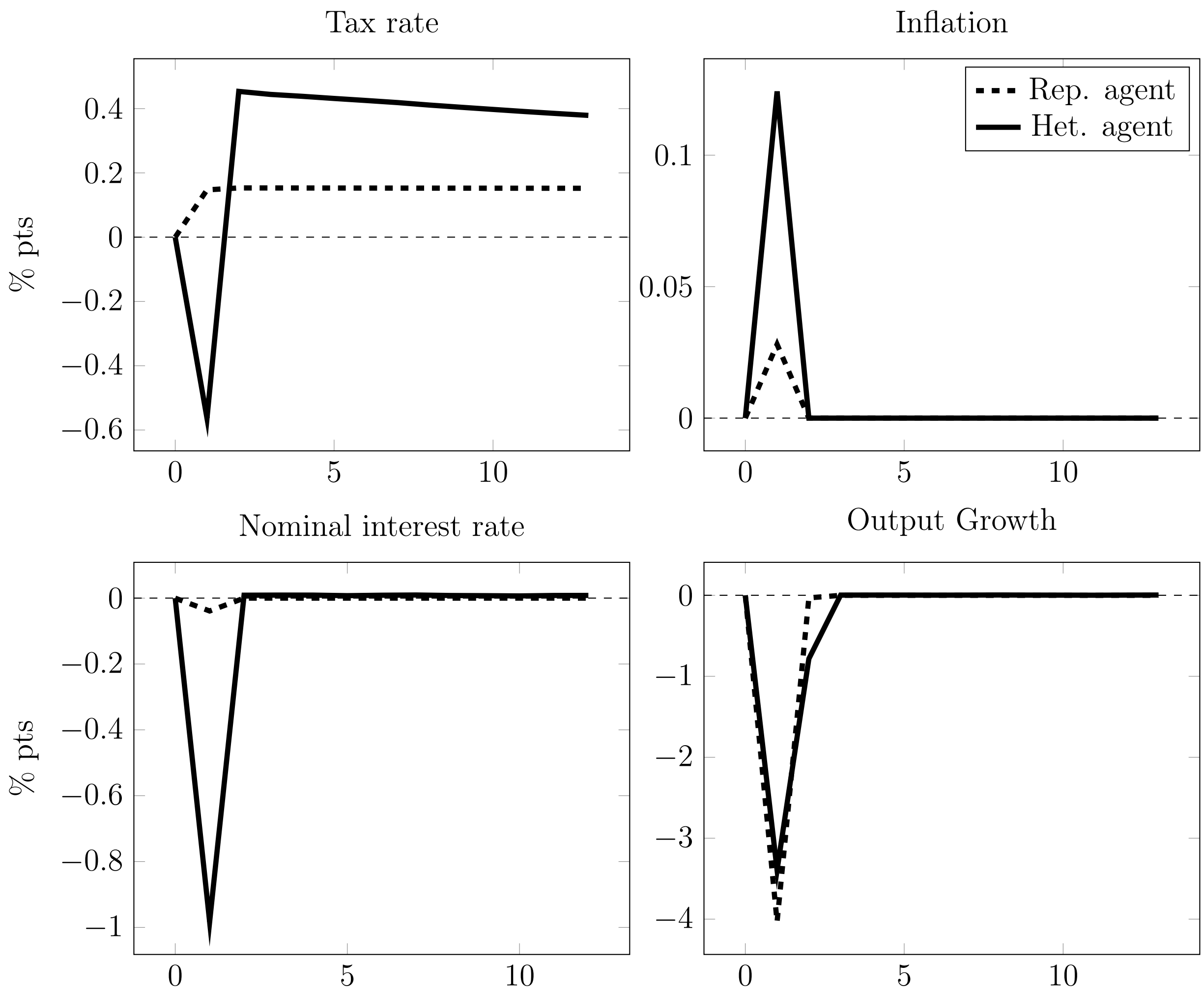

Figure II: Impulse response functions for the baseline calibration

difference is in behavior of the nominal rates. Nominal and real rates rates in the heterogeneous agent economy are about 100 basis point lower, where as they barely move in the representative agent case.

Redistributive hedging

Redistribution concerns explain the difference between our heterogeneous agent baseline and our representative agent comparison model in how optimal fiscal and monetary policies respond to a recession that is accompanied by an increase in inequality.

In the representative agent economy, there are no redistribution concerns, so the key margin confronting the planner is how to trade off benefits of using inflation to hedge aggregate shocks against the welfare costs of inflation volatility coming from the Rotemberg price adjustment costs. Our findings indicate that for business cycle sized shocks, welfare costs arising from accepting imperfect spanning of aggregate shocks are small relative to those that would come from making inflation more volatile. Accordingly, like Siu (2004) and Schmitt-Grohé and Uribe (2004a), we find small responses of the tax rate and smaller responses of inflation to a technology shock.

In our heterogeneous agent economy, a new force becomes active: with our setting of Pareto weights, optimal fiscal-monetary policy redistributes towards agents who are especially adversely affected by aggregate shocks. We call this "redistributive hedging". To see how it operates, consider a recession that increases inequality, as occurs in our baseline economy. The planner wants to redistribute both asset and labor income income, making it optimal to lower the real return on bonds. That can be achieved in two ways: either higher inflation which lowers the ex-post return or a lower contemporaneous tax rate combined with a higher future tax rates which lowers the ex-ante real rate. The planner also redistributes labor income by increasing the labor tax rate and giving larger transfers. Our findings suggest that redistributive hedging is quantitatively more important than the spanning vs inflation costs trade-off that takes center stage in the representative agent model.

As a comparison, we study an economy in which $\psi = 0$. As emphasized by Chari et al. (1991), without costs of inflation, the planner in a representative agent economy can implement a complete markets optimal plan: it sets the tax rate to be constant and generates the appropriate Ramsey real state-contingent returns by varying inflation systematically with respect to aggregate shocks. This raises inflation volatility. In a heterogeneous agent economy, the planner still uses inflation to obtain state-contingent returns, but here the motive is redistributive hedging and not state-contingent spanning. In the top panel of Figure III we see that when $\psi = 0$ the tax rate no longer dips on impact but stays high for a long time. The planner need not move the tax rate in order to lower real returns but can instead raise inflation to achieve the same outcome at a lower welfare cost. The inflation rate jumps by 8 percentage points as against 0.15 percentage points in our baseline heterogeneous agent economy.

Role of monetary policy

To isolate the role of monetary policy, we study optimal policy keeping tax rates fixed. This case is also interesting on political economy grounds that it is difficult to adjust fiscal policy at business cycle frequencies. The outcomes are shown in Figure IV. We see that the inflation response is larger than the baseline. Absent fiscal policy, the planner achieves the desired redistribution using higher inflation of about 30 basis points compared to 15 basis points in the previous case where both fiscal and monetary tools were available. The higher inflation is implemented by lowering interest rates on impact, stimulating the aggregate demand and

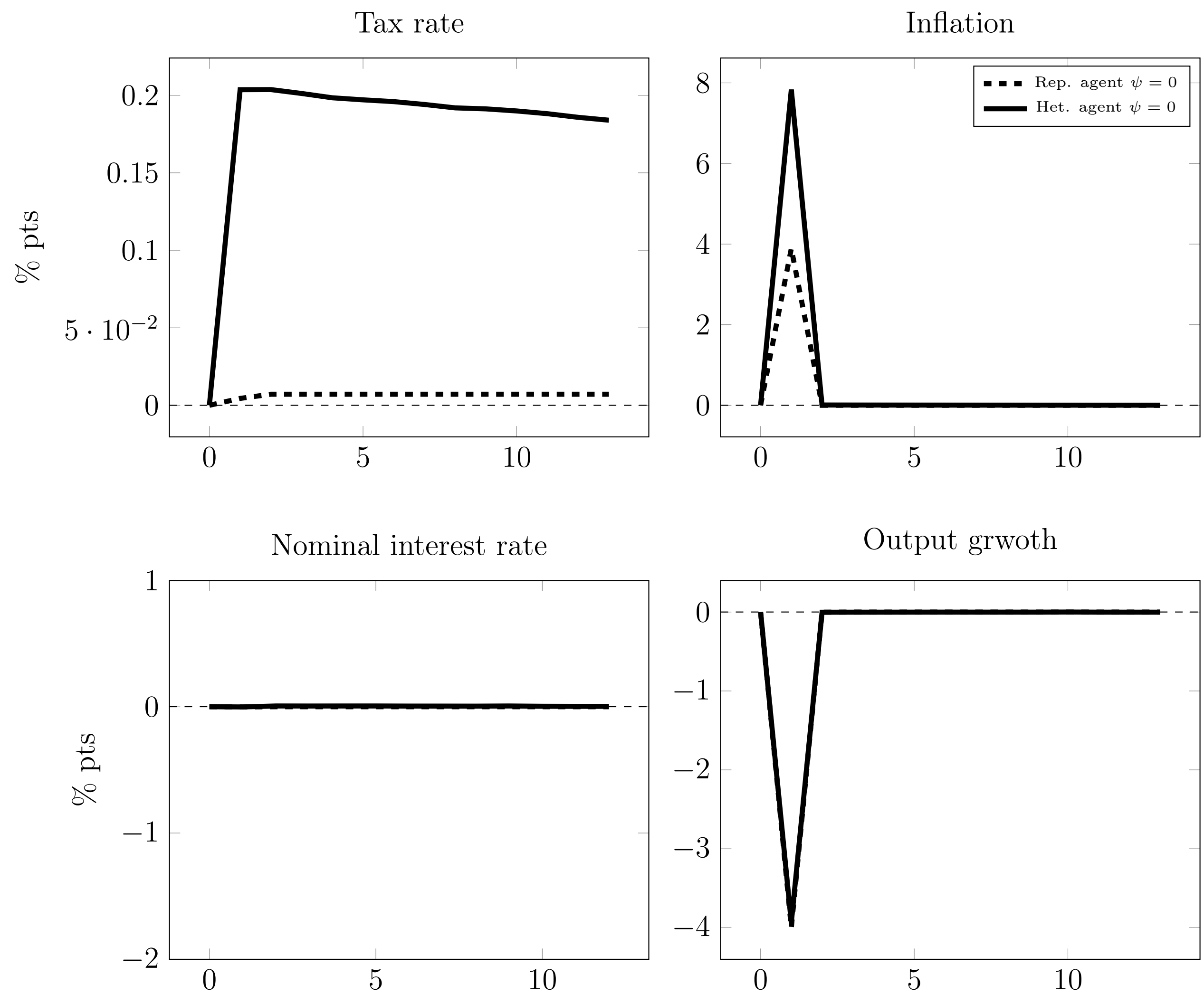

Figure III: Impulse response functions for flexible prices, $\psi = 0$.

