Optimal Monetary Policy with Redistribution*

Jennifer La'O†

Wendy Morrison‡

July 2024

Abstract

We study optimal monetary policy in a dynamic, general equilibrium economy with heterogeneous agents. All heterogeneity is ex-ante: workers differ in type-specific, state-contingent labor productivity, yet markets are complete. The fiscal authority has access to a uniform, state-contingent lump-sum tax (or transfer), while linear taxes are restricted to be non-state contingent. We derive necessary and sufficient conditions under which implementing flexible-price allocations is optimal. We show that such allocations are not optimal when the relative labor income distribution varies with the business cycle; in such cases, optimal monetary policy implements a state-contingent mark-up that covaries positively with a sufficient statistic for labor income inequality.

Keywords: monetary policy, inequality, redistribution, household heterogeneity, fiscal policy, nominal rigidity, information frictions.

JEL codes: E52, D63, H23

---

*We thank Joao Guerreiro, Victor Rios-Rull, and Elisa Rubbo for their excellent and constructive discussions of our paper. We furthermore thank Fernando Alvarez, Adrien Auclert, Marina Azzimonti, Marco Bassetto, Anmol Bhandari, V.V. Chari, Mike Golosov, Greg Kaplan, Juanpa Nicolini, Andreas Schaab, Balint Szoke, Harald Uhlig, Ivan Werning, Nicolas Werquin, and Mike Woodford for insightful comments, feedback, and suggestions.

†Columbia University, NBER, and CEPR. jenlao@columbia.edu.

‡Columbia University. wam2133@columbia.edu.

1 Introduction

Empirical studies find large, systematic, and forecastable differences in labor earnings profiles across households. One prominent feature of the data is the unequal exposure of household earnings to the business cycle: the labor income of low-income households exhibit a greater covariance with aggregate fluctuations than that of high-income households (Parker and Vissing-Jorgensen, 2009; Guvenen et al., 2017; Alves et al., 2020). Heterogeneity in the covariance of individual earnings with aggregate fluctuations contributes to countercyclical earnings inequality. Notably, these differences are, to a large extent, systematic: "fortunes during recessions are predictable by observable characteristics before the recession," (Guvenen, Ozkan and Song, 2014).

Monetary policy is not a tool best suited for achieving distributional objectives—questions of redistribution typically fall under the purview of the fiscal authority. Yet countercyclical earnings inequality calls this division of labor into question: if fiscal instruments cannot respond to movements at short-run, business cycle frequency, monetary instruments might be the next best alternative. That said, even if monetary policymakers were to respond to short-run fluctuations in the distribution of earnings, it is not obvious from a theoretical perspective whether monetary policy should do so and in what manner.

In this paper we seek to answer this question. We study the optimal conduct of monetary and fiscal policy in a dynamic, general equilibrium model in which households are *ex-ante* heterogeneous and markets are complete. We focus on ex ante heterogeneity, rather than ex post, and therefore on the question of redistribution rather than lack of insurance.[1] In this context, we adopt a utilitarian welfare function and follow the Ramsey approach. Given a restricted set of fiscal instruments, we ask under what conditions should monetary policy play an active role in redistribution? And if such conditions are met, in what manner should monetary policy be conducted in order to maximize welfare?

Our Framework and Methodology. We study a general equilibrium, heterogeneous agent economy with nominal rigidities. We model household heterogeneity following Werning (2007). Households are assigned a "type" at birth and remain that type throughout their lifetime. Type-specific labor productivities are stochastic and contingent on the aggregate state; we allow these contingencies to be fully general and can therefore nest any exogenous labor income process. We assume that markets are complete: in every period, households can trade a complete set of Arrow securities. It follows that there are no missing insurance markets.

[1] By focusing on ex ante heterogeneity and complete markets, our framework stands in contrast to heterogeneous-agent New Keynesian models (HANK), e.g. Kaplan, Moll and Violante (2018); Auclert (2019); Auclert, Rognlie and Straub (2018), that typically feature idiosyncratic labor income risk and incomplete asset markets. We discuss the relationship to the HANK literature below.

A continuum of intermediate-good firms employ workers, produce differentiated goods, and are subject to aggregate productivity shocks. These firms are monopolistically-competitive and set prices subject to nominal rigidities. We model the nominal rigidity as an informational friction as in Woodford (2003a); Mankiw and Reis (2002); Mackowiak and Wiederholt (2009); Angeletos and La'O (2020). For tractability we adopt a particular specification of Correia, Nicolini and Teles (2008): we assume that a fixed fraction of randomly-selected firms set their nominal prices before perfectly observing realized demand. In our benchmark we assume that equity shares of the intermediate-good firms are evenly distributed among all households, but we relax this assumption in our extended model.

The desirability of monetary policy in any context depends on the available set of fiscal instruments. We consider a consolidated government that controls both fiscal and monetary policy. The government raises tax revenue and issues nominal bonds in order to finance redistribution via uniform, lump-sum transfers.

We follow the Ramsey approach and allow for affine taxes on consumption, labor income, firm revenue (sales), and profits. We assume that all tax rates are non-state-contingent, in line with the New Keynesian literature. One can think of this lack of fiscal state-contingency as a political constraint: the fiscal authority cannot change tax rates at business cycle frequency. Furthermore, and in contrast to the typical restriction imposed in the Ramsey literature, we allow for state-contingent, lump-sum taxes or transfers as in Werning (2007). That is, while the fiscal authority cannot change the *slope* of the tax schedule in response to shocks, it can freely adjust the intercept. Crucially, however, we restrict lump-sum transfers (or taxes) to be uniform across household types.[2]

Finally, we adopt a utilitarian welfare function with arbitrary Pareto weights. We solve for optimal fiscal and monetary policy jointly using the primal approach (Lucas and Stokey, 1983; Chari, Christiano and Kehoe, 1991, 1994; Chari and Kehoe, 1999). In particular we adapt the primal approach used in Werning (2007) for a flexible-price economy with heterogeneous agents, and that employed in Correia, Nicolini and Teles (2008) for a representative agent economy with nominal rigidities, to our setting that features both ex ante heterogeneous households and nominal rigidities.

Results. We first derive sufficient conditions under which implementation of flexible-price allocations is optimal. Specifically, we show that when shocks to the labor skill distribution affect all households proportionally—that is, when there are no movements in workers' *relative* skills—the optimal level of redistribution is achieved through the tax system. In this case, non-state-contingent distortionary taxes and uniform lump-sum transfers are sufficient to implement the Ramsey optimum, and monetary policy should play no redistributive role.

[2] One can motivate this restriction with an informational constraint on the government: the fiscal authority cannot tell apart high-type households from low-types.

A distortionary tax on consumption or on labor income implies that high-skilled, wealthy households pay more taxes (in levels) than low-skilled, poor households. In combination with a uniform lump-sum transfer, a higher tax rate lowers wealth inequality (Werning, 2007; Correia, 2010). The planner in our environment optimally trades off the redistributional benefit of distortionary taxation with its efficiency cost. When shocks to the labor skill distribution affect all workers proportionally (and preferences are homothetic), both the marginal benefit and the marginal cost to this tax are invariant to the aggregate state. It follows that the optimal wedge is invariant to the the business cycle and, as a result, the restricted set of fiscal instruments is sufficient to implement the planner's optimum. The best that monetary policy can do, in this case, is to replicate flexible-price allocations; it can do so by targetting price stability.

We show that this is not the case when shocks alter the workers' relative skill distribution. When the labor income of certain households are disproportionally affected by business cycle fluctuations, the available set of fiscal instruments is insufficient to implement the Ramsey optimum on its own. It is then optimal for monetary policy to deviate from implementing flexible-price allocations and play an active role in redistribution. In particular, we find that optimal monetary policy targets a state-contingent markup that co-varies positively with a sufficient statistic for labor income inequality.

To understand this result, consider again the case in which labor skill shocks are proportional. A constant tax rate is sufficient to implement the planner's optimum because both the marginal benefit of distortionary taxation (greater redistribution) and the marginal cost (efficiency) are invariant to the aggregate state. When instead labor skill shocks are disproportional, the marginal redistributional benefit of distortionary taxation increases with labor income inequality, while the marginal cost remains the same; it follows that the optimal tax rate in such states should increase.

However, tax rates are restricted to be non-state-contingent. This restriction on fiscal state-contingency is what opens the door for monetary policy to step in and play a redistributive role. In particular, we show that it is optimal for monetary policy to target a higher aggregate mark-up when labor market inequality is high and, conversely, a lower mark-up when labor market inequality is low. In doing so, monetary policy imperfectly replicates the missing tax instrument with a "monetary tax" in high inequality states and a "monetary subsidy" in low inequality states.

Our results are robust to heterogeneous equity shares. When monetary policy increases the "monetary tax" by targeting a higher mark-up, firm profits increase. Depending on how profit shares co-vary with lifetime income, this can either curb or exacerbate overall income inequality. We show that the presence of heterogeneous equity shares changes both the slope and the intercept of the response of monetary policy to labor income inequality, depending in part on this covariance as well as the degree of firm market power, but it does not alter the general lesson that the optimal markup should covary positively with a sufficient statistic for labor income inequality.

Related literature. The most widely used framework for analyzing monetary policy is the New Keynesian (NK) model (Woodford, 2003b; Galí, 2008) in which a single representative household is typically assumed. However, the more recent Heterogeneous Agent New Keynesian (HANK) literature explicitly incorporates heterogeneity into the NK framework, but does so by introducing uninsurable idiosyncratic income risk (Kaplan, Moll and Violante, 2018; Auclert, Rognlie and Straub, 2018). In these models of the Bewley-Imrohoroglu-Huggett-Aiyagari variety, households use precautionary savings to self-insure against income shocks and smooth their consumption. HANK models can therefore generate an endogenous wealth distribution with heterogeneous marginal propensities to consume, affecting both the amplification and transmission of monetary shocks. Furthermore, monetary policy can play a novel role of providing insurance by transferring resources from savers to borrowers (Acharya, Challe and Dogra, 2020; Dávila and Schaab, 2022; McKay and Wolf, 2022; Bhandari, Evans, Golosov and Sargent, 2021).

In contrast to the canonical HANK model, in our framework markets are complete: households can fully insure themselves against idiosyncratic labor income shocks. We thus focus solely on ex-ante heterogeneity rather than ex-post. In doing so, we abstract entirely from the insurance motive for monetary policy and focus solely on the *redistributive* motive. As noted above, empirical evidence suggests that systematic and forecastable differences in household earnings profiles are quantitatively important. Guvenen and Smith (2014) and Schulhofer-Wohl (2011) furthermore find that households are able to smooth their consumption to a large degree and that systematic differences between households account for a large share of differences in household income growth.

Within this context, we show that monetary policy can exploit the redistributional benefits of an "inflation tax" in a manner that is similar to a distortionary tax rate coupled with a lump-sum transfer. The theoretical insight that a flat tax with a lump-sum transfer can reduce inequality has theoretical underpinnings in the macro-public finance literature. In particular, we build on previous insights found in Werning (2007) and Correia (2010).

Our paper is most closely related to the Ramsey literature on optimal taxation, in particular those that apply the primal approach (Lucas and Stokey, 1983; Chari, Christiano and Kehoe, 1991, 1994; Chari and Kehoe, 1999). A subset uses the primal approach to characterize optimal monetary policy in economies with nominal rigidities (Correia, Nicolini and Teles, 2008; Correia, Farhi, Nicolini and Teles, 2013; Angeletos and La'O, 2020). As a methodological contribution, to the best of our knowledge we are the first to show how the primal approach can be used to characterize optimal monetary policy even when the Ramsey optimum cannot be implemented under flexible prices.

Layout. This paper is organized as follows. In Section 2 we describe the economic environment and in Section 3 we characterize equilibrium allocations. In Sections 4 and 5 we solve two versions of the Ramsey problem: a relaxed problem and the full version, respectively. In Section 6 we discuss implementation. In Section 7 we analyze an extension of the model in which households hold heterogeneous profit shares. Finally, in Section 8 we explore a simple, calibrated version of the baseline model. Section 9 concludes. All proofs, except for those explicitly provided in the text, are found in the appendix.

2 The Model

We study a dynamic, stochastic, general equilibrium economy with heterogeneous agents and a form of nominal rigidity.

2.1 The Environment

Time is discrete, indexed by $t = 0, 1, \dots, \infty$. We denote the aggregate state at time t by $s_t \in S$ where S is a finite set. We let $s^t = \{s_0, ..., s_t\} \in S^t$ denote a history of states up to and including time t. We let $\mu(s^t|s^{t-1})$ denote the probability of history s^t conditional on s^{t-1}. With slight abuse of notation, we denote the unconditional probability of history s^t by $\mu(s^t)$.

Households. There is a measure one continuum of households. Households have identical preferences; in each period, a household receives flow utility $U(c, h)$ from consumption c and work effort h. We assume throughout that preferences are additively-separable and iso-elastic:

$$U(c,h) = \frac{c^{1-\gamma}}{1-\gamma} - \frac{h^{1+\eta}}{1+\eta}, \qquad \text{with} \qquad \gamma, \eta > 0.$$

The parameters γ and η denote the inverse elasticity of intertemporal substitution and the inverse Frisch elasticity of labor supply, respectively.

Households are divided into a finite number of types $i \in I$ of relative size π^i, with $\sum_{i \in I} \pi^i = 1$. Households are born a type and remain that type throughout their (infinite) lifetime. The worker of a type-i household has "skill" level $\theta^i(s_t)$ in time t, state s_t. If the worker puts in $h^i(s^t)$ units of effort, then its labor in efficiency units are given by: $\ell^i(s^t) = \theta^i(s_t)h^i(s^t)$. Thus, the household maximizes lifetime expected utility given by:

$$\sum_t \sum_{s^t} \beta^t \mu(s^t) U(c^i(s^t), \ell^i(s^t)/\theta^i(s_t)). \tag{1}$$

The household's budget constraint at time t, history s^t is written in nominal terms by:

$$\begin{aligned}(1+\tau_c)P(s^t)c^i(s^t) + b^i(s^t) - (1+i(s^{t-1}))b^i(s^{t-1}) + \sum_{s^{t+1}|s^t} Q(s^{t+1}|s^t)z^i(s^{t+1}|s^t) \leq \\ (1-\tau_\ell)W(s^t)\ell^i(s^t) + (1-\tau_\Pi)\Pi(s^t) + P(s^t)T(s^t) + z^i(s^t|s^{t-1})\end{aligned} \tag{2}$$

where $P(s^t)$ is the nominal price of the final good at time t and $W(s^t)$ is the nominal wage per efficiency unit. The household faces constant consumption and labor tax rates, τ_c and τ_ℓ, respectively.

For our baseline analysis we assume that households own equal shares of the intermediate good firms. Equity ownership is a claim to intermediate good firm profits, denoted in nominal terms by $\Pi(s^t)$ and taxed at a constant rate of $\tau_\Pi \in [0, 1]$. We relax this assumption and consider heterogeneous equity shares in Section 7.

The household may choose to borrow or save via two separate instruments. The first is a one-period, non-state-contingent bond, $b^i(s^t)$ which the household may buy or sell at time t, history s^t,and which pay $(1 + i(s^t))b^i(s^t)$ units of money one period later. The second is a complete set of state-contingent Arrow securities, indexed by $s^{t+1} \in S^{t+1}$. We let $Q(s^{t+1}|s^t)$ denote the price at time t, history s^t, of an Arrow security that pays 1 unit of money in period $t + 1$ if state s^{t+1} is realized and 0 otherwise. We denote the corresponding quantities purchased of this Arrow security by $z^i(s^{t+1}|s^t)$. Note that the non-state-contingent bond is a redundant asset but allows us to represent the one-period interest rate, $i(s^t)$.

Finally, $T(s^t)$ is a real, lump-sum transfer and is unrestricted; it can be either positive (a transfer) or negative (a tax) and can depend on the realized history of aggregate states s^t. We assume that initial wealth is zero and we state the household's problem as follows.

Household's Problem. *Given initial bond holdings $b^i(s^0) = 0$ and Arrow securities $z^i(s^0) = 0$, the type-i household chooses a complete contingent plan for consumption, efficiency units of labor, bond holdings, and Arrow security holdings: $\{c^i(s^t), \ell^i(s^t), b^i(s^t), (z^i(s^{t+1}|s^t))_{s^{t+1}}\}_{t\geq 0, s^t \in S^t}$, in order to maximize its lifetime expected utility* (2) *subject to its budget constraint* (2) *and no-Ponzi conditions.*

Intermediate good production. There is a measure one continuum of intermediate-good firms, indexed by $j \in \mathcal{J} \equiv [0, 1]$, with identical technologies. The production function of intermediate-good firm j is given by the constant returns to scale production function

$$y^j(s^t) = A(s_t)n^j(s^t), \tag{3}$$

where $A(s_t)$ is an exogenous, aggregate productivity shock and $n^j(s^t)$ is firm j's input of efficiency units of labor.

Intermediate-good firms are monopolistically-competitive: they produce differentiated goods and set nominal prices. The nominal profits of firm j in history s^t are given by $f^j(s^t) = (1-\tau_r)p^j_t(\cdot)y^j(s^t) - W(s^t)n^j(s^t)$ where τ_r is a constant marginal tax on firm revenue. We postpone for the moment our discussion of the nominal rigidity—that is, how the price $p^j_t(\cdot)$ is set.

Final good production. A representative firm produces the final good with the following constant elasticity of substitution (CES) technology over intermediate-good varieties:

$$Y(s^t) = \left[\int_{j\in\mathcal{J}} y^j(s^t)^{\frac{\rho-1}{\rho}} \mathrm{d}j\right]^{\frac{\rho}{\rho-1}},$$

with constant elasticity of substitution parameter $\rho > 1$. The final good producer is perfectly competitive and takes prices as given. Its nominal profits are given by $P(s^t)Y(s^t) - \int_{j\in\mathcal{J}} p_t^j(\cdot)y^j(s^t)\mathrm{d}j$ where $p_t^j(\cdot)$ is the price of variety j at time t and $P(s^t)$ is the nominal price of the final good.

Given nominal prices, profit maximization of the representative final good producer implies the standard downward-sloping CES demand function for intermediate good j given by:

$$y^j(s^t) = \left(\frac{p_t^j(\cdot)}{P(s^t)}\right)^{-\rho} Y(s^t), \qquad \forall s^t \in S^t. \tag{4}$$

At its optimum, the representative final good producer makes zero profits.

The government. The government consists of a consolidated monetary and fiscal authority with commitment. Let $\mathcal{T}(s^t)$ denote the nominal tax revenue collected at time t, history s^t, given by:

$$\mathcal{T}(s^t) \equiv \tau_c P(s^t)C(s^t) + \tau_\ell W(s^t)L(s^t) + \tau_r P(s^t)Y(s^t) + \tau_\Pi \Pi(s^t),$$

where

$$C(s^t) \equiv \sum_{i\in I} \pi^i c^i(s^t), \qquad L(s^t) \equiv \sum_{i\in I} \pi^i \ell^i(s^t), \qquad \text{and} \qquad \Pi(s^t) \equiv \int_{j\in\mathcal{J}} f^j(s^t)\mathrm{d}j$$

denote aggregate consumption, aggregate labor supply in efficiency units, and aggregate profits of the intermediate-good firms, respectively.

The government's period-t nominal budget constraint is given by:

$$(1+i(s^{t-1}))B(s^{t-1}) + Z(s^t) + P(s^t)T(s^t) \le B(s^t) + \sum_{s^{t+1}|s^t} Q(s^{t+1}|s^t)Z(s^{t+1}) + \mathcal{T}(s^t), \tag{5}$$

where $B(s^t) \equiv \sum_{i\in I} \pi^i b^i(s^t)$ and $Z(s^t) \equiv \sum_{i\in I} \pi^i z^i(s^t|s^{t-1})$ denote aggregate bond holdings and aggregate Arrow security holdings, respectively. Finally, we abstract from the zero lower bound on the nominal interest rate.

Market Clearing. Market clearing in the goods and labor markets imply $C(s^t) = Y(s^t)$ and $L(s^t) = \int_{j\in J} n^j(s^t)\mathrm{d}j$, respectively. That is, aggregate consumption is equated with final good production, and aggregate labor supply (in efficiency units) is equated with labor demand.

2.2 Shocks and the Nominal Rigidity

At each date t, Nature draws the aggregate state $s_t \in S$ according to the probability distribution μ. The aggregate state determines period t total factor productivity and the relative skills for each type $i \in I$. Formally, we define functions $A : S \to \mathbb{R}_+$ and $\theta^i : S \to \mathbb{R}_+$, for all $i \in I$, mapping the state s_t at time t to aggregate productivity and the relative skill distribution.

The nominal rigidity. Intermediate good firms are price-setters. We model the nominal rigidity as an informational friction as in Woodford (2003a), and Mankiw and Reis (2002). For tractability we follow a particular specification assumed in Correia, Nicolini and Teles (2008); that is, we assume that all firms can set their nominal prices in every period, but in each period a fraction of firms are inattentive to the current state.

Formally, we assume that in every period a mass κ of firms are inattentive, or "sticky," and a mass $1 - \kappa$ firms are attentive, or "flexible." We let $\mathcal{J}^s \subset \mathcal{J}$ denote the set of firms that are sticky and $\mathcal{J}^f \subset \mathcal{J}$ denote the set of firms that are flexible, with $\mathcal{J}^f = (\mathcal{J}^s)'$.

Sticky-price firms are inattentive to the current state s_t at time t. They choose their price based only on their knowledge of the history of previous states, s^{t-1}. We denote the price they set by $p_t^s(s^{t-1})$. The subscript t on the price indicates that this is the nominal price set *at time* t by the sticky-price firm, however, the price itself is a function of the history of states only up to time $t-1$. That is, we impose that this price is measurable only up to history s^{t-1}.

The flexible-price firms, on the other hand, are attentive to the current state s_t as well as the entire history of previous states, s^{t-1}. Hence, these firms can set their price as functions of s^t. We denote the price they set by $p_t^f(s^t)$. The subscript t on the price similarly indicates that this is the nominal price set *at time* t by the flexible-price firm. However, unlike the sticky price firms, the flexible-price firms are attentive to the current state s_t, and hence their price is measurable in the current history s^t.

Implicit Timing Assumption. Implicit in the above measurability constraints is the following within-period timing assumption. Nature draws the aggregate state $s_t \in S$ at the beginning of the period and randomly selects which firms are sticky, or "inattentive," $j \in \mathcal{J}^s$, and which firms are flexible, $j \in \mathcal{J}^f$. Intermediate good firms make their nominal pricing decisions given their available information set: s^{t-1}if sticky, s^t if flexible. Once nominal prices are set, the aggregate state becomes common knowledge. Given intermediate good prices, households and final good firms make their respective decisions; specifically, the final good firm purchases inputs and produces, and the households make their consumption, savings, and effort choices. All allocations adjust so that supply equals demand and markets clear.[3]

[3]We make the simplifying assumption that all intermediate-good firms learn the aggregate state at the end of each period. This assumption is compatible with the notion that all firms can observe end-of-

2.3 Remarks on the model

This concludes our description of the model. That said, we have made several modeling choices that depart from the standard New Keynesian model, the typical Ramsey framework, as well as more the recent HANK models. We discuss these choices below.

Heterogeneity with market completeness. Household types remain fixed, however household labor income can vary over time and over states in a general and flexible manner characterized by the arbitrary function $\theta^i : S \to \mathbb{R}_+$. This formulation nests all labor income processes, including those with a high degree of heterogeneity in the covariance of household labor income with aggregate shocks. However, in the proceeding analysis we show that the complete markets assumption implies that households fully insure themselves against idiosyncratic consumption risk: equilibrium household consumption varies only with aggregate consumption. In this sense there are no missing insurance markets; household heterogeneity in consumption is entirely "ex-ante" rather than "ex post."

Lump-sum transfers. In the standard, representative-agent Ramsey framework, lump-sum taxes or transfers—or any combination of taxes that may replicate them—are a priori ruled out. Were it not the case, the first best would be achievable. When instead households are heterogeneous, Werning (2007) shows that one can incorporate a lump-sum tax or transfer into a Ramsey taxation-style model without sacrificing the earlier lessons from the optimal taxation literature. In such a framework, it is the uniformity of the lump-sum transfer *across* types that ensures that the first best is unattainable. We follow Werning (2007) in this vein and assume the existence of a lump-sum transfer that is uniform across household types. One can think of the uniformity restriction as an informational constraint on the government: the fiscal authority cannot distinguish household types.

The lack of fiscal state-contingency. The nature of optimal monetary policy often depends on the set of available fiscal instruments. We assume linear taxation as in Ramsey; accordingly we allow for consumption, labor income, sales, and profit taxes. We therefore do not artificially restrict the *type* of linear taxes in our model.

However, we constrain these tax rates to be fixed, i.e. non-state-contingent. This lack of state-contingency is what opens the door for a potential role for monetary policy. State-contingency of monetary policy but non-state-contingency of taxes is the typical assumption made in New Keynesian models; it is motivated by the idea that the monetary authority is better suited for responding to shocks at business cycle frequency than the fiscal authority.

period equilibrium outcomes and from these endogenous objects infer the realized state at time t.

At the same time we allow the uniform lump-sum transfer to be state-contingent. We find this particular choice of fiscal state-contingency to be reasonable: while legislation of tax rates is often a prolonged and difficult political process, the same is not necessarily true for fiscal transfers. However, we will show that the state-contingency of the lump-sum transfer is in fact without loss of generality–due to the complete markets assumption, all agents in our model are Ricardian.

Nominal Rigidity. Finally, we equate the nominal rigidity in our model with an informational friction. We do so for tractability. By assuming only a measurability constraint on firm pricing, the firm's problem becomes static. Every firm is free to adjust its price in every period; it follows that no firm needs to take into account future periods when setting its current period price.[4]

3 Equilibrium Definition and Characterization

In this section we define a competitive equilibrium in our economy and characterize the set of equilibrium allocations. We close this section with a definition of the Ramsey problem.

3.1 Equilibrium Definition

We denote an allocation in this economy by:

$$x \equiv \{(c^i(s^t), \ell^i(s^t))_{i\in I}, (y^j(s^t), n^j(s^t))_{j\in\mathcal{J}}, C(s^t), Y(s^t), L(s^t)\}_{t\geq 0, s^t\in S^t}$$

Formally, we say that an allocation x is feasible if it satisfies the economy's technology and resource constraints.

Definition 1. *An allocation x is* feasible *if, for all $s^t \in S^t$:*

$$y^j(s^t) = A(s_t)n^j(s^t), \qquad \forall j \in \mathcal{J}; \tag{6}$$

$$C(s^t) = Y(s^t) = \left[\int_{j\in\mathcal{J}} y^j(s^t)^{\frac{\rho-1}{\rho}}\mathrm{d}j\right]^{\frac{\rho}{\rho-1}}; \qquad L(s^t) = \int_{j\in\mathcal{J}} n^j(s^t)\mathrm{d}j; \tag{7}$$

$$C(s^t) = \sum_{i\in I}\pi^i c^i(s^t); \quad \textit{and} \qquad L(s^t) = \sum_{i\in I}\pi^i \ell^i(s^t). \tag{8}$$

Let $\mathcal{X}$ denote the set of all feasible allocations. We are interested in the set of feasible allocations $x \in \mathcal{X}$ that can be supported as part of a competitive equilibrium in our economy. Prior

[4]Furthermore, by assuming the exact same nominal rigidity present in Correia, Nicolini and Teles (2008), we can directly tie our results to the relevant literature.

to defining our equilibrium concept(s), we introduce some simplifying notation. We denote a policy by:

$$\Omega \equiv \{\tau_c, \tau_\ell, \tau_r, \tau_\Pi, T(s^t), i(s^t)\}_{t\geq 0, s^t\in S^t},$$

a price system by:

$$\varrho \equiv \{p_t^f(s^t), p_t^s(s^{t-1}), P(s^t), W(s^t), (Q(s^{t+1}|s^t))_{s^{t+1}\in S^{t+1}}\}_{t\geq 0, s^t\in S^t},$$

and a set of financial asset positions by:

$$\zeta \equiv \{(b^i(s^t))_{i\in I}, B(s^t), (z^i(s^{t+1}|s^t), Z(s^{t+1}))_{s^{t+1}\in S^{t+1}}\}_{t\geq 0, s^t\in S^t}.$$

We define an equilibrium in this economy as follows.

Definition 2. *A sticky-price equilibrium is an allocation x, a price system ϱ, a policy Ω, and asset holdings ζ such that: (i) at time t, history s^t, the price $p_t^s(s^{t-1})$ is optimal for all sticky-price firms $j \in \mathcal{J}^s$, the price $p_t^f(s^t)$ is optimal for all flexible-price firms $j \in \mathcal{J}^f$, and the aggregate price level given by:*

$$P(s^t) = \left[\kappa p_t^s(s^{t-1})^{1-\rho} + (1-\kappa)p_t^f(s^t)^{1-\rho}\right]^{\frac{1}{1-\rho}}; \tag{9}$$

(ii) prices and allocations satisfy the CES demand function (4) for all $j \in \mathcal{J}$ at time t; (iii) given the price system and the policy, the allocation and financial asset holdings of type i solve the household problem of type i, for every $i \in I$; (iv) the government budget constraint is satisfied; and (v) markets clear.

In addition to sticky-price equilibria, we will also consider a hypothetical benchmark economy in which we abstract from nominal rigidities. To construct this benchmark we relax the measurability constraints on firms so that all firms have complete information about current fundamentals s_t when making their respective decisions. Formally we call this the "flexible-price" environment and define a competitive equilibrium in this environment accordingly.

Definition 3. *A flexible-price equilibrium is an allocation x, a price system ϱ, a policy Ω, and asset holdings ζ such that: (i) at time t, history s^t, the price $p_t^f(s^t)$ is optimal for all firms $j \in \mathcal{J}^f = \mathcal{J}$, and the aggregate price level given by:*

$$P(s^t) = p_t^f(s^t), \qquad \forall s^t \in S^t; \tag{10}$$

and parts (ii)-(v) of Definition 2 hold.