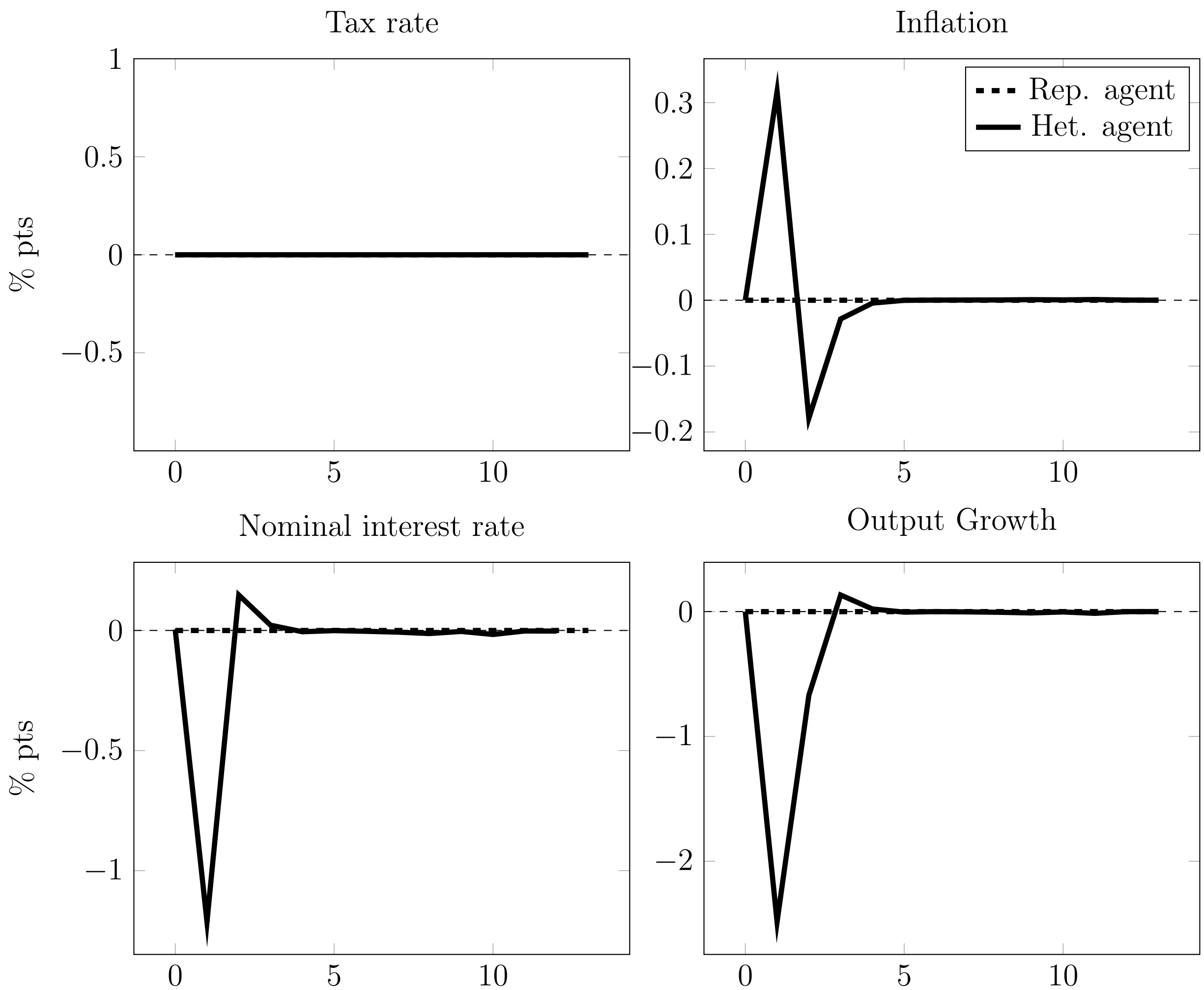

Figure IV: Impulse response functions for the baseline calibration with no fiscal policy

thereby inducing firms to raise prices. The planner also commits to higher interest rates in the future, which lower expected output and inflation. Together the combination of contractionary monetary policy and a promise to future tightening lowers real int rates rates which compensates partially the contemporaneous tax subsidy and higher future taxes that the planner used in the baseline case.

5 Extensions

[TBA]

6 Concluding remarks

James Tobin summarized macroeconomics as a field that explains aggregate quantities and prices while ignoring distribution effects. Tobin's characterization also describes much work subsequent to his in the real business cycle, asset pricing, Ramsey tax and debt, and New Keynesian research traditions. In each of these, the assumptions of complete markets and a representative consumer allow the analyst to compute aggregate quantities and prices without also determining distributions across agents.

This paper departs from such "aggregative economics" in two ways. First, we assume incomplete markets – which means that aggregate quantities, prices, and allocations across agents must be determined jointly, not recursively as in complete markets models. And second, we specify technology and relative skills shocks in a way that makes contact with findings of Guvenen et al. (2014) that US cross section distributions of labor earnings have moved systematically over business cycles. A common shock affects *both* an aggregate technology shock and the cross-section distribution of skills. Cross-section dispersions in labor earnings and asset holdings shape both aggregate outcomes and choices confronting a Ramsey planner.

Finally, an incomplete markets model goes a long way toward framing an optimal policy problem when it sets the menu of assets. By specifying that the only asset traded in our model is a risk-free *nominal* bond, we activate a beneficial role for fiscal and monetary policy to make nominal interest rates fluctuate in ways that hedge inequality-increasing shocks to distributions of labor earnings.[8]

[8]It is fruitful to compare our assumptions with those of Musto and Yilmaz (2003) who focus on how markets that allow citizens to insure outcomes of voting affect the efficacy of redistribution.

References

Ahn, SeHyoun, Greg Kaplan, Benjamin Moll, Thomas Winberry, and Christian Wolf. 2017. "When Inequality Matters for Macro and Macro Matters for Inequality." In *NBER Macroeconomics Annual 2017, volume 32.*: University of Chicago Press.

Aiyagari, S. Rao, Albert Marcet, Thomas J. Sargent, and Juha Seppala. 2002. "Optimal Taxation without State-Contingent Debt." *Journal of Political Economy*, 110(6): 1220–1254.

Algan, Yann, Olivier Allais, Wouter J Den Haan, and Pontus Rendahl. 2010. "Solving and simulating models with heterogeneous agents and aggregate uncertainty." *Handbook of Computational Economics*.

Anderson, Evan W, Lars Peter Hansen, and Thomas J Sargent. 2012. "Small noise methods for risk-sensitive/robust economies." *Journal of Economic Dynamics and Control*, 36(4): 468–500.

Barro, Robert J. 1979. "On the Determination of the Public Debt." *Journal of Political Economy*, 87(5): 940–971.

Bhandari, Anmol, David Evans, Mikhail Golosov, and Thomas Sargent. 2017. "Fiscal Policy and Debt Management with Incomplete Markets." *Quarterly Journal of Economics*.

Campbell, Jeffrey R. 1998. "Entry, exit, embodied technology, and business cycles." *Review of economic dynamics*, 1(2): 371–408.

Chari, V. V., Lawrence J. Christiano, and Patrick J. Kehoe. 1991. "Optimal Fiscal and Monetary Policy: Some Recent Results." *Journal of Money, Credit and Banking*, 23(3): 519–539.

Correia, Isabel, Juan Pablo Nicolini, and Pedro Teles. 2008. "Optimal fiscal and monetary policy: Equivalence results." *Journal of political Economy*, 116(1): 141–170.

Den Haan, Wouter J. 1997. "Solving dynamic models with aggregate shocks and heterogeneous agents." *Macroeconomic dynamics*, 1 355–386.

Farhi, Emmanuel. 2010. "Capital Taxation and Ownership When Markets Are Incomplete." *Journal of Political Economy*, 118(5): 908–948.

Fleming, Wendell H. 1971. "Stochastic Control for Small Noise Intensities." *SIAM Journal on Control*, 9(3): 473–517.

Fleming, Wendell H, and PE Souganidis. 1986. "Asymptotic series and the method of vanishing viscosity." *Indiana University Mathematics Journal*, 35(2): 425–447.

Guvenen, Fatih, Serdar Ozkan, and Jae Song. 2012. "The Nature of Countercyclical Income Risk." Working Paper 18035, National Bureau of Economic Research.

Guvenen, Fatih, Serdar Ozkan, and Jae Song. 2014. "The Nature of Countercyclical Income Risk." *Journal of Political Economy*, 122(3): 621–660.

Judd, Kenneth L. 1996. "Approximation, perturbation, and projection methods in economic analysis." *Handbook of computational economics*, 1 509–585.

Judd, Kenneth L. 1998. *Numerical methods in economics.*: MIT press.

Judd, Kenneth L, and Sy-Ming Guu. 1993. "Perturbation solution methods for economic growth models." In *Economic and Financial Modeling with Mathematica®.*: Springer, 80–103.

Judd, Kenneth L, and Sy-Ming Guu. 1997. "Asymptotic methods for aggregate growth models." *Journal of Economic Dynamics and Control*, 21(6): 1025–1042.

Kaplan, Greg, Benjamin Moll, and Giovanni L. Violante. 2016. "Monetary Policy According to HANK." Working Papers 1602, Council on Economic Policies.

Krusell, Per, and Anthony A Smith, Jr. 1998. "Income and wealth heterogeneity in the macroeconomy." *Journal of political Economy*, 106(5): 867–896.

Kydland, Finn E, and Edward C Prescott. 1980. "Dynamic optimal taxation, rational expectations and optimal control." *Journal of Economic Dynamics and Control*, 2(0): 79–91.

Legrand, Francois, and Xavier Ragot. 2017. "Optimal policy with heterogeneous agents and aggregate shocks : An application to optimal public debt dynamics."Technical report.

Low, Hamish, Costas Meghir, and Luigi Pistaferri. 2010. "Wage risk and employment risk over the life cycle." *The American economic review*, 100(4): 1432–1467.

Marcet, Albert, and Ramon Marimon. 2011. "Recursive contracts."

Mertens, Thomas M, and Kenneth L Judd. 2013. "Equilibrium existence and approximation for incomplete market models with substantial heterogeneity." *SSRN 1859650.*

Musto, David K., and Bilge Yilmaz. 2003. "Trading and Voting." *Journal of Political Economy*, 111(5): 990–1003.

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

Rotemberg, Julio J. 1982. "Monopolistic Price Adjustment and Aggregate Output." *The Review of Economic Studies*, 49(4): , p. 517.

Sbordone, Argia M. 2002. "Prices and unit labor costs: a new test of price stickiness." *Journal of Monetary Economics*, 49(2): 265–292.

Schmitt-Grohé, Stephanie, and Martın Uribe. 2004a. "Optimal fiscal and monetary policy under sticky prices." *Journal of economic Theory*, 114(2): 198–230.

Schmitt-Grohe, Stephanie, and Martin Uribe. 2004b. "Solving dynamic general equilibrium models using a second-order approximation to the policy function." *Journal of Economic Dynamics and Control*, 28(4): 755–775.

Siu, Henry E. 2004. "Optimal fiscal and monetary policy with sticky prices." *Journal of Monetary Economics*, 51(3): 575–607.

Storesletten, Kjetil, Chris I Telmer, and Amir Yaron. 2001. "How important are idiosyncratic shocks? Evidence from labor supply." *The American Economic Review*, 91(2): 413–417.

Storesletten, Kjetil, Christopher I Telmer, and Amir Yaron. 2004. "Consumption and risk sharing over the life cycle." *Journal of monetary Economics*, 51(3): 609–633.