The flexible-price environment will serve as a natural benchmark for separating the roles of fiscal and monetary policy in our model.

3.2 Household and Firm optimality

Households. Consider the individual household's problem.[5] Markets are complete and taxes are linear; this implies that all households face the same after-tax prices. As a result, marginal rates of substitution across all goods and states are equated across households. The Negishi (1960) characterization of competitive equilibria then follows.

Lemma 1. *(Negishi, 1960; Werning, 2007). For any equilibrium there exist "market" or "Negishi" weights $\varphi \equiv (\varphi^i)_{i\in I}$ with $\varphi^i \geq 0$ so that the individual assignments of consumption and labor in each history s^t solve the following static sub-problem*

$$U^m(C(s^t), L(s^t); \varphi) \equiv \max_{(c^i(s^t), \ell^i(s^t))_{i\in I}} \sum_{i\in I} \varphi^i \pi^i U(c^i(s^t), \ell^i(s^t)/\theta^i(s_t)) \tag{11}$$

subject to

$$C(s^t) = \sum_{i\in I} \pi^i c^i(s^t), \qquad \text{and} \qquad L(s^t) = \sum_{i\in I} \pi^i \ell^i(s^t) \tag{12}$$

where the superscript m stands for "market."

Proof. See Appendix A.2. □

That is, any equilibrium delivers an efficient assignment of individual consumption and labor $(c^i(s^t), \ell^i(s^t))_{i\in I}$ given aggregates $(C(s^t), L(s^t))$ and market weights φ. The economy thus behaves *as if* there exists a representative household with utility function $U^m(C, L; \varphi)$. Relative prices satisfy the representative household's intratemporal condition:

$$-\frac{U_L^m(s^t)}{U_C^m(s^t)} = \left(\frac{1-\tau_\ell}{1+\tau_c}\right)\frac{W(s^t)}{P(s^t)}, \qquad \forall s^t \in S^t, \tag{13}$$

and intertemporal conditions:

$$Q(s^{t+1}|s^t) = \beta\mu(s^{t+1}|s^t)\frac{U_C^m(s^{t+1})}{U_C^m(s^t)}\frac{P(s^t)}{P(s^{t+1})}, \qquad \forall s^{t+1} \in S^{t+1}, \tag{14}$$

$$\frac{U_C^m(s^t)}{P(s^t)} = \beta(1+i(s^t))\sum_{s^{t+1}|s^t}\mu(s^{t+1}|s^t)\frac{U_C^m(s^{t+1})}{P(s^{t+1})}, \qquad \forall s^t \in S^t. \tag{15}$$

where we let $U_C^m(s^t) \equiv \partial U^m(\cdot)/\partial C(s^t)$ and $U_L^m(s^t) \equiv \partial U^m(\cdot)/\partial L(s^t)$ denote the representative household's marginal utilities with respect to aggregate consumption and aggregate labor. Condition (13) indicates that the representative household's marginal rate of substitution between consumption and labor is equal to the after-tax real wage, condition (15) is the bond Euler equation, and conditions (14) are the Euler equations for each specific Arrow security.

[5] See Appendix A.1 for the full derivation of the households' optimality conditions.

From the envelope condition of the static sub-problem, $U_C^m(s^t) = \varphi^i U_c^i(s^t)$ and $U_L^m(s^t) = \varphi^i U_\ell^i(s^t)$, where we let $U_c^i(s^t) \equiv \partial U(\cdot)/\partial c^i(s^t)$ and $U_\ell^i(s^t) \equiv \partial U(\cdot)/\partial \ell^i(s^t)$ denote household i's marginal utilities with respect to individual consumption and labor.[6] Therefore equations (13)-(15) hold with U^i in place of U^m, and individual household's marginal rates of substitution are equated to after-tax prices.

With general preferences, the unique solution to the static sub-problem in Lemma 1 implies that individual household consumption and labor can be written as functions of aggregates $(C(s^t), L(s^t))$, the Negishi weights φ, and the distribution $(\theta^i(s_t))_{i\in I}$ alone; see Werning (2007). With the additively-separable and iso-elastic preferences assumed in (1), the solution can be written in closed form:

$$c^i(s^t) = \omega_C^i(\varphi)C(s^t) \qquad \text{and} \qquad \ell^i(s^t) = \omega_L^i(\varphi, s_t)L(s^t), \tag{16}$$

with

$$\omega_C^i(\varphi) \equiv \frac{(\varphi^i)^{1/\gamma}}{\sum_{k\in I} \pi^k (\varphi^k)^{1/\gamma}} \qquad \text{and} \qquad \omega_L^i(\varphi, s_t) \equiv \frac{(\varphi^i)^{-1/\eta}\theta^i(s_t)^{\frac{1+\eta}{\eta}}}{\sum_{k\in I} \pi^k (\varphi^k)^{-1/\eta}\theta^k(s_t)^{\frac{1+\eta}{\eta}}}. \tag{17}$$

Therefore, with these preferences, individual consumption and labor are proportional to their aggregates.

The household's shares of aggregate consumption and aggregate labor are given by $\omega_C^i(\varphi)$ and $\omega_L^i(\varphi, s_t)$, respectively. The consumption share is fixed and depends only on the market weights, φ, and the risk aversion parameter, γ. Markets are complete—as a result, individual households insure away all idiosyncratic risk in consumption and face only aggregate risk. In contrast, the share of labor is a function of the market weights, φ, the Frisch elasticity of labor supply, η, as well as the entire distribution of worker productivities $(\theta^i(s_t))_{i\in I}$. The household's share of labor supply is thereby state-contingent: it depends on the household's relative skill in state s_t.[7]

In equilibrium, each household's budget constraint (2) holds with equality. Using equations (13)-(15) to substitute out after-tax prices, we obtain the following implementability conditions:

$$\sum_t \sum_{s^t} \beta^t \mu(s^t) \left[U_C^m(s^t)\omega_C^i(\varphi)C(s^t) + U_L^m(s^t)\omega_L^i(\varphi, s_t)L(s^t)\right] = U_C^m(s_0)\bar{T}, \qquad \forall i \in I, \tag{18}$$

where

$$\bar{T} \equiv \frac{(1+\tau_c)^{-1}}{U_c^m(s_0)} \sum_t \sum_{s^t} \beta^t \mu(s^t) U_C^m(s^t) \left[T(s^t) + (1-\tau_\Pi)\frac{\Pi(s^t)}{P(s^t)}\right]. \tag{19}$$

The above implementability conditions are expressed entirely in terms of the aggregate allocation $(C(s^t), L(s^t))$ and the market weights φ. See Appendix A.3 for their derivation.

[6] Note that $\frac{\partial U(\cdot)}{\partial \ell^i(s^t)} = \frac{1}{\theta^i(s_t)}\frac{\partial U(\cdot)}{\partial h^i(s^t)}$.

[7] Although our model features ex-post differences in labor supply across households, these differences reflect an efficient allocation of labor supply for a given level of aggregate labor. For a discussion of this point, see Werning (2007).

Condition (18) corresponds to household i's lifetime budget constraint and is similar to the standard implementability condition found in the Ramsey taxation literature; see Lucas and Stokey (1983); Chari and Kehoe (1999). However, in contrast to representative agent economies, rather than equilibrium imposing just one implementability condition of the form in (18), in our economy there exists a *set* of conditions: one for each type $i \in I$.

As noted previously, one stark difference between our framework and the representative-agent Ramsey framework is the existence of lump-sum taxes and transfers, as in Werning (2007). When coupled with labor income taxes, these lump-sum transfers give the planner some ability to redistribute. This power, however, is limited: the planner cannot achieve *any* desired distribution of resources across households because lump-sum transfers are non-targeted. To see this, note that the right hand side of equation (18) represents the present discounted value of lifetime transfers and after-tax profits, denoted by $\bar{T}$, and this value is the same across all types $i \in I$. It follows that the conditions in (18) are joint restrictions on the planner's problem.

Furthermore, in the representative agent Ramsey framework, not only does one typically rule out lump-taxes, but also any combination of taxes that may replicate them. When consumption and labor income taxes are available, this applies to the initial period consumption tax—one can set the initial period consumption tax arbitrarily high and achieve the undistorted optimum. Typically to rule this out, one must treat the initial consumption tax as separate from all other period consumption taxes and impose a binding upper bound; see Chari and Kehoe (1999). Here we have no such issue because we assume the existence of lump sum taxes. It follows that we need no such restriction on the initial period consumption tax; in fact, we simply subsume it into our definition of $\bar{T}$. Furthermore, as noted above, one can see that state-contingency of the lump-sum transfer is unnecessary–due to the complete markets assumption, all agents in our model are Ricardian.

Finally, in our framework, due to the monopolistic competition assumption, intermediate-good firms earn equilibrium profits. Equilibrium profits would presumably complicate our analysis as they enter endogenously into household budget constraints as dividend payouts. However, from condition (19) it is evident that profits are isomorphic to lump-sum transfers. This equivalence relies on the assumption of homogeneous equity shares across households; we relax this assumption in Section 7.

Firms. We now turn to the firms' problems and begin by considering that of the flexible-price firms, $j \in \mathcal{J}^f$. A flexible-price firm is attentive to the current state and can therefore choose a price measurable in the history s^t. We state the firms' problem as follows: firm $j \in \mathcal{J}^f$ chooses a nominal price $p_t^j(s^t)$ to maximize firm profits:

$$p_t^j(s^t) \in \arg\max_{p'} \left\{ (1-\tau_r)p' y^j(s^t) - \frac{W(s^t)}{A(s_t)} y^j(s^t) \right\}$$

subject to demand function:

$$y^j(s^t) = \left(\frac{p'}{P(s^t)}\right)^{-\rho} Y(s^t), \qquad \forall s^t \in S^t, \tag{20}$$

The solution to this problem is given by:

$$p_t^f(s^t) = \left[(1-\tau_r)\left(\frac{\rho-1}{\rho}\right)\right]^{-1} \frac{W(s^t)}{A(s_t)}, \qquad \forall s^t \in S^t. \tag{21}$$

That is, the firm optimally equates its marginal cost with its after-tax marginal revenue. This implies that the firm's optimal nominal price equal to a constant mark-up over its nominal marginal cost $W(s^t)/A(s_t)$. The mark-up is a function of the CES parameter ρ and the marginal tax (or subsidy) on revenue.

Consider next the problem of the sticky-price firms, $j \in \mathcal{J}^s$. Sticky-price firms are inattentive to the current state s_t and hence make their nominal pricing decisions based only on their knowledge of the history of previous states, s^{t-1}. Recall that all firms are owned by the households; and the fictitious representative household's stochastic discount factor in state s^t is given by $U_C^m(s^t)/P(s^t)$. From our previous derivation of household optimality, the Arrow security price $Q(s^t|s^{t-1})$ satisfies (14) and hence can be interpreted as the firm's pricing kernel. We may therefore write the sticky price firm's problem as follows.

Firm $j \in \mathcal{J}^s$ chooses a nominal price $p_t^j(s^{t-1})$ such that it maximizes the expected value of firm profits (weighted appropriately by the market's stochastic discount factor):

$$p_t^j(s^{t-1}) \in \arg\max_{p'} \sum_{s^t|s^{t-1}} Q(s^t|s^{t-1}) \left\{(1-\tau_r)p' y^j(s^t) - \frac{W(s^t)}{A(s_t)} y^j(s^t)\right\}$$

subject to (20). The solution to the sticky-price firm's problem is given by:

$$p_t^s(s^{t-1}) = \left[(1-\tau_r)\left(\frac{\rho-1}{\rho}\right)\right]^{-1} \sum_{s^t|s^{t-1}} \frac{W(s^t)}{A(s_t)} q(s^t|s^{t-1}) \tag{22}$$

where we let

$$q(s^t|s^{t-1}) \equiv \frac{\mu(s^t|s^{t-1}) U_C^m(s^t) Y(s^t) P(s^t)^{\rho-1}}{\sum_{s^t|s^{t-1}} \mu(s^t|s^{t-1}) U_C^m(s^t) Y(s^t) P(s^t)^{\rho-1}} \tag{23}$$

denote the risk-adjusted conditional probabilities of sticky-price firm j, conditional on history s^{t-1}. Note that these probabilities satisfy for all s^{t-1}, $\sum_{s^t} q(s^t|s^{t-1}) = 1$, by construction. Therefore, the firm's optimal price is equal to a markup over its risk-weighted expectation of its nominal marginal cost, $W(s^t)/A(s_t)$, conditional on information set s^{t-1}.[8]

Comparing this to the optimal price of the flexible-price firm, (21), one can rewrite (22) in the following manner: $p_t^s(s^{t-1}) = \sum_{s^t} q(s^t|s^{t-1}) p_t^f(s^t)$. That is, the optimal price of the sticky-price firm is equal to its risk-weighted expectation, conditional on information set s^{t-1}, of the optimal price of the flexible-price firm (Correia, Nicolini and Teles, 2008).

[8]See Appendix A.4 for a derivation of (22) and (23).

3.3 Equilibrium Allocations

We now characterize the set of allocations that can be implemented as a competitive equilibrium under flexible prices as well as under sticky prices.

Flexible-price equilibria. Consider first the equilibrium under flexible prices. In any such equilibrium, all firms set their price according to (21). Combining this with the fictitious representative household's optimality conditions, we obtain the following result.

Proposition 1. *A feasible allocation $x \in \mathcal{X}$ can be implemented as a flexible-price equilibrium if and only if there exist market weights $\varphi \equiv (\varphi^i)$, a scalar $\bar{T} \in \mathbb{R}$, and a strictly positive scalar $\chi \in \mathbb{R}_+$, such that the following three sets of conditions are jointly satisfied:*

(i) for all $s^t \in S^t$, $y^j(s^t) = Y(s^t)$ for all $j \in \mathcal{J}$;

(ii) for all $s^t \in S^t$,

$$-\frac{U_L^m(s^t)}{U_C^m(s^t)} = \chi A(s_t); \qquad \textit{and} \tag{24}$$

(iii) condition (18) holds for every $i \in I$.

Proof. See Appendix A.5. □

Proposition 1 characterizes the entire set of allocations that can be supported as a flexible-price equilibrium; for shorthand we call such allocations "flexible-price allocations." In addition to resource and technology constraints, any flexible-price allocation satisfies three sets of constraints described in parts (i)-(iii) of the proposition.

Part (i) of Proposition 1 indicates that in any flexible-price equilibrium, there is no output dispersion across firms. Firms share the same technology and face the same nominal wages; as a result they choose the same prices as in (21). It follows from their demand functions (4) that, in any flexible-price equilibrium, all firms produce identical levels of output.

Next, part (ii) of Proposition 1 states that in any flexible-price equilibrium, condition (24) must hold in every history. This condition follows from taking the optimality condition of the flexible-price firms, 21, noting that in equilibrium all firms set the same nominal price: $p_t^f(s^t) = P(s^t)$, and combining this with the representative household's intratemporal condition in (13).

Therefore, in any flexible-price equilibrium, the marginal rate of substitution between aggregate consumption and aggregate labor is equated with the marginal rate of transformation, $A(s_t)$, modulo a constant labor wedge, denoted by χ. This wedge is given by:

$$\chi \equiv \left(\frac{\rho - 1}{\rho}\right) \frac{(1-\tau_\ell)(1-\tau_r)}{1+\tau_c}. \tag{25}$$

The wedge is the product of multiple terms: the consumption, sales, and labor income taxes levied by the government, and the markup that arises due to monopolistic-competition among intermediate-good producers. It is important to note that χ is a time and state-invariant constant—this follows from the assumption that the tax rates, as well as the elasticity of substitution parameter, ρ, are not contingent on the aggregate state.

Finally part (iii) states that in any flexible-price equilibrium, condition (18) must hold for every $i \in I$. These implementability conditions ensure that every households' lifetime budget constraint is satisfied. The government's budget constraint holds by Walras's Law.

The power of fiscal policy. The flexible-price economy allows us to isolate the role of fiscal policy in our environment. In particular, the power of the fiscal authority is parameterized by the scalars χ and $\bar{T}$. Consider χ: the fiscal policy can control, via the linear taxes in (25), this wedge. However, note that the fiscal authority's power to influence allocations using this instrument is limited: χ is a scalar, but condition (24) must hold in every history, $s^t \in S^t$. Therefore, because we have assumed non-state-contingent tax rates, the set of feasible allocations that can be implemented as a flexible price equilibrium is constrained.

Next, consider the scalar $\bar{T}$. The fiscal policy can use lump-sum transfers (or taxes) to control the level of the households' budget constraints. However, again the fiscal authority's power to influence allocations using this instrument is limited: condition (18) must hold for every household type $i \in I$, as we have assumed lump-sum transfers are non-targeted. Therefore conditions (18) constrain the set of feasible allocations that can be implemented as a flexible price equilibrium.

Sticky-Price Equilibria. We turn now to the set of allocations that can be supported as part of an equilibrium under sticky prices. In any sticky-price equilibrium, all sticky-price firms set their prices according to (22) and all flexible-price firms set their prices according to (21). It follows from the demand functions (4) that all sticky-price firms produce the same level of output, hire the same amount of labor, and earn the same level of profits; we henceforth denote these objects by $y^s(s^t)$, $n^s(s^t)$, and $\pi^s(s^t)$, respectively. By the same logic, we denote their output, labor, and profits of the flexible-price firms by $y^f(s^t)$, $n^f(s^t)$, and $\pi^f(s^t)$, respectively. This brings us to the following equilibrium characterization.

Proposition 2. *A feasible allocation $x \in \mathcal{X}$ can be implemented as a sticky-price equilibrium if and only if there exist market weights $\varphi \equiv (\varphi^i)$, a scalar $\bar{T} \in \mathbb{R}$, and a strictly positive scalar $\chi \in \mathbb{R}_+$, such that the following three sets of conditions are jointly satisfied:*

(i) for all $s^t \in S^t$, $y^j(s^t) = y^f(s^t)$ for all $j \in \mathcal{J}^f$, and $y^j(s^t) = y^s(s^t)$ for all $j \in \mathcal{J}^s$;

(ii) for all $s^t \in S^t$,

$$\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho} + \frac{U_L^m(s^t)}{U_C^m(s^t)}\frac{1}{\chi A(s_t)} = 0; \tag{26}$$

and for all $s^{t-1} \in S^{t-1}$,

$$\sum_{s^t|s^{t-1}} U_C^m(s^t) y^s(s^t) \left\{ \left[\frac{y^s(s^t)}{Y(s^t)} \right]^{-1/\rho} + \frac{U_L^m(s^t)}{U_C^m(s^t)} \frac{1}{\chi A(s_t)} \right\} \mu(s^t|s^{t-1}) = 0; \qquad \textit{and} \tag{27}$$

(iii) condition (18) *holds for every* $i \in I$.

Proof. See Appendix A.6. □

Proposition 2 characterizes the entire set of allocations that can be supported as a sticky-price equilibrium; for shorthand we call such allocations "sticky-price allocations." Similar to Proposition 1, Proposition 2 states that, aside from satisfying resource and technology constraints, any sticky-price allocation satisfies three additional sets of constraints.

Part (i) indicates that in any sticky-price equilibrium, there is no output dispersion within the set of sticky-price firms and similarly no output dispersion within the set of flexible-price firms. However, there can be differences in production across the two sets of firms.

Part (ii) states that in any sticky-price equilibrium, condition (26) must hold in every history. This condition follows from combining the optimality condition of the flexible-price firms with the fictitious representative household's intratemporal condition (13). The resulting condition simply states that the marginal cost of producing an extra unit of output of the flexible-price firm is equated with its marginal revenue. Note that this condition is similar to condition (24) in Proposition 1, and in fact is identical when $y^f(s^t) = Y(s^t)$.

Condition (27) similarly follows from combining the optimality condition for the sticky-price firms with the fictitious representative household's intratemporal optimality condition. This condition states that the marginal cost of producing an extra unit of output of the sticky-price firm is equated with its marginal revenue "on average." It is essentially the same as condition (26) corresponding to flexible-price firm optimality, the only difference being that in (27), marginal cost and marginal revenue are equated in expectation, conditional on information set s^{t-1}.

Finally, part (iii) of Proposition 2 is identical to part (iii) of Proposition 1; these conditions ensure that the budget constraint is satisfied for every household in the economy.

3.4 The power of monetary policy.

To understand the power of monetary policy vis-a-vis fiscal policy in this economy, it is easier to rewrite the equilibrium conditions for the sticky price economy in the following manner. First, we can rewrite the optimal price of the sticky-price firm in (22) as follows:

$$p_t^s(s^{t-1}) = \epsilon(s^t) \left[(1-\tau_r) \left(\frac{\rho-1}{\rho} \right) \right]^{-1} \frac{W(s^t)}{A(s_t)}, \tag{28}$$

where $\epsilon(s^t)$ is defined as:

$$\epsilon(s^t) \equiv \frac{\sum_{s^t|s^{t-1}} q(s^t|s^{t-1}) W(s^t)/A(s_t)}{W(s^t)/A(s_t)}. \tag{29}$$

Thus $\epsilon(s^t)$ denotes a stochastic wedge between the optimal prices of the sticky- and flexible-price firms: $p_t^s(s^{t-1}) = \epsilon(s^t) p_t^f(s^t)$. Because the sticky-price firm has incomplete information, it cannot perfectly forecast its ex-post optimal price, i.e. a markup over its nominal marginal cost. The wedge $\epsilon(s^t)$ can therefore be interpreted as the sticky-price firm's "pricing mistake." Formally, $\epsilon(s^t)$ is defined in (29) as the firm's optimal "forecast error" of its nominal marginal cost, $W(s^t)/A(s_t)$, given its incomplete information set s^{t-1}.

Next, aggregating over the sticky-price and flexible-price firm prices according to (9) and combining the aggregate price level with the representative household's intratemporal optimality condition (13), we obtain the following equilibrium condition in the sticky-price economy:

$$-\frac{U_L^m(s^t)}{U_C^m(s^t)} = \chi \left[\kappa \epsilon(s^t)^{1-\rho} + (1-\kappa)\right]^{-\frac{1}{1-\rho}} A(s_t).$$

This condition looks similar to condition (24) for the flexible price economy. As in the flexible-price economy, it indicates that the marginal rate of substitution between aggregate consumption and aggregate labor is equated with the marginal rate of transformation, $A(s_t)$, modulo a wedge. In this case, though, the labor wedge is the product of two components. The first is the constant scalar denoted by χ that corresponds to the mark-up and taxes (25). The second is a new, state-contingent component given by $\left[\kappa \epsilon(s^t)^{1-\rho} + (1-\kappa)\right]^{-\frac{1}{1-\rho}}$. This component contains the state-contingent "pricing mistakes" made by the fraction κ of inattentive firms.

Therefore, relative to the flexible-price economy, there is a state-contingent component of the labor wedge. This state-contingent component, in particular the forecast error $\epsilon(s^t)$, represents an *additional* control variable for the planner in the sticky-price economy, one that encompasses the power of monetary policy over real allocations. By controlling $\epsilon(s^t)$, the monetary authority can move around allocations in a manner that the fiscal authority cannot.

However, the power of monetary policy is limited by parts (i) and (ii) of Proposition 2. As the forecast error of the sticky-price firms, $\epsilon(s^t)$ introduces a wedge between the sticky-price and flexible-price firms' prices. This in turn drives a wedge between the sticky-price and flexible-price firms' output: $y^s(s^t) = \epsilon(s^t)^{-\rho} y^f(s^t)$, leading to a loss in production efficiency.

Second, by definition of $\epsilon(s^t)$ in (29), the forecast error must "on average" be equal to 1. This in fact is the meaning of the constraint in (27): it is an adding up constraint that states that monetary policy cannot surprise firms "on average." This follows from the optimal price-setting behavior of the sticky-price firms.

Finally, to drive home the point that sticky prices and monetary policy enlarges the set of implementable allocations, we provide the following lemma.

Lemma 2. *Let $\mathcal{X}^f$ denote the set of all flexible-price allocations and let $\mathcal{X}^s$ denote the set of all sticky-price allocations.*

$$\mathcal{X}^f \subset \mathcal{X}^s.$$

Proof. Take any allocation $x \in \mathcal{X}^f$; that is, x satisfies the conditions stated in Proposition 1. This allocation satisfies all conditions stated in Proposition 2 with $\frac{y^s(s^t)}{Y(s^t)} = \frac{y^f(s^t)}{Y(s^t)} = 1$ for all $s^t \in S^t$. Therefore, $x \in \mathcal{X}^s$. □

Therefore, any allocation that can be implemented under flexible-prices can also be implemented under sticky prices. This can be implemented with a monetary policy that targets price stability.

3.5 Welfare function and Ramsey problem definition

Finally, the goal of this paper is to solve the Ramsey problem in this economy. We consider a utilitarian planner with social welfare function given by:

$$\mathcal{U} \equiv \sum_{i\in I} \lambda^i \pi^i \sum_t \sum_{s^t} \beta^t \mu(s^t) U(c^i(s^t), \ell^i(s^t)/\theta^i(s_t)) \tag{30}$$

where $\lambda \equiv (\lambda^i)$ denote an arbitrary set of Pareto weights, with $\lambda^i > 0$ for all $i \in I$. We define the Ramsey problem as follows.

Definition 4. *A Ramsey optimum x^* is an allocation x that maximizes social welfare* (30) *subject to $x \in \mathcal{X}^s$.*

4 The Relaxed Ramsey Problem

The goal of our analysis is to characterize the social welfare-maximizing allocation among the set of sticky-price allocations. However, the set of sticky-price allocations, $\mathcal{X}^s$, is fairly complicated: there are a number of constraints that must be satisfied in order for an allocation to be supported as a sticky-price equilibrium. We thus proceed in this section by first solving an *easier* problem, that of a "relaxed" Ramsey planner.

The relaxed Ramsey planning problem is one in which we maximize over a larger, relaxed set of allocations relative to the set of sticky-price allocations; see Correia, Nicolini and Teles (2008) and Angeletos and La'O (2020) for similar analyses. We define the relaxed set of allocations and an optimum within this set as follows.

Definition 5. *The relaxed set of allocations $\mathcal{X}^R$ is the set of all feasible allocations $x \in \mathcal{X}$ for which there exists a set of market weights $\varphi \equiv (\varphi^i)$ such that condition* (18) *holds for all $i \in I$. A relaxed Ramsey optimum x^{R*} is an allocation x that maximizes social welfare (30) subject to $x \in \mathcal{X}^R$.*

Relative to the set of sticky-price allocations characterized in Proposition 2, the relaxed set is constructed by dropping all equilibrium conditions stated in parts (i) and (ii) of the proposition, but maintaining those stated in part (iii). The following corollary is the direct result of Proposition 2, Lemma 2, and Definition 5.

Corollary 1. $\mathcal{X}^f \subset \mathcal{X}^s \subset \mathcal{X}^R \subset \mathcal{X}$.

The relaxed set is a strict superset of $\mathcal{X}^s$, the set of sticky-price allocations, and by implication, $\mathcal{X}^f$, the set of flexible-price allocations. One can think of the relaxed Ramsey planner as a planner that has access to a complete set of state-contingent, firm- and/or good-specific tax instruments, and can thus freely choose the equilibrium price of *any* good in *any* state, but does not have access to type-specific lump-sum transfers, therefore must respect the lifetime budget constraints of the households.[9]

Why do we study the relaxed Ramsey planning problem? This problem is useful for our analysis in the following sense. We will first characterize the relaxed Ramsey optimum x^{R*}. We will then derive necessary and sufficient conditions under which $x^{R*} \in \mathcal{X}^f$, and by implication, $x^{R*} \in \mathcal{X}^s$. Finally, because the relaxed set is a strict superset of the set of sticky-price allocations, it follows that under these conditions, x^{R*} is both the relaxed Ramsey optimum and the *unrelaxed* Ramsey optimum!

Let $\pi^i \nu^i$ denote the Lagrange multiplier on the implementability condition (18) of type $i \in I$; let $\nu \equiv (\nu^i)_{i\in I}$ denote the set of multipliers. Following Werning (2007), we incorporate these constraints into the planner's maximand and define the pseudo-welfare function $\mathcal{W}(\cdot)$:

$$\begin{aligned}\mathcal{W}(C(s^t), L(s^t), s_t; \varphi, \nu, \lambda) \equiv \sum_{i\in I} \pi^i \{ & \lambda^i U(\omega_C^i(\varphi)C(s^t), \omega_L^i(\varphi, s_t)L(s^t)/\theta^i(s_t)) \\ & + \nu^i \left[U_C^m(s^t)\omega_C^i(\varphi)C(s^t) + U_L^m(s^t)\omega_L^i(\varphi, s_t)L(s^t) \right] \}\end{aligned} \tag{31}$$

We then write the relaxed Ramsey planning problem as follows.

Relaxed Ramsey Planner's Problem. *The Relaxed Ramsey planner chooses an allocation x, market weights $\varphi \equiv (\varphi^i)$, and $\bar{T} \in \mathbb{R}$, in order to maximize*

$$\sum_t \sum_{s^t} \beta^t \mu(s^t) \mathcal{W}(C(s^t), L(s^t), s_t; \varphi, \nu, \lambda) - U_C^m(s_0)\bar{T} \sum_{i\in I} \pi^i \nu^i \tag{32}$$

subject to feasibility: $x \in \mathcal{X}$.

The pseudo-utility function is stated in terms of aggregates alone, making the relaxed Ramsey planning problem quite tractable. One can think of the pseudo-welfare function as a social welfare function that incorporates not only the distributive motives of society, as captured by the Pareto weights, but also the constraints imposed by the households' heterogeneous budget sets.

[9]With a complete set of state-contingent, firm- and/or good-specific tax instruments, parts (i) and (ii) of Proposition 2 are no longer necessary conditions.

4.1 Relaxed Ramsey optimum.

The following proposition characterizes a relaxed Ramsey optimum given an arbitrary set of Pareto weights. For shorthand, we let $\mathcal{W}_C(s^t) \equiv \partial\mathcal{W}(\cdot)/\partial C(s^t)$ and $\mathcal{W}_L(s^t) \equiv \partial\mathcal{W}(\cdot)/\partial L(s^t)$ denote the marginal pseudo-utility of aggregate consumption and of aggregate labor, respectively.

Proposition 3. *An allocation is a relaxed Ramsey optimum x^{R*} if (i) for all $s^t \in S^t$, $y^j(s^t) = Y(s^t)$ for all $j \in \mathcal{J}$; and (ii)*

$$-\frac{\mathcal{W}_L(s^t)}{\mathcal{W}_C(s^t)} = A(s_t), \qquad \forall s^t \in S^t. \tag{33}$$

Proof. See Appendix A.7. □

Consider first part (ii) of Proposition 3. It is optimal, from the relaxed Ramsey planner's perspective, to set the social marginal rate of substitution between consumption and labor equal to the marginal rate of transformation, $A(s_t)$. In this formulation, the social marginal rate of substitution between consumption and labor is given by the ratio of the marginal pseudo-utility of labor to the marginal pseudo-utility of consumption. The social marginal rate of substitution therefore takes into account the Pareto weights, i.e. the planner's appetite for redistribution, as well as the implementability constraints imposed by household budget sets.

Consider now part (i). It is furthermore optimal, from the relaxed Ramsey planner's perspective, that there be zero output dispersion across intermediate good firms. The relaxed Ramsey optimum thereby preserves production efficiency in the sense of Diamond and Mirrlees (1971). Although the planner chooses to tax certain margins in order to raise money to support lump-sum transfers (or taxes), it does so at the intratemporal margin.

Preservation of production efficiency indicates that a relaxed Ramsey optimum *could be* a flexible-price allocation—in any flexible-price equilibrium, there is zero cross-sectional dispersion in output—but it does not yet tell us *when* such an allocation is implementable under flexible prices. The following result provides an answer.

Theorem 1. *If there exist positive scalars $(\vartheta^1, \vartheta^2, \dots \vartheta^I) \in \mathbb{R}_+^I$ and a positively-valued function $\Theta : S \to \mathbb{R}_+$ such that the skill distribution satisfies:*

$$\theta^i(s_t) = \vartheta^i \Theta(s_t), \qquad \forall s_t \in S, \tag{34}$$

then:

(i) the relaxed Ramsey optimum is implementable as a flexible-price allocation, $x^{R*} \in \mathcal{X}^f$;
(ii) the relaxed Ramsey optimum is implementable as a sticky-price allocation, $x^{R*} \in \mathcal{X}^s$; and
(iii) the relaxed Ramsey optimum x^{R*} is an (unrelaxed) Ramsey optimum, x^*.

Proof. Suppose there exists positive scalars $(\vartheta^1, \vartheta^2, \ldots \vartheta^I) \in \mathbb{R}^I_+$ and a function $\Theta : S \to \mathbb{R}_+$ such that (34) is satisfied. The individual household shares defined in (17) reduce to:

$$\omega^i_C(\varphi) \equiv \frac{(\varphi^i)^{1/\gamma}}{\sum_{j\in I} \pi^j (\varphi^j)^{1/\gamma}} \qquad \text{and} \qquad \omega^i_L(\varphi) \equiv \frac{(\varphi^i)^{-1/\eta}(\vartheta^i)^{\frac{1+\eta}{\eta}}}{\sum_{k\in I} \pi^k (\varphi^k)^{-1/\eta}(\vartheta^k)^{\frac{1+\eta}{\eta}}},$$

and are therefore non-state-contingent. The relaxed Ramsey optimality condition in (33) can then be written as follows:

$$-\frac{U^m_L(s^t)}{U^m_C(s^t)} \left[\frac{\sum_{i\in I} \pi^i \omega^i_L(\varphi)\left(\lambda^i/\varphi^i + \nu^i(1+\eta)\right)}{\sum_{i\in I} \pi^i \omega^i_C(\varphi)\left(\lambda^i/\varphi^i + \nu^i(1-\gamma)\right)} \right] = A(s_t) \tag{35}$$

Comparing this to the flexible-price intratemporal condition (24), it is clear that (35) can be replicated under flexible prices with an appropriate choice of scalar χ. This proves part (i) of the theorem; part (ii) follows directly from Lemma 2. Finally, part (iii) follows from the fact that x^{R*} is the welfare-maximizing allocation in $\mathcal{X}^R$, and $x^{R*} \in \mathcal{X}^s \subset \mathcal{X}^R$. □

Theorem 1 provides a sufficient condition under which a relaxed Ramsey optimum can be implemented under flexible prices. We henceforth refer to a skill distribution that satisfies this property as one that "exhibits only proportional aggregate shocks."

To understand the intuition behind Theorem 1, it is helpful to first think about the problem of the relaxed Ramsey planner, a planner constrained only by the feasibility of allocations and the household budget implementability conditions (18). This planner faces a trade-off between the benefit of redistribution and its cost. The cost of redistribution is efficiency: if the planner would like to achieve a more equal distribution of resources across households than under laissez-faire, the planner must distort the fictitious household's intratemporal margin between aggregate consumption and aggregate labor, i.e. the after-tax real wage, in order to raise tax revenue. The relaxed Ramsey planner's optimum is thus the point at which, in every state, the marginal benefit of redistribution is equal to the marginal cost; this state-by-state trade-off is captured in the planner's intratemporal optimality condition (33).

Now consider whether this optimum can be achieved under flexible prices. Suppose first that there are no shocks in the economy: TFP and the labor skill distribution is fixed; one can think of this as the economy's "non-stochastic steady state." In the absence of shocks, the marginal benefit of redistribution and its marginal cost are both constant over time. It follows that the relaxed Ramsey optimum can be implemented as an equilibrium under flexible prices with some constant level of distortion χ. A higher tax rate (equivalently, a lower χ) means that greater tax revenue can be collected from high-skilled, wealthy households than from low-skilled, poor households. Because such tax revenue is redistributed via a uniform, lump sum transfer, a greater tax rate implies greater redistribution. The optimal, constant tax rate thus balances the relaxed Ramsey planner's distributional concerns against efficiency in the "non-stochastic steady state" of the economy.

Now consider the case in which there are shocks to TFP and to the labor skill distribution, but we restrict the latter to feature only proportional aggregate shocks $\Theta(s_t)$ as described in Lemma 1. When such is the case, the ratio of labor productivity between any two household types remains constant over time and over states:

$$\frac{\theta^i(s_t)}{\theta^j(s_t)} = \frac{\vartheta^i}{\vartheta^j}, \qquad \forall s_t \in S.$$

As a result, because there are no shocks to the *relative* skill distribution, the marginal benefit from redistribution *does not vary* over the business cycle. Because the marginal cost also does not vary over the business cycle (technology and preferences are homothetic), the optimum at which the marginal benefit of redistribution equals the marginal cost is invariant to the aggregate state. It follows that the optimal level of redistribution can be achieved under flexible prices with a constant level of distortion χ; this is the result described in Lemma 1.

Finally, when the relaxed Ramsey optimum can be achieved under flexible prices—that is, when the tax system is sufficient to achieve the optimal level of redistribution—then the best that monetary policy can do is to replicate flexible price allocations. We show in Section 6 that it can do so by targeting a constant price level.[10]

Note that the key property that drives this result is the preservation of Diamond and Mirrlees (1971) production efficiency at the relaxed Ramsey optimum. In this sense Theorem 1 is similar to the key insight in Correia, Nicolini and Teles (2008). Although the planner in our environment trades off redistribution with a wedge that distorts the fictitious representative household's intratemporal margin, under no circumstances does the relaxed planner find it optimal to misallocate resources *across* firms. Thus, with homothetic preferences and proportional aggregate shocks to the labor skill distribution, there is no reason for monetary policy to introduce such distortions.

Homotheticity. The homotheticity assumption on preferences plays a role in generating the above results. In the proof of Theorem 1, we use the fact that the equilibrium allocation of consumption and labor across households take the form given in (16), which itself relies on the iso-elastic preference specification.

Necessity. Theorem 1 provides a sufficient condition under which a relaxed Ramsey optimum can be implemented under flexible prices but this condition is not necessary. To see this, note that there exists a degenerate case in which the Pareto weights are exactly equal to the Negishi weights at the laissez-faire equilibrium. In this case, $\lambda^i = \varphi^i$ for all $i \in I$, in which case $\nu^i = 0$ for all $i \in I$. In other words, given any stochastic process of the skill distribution, there exists a knife edge case of Pareto weights such that the laissez-faire equilibrium under flexible prices

[10] Equivalently, by setting $\epsilon(s^t) = 1$ for all $s^t \in S^t$.

is always a relaxed Ramsey optimum. Furthermore, this allocation is first best efficient: the implementability conditions in (18) are slack at this optimum.

4.2 Proportional shocks but suboptimal fiscal policy.

We can in fact strengthen the result on monetary policy provided in Theorem 1. Theorem 1 assumes that the planner can choose both monetary *and* fiscal policy. Consider now the special case in which labor skill shocks are proportional, but the Ramsey planner has no longer the ability to control fiscal policy and can only control monetary policy. Even if fiscal policy is set suboptimally, our next result shows that our conclusion on monetary policy in this environment remains unchanged.

Proposition 4. *If there exist positive scalars* $(\vartheta^1, \vartheta^2, \ldots \vartheta^I) \in \mathbb{R}^I_+$ *and a positively-valued function* $\Theta : S \to \mathbb{R}_+$ *such that the skill distribution satisfies* (34)*, and taxes are set such that:*

$$\chi = \left(\frac{\rho - 1}{\rho}\right) \frac{(1-\tau_\ell)(1-\tau_r)}{1+\tau_c} \neq \chi^*,$$

then it remains optimal for monetary policy to implement the flexible-price allocation.

Proof. See Appendix A.9. □

If the labor skill distribution exhibits only proportional aggregate shocks, then it is optimal for monetary policy to implement flexible price allocations, regardless of the fiscal policy. This generalizes the result provided in Theorem 1 to all fiscal policies (within the affine tax structure), including sub-optimal policies.

Why is this the case? The tax rate is set suboptimally, $\chi \neq \chi^*$, in which case one might presume that monetary policy ought to substitute for the missing tax (wedge). Yet this reasoning would be wrong: with proportional shocks monetary policy should continue to target flexible-price allocations regardless of the fiscal authority's mistake.

To understand why, note that the missing tax wedge is constant across all histories s^t. But recall that the only power monetary policy has over allocations is through the forecast error, $\epsilon(s^t)$. Because this is a forecast error (or wedge), it must "add up" by definition to one over all states conditional on the history: $s^t|s^{t-1}$. Therefore, if the monetary authority were to raise this wedge in one state, it cannot do so unless it lowers it in another state. If the missing tax wedge is a constant across all states, there is no reason one state should be favored over another, and moving around $\epsilon(s^t)$ in such a fashion only worsens the allocation. As result, it is best for monetary policy to do absolutely nothing at all and target price stability.

5 The Ramsey Problem

We have shown that the relaxed Ramsey optimum can be implemented as a sticky-price equilibrium under very special circumstances: proportional aggregate shocks to the labor skill distribution. Away from the proportional shock case, though, it is not readily obvious what the optimal sticky-price allocation is, and hence what monetary policy should do. In order to answer this question, we now return to our original problem, that of the "unrelaxed" Ramsey planner, as defined in 4.

Again letting $\pi^i \nu^i$ denote the Lagrange multipliers on the budget implementability conditions (18), we incorporate the households' budget implementability conditions into the planner's maximand via the pseudo-utility function. The Ramsey planning problem can then be written in the following manner.

Ramsey Planner's Problem. *The Ramsey planner chooses an allocation:*

$$x \equiv \{y^s(s^t), y^f(s^t), C(s^t), Y(s^t), L(s^t)\}_{t\geq 0, s^t \in S^t}$$

market weights $\varphi \equiv (\varphi^i)$*, constants* $\bar{T} \in \mathbb{R}$ *and* $\chi \in \mathbb{R}_+$*, in order to maximize* (32)*, subject to:*

$$C(s^t) = Y(s^t) = \left[\kappa y^s(s^t)^{\frac{\rho-1}{\rho}} + (1-\kappa) y^f(s^t)^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}}, \qquad L(s^t) = \kappa \frac{y^s(s^t)}{A(s_t)} + (1-\kappa)\frac{y^f(s^t)}{A(s_t)}, \quad (36)$$

(26)*, and* (27).

Unlike the relaxed planner, the unrelaxed Ramsey planner is subject to *all* equilibrium implementability conditions (Proposition 2), including conditions (26) and (27). In Section A.8 of the Appendix we provide a complete characterization of the Ramsey optimum.[11] While we do not provide the full characterization here for exposition and conciseness, the essential necessary condition of the planner's optimum appears as follows:

$$-\frac{\mathcal{W}_L(s^t)}{\mathcal{W}_C(s^t)}\left[\text{Ramsey wedge}(s^t)\right] = \frac{Y(s^t)}{L(s^t)}. \quad (37)$$

Condition (37) is the Ramsey planner's intratemporal optimality condition; it is the counterpart to condition (33) of the relaxed Ramsey optimum. The Ramsey planner sets the social marginal rate of substitution between consumption and labor, $-\mathcal{W}_L(s^t)/\mathcal{W}_C(s^t)$ equal to the marginal rate of transformation, $Y(s^t)/L(s^t)$, modulo a state-contingent wedge. Relative to the relaxed planner, the Ramsey planner is subject to implementability conditions (26) and (27); the wedge in condition (37) is a function of their state-contingent Lagrange multipliers. When these conditions are non-binding, the multipliers on them are equal to zero and condition (73)

[11] See Proposition 8 in Appendix A.8 and its proof.

reduces to (33). In this case, the social MRS is equalized with the MRT. When these conditions are binding, the ratio $\mathcal{W}_L(s^t)/\mathcal{W}_C(s^t)$ departs from the MRT at the Ramsey optimum.

Furthermore, in contrast to the relaxed Ramsey optimum, note that the marginal rate of transformation between labor and consumption at the Ramsey optimum is no more $A(s_t)$, but instead $Y(s^t)/L(s^t)$. As long as the planner finds it optimal to deviate from flexible-price allocations, and can do so using monetary policy, there will be misallocation between sticky- and flexible-price firms and thereby a loss in production efficiency.

We are interested in what the Ramsey optimum implies for optimal monetary policy. To that extent, we follow the Ramsey literature and analyze the implicit labor wedge that implements the Ramsey optimum in equilibrium. That is, given a Ramsey optimal allocation x^*, we define the implicit "monetary wedge," $1-\tau_M^*(s^t)$, by:

$$-\frac{U_L^m(s^t)}{U_C^m(s^t)} = \chi^*(1-\tau_M^*(s^t))\frac{Y(s^t)}{L(s^t)}, \tag{38}$$

where χ^* denotes the implicit fiscal wedge at this allocation. The following theorem provides a characterization of $\tau_M^*(s^t)$, the optimal "monetary tax" at the Ramsey optimum x^*.

Theorem 2. *Let $\mathcal{I}: S \to \mathbb{R}_+$ be a positively-valued function defined by:*

$$\mathcal{I}(s_t) \equiv \frac{\sum_{i\in I}\tilde{\pi}^i(\varphi^i)^{-1/\eta}(\theta^i(s_t))^{\frac{1+\eta}{\eta}}}{\sum_{i\in I}\pi^i(\varphi^i)^{-1/\eta}(\theta^i(s_t))^{\frac{1+\eta}{\eta}}} > 0, \qquad \textit{where} \qquad \tilde{\pi}^i \equiv \pi^i\left[\frac{\lambda^i}{\varphi^i}+\nu^i(1+\eta)\right]. \tag{39}$$

There exists a threshold $\bar{\mathcal{I}}(s^{t-1}) > 0$, measurable in history s^{t-1}, such that the optimal implicit monetary tax $\tau_M^*(s^t)$ satisfies:

$$\begin{aligned}
\tau_M^*(s^t) > 0 \qquad & \textit{if and only if} \quad \mathcal{I}(s_t) > \bar{\mathcal{I}}(s^{t-1}),\\
\tau_M^*(s^t) = 0 \qquad & \textit{if and only if} \quad \mathcal{I}(s_t) = \bar{\mathcal{I}}(s^{t-1}),\\
\tau_M^*(s^t) < 0 \qquad & \textit{if and only if} \quad \mathcal{I}(s_t) < \bar{\mathcal{I}}(s^{t-1}).
\end{aligned}$$

Proof. See Appendix A.10. □

The function $\mathcal{I}(s_t)$ can be interpreted as a sufficient statistic for the level of labor income inequality in this economy. First, recall that λ^i are the Pareto weights, φ^i are the market weights, and ν^i are the multipliers on the implementability conditions in (18). Notably these are all non-state-contingent scalars. As a result, $\mathcal{I}: S \to \mathbb{R}_+$ is a function of the current state alone, not the entire history, and in particular depends only on the labor skill distribution, $(\theta^i(s_t))_{i\in I}$.

Furthermore, as we show in an example below, the term $\lambda^i/\varphi^i + \nu^i(1+\eta)$ is increasing in the household's human wealth: households with high human wealth have a larger weight, $\lambda^i/\varphi^i + \nu^i(1+\eta)$ at the Ramsey optimum than households with lower labor market productivity.[12] As the

[12] While it is true that high-type households have high market weights, φ^i, at the Ramsey optimum, their multipliers ν^i, are also high and dominate the overall direction of this term.

labor productivity $\theta^i(s_t)$ of the high-type households increases relative to that of the low types, the numerator of $\mathcal{I}(s_t)$ grows relative to its denominator. As a result, $\mathcal{I}(s_t)$ is high in states in which high human-wealth households are relatively more productive than low human-wealth thouseholds, and $\mathcal{I}(s_t)$ is low in states where the converse is true.

Theorem 2 states that the optimal monetary tax varies with the state and depends on the level of labor income inequality, as proxied by $\mathcal{I}(s_t)$. There exists a threshold $\bar{\mathcal{I}}(s^{t-1})$ such that when labor income inequality is strictly greater than this threshold, the implicit monetary tax is positive. On the other hand, when labor income inequality is below this threshold, the implicit monetary tax is negative (i.e. a subsidy). When $\mathcal{I}(s_t)$ is exactly equal to the threshold, the optimal monetary tax is equal to zero.

Intuition. To understand this result, recall the intuition behind Theorem 1. When preferences are homothetic and the labor skill distribution exhibits only proportional aggregate shocks, the tax system is sufficient to achieve the optimal level of redistribution. In this case, fiscal policy optimally trades off the benefit of redistribution with its cost—that is, the efficiency cost from distorting aggregate consumption and aggregate labor—and this trade-off does not vary over time. Note that Theorem 2 nests this as a special case: when the labor skill distribution satisfies (34), the function $\mathcal{I}(s_t)$ reduces to a constant equal to $\bar{\mathcal{I}}$ in all states, and the optimal monetary tax is zero.

Starting from $\mathcal{I}(s_t) = \bar{\mathcal{I}}(s^{t-1})$, now consider a small shock to the *relative* skill distribution that raises $\mathcal{I}(s_t)$ above $\bar{\mathcal{I}}(s^{t-1})$. The marginal benefit to redistribution increases in response to this shock while the marginal cost remains the same. It follows that if state-contingent tax rates were available, the optimal tax rate would increase. A higher tax rate implies that high-skilled, rich workers pay more taxes than low-skilled, poor workers, but everyone receives the same lump-sum transfer; hence an increase in the optimal tax rate provides greater redistribution.

When state-contingent taxes are unavailable, however, it becomes optimal for monetary policy to step in and fill this role. Thus, when $\mathcal{I}(s_t)$ rises above $\bar{\mathcal{I}}(s^{t-1})$, it is optimal for monetary policy to abandon implementing flexible-price allocations and instead *mimic* an increase in the tax rate. Conversely, when $\mathcal{I}(s_t)$ falls below $\bar{\mathcal{I}}(s^{t-1})$, it is optimal for monetary policy to mimic a fall in the labor income tax, that is, to act as a labor income subsidy: $\tau_M^*(s^t) < 0$.

There is a clear distinction between using monetary and fiscal policy. Both monetary and fiscal policy can be used to drive a wedge between the MRS and the MRT of aggregate consumption and aggregate labor. However, unlike state-contingent fiscal policy, state-contingent monetary policy leads to an additional type of distortion: a wedge between the prices of sticky-price firms and those of flexible-price firms. Equilibrium price dispersion results in misallocation and, ultimately, a loss in production efficiency.

For this reason, monetary policy should be considered an imperfect substitute for missing tax instruments. Were state-contingent tax instruments readily available, one could use these instruments to implement a relaxed Ramsey optimum x^{R*} without any corresponding loss in production efficiency. However, when such fiscal state-contingency is ruled out, as we have assumed a priori, the next best tool is monetary policy. In this case, the best possible allocation is a Ramsey optimum x^* which necessarily features misallocation across sticky- and flexible-price firms whenever $\mathcal{I}(s_t) \neq \bar{\mathcal{I}}(s^{t-1})$.

Note that this additional efficiency cost of using monetary policy does not negate the intuition provide above. To see this, start again from $\mathcal{I}(s_t) = \bar{\mathcal{I}}(s^{t-1})$. Here, the fiscal policy is set so that the marginal benefit of redistribution is equal to the marginal cost of distorting the intratemporal margin, and monetary policy implements the flexible price allocation ($\tau_M^*(s^t) = 0$).

Now consider a small deviation of $\mathcal{I}(s_t)$ above $\bar{\mathcal{I}}(s^{t-1})$. The marginal benefit to redistribution increases, the marginal cost of distorting the intratemporal margin stays the same, which implies that monetary policy should abandon the flexible-price benchmark. But the marginal cost in production efficiency due to such abandonment is, to a first-order, zero. This is because at the flexible-price allocation, production efficiency is maximized. Therefore, any loss in production efficiency due to misallocation of intermediate goods around this benchmark must be of second-order. It follows that for small deviations of $\mathcal{I}(s_t)$ above $\bar{\mathcal{I}}(s^{t-1})$, the optimal monetary tax must be strictly positive: $\tau_M^*(s^t) > 0$.

The intuition provided above holds for small shocks around $\bar{\mathcal{I}}(s^{t-1})$, but what about for large shocks? In fact, Theorem 2 has makes no provision that $\mathcal{I}(s_t)$ be close to $\bar{\mathcal{I}}(s^{t-1})$. Yet, the intuition does not change when $\mathcal{I}(s_t)$ is far from $\bar{\mathcal{I}}(s^{t-1})$. As monetary policy moves further and further away from implementing flexible-price allocations, losses in production efficiency become larger. However, these losses can never be strong enough to reverse the sign of monetary policy—such an occurrence would lead to a contradiction. This is because the only reason monetary policy abandons the flexible-price benchmark in the first place is the planner's distributional motive. Even if the loss in production efficiency may dampen the extent to which monetary policy mimics a missing tax instrument as $\mathcal{I}(s_t)$ moves further and further away from $\bar{\mathcal{I}}$, it can never force monetary policy to reverse its sign: were that the case, monetary policy could always do better by reverting to implementation of flexible-price allocations, therefore contradicting the Ramsey optimality of the proposed allocation.

A Numerical Illustration. We illustrate the mechanism underlying Theorem 2 with a simple numerical example with two household types, $i \in \{H, L\}$, of equal sizes ($\pi^H = \pi^L = 1/2$). We consider a labor skill distribution that features non-proportional shocks: in particular, we let the ratio of θ^H/θ^L fluctuate across $N = 10$ possible states. We assume these states are i.i.d. and uniformly distributed so that $\mu(s'|s) = 1/N$ for all $s, s' \in S$. Finally, we set $\eta = .5$, $\gamma = 2$, $\beta = .98$, $\kappa = .5$ and $\rho = 2$.

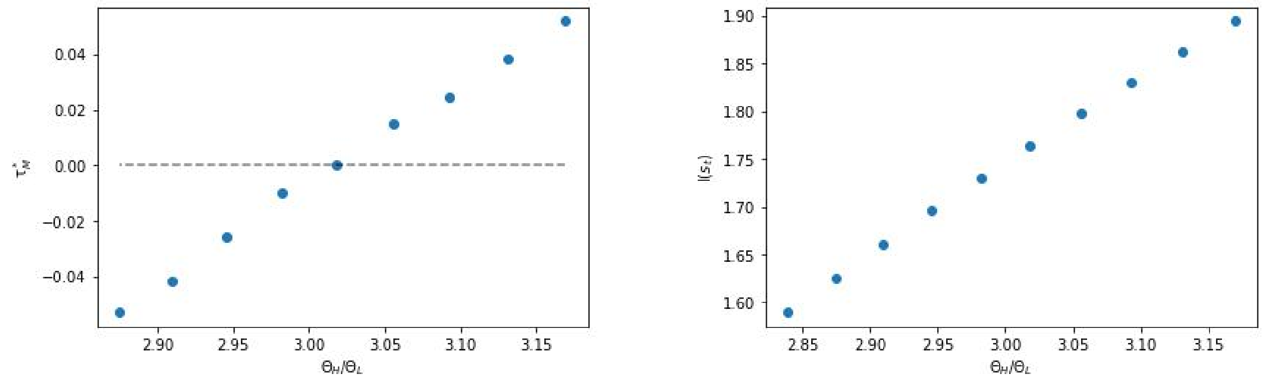

Figure 1. Left panel: $\tau_M^*(s^t)$ as a function of $\theta^H(s_t)/\theta^L(s_t)$. Right panel: $\mathcal{I}(s_t)$ as a function of $\theta^H(s_t)/\theta^L(s_t)$.

We numerically solve for optimal fiscal and monetary policy with equal Pareto weights: $\lambda^H = \lambda^L = 1$. The left panel of Figure 1 plots the implied optimal monetary tax for different values of $\theta^H(s_t)/\theta^L(s_t)$. As this ratio increases, the optimal monetary tax increases. In this simple example, the weight $\lambda^i/\varphi^i + \nu^i(1+\eta)$ of the high-productivity type is greater than that of the low-productivity type. As a result, our sufficient statistic for inequality, $\mathcal{I}(s_t)$, is increasing in the ratio $\theta^H(s_t)/\theta^L(s_t)$, as illustrated in the right panel of Figure 1.

5.1 Partial State-Contingency of Taxes

In our previous analysis we made the stark assumption that monetary policy is state-contingent while fiscal policy is not—this is the typical assumption found throughout the New Keynesian literature. We now partially relax this restriction on fiscal tools and allow tax rates to be set one period in advance; specifically, we let τ_c, τ_ℓ, and τ_r at time t be contingent on s^{t-1}. While this does not go all the way to full state-contingency of tax rates, it provides the fiscal authority with some ability to respond to shocks—it can respond with a one-period lag. To the extent that shocks are persistent, this allows fiscal policy to absorb some of the state-contingency pressure placed on monetary policy in the previous analysis.

In this case we can more cleanly characterize the behavior of $\tau_M^*(s^t)$ around zero.

Theorem 3. *Let tax rates be set one period in advance. There exists a threshold $\bar{\mathcal{I}}(s^{t-1}) > 0$ such that $\tau_M^*(s^t) = 0$ if and only if $\mathcal{I}(s_t) = \bar{\mathcal{I}}(s^{t-1})$, and $\tau_M^*(s^t) > 0$ if and only if $\mathcal{I}(s_t) > \bar{\mathcal{I}}(s^{t-1})$. To a first-order Taylor approximation around $\mathcal{I}(s_t) = \bar{\mathcal{I}}(s^{t-1})$, the optimal monetary tax satisfies:*

$$\tau_M^*(s^t) \approx \delta_0[\mathcal{I}(s_t)/\bar{\mathcal{I}}(s^{t-1}) - 1] \tag{40}$$

where

$$\delta_0 = \frac{1}{1+\rho(\eta+\gamma)\frac{1-\kappa}{\kappa}} \in (0,1). \tag{41}$$

Proof. See Appendix B.3. □

When we allow tax rates to be set one-period in advance, our main result on the optimal conduct of monetary policy remains intact. However, with this greater level of fiscal flexibility, we obtain a sharper characterization of the optimal monetary tax near the benchmark of zero, i.e. when $\mathcal{I}(s_t) = \bar{\mathcal{I}}(s^{t-1})$. In particular, we show that to a first-order Taylor approximation around $\mathcal{I}(s_t)/\bar{\mathcal{I}}(s^{t-1}) = 1$, the optimal monetary tax is strictly increasing in $\mathcal{I}(s_t)/\bar{\mathcal{I}}(s^{t-1})$, with a slope of $\delta_0 \in (0, 1)$.

The slope δ_0 characterizes the extent to which the optimal monetary tax responds to an increase in $\mathcal{I}(s_t)$: a larger value for δ_0 indicates a more aggressive response, whereas a lower value indicates a less agressive response. An explicit, closed-form expression of this derivative is given in (41); in particular, δ_0 is strictly positive, strictly less than 1, and a function of the primitives ρ, η, γ, and κ.

First, note that δ_0 is decreasing in ρ, the elasticity of substitution across intermediate goods. Deviations of monetary policy away from flexible prices results in intermediate good price dispersion: sticky-price firms charge prices that differ from those of flexible-price firms. In response, the final good firm substitutes away from high-priced intermediate goods towards low-priced intermediates, resulting in the misallocation of inputs and a loss in overall production efficiency. The greater the substitutability across goods, the greater the loss in production efficiency. It follows that when ρ is high, monetary policy responds less aggressively to movements in $\mathcal{I}(s_t)$.

Next, δ_0 is increasing in $\kappa/(1-\kappa)$, the ratio of sticky price firms relative to flexible price firms in the economy. Consider the limit in which $\kappa \to 1$. In this case $\delta_0 \to 1$. In the limit in which all firms in the economy are sticky, movements in monetary policy away from flexible-rice allocations result in near zero losses in production effiiency; monetary policy therefore approximates a true labor income tax as it only distorts the overall labor wedge. In this limit monetary policy thereby mimics the optimal state-contingent tax rate, which responds one-for-one with changes in $\mathcal{I}(s_t)$. Consider instead the case of $\kappa \to 0$. In the limit in which all firms in the economy are flexible, monetary policy has no power and $\delta_0 \to 0$.