Winberry, Thomas. 2016. "A Toolbox for Solving and Estimating Heterogeneous Agent Macro Models." *Forthcoming Quantitative Economics.*

A Appendix for deriving the scaled Bellman equation

After history $\{\mathcal{E}_{\Theta,s}, \mathcal{E}_{G,s}, \varsigma_{i,s}, \eta_{i,s}\}_{s=0}^{t-1}$ the continuation value for agent i for an allocation $\{c_{i,t}, n_{it}\}_t$ is

$$W_{t-1}^{HH}\left(\{c_{i,t}, n_{it}\}_t\right) \equiv \mathbb{E}_{t-1} \sum_{j=0}^{\infty} \beta^{t+j} U\left(c_{i,t+j}, n_{i,t+j}, \Theta_{t+j}\right).$$

Scaling $\hat{W}_{t-1}^{HH} = e^{-(1-\nu)\Theta_{t-1}} W_{t-1}^{HH}$, we then have

$$\hat{W}_{t-1}^{HH}\left(\{c_{i,t}, n_{it}\}_t\right) = \mathbb{E}_{t-1} e^{(1-\nu)\left[\bar{\Theta}+\mathcal{E}_{\Theta,t}\right]} \left[U\left(\hat{c}_t, n_t, 1\right) + \beta \hat{W}_t^{HH}\left(\{c_{i,t}, n_{it}\}_t\right)\right]. \tag{25}$$

Household budget constraint at date t with scaled variables is

$$\hat{c}_{i,t} + \hat{b}_{i,t} = (1-\tau_t) n_{i,t} W_t \exp(\theta_{i,t} - \Theta_t) + \hat{T}_t + \hat{D}_t + \left(\frac{1+\iota_{t-1}}{1+\pi_t}\right) \exp\left\{-\left[\bar{\Theta} + \mathcal{E}_{\Theta,t}\right]\right\} \hat{b}_{i,t-1}. \tag{26}$$

For a given fiscal-monetary policy, the household maximizes $\hat{W}_0^{HH}\left(\{c_{i,t}, n_{it}\}_t\right)$ subject to (26) for all $t \geq 0$ and natural debt limits. The optimality conditions are

$$\mathbb{E}_{t-1}\left\{\beta\left(\frac{1+\iota_{t-1}}{1+\pi_t}\right) \exp\left\{-\nu\left[\bar{\Theta} + \mathcal{E}_{\Theta,t}\right]\right\} \frac{U_{\hat{c}_{it}}}{U_{\hat{c}_{it-1}}}\right\} = 1, \tag{27}$$

$$U_{n_{i,t}} = -(1-\tau_t) W_t U_{\hat{c}_{it}} \exp\left(\theta_{i,t} - \Theta_t\right). \tag{28}$$

The firms' optimal pricing and symmetry after dividing by $\exp(\Theta_t)$ gives us

$$\left(\frac{\hat{Y}_t\left[1-\epsilon\left(1-W_t\right)\right]}{\psi}\right) - \pi_t(1+\pi_t) + \beta \mathbb{E}_t \exp\left((1-\nu)\left[\bar{\Theta} + \mathcal{E}_{\Theta,t+1}\right]\right)\left(\frac{\hat{C}_{t+1}}{\hat{C}_t}\right)^{-\nu} \pi_{t+1}(1+\pi_{t+1}) = 0. \tag{29}$$

The market clearing conditions using the scaled variables are

$$\hat{G}_t + \hat{T}_t + \left(\frac{1+\iota_{t-1}}{1+\pi_t}\right) \exp\left\{-\left[\bar{\Theta} + \mathcal{E}_{\Theta,t}\right]\right\} \hat{B}_{t-1} = \tau_t \int_i n_{i,t} \exp(\theta_{i,t} - \Theta_t) W_t di + \hat{B}_t. \tag{30}$$

$$\hat{Y}_t = \int n_{i,t} \exp(\theta_{i,t} - \Theta_t) di, \tag{31}$$

$$\hat{C}_t + \hat{G}_t = \hat{Y}_t - \frac{\psi}{2}\pi_t^2, \tag{32}$$

$$\int \hat{b}_{i,t} di = \hat{B}_t, \tag{33}$$

$$\hat{D}_t = (1 - W_t)\hat{Y}_t - \frac{\psi}{2}{\pi_t}^2. \tag{34}$$

The planner maximizes (10) subject to (26) - (34). The next lemma shows that (29) and (34) are slack at the optimal allocation.

Lemma 2. *Constraint (29) and (34) do not bind at the optimal allocation.*

Proof. Consider an allocation, bond profile, price system, and fiscal-monetary policy

$$\{\hat{c}_{i,t}, n_{i,t}\}_{i,t}, \left\{\left\{\hat{b}_{i,t}\right\}_i, \hat{B}_t\right\}_t, \{W_t, R_t, P_t\}_t, \{\tau_t, \hat{T}_t, i_t\}$$

that satisfies constraints (26) - (34) except (29) and (34). We can construct an alternative price system and fiscal-monetary policy that does satisfy all equations (26) - (34) and attains the same value to the Planner.

First, choose a sequence $\{\check{W}_t\}_t$ that makes constraint (29) satisfied. Then choose $\left\{\check{\tau}_t, \check{T}_t, \check{D}_t\right\}$ so that

$$(1 - \check{\tau}_t)\,\check{W}_t = (1 - \tau_t)\,W_t$$

$$\check{T}_t + \check{W}\hat{Y}_t = T_t + W_t\hat{Y}_t.$$

$$\check{D}_t = (1 - \check{W}_t)\hat{Y}_t - \frac{\psi}{2}{\pi_t}^2$$

Evidently $\{\hat{c}_{i,t}, n_{i,t}\}_{i,t}, \left\{\left\{\hat{b}_{i,t}\right\}_i, \hat{B}_t\right\}_t, \{\check{W}_t, R_t, P_t\}_t, \{\check{\tau}_t, \check{T}_t, i_t\}$ satisfies (26) - (34) and is thus implementable. Furthermore, since the allocation, $\{\hat{c}_{i,t}, n_{i,t}\}_{i,t}$ is unchanged, the value that the planner assigns to the equilibrium allocation is also unchanged. □

B First Order Conditions for the $t \geq 1$ Planner's problem

For brevity we will use s to denote the joint of states and shocks $(z, \varepsilon, \mathcal{E})$. Let $\mu(s), \rho(s), \phi(s)$ be the multipliers on the individual constraints (12) - (14) and $\chi(\mathcal{E}), \xi(\mathcal{E}), \lambda(\mathcal{E})$ be the multipliers on the aggregate constraints (15) - (17).

Lemma 2 shows that (18) does not bind and hence $W(\mathcal{E})$ only appears in the form of term $(1 - \tau(\mathcal{E}))W(\mathcal{E})$ and $\hat{D}$ in the form of term $\hat{D} + \hat{T}$. Similarly the state variable $\hat{m}$ only enters in equation (13). This means that in order to solve for the optimal allocation we need to find the product $(1 - \tau(\mathcal{E}))W(\mathcal{E})$ which we denote by $\mathrm{W}(\mathcal{E})$, the sum $\hat{D} + \hat{T}$ denoted by $\hat{\mathrm{T}}$ and scale $\hat{m}'$ by a constant of proportionality such that $\hat{m}'(s)U_{\hat{c}}(s) = \aleph(\mathcal{E})$ and

$$\int \hat{m}'(s)d\phi dZ = 1.$$

Following Marcet and Marimon (2011) we also replace $\hat{a}$ in z with a transformation of its associated co-state variable $\hat{\mu} \equiv \frac{\mathbb{E}_z\left[\exp\left\{-\left[\bar{\Theta}+\mathcal{E}_\Theta\right]\right\}\hat{U}_c(1+\pi)^{-1}\mu\right]}{\beta\mathbb{E}_z[\exp\left\{-\nu\left[\bar{\Theta}+\mathcal{E}_\Theta\right]\right\}\hat{U}_c(1+\pi)^{-1}]}$.

The list of equations that comprise functions F are

$$\begin{aligned}
0 =& \frac{\exp\left\{-\left[\bar{\Theta}+\mathcal{E}_\Theta\right]\right\}\hat{a}\,(s)\,U_{\hat{c}}(s)(1+\pi(\mathcal{E}))^{-1}}{\beta\mathbb{E}_z[\exp\left\{-\nu\left[\bar{\Theta}+\mathcal{E}_\Theta\right]\right\}U_{\hat{c}}(1+\pi)^{-1}]} \\
&- U_{\hat{c}}(s)\left[\hat{c}(s)-\hat{\mathrm{T}}(\mathcal{E})\right] - U_n(s)n(s) + \hat{a}(s'), && \text{(35a)}\\
0 =& -\alpha + \hat{m}(z)\mathbb{E}_z[\exp\left\{-\nu\left[\bar{\Theta}+\mathcal{E}_\Theta\right]\right\}U_{\hat{c}}(1+\pi)^{-1}], && \text{(35b)}\\
0 =& -U_n(s) - \mathrm{W}(\mathcal{E})U_{\hat{c}}(s)g(s), && \text{(35c)}\\
0 =& -\aleph(\mathcal{E}) + \hat{m}'(s)U_{\hat{c}}(s), && \text{(35d)}\\
0 =& \frac{\exp\left\{-\nu\left[\bar{\Theta}+\mathcal{E}_\Theta\right]\right\}\hat{a}(s)U_{\hat{c}\hat{c}}(s)(1+\pi(\mathcal{E}))^{-1}}{\beta\mathbb{E}_z[\exp\left\{-\nu\left[\bar{\Theta}+\mathcal{E}_\Theta\right]\right\}U_{\hat{c}}(1+\pi)^{-1}]}\left(\hat{\mu}'(s)-\hat{\mu}(z)\right) \\
&- \exp\left\{(1-\nu)\left[\bar{\Theta}+\mathcal{E}_\Theta\right]\right\}\hat{\mu}'(s)\left(U_{\hat{c}\hat{c}}(s)\left[\hat{c}(s)-\hat{\mathrm{T}}(\mathcal{E})\right]+U_{\hat{c}}(s)\right) \\
&- \rho(s)\exp\left\{-\nu\left[\bar{\Theta}+\mathcal{E}_{\Theta,t}\right]\right\}U_{\hat{c}}(s)(1+\pi(\mathcal{E}))^{-1} \\
&+\mathrm{W}(\mathcal{E})U_{\hat{c}\hat{c}}(s)g(s)\phi(s) - \chi(\mathcal{E}) + e^{(1-\nu)\left[\bar{\Theta}+\mathcal{E}_\Theta\right]}\left(w(z)U_{\hat{c}} - \beta\varrho(s')\aleph(\mathcal{E})\frac{U_{\hat{c}\hat{c}}(s)}{(U_{\hat{c}}(s))^2}\right), && \text{(35e)}\\
0 =& w(z)e^{(1-\nu)\left[\bar{\Theta}+\mathcal{E}_\Theta\right]}\hat{U}_n(s) - \exp\left\{(1-\nu)\left[\bar{\Theta}+\mathcal{E}_\Theta\right]\right\}\hat{\mu}'(s)\left(U_{nn}(s)n(s)+U_n(s)\right) \\
&+ \phi(s)U_{nn}(s) - g(s)\xi(\mathcal{E}), && \text{(35f)}\\
0 =& -\hat{\mu}(z) + \frac{\mathbb{E}_z\left[\exp\left\{-\nu\left[\bar{\Theta}+\mathcal{E}_{\Theta,t}\right]\right\}U_{\hat{c}}(1+\pi)^{-1}\hat{\mu}'\right]}{\mathbb{E}_z[\exp\left\{-\nu\left[\bar{\Theta}+\mathcal{E}_{\Theta,t}\right]\right\}U_{\hat{c}}(1+\pi)^{-1}]}, && \text{(35g)}\\
0 =& -\varrho(s) - \rho(s)\mathbb{E}_z[\exp\left\{-\nu\left[\bar{\Theta}+\mathcal{E}_\Theta\right]\right\}U_{\hat{c}}(1+\pi)^{-1}], && \text{(35h)}\\
0 =& -e'(s) + \rho_e e(z) + f\left(z, Z, \mathcal{E}\right)\mathcal{E}_\Theta + \eta(\varepsilon), && \text{(35i)}\\
0 =& -\omega'(s) + \omega(z). && \text{(35j)}
\end{aligned}$$