Finally, δ_0 is decreasing in η and γ. These parameters govern the curvature of the households' utility function in consumption and labor, and hence the extent to which marginal utility changes when distortionary taxation increases. When η and γ are high, households are more inelastic, in which case the marginal "cost" of distorting the households' intratemporal margin is larger. It follows that monetary policy should respond less aggressively to movements in $\mathcal{I}(s_t)$ away from $\bar{\mathcal{I}}(s^{t-1})$.

6 Implementation

In this section we discuss how the Ramsey optimum can be implemented with the available set of fiscal and monetary tools.

Fiscal policy. The optimal fiscal wedge is χ^*. Clearly there is no unique implementation of this wedge, and any implementation of χ^* results in the same behavior for optimal monetary policy. For the sake of exposition, in this section we set the sales tax (subsidy) such that it directly neutralizes the monopolistic markup and let the labor income and consumption tax rates jointly implement the optimal fiscal wedge. Specifically:

$$1-\tau_r = \frac{\rho}{\rho-1}, \qquad \text{and} \qquad \frac{1-\tau_\ell}{1+\tau_c} = \chi^*. \tag{42}$$

Monetary Target. We define the aggregate markup $\mathcal{M}(s^t)$ in the economy to be the aggregate price level over the firms' nominal marginal cost; in logs:

$$\log \mathcal{M}(s^t) \equiv \log P(s^t) - \log(W(s^t)/A(s^t)). \tag{43}$$

Note that if we shut down aggregate productivity shocks, i.e. $A(s^t) = 1$ for all s^t,then the aggregate markup is equal to the inverse of the real wage, $W(s^t)/P(s^t)$. We express optimal monetary policy in terms of the aggregate markup as follows.

Proposition 5. *With tax rates set according to* (42)*, the optimal markup satisfies:*

$$\begin{array}{lll} \log \mathcal{M}(s^t) > 0 & \textit{if and only if} & \mathcal{I}(s_t) > \bar{\mathcal{I}}(s^{t-1}), \\ \log \mathcal{M}(s^t) = 0 & \textit{if and only if} & \mathcal{I}(s_t) = \bar{\mathcal{I}}(s^{t-1}), \\ \log \mathcal{M}(s^t) < 0 & \textit{if and only if} & \mathcal{I}(s_t) < \bar{\mathcal{I}}(s^{t-1}). \end{array}$$

Proof. See Appendix A.12. □

Proposition 5 is essentially a restatement of Theorem 3 in terms of the aggregate markup instead of the monetary wedge. In fact, the two are nearly equivalent: if households purchase the final good at a higher price than the marginal cost to produce it, it is as if they are paying a tax. Conversely, if households purchase the final good at a lower price than its marginal cost, it is as if they are receiving a subsidy.

When the labor income distribution exhibits only proportional aggregate shocks, it is optimal for monetary policy to implement flexible-price allocations. This possibility is nested in Proposition 5 as the case in which $\mathcal{I}(s_t) = \bar{\mathcal{I}}(s^{t-1})$ in all states, in which case optimal monetary policy targets a constant mark-up: $\log \mathcal{M}(s^t) = 0$.[13] Away from this special case, when there are

[13] Here, the level of zero for the log markup under flexible-prices is arbitrary: it is only equal to zero because we have set the sales tax to exactly cancel out the monopolistic markup (42). Had we not made that choice, the markup under flexible prices would be a non-zero constant. More generally, under flexible prices: $\mathcal{M}(s^t) = \left[(1-\tau_r)\left(\frac{\rho-1}{\rho}\right)\right]^{-1}$ for all $s^t \in S^t$.

movements in the relative labor skill distribution, the tax system is insufficient to implement the optimal level of redistribution, and it is then optimal for monetary policy to deviate from implementing flexible price allocations and target a state-contingent markup. As with the monetary tax, we find that when $\mathcal{I}(s_t)$ rises above the threshold, the optimal mark-up is positive, meaning that the aggregate price level should rise above the nominal marginal cost of production. Conversely when $\mathcal{I}(s_t)$ falls below the threshold, the optimal mark-up is negative, meaning that the aggregate price level should fall below the nominal marginal cost of production.

As noted, variation in the mark-up (and conversely the real wage) works much like variation in the tax rate. Although all households face the same marginal tax rate, they face different average tax rates: given a positive marginal tax, high-skilled households pay more taxes (in levels) than low-skilled households. The same holds true with the mark-up. Note, however, that the manner by which the proceeds of the markup are collected and distributed to the households differs from how tax revenue is collected and distributed. With standard fiscal instruments, tax revenue is collected by the government and redistributed to households uniformly via the lump-sum transfer. In contrast, when prices rise above marginal costs, firms make positive profits. Profits, in turn, are either taxed directly and distributed equally across households as a transfer, or if not taxed, distributed as dividends across shareholders. It follows that a higher markup, like a higher marginal tax, is more redistributive.

In the baseline model, firm profits are isomorphic to lump-sum transfers.[14] This equivalence relies clearly on the assumption that equity shares are uniform across households. Admittedly, this is an unrealistic assumption—it allows for a clean benchmark for our analysis. In the following section, Section 7, we relax this assumption and analyze an extension of our model in which households own unequal shares of the firms.

The Price Level and Interest Rates. Finally, we consider the behavior of aggregate price levels and nominal interest rates consistent with the Ramsey optimum. Let $\mathcal{B}_t(s^{t-1}) > 0$ denote the common belief of the aggregate price level at time t based on history s^{t-1}. Aside from being strictly positive, $\mathcal{B}_t(s^{t-1}) > 0$ is a free parameter in our model. For a given $\mathcal{B}_t(s^{t-1})$, when $P(s^t) = \mathcal{B}_t(s^{t-1})$, the economy replicates the flexible price outcome. Let $\hat{\imath}(s^t)$ denote the nominal interest rate consistent with the flexible-price outcome; one can think of $\hat{\imath}(s^t)$ as the "natural" rate of interest.

Proposition 6. *Given a common belief $\mathcal{B}_t(s^{t-1}) > 0$, the aggregate price level, $P(s^t)$, at the Ramsey optimum and the nominal interest rate, $i(s^t)$, consistent with that price level satisfies:*

[14] This can be seen directly in the budget implementability conditions (18), in particular our definition of $\bar{T}$ in (19).

$$
\begin{array}{llll}
P(s^t) < \mathcal{B}_t(s^{t-1}) & \textit{and} \quad i(s^t) > \hat{\imath}(s^t) & \textit{if and only if} & \mathcal{I}(s_t) > \bar{\mathcal{I}}(s^{t-1}), \\
P(s^t) = \mathcal{B}_t(s^{t-1}) & \textit{and} \quad i(s^t) = \hat{\imath}(s^t) & \textit{if and only if} & \mathcal{I}(s_t) = \bar{\mathcal{I}}(s^{t-1}), \\
P(s^t) > \mathcal{B}_t(s^{t-1}) & \textit{and} \quad i(s^t) < \hat{\imath}(s^t) & \textit{if and only if} & \mathcal{I}(s_t) < \bar{\mathcal{I}}(s^{t-1}).
\end{array}
$$

Proof. See Appendix A.13. □

The behavior of the aggregate price level and nominal interest rate described in Proposition 6 is consistent with the optimal markup described in Proposition 5. To understand this, recall that prices are "sticky" in our model, while nominal wages are fully flexible. Specifically, given a common belief $\mathcal{B}_t(s^{t-1})$, the prices of the sticky-price firms are "stuck" at $p^s_t(s^{t-1}) = \mathcal{B}_t(s^{t-1})$, while the prices of the flexible-price firms and the nominal wage respond to the realized state.

Therefore, in order to generate an increase in the aggregate markup, the nominal wage must (unexpectedly) fall. In this case, the prices of the flexible-price firms fall, in line with the realized nominal wage, while those of the sticky price firms remain stuck. As a result, the aggregate price level falls, but not to the same extent as the nominal wage, and as a result the mark-up increases. Conversely, in order to generate a fall in the aggregate markup, the nominal wage must (unexpectedly) rise. In this case, the aggregate price level also rises, but less so than wages, so that the aggregate markup indeed falls.

Furthermore, consistent with the price movements described in Proposition 6 and the optimal markup described in Proposition 6 is a nominal interest rate that satisfies the households' Euler equation in (15). An increase in the aggregate markup in our model is consistent with a tightening of the nominal interest rate relative to the natural rate; conversely a fall in the markup is consistent with a loosening of the nominal interest rate.

Robustness. Finally, it is important to note that the price level and interest rate movement consistent with the Ramsey optimum depend on the relative stickiness of prices versus wages. More specifically, our characterization of the aggregate price level (and interest rate) in Proposition 6 relies on our assumption of sticky prices but flexible wages. If instead wages were sticky and prices were flexible, an increase in the aggregate markup would require an unexpected *increase* in the aggregate price level rather than a fall.

For this reason, in terms of monetary implementation we wish to put less emphasis on Proposition 6 and more emphasis on Proposition 5—what is robust to the relative stickiness of prices versus wages is the movement in the optimal markup. Furthermore, although monetary policy typically centers on changes in nominal interest rates, the emphasis of the Ramsey literature on characterizing allocations proves highly useful in our analysis. In particular, the economics behind (and robustness of) optimal monetary policy in our environment can best be understood by studying the state-contingent wedge at the Ramsey optimum (Section 5).

7 Heterogeneous Equity Shares

In this section we relax our assumption that households own equal shares of the firms. We let $1+\sigma^i$ denote the share of equity held by a household of type $i \in I$, with $\sum_{i \in I} \pi^i \sigma^i = 0$, and assume that equity shares are fixed. The budget constraint of household i is given by:

$$\begin{aligned}&(1+\tau_c)P(s^t)c^i(s^t) + b^i(s^t) - (1+i(s^{t-1}))b^i(s^{t-1}) + \sum_{s^{t+1}|s^t} Q(s^{t+1}|s^t)z^i(s^{t+1}|s^t) \\ &\leq (1-\tau_\ell)W(s^t)\ell^i(s^t) + (1-\tau_\Pi)(1+\sigma^i)\Pi(s^t) + P(s^t)T(s^t) + z^i(s^t|s^{t-1})\end{aligned} \tag{44}$$

For now, we impose no restrictions on the cross-sectional covariance between household labor skill type and equity share, but we discuss the implications of this covariance for optimal policy. In particular we are interested in the case in which shares σ^i covary positively with human wealth (lifetime labor earnings).

7.1 Equilibrium Characterization

We begin by characterizing the set of equilibrium allocations in this economy. The households' intratemporal and intertemporal optimality conditions, (13)-(15), and the firms' optimal pricing equations, (21)-(22), are unaltered relative to the baseline model. What must be adjusted are the implementability conditions corresponding to the households' budget constraints.

The lifetime budget constraint of type $i \in I$ is now equivalent to the following condition:

$$\begin{aligned}&\sum_t \sum_{s^t} \beta^t \mu(s^t) \left[U_C^m(s^t)\omega_C^i(\varphi)C(s^t) + U_L^m(s^t)\omega_L^i(\varphi, s_t)L(s^t)\right] \\ &= U_C^m(s_0)\bar{T} + \sigma^i(1-\tau_\Pi)\sum_t \sum_{s^t} \beta^t \mu(s^t)\frac{U_C^m(s^t)}{1+\tau_c}\frac{\Pi(s^t)}{P(s^t)}\end{aligned} \tag{45}$$

where $\bar{T}$ is as in (19) and $\Pi(s^t)/P(s^t)$ are real profits in history s^t.[15] The only difference between the above condition and the corresponding condition in our baseline model is the final term on the right-hand side of the equation. This term, the present discounted value of lifetime after-tax real profits, emerges as a direct result of the heterogeneity in equity shares; note that it disappears when either there is no heterogeneity as in our baseline model ($\sigma^i = 0$ for all $i \in I$), or when profits are taxed fully ($\tau_\Pi = 1$).

Following the primal approach, we derive an expression for real profits, $\Pi(s^t)/P(s^t)$, in terms of allocations alone, and in so doing we characterize the set of sticky-price allocations as follows.

[15] We derive this in the usual way using the household's budget constraint; see Appendix C.1 for the specific derivation. Furthermore, we show in Appendix C.1 that aggregate profits satisfy:

$$\Pi(s^t) = (1-\tau_r)P(s^t)Y(s^t) - W(s^t)L(s^t).$$

Proposition 7. *A feasible allocation $x \in \mathcal{X}$ can be implemented as a sticky-price equilibrium if and only if there exist $\varphi \equiv (\varphi^i)$, $\bar{T} \in \mathbb{R}$, $\chi \in \mathbb{R}_+$, and a weakly positive scalar $\vartheta \in \mathbb{R}_{\geq 0}$, such that parts (i)-(ii) of Proposition 2 are satisfied, and*

$$\sum_t \sum_{s^t} \beta^t \mu(s^t) \left[U_C^m(s^t)\omega_C^i(\varphi)C(s^t) + U_L^m(s^t)\omega_L^i(\varphi, s_t)L(s^t)\right] \\ = U_c^i(s_0)\bar{T} + \sigma^i \vartheta \sum_t \sum_{s^t} \beta^t \mu(s^t) \left[\chi \frac{\rho}{\rho - 1} U_C^m(s^t)C(s^t) + U_L^m(s^t)L(s^t)\right] \tag{46}$$

holds for all $i \in I$.

Proof. See Appendix C.2. □

Proposition 7 is the heterogeneous share economy analog of Proposition 2 in our baseline economy. Parts (i) and (ii) of Proposition 2 remain intact. The only difference between the two characterizations is that we replace the baseline economy's budget set implementability conditions (18) with the corresponding implementability conditions in the heterogeneous share economy, (46).

The final term in (46) is the one that corresponds to the household's lifetime value of after-tax real profits. Note that this term includes a weakly-positive scalar denoted by ϑ. This scalar is given by $\vartheta \equiv (1 - \tau_\Pi)/(1 - \tau_\ell)$; it therefore parameterizes an additional lever by which fiscal policy can influence real allocations in the heterogeneous share economy.

Note that if profits are fully taxed, $\tau_\Pi = 1$, then (46) is identical to (18). In this case it is irrelevant as to whether equity shares are homogeneous or heterogeneous: the implementability conditions are the same across the two economies. Furthermore, to the extent that equity shares covary positively with human wealth, and the Ramsey planner wishes to redistribute wealth from high labor income households to low labor income households, it is in fact optimal to fully tax profits.

Therefore, in order to make the analysis interesting, we make the following ad hoc assumption on τ_Π and τ_ℓ.

Assumption 1. *Let $\bar{\vartheta} > 0$ be a strictly positive scalar. The tax rates τ_Π and τ_ℓ are such that $\vartheta = \bar{\vartheta}$.*

We assume that there exists an ad hoc constraint on the fiscal authority such that it cannot fully tax profits nor drive τ_ℓ to negative infinity. For the remainder of our analysis, we impose Assumption 1 and set $\vartheta = \bar{\vartheta}$. We let $\mathcal{X}_\sigma^s$ denote the set of sticky-price allocations: those that satisfy Proposition 7.

7.2 The Ramsey Problem and Optimal Monetary Policy

We turn now to the Ramsey problem in the heterogeneous shares economy. We adopt the utilitarian social welfare function in (30); the Ramsey planner chooses an allocation x that maximizes social welfare subject to $x \in \mathcal{X}^s_\sigma$. We solve this problem in Appendix C.3 and characterise the Ramsey optimum. Here, we present our main result on optimal monetary policy.

Theorem 4. *There exists a threshold $\hat{\mathcal{I}}(s^{t-1}) > 0$, such that $\tau^*_M(s^t) = 0$ if and only if $\mathcal{I}(s_t) = \hat{\mathcal{I}}(s^{t-1})$, and $\tau^*_M(s^t) > 0$ if and only if $\mathcal{I}(s_t) > \hat{\mathcal{I}}(s^{t-1})$. The threshold $\hat{\mathcal{I}}(s^{t-1})$ satisfies:*

$$\hat{\mathcal{I}}(s^{t-1}) = \bar{\mathcal{I}}(s^{t-1}) + \bar{\vartheta}\left[(1+\eta) - (1-\gamma)\frac{\rho}{\rho-1}\right]\sum_{i\in I}\pi^i\nu^i\sigma^i. \tag{47}$$

where $\bar{\mathcal{I}}(s^{t-1})$ is the corresponding threshold in the homogeneous equity shares economy.

Proof. See Appendix C.4. □

When profit shares are heterogeneous, the behavior of the optimal monetary tax resembles that in the baseline economy with homogeneous shares. In particular, the optimal monetary tax depends on $\mathcal{I}(s_t)$, the sufficient statistic of labor income inequality. There exists a threshold $\hat{\mathcal{I}}(s^{t-1})$ such that when labor income inequality is strictly greater than this threshold, the implicit monetary tax is positive. On the other hand, when labor income inequality is below this threshold, the implicit monetary tax is negative (i.e. a subsidy). When $\mathcal{I}(s_t)$ is exactly equal to the threshold, the optimal monetary tax is equal to zero.

Therefore, our main *qualitative* result on optimal monetary policy is robust to heterogeneity in equity shares. What heterogeneity in shares alters, however, is the threshold at which the optimal monetary tax changes from a subsidy to a tax. To see this, note that the new threshold $\hat{\mathcal{I}}(s^{t-1})$ differs from $\bar{\mathcal{I}}(s^{t-1})$, the threshold in the homogeneous shares economy. In order to explore this difference, as a direct corollary to Theorem 4, we state the following.

Corollary 2. *Under Assumption 1, the threshold $\hat{\mathcal{I}}(s^{t-1})$ in 47 is strictly increasing in $\sum_{i\in I}\pi^i\nu^i\sigma^i$ if and only if*

$$\rho > \frac{1+\eta}{\gamma+\eta}. \tag{48}$$

Furthermore, if (ρ,γ,η) satisfy 48, then:

$$\hat{\mathcal{I}}(s^{t-1}) = \bar{\mathcal{I}}(s^{t-1}) \qquad \textit{if and only if} \qquad \sum_{i\in I}\pi^i\nu^i\sigma^i = 0$$

A sufficient condition for 48 to hold is $\gamma > 1$. This is a natural restriction on γ; hence for the remainder of this discussion we will take as given that (ρ,γ,η) satisfy 48.

In this case, the threshold $\hat{\mathcal{I}}(s^{t-1})$ is strictly increasing in $\sum_{i\in I}\pi^i\nu^i\sigma^i$. This term is the cross-sectional covariance between σ^i (profit shares) and ν^i, the Ramsey planner's multipliers on the implementability conditions in (46). Recall from our earlier discussion that households with high lifetime labor income have high values of ν^i, and households with low lifetime labor income have low values of ν^i. Therefore, a positive value for $\sum_{i\in I}\pi^i\nu^i\sigma^i$ indicates a positive covariance between equity shares and lifetime labor income—high labor productivity households own greater shares of the firm.

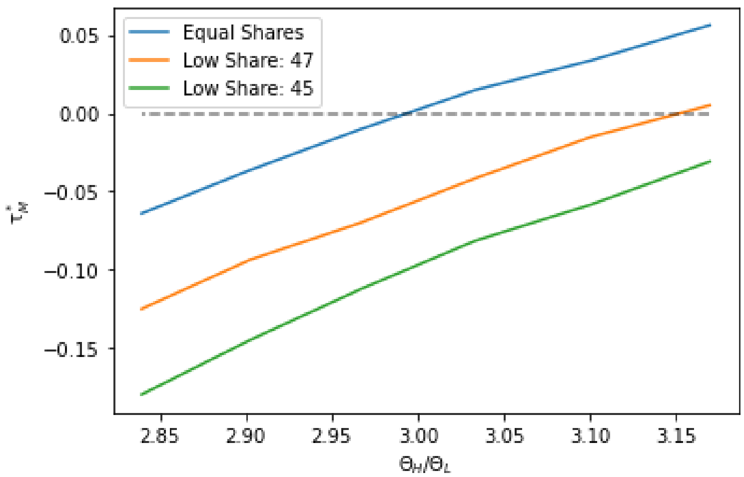

Figure 2. $\tau^*_M(s^t)$ as a function of $\theta^H(s_t)/\theta^L(s_t)$.

When this covariance is equal to zero, $\hat{\mathcal{I}}(s^{t-1}) = \bar{\mathcal{I}}(s^{t-1})$; the threshold at which $\tau^*_M(s^t) = 0$ is the same across the two economies. However, when this covariance is strictly positive, then these thresholds differ: $\hat{\mathcal{I}}(s^{t-1}) > \bar{\mathcal{I}}(s^{t-1})$. In Figure 2, the threshold in the heterogeneous share economy shifts to the right as the profit share of the low type falls (and the profit share of the high type increases). Conversely, if the covariance term is strictly negative, then $\hat{\mathcal{I}}(s^{t-1}) < \bar{\mathcal{I}}(s^{t-1})$.

This implies that when the covariance term is strictly positive, there exists a non-empty interval $[\bar{\mathcal{I}}(s^{t-1}), \hat{\mathcal{I}}(s^{t-1})]$ such that for any $\mathcal{I}(s_t)$ within this interval, the optimal monetary tax differs in sign across the the two economies: $\tau^*_M(s^t) < 0$ in the heterogeneous share economy and $\tau^*_M(s^t) > 0$ in the homogeneous share economy.

To understand this, note that when the "monetary tax" increases, firm profits increase. If profit shares co-vary possibly with lifetime income, then even if the monetary tax lowers labor income inequality in a particular state, the increase in profits exacerbates dividend payout inequality. In this case, it is optimal to begin taxing (and increasing profits) at a higher threshold of labor market inequality.

Despite this, the optimal monetary tax is positive for sufficiently large $\mathcal{I}(s_t)$. Intuitively, as profit shares are fixed, if labor market inequality becomes sufficiently severe, the value to the planner of decreasing labor market inequality outweights the cost of increasing dividend payout inequality.

Partial state contingency of taxes. Our analysis so far has focused on the comparative statics of the threshold with respect to heterogeneity in profit shares—in particular, how the threshold depends on $\sum_{i\in I} \pi^i \nu^i \sigma^i$. Another question is whether the slope of $\tau^*_M(s^t)$ with respect to $\mathcal{I}(s_t)$ changes with this covariance. While it is difficult to characterize the behavior of the monetary tax in general, we can provide a partial answer to this question in the particular version of our economy in which the fiscal authority can set tax rates one period in advance.

Theorem 5. *Let tax rates be set one period in advance. There exists a threshold $\hat{\mathcal{I}}(s^{t-1}) > 0$, such that $\tau_M^*(s^t) = 0$ if and only if $\mathcal{I}(s_t) = \hat{\mathcal{I}}(s^{t-1})$, and $\tau_M^*(s^t) > 0$ if and only if $\mathcal{I}(s_t) > \hat{\mathcal{I}}(s^{t-1})$. The threshold $\hat{\mathcal{I}}(s^{t-1})$ satisfies (47) where, in this case, $\bar{\mathcal{I}}(s^{t-1})$ is the optimal threshold in the homogeneous shares economy with tax rates set one period in advance.*

To a first-order Taylor approximation around $\mathcal{I}(s_t) = \hat{\mathcal{I}}(s^{t-1})$, the optimal monetary tax satisfies:

$$\tau_M^*(s^t) \approx \delta_0 \frac{1}{\mathcal{H}(s^{t-1}) + \bar{\vartheta}(\gamma - 1)\frac{\rho}{\rho-1}\sum_{i\in I}\pi^i\nu^i\sigma^i}[\mathcal{I}(s_t) - \hat{\mathcal{I}}(s^{t-1})] \tag{49}$$

where $\delta_0 \in (0,1)$ is given by (41), and

$$\mathcal{H}(s^{t-1}) = \chi^*(s^{t-1})^{-1}\sum_{i\in I}\pi^i\omega_C^i(\varphi)\left[\frac{\lambda^i}{\varphi^i} + \nu^i(1-\gamma)\right]. \tag{50}$$

Proof. See Appendix C.6. □

We again find that when profit shares are heterogeneous, the behavior of the optimal monetary tax resembles that in the baseline economy with homogeneous shares. Again taking as given that $\gamma > 1$, the threshold $\hat{\mathcal{I}}(s^{t-1})$ is strictly increasing in $\sum_{i\in I}\pi^i\nu^i\sigma^i$.

Theorem 5 furthermore characterizes the slope of $\tau_M^*(s^t)$ with respect to $\mathcal{I}(s_t)$ around $\tau_M^*(s^t) = 0$. It is helpful to compare this slope to that in the baseline homogeneous share economy (with tax rates set one period in advance. In the baseline economy, to a first order around $\tau_M^*(s^t) = 0$, the optimal monetary tax satisfies:

$$\tau_M^*(s^t) \approx \delta_0 \frac{1}{\mathcal{H}(s^{t-1})}[\mathcal{I}(s_t) - \bar{\mathcal{I}}(s^{t-1})]. \tag{51}$$

This corresponds to equation (40) in Theorem 3, noting that in the baseline economy, $\bar{\mathcal{I}}(s^{t-1}) = \mathcal{H}(s^{t-1})$.

We can compare the slope in (49) to that in (51). First, when the covariance term $\sum_{i\in I}\pi^i\nu^i\sigma^i$ is equal to zero, the slopes across the two economies coincide. Furthermore, taking for granted that $\gamma > 1$, the slope in (49) is strictly decreasing in the covariance. Therefore, when $\sum_{i\in I}\pi^i\nu^i\sigma^i$ is strictly positive, the slope in the heterogeneous share economy is strictly lower than the slope in the homogeneous share economy. This seems consistent with the intuition that in an economy in which high labor income households own greater shares of the firm, an increase in the monetary tax lowers labor income inequality but at the same leads to an increase in dividend payout inequality. As a result, optimal monetary policy is less aggresive in response to changes in labor earnings inequality.

To conclude, the presence of heterogeneous equity shares changes both the slope and the intercept of the response of optimal monetary policy to labor income inequality. However, it does not alter the general lesson that, near zero, the optimal markup covaries positively with a sufficient statistic for labor income inequality.

8 Calibration

In this final section we consider a simple, calibrated version of the baseline model. In this exercise, we map changes in the labor productivity distribution over the business cycle to reflect empirical estimates of the incidence of GDP fluctuations on labor income. In this calibration, we assume 5 household types and 4 economic states. The first 4 equally-sized types capture the bottom 90 percent of the income distribution, while the last type captures the top decile of the labor income distribution. We use estimates of worker betas—the percent change in the growth rate of labor income associated with a percent change in GDP growth—from Guvenen, Schulhofer-Wohl, Song and Yogo (2017) to construct the functions $\theta^i : S \to \mathbb{R}_+$.

We use Treasury Department estimates of the labor income distribution in 2019 to construct the long-run labor productivity distribution. We furthermore use real GDP data from the Bureau of Economic Analysis to estimate both the long run growth rate (3 percent) along with growth rates and unconditional probabilities for 4 distinct economic states: a severe recession (-2.2 percent growth), a mild recession with (-.4 percent growth), normal growth (2.5 percent), and fast growth (5.4 percent). Assuming that the average labor income growth rate is also 3 percent, a worker's labor productivity in a particular state can be calculated by using their worker beta to translate changes in the GDP in that state into changes in labor productivity. In line with that data, we interpret a period as a year and set the households' discount factor $\beta = .98$. We set the elasticity of substitution between goods $\rho = 6$, the inverse Frisch elasticity, $\eta = 1$, the inverse elasticity of intertemporal substitution $\gamma = 2$, and assume that 20 percent of firms are sticky price firms.

We solve numerically for the Ramsey optimum with equal Pareto weights: $\lambda_i = 1$ for all i. In terms of the optimum's fiscal implementation, we follow the one suggested in (42), so that the mark-up under flexible prices is zero.

We start by first excluding the top decile of the income distribution. In this case, worker betas are monotonically decreasing in income, with low productivity workers being the most exposed to fluctuations in GDP.[16] Optimal monetary policy, expressed in terms of the log markup, is reported in the left panel of Figure 3. We find that the relationship between the optimal markup and output is monotonically decreasing in $A(s_t)$. As $A(s_t)$ falls and the economy enters a recession, the productivity of the lowest type households falls disproportionately relative to the high type households. As a result, labor income inequality increases in the recession, which leads to an increase in the optimal monetary tax (or mark-up).

We then solve for optimal policy including the top income decile. Importantly, the worker betas are no longer monotonic in this case. In particular, the labor income of the top income decile is more exposed to output fluctuations than the previous 2 income groups. At the same

[16]See Guvenen et al. (2017) Figure 1.

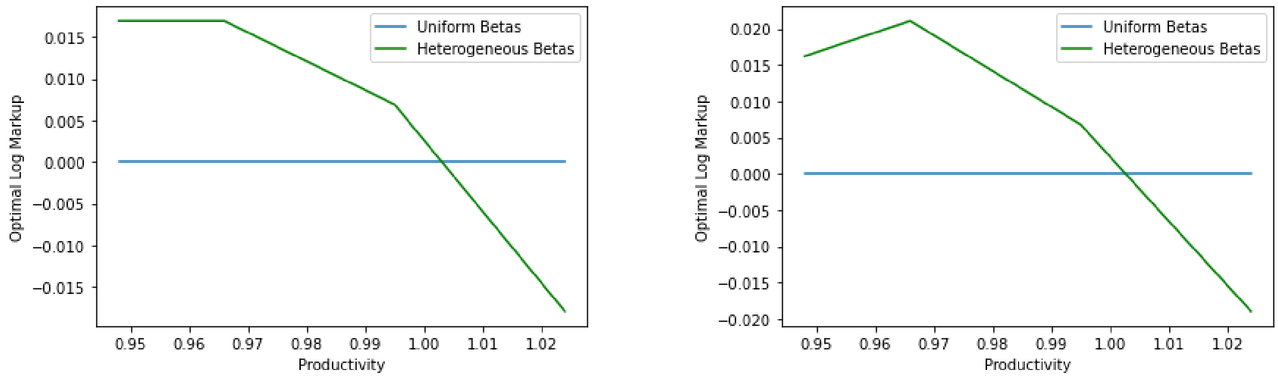

Figure 3. $\mathcal{M}^*(s^t)$ as a function of $A(s_t)$. Left panel: Monotonic worker betas. Right panel: Non-monotonic worker betas.

time, the top decile has a much higher average labor productivity. Optimal monetary policy in this case is reported in the right panel of Figure 3.

We find that optimal policy is no longer strictly monotonic in productivity. Because the productivity of top earners is more highly exposed to GDP fluctations, the productivity distribution in severe recessions is more compressed than in mild recessions. In this case, the redistributive power of the monetary tax is lessened, and the optimal markup is slightly smaller.

9 Conclusion

In this paper we study optimal monetary policy in a dynamic, general equilibrium economy with heterogeneous agents. Markets are complete, but there is room for redistribution. We find that when preferences are iso-elastic and there are no shocks to the relative skill distribution, all redistribution is done via the tax system. In this case it is optimal for monetary policy to implement flexible-price allocations and it does so by targeting price stability. Furthermore, we find that it is optimal to implement flexible-price allocations even if taxes are set suboptimally.

On the other hand, when there are shocks to the relative skill distribution, the available tax instruments are insufficient for implementing the Ramsey optimum. In this case it is optimal for monetary policy to deviate from implementing flexible-price allocations and instead target a state-contingent markup. We find that the optimal markup co-varies positively with a sufficient statistic for labor income inequality. Finally, our results are robust to heterogeneity in profit shares.

There are many interesting channels that we have abstracted from in this paper. First, although we allow for differences in labor productivity, there is perfect reallocation of efficiency units of labor across firms in our economy. Furthermore, labor productivity is modeled as a type-specific, stochastic process that is contingent on the aggregate state, but this process is exogenous—we do not allow for endogeneity of labor productivity to the monetary policy. Finally, several papers document the distributional effects of monetary policy shocks; see e.g. Doepke and Schneider (2006); Coibion, Gorodnichenko, Kueng and Silvia (2017); Auclert (2019). We abstract from the heterogeneous effects of monetary shocks, and focus solely on how an inflation tax that affects all households uniformly can be useful for redistribution. We leave exploration of these channels open for future work.

References

Acharya, S., Challe, E., Dogra, K., 2020. Optimal monetary policy according to hank .

Alves, F., Kaplan, G., Moll, B., Violante, G.L., 2020. A further look at the propagation of monetary policy shocks in hank. Journal of Money, Credit and Banking 52, 521–559.

Angeletos, G.M., La'O, J., 2020. Optimal monetary policy with informational frictions. Journal of Political Economy 128, 1027–1064.

Auclert, A., 2019. Monetary policy and the redistribution channel. American Economic Review 109, 2333–67.

Auclert, A., Rognlie, M., Straub, L., 2018. The intertemporal keynesian cross. NBER Working Paper 25020 .

Bhandari, A., Evans, D., Golosov, M., Sargent, T.J., 2021. Inequality, business cycles, and monetary-fiscal policy. Econometrica 89, 2559–2599.

Chari, V., Christiano, L., Kehoe, P., 1994. Optimal fiscal policy in a business cycle model. Journal of Political Economy 102, 617–52.

Chari, V., Christiano, L.J., Kehoe, P.J., 1991. Optimal fiscal and monetary policy: Some recent results. Journal of Money, Credit and Banking 23, 519–539.

Chari, V.V., Kehoe, P.J., 1999. Optimal fiscal and monetary policy, in: Handbook of Macroeconomics. Elsevier. volume 1, pp. 1671–1745.

Coibion, O., Gorodnichenko, Y., Kueng, L., Silvia, J., 2017. Innocent bystanders? monetary policy and inequality. Journal of Monetary Economics 88, 70–89.

Correia, I., 2010. Consumption taxes and redistribution. American Economic Review 100, 1673–94.

Correia, I., Farhi, E., Nicolini, J.P., Teles, P., 2013. Unconventional fiscal policy at the zero bound. American Economic Review 103, 1172–1211.

Correia, I., Nicolini, J.P., Teles, P., 2008. Optimal fiscal and monetary policy: Equivalence results. Journal of Political Economy 116, 141–170.

Dávila, E., Schaab, A., 2022. Optimal monetary policy with heterogeneous agents: A timeless ramsey approach. Available at SSRN 4102028.

Diamond, P., Mirrlees, J., 1971. Optimal taxation and public production: I–production efficiency. American Economic Review 61, 8–27.

Doepke, M., Schneider, M., 2006. Inflation and the redistribution of nominal wealth. Journal of Political Economy 114, 1069–1097.

Galí, J., 2008. Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework. Princeton University Press.

Guvenen, F., Ozkan, S., Song, J., 2014. The nature of countercyclical income risk. Journal of Political Economy 122, 621–660.

Guvenen, F., Schulhofer-Wohl, S., Song, J., Yogo, M., 2017. Worker betas: Five facts about systematic earnings risk. American Economic Review 107, 398–403.

Guvenen, F., Smith, A.A., 2014. Inferring labor income risk and partial insurance from economic choices. Econometrica 82, 2085–2129.

Kaplan, G., Moll, B., Violante, G.L., 2018. Monetary policy according to hank. American Economic Review 108, 697–743.

Lucas, R.J., Stokey, N.L., 1983. Optimal fiscal and monetary policy in an economy without capital. Journal of Monetary Economics 12, 55–93.

Mackowiak, B., Wiederholt, M., 2009. Optimal sticky prices under rational inattention. American Economic Review 99, 769–803.

Mankiw, N.G., Reis, R., 2002. Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve. Quarterly Journal of Economics 117, 1295–1328.

McKay, A., Wolf, C., 2022. Optimal policy rules in hank. Technical Report. Working Paper, FRB Minneapolis.

Negishi, T., 1960. Welfare economics and existence of an equilibrium for a competitive economy. Metroeconomica 12, 92–97.

Parker, J.A., Vissing-Jorgensen, A., 2009. Who bears aggregate fluctuations and how? American Economic Review 99, 399–405.

Schulhofer-Wohl, S., 2011. Heterogeneity and tests of risk sharing. Journal of Political Economy 119, 925–958.

Werning, I., 2007. Optimal fiscal policy with redistribution. The Quarterly Journal of Economics 122, 925–967.

Woodford, M., 2003a. Imperfect common knowledge and the effects of monetary policy. Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps .

Woodford, M., 2003b. Interest and Prices: Foundations of a Theory of Monetary Policy. Princeton University Press.

A Appendix: Proofs for Baseline Economy

A.1 Household optimality

In this section of the appendix, we derive the optimality conditions for household i. We let $\beta^t \mu(s^t)\Lambda^i(s^t)$ denote the Lagrange multiplier on household i's budget set at time t, history s^t.

The first-order conditions for household i with respect to consumption and labor are given by, respectively:

$$\mu(s^t)U_c^i(s^t) - \mu(s^t)\Lambda^i(s^t)(1+\tau_c)P(s^t) = 0, \tag{52}$$

$$\mu(s^t)\frac{1}{\theta^i(s_t)}U_\ell^i(s^t) + \mu(s^t)\Lambda^i(s^t)(1-\tau_\ell)W(s^t) = 0, \tag{53}$$

where $U_c^i(s^t) \equiv \partial U(\cdot)/\partial c^i(s^t)$ and $U_\ell^i(s^t) \equiv \partial U(\cdot)/\partial h^i(s^t)$ denote the marginal utilities of the household of type i with respect to individual consumption and work effort. The first-order condition with respect to nominal bonds $b^i(s^t)$ is given by:

$$-\beta^t\mu(s^t)\Lambda^i(s^t) + \beta^{t+1}\sum_{s^{t+1}|s^t}\mu(s^{t+1})\Lambda^i(s^{t+1})(1+i(s^t)) = 0. \tag{54}$$

The first-order condition with respect to Arrow security $z^i(s^{t+1})$ is given by:

$$-\beta^t\mu(s^t)\Lambda^i(s^t)Q(s^{t+1}|s^t) + \beta^{t+1}\mu(s^{t+1})\Lambda^i(s^{t+1}) = 0. \tag{55}$$

The household's transversality conditions for nominal bonds and Arrow securities are given by:

$$\lim_{t\to\infty}\beta^t\mu(s^t)\Lambda^i(s^t)b^i(s^t) = 0 \qquad \text{and} \qquad \lim_{t\to\infty}\beta^t\mu(s^t)\Lambda^i(s^t)Q(s^{t+1}|s^t)z^i(s^{t+1}) = 0$$

Combining (52) and (53), we obtain the household's intratemporal condition:

$$-\frac{1}{\theta^i(s_t)}\frac{U_\ell^i(s^t)}{U_c^i(s^t)} = \frac{(1-\tau_\ell)W(s^t)}{(1+\tau_c)P(s^t)} \tag{56}$$

Using the fact that $U_c^i(s^t) = \Lambda(s^t)(1+\tau_c)P(s^t)$, we may rewrite the Euler equation for bonds as

$$\frac{U_c^i(s^t)}{P(s^t)} = \beta(1+i(s^t))\sum_{s^{t+1}|s^t}\mu(s^{t+1}|s^t)\frac{U_c^i(s^{t+1})}{P(s^{t+1})}, \tag{57}$$

Finally, the Arrow security price must satisfy

$$Q(s^{t+1}|s^t) = \beta\mu(s^{t+1}|s^t)\frac{U_c^i(s^{t+1})}{P(s^{t+1})}\frac{P(s^t)}{U_c^i(s^t)} \tag{58}$$

where $\mu(s^{t+1}|s^t) \equiv \mu(s^{t+1})/\mu(s^t)$ is the probability of s^{t+1} conditional on s^t.

A.2 Proof of Lemma 1

Condition (56) of the household's problem implies that in any equilibrium, the following condition must hold:

$$-\frac{1}{\theta^i(s_t)}\frac{U_\ell^i(s^t)}{U_c^i(s^t)} = \frac{(1-\tau_\ell)W(s^t)}{(1+\tau_c)P(s^t)} = -\frac{1}{\theta^k(s_t)}\frac{U_\ell^k(s^t)}{U_c^k(s^t)}$$

for all types $i, k \in I$. Consider now the static subproblem described in Lemma 1. Let $\rho_C(s^t)$ and $\rho_L(s^t)$ be the Lagrange multipliers on the constraints in (12). The first-order conditions of this subproblem are given by

$$\varphi^i U_c^i(s^t) - \rho_C(s^t) = 0, \qquad \forall i \in I \tag{59}$$

$$\varphi^i \frac{1}{\theta^i(s_t)} U_\ell^i(s^t) + \rho_L(s^t) = 0, \qquad \forall i \in I \tag{60}$$

Therefore

$$-\frac{1}{\theta^i(s_t)}\frac{U_\ell^i(s^t)}{U_c^i(s^t)} = \frac{\rho_L(s^t)}{\rho_C(s^t)} = -\frac{1}{\theta^k(s_t)}\frac{U_\ell^k(s^t)}{U_c^k(s^t)}$$

for all types $i, k \in I$. It follows that the solution to the sub-problem coincides with the equilibrium allocation. Finally, the envelope conditions for this static sub-problem are given by:

$$U_C^m(s^t) = \varphi^i U_c^i(\mathcal{C}^i(C(s^t), L(s^t); \varphi)),$$
$$U_L^m(s^t) = \varphi^i \frac{1}{\theta^i(s_t)} U_\ell^i(\mathcal{L}^i(C(s^t), L(s^t); \varphi)),$$

for all $i \in I$. Next, with the separable and isoelatic preferences assumed in (1), the FOCs in (59)-(60) can be written as

$$\varphi^i c^i(s^t)^{-\gamma} - \rho_C(s^t) = 0$$
$$\varphi^i \frac{1}{\theta^i(s_t)} \left[\frac{\ell^i(s^t)}{\theta^i(s_t)}\right]^\eta + \rho_L(s^t) = 0$$

Combining these conditions with the resource constraints in (12), we obtain the linear expressions in (16) for individual consumption and labor with shares given by (17).

A.3 Derivation of Budget Implementability Conditions

We derive condition (18). We take the household's budget constraint in (2) for type $i \in I$, multiply both sides by $\Lambda^i(s^t)$, and use the household's FOCs in (52) and (53) to substitute out consumption and labor prices. Doing so, we obtain:

$$U_c^i(s^t)c^i(s^t) + \frac{1}{\theta^i(s_t)}U_\ell^i(s^t)\ell^i(s^t) = \Lambda^i(s^t)z^i(s^t|s^{t-1}) - \Lambda^i(s^t)\sum_{s^{t+1}|s^t} Q(s^{t+1}|s^t)z^i(s^{t+1}|s^t) - \Lambda^i(s^t)b^i(s^t)$$
$$+ \Lambda^i(s^t)(1+i(s^{t-1}))b^i(s^{t-1}) + \Lambda^i(s^t)P(s^t)\bar{T}(s^t)$$

where we let

$$\bar{T}(s^t) \equiv T(s^t) + \frac{1}{P(s^t)}(1-\tau_\Pi)\Pi(s^t). \tag{61}$$

Multiplying both sides by $\beta^t\mu(s^t)$, summing over t and s^t, and using the household's intertemporal optimality conditions (57)-(56) to cancel terms, we obtain:

$$\sum_t\sum_{s^t}\beta^t\mu(s^t)\left[U_c^i(s^t)c^i(s^t)+\frac{1}{\theta^i(s_t)}U_\ell^i(s^t)\ell^i(s^t)\right]\le U_c^i(s_0)\bar{T},$$

where

$$\bar{T}\equiv\frac{1}{\Lambda^i(s_0)(1+\tau_c)P(s_0)}\sum_t\sum_{s^t}\beta^t\mu(s^t)\Lambda^i(s^t)P(s^t)\bar{T}(s^t).$$

Therefore

$$\bar{T}=\frac{1}{U_c^i(s_0)(1+\tau_c)}\sum_t\sum_{s^t}\beta^t\mu(s^t)U_c^i(s^t)\bar{T}(s^t)$$

for all $i\in I$. This corresponds with the definition in 19. Finally, using the solution and the envelope conditions for the static sub-problem described in Lemma 1, as well as the fact that individual allocations satisfy (16), we can rewrite the above conditions as:

$$\sum_t\sum_{s^t}\beta^t\mu(s^t)\left[U_C^m(s^t)\omega_C^i(\varphi)C(s^t)+U_L^m(s^t)\omega_L^i(\varphi,s_t)L(s^t)\right]\le U_C^m(s_0)\bar{T}$$

where

$$\bar{T}\equiv\frac{(1+\tau_c)^{-1}}{U_c^m(s_0)}\sum_t\sum_{s^t}\beta^t\mu(s^t)U_C^m(s^t)\left[T(s^t)+(1-\tau_\Pi)\frac{\Pi(s^t)}{P(s^t)}\right]$$

for all $i\in I$, as was to be shown.

A.4 Derivation of Sticky-Price Firm Optimality

The sticky-price firm solves the following problem:

$$\max_{p'}\sum_{s^t|s^{t-1}}Q(s^t|s^{t-1})\left\{(1-\tau_r)p'\left(\frac{p'}{P(s^t)}\right)^{-\rho}Y(s^t)-\frac{W(s^t)}{A(s_t)}\left(\frac{p'}{P(s^t)}\right)^{-\rho}Y(s^t)\right\}.$$

The first-order condition with respect to p' is given by

$$\sum_{s^t|s^{t-1}}Q(s^t|s^{t-1})\left\{(1-\tau_r)(\rho-1)\left(\frac{p_t^s(s^{t-1})}{P(s^t)}\right)^{-\rho}Y(s^t)-\rho\frac{1}{p_t^s(s^{t-1})}\frac{W(s^t)}{A(s_t)}\left(\frac{p_t^s(s^{t-1})}{P(s^t)}\right)^{-\rho}Y(s^t)\right\}=0.$$

Rearranging gives us:

$$\sum_{s^t|s^{t-1}}Q(s^t|s^{t-1})Y(s^t)\left(\frac{p_t^s(s^{t-1})}{P(s^t)}\right)^{-\rho}\left\{p_t^s(s^{t-1})-\left[(1-\tau_r)\left(\frac{\rho-1}{\rho}\right)\right]^{-1}\frac{W(s^t)}{A(s_t)}\right\}=0.$$

Substituting in the equilibrium Arrow prices $Q(s^t|s^{t-1})$ from (14) yields:

$$\sum_{s^t|s^{t-1}} \mu(s^t|s^{t-1}) \frac{U_C^m(s^t)}{P(s^t)} Y(s^t) P(s^t)^\rho \left\{ p_t^s(s^{t-1}) - \left[(1-\tau_r) \left(\frac{\rho-1}{\rho} \right) \right]^{-1} \frac{W(s^t)}{A(s_t)} \right\} = 0.$$

Solving this for $p_t^s(s^{t-1})$ gives us (22) with $q(s^t|s^{t-1})$ defined in (23).

A.5 Proof of Proposition 1

Necessity. In any flexible-price equilibrium, all firms set the same nominal price. The demand functions in (4) imply that all firms produce the same level of output, proving necessity of $y^j(s^t) = Y(s^t)$ for all $j \in \mathcal{J}$.

Aggregation over the optimal price (21) implies that the aggregate price level is given by:

$$P(s^t) = \left[(1-\tau_r) \left(\frac{\rho-1}{\rho} \right) \right]^{-1} \frac{W(s^t)}{A(s_t)}, \qquad \forall s^t \in S^t. \tag{62}$$

Condition (24) follows from combining (62) with the household's intratemporal optimality condition (13), and letting χ denote the labor wedge as follows:

$$\chi \equiv \left(\frac{\rho-1}{\rho} \right) \frac{(1-\tau_\ell)(1-\tau_r)}{1+\tau_c}. \tag{63}$$

Finally, the derivation of the set of necessary conditions (18) is provided in Appendix A.3.

Sufficiency. Take any feasible allocation $x \in \mathcal{X}$, vector $\varphi \equiv (\varphi^i)$, and scalars $\bar{T} \in \mathbb{R}$ and $\chi \in \mathbb{R}_+$ that satisfy conditions (i)-(iii) of Proposition 1. We show that there exists a price system ϱ, a policy Ω, and asset holdings ζ, that support x as a flexible-price equilibrium; we construct these objects as follows.

First, for all $s^t \in S^t$, we set intermediate-good prices according to:

$$p_t^j(s^t) = p_t^f(s^t) = P(s^t) = 1, \qquad \forall j \in \mathcal{J},$$

in which we normalize the aggregate price level to one. These prices, combined with condition (i) of Proposition 1 ensure that the CES demand function (4) is satified for all goods, $j \in \mathcal{J}$.

Second, we set the tax rates $(\tau_\ell, \tau_c, \tau_r)$ such that they jointly satisfy:

$$\frac{(1-\tau_\ell)(1-\tau_r)}{1+\tau_c} = \left(\frac{\rho-1}{\rho} \right)^{-1} \chi. \tag{64}$$

For any strictly positive χ and $\rho > 1$, such tax rates exist. Combining this with condition (24), we obtain the following:

$$-\frac{U_L^m(s^t)}{U_C^m(s^t)} \left(\frac{1+\tau_c}{1-\tau_\ell} \right) = \left(\frac{\rho-1}{\rho} \right) (1-\tau_r) A(s_t). \tag{65}$$

Given tax rates $(\tau_\ell, \tau_c, \tau_r)$, we set the real wage $W(s^t)$ as follows:

$$W(s^t) = -\frac{U_L^m(s^t)}{U_C^m(s^t)}\left(\frac{1+\tau_c}{1-\tau_\ell}\right), \tag{66}$$

and therefore satisfy the household's intratemporal condition in (13). Substituting the above expression for the real wage into (65) and re-arranging gives us:

$$1 = \left[(1-\tau_r)\left(\frac{\rho-1}{\rho}\right)\right]^{-1}\frac{W(s^t)}{A(s_t)}. \tag{67}$$

Therefore the flexible-price firm's optimality condition (21) is satisfied.

Next, we set Arrow prices and the nominal interest rate as follows:

$$Q(s^{t+1}|s^t) = \beta\mu(s^{t+1}|s^t)\frac{U_C^m(s^{t+1})}{U_C^m(s^t)} \qquad \text{and} \qquad 1 = \beta(1+i(s^t))\sum_{s^{t+1}|s^t}\mu(s^{t+1}|s^t)\frac{U_C^m(s^{t+1})}{U_C^m(s^t)},$$

and therefore satisfy the household's intertemporal conditions in (14) and (14).

What remains to be shown is that we may construct bond holdings such that the household's budget constraints are satisfied at this allocation in every history. To do so, we first choose any sequence $\bar{T}(s^t)$ that satifies the following condition:

$$\bar{T} = \frac{1}{U_c^m(s_0)(1+\tau_c)}\sum_t\sum_{s^t}\beta^t\mu(s^t)U_c^m(s^t)\bar{T}(s^t).$$

Next we take the household's budget constraint in (2) for type $i \in I$. Multiplying both sides by $\beta^t\mu(s^t)\Lambda^i(s^t)$ and summing over all periods and states following period r, history s^r, we get:

$$\sum_{t=r+1}^{\infty}\sum_{s^t}\beta^t\mu(s^t)\Lambda^i(s^t)\left[(1+\tau_c)c^i(s^t) + b^i(s^t) + \sum_{s^{t+1}|s^t}Q(s^{t+1}|s^t)z^i(s^{t+1}|s^t)\right]$$
$$= \sum_{t=r+1}^{\infty}\sum_{s^t}\beta^t\mu(s^t)\Lambda^i(s^t)\left[(1-\tau_\ell)W(s^t)\ell^i(s^t) + \bar{T}(s^t) + (1+i(s^{t-1}))b^i(s^{t-1}) + z^i(s^t|s^{t-1})\right]$$

where we let $T(s^t) + (1-\tau_\Pi)\Pi(s^t) = \bar{T}(s^t)$. Substituting in the household's FOCs for bonds (54) and Arrow securities (55) we get:

$$\sum_{t=r+1}^{\infty}\sum_{s^t}\beta^t\mu(s^t)\Lambda^i(s^t)\left[(1+\tau_c)c^i(s^t)\right] = \sum_{t=r+1}^{\infty}\sum_{s^t}\beta^t\mu(s^t)\Lambda^i(s^t)\left[(1-\tau_\ell)W(s^t)\ell^i(s^t) + \bar{T}(s^t)\right]$$
$$+ \sum_{s^{r+1}|s^r}\beta^{r+1}\mu(s^{r+1})\Lambda^i(s^{r+1})(1+i(s^r))b^i(s^r)$$

Rearranging gives us:

$$\beta^r\mu(s^r)\Lambda^i(s^r)b^i(s^r) = \sum_{t=r+1}^{\infty}\sum_{s^t}\beta^t\mu(s^t)\Lambda^i(s^t)\left[(1+\tau_c)c^i(s^t) - (1-\tau_\ell)W(s^t)\ell^i(s^t) - \bar{T}(s^t)\right]$$

Next, using the household's FOCs for consumption and labor (52) and (53), we obtain:

$$\frac{\beta^r\mu(s^r)U_c^i(s^r)}{1+\tau_c}b^i(s^r) = \sum_{t=r+1}^{\infty}\sum_{s^t}\beta^t\mu(s^t)\left[U_c^i(s^t)c^i(s^t) + \frac{1}{\theta^i(s_t)}U_\ell^i(s^t)\ell^i(s^t) - \frac{U_c^i(s^t)}{(1+\tau_c)}\bar{T}(s^t)\right]$$

which we may rewrite as follows:

$$\frac{U_c^i(s^r)}{1+\tau_c}b^i(s^r) = \sum_{t=r+1}^{\infty}\sum_{s^t}\beta^{t-r}\mu(s^t|s^r)\left[U_c^i(s^t)c^i(s^t) + \frac{1}{\theta^i(s_t)}U_\ell^i(s^t)\ell^i(s^t) - \frac{U_c^i(s^t)}{(1+\tau_c)}\bar{T}(s^t)\right].$$

Therefore real bond holdings of household i are given by

$$b^i(s^r) = \left(\frac{U_c^i(s^r)}{1+\tau_c}\right)^{-1}\left\{\sum_{t=r+1}^{\infty}\sum_{s^t}\beta^{t-r}\mu(s^t|s^r)\left[U_c^i(s^t)c^i(s^t) + \frac{1}{\theta^i(s_t)}U_\ell^i(s^t)\ell^i(s^t) - \frac{U_c^i(s^t)}{(1+\tau_c)}\bar{T}(s^t)\right]\right\}$$

for any period r, history s^r.

A.6 Proof of Proposition 2

Necessity. Condition (21) indicates that all flexible-price firms set the same nominal price; similarly condition (28) indicates that all sticky-price firms set the same nominal price. Combining this observation with the demand functions,

$$\frac{y^f(s^t)}{Y(s^t)} = \left[\frac{p_t^f(s^t)}{P(s^t)}\right]^{-\rho} \qquad \text{and} \qquad \frac{y^s(s^t)}{Y(s^t)} = \left[\frac{p_t^s(s^{t-1})}{P(s^t)}\right]^{-\rho}. \tag{68}$$

we infer that all flexible-price firms produce the same level of output and all sticky-price firms produce the same level of output, denoted by $y^f(s^t)$ and $y^s(s^t)$, respectively.

The flexible price firm sets its price according to (21). Rearranging and dividing through by $P(s^t)$ gives us:

$$\frac{p_t^f(s^t)}{P(s^t)} - \left[(1-\tau_r)\left(\frac{\rho-1}{\rho}\right)\right]^{-1}\frac{W(s^t)}{P(s^t)A(s_t)} = 0.$$

The flexible-price firm optimality condition can be written as follows:

$$\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho} - \left[(1-\tau_r)\left(\frac{\rho-1}{\rho}\right)\right]^{-1}\frac{W(s^t)}{P(s^t)A(s_t)} = 0.$$

Combining the above condition with the household's intratemporal optimality condition (13) yields equilibrium necessary condition (26) with χ defined in (63).

As shown in Section A.4 of the Appendix, the sticky price firm sets its price according to

$$\sum_{s^t|s^{t-1}}\mu(s^t|s^{t-1})U_C^m(s^t)Y(s^t)\left(\frac{p_t^s(s^{t-1})}{P(s^t)}\right)^{-\rho}\left\{\frac{p_t^s(s^{t-1})}{P(s^t)} - \left[(1-\tau_r)\left(\frac{\rho-1}{\rho}\right)\right]^{-1}\frac{W(s^t)}{P(s^t)A(s_t)}\right\} = 0.$$

Using the CES demand function (68), the sticky-price firm optimality condition can be written as:

$$\sum_{s^t|s^{t-1}} \mu(s^t|s^{t-1}) U_C^m(s^t) Y(s^t) \frac{y^s(s^t)}{Y(s^t)} \left\{ \left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho} - \left[(1-\tau_r)\left(\frac{\rho-1}{\rho}\right)\right]^{-1} \frac{W(s^t)}{P(s^t)A(s_t)} \right\} = 0$$

Combining the above condition with the household's intratemporal optimality condition (13) yields the equilibrium necessary condition (27) with χ defined in (63). Finally, the derivation of the set of necessary conditions (18) is provided in Appendix A.3.

Sufficiency. Take any feasible allocation $x \in \mathcal{X}$, vector $\varphi \equiv (\varphi^i)$, and scalars $\bar{T} \in \mathbb{R}$ and $\chi \in \mathbb{R}_+$ that satisfy conditions (i)-(iii) of Proposition 2. We show that there exists a price system ϱ, a policy Ω, and asset holdings ζ, that support x as a sticky-price equilibrium; we construct these as follows.

First, we construct nominal prices as follows. First, without loss of generality we can decompose the sticky price firm's output into two components:

$$y^s(s^t) = \phi^s(s^{t-1})\Phi(s^t).$$

[This is without loss because one can always set $\phi^s(s^{t-1}) = 1$ or any arbitary constant.] Next we define a function $\phi^f(s^t) \equiv y^f(s^t)/\Phi(s^t)$ and set prices as follows:

$$p_t^s(s^{t-1}) = \phi^s(s^{t-1})^{-1/\rho}$$

for all firms $j \in \mathcal{J}^s$, and

$$p_t^f(s^t) = \phi^f(s^t)^{-1/\rho}$$

for all firms $j \in \mathcal{J}^f$. Note that these imply that the aggregate price level satisfies:

$$P(s^t) = \left[\frac{Y(s^t)}{\Phi(s^t)}\right]^{-1/\rho}.$$

As such, the CES demand curves in (68) are satisfied. Given this price level, we set the nominal interest rate and the Arrow security prices such that conditions (14) and (15) are satisfied.