To impose a measurability restriction that $\hat{a}$ and ρ are choice variables that do not depend on shocks $\varepsilon, \mathcal{E}$ we require

$$\hat{a}(s) = \mathbb{E}_z\hat{a}(s), \quad \rho(s) = \mathbb{E}_z\rho(s). \tag{35k}$$

and include them in F. Next, the list of equations in function R are

$$0 = -\hat{C}(\mathcal{E}) + \int \hat{c}(s) d\phi dZ, \tag{36a}$$

$$0 = -\hat{Y}(\mathcal{E}) + \int n(s) g(s) d\phi dZ, \tag{36b}$$

$$0 = -\hat{Y}(\mathcal{E}) + \frac{\psi}{2}\pi(\mathcal{E})^2 + \hat{C}(\mathcal{E}) + \hat{G}(\mathcal{E}), \tag{36c}$$

$$0 = \chi(\mathcal{E}) - \lambda(\mathcal{E}), \tag{36d}$$

$$0 = \xi(\mathcal{E}) + \lambda(\mathcal{E}), \tag{36e}$$

$$0 = \int \phi(s) U_{\hat{c}}(s) g(s) d\phi dZ, \tag{36f}$$

$$0 = \int \hat{\mu}'(s) U_{\hat{c}}(s) d\phi dZ, \tag{36g}$$

$$0 = \int \rho(s) d\phi dZ \tag{36h}$$

$$\begin{aligned} 0 = &- \int \frac{\exp\left\{-\nu\left[\bar{\Theta} + \mathcal{E}_{\Theta,t}\right]\right\} \hat{a}(s) U_{\hat{c}}(s)(1+\pi(\mathcal{E}))^{-2}}{\beta \mathbb{E}_z[\exp\left\{-\nu\left[\bar{\Theta} + \mathcal{E}_{\Theta,t}\right]\right\} U_{\hat{c}}(1+\pi)^{-1}]} \left(\hat{\mu}'(s) - \hat{\mu}(z)\right) d\phi dZ, \\ &+ \int \rho(z) m(z) \exp\left\{-\nu\left[\bar{\Theta} + \mathcal{E}_{\Theta,t}\right]\right\} U_{\hat{c}}(s)(1+\pi(\mathcal{E}))^{-2} d\phi dZ - \psi\pi(\mathcal{E})\lambda(\mathcal{E}). \end{aligned} \tag{36i}$$

C Taylor Expansion

C.1 Proof of Proposition 1

Let $\tilde{\mu}'(z, Z, \sigma\varepsilon, \sigma\mathcal{E}; \sigma)$ and $\tilde{m}'(z, Z, \sigma\varepsilon, \sigma\mathcal{E}; \sigma)$ be optimal policies for the state variable $\hat{\mu}'$ and $\hat{m}'$. From equation (35g) when when $\sigma = 0$

$$\hat{\mu}(z) = \frac{\tilde{\mu}'(z, Z, 0, 0; 0)\tilde{U}_c(z, Z, 0, 0; 0)(1 + \tilde{\pi}(Z, 0; 0))^{-1}}{\tilde{U}_c(z, Z, 0, 0; 0)(1 + \tilde{\pi}(Z, 0; 0))^{-1}} = \tilde{\mu}'(z, Z, 0, 0; 0).$$

From (35b) we have, when $\sigma = 0$,

$$\tilde{\alpha}(Z, 0; 0) = \hat{m}(z)\tilde{U}_{\hat{c}}(z, Z, 0, 0; 0)(1 + \tilde{\pi}(Z, 0; 0))^{-1}.$$

and (35d) implies

$$\tilde{m}'(z, Z, 0, 0; 0)\tilde{U}_{\hat{c}}(z, Z, 0, 0; 0) = \tilde{\aleph}(Z, 0; 0), \qquad \int \tilde{m}'(z, Z, 0, 0; 0) dZ = 1.$$

Thus

$$\tilde{m}'(z, Z, 0, 0; 0) = \frac{\tilde{\aleph}(Z, 0; 0)(1 + \tilde{\pi}(Z, 0; 0))^{-1}}{\tilde{\alpha}(Z, 0; 0)} m(z)$$

and integrating both sides with respect to dZ we find $\frac{\tilde{\aleph}(Z,0;0)(1+\tilde{\pi}(Z,0;0))^{-1}}{\tilde{\alpha}(Z,0;0)} = 1$ and hence

$$\tilde{m}'(z, Z, 0, 0; 0) = m(z).$$

Finally, the stochastic process (35i) for shocks implies

$$e'(z, Z, 0, 0; 0) = e(z).$$

We conclude that

$$\tilde{z}(z, Z, 0, 0; 0) = z$$

and therefore the $\sigma = 0$ allocation is stationary.

C.2 First Order Terms

We first prove Lemma 1:

Proof. Proposition 1 implies that $\frac{\partial}{\partial z}\tilde{z}(z, Z, 0, 0; 0) = 1$ for all (z, Z) and therefore $z_0^l = 1$ for all l. Using this fact, differentiate (19) with respect to z, evaluated at z^l, and re-arrange to get

$$x_0^l = -\left[F_{x-}^l + F_x^l + F_{x+}^l\right]^{-1} F_0^l \text{ for all } l.$$

Differentiation of (19) and (20) with respect to the k^{th} argument of Z gives

$$F_{x-}^l x_k^l + F_x^l x_k^l + F_{x+}^l x_k^l + F_X X_k = 0 \text{ for all } l, k, \tag{37}$$

$$\left(R_0^k + R_x^k x_0^k\right) + \sum_{l=1}^{K}\left(R_x^l x_k^l + R_X^l X_k\right) = 0 \text{ for all } k. \tag{38}$$

Then solving (37) for x_k^l we get

$$x_k^l = -\left[F_{x-}^l + F_x^l + F_{x+}^l\right]^{-1} F_X^l X_k \equiv x_X^l X_k. \tag{39}$$

Substituting x_k^l using (39) allows us to solve for Y_k as

$$X_k = \left(\sum_l \left(R_x^l x_X^l + R_X^l\right)\right)^{-1}\left[R_0^k + R_x^k x_0^k\right]. \tag{40}$$

After we have found X_k, we compute x_k^l from (39). □

We now are ready to prove Proposition 2

Proof. The total derivative of (19) and (20) with respect to σ along with $\mathbb{E}_z[\varepsilon] = \mathbb{E}_z[\mathcal{E}] = 0$ yields[9]

$$\begin{aligned} 0 =& F_x^l(x_\varepsilon^l \varepsilon + x_{\mathcal{E}}^l \mathcal{E} + x_\sigma^l) + F_X^l(X_{\mathcal{E}}\mathcal{E} + X_\sigma) + F_{\mathcal{E}}^l \mathcal{E} + F_\varepsilon^l \varepsilon \\ &+ F_{x+}^l \left[x_0^l Q(x_\varepsilon^l \varepsilon + x_{\mathcal{E}}^l \mathcal{E} + x_\sigma^l) + \sum_k x_k^l Q(x_{\mathcal{E}}^k + x_\sigma^k) + x_\sigma^l \right] \end{aligned} \tag{41}$$

and

$$\sum_k \left(R_x^k(x_{\mathcal{E}}^k \mathcal{E} + x_\sigma^k) + R_{\mathcal{E}}^k \mathcal{E} + R_X^k \left(X_{\mathcal{E}} \mathcal{E} + X_\sigma \right) \right) = 0. \tag{42}$$

As equations (41) and (42) must hold for all $\mathcal{E}$ and ε, combining the terms loading on ε yields the first result of Proposition 2

$$x_\varepsilon^l = \left[F_x^l + F_{x+}^l x_0^l Q \right]^{-1} F_\varepsilon^l.$$

The terms multiplying $\mathcal{E}$ produce the equations

$$F_x^l x_{\mathcal{E}}^l + F_X^l X_{\mathcal{E}} + F_{\mathcal{E}}^l + F_{x+}^l \left[x_0^l Q x_{\mathcal{E}}^l + \sum_k x_k^l Q x_{\mathcal{E}}^k \right] = 0 \tag{43}$$

and

$$\sum_k \left(R_x^k x_{\mathcal{E}}^k + R_{\mathcal{E}}^k + R_X^k X_{\mathcal{E}} \right) = 0. \tag{44}$$

After substituting $x_k^l = x_X^l X_k$ and solving equation (43) for $x_{\mathcal{E}}^l$

$$\begin{aligned} x_{\mathcal{E}}^l &= - \left[F_x^l + F_{x+}^l x_0^l Q \right]^{-1} \left(F_{\mathcal{E}}^l + F_X^l X_{\mathcal{E}} + F_{x+}^l x_X^l \sum_k X_k Q x_{\mathcal{E}}^k \right) \\ &\equiv x_{\mathcal{E},1}^l + x_{\mathcal{E},2}^l X_{\mathcal{E}} + x_{\mathcal{E},3}^l \sum_k X_k Q x_{\mathcal{E}}^k. \end{aligned}$$

Applying $\sum_l X_l Q$ to both sides of this equation gives

$$\sum_k X_k Q x_{\mathcal{E}}^k = \left(I - \sum_l X_l Q x_{\mathcal{E},3}^l \right)^{-1} \left(\sum_l X_l Q x_{\mathcal{E},1}^l + \sum_l X_l Q x_{\mathcal{E},2}^l X_{\mathcal{E}} \right),$$

[9] As ε is mean 0, to first order it will not affect the distribution Z, thus there is no x_ε^k in the laster term of the derivative of F or R.

which can be substituted back into $x^l_{\mathcal{E}}$ to yield the next result of the proposition

$$x^l_{\mathcal{E}} = \left(x^l_{\mathcal{E},1} + x^l_{\mathcal{E},3} \left(I - \sum_k X_k Q x^k_{\mathcal{E},3} \right)^{-1} \left(\sum_k X_k Q x^k_{\mathcal{E},1} \right) \right)$$
$$+ \left(x^l_{\mathcal{E},2} + x^l_{\mathcal{E},3} \left(I - \sum_k X_k Q x^k_{\mathcal{E},3} \right)^{-1} \left(\sum_k X_k Q x^k_{\mathcal{E},2} \right) \right) X_{\mathcal{E}}.$$