Next, we set the tax rates $(\tau_\ell, \tau_c, \tau_r)$ such that they jointly satisfy (64). For any strictly positive χ and $\rho > 1$, such tax rates exist. Combining this with conditions (26) and (27), we obtain the following two conditions:

$$\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho} + \frac{U_L^m(s^t)}{U_C^m(s^t)} \frac{1+\tau_c}{1-\tau_\ell} \left[(1-\tau_r)\left(\frac{\rho-1}{\rho}\right)\right]^{-1} \frac{1}{A(s_t)} = 0, \tag{69}$$

and

$$\sum_{s^t|s^{t-1}} \mu(s^t|s^{t-1}) U_C^m(s^t) y^s(s^t) \left\{ \left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho} + \frac{U_L^m(s^t)}{U_C^m(s^t)} \frac{1+\tau_c}{1-\tau_\ell} \left[(1-\tau_r)\left(\frac{\rho-1}{\rho}\right)\right]^{-1} \frac{1}{A(s_t)} \right\} = 0. \tag{70}$$

Given tax rates $(\tau_\ell, \tau_c, \tau_r)$, we set the nominal wage as follows:

$$W(s^t) = -\frac{U_L^m(s^t)}{U_C^m(s^t)}\left(\frac{1+\tau_c}{1-\tau_\ell}\right)P(s^t), \tag{71}$$

and therefore satisfy the household's intratemporal condition in (13). Substituting the above expression for the real wage into (69) and (70) rearranging gives us:

$$\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho} - \frac{W(s^t)}{P(s^t)}\left[(1-\tau_r)\left(\frac{\rho-1}{\rho}\right)\right]^{-1}\frac{1}{A(s_t)} = 0.$$

and

$$\sum_{s^t|s^{t-1}} \mu(s^t|s^{t-1})U_C^m(s^t)y^s(s^t)\left\{\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho} - \frac{W(s^t)}{P(s^t)}\left[(1-\tau_r)\left(\frac{\rho-1}{\rho}\right)\right]^{-1}\frac{1}{A(s_t)}\right\} = 0.$$

Combining these with the CES demand functions in (68) we get the following two conditions:

$$\frac{p_t^f(s^t)}{P(s^t)} - \frac{W(s^t)}{P(s^t)}\left[(1-\tau_r)\left(\frac{\rho-1}{\rho}\right)\right]^{-1}\frac{1}{A(s_t)} = 0.$$

and

$$\sum_{s^t|s^{t-1}} \mu(s^t|s^{t-1})U_C^m(s^t)Y(s^t)\left(\frac{p_t^s(s^{t-1})}{P(s^t)}\right)^{-\rho}\left\{\frac{p_t^s(s^{t-1})}{P(s^t)} - \frac{W(s^t)}{P(s^t)}\left[(1-\tau_r)\left(\frac{\rho-1}{\rho}\right)\right]^{-1}\frac{1}{A(s_t)}\right\} = 0.$$

Finally, with some rearrangement, these imply:

$$p_t^f(s^t) - \left[(1-\tau_r)\left(\frac{\rho-1}{\rho}\right)\right]^{-1}\frac{W(s^t)}{A(s_t)} = 0.$$

and

$$\sum_{s^t|s^{t-1}} \mu(s^t|s^{t-1})\frac{U_C^m(s^t)}{P(s^t)}Y(s^t)P(s^t)^\rho\left\{p_t^s(s^{t-1}) - \left[(1-\tau_r)\left(\frac{\rho-1}{\rho}\right)\right]^{-1}\frac{W(s^t)}{A(s_t)}\right\} = 0.$$

Therefore both the flexible-price and the sticky-price firm's optimality conditions, (21) and (22), are satisfied.

Finally, what remains to be shown is that we may construct bond holdings such that the household's budget sets are satisfied at this allocation at every history. For this we follow the exact same steps above used to obtain equilibrium bond holdings in the sufficiency portion of the proof of Proposition 1. Following these steps, real bond holdings of household i are given by

$$\frac{b^i(s^r)}{P(s^t)} = \left(\frac{U_c^i(s^r)}{1+\tau_c}\right)^{-1}\left\{\sum_{t=r+1}^{\infty}\sum_{s^t}\beta^{t-r}\mu(s^t|s^r)\left[U_c^i(s^t)c^i(s^t) + \frac{1}{\theta^i(s_t)}U_\ell^i(s^t)\ell^i(s^t) - \frac{U_c^i(s^t)}{(1+\tau_c)}\bar{T}(s^t)\right]\right\}$$

for any period r, history s^r.

A.7 Proof of Proposition 3

The Relaxed Ramsey planner's problem is to choose an allocation $x \in \mathcal{X}$, market weights $\varphi \equiv (\varphi^i)$, and scalar $\bar{T} \in \mathbb{R}$, in order to maximize the pseudo-welfare function in (32) subject to technology and resource constraints (6)-(8). First, note that in any history s^t, the planner can solve a static sub-problem: maximize final good output $Y(s^t)$ given productivity $A(s_t)$ and aggregate labor supply, $L(s^t)$. Specifically:

$$Y(s^t) = \max_{(n^j(s^t))_{j\in\mathcal{J}}} \left[\int_{j\in\mathcal{J}} (A(s_t)n^j(s^t))^{\frac{\rho-1}{\rho}} \mathrm{d}j\right]^{\frac{\rho}{\rho-1}} \qquad \text{subject to} \qquad L(s^t) = \int_{j\in\mathcal{J}} n^j(s^t)\mathrm{d}j.$$

The first-order conditions for this sub-problem yield: $n^j(s^t) = n^{j'}(s^t) = L(s^t)$ for all $j, j' \in \mathcal{J}$, which implies that at the planner's optimum $y^j(s^t) = Y(s^t) = A(s_t)L(s^t)$ for all $j \in \mathcal{J}$. Using this, we can rewrite the relaxed planner's problem in terms of aggregates alone:

$$\max_{\{C(s^t),L(s^t)\},\varphi,\bar{T}} \sum_t \sum_{s^t} \beta^t \mu(s^t) \mathcal{W}(C(s^t), L(s^t); \varphi, \nu, \lambda) - U_C^m(s_0) \sum_{i\in I} \pi^i \nu^i \bar{T}$$

subject to

$$C(s^t) = A(s_t)L(s^t), \qquad \forall s^t \in S^t. \tag{72}$$

We let $\beta^t\mu(s^t)\hat{\varsigma}(s^t)$ denote the Lagrange multiplier on the time t, history s^t resource constraint (72). The first-order conditions of this problem are given by:

$$\beta^t \mu(s^t)\mathcal{W}_C(s^t) - \beta^t\mu(s^t)\hat{\varsigma}(s^t) = 0,$$
$$\beta^t \mu(s^t)\mathcal{W}_L(s^t) + \beta^t\mu(s^t)\hat{\varsigma}(s^t)A(s_t) = 0.$$

Combining, we obtain the relaxed planner's optimality condition in (33).

A.8 The Ramsey Optimum

In this section of the appendix, we solve the Ramsey problem stated in Section 5. We let $\beta^t\mu(s^t)(1-\kappa)\xi(s^t)$ and $\beta^t\mu(s^{t-1})\kappa\upsilon(s^{t-1})$ denote the Lagrange multipliers on the implementability conditions (26) and (27), respectively. We obtain the following Ramsey optimality condition.

Proposition 8. *A Ramsey optimum x^* satisfies, for all $s^t \in S^t$,*

$$-\frac{\mathcal{W}_L(s^t) + (U_L^m(s^t) + U_{LL}^m(s^t)L(s^t))\left[\kappa\upsilon(s^{t-1})\frac{y^s(s^t)}{A(s_t)L(s^t)} + (1-\kappa)\xi(s^t)\frac{y^f(s^t)}{A(s_t)L(s^t)}\right]}{\mathcal{W}_C(s^t) + \chi(U_C^m(s^t) + U_{CC}^m(s^t)C(s^t))\left[\kappa\upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{1-1/\rho} + (1-\kappa)\xi(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1-1/\rho}\right]} = \frac{Y(s^t)}{L(s^t)}. \tag{73}$$

The proof of Proposition 8 is found below. Note first that we can rewrite condition (73) as it is stated in the main text in equation (37), with

$$\left[\text{Ramsey wedge}(s^t)\right] \equiv \frac{1+\left(\frac{U_L^m(s^t)+U_{LL}^m(s^t)L(s^t)}{\mathcal{W}_L(s^t)}\right)\left[\kappa\upsilon(s^{t-1})\frac{y^s(s^t)}{A(s_t)L(s^t)}+(1-\kappa)\xi(s^t)\frac{y^f(s^t)}{A(s_t)L(s^t)}\right]}{1+\chi\left(\frac{U_C^m(s^t)+U_{CC}^m(s^t)C(s^t)}{\mathcal{W}_C(s^t)}\right)\left[\kappa\upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{1-1/\rho}+(1-\kappa)\xi(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1-1/\rho}\right]}.$$

Proof. We again incorporate the budget implementability conditions into the planner's maximand via the pseudo-utility function in (32). We write the planner's Lagrangian as follows:

$$\begin{aligned}
\mathcal{L} =& \sum_t \sum_{s^t} \beta^t \mu(s^t) \mathcal{W}(C(s^t), L(s^t); \varphi, \nu, \lambda) \\
&+ \sum_t \sum_{s^t} \beta^t \mu(s^t) \varsigma^Y(s^t) \left\{ \left[\kappa y^s(s^t)^{\frac{\rho-1}{\rho}} + (1-\kappa) y^f(s^t)^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}} - Y(s^t) \right\} \\
&+ \sum_t \sum_{s^t} \beta^t \mu(s^t) \varsigma^L(s^t) \left\{ \kappa \frac{y^s(s^t)}{A(s_t)} + (1-\kappa) \frac{y^f(s^t)}{A(s_t)} - L(s^t) \right\} \\
&+ \sum_t \sum_{s^t} \beta^t \mu(s^t) \varsigma^C(s^t) \left\{ Y(s^t) - C(s^t) \right\} \\
&+ \sum_t \sum_{s^t} \beta^t \mu(s^{t-1}) \kappa \upsilon(s^{t-1}) \sum_{s^t|s^{t-1}} \mu(s^t|s^{t-1}) y^s(s^t) \left\{ \chi U_C^m(s^t) \left[\frac{y^s(s^t)}{Y(s^t)} \right]^{-1/\rho} + U_L^m(s^t) \frac{1}{A(s_t)} \right\} \\
&+ \sum_t \sum_{s^t} \beta^t \mu(s^t) (1-\kappa) \xi(s^t) y^f(s^t) \left\{ \chi U_C^m(s^t) \left(\frac{y^f(s^t)}{Y(s^t)} \right)^{-1/\rho} + U_L^m(s^t) \frac{1}{A(s_t)} \right\}
\end{aligned}$$

The FOC with respect to $y^s(s^t)$ is given by:

$$\begin{aligned}
0 =& \kappa \varsigma^Y(s^t) \left[\kappa y^s(s^t)^{\frac{\rho-1}{\rho}} + (1-\kappa) y^f(s^t)^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}-1} y^s(s^t)^{\frac{\rho-1}{\rho}-1} + \kappa \varsigma^L(s^t) \frac{1}{A(s_t)} \\
&+ \kappa \upsilon(s^{t-1}) \left\{ \chi U_C^m(s^t) \left[\frac{y^s(s^t)}{Y(s^t)} \right]^{-1/\rho} + U_L^m(s^t) \frac{1}{A(s_t)} - \frac{1}{\rho} \chi U_C^m(s^t) \left[\frac{y^s(s^t)}{Y(s^t)} \right]^{-1/\rho} \right\}
\end{aligned} \tag{74}$$

The FOC with respect to $y^f(s^t)$

$$\begin{aligned}
0 =& (1-\kappa) \varsigma^Y(s^t) \left[\kappa y^s(s^t)^{\frac{\rho-1}{\rho}} + (1-\kappa) y^f(s^t)^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}-1} y^f(s^t)^{\frac{\rho-1}{\rho}-1} + (1-\kappa) \varsigma^L(s^t) \frac{1}{A(s_t)} \\
&+ (1-\kappa) \xi(s^t) \left\{ \chi U_C^m(s^t) \left(\frac{y^f(s^t)}{Y(s^t)} \right)^{-1/\rho} + U_L^m(s^t) \frac{1}{A(s_t)} - \frac{1}{\rho} \chi U_C^m(s^t) \left[\frac{y^f(s^t)}{Y(s^t)} \right]^{-1/\rho} \right\}
\end{aligned} \tag{75}$$

Note that (74) can equivalently be written as:

$$\begin{aligned}
0 =& \kappa \varsigma^Y(s^t) \left[\kappa y^s(s^t)^{\frac{\rho-1}{\rho}} + (1-\kappa) y^f(s^t)^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}-1} y^s(s^t)^{\frac{\rho-1}{\rho}} + \kappa \varsigma^L(s^t) \frac{y^s(s^t)}{A(s_t)} \\
&+ \kappa \upsilon(s^{t-1}) y^s(s^t) \left\{ \chi U_C^m(s^t) \left[\frac{y^s(s^t)}{Y(s^t)} \right]^{-1/\rho} \frac{\rho-1}{\rho} + U_L^m(s^t) \frac{1}{A(s_t)} \right\}
\end{aligned} \tag{76}$$

or

$$0 = \varsigma^Y(s^t)\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho} + \varsigma^L(s^t)\frac{1}{A(s_t)} + \upsilon(s^{t-1})\left\{\chi U_C^m(s^t)\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho}\frac{\rho-1}{\rho} + U_L^m(s^t)\frac{1}{A(s_t)}\right\} \tag{77}$$

Similarly, note that (75) can equivalently be written as:

$$\begin{aligned} 0 =& (1-\kappa)\varsigma^Y(s^t)\left[\kappa y^s(s^t)^{\frac{\rho-1}{\rho}} + (1-\kappa)y^f(s^t)^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}-1} y^f(s^t)^{\frac{\rho-1}{\rho}} + (1-\kappa)\varsigma^L(s^t)\frac{y^f(s^t)}{A(s_t)} \\ &+ (1-\kappa)\xi(s^t)y^f(s^t)\left\{\chi U_C^m(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho}\frac{\rho-1}{\rho} + U_L^m(s^t)\frac{1}{A(s_t)}\right\} \end{aligned} \tag{78}$$

or

$$0 = \varsigma^Y(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho} + \varsigma^L(s^t)\frac{1}{A(s_t)} + \xi(s^t)\left\{\chi U_C^m(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho}\frac{\rho-1}{\rho} + U_L^m(s^t)\frac{1}{A(s_t)}\right\} \tag{79}$$

Adding (76) to (77) gives us:

$$\begin{aligned} 0 =& \varsigma^Y(s^t)Y(s^t) + \varsigma^L(s^t)L(s^t) \\ &+ \chi\frac{\rho-1}{\rho}U_C^m(s^t)Y(s^t)\left[\kappa\upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{1-1/\rho} + (1-\kappa)\xi(s^t)\left(\frac{y^f(s^t)}{Y(s^t)}\right)^{1-1/\rho}\right] \\ &+ U_L^m(s^t)\frac{1}{A(s_t)}\left[\kappa\upsilon(s^{t-1})y^s(s^t) + (1-\kappa)\xi(s^t)y^f(s^t)\right] \end{aligned} \tag{80}$$

We can rewrite the above condition as follows:

$$-\frac{\varsigma^L(s^t) + U_L^m(s^t)\frac{1}{A(s_t)L(s^t)}\left[\kappa\upsilon(s^{t-1})y^s(s^t) + (1-\kappa)\xi(s^t)y^f(s^t)\right]}{\varsigma^Y(s^t) + \chi\left(1-\frac{1}{\rho}\right)U_C^m(s^t)\left[\kappa\upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{1-1/\rho} + (1-\kappa)\xi(s^t)\left(\frac{y^f(s^t)}{Y(s^t)}\right)^{1-1/\rho}\right]} = \frac{Y(s^t)}{L(s^t)} \tag{81}$$

Next, the FOC with respect to $C(s^t)$ is given by:

$$0 = \mathcal{W}_C(s^t) - \varsigma^C(s^t) + \kappa\upsilon(s^{t-1})\chi y^s(s^t)U_{CC}^m(s^t)\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho} + (1-\kappa)\xi(s^t)\chi y^f(s^t)U_{CC}^m(s^t)\left(\frac{y^f(s^t)}{Y(s^t)}\right)^{-1/\rho}, \tag{82}$$

The FOC with respect to $Y(s^t)$ is given by:

$$0 = -\varsigma^Y(s^t) + \varsigma^C(s^t) + \frac{1}{\rho}\kappa\upsilon(s^{t-1})\chi U_C^m(s^t)\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho}\frac{y^s(s^t)}{Y(s^t)} + \frac{1}{\rho}(1-\kappa)\xi(s^t)\chi U_C^m(s^t)\left(\frac{y^f(s^t)}{Y(s^t)}\right)^{-1/\rho}\frac{y^f(s^t)}{Y(s^t)} \tag{83}$$

The FOC with respect to $L(s^t)$ is given by:

$$0 = \mathcal{W}_L(s^t) - \varsigma^L(s^t) + \kappa\upsilon(s^{t-1})y^s(s^t)U_{LL}^m(s^t)\frac{1}{A(s_t)} + (1-\kappa)\xi(s^t)y^f(s^t)U_{LL}^m(s^t)\frac{1}{A(s_t)}, \tag{84}$$

Combining (82) and (83) we get:

$$\varsigma^Y(s^t) = \mathcal{W}_C(s^t) + \kappa\upsilon(s^{t-1})\chi y^s(s^t)U^m_{CC}(s^t)\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho} + (1-\kappa)\xi(s^t)\chi y^f(s^t)U^m_{CC}(s^t)\left(\frac{y^f(s^t)}{Y(s^t)}\right)^{-1/\rho}$$
$$+\frac{1}{\rho}\kappa\upsilon(s^{t-1})\chi U^m_C(s^t)\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho}\frac{y^s(s^t)}{Y(s^t)} + \frac{1}{\rho}(1-\kappa)\xi(s^t)\chi U^m_C(s^t)\left(\frac{y^f(s^t)}{Y(s^t)}\right)^{-1/\rho}\frac{y^f(s^t)}{Y(s^t)}$$

This reduces to:

$$\varsigma^Y(s^t) = \mathcal{W}_C(s^t) + \chi\left\{U^m_{CC}(s^t)Y(s^t) + \frac{1}{\rho}U^m_C(s^t)\right\}\left[\kappa\upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{1-1/\rho} + (1-\kappa)\xi(s^t)\left(\frac{y^f(s^t)}{Y(s^t)}\right)^{1-1/\rho}\right] \tag{85}$$

and from (84) we have:

$$\varsigma^L(s^t) = \mathcal{W}_L(s^t) + U^m_{LL}(s^t)\frac{1}{A(s_t)}\left[\kappa\upsilon(s^{t-1})y^s(s^t) + (1-\kappa)\xi(s^t)y^f(s^t)\right] \tag{86}$$

Substituting these into (81) and noting that $Y(s^t) = C(s^t)$, for all $s^t \in S^t$, we obtain:

$$-\frac{\mathcal{W}_L(s^t) + \left\{U^m_{LL}(s^t)\frac{1}{A(s_t)} + U^m_L(s^t)\frac{1}{A(s_t)L(s^t)}\right\}\left[\kappa\upsilon(s^{t-1})y^s(s^t) + (1-\kappa)\xi(s^t)y^f(s^t)\right]}{\mathcal{W}_C(s^t) + \chi\left\{U^m_{CC}(s^t)C(s^t) + U^m_C(s^t)\right\}\left[\kappa\upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}} + (1-\kappa)\xi(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}}\right]} = \frac{Y(s^t)}{L(s^t)},$$

which can be rewritten as follows:

$$-\frac{\mathcal{W}_L(s^t) + U^m_L(s^t)\left\{\frac{U^m_{LL}(s^t)L(s^t)}{U^m_L(s^t)} + 1\right\}\left[\kappa\upsilon(s^{t-1})\frac{y^s(s^t)}{A(s_t)L(s^t)} + (1-\kappa)\xi(s^t)\frac{y^f(s^t)}{A(s_t)L(s^t)}\right]}{\mathcal{W}_C(s^t) + \chi U^m_C(s^t)\left\{\frac{U^m_{CC}(s^t)C(s^t)}{U^m_C(s^t)} + 1\right\}\left[\kappa\upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}} + (1-\kappa)\xi(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}}\right]} = \frac{Y(s^t)}{L(s^t)} \tag{87}$$

as in (73). □

A.9 Proof of Proposition 4

Iso-elastic preferences satisfy:

$$\frac{U^m_{CC}(s^t)C(s^t)}{U^m_C(s^t)} = -\gamma \qquad \text{and} \qquad \frac{U^m_{LL}(s^t)L(s^t)}{U^m_L(s^t)} = \eta$$

This implies that (87) can be written as follows:

$$-\frac{\mathcal{W}_L(s^t) + (1+\eta)U^m_L(s^t)\frac{Y(s^t)}{A(s_t)L(s^t)}\left[\kappa\upsilon(s^{t-1})\frac{y^s(s^t)}{Y(s^t)} + (1-\kappa)\xi(s^t)\frac{y^f(s^t)}{Y(s^t)}\right]}{\mathcal{W}_C(s^t) + \chi(1-\gamma)U^m_C(s^t)\left[\kappa\upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}} + (1-\kappa)\xi(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}}\right]} = \frac{Y(s^t)}{L(s^t)} \tag{88}$$

This holds for any arbitrary χ. We combine this with the implementability condition (26) and obtain:

$$\frac{\frac{\mathcal{W}_L(s^t)}{U_L^m(s^t)}+(1+\eta)\frac{Y(s^t)}{A(s_t)L(s^t)}\left[\kappa\upsilon(s^{t-1})\frac{y^s(s^t)}{Y(s^t)}+(1-\kappa)\xi(s^t)\frac{y^f(s^t)}{Y(s^t)}\right]}{\chi^{-1}\frac{\mathcal{W}_C(s^t)}{U_C^m(s^t)}+(1-\gamma)\left[\kappa\upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}}+(1-\kappa)\xi(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}}\right]}=\left(\frac{y^f(s^t)}{Y(s^t)}\right)^{1/\rho}\frac{Y(s^t)}{A(s_t)L(s^t)} \tag{89}$$

With proportional aggregate shocks we have that:

$$\mathcal{W}_C(s^t)=U_C^m(s^t)\Omega_C(\varphi) \qquad \text{and} \qquad \mathcal{W}_L(s^t)=U_L^m(s^t)\Omega_L(\varphi).$$

where

$$\Omega_C(\varphi)=\sum_{i\in I}\pi^i\omega_C^i(\varphi)\left[\frac{\lambda^i}{\varphi^i}+\nu^i(1-\gamma)\right] \qquad \text{and} \qquad \Omega_L\equiv\sum_{i\in I}\pi^i\omega_L^i(\varphi)\left[\frac{\lambda^i}{\varphi^i}+\nu^i(1+\eta)\right]$$

are constants. Substituting this into (89) gives us the following optimality condition:

$$\frac{\Omega_L(\varphi)+(1+\eta)\frac{Y(s^t)}{A(s_t)L(s^t)}\left[\kappa\upsilon(s^{t-1})\frac{y^s(s^t)}{Y(s^t)}+(1-\kappa)\xi(s^t)\frac{y^f(s^t)}{Y(s^t)}\right]}{\chi^{-1}\Omega_C(\varphi)+(1-\gamma)\left[\kappa\upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}}+(1-\kappa)\xi(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}}\right]}=\left(\frac{y^f(s^t)}{Y(s^t)}\right)^{1/\rho}\frac{Y(s^t)}{A(s_t)L(s^t)} \tag{90}$$

Next we combine FOCs (77) and (79) in order to obtain:

$$\frac{\varsigma^Y(s^t)+\upsilon(s^{t-1})\left\{\chi U_C^m(s^t)\frac{\rho-1}{\rho}+U_L^m(s^t)\frac{1}{A(s_t)}\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{1/\rho}\right\}}{\varsigma^Y(s^t)+\xi(s^t)\left\{\chi U_C^m(s^t)\frac{\rho-1}{\rho}+U_L^m(s^t)\frac{1}{A(s_t)}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1/\rho}\right\}}=\frac{\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho}}{\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho}} \tag{91}$$

Furthermore, condition (85) can be written as:

$$\varsigma^Y(s^t)=\mathcal{W}_C(s^t)+\chi U_C^m(s^t)\left\{\frac{U_{CC}^m(s^t)C(s^t)}{U_C^m(s^t)}+\frac{1}{\rho}\right\}\left[\kappa\upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{1-1/\rho}+(1-\kappa)\xi(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1-1/\rho}\right]$$

Therefore:

$$\varsigma^Y(s^t)=U_C^m(s^t)\Omega_C(\varphi)+\chi U_C^m(s^t)\left\{\frac{1}{\rho}-\gamma\right\}\left[\kappa\upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}}+(1-\kappa)\xi(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}}\right]$$

Substituting this into (91) we get:

$$\frac{\chi^{-1}\Omega_C(\varphi)+\left\{\frac{1}{\rho}-\gamma\right\}\left[\kappa\upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}}+(1-\kappa)\xi(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}}\right]+\upsilon(s^{t-1})\left\{\frac{\rho-1}{\rho}+\frac{U_L^m(s^t)}{U_C^m(s^t)}\frac{1}{\chi A(s_t)}\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{1/\rho}\right\}}{\chi^{-1}\Omega_C(\varphi)+\left\{\frac{1}{\rho}-\gamma\right\}\left[\kappa\upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}}+(1-\kappa)\xi(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}}\right]+\xi(s^t)\left\{\frac{\rho-1}{\rho}+\frac{U_L^m(s^t)}{U_C^m(s^t)}\frac{1}{\chi A(s_t)}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1/\rho}\right\}}$$

Again using the implementability condition (26) we obtain:

$$\frac{\chi^{-1}\Omega_C(\varphi) + \left\{\frac{1}{\rho} - \gamma\right\}\left[\kappa\upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}} + (1-\kappa)\xi(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}}\right] + \upsilon(s^{t-1})\left\{\frac{\rho-1}{\rho} - \left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho}\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{1/\rho}\right\}}{\chi^{-1}\Omega_C(\varphi) + \left\{\frac{1}{\rho} - \gamma\right\}\left[\kappa\upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}} + (1-\kappa)\xi(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}}\right] - \frac{1}{\rho}\xi(s^t)} = \tag{92}$$

We thus have a system of four equations in four unknowns. The four equations are (90), (92), along with the two resource constraints in (36). The four unknowns are:

$$\left\{\frac{Y(s^t)}{A(s_t)L(s^t)}, \frac{y^s(s^t)}{Y(s^t)}, \frac{y^f(s^t)}{Y(s^t)}, \xi(s^t)\right\}$$

We may relabel these as variables as follows:

$$\left\{\tilde{Y}(s^t), \tilde{y}^s(s^t), \tilde{y}^f(s^t), \xi(s^t)\right\} \equiv \left\{\frac{Y(s^t)}{A(s_t)L(s^t)}, \frac{y^s(s^t)}{Y(s^t)}, \frac{y^f(s^t)}{Y(s^t)}, \xi(s^t)\right\} \tag{93}$$

We rewrite the four equations with the relabeled variables as follows:

$$\frac{\Omega_L(\varphi) + (1+\eta)\tilde{Y}(s^t)\left[\kappa\upsilon(s^{t-1})\tilde{y}^s(s^t) + (1-\kappa)\xi(s^t)\tilde{y}^f(s^t)\right]}{\chi^{-1}\Omega_C(\varphi) + (1-\gamma)\left[\kappa\upsilon(s^{t-1})\tilde{y}^s(s^t)^{\frac{\rho-1}{\rho}} + (1-\kappa)\xi(s^t)\tilde{y}^f(s^t)^{\frac{\rho-1}{\rho}}\right]} = \tilde{y}^f(s^t)^{1/\rho}\tilde{Y}(s^t)$$

$$\frac{\chi^{-1}\Omega_C(\varphi) + \left(\frac{1}{\rho} - \gamma\right)\left[\kappa\upsilon(s^{t-1})\tilde{y}^s(s^t)^{\frac{\rho-1}{\rho}} + (1-\kappa)\xi(s^t)\tilde{y}^f(s^t)^{\frac{\rho-1}{\rho}}\right] + \upsilon(s^{t-1})\left\{\frac{\rho-1}{\rho} - \tilde{y}^f(s^t)^{-1/\rho}\tilde{y}^s(s^t)^{1/\rho}\right\}}{\chi^{-1}\Omega_C(\varphi) + \left(\frac{1}{\rho} - \gamma\right)\left[\kappa\upsilon(s^{t-1})\tilde{y}^s(s^t)^{\frac{\rho-1}{\rho}} + (1-\kappa)\xi(s^t)\tilde{y}^f(s^t)^{\frac{\rho-1}{\rho}}\right] - \frac{1}{\rho}\xi(s^t)} = \frac{\tilde{y}^f(s^t)}{\tilde{y}^s(s^t)}$$

$$1 = \kappa\tilde{y}^s(s^t)^{\frac{\rho-1}{\rho}} + (1-\kappa)\tilde{y}^f(s^t)^{\frac{\rho-1}{\rho}},$$

$$1 = \kappa\tilde{y}^s(s^t)\tilde{Y}(s^t) + (1-\kappa)\tilde{y}^f(s^t)\tilde{Y}(s^t).$$

Note that these equations are identical across all states s, s' conditional on s^{t-1}. Therefore, conditional on s^{t-1}, the quadruplet in (93) satisfies:

$$\left\{\tilde{Y}(s), \tilde{y}^s(s), \tilde{y}^f(s), \xi(s)|s^{t-1}\right\} = \left\{\tilde{Y}(s'), \tilde{y}^s(s'), \tilde{y}^f(s'), \xi(s')|s^{t-1}\right\}, \qquad \forall s, s' \in S|s^{t-1} \tag{94}$$

In other words, conditional on history s^{t-1}, there is no variation of these endogenous variables across states.

Finally we use the implementability condition (27). By combining it with (26) it can be written as:

$$\sum_{s^t|s^{t-1}} \mu(s^t|s^{t-1})U_C^m(s^t)y^s(s^t)\left\{\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho} - \left(\frac{y^f(s^t)}{Y(s^t)}\right)^{-1/\rho}\right\} = 0$$

or,

$$\sum_{s^t|s^{t-1}} \mu(s^t|s^{t-1})U_C^m(s^t)y^s(s^t)\left\{\tilde{y}^s(s^t)^{-1/\rho} - \tilde{y}^f(s^t)^{-1/\rho}\right\} = 0$$

This is consistent with the property stated in (94) if and only if:

$$\tilde{y}^s(s^t) = \tilde{y}^f(s^t) = 1, \qquad \forall s^t | s^{t-1}.$$

It is therefore optimal for monetary policy to implement the flexible-price allocation given any arbitrary χ.