The response of aggregate variables can be found by substituting for $x^l_{\mathcal{E}}$ in (44) and then solving for $X_{\mathcal{E}}$

$$X_{\mathcal{E}} = -\left(\sum_k \left[R^k_X + R^k_x x^k_{\mathcal{E},2} + R^k_x x^k_{\mathcal{E},3} \left(I - \sum_l X_l Q x^l_{\mathcal{E},3} \right)^{-1} \left(\sum_l X_l Q x^l_{\mathcal{E},2} \right) \right] \right)^{-1}$$
$$\times \left(\sum_k \left[R^k_{\mathcal{E}} + R^k_x x^k_{\mathcal{E},1} + R^k_x x^k_{\mathcal{E},3} \left(I - \sum_l X_l Q x^l_{\mathcal{E},3} \right)^{-1} \left(\sum_l X_l Q x^l_{\mathcal{E},1} \right) \right] \right).$$

The remaining terms of (41) and (42) give

$$F_\sigma + F^l_{x-} x^l_\sigma + F^l_x x^l_\sigma + F^l_X X_\sigma + F^l_{x+} \left[x^l_0 Q x^l_\sigma + \sum_k x^l_k Q x^k_\sigma + x^l_\sigma \right] = 0 \tag{45}$$

and

$$\sum_k \left(R^k_x x^k_\sigma + R^k_X X_\sigma \right) = 0. \tag{46}$$

Solving for (45) for x^l_σ yields

$$x^l_\sigma = -\left(F^l_{x-} + F^l_x + F^l_{x+} + F^l_{x+} x^l_0 Q \right)^{-1} \left(F_\sigma + F^l_X X_\sigma + F^l_{x+} x^l_X \sum_k X_k Q x^k_\sigma \right)$$
$$\equiv x^l_{\sigma,1} + x^l_{\sigma,2} X_\sigma + x^l_{\sigma,3} \left(\sum_k X_k Q x^k_\sigma \right).$$

Applying $\sum_l X_l Q$ to both sides of this equation gives

$$\sum_k X_k Q x^k_\sigma = \left(I - \sum_l X_l Q x^l_{\sigma,3} \right)^{-1} \left(\sum_l X_l Q x^l_{\sigma,1} + \sum_l X_l Q x^l_{\sigma,2} X_\sigma \right),$$

which can be substituted back into $x^l_{\mathcal{E}}$ to yield the next result of the proposition

$$x^l_\sigma = \left(x^l_{\sigma,1} + x^l_{\sigma,3} \left(I - \sum_k X_k Q x^k_{\sigma,3} \right)^{-1} \left(\sum_k X_k Q x^k_{\sigma,1} \right) \right) + \left(x^l_{\sigma,2} + x^l_{\sigma,3} \left(I - \sum_k X_k Q x^k_{\sigma,3} \right)^{-1} \left(\sum_k X_k Q x^k_{\sigma,2} \right) \right) X_\sigma.$$

The response of aggregate variables can be found by substituting for x^l_σ in (44) and then solving for X_σ

$$X_\sigma = - \left(\sum_k \left[R^k_X + R^k_x x^k_{\sigma,2} + R^k_x x^k_{\sigma,3} \left(I - \sum_l X_l Q x^l_{\sigma,3} \right)^{-1} \left(\sum_l X_l Q x^l_{\sigma,2} \right) \right] \right)^{-1} \times \left(\sum_k \left[R^k_x x^k_{\sigma,1} + R^k_x x^k_{\sigma,3} \left(I - \sum_l X_l Q x^l_{\sigma,3} \right)^{-1} \left(\sum_l X_l Q x^l_{\sigma,1} \right) \right] \right).$$

□

C.3 Proof of Proposition 3

In this section we document the properties of the second order approximation. To express the terms compactly we will use tensor notation. In particular, suppose that A is a $n_1 \times n_2 \times n_3$ dimensional tensor, H is a $n_2 \times n_4$ dimensional tensor and L is a $n_3 \times n_5$ dimensional tensor. Let $A_{\mathtt{ijk}}$ a particular element of the tensor A and $H_{\mathtt{jl}}$ be an element of the tensor H and $L_{\mathtt{km}}$ be an element of tensor L. Define $\langle A, H, L \rangle$ as the $n_1 \times n_4 \times n_5$ tensor given by

$$\langle A, H, L \rangle_{\mathtt{ilm}} \equiv \sum_{j,k} A_{\mathtt{ijk}} H_{\mathtt{jl}} L_{\mathtt{km}},$$

and similarly $\langle A, \cdot, L \rangle$, $\langle A, H, \cdot \rangle$ as

$$\langle A, \cdot, L \rangle_{\mathtt{ijl}} \equiv \sum_l A_{\mathtt{ijk}} L_{\mathtt{kl}} \quad \langle A, H, \cdot \rangle_{\mathtt{ikl}} \equiv \sum_j A_{\mathtt{ijk}} H_{\mathtt{jl}}$$

For convenience, we also define z^l_0 to be the identity matrix, $(x-)^l_0 = x^l_0$, $(x-)^l_X = x^l_X$, $(x+)^l_0 = x^l_0$, $(x+)^l_X = x^l_X$, and X^l_X to be the identity matrix. We begin with a lemma that is the counterpart of Lemma 1 for higher order derivatives.

Lemma 3. $\{x^l_{jk}, X_{jk}\}_{k,l,j}$ *satisfy*

$$x^l_{00} = -\left(F^l_{x-} + F^l_x + F^l_{x+}\right)^{-1} \left(\sum_{\alpha\in\{z,x-,x,x+\}} \sum_{\beta\in\{z,x-,x,x+\}} \langle F^l_{\alpha\beta}, \alpha^l_0, \beta^l_0\rangle \right), \tag{47a}$$

$x^l_{0k} = \langle x^l_{0X}, \cdot, X_k\rangle$ *for* $k \geq 1$ *(symmetrically for* x^l_{j0}*) with*

$$x^l_{0X} = -\left(F^l_{x-} + F^l_x + F^l_{x+}\right)^{-1} \left(\sum_{\alpha\in\{z,x-,x,x+\}} \sum_{\beta\in\{x-,x,x+,X\}} \langle F^l_{\alpha\beta}, \alpha^l_0, \beta^l_X\rangle \right), \tag{47b}$$

and $x^l_{jk} = \langle x^l_{XX}, X_j, X_k\rangle + x^l_X X_{jk}$ *with*

$$x^l_{XX} = -\left(F^l_{x-} + F^l_x + F^l_{x+}\right)^{-1} \left(\sum_{\alpha\in\{x-,x,x+,X\}} \sum_{\beta\in\{x-,x,x+,X\}} \langle F^l_{\alpha\beta}, \alpha^l_X, \beta^l_X\rangle \right). \tag{47c}$$

Finally

$$X_{jk} = 1_{jk} X^j_{00} + \langle X_{jX}, \cdot, X_k\rangle + \langle X_{Xk} X_j, \cdot\rangle + \langle X_{XX}, X_j, X_k\rangle \tag{48a}$$

where 1_{jk} *is* 1 *if* $j = k$ *and* 0 *otherwise*

$$X^j_{00} = -\left[\sum_l \left(R^l_x x^l_X + R^i_X\right)\right]^{-1} \left(R^j_x x^j_{00} + \sum_{\alpha\in\{z,x\}} \sum_{\beta\in\{z,x\}} \langle R^j_{\alpha\beta}, \alpha^j_0, \beta^j_0\rangle \right) \tag{48b}$$

$$X_{jX} = -\left[\sum_l \left(R^l_x x^l_X + R^i_X\right)\right]^{-1} \left(R^j_x x^j_{0X} + \sum_{\alpha\in\{z,x\}} \sum_{\beta\in\{x,X\}} \langle R^j_{\alpha\beta}, \alpha^j_0, \beta^j_X\rangle \right) \tag{48c}$$

$$X_{XX} = -\left[\sum_l \left(R^l_x x^l_X + R^i_X\right)\right]^{-1} \left(\sum_l \left[R^l_x x^l_{XX} + \sum_{\alpha\in\{x,X\}} \sum_{\beta\in\{x,X\}} \langle R^l_{\alpha\beta}, \alpha^l_X, \beta^l_X\rangle \right] \right) \tag{48d}$$

and symmetrically for X_{Xk}.

Proof. From Proposition 1, we know several key features of the policy functions when $\sigma = 0$. For the first order terms $z_0 = I$, $z_k = 0$ for all $k \geq 1$, and $Z^l_k = I$ if $l = k$ and 0 otherwise. For the second order terms, $z_{jk} = 0$ and $Z^l_{jk} = 0$ for all j, k, l. With this, total differentiation of (19), evaluated at z^l, twice with respect to z yields

$$F^l_{x-} x^l_{00} + F^l_x x^l_{00} + F^l_{x+} x^l_{00} + \sum_{\alpha\in\{z,x-,x,x+\}} \sum_{\beta\in\{z,x-,x,x+\}} \langle F^l_{\alpha\beta}, \alpha^l_0, \beta^l_0\rangle = 0,$$

where latter sum captures the contribution of all the first order terms. Solving for x^l_{00} produces equation (47a). Following similar procedures we can produce the other terms (47b) and (47c). To find X_{jk}, total differentiate equation (20) twice with respect to the jth and kth arguments of Z to obtain

$$\begin{aligned}
0 = \sum_l \Bigg(& R^l_x x^l_{jk} + R^l_X X_{jk} + \sum_{\alpha\in\{x,X\}} \sum_{\beta\in\{x,X\}} \langle R^l_{\alpha\beta}, \alpha^l_X X_j, \beta^l_X X_k\rangle \Bigg) \\
& + R^j_x \langle x^j_{0X}, \cdot, X_k\rangle + \sum_{\alpha\in\{z,x\}} \sum_{\beta\in\{x,X\}} \langle R^j_{\alpha\beta}, \alpha^j_0, \beta^j_X X_k\rangle \\
& + R^k_x \langle x^k_{X0}, X_j, \cdot\rangle + \sum_{\alpha\in\{x,X\}} \sum_{\beta\in\{z,x\}} \langle R^k_{\alpha\beta}, \alpha^k_X X_j, \beta^k_0\rangle \\
& + 1_{jk} R^j_x x^j_{00} + 1_{jk} \sum_{\alpha\in\{z,x\}} \sum_{\beta\in\{z,x\}} \langle R^j_{\alpha\beta}, \alpha^j_0, \beta^j_0\rangle.
\end{aligned}$$

Substituting for $x^l_{jk} = \langle x^l_{XX}, X_j, X_k\rangle + x^l_X X_{jk}$ and then solving for X_{jk} gives the expressions in equations (48). □

In addition to Lemma 3 we require expressions governing the interactions of the individual and aggregate states (z, Z) with the perturbation parameter σ. For convenience, we also define $\sigma^l_\sigma = 1, (x-)^l_\sigma = x_\sigma, X^l_\sigma = X_\sigma$ and $(x+)^l_\sigma = x^l_\sigma + x^l_0 Q x^l_\sigma + \sum_k x^l_k Q x^k_\sigma$.

Lemma 4. $\{x^l_{k\sigma}, X_{k\sigma}\}_{k,l}$ *satisfy*

$$x^l_{0\sigma} = -\left(F^l_{x-} + F^l_x + F^l_{x+} + F^l_{x+} x^l_0 Q\right)^{-1} \left(\sum_{\alpha\in\{z,x-,x,x+\}} \sum_{\beta\in\{\sigma,x-,x,X,x+\}} \langle F^l_{\alpha\beta}, \alpha^l_0, \beta^l_\sigma\rangle \right)$$

and

$$\begin{aligned}
x^l_{k\sigma} = & \left(x^l_{Z\sigma,1} + x^l_{\sigma,3} \left(I - \sum_m X_m Q x^m_{\sigma,3}\right)^{-1} \left(\sum_m X_m Q x^m_{Z\sigma,1}\right)\right) X_k \\
& + \left(x^l_{Z\sigma,2} + x^l_{\sigma,3} \left(I - \sum_m X_m Q x^m_{\sigma,3}\right)^{-1} \left(\sum_m X_m Q x^m_{Z\sigma,2}\right)\right) \left(\langle X^k_{00}, \cdot, Q x^k_\sigma\rangle + \langle X_{kX}, \cdot, X'_\sigma\rangle\right) \\
& + \left(x^l_{\sigma,2} + x^l_{\sigma,3} \left(I - \sum_m X_m Q x^m_{\sigma,3}\right)^{-1} \left(\sum_m X_m Q x^m_{\sigma,2}\right)\right) X_{k\sigma}
\end{aligned}$$

where $X'_\sigma = \sum_j X_j Q x^j_\sigma$ *and*

$$z^l_{Z\sigma,1} = -\left(F^l_{x-} + F^l_x + F^l_{x+} + F^l_{x+} x^l_0 Q\right)^{-1} \Bigg(F^l_{x+} \langle x^l_{XX}, \cdot, X'_\sigma \rangle + F^l_{x+} x^l_X \langle X_{XX}, \cdot, X'_\sigma \rangle$$
$$+ F^l_{x+} \langle x^l_{X0}, \cdot, Q x^l_\sigma \rangle + \sum_{\alpha \in \{x-, x, X,, x+\}} \sum_{\beta \in \{\sigma, x-, x, X, x+\}} \langle F^l_{\alpha\beta}, \alpha^l_X, \beta^l_\sigma \rangle \Bigg)$$
$$z^l_{Z\sigma,2} = -\left(F^l_{x-} + F^l_x + F^l_{x+} + F^l_{x+} x^l_0 Q\right) F^l_{x+} x^l_X.$$

Finally

$$X_{k\sigma} = -\left(\sum_l \left[R^l_x x^l_{\sigma,2} + R^l_x x^l_{\sigma,3} \left(I - \sum_m X_m Q x^m_{\sigma,3} \right)^{-1} \left(\sum_m X_m Q x^m_{\sigma,2} \right) + R^l_X \right] \right)^{-1}$$
$$\times \left(\sum_l \left[R^l_x x^l_{Z\sigma,1} + R^l_x x^l_{\sigma,3} \left(I - \sum_m X_m Q x^m_{\sigma,3} \right)^{-1} \left(\sum_m X_m Q x^m_{Z\sigma,1} \right) + \sum_{\alpha,\beta \in \{x,X\}} \langle R^l_{\alpha\beta}, \alpha^l_X, \beta^l_\sigma \rangle \right] X_k \right.$$
$$+ \sum_l \left[x^l_{Z\sigma,2} + x^l_{\sigma,3} \left(I - \sum_m X_m Q x^m_{\sigma,3} \right)^{-1} \left(\sum_m X_m Q x^m_{Z\sigma,2} \right) \right] \left(\langle X^k_{00}, \cdot, Q x^k_\sigma \rangle + \langle X_{kX}, \cdot, X'_\sigma \rangle \right)$$
$$\left. + R^k_x x^k_{0\sigma} + \sum_{\alpha \in \{z,x\}} \sum_{\beta \in \{x,X\}} \langle R^k_{\alpha\beta}, \alpha^k_0, \beta^l_\sigma \rangle \right).$$

and symmetrically for $\{x^l_{\sigma k}, X_{\sigma k}\}_{k,l}$.

Proof. Total differentiation of (19) with respect to z and σ gives

$$F^l_{x-} x^l_{0\sigma} + F^l_x x^l_{0\sigma} + F^l_{x+} x^l_{0\sigma} + F^l_{x+} x^l_0 Q x^l_{0\sigma} + \sum_{\alpha \in \{z, x-, x, x+\}} \sum_{\beta \in \{\sigma, x-, x, x+\}} \langle F^l_{\alpha\beta}, \alpha^l_0, \beta^l_\sigma \rangle = 0.$$

Directly solving for $x^l_{0\sigma}$ yields the expression in the Lemma. For $x^l_{k\sigma}$ and $X^l_{k\sigma}$, total differentiation of (19) and (20) with respect to Z_k and σ obtains

$$0 = F^l_{x-} x^l_{k\sigma} + F^l_x x^l_{k\sigma} + F^l_X X_{k\sigma} + F^l_{x+} x^l_{k\sigma} + F^l_{x+} x^l_0 Q x^l_{k\sigma} + F^l_{x+} \sum_m \langle x^l_{km}, \cdot, Q x^m_\sigma \rangle \tag{49}$$
$$+ F^l_{x+} \langle x^l_{k0}, \cdot, Q x^l_\sigma \rangle + F^l_{x+} \sum_m x^l_m Q x^m_{k\sigma} + \sum_{\alpha \in \{x-, x, X,, x+\}} \sum_{\beta \in \{\sigma, x-, x, X, x+\}} \langle F^l_{\alpha\beta}, \alpha^l_k, \beta^l_\sigma \rangle$$

and

$$0 = \sum_l \left(R_x^l x_{k\sigma}^l + R_X^l X_{k\sigma} + \sum_{\alpha,\beta \in \{x,X\}} \langle R_{\alpha\beta}^l, \alpha_X^l X_k, \beta_\sigma^l \rangle \right) + R_x^k x_{0\sigma}^k + \sum_{\alpha \in \{z,x\}} \sum_{\beta \in \{x,X\}} \langle R_{\alpha\beta}^k, \alpha_0^k, \beta_\sigma^l \rangle. \tag{50}$$

Solving equation (49) for $x_{k\sigma}^l$ gives

$$x_{k\sigma}^l = x_{Z\sigma,1}^l X_k + x_{Z\sigma,2}^l \left(\langle X_{00}^k, \cdot, Q x_\sigma^k \rangle + \langle X_{kX}, \cdot, X_\sigma' \rangle \right) + x_{\sigma,2}^l X_{k\sigma} + x_{\sigma,3}^l \sum_m X_m Q x_{k\sigma}^m$$

where $x_{Z\sigma,1}^l$ and $x_{Z\sigma,2}^l$ are given in the statement of the Lemma and $x_{\sigma,2}^l$ and $x_{\sigma,3}^l$ are in the statement of Proposition 2. Applying $\sum_l X_l Q$ to both sides of this equation and solving for $\sum_l X_l Q x_{k\sigma}^l$ yields the expression for $x_{k\sigma}^l$ found in the Lemma. Substituting for $x_{k\sigma}^l$ in (50) and solving for $X_{k\sigma}$ generates the expression for $X_{k\sigma}$ in the statement of the Lemma. □

Once again we are able to decompose complicated terms such as x_{jk}^l and $x_{k\sigma}^l$ which depend on multiple agents into terms that only depend on a single agent. We exploit this in the following lemma that gives the quadratic terms in the Taylor expansion of $\tilde{x}$ and $\tilde{X}$. For convenience, define $\mathcal{E}_\mathcal{E}^l$ as the identity matrix, $X_\mathcal{E}^l = X_\mathcal{E}$ and $(x+)_\mathcal{E}^l = x_\mathcal{E}^l + x_0^l Q x_\mathcal{E}^l + \sum_k x_k^l Q x_\mathcal{E}^k$.

Lemma 5. *The terms* $\left\{ x_{\varepsilon\varepsilon}^l, x_{\varepsilon\mathcal{E}}^l, x_{\varepsilon\sigma}^l, x_{\mathcal{E}\mathcal{E}}^l, x_{\mathcal{E}\sigma}^l, x_{\sigma\sigma}^l, X_{\mathcal{E}\mathcal{E}}, X_{\mathcal{E}\sigma} X_{\sigma\sigma} \right\}$ *in the second order expan-*

sion for the individual policies $\tilde{x}$ are given by

$$x^l_{\varepsilon\varepsilon} = -\left[F^l_x + F^l_{x+}x^l_0 Q\right]^{-1}\left(\sum_{\alpha,\beta\in\{\varepsilon,x,x+\}}\langle F^l_{\alpha\beta},\alpha^l_\varepsilon,\beta^l_\varepsilon\rangle + F^l_{x+}\langle x^l_{00},Qx^l_\varepsilon,Qx^l_\varepsilon\rangle\right)$$

$$x^l_{\varepsilon\sigma} = -\left[F^l_x + F^l_{x+}x^l_0 Q\right]^{-1}\Bigg(\sum_{\alpha\in\{\varepsilon,x,x+\}}\sum_{\beta\in\{\sigma,x-,x,X,x+\}}\langle F^l_{\alpha\beta},\alpha^l_\varepsilon,\beta^l_\sigma\rangle$$

$$+ F^l_{x+}\langle x^l_{0\sigma},Qx^l_\varepsilon,\cdot\rangle + F^l_{x+}\langle x^l_{00},Qx^l_\varepsilon,Qx^l_\sigma\rangle + F^l_{x+}\sum_k\langle x^l_{0k},Qx^l_\varepsilon,Qx^k_\sigma\rangle\Bigg)$$

$$x^l_{\varepsilon\mathcal{E}} = -\left[F^l_x + F^l_{x+}x^l_0 Q\right]^{-1}\Bigg(\sum_{\alpha\in\{\varepsilon,x,x+\}}\sum_{\beta\in\{\mathcal{E},x,X,x+\}}\langle F^l_{\alpha\beta},\alpha^l_\varepsilon,\beta^l_{\mathcal{E}}\rangle$$

$$+ F^l_{x+}\langle x^l_{00},Qx^l_\varepsilon,Qx^l_{\mathcal{E}}\rangle + F^l_{x+}\sum_k\langle x^l_{0k},Qx^l_\varepsilon,Qx^k_{\mathcal{E}}\rangle$$

$$x^l_{\mathcal{E}\mathcal{E}} = \left(x^l_{\mathcal{E}\mathcal{E},1} + x^l_{\mathcal{E},3}\left(I-\sum_k X_kQx^k_{\mathcal{E},3}\right)^{-1}\left(\sum_k X_kQx^k_{\mathcal{E}\mathcal{E},1}\right)\right)$$

$$+ \left(x^l_{\mathcal{E},2} + x^l_{\mathcal{E},3}\left(I-\sum_k X_kQx^k_{\mathcal{E},3}\right)^{-1}\left(\sum_k X_kQx^k_{\mathcal{E},2}\right)\right)X_{\mathcal{E}\mathcal{E}},$$