A.10 Proof of Theorem 2

At the Ramsey optimum, we have:

$$-\frac{\mathcal{W}_L(s^t) + (1+\eta)U_L^m(s^t)\frac{Y(s^t)}{A(s_t)L(s^t)}\left[\kappa\upsilon(s^{t-1})\frac{y^s(s^t)}{Y(s^t)} + (1-\kappa)\xi(s^t)\frac{y^f(s^t)}{Y(s^t)}\right]}{\mathcal{W}_C(s^t) + \chi(1-\gamma)U_C^m(s^t)\left[\kappa\upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}} + (1-\kappa)\xi(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}}\right]} = \frac{Y(s^t)}{L(s^t)}. \tag{95}$$

With separable and iso-elastic utility, $\mathcal{W}_C(s^t)$ and $\mathcal{W}_L(s^t)$ satisfy:

$$\mathcal{W}_C(s^t) = U_C^m(s^t)\sum_{i\in I}\pi^i\omega_C^i(\varphi)\left[\frac{\lambda^i}{\varphi^i} + \nu^i(1-\gamma)\right] \tag{96}$$

$$\mathcal{W}_L(s^t) = U_L^m(s^t)\sum_{i\in I}\pi^i\omega_L^i(\varphi, s_t)\left[\frac{\lambda^i}{\varphi^i} + \nu^i(1+\eta)\right]. \tag{97}$$

Substituting these expressions for $\mathcal{W}_C(s^t)$ and $\mathcal{W}_L(s^t)$ into (95), we obtain:

$$-\frac{U_L^m(s^t)}{U_C^m(s^t)}\left\{\frac{\sum_{i\in I}\pi^i\omega_L^i(\varphi, s_t)\left[\frac{\lambda^i}{\varphi^i} + \nu^i(1+\eta)\right] + (1+\eta)\frac{Y(s^t)}{A(s_t)L(s^t)}\left[\kappa\upsilon(s^{t-1})\frac{y^s(s^t)}{Y(s^t)} + (1-\kappa)\xi(s^t)\frac{y^f(s^t)}{Y(s^t)}\right]}{\sum_{i\in I}\pi^i\omega_C^i(\varphi)\left[\frac{\lambda^i}{\varphi^i} + \nu^i(1-\gamma)\right] + \chi(1-\gamma)\left[\kappa\upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}} + (1-\kappa)\xi(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}}\right]}\right\} = \frac{Y(s^t)}{L(s^t)}$$

Therefore the optimal monetary wedge satisfies:

$$1-\tau_M^*(s^t) = \frac{(\chi^*)^{-1}\sum_{i\in I}\pi^i\omega_C^i(\varphi)\left[\frac{\lambda^i}{\varphi^i} + \nu^i(1-\gamma)\right] + (1-\gamma)\left[\kappa\upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}} + (1-\kappa)\xi(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}}\right]}{\sum_{i\in I}\pi^i\omega_L^i(\varphi, s_t)\left[\frac{\lambda^i}{\varphi^i} + \nu^i(1+\eta)\right] + (1+\eta)\frac{Y(s^t)}{A(s_t)L(s^t)}\left[\kappa\upsilon(s^{t-1})\frac{y^s(s^t)}{Y(s^t)} + (1-\kappa)\xi(s^t)\frac{y^f(s^t)}{Y(s^t)}\right]}$$

Next we define a function $\mathcal{I}(s_t)$ and a scalar $\mathcal{H}$ as follows:

$$\mathcal{I}(s_t) \equiv \sum_{i\in I}\pi^i\omega_L^i(s^t)\left[\frac{\lambda^i}{\varphi^i} + \nu^i(1+\eta)\right], \qquad \text{and} \qquad \mathcal{H} \equiv (\chi^*)^{-1}\Omega_C.$$

The optimal monetary wedge can then be written as follows:

$$1-\tau_M^*(s^t) = \frac{\mathcal{H} + (1-\gamma)\left[\kappa\upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}} + (1-\kappa)\xi(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}}\right]}{\mathcal{I}(s_t) + (1+\eta)\frac{Y(s^t)}{A(s_t)L(s^t)}\left[\kappa\upsilon(s^{t-1})\frac{y^s(s^t)}{Y(s^t)} + (1-\kappa)\xi(s^t)\frac{y^f(s^t)}{Y(s^t)}\right]}. \tag{98}$$

Threshold. We first consider the conditions under which $\tau_M^*(s^t) = 0$. In this state: $y^s(s^t) = y^f(s^t) = Y(s^t) = A(s_t)L(s^t)$. Condition (98) reduces to:

$$1 = \frac{\mathcal{H} + (1-\gamma)\left[\kappa\upsilon(s^{t-1}) + (1-\kappa)\xi(s^t)\right]}{\mathcal{I}(s_t) + (1+\eta)\left[\kappa\upsilon(s^{t-1}) + (1-\kappa)\xi(s^t)\right]}$$

Furthermore, conditions (77) and (79) imply that $\xi(s^t) = \upsilon(s^{t-1})$ in this state. Therefore:

$$1 = \frac{\mathcal{H} + (1-\gamma)\upsilon(s^{t-1})}{\mathcal{I}(s_t) + (1+\eta)\upsilon(s^{t-1})}$$

Solving this for $\mathcal{I}(s_t)$ we obtain the following threshold:

$$\bar{\mathcal{I}}(s^{t-1}) = \mathcal{H} - (\eta+\gamma)\upsilon(s^{t-1})$$

When $\mathcal{I}(s_t) = \bar{\mathcal{I}}(s^{t-1})$ the optimal monetary tax is equal to zero: $\tau_M^*(s^t) = 0$.

The fictitious tax wedge. We next define a fictitious tax wedge as follows:

$$1 - \hat{\tau}(s^t) \equiv \frac{\mathcal{H} + (1-\gamma)\upsilon(s^{t-1})}{\mathcal{I}(s_t) + (1+\eta)\upsilon(s^{t-1})} \tag{99}$$

This wedge is unambiguously falling in $\mathcal{I}(s_t)$, as all other terms are constants. Furthermore, note that when $\mathcal{I}(s_t) = \bar{\mathcal{I}}(s^{t-1})$, this wedge is equal to one. As a result, the fictitious tax $\hat{\tau}(s^t)$ trivially satisfies:

$$\begin{array}{lll} \hat{\tau}(s^t) > 0 & \text{if and only if} & \mathcal{I}(s_t) > \bar{\mathcal{I}}(s^{t-1}), \\ \hat{\tau}(s^t) = 0 & \text{if and only if} & \mathcal{I}(s_t) = \bar{\mathcal{I}}(s^{t-1}), \\ \hat{\tau}(s^t) < 0 & \text{if and only if} & \mathcal{I}(s_t) < \bar{\mathcal{I}}(s^{t-1}). \end{array}$$

The optimal monetary wedge. The next step of our proof involves characterizing the multipliers $\xi(s^t)$ and $\upsilon(s^{t-1})$.[17]

Lemma 3. *In the optimal allocation,*

$$\begin{array}{lll} y^f(s^t) > y^s(s^t) & \textit{if and only if} & \tau_M^*(s^t) > 0; \\ y^f(s^t) = y^s(s^t) & \textit{if and only if} & \tau_M^*(s^t) = 0; \\ y^f(s^t) < y^s(s^t) & \textit{if and only if} & \tau_M^*(s^t) < 0. \end{array}$$

Furthermore, $\tau_M^*(s^t) = 0$ if and only if

$$\xi(s^t) = \upsilon(s^{t-1}).$$

[17] Note that if $\upsilon(s^{t-1}) < 0$, then $\xi(s^t)/\upsilon(s^{t-1}) < 1$ implies:

$$\upsilon(s^{t-1}) < \xi(s^t) < 0.$$

Proof. See Section A.11 of this appendix. □

Finally we consider the optimal monetary wedge (98), which can be written as

$$1-\tau_M^*(s^t)=\frac{\mathcal{H}+(1-\gamma)\upsilon(s^{t-1})\left[\kappa\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}}+(1-\kappa)\frac{\xi(s^t)}{\upsilon(s^{t-1})}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}}\right]}{\mathcal{I}(s_t)+(1+\eta)\upsilon(s^{t-1})\left[\kappa\frac{y^s(s^t)}{A(s_t)L(s^t)}+(1-\kappa)\frac{\xi(s^t)}{\upsilon(s^{t-1})}\frac{y^f(s^t)}{A(s_t)L(s^t)}\right]}$$

From our resource constraints, note that aggregate labor and aggregate output satisfy:

$$L(s^t)=\kappa\frac{y^s(s^t)}{A(s_t)}+(1-\kappa)\frac{y^f(s^t)}{A(s_t)}, \tag{100}$$

and

$$1=\kappa\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}}+(1-\kappa)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}}, \tag{101}$$

respectively.

Substituting this into our above expression for the optimal monetary wedge and rearranging, we obtain the following expression for the optimal monetary wedge:

$$1-\tau_M^*(s^t)=\frac{\mathcal{H}+(1-\gamma)\upsilon(s^{t-1})+(1-\kappa)(1-\gamma)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}}(\xi(s^t)-\upsilon(s^{t-1}))}{\mathcal{I}(s_t)+(1+\eta)\upsilon(s^{t-1})+(1-\kappa)(1+\eta)\frac{y^f(s^t)}{Y(s^t)}\frac{Y(s^t)}{A(s_t)L(s^t)}(\xi(s^t)-\upsilon(s^{t-1}))} \tag{102}$$

We want to compare this to the fictitious tax wedge defined in (99).

We first consider the state in which:

$$y^s(s^t)=y^f(s^t)=Y(s^t)$$

From Lemma (3) we have that in this state, $\tau_M^*(s^t)=0$ and $\xi(s^t)=\upsilon(s^{t-1})$. Therefore:

$$1-\tau_M^*(s^t)=\frac{\mathcal{H}+(1-\gamma)\upsilon(s^{t-1})}{\mathcal{I}(s_t)+(1+\eta)\upsilon(s^{t-1})}=1-\hat{\tau}(s^t)=1.$$

Therefore $\tau_M^*(s^t)=0$ if and only if $\hat{\tau}(s^t)=0$.

Next we define $x^*(s^t)$ as the inverse of the monetary wedge:

$$x^*(s^t)\equiv\frac{1}{1-\tau_M^*(s^t)},\qquad\text{and}\qquad\hat{x}(s^t)\equiv\frac{1}{1-\hat{\tau}(s^t)}.$$

From (102) we have that:

$$x^*(s^t)=\frac{\mathcal{I}(s_t)+(1+\eta)\upsilon(s^{t-1})+(1-\kappa)(1+\eta)\frac{y^f(s^t)}{Y(s^t)}\frac{Y(s^t)}{A(s_t)L(s^t)}(\xi(s^t)-\upsilon(s^{t-1}))}{\mathcal{H}+(1-\gamma)\upsilon(s^{t-1})+(1-\kappa)(1-\gamma)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}}(\xi(s^t)-\upsilon(s^{t-1}))},$$

and from (99) we have that

$$\hat{x}(s^t) = \frac{\mathcal{I}(s_t) + (1+\eta)\upsilon(s^{t-1})}{\mathcal{H} + (1-\gamma)\upsilon(s^{t-1})}.$$

Therefore:

$$x^*(s^t)\left\{1 + \frac{(1-\kappa)(1-\gamma)}{\mathcal{H} + (1-\gamma)\upsilon(s^{t-1})}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}}(\xi(s^t) - \upsilon(s^{t-1}))\right\} = \hat{x}(s^t) + \frac{(1-\kappa)(1+\eta)}{\mathcal{H} + (1-\gamma)\upsilon(s^{t-1})}\frac{y^f(s^t)}{Y(s^t)}\frac{Y(s^t)}{A(s_t)L(s^t)}(\xi(s$$

which implies:

$$x^*(s^t) = \hat{x}(s^t) + \frac{(1-\kappa)}{\mathcal{H} + (1-\gamma)\upsilon(s^{t-1})}\left[(1+\eta)\frac{y^f(s^t)}{Y(s^t)}\frac{Y(s^t)}{A(s_t)L(s^t)} - (1-\gamma)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}} x^*(s^t)\right](\xi(s^t) - \upsilon(s^{t-1})) \tag{103}$$

Next, we combine the monetary wedge defined in (38) with the implementability condition (26). Doing so yields the following equation:

$$\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho} = (1 - \tau^*_M(s^t))\frac{Y(s^t)}{A(s_t)L(s^t)} \tag{104}$$

Therefore:

$$x^*(s^t) = \left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1/\rho}\frac{Y(s^t)}{A(s_t)L(s^t)}$$

Substituting this expression into (103) we get:

$$x^*(s^t) = \hat{x}(s^t) + \frac{(1-\kappa)(\eta+\gamma)}{\mathcal{H} + (1-\gamma)\upsilon(s^{t-1})}\frac{y^f(s^t)}{Y(s^t)}\frac{Y(s^t)}{A(s_t)L(s^t)}(\xi(s^t) - \upsilon(s^{t-1})) \tag{105}$$

It is clear from (105) that, to a first order around $\hat{\tau}(s^t) = 0$,

$$x^*(s^t) \approx \hat{x}(s^t).$$

Therefore, for small ϵ, if $\hat{\delta}(s^t) < 1$ then $\delta^*(s^t) < 1$.

$$x^*(\epsilon) = \hat{x}(\epsilon) + \frac{(1-\kappa)(\eta+\gamma)}{\mathcal{H} + (1-\gamma)\upsilon}\frac{y^f(\epsilon)}{Y(\epsilon)}\frac{Y(\epsilon)}{AL(\epsilon)}(\xi(\epsilon) - \upsilon)$$

$$x^*(\epsilon) = \hat{x}(\epsilon) + g(\epsilon)$$

For ϵ small:

$$x^*(\epsilon) \approx \hat{x}(\epsilon)$$

Now consider an arbitrary $\epsilon' > 0$. There exists an $\epsilon > 0$ such that:

$$0 < \epsilon < \epsilon'$$

and

$$x^*(\epsilon) < 1.$$

Suppose that for some ϵ

This follows from the fact that

$$x^*(\epsilon') > 1$$

implies that

$$\hat{x}(\epsilon') + g(\epsilon') > 1$$

Then

$$g(\epsilon') > 1 - \hat{x}(\epsilon') > 0$$

But note that $g(\epsilon')$ is a continuous function. Therefore..

****

To conclude, the optimal monetary tax rate $\tau_M^*(s^t)$ satisfies:

$$\begin{array}{lll} \tau_M^*(s^t) > 0 & \text{if and only if} & \mathcal{I}(s_t) > \bar{\mathcal{I}}(s^{t-1}), \\ \tau_M^*(s^t) = 0 & \text{if and only if} & \mathcal{I}(s_t) = \bar{\mathcal{I}}(s^{t-1}), \\ \tau_M^*(s^t) < 0 & \text{if and only if} & \mathcal{I}(s_t) < \bar{\mathcal{I}}(s^{t-1}). \end{array}$$

A.11 Proof of Lemma 3

Part (i). We combine the monetary wedge defined in (38) with the implementability condition (26). Next note that aggregate labor and aggregate output satisfy (100) and (101).respectively. Substituting the expression for $L(s^t)$ from (100) into (104) we get:

$$\kappa \frac{y^s(s^t)}{Y(s^t)} \left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho} + (1-\kappa)\frac{y^f(s^t)}{Y(s^t)} \left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho} = 1 - \tau_M^*(s^t)$$

which we may rewrite as:

$$\kappa \left[\frac{y^s(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}} \left[\frac{y^s(s^t)}{y^f(s^t)}\right]^{1/\rho} + (1-\kappa) \left[\frac{y^f(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}} = 1 - \tau_M^*(s^t).$$

Next we combine this with (101) and obtain:

$$\kappa \left[\frac{y^s(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}} \left[\frac{y^s(s^t)}{y^f(s^t)}\right]^{1/\rho} + 1 - \kappa \left[\frac{y^s(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}} = 1 - \tau_M^*(s^t)$$

It follows that:

$$\tau_M^*(s^t) = \kappa \left[\frac{y^s(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}} \left\{1 - \left[\frac{y^s(s^t)}{y^f(s^t)}\right]^{1/\rho}\right\}$$

Therefore

$$\text{sign}(\tau_M^*(s^t)) = \text{sign}\left\{1 - \left[\frac{y^s(s^t)}{y^f(s^t)}\right]^{1/\rho}\right\}$$

It follows that:

$$\begin{array}{lll} y^f(s^t) > y^s(s^t) & \text{if and only if} & \tau_M^*(s^t) > 0, \\ y^f(s^t) = y^s(s^t) & \text{if and only if} & \tau_M^*(s^t) = 0, \\ y^f(s^t) < y^s(s^t) & \text{if and only if} & \tau_M^*(s^t) < 0. \end{array}$$

Part (ii). Combining the planner optimality conditions (77) and (79), we obtain the following condition which must hold at the planner's optimum:

$$\frac{\varsigma^Y(s^t) + \chi U_C^m(s^t)\upsilon(s^{t-1})\left\{\frac{\rho-1}{\rho} + \frac{U_L^m(s^t)}{\chi U_C^m(s^t)}\frac{1}{A(s_t)}\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{1/\rho}\right\}}{\varsigma^Y(s^t) + \chi U_C^m(s^t)\xi(s^t)\left\{\frac{\rho-1}{\rho} + \frac{U_L^m(s^t)}{\chi U_C^m(s^t)}\frac{1}{A(s_t)}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1/\rho}\right\}} = \left[\frac{y^s(s^t)}{y^f(s^t)}\right]^{1/\rho}$$

Next, using the implementability condition (26), we rewrite the above equation as follows:

$$\frac{\varsigma^Y(s^t) + \chi U_C^m(s^t)\upsilon(s^{t-1})\left\{\frac{\rho-1}{\rho} - \left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho}\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{1/\rho}\right\}}{\varsigma^Y(s^t) + \chi U_C^m(s^t)\xi(s^t)\left\{\frac{\rho-1}{\rho} - 1\right\}} = \left[\frac{y^s(s^t)}{y^f(s^t)}\right]^{1/\rho}$$

which reduces to:

$$\frac{\frac{\varsigma^Y(s^t)}{\chi U_C^m(s^t)} + \upsilon(s^{t-1})\left\{\frac{\rho-1}{\rho} - \left[\frac{y^s(s^t)}{y^f(s^t)}\right]^{1/\rho}\right\}}{\frac{\varsigma^Y(s^t)}{\chi U_C^m(s^t)} + \xi(s^t)\left\{\frac{\rho-1}{\rho} - 1\right\}} = \left[\frac{y^s(s^t)}{y^f(s^t)}\right]^{1/\rho}$$

Rearranging, we obtain the following equilibrium condition:

$$\upsilon(s^{t-1})\left\{\frac{\rho-1}{\rho} - \left[\frac{y^s(s^t)}{y^f(s^t)}\right]^{1/\rho}\right\} - \xi(s^t)\left\{\frac{\rho-1}{\rho} - 1\right\}\left[\frac{y^s(s^t)}{y^f(s^t)}\right]^{1/\rho} = \frac{\varsigma^Y(s^t)}{\chi U_C^m(s^t)}\left\{\left[\frac{y^s(s^t)}{y^f(s^t)}\right]^{1/\rho} - 1\right\} \tag{106}$$

We consider the case in which $y^s(s^t) = y^f(s^t)$. It is clear from (106) that if $y^s(s^t) = y^f(s^t)$, then

$$\upsilon(s^{t-1})\left\{\frac{\rho-1}{\rho} - 1\right\} - \xi(s^t)\left\{\frac{\rho-1}{\rho} - 1\right\} = 0.$$

It follows that $\xi(s^t) = \upsilon(s^{t-1})$.

A.12 Proof of Proposition 5

The optimal monetary wedge is defined in (38). We combine this with the household's intratemporal condition in (13) and obtain the following condition:

$$\frac{W(s^t)}{P(s^t)} = (1-\tau_r)\left(\frac{\rho-1}{\rho}\right)(1-\tau_M^*(s^t))\frac{Y(s^t)}{L(s^t)}$$

To simplify, we consider the fiscal implementation that sets $(1-\tau_r)\left(\frac{\rho-1}{\rho}\right) = 1$. This implies:

$$\frac{W(s^t)}{P(s^t)} = (1-\tau_M^*(s^t))\frac{Y(s^t)}{L(s^t)}$$

We combine the above expression with the implementability condition (104), and infer that the price level satisfies:

$$P(s^t) = \left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1/\rho} \frac{W(s^t)}{A(s_t)}.$$

Therefore the optimal markup satisfies:

$$\log \mathcal{M}(s^t) = \frac{1}{\rho}(\log y^f(s^t) - \log Y(s^t))$$

with $\rho > 1$. Therefore

$$\begin{array}{lll} \log \mathcal{M}(s^t) > 0 & \text{if and only if} & y^f(s^t) > y^s(s^t), \\ \log \mathcal{M}(s^t) = 0 & \text{if and only if} & y^f(s^t) = y^s(s^t), \\ \log \mathcal{M}(s^t) < 0 & \text{if and only if} & y^f(s^t) < y^s(s^t). \end{array}$$

Combining this with our characterization of the optimal monetary tax in Lemma 3, we obtain:

$$\begin{array}{lll} \log \mathcal{M}(s^t) > 0 & \text{if and only if} & \tau_M^*(s^t) > 0, \\ \log \mathcal{M}(s^t) = 0 & \text{if and only if} & \tau_M^*(s^t) = 0, \\ \log \mathcal{M}(s^t) < 0 & \text{if and only if} & \tau_M^*(s^t) < 0. \end{array}$$

The result stated in Proposition 5 follows by combining the above with Theorem 2.

A.13 Proof of Proposition 6

Next, recall from the sufficiency argument of Proposition 2 that for any sticky-price allocation, we can construct nominal prices as follows. Without loss of generality we set:

$$p_t^s(s^{t-1}) = \mathcal{B}_t(s^{t-1}).$$

This implies, in terms of our decomposition of $y^s(s^t) = \phi^s(s^{t-1})\Phi(s^t)$, that

$$\phi^s(s^{t-1})^{-1/\rho} = \mathcal{B}_t(s^{t-1}),$$

Therefore:

$$\phi^s(s^{t-1}) = \mathcal{B}_t(s^{t-1})^{-\rho} \qquad \text{and} \qquad y^s(s^t) = \mathcal{B}_t(s^{t-1})^{-\rho}\Phi(s^t).$$

which implies that $\Phi(s^t) = y^s(s^t)/\mathcal{P}(s^{t-1})^{-\rho}$. The aggregate price level thereby satisfies:

$$P(s^t) = \left[\frac{Y(s^t)}{\Phi(s^t)}\right]^{-1/\rho} = \left[\frac{Y(s^t)}{y^s(s^t)}\mathcal{B}_t(s^{t-1})^{-\rho}\right]^{-1/\rho} = \left[\frac{Y(s^t)}{y^s(s^t)}\right]^{-1/\rho} \mathcal{B}_t(s^{t-1}).$$

Therefore the deviation of the price level from the expected price satisfies:

$$\log P(s^t) - \log \mathcal{B}_t(s^{t-1}) = -\frac{1}{\rho}(\log Y(s^t) - \log y^s(s^t)) \tag{107}$$

with $\rho > 1$. Therefore

$$\begin{array}{lll} P(s^t) < \mathcal{B}_t(s^{t-1}) & \text{if and only if} & y^f(s^t) > y^s(s^t), \\ P(s^t) = \mathcal{B}_t(s^{t-1}) & \text{if and only if} & y^f(s^t) = y^s(s^t), \\ P(s^t) > \mathcal{B}_t(s^{t-1}) & \text{if and only if} & y^f(s^t) < y^s(s^t). \end{array}$$

Combining this with our characterization of the optimal monetary tax in Lemma 3, we find:

$$\begin{array}{lll} P(s^t) < \mathcal{B}_t(s^{t-1}) & \text{if and only if} & \tau_M^*(s^t) > 0, \\ P(s^t) = \mathcal{B}_t(s^{t-1}) & \text{if and only if} & \tau_M^*(s^t) = 0, \\ P(s^t) > \mathcal{B}_t(s^{t-1}) & \text{if and only if} & \tau_M^*(s^t) < 0. \end{array}$$

Combining this with our result in Theorem 2, the result about the aggregate price level stated in Proposition 5 then follows.

Next we turn to the nominal interest rate. The nominal interest rate satisfies the Euler equation in 15:

$$\frac{C(s^t)^{-\gamma}}{P(s^t)} = \beta(1+i(s^t)) \sum_{s^{t+1}|s^t} \mu(s^{t+1}|s^t) \frac{U_C^m(s^{t+1})}{P(s^{t+1})}.$$

Let $\hat{C}(s^t) = \hat{Y}(s^t)$ denote the flexible-price level of output. The natural, flexible-price, interest rate satisfies:

$$\frac{\hat{C}(s^t)^{-\gamma}}{\mathcal{B}_t(s^{t-1})} = \beta(1+\hat{\imath}(s^t)) \sum_{s^{t+1}|s^t} \mu(s^{t+1}|s^t) \frac{U_C^m(s^{t+1})}{P(s^{t+1})}.$$

Therefore

$$\frac{1+i(s^t)}{1+\hat{\imath}(s^t)} = \frac{C(s^t)^{-\gamma}/P(s^t)}{\hat{C}(s^t)^{-\gamma}/\mathcal{B}_t(s^{t-1})}$$

In logs:

$$\log\left[\frac{1+i(s^t)}{1+\hat{\imath}(s^t)}\right] = -\gamma[\log Y(s^t) - \log \hat{Y}(s^t)] - [\log P(s^t) - \log \mathcal{B}_t(s^{t-1})]$$

Next we substitute the price level from 107 into the above expression. Doing so, we obtain the following expression:

$$\log(1+i(s^t)) - \log(1+\hat{\imath}(s^t)) = \frac{1}{\rho}(\log Y(s^t) - \log y^s(s^t)) - \gamma(\log Y(s^t) - \log \hat{Y}(s^t)).$$

First, from our characterization of the optimal monetary tax in Lemma 3, we have that:

$$\begin{array}{lll} \log Y(s^t) > \log y^s(s^t) & \text{if and only if} & \tau_M^*(s^t) > 0, \\ \log Y(s^t) = \log y^s(s^t) & \text{if and only if} & \tau_M^*(s^t) = 0, \\ \log Y(s^t) < \log y^s(s^t) & \text{if and only if} & \tau_M^*(s^t) < 0. \end{array}$$

The next step in our proof requires the following lemma:

Lemma 4. *Let $\hat{Y}(s^t)$ denote the level of output in history s^t under flexible prices. Then*

$$\begin{array}{lll} \log Y(s^t) < \log \hat{Y}(s^t) & \textit{if and only if} & \tau_M^*(s^t) > 0, \\ \log Y(s^t) = \log \hat{Y}(s^t) & \textit{if and only if} & \tau_M^*(s^t) = 0, \\ \log Y(s^t) > \log \hat{Y}(s^t) & \textit{if and only if} & \tau_M^*(s^t) < 0. \end{array}$$

Proof. See Section A.15 of this appendix. □

Therefore, using Lemmas 3 and 4, it follows that:

$$\begin{array}{lll} i(s^t) > \hat{\imath}(s^t) & \text{if and only if} & \tau_M^*(s^t) > 0, \\ i(s^t) = \hat{\imath}(s^t) & \text{if and only if} & \tau_M^*(s^t) = 0, \\ i(s^t) < \hat{\imath}(s^t) & \text{if and only if} & \tau_M^*(s^t) < 0. \end{array}$$

The result stated in Proposition 6 follows by combining the above with Theorem 2.

A.14 Equivalent Equilibrium Representation

In this section of the appendix we provide an equivalent characterization of the equilibrium using the forecast errors $\epsilon(s^t)$ defined in (29). Such a representation gives rise to a few auxiliary results, Lemmas (5), (6), and (7), that we use in later proofs.

Note that with $\epsilon(s^t)$ so defined, equation (28) can be rewritten as $p_t^s(s^{t-1}) = \epsilon(s^t) p_t^f(s^t)$. The CES demand function in equation (4) implies that relative quantities across the two types of firms must satisfy:

$$\frac{y^s(s^t)}{y^f(s^t)} = \left(\frac{p_t^s(s^{t-1})}{p_t^f(s^t)}\right)^{-\rho}.$$

Therefore:

$$y^s(s^t) = \epsilon(s^t)^{-\rho} y^f(s^t) \tag{108}$$

Lemma 5. *Given $\epsilon(s^t)$, aggregate output and labor satisfy:*

$$Y(s^t) = A(s_t)\Delta(\epsilon(s^t))L(s^t) \tag{109}$$

where $\Delta : \mathbb{R}_+ \to \mathbb{R}_+$ is a function defined by:

$$\Delta(\epsilon) \equiv \left\{ \frac{\left[\kappa\epsilon^{1-\rho} + (1-\kappa)\right]^{-\frac{1}{1-\rho}}}{\left[\kappa\epsilon^{-\rho} + (1-\kappa)\right]^{1/\rho}} \right\}^{\rho} > 0. \tag{110}$$

The function $\Delta : \mathbb{R}_+ \to \mathbb{R}_+$ is continuous, differentiable, strictly concave, and satisfies $\max_{\epsilon>0} \Delta(\epsilon) = 1$. Furthermore, it attains its unique maximum at $\epsilon = 1$.