$$x^l_{\sigma\mathcal{E}} = \left(x^l_{\sigma\mathcal{E},1} + x^l_{\mathcal{E},3}\left(I-\sum_k X_kQx^k_{\mathcal{E},3}\right)^{-1}\left(\sum_k X_kQx^k_{\sigma\mathcal{E},1}\right)\right)$$

$$+ \left(x^l_{\mathcal{E},2} + x^l_{\mathcal{E},3}\left(I-\sum_k X_kQx^k_{\mathcal{E},3}\right)^{-1}\left(\sum_k X_kQx^k_{\mathcal{E},2}\right)\right)X_{\sigma\mathcal{E}},$$

$$x^l_{\sigma\sigma} = \left(x^l_{\sigma\sigma,1} + x^l_{\sigma,3}\left(I-\sum_k X_kQx^k_{\sigma,3}\right)^{-1}\left(\sum_k X_kQx^k_{\sigma\sigma,1}\right)\right)$$

$$+ \left(x^l_{\sigma,2} + x^l_{\sigma,3}\left(I-\sum_k X_kQx^k_{\sigma\sigma,3}\right)^{-1}\left(\sum_k X_kQx^k_{\sigma,2}\right)\right)X_{\sigma\sigma}$$

and for the aggregate policy function $\tilde{X}$ *are given by*

$$X_{\mathcal{E}\mathcal{E}} = -\left(\sum_k \left[R_X^k + R_x^k x_{\mathcal{E},2}^k + R_x^k x_{\mathcal{E},3}^k \left(I - \sum_l X_l Q x_{\mathcal{E},3}^l\right)^{-1} \left(\sum_l X_l Q x_{\mathcal{E},2}^l\right)\right]\right)^{-1}$$
$$\times \left(\sum_k \left[R_x^k x_{\mathcal{E}\mathcal{E},1}^k + R_x^k x_{\mathcal{E},3}^k \left(I - \sum_l X_l Q x_{\mathcal{E},3}^l\right)^{-1} \left(\sum_l X_l Q x_{\mathcal{E}\mathcal{E},1}^l\right) + \sum_{\alpha,\beta \in \{\mathcal{E},x,X\}} \langle R_{\alpha\beta}^k, \alpha_{\mathcal{E}}^k, \beta_{\mathcal{E}}^k \rangle\right]\right)$$

$$X_{\sigma\mathcal{E}} = -\left(\sum_k \left[R_X^k + R_x^k x_{\mathcal{E},2}^k + R_x^k x_{\mathcal{E},3}^k \left(I - \sum_l X_l Q x_{\mathcal{E},3}^l\right)^{-1} \left(\sum_l X_l Q x_{\mathcal{E},2}^l\right)\right]\right)^{-1}$$
$$\times \left(\sum_k \left[R_x^k x_{\sigma\mathcal{E},1}^k + R_x^k x_{\mathcal{E},3}^k \left(I - \sum_l X_l Q x_{\mathcal{E},3}^l\right)^{-1} \left(\sum_l X_l Q x_{\mathcal{E}\mathcal{E},1}^l\right) + \sum_{\alpha \in \{x,X\}} \sum_{\beta \in \{\mathcal{E},x,X\}} \langle R_{\alpha\beta}^k, \alpha_{\sigma}^k, \beta_{\mathcal{E}}^k \rangle\right]\right)$$

$$X_{\sigma\sigma} = -\left(\sum_k \left[R_X^k + R_x^k x_{\sigma,2}^k + R_x^k x_{\sigma,3}^k \left(I - \sum_l X_l Q x_{\sigma,3}^l\right)^{-1} \left(\sum_l X_l Q x_{\sigma,2}^l\right)\right]\right)^{-1}$$
$$\times \left(\sum_k \left[R_x^k x_{\sigma\sigma,1}^k + R_x^k x_{\sigma,3}^k \left(I - \sum_l X_l Q x_{\sigma,3}^l\right)^{-1} \left(\sum_l X_l Q x_{\sigma\sigma,1}^l\right)\right.\right.$$
$$\left.\left.\sum_{\alpha,\beta \in \{\varepsilon,x\}} \mathbb{E}_z \left[\langle R_{\alpha\beta}^k, \alpha_{\varepsilon}^k \varepsilon, \beta_{\varepsilon}^k \varepsilon \rangle\right] + R_x^k \mathbb{E}_z \left[\langle x_{00}^k, \varepsilon, \varepsilon \rangle\right]\right] + \sum_{\alpha,\beta \in \{x,X\}} \langle R_{\alpha\beta}^k, \alpha_{\sigma}^k, \beta_{\sigma}^k \rangle\right).$$

Where

$$x^l_{\mathcal{E}\mathcal{E},1} = -(F^l_x + F^l_{x+}x^l_0 Q)^{-1}\Bigg(F^l_{x+}\sum_{j,k}\langle x^l_{jk}, Qx^j_{\mathcal{E}}, Qx^k_{\mathcal{E}}\rangle + F^l_{x+}\sum_{j}\langle x^l_{j0}, Qx^j_{\mathcal{E}}, Qx^l_{\mathcal{E}}\rangle$$
$$+ F^l_{x+}\sum_{k}\langle x^l_{0k}, Qx^l_{\mathcal{E}}, Qx^k_{\mathcal{E}}\rangle + F^l_{x+}\langle x^l_{00}, Qx^l_{\mathcal{E}}, Qx^l_{\mathcal{E}}\rangle$$
$$+ \sum_{\alpha,\beta\in\{\mathcal{E},x,X,x+\}}\langle F^l_{\alpha\beta}, \alpha^l_{\mathcal{E}}, \beta^l_{\mathcal{E}}\rangle\Bigg),$$

$$x^l_{\sigma\mathcal{E},1} = -(F^l_x + F^l_{x+}x^l_0 Q)^{-1}\Bigg(F^l_{x+}\sum_{j,k}\langle x^l_{jk}, Qx^j_{\sigma}, Qx^k_{\mathcal{E}}\rangle + F^l_{x+}\sum_{j}\langle x^l_{j0}, Qx^j_{\sigma}, Qx^l_{\mathcal{E}}\rangle$$
$$+ F^l_{x+}\sum_{k}\langle x^l_{0k}, Qx^l_{\sigma}, Qx^k_{\mathcal{E}}\rangle + F^l_{x+}\langle x^l_{00}, Qx^l_{\sigma}, Qx^l_{\mathcal{E}}\rangle$$
$$+ F^l_{x+}\sum_{k} x^l_{\sigma k}Qx^k_{\mathcal{E}} + F_{x+}x^l_{\sigma 0}Qx^l_{\mathcal{E}}$$
$$+ \sum_{\alpha\in\{\sigma,x-,x,X,x+\}}\ \sum_{\beta\in\{\mathcal{E},x,X,x+\}}\langle F^l_{\alpha\beta}, \alpha^l_{\sigma}, \beta^l_{\mathcal{E}}\rangle\Bigg),$$

$$x^l_{\sigma\sigma,1} = -(F^l_{x-} + F^l_x + F^l_{x+} + F^l_{x+}x^l_0 Q)^{-1}\Bigg(F^l_{x+}\sum_{j,k}\langle x^l_{jk}, Qx^j_{\sigma}, Qx^k_{\sigma}\rangle + F^l_{x+}\sum_{j}\langle x^l_{j0}, Qx^j_{\sigma}, Qx^l_{\sigma}\rangle$$
$$+ F^l_{x+}\sum_{k}\langle x^l_{0k}, Qx^l_{\sigma}, Qx^k_{\sigma}\rangle + F^l_{x+}\langle x^l_{00}, Qx^l_{\sigma}, Qx^l_{\sigma}\rangle$$
$$+ F^l_{x+}\sum_{k} x^l_{\sigma k}Qx^k_{\sigma} + F^l_{x+}x^l_{\sigma 0}Qx^l_{\sigma} + F^l_{x+}x^l_{0\sigma}Qx^l_{\sigma}$$
$$+ F^l_{x+}\sum_{k} x^l_{k\sigma}Qx^k_{\sigma} + \left(F^l_{x-} + F^l_{x+}\right)\mathbb{E}_z\left[\langle x^l_{\varepsilon\varepsilon}, \varepsilon, \varepsilon\rangle + \langle x^l_{\mathcal{E}\mathcal{E}}, \mathcal{E}, \mathcal{E}\rangle\right]$$
$$+ F^l_{x+}x^l_X\sum_{k} X_k Q\mathbb{E}_z\left[\langle x^k_{\varepsilon\varepsilon}, \varepsilon, \varepsilon\rangle\right]$$
$$+ F^l_{x+}x^l_X\sum_{k}\mathbb{E}_z\left[\langle X^k_{00}, Qx^k_{\varepsilon}\varepsilon, Qx^k_{\varepsilon}\varepsilon\rangle\right]$$
$$+ \sum_{\alpha,\beta\in\{\sigma,x-,x,X,x+\}}\langle F^l_{\alpha\beta}, \alpha^l_{\sigma}, \beta^l_{\sigma}\rangle\Bigg).$$

Proof. As in Proposition 2, proceed by total differentiating equations (19) and (20) twice with respect to σ. To simplify exposition we will only report the component parts of this derivatives. Combining the terms loading on $\varepsilon\varepsilon$ yields

$$F^l_x x^l_{\varepsilon\varepsilon} + F^l_{x+}x^l_0 Qx^l_{\varepsilon\varepsilon} + \sum_{\alpha,\beta\in\{\varepsilon,x,x+\}}\langle F^l_{\alpha\beta}, \alpha^l_{\varepsilon}, \beta^l_{\varepsilon}\rangle + F^l_{x+}\langle x^l_{00}, Qx^l_{\varepsilon}, Qx^l_{\varepsilon}\rangle = 0.$$

Solving this equation for $x^l_{\varepsilon\varepsilon}$ obtains the expressions in the proposition. A similar procedure with produces the expression for $x^l_{\varepsilon\mathcal{E}}$ and $x^l_{\varepsilon\sigma}$. After combining the terms loading on $\mathcal{E}\mathcal{E}$, from the derivative of (19), we see

$$\begin{aligned}0 =& F^l_x x^l_{\mathcal{E}\mathcal{E}} + F^l_X X_{\mathcal{E}\mathcal{E}} + F^l_{x+} x^l_0 Q x^l_{\mathcal{E}\mathcal{E}} + F^l_{x+} \sum_k x^l_k Q x^k_{\mathcal{E}\mathcal{E}} \\ &+ F^l_{x+} \langle x^l_{00}, Q x^l_{\mathcal{E}}, Q x^l_{\mathcal{E}} \rangle + F^l_{x+} \sum_j \langle x^l_{0k}, Q x^l_{\mathcal{E}}, Q x^k_{\mathcal{E}} \rangle \\ &+ F^l_{x+} \sum_k \langle x^l_{k0}, Q x^k_{\mathcal{E}}, Q x^l_{\mathcal{E}} \rangle + F^l_{x+} \sum_{j,k} \langle x^l_{jk}, Q x^j_{\mathcal{E}}, Q x^k_{\mathcal{E}} \rangle \\ &+ \sum_{\alpha,\beta \in \{\mathcal{E}, x, X, x+\}} \langle F^l_{\alpha\beta}, \alpha^l_{\mathcal{E}}, \beta^l_{\mathcal{E}} \rangle.\end{aligned}$$

After substituting $x^l_k = x^l_X X_k$ and solving for $x^l_{\mathcal{E}\mathcal{E}}$ gives

$$x^l_{\mathcal{E}\mathcal{E}} = x^l_{\mathcal{E}\mathcal{E},1} + x^l_{\mathcal{E},2} X_{\mathcal{E}\mathcal{E}} + x^l_{\mathcal{E},3} \sum_k X_k Q x^k_{\mathcal{E}\mathcal{E}},$$

where $x^l_{\mathcal{E}\mathcal{E},1}$ is the expression from the Lemma and $x^l_{\mathcal{E},2}, x^l_{\mathcal{E},3}$ are the same terms from Proposition 2. Applying $\sum_l X_l Q\cdot$ to both sides and then solving for $\sum_k X_k Q x^k_{\mathcal{E}\mathcal{E}}$ yields the expression for $x^l_{\mathcal{E}\mathcal{E}}$ in the proposition. The terms loading on $\mathcal{E}\mathcal{E}$ in the derivative of (20) imply

$$0 = \sum_k \left(R^k_X X_{\mathcal{E}\mathcal{E}} + R^k_x x^k_{\mathcal{E}\mathcal{E}} + \sum_{\alpha,\beta \in \{\mathcal{E}, x, X\}} \langle R^k_{\alpha\beta}, \alpha^k_{\mathcal{E}}, \beta^k_{\mathcal{E}} \rangle \right).$$

Substituting for $x^l_{\mathcal{E}\mathcal{E}}$ and solving for $X_{\mathcal{E}\mathcal{E}}$, produces the expression for $X_{\mathcal{E}\mathcal{E}}$ in the Lemma.

Combining the loading on $\mathcal{E}\sigma$, from the derivative of (19) gives

$$\begin{aligned}0 =& F^l_x x^l_{\sigma\mathcal{E}} + F^l_X X_{\sigma\mathcal{E}\sigma} + F^l_{x+} x^l_0 Q x^l_{\sigma\mathcal{E}} + F^l_{x+} \sum_k x^l_k Q x^k_{\sigma\mathcal{E}} \\ &+ F^l_{x+} \sum_{j,k} \langle x^l_{jk}, Q x^j_{\sigma}, Q x^k_{\mathcal{E}} \rangle + F^l_{x+} \sum_j \langle x^l_{j0}, Q x^j_{\sigma}, Q x^l_{\mathcal{E}} \rangle \\ &+ F^l_{x+} \sum_k \langle x^l_{0k}, Q x^l_{\sigma}, Q x^k_{\mathcal{E}} \rangle + F^l_{x+} \langle x^l_{00}, Q x^l_{\sigma}, Q x^l_{\mathcal{E}} \rangle \\ &+ F^l_{x+} \sum_k x^l_{\sigma k} Q x^k_{\mathcal{E}} + F_{x+} x^l_{\sigma 0} Q x^l_{\mathcal{E}} \\ &+ \sum_{\alpha \in \{\sigma, x-, x, X, x+\}} \sum_{\beta \in \{\mathcal{E}, x, X, x+\}} \langle F^l_{\alpha\beta}, \alpha^l_{\sigma}, \beta^l_{\mathcal{E}} \rangle.\end{aligned}$$

After substituting $x^l_k = x^l_X X_k$ and solving for $x^l_{\sigma\mathcal{E}}$ gives

$$x^l_{\sigma\mathcal{E}} = x^l_{\sigma\mathcal{E},1} + x^l_{\mathcal{E},2} X_{\mathcal{E}\mathcal{E}} + x^l_{\mathcal{E},3} \sum_k X_k Q x^k_{\mathcal{E}\mathcal{E}},$$

where $x^l_{\mathcal{E}\mathcal{E},1}$ is the expression from the Lemma and $x^l_{\mathcal{E},2}, x^l_{\mathcal{E},3}$ are the same terms from Proposition 2. Applying $\sum_l X_l Q\cdot$ to both sides and then solving for $\sum_k X_k Q x^k_{\sigma\mathcal{E}}$ yields the expression for $x^l_{\sigma\mathcal{E}}$ in the proposition. The terms loading on $\sigma\mathcal{E}$ in the derivative of (20) imply

$$0 = \sum_k \left(R^k_X X_{\sigma\mathcal{E}} + R^k_x x^k_{\sigma\mathcal{E}} + \sum_{\alpha\in\{x,X\}} \sum_{\beta\in\{\mathcal{E},x,X\}} \langle R^k_{\alpha\beta}, \alpha^k_\sigma, \beta^k_{\mathcal{E}} \rangle \right).$$

Substituting for $x^l_{\sigma\mathcal{E}}$ and solving for $X_{\sigma\mathcal{E}}$ yields the expression for $X_{\mathcal{E}\mathcal{E}}$ in the Lemma.

The remaining terms capture the direct dependence on $\sigma\sigma$. From the derivative of (19) :

$$\begin{aligned}
0 =& F^l_x x^l_{\sigma\sigma} + F^l_X X_{\sigma\sigma} + F^l_{x+} x^l_{\sigma\sigma} + F^l_{x+} x^l_0 Q x^l_{\sigma\sigma} + F^l_{x+} \sum_k x^l_k Q x^k_{\sigma\sigma} \\
&+ F_{x-} \left(\mathbb{E}_z \left[\langle x^l_{\varepsilon\varepsilon}, \varepsilon, \varepsilon \rangle \right] + \mathbb{E}_z \left[\langle x^l_{\mathcal{E}\mathcal{E}}, \mathcal{E}, \mathcal{E} \rangle \right] \right) + F_{x+} \left(\mathbb{E}_z \left[\langle x^l_{\varepsilon\varepsilon}, \varepsilon, \varepsilon \rangle \right] + \mathbb{E}_z \left[\langle x^l_{\mathcal{E}\mathcal{E}}, \mathcal{E}, \mathcal{E} \rangle \right] \right) \\
&+ F^l_{x+} x^l_X \sum_k X_k Q \mathbb{E}_z \left[\langle x^k_{\varepsilon\varepsilon}, \varepsilon, \varepsilon \rangle \right] + F^l_{x+} x^l_X \sum_k \mathbb{E}_z \left[\langle X^k_{00}, Q x^k_\varepsilon \varepsilon, Q x^k_\varepsilon \varepsilon \rangle \right] \\
&+ F^l_{x+} \sum_{j,k} \langle x^l_{jk}, Q x^j_\sigma, Q x^k_\sigma \rangle + F^l_{x+} \sum_j \langle x^l_{j0}, Q x^j_\sigma, Q x^l_\sigma \rangle \\
&+ F^l_{x+} \sum_k \langle x^l_{0k}, Q x^l_\sigma, Q x^k_\sigma \rangle + F^l_{x+} \langle x^l_{00}, Q x^l_\sigma, Q x^l_\sigma \rangle \\
&+ F^l_{x+} \sum_k x^l_{\sigma k} Q x^k_\sigma + F^l_{x+} x^l_{\sigma 0} Q x^l_\sigma + F^l_{x+} x^l_{0\sigma} Q x^l_\sigma \\
&+ F^l_{x+} \sum_k x^l_{k\sigma} Q x^k_\sigma + \sum_{\alpha,\beta\in\{\sigma,x-,x,X,x+\}} \langle F^l_{\alpha\beta}, \alpha^l_\sigma, \beta^l_\sigma \rangle.
\end{aligned}$$

The final line line captures the effect of the idiosyncratic shocks on the distribution Z and hence of future policies. Solving this equation for $x^l_{\sigma\sigma}$ gives

$$x^l_{\sigma\sigma} = x^l_{\sigma\sigma,1} + x^l_{\sigma\sigma,2} X_{\sigma\sigma} + x^l_{\sigma\sigma,3} \sum_k X_k Q x^k_{\sigma\sigma},$$

where $x^l_{\sigma\sigma,1}, x^l_{\sigma\sigma,2}$ and $x^l_{\sigma\sigma,3}$ are the terms given in the Lemma. Applying $\sum_l X_l Q\cdot$ to both sides and then solving for $\sum_k X_k Q x^k_{\sigma\sigma}$ yields the expression for $x^l_{\sigma\sigma}$. Finally the remaining terms from the derivative of (20) are

$$0 = \sum_k \left(R^k_x x^k_{\sigma\sigma} + R^k_X X_{\sigma\sigma} + \sum_{\alpha,\beta\in\{\varepsilon,x\}} \mathbb{E}_z \left[\langle R^k_{\alpha\beta}, \alpha^k_\varepsilon \varepsilon, \beta^k_\varepsilon \varepsilon \rangle \right] + R^k_x \mathbb{E}_z \left[\langle x^k_{00}, \varepsilon, \varepsilon \rangle \right] \right).$$

The expression for $X_{\sigma\sigma}$ in the proposition is obtained by substituting for $x^l_{\sigma\sigma}$ and solving for $X_{\sigma\sigma}$. □

Finally we note that all the expressions in Lemma 5 involve expressions that inverses of matricies of order at most $\max\{N_x, N_X\}$. And sum over at most K elements. The latter can be seen as all of the subcomponents of any sum can be decomposed into terms depending on only one group of agents Z_k. For example

$$\begin{aligned}
\sum_{j,k} \langle x^l_{jk}, Qx^j_\sigma, Qx^k_{\mathcal{E}} \rangle =& \sum_{j,k} \left(\langle x^l_{XX}, X_j Q x^j_\sigma, X_k Q x^k_\sigma \rangle + x^l_X \langle X_{jk}, Qx^j_\sigma, Qx^k_{\mathcal{E}} \rangle \right) \\
=& \langle x^l_{XX}, \sum_j X_j Q x^j_\sigma, \sum_k X_k Q x^k_\sigma \rangle + x^l_X \sum_{j,k} 1_{jk} \Bigg(\langle X^j_{00}, Qx^j_\sigma, Qx^k_{\mathcal{E}} \rangle \\
& + \langle X_{Xk}, X_j Q x^j_\sigma, Qx^k_\sigma \rangle + \langle X_{jX}, Qx^j_\sigma, X_k Q x^k_\sigma \rangle + \langle X_{XX}, X_j Q x^j_\sigma, X_k Q x^k_\sigma \rangle \Bigg) \\
=& \langle x^l_{XX}, X'_\sigma, X'_\sigma \rangle + x^l_X \Bigg(\sum_k \langle X^k_{00}, Qx^k_\sigma, Qx^k_{\mathcal{E}} \rangle + \langle X'_{X\sigma}, X'_\sigma, \cdot \rangle \\
& + \langle X'_{\sigma X}, \cdot, X'_\sigma \rangle + \langle X_{XX}, X'_\sigma, X'_\sigma \rangle \Bigg).
\end{aligned}$$

where $X'_\sigma = \sum_k X_k Q x^k_\sigma$, $X'_{X\sigma} = \sum_k \langle X_{Xk}, \cdot, Qx^k_\sigma \rangle$ and symmetrically for $X'_{\sigma X}$.