Proof. From the individual firm production functions:

$$n^s(s^t) = \frac{y^s(s^t)}{A(s_t)}, \qquad \text{and} \qquad n^f(s^t) = \frac{y^f(s^t)}{A(s_t)}.$$

Using (108), we have that aggregate output and aggregate labor satisfy:

$$Y(s^t) = \left[\kappa y^s(s^t)^{\frac{\rho-1}{\rho}} + (1-\kappa) y^f(s^t)^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}} = \left[\kappa\epsilon(s^t)^{-(\rho-1)} y^f(s^t)^{\frac{\rho-1}{\rho}} + (1-\kappa) y^f(s^t)^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}}$$

and

$$L(s^t) = \kappa \frac{y^s(s^t)}{A(s_t)} + (1-\kappa)\frac{y^f(s^t)}{A(s_t)} = \kappa \frac{\epsilon(s^t)^{-\rho} y^f(s^t)}{A(s_t)} + (1-\kappa)\frac{y^f(s^t)}{A(s_t)},$$

respectively. Therefore

$$Y(s^t) = y^f(s^t)\left[\kappa\epsilon(s^t)^{-(\rho-1)} + (1-\kappa)\right]^{\frac{\rho}{\rho-1}} \qquad \text{and} \qquad L(s^t) = \frac{y^f(s^t)}{A(s_t)}\left[\kappa\epsilon(s^t)^{-\rho} + (1-\kappa)\right]$$

Taking the ratio of aggregate output to aggregate labor, we get:

$$\frac{Y(s^t)}{L(s^t)} = \frac{y^f(s^t)\left[\kappa\epsilon(s^t)^{-(\rho-1)} + (1-\kappa)\right]^{\frac{\rho}{\rho-1}}}{\frac{y^f(s^t)}{A(s_t)}\left[\kappa\epsilon(s^t)^{-\rho} + (1-\kappa)\right]} = A(s_t)\frac{\left[\kappa\epsilon(s^t)^{-(\rho-1)} + (1-\kappa)\right]^{\frac{\rho}{\rho-1}}}{\left[\kappa\epsilon(s^t)^{-\rho} + (1-\kappa)\right]}$$

It follows that the aggregate production function can be expressed as (109) with

$$\Delta(\epsilon) = \frac{\left[\kappa\epsilon^{-(\rho-1)} + (1-\kappa)\right]^{\frac{\rho}{\rho-1}}}{\left[\kappa\epsilon^{-\rho} + (1-\kappa)\right]} = \left\{\frac{\left[\kappa\epsilon^{1-\rho} + (1-\kappa)\right]^{-\frac{1}{1-\rho}}}{\left[\kappa\epsilon^{-\rho} + (1-\kappa)\right]^{1/\rho}}\right\}^{\rho}.$$

Next, note that $\Delta(\epsilon)$ is a continuous and differentiable function. The first derivative of $\Delta(\epsilon)$ with respect to ϵ is given by:

$$\frac{d\Delta(\epsilon)}{d\epsilon} = \rho\Delta(\epsilon)^{1-\frac{1}{\rho}}\frac{d}{d\epsilon}\left\{\frac{\left[\kappa\epsilon^{-(\rho-1)} + (1-\kappa)\right]^{\frac{1}{\rho-1}}}{\left[\kappa\epsilon^{-\rho} + (1-\kappa)\right]^{1/\rho}}\right\},$$

where the last term satisfies:

$$\frac{d}{d\epsilon}\left\{\frac{\left[\kappa\epsilon^{-(\rho-1)} + (1-\kappa)\right]^{\frac{1}{\rho-1}}}{\left[\kappa\epsilon^{-\rho} + (1-\kappa)\right]^{1/\rho}}\right\} = \kappa\Delta(\epsilon)^{\frac{1}{\rho}}\epsilon^{-\rho-1}\left\{\left[\kappa\epsilon^{-\rho} + (1-\kappa)\right]^{-1} - \left[\kappa\epsilon^{-\rho+1} + (1-\kappa)\right]^{-1}\epsilon\right\}.$$

Therefore:

$$\frac{d\Delta(\epsilon)}{d\epsilon} = \kappa\rho\Delta(\epsilon)\epsilon^{-\rho-1}\left\{\left[\kappa\epsilon^{-\rho} + (1-\kappa)\right]^{-1} - \left[\kappa\epsilon^{-\rho+1} + (1-\kappa)\right]^{-1}\epsilon\right\} \tag{111}$$

To obtain a maxima or minima, we set the first derivative equal to zero as follows:

$$\Delta(\epsilon)\epsilon^{-\rho-1}\left\{\left[\kappa\epsilon^{-\rho} + (1-\kappa)\right]^{-1} - \left[\kappa\epsilon^{-\rho+1} + (1-\kappa)\right]^{-1}\epsilon\right\} = 0.$$

Noting that both $\Delta(\epsilon)$ and $\epsilon^{-\rho-1}$ are strictly positive, this implies:

$$\left[\kappa\epsilon^{-\rho} + (1-\kappa)\right]^{-1} - \left[\kappa\epsilon^{-\rho+1} + (1-\kappa)\right]^{-1}\epsilon = 0.$$

Solving this for ϵ, we obtain a unique solution of $\epsilon = 1$. Furthermore, note that from (111), $d\Delta(\epsilon)/d\epsilon > 0$ if and only if $\epsilon < 1$. Finally, we evaluate the second derivative of $\Delta(\epsilon)$ at $\epsilon = 1$, and find that it is unambiguously negative:

$$\Delta''(1) = -\rho\kappa(1-\kappa) < 0$$

We conclude that the function $\Delta(\epsilon)$ attains a global maximum at $\epsilon = 1$. The function $\Delta(\epsilon)$ is strictly increasing in ϵ when $\epsilon < 1$ and is strictly decreasing in ϵ when $\epsilon > 1$. Finally, the maximal value of this function is given by:

$$\max_{\epsilon>0} \Delta(\epsilon) = \Delta(1) \equiv \left\{ \frac{[\kappa + (1-\kappa)]^{\frac{1}{\rho-1}}}{[\kappa + (1-\kappa)]^{1/\rho}} \right\}^{\rho} = 1$$

as was to be shown. □

When monetary policy implements flexible-price allocations—that is, when it sets $\epsilon = 1$ in all states—then $\Delta(\epsilon)$ attains its unique maximum of 1. In this case, there is no misallocation across firms and therefore no loss in production efficiency. On the other hand, when monetary policy deviates from implementing flexible-price allocations—that is, when it sets $\epsilon \neq 1$ in some or all states—then in those states, $\Delta(\epsilon)$ is strictly below 1. In this case, the "active" use of monetary policy leads to forecast errors of the sticky-price firms. Dispersion of prices across sticky- and flexible-price firms implies misallocation of inputs and results in an efficiency wedge, or TFP loss. The term $\Delta(\epsilon)$ represents this efficiency wedge.

Lemma (5) provides a succinct characterization of the efficiency wedge in this economy. The following lemma provides a similar result for the labor wedge.

Lemma 6. *For a given $\epsilon(s^t)$, aggregate output and labor joint satisfy:*

$$-\frac{U_L^m(s^t)}{U_C^m(s^t)} = \chi\Gamma(\epsilon(s^t))A(s_t) \tag{112}$$

where $\Gamma : \mathbb{R}_+ \to \mathbb{R}_+$ is a function defined by:

$$\Gamma(\epsilon) \equiv \left[\kappa\epsilon^{1-\rho} + (1-\kappa)\right]^{\frac{1}{\rho-1}} > 0. \tag{113}$$

The function $\Gamma : \mathbb{R}_+ \to \mathbb{R}_+$ is continuous, differentiable, and satisfies the following two properties (i) $\Gamma(1) = 1$ and (ii) $\Gamma'(\epsilon) < 0$ for all $\epsilon > 0$. It follows that:

$$\begin{array}{lcr} \Gamma(\epsilon) < 1 & \textit{if and only if} & \epsilon > 1, \\ \Gamma(\epsilon) = 1 & \textit{if and only if} & \epsilon = 1, \\ \Gamma(\epsilon) > 1 & \textit{if and only if} & \epsilon \in (0,1). \end{array}$$

Proof. The aggregate price level satisfies:

$$P(s^t) = \left[\kappa p_t^s(s^{t-1})^{1-\rho} + (1-\kappa)p_t^f(s^t)^{1-\rho}\right]^{\frac{1}{1-\rho}}.$$

Substituting in for the firms' optimal prices, we obtain:

$$P(s^t) = \left[\kappa\epsilon(s^t)^{1-\rho} + (1-\kappa)\right]^{\frac{1}{1-\rho}} \left[(1-\tau_r)\left(\frac{\rho-1}{\rho}\right)\right]^{-1} \frac{W(s^t)}{A(s_t)}$$

Combining the above equation with the household's intratemporal condition (13), we get:

$$-\frac{U_L^m(s^t)}{U_C^m(s^t)}\left[\kappa\epsilon(s^t)^{1-\rho}+(1-\kappa)\right]^{\frac{1}{1-\rho}}=\chi A(s_t)$$

It follows that in equilibrium, aggregate output and aggregate labor jointly satisfy (112) with

$$\Gamma(\epsilon)\equiv\left[\kappa\epsilon(s^t)^{1-\rho}+(1-\kappa)\right]^{-\frac{1}{1-\rho}}>0.$$

Next, note that $\Gamma(\epsilon)$ is a continuous and differentiable function. Furthermore, $\Gamma(1) = [\kappa+(1-\kappa)]^{\frac{1}{\rho-1}}=1$. Finally, the first derivative is given by:

$$\frac{d\Gamma(\epsilon)}{d\epsilon}\equiv-\kappa\left[\kappa\epsilon^{-(\rho-1)}+(1-\kappa)\right]^{\frac{1}{\rho-1}-1}\epsilon^{-\rho}$$

Therefore $d\Gamma(\epsilon)/d\epsilon<0$ for all $\epsilon>0$. □

Finally, we relate the forecast error to the monetary tax $\tau_M(s^t)$ defined in ().

Lemma 7. *The monetary tax satisfies:*

$$\tau_M(\epsilon)=1-\frac{\Gamma(\epsilon)}{\Delta(\epsilon)}=\frac{\kappa\epsilon^{-\rho}(\epsilon-1)}{\kappa\epsilon^{-(\rho-1)}+(1-\kappa)}.$$

It follows that:

$$\begin{array}{lcl}\tau_M(\epsilon)>0 & \textit{if and only if} & \epsilon>1,\\ \tau_M(\epsilon)=0 & \textit{if and only if} & \epsilon=1,\\ \tau_M(\epsilon)<0 & \textit{if and only if} & \epsilon\in(0,1).\end{array}$$

Proof. Recall that the monetary tax satisfies:

$$-\frac{U_L^m(s^t)}{U_C^m(s^t)}=\chi(1-\tau_M(s^t))\frac{Y(s^t)}{L(s^t)}$$

Combining this with (112) we get:

$$(1-\tau_M(s^t))\frac{Y(s^t)}{L(s^t)}=\Gamma(\epsilon(s^t))A(s_t)$$

Combining this with (109) we infer that in equilibrium the monetary tax satisfies:

$$1-\tau_M(s^t)=\frac{\Gamma(\epsilon(s^t))}{\Delta(\epsilon(s^t))}$$

Substituting in for the functions Γ and Δ, we get:

$$1-\tau_M(\epsilon)=\frac{\left[\kappa\epsilon^{-(\rho-1)}+(1-\kappa)\right]^{\frac{1}{\rho-1}}\left[\kappa\epsilon^{-\rho}+(1-\kappa)\right]}{\left[\kappa\epsilon^{-(\rho-1)}+(1-\kappa)\right]^{\frac{\rho}{\rho-1}}},$$

which implies:

$$1-\tau_M(\epsilon)=\frac{\kappa\epsilon^{-\rho}+(1-\kappa)}{\kappa\epsilon^{-(\rho-1)}+(1-\kappa)}. \tag{114}$$

Solving this for $\tau_M(\epsilon)$, we get:

$$\tau_M(\epsilon)=\frac{\kappa\epsilon^{-\rho}(\epsilon-1)}{\kappa\epsilon^{-(\rho-1)}+(1-\kappa)}$$

With this expression we can prove the last part of Lemma 7. Note that the denominator is strictly positive for all $\epsilon>0$. Furthermore $\kappa\epsilon^{-\rho}>0$ for all $\epsilon>0$. Therefore: sign$(\tau_M(\epsilon))=$ sign$(\epsilon-1)$. The rest then follows. □

A.15 Proof of Lemma 4

First, we solve for the natural level of output. Let $\hat{Y}(s^t)$ and $\hat{L}(s^t)$ denote the flexible-price level of output and employment, respectively. These jointly satisfy:

$$\frac{\hat{L}(s^t)^{\eta}}{\hat{Y}(s^t)^{-\gamma}}=\chi A(s_t) \qquad \text{and} \qquad \frac{\hat{Y}(s^t)}{\hat{L}(s^t)}=A(s_t)$$

This is two equations in two unknowns. Solving for $\hat{Y}(s^t)$, we obtain the following expression for the flex-price level of output:

$$\hat{Y}(s^t)^{\eta+\gamma}=\chi A(s_t)^{1+\eta}.$$

Next, we solve for the realized level of output. Using the $\epsilon(s^t)$ notation articulated in Section of the appendix, realized output $Y(s^t)$ and employment $L(s^t)$ jointly satisfy:

$$\frac{L(s^t)^{\eta}}{Y(s^t)^{-\gamma}}=\chi\Gamma(\epsilon(s^t))A(s_t) \qquad \text{and} \qquad Y(s^t)=A(s_t)\Delta(\epsilon(s^t))L(s^t).$$

This is two equations in two unknowns. Solving for $Y(s^t)$, we obtain the following expression for realized output:

$$Y(s^t)^{\eta+\gamma}=\chi A(s_t)^{1+\eta}\Gamma(\epsilon(s^t))\Delta(\epsilon(s^t))^{\eta}.$$

Combining this with the flexible-price level of output we get:

$$\frac{Y(s^t)^{\eta+\gamma}}{\hat{Y}(s^t)^{\eta+\gamma}}=\Gamma(\epsilon(s^t))\Delta(\epsilon(s^t))^{\eta}.$$

In logs:

$$\log Y(s^t)-\log\hat{Y}(s^t)=\frac{1}{\eta+\gamma}\log\Gamma(\epsilon(s^t))+\frac{\eta}{\eta+\gamma}\log\Delta(\epsilon(s^t)).$$

First, recall from Lemmas (5) and (6) that $\Gamma(1)=1$ and $\Delta(1)=1$. It follows that if $\epsilon(s^t)=1$, then $Y(s^t)=\hat{Y}(s^t)$.

Second, note that, to a first order around $\epsilon(s^t) = 1$,

$$\log \Delta(\epsilon(s^t)) \approx 0.$$

To see this, note that:

$$\log \Delta(\epsilon(s^t)) \approx \log \Delta(\epsilon(s^t))\big|_{\epsilon=1} + \left.\frac{d \log \Delta(\epsilon)}{d\epsilon}\right|_{\epsilon=1} \epsilon(s^t) = 0$$

since:

$$\Delta(1) = 1, \qquad \frac{d \log \Delta(\epsilon)}{d\epsilon} = \frac{1}{\Delta(\epsilon)} \frac{d\Delta(\epsilon)}{d\epsilon}, \qquad \text{and} \qquad \left.\frac{d\Delta(\epsilon)}{d\epsilon}\right|_{\epsilon=1} = 0.$$

This implies that for small shocks around $\epsilon(s^t) = 1$,

$$\log Y(s^t) - \log \hat{Y}(s^t) \approx \frac{1}{\eta + \gamma} \log \Gamma(\epsilon(s^t)) \tag{115}$$

From Lemma (6), we have that the function $\Gamma : \mathbb{R}_+ \to \mathbb{R}_+$ satisfies the following:

$$\begin{array}{lcr} \log \Gamma(\epsilon) < 0 & \text{if and only if} & \epsilon > 1, \\ \log \Gamma(\epsilon) = 0 & \text{if and only if} & \epsilon = 1, \\ \log \Gamma(\epsilon) > 0 & \text{if and only if} & \epsilon \in (0, 1). \end{array}$$

Finally, using this in equation (115) and the result of Lemma (7), it follows that:

$$\begin{array}{lcl} \log Y(s^t) < \log Y^n(s^t) & \text{if and only if} & \tau_M^*(s^t) > 0, \\ \log Y(s^t) = \log Y^n(s^t) & \text{if and only if} & \tau_M^*(s^t) = 0, \\ \log Y(s^t) > \log Y^n(s^t) & \text{if and only if} & \tau_M^*(s^t) < 0. \end{array}$$

B Appendix: One-Period-Ahead Tax Rates

In this section of the appendix, we characterize the economy with one-period-ahead tax rates. We first state some auxiliary results, followed by their proofs. We begin with our characterization of the set of sticky price allocations, $\mathcal{X}^s$.

Proposition 9. *A feasible allocation $x \in \mathcal{X}$ can be implemented as a sticky-price equilibrium with one-period-ahead taxes if and only if there exist market weights $\varphi \equiv (\varphi^i)$ and a scalar $\bar{T} \in \mathbb{R}$, such that the following three sets of conditions are satisfied:*

(i) $y^j(s^t) = y^f(s^t)$ for all $j \in \mathcal{J}^f$, and $y^j(s^t) = y^s(s^t)$ for all $j \in \mathcal{J}^s$, for all $s^t \in S^t$;

(ii) for all $s^{t-1} \in S^{t-1}$,

$$\left[\frac{y^f(s|s^{t-1})}{Y(s|s^{t-1})}\right]^{-1/\rho} \frac{A(s)U_C^m(s|s^{t-1})}{-U_L^m(s|s^{t-1})} = \left[\frac{y^f(s'|s^{t-1})}{Y(s'|s^{t-1})}\right]^{-1/\rho} \frac{A(s')U_C^m(s'|s^{t-1})}{-U_L^m(s'|s^{t-1})}, \qquad \forall s, s'|s^{t-1}; \tag{116}$$

(iii) condition (18) holds for every $i \in I$.

Proof. See Appendix B.1. □

Proposition 9 characterizes the set $\mathcal{X}^s$ when tax rates can be set one period in advance. Note that part (ii) of the proposition, namely, the conditions stated in (116), are equivalent to the following statement: for all $s^{t-1} \in S^{t-1}$, there exists a positive scalar $\chi(s^{t-1}) \in \mathbb{R}_+$ such that:

$$\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho} \frac{A(s_t)U_C^m(s^t)}{-U_L^m(s^t)} = \frac{1}{\chi(s^{t-1})}, \qquad \forall s^t|s^{t-1}. \tag{117}$$

This allows us to state the Ramsey planner's problem as follows.

Ramsey Planner's Problem. *The Ramsey planner chooses an allocation,*

$$x \equiv \{y^s(s^t), y^f(s^t), C(s^t), Y(s^t), L(s^t)\}_{t\geq 0, s^t \in S^t},$$

market weights $\varphi \equiv (\varphi^i)$, and scalar $\bar{T} \in \mathbb{R}$, in order to maximize (32)*, subject to*

$$C(s^t) = Y(s^t) = \left[\kappa y^s(s^t)^{\frac{\rho-1}{\rho}} + (1-\kappa)y^f(s^t)^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}}, \qquad L(s^t) = \kappa\frac{y^s(s^t)}{A(s_t)} + (1-\kappa)\frac{y^f(s^t)}{A(s_t)}, \tag{118}$$

and (117).

We let $\beta^t \mu(s^t)(1-\kappa)\xi(s^t)$ denote the Lagrange multiplier on the implementability condition (117). The Ramsey optimum can be characterized as follows.

Proposition 10. *A Ramsey optimum x^* satisfies*

$$-\frac{\mathcal{W}_L(s^t) + \xi(s^t)U_{LL}^m(s^t)\frac{1}{A(s_t)}}{\mathcal{W}_C(s^t) + \xi(s^t)\chi(s^{t-1})U_{CC}^m(s^t)\left(\frac{y^f(s^t)}{Y(s^t)}\right)^{-1/\rho}} = \frac{Y(s^t)}{L(s^t)}, \qquad \forall s^t \in S^t. \tag{119}$$

Proof. See Appendix B.2. □

This result is the counterpart to Proposition 8 for the economy with one-period ahead tax rates. Finally, we use this result to characterize the optimal monetary wedge in this economy; this characterization is presented in Theorem 3 in the main text and its proof is in Appendix B.3.

B.1 Proof of Proposition 9

Necessity. The necessity argument follows similar steps as the proof of Proposition 2 [Section A.6 of the Appendix]. In particular, we combine the flexible-price firm's optimality condition with the CES demand function (68) and the household's intratemporal optimality condition (13) and obtain the following equilibrium condition:

$$\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho} + \left(\frac{\rho-1}{\rho}\right)^{-1}\left[\frac{1+\tau_c(s^{t-1})}{(1-\tau_r(s^{t-1}))(1-\tau_\ell(s^{t-1}))}\right]\frac{U_L^m(s^t)}{U_C^m(s^t)}\frac{1}{A(s_t)} = 0.$$

We let

$$\chi(s^{t-1}) \equiv \left(\frac{\rho-1}{\rho}\right) \frac{(1-\tau_\ell(s^{t-1}))(1-\tau_r(s^{t-1}))}{1+\tau_c(s^{t-1})}.$$

denote the wedge due to the mark-up and taxes. Therefore, the following condition:

$$\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho} + \frac{1}{\chi(s^{t-1})}\frac{U_L^m(s^t)}{U_C^m(s^t)}\frac{1}{A(s_t)} = 0.$$

is a necessary condition for an allocation to be supported in equilibrium. Note that the above is equivalent to the conditions stated in (116).

We can similarly combine the sticky-price firm optimality condition with the CES demand function (68) and the household's intratemporal optimality condition (13) and obtain the following equilibrium condition:

$$\sum_{s^t|s^{t-1}} U_C^m(s^t) y^s(s^t) \left\{ \left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho} + \frac{U_L^m(s^t)}{U_C^m(s^t)}\frac{1}{\chi(s^{t-1})A(s_t)} \right\} \mu(s^t|s^{t-1}) = 0; \tag{120}$$

for all $s^{t-1} \in S^{t-1}$. Therefore (120) is a necessary condition for an allocation to be supported in equilibrium. The remainder of the proof of necessity follows the same steps as in the proof of Proposition 2.

Sufficiency. Take any feasible allocation $x \in \mathcal{X}$, vector $\varphi \equiv (\varphi^i)$, and scalar $\bar{T} \in \mathbb{R}$, that satisfy conditions (i)-(iii) of Proposition 9. We show that there exists a price system ϱ, a policy Ω, and asset holdings ζ, that support x as a sticky-price equilibrium; we construct these as follows.

First, we set prices such that: $p_t^j(s^t) = p_t^f(s^t)$ for all $j \in \mathcal{J}^f$ and $p_t^j(s^t) = p_t^s(s^{t-1})$ for all $j \in \mathcal{J}^s$. That is, all flexible-price firms set the same price and all sticky-price firms set the same price. Given the allocation, we set $p_t^f(s^t)/P(s^t)$ and $p_t^s(s^{t-1})/P(s^t)$ such that the CES demand curves are satisfied: (68).

Next, note that part (ii) of the proposition, namely, the conditions stated in (116), are equivalent to the following statement: for all $s^{t-1} \in S^{t-1}$, there exists a positive scalar $\chi(s^{t-1}) \in \mathbb{R}_+$ such that:

$$\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho} + \frac{1}{\chi(s^{t-1})}\frac{U_L^m(s^t)}{U_C^m(s^t)}\frac{1}{A(s_t)} = 0. \tag{121}$$

Note that the conditions in (116) imply that such a constant exists but is not unique. In fact, for any $s^{t-1} \in S^{t-1}$, we can choose $\chi(s^{t-1})$ freely, provided it remain strictly positive. In particular, we set $\chi(s^{t-1})$ as follows:

$$\chi(s^{t-1}) = -\frac{\sum_{s^t|s^{t-1}} y^s(s^t) U_L^m(s^t) \frac{1}{A(s_t)} \mu(s^t|s^{t-1})}{\sum_{s^t|s^{t-1}} y^s(s^t) U_C^m(s^t) \left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho} \mu(s^t|s^{t-1})} > 0. \tag{122}$$

Next, we set tax rates $\{\tau_\ell(s^{t-1}), \tau_c(s^{t-1}), \tau_r(s^{t-1})\}$ such that they jointly satisfy:

$$\frac{(1-\tau_\ell(s^{t-1}))(1-\tau_r(s^{t-1}))}{1+\tau_c(s^{t-1})} = \left(\frac{\rho-1}{\rho}\right)^{-1} \chi(s^{t-1}). \tag{123}$$

For any strictly positive $\chi(s^{t-1})$ and $\rho > 1$, such tax rates exist.

Combining (123) with condition (121), we obtain:

$$\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho} + \left(\frac{\rho-1}{\rho}\right)^{-1} \left[\frac{1+\tau_c(s^{t-1})}{(1-\tau_r(s^{t-1}))(1-\tau_\ell(s^{t-1}))}\right] \frac{U_L^m(s^t)}{U_C^m(s^t)} \frac{1}{A(s_t)} = 0, \tag{124}$$

Furthermore, combining (123) with condition (122) and rearranging, we obtain:

$$\sum_{s^t|s^{t-1}} \mu(s^t|s^{t-1}) U_C^m(s^t) y^s(s^t) \left\{ \left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho} + \left(\frac{\rho-1}{\rho}\right)^{-1} \left[\frac{1+\tau_c(s^{t-1})}{(1-\tau_r(s^{t-1}))(1-\tau_\ell(s^{t-1}))}\right] \frac{U_L^m(s^t)}{U_C^m(s^t)} \frac{1}{A(s_t)} \right\} = 0. \tag{125}$$

Next, we set the real wage as follows:

$$\frac{W(s^t)}{P(s^t)} = -\frac{U_L^m(s^t)}{U_C^m(s^t)} \left(\frac{1+\tau_c(s^{t-1})}{1-\tau_\ell(s^{t-1})}\right), \tag{126}$$

and therefore satisfy the household's intratemporal condition in (13). Substituting the above expression for the real wage into (124) and (125), we obtain:

$$\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho} - \frac{W(s^t)}{P(s^t)} \left[(1-\tau_r(s^{t-1})) \left(\frac{\rho-1}{\rho}\right)\right]^{-1} \frac{1}{A(s_t)} = 0.$$

and

$$\sum_{s^t|s^{t-1}} \mu(s^t|s^{t-1}) U_C^m(s^t) y^s(s^t) \left\{ \left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho} - \frac{W(s^t)}{P(s^t)} \left[(1-\tau_r(s^{t-1})) \left(\frac{\rho-1}{\rho}\right)\right]^{-1} \frac{1}{A(s_t)} \right\} = 0.$$

Combining these with the CES demand functions in (68), and with some rearrangement, we derive the following two conditions:

$$p_t^f(s^t) - \left[(1-\tau_r(s^{t-1})) \left(\frac{\rho-1}{\rho}\right)\right]^{-1} \frac{W(s^t)}{A(s_t)} = 0.$$

and

$$\sum_{s^t|s^{t-1}} \mu(s^t|s^{t-1}) \frac{U_C^m(s^t)}{P(s^t)} Y(s^t) P(s^t)^\rho \left\{ p_t^s(s^{t-1}) - \left[(1-\tau_r(s^{t-1})) \left(\frac{\rho-1}{\rho}\right)\right]^{-1} \frac{W(s^t)}{A(s_t)} \right\} = 0.$$

Therefore both the flexible-price and the sticky-price firm's optimality conditions are satisfied. The remainder of the proof of sufficiency follows the same steps as in the proof of Proposition 2 [Section A.6 of the Appendix].

B.2 Proof of Proposition 10

We write the planner's Lagrangian as follows:

$$
\begin{aligned}
\mathcal{L} = & \sum_t \sum_{s^t} \beta^t \mu(s^t) \mathcal{W}(C(s^t), L(s^t); \varphi, \nu, \lambda) \\
& + \sum_t \sum_{s^t} \beta^t \mu(s^t) \varsigma^Y(s^t) \left\{ \left[\kappa y^s(s^t)^{\frac{\rho-1}{\rho}} + (1-\kappa) y^f(s^t)^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}} - Y(s^t) \right\} \\
& + \sum_t \sum_{s^t} \beta^t \mu(s^t) \varsigma^L(s^t) \left\{ \kappa \frac{y^s(s^t)}{A(s_t)} + (1-\kappa) \frac{y^f(s^t)}{A(s_t)} - L(s^t) \right\} \\
& + \sum_t \sum_{s^t} \beta^t \mu(s^t) \varsigma^C(s^t) \left\{ Y(s^t) - C(s^t) \right\} \\
& + \sum_t \sum_{s^t} \beta^t \mu(s^t) \xi(s^t) \left\{ \chi(s^{t-1}) U_C^m(s^t) \left(\frac{y^f(s^t)}{Y(s^t)} \right)^{-1/\rho} + U_L^m(s^t) \frac{1}{A(s_t)} \right\}
\end{aligned}
$$

The FOC with respect to $y^s(s^t)$ is given by:

$$
0 = \kappa \varsigma^Y(s^t) \left[\kappa y^s(s^t)^{\frac{\rho-1}{\rho}} + (1-\kappa) y^f(s^t)^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}-1} y^s(s^t)^{\frac{\rho-1}{\rho}-1} + \kappa \varsigma^L(s^t) \frac{1}{A(s_t)}, \tag{127}
$$

and the FOC with respect to $y^f(s^t)$ is given by:

$$
\begin{aligned}
0 = & (1-\kappa) \varsigma^Y(s^t) \left[\kappa y^s(s^t)^{\frac{\rho-1}{\rho}} + (1-\kappa) y^f(s^t)^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}-1} y^f(s^t)^{\frac{\rho-1}{\rho}-1} + (1-\kappa) \varsigma^L(s^t) \frac{1}{A(s_t)} \\
& - \frac{1}{\rho} \xi(s^t) \chi(s^{t-1}) U_C^m(s^t) \left[\frac{y^f(s^t)}{Y(s^t)} \right]^{-1/\rho} \frac{1}{y^f(s^t)}.
\end{aligned} \tag{128}
$$

Note that we can rewrite (127) as

$$
0 = \kappa \varsigma^Y(s^t) Y(s^t)^{1/\rho} y^s(s^t)^{\frac{\rho-1}{\rho}} + \kappa \varsigma^L(s^t) \frac{y^s(s^t)}{A(s_t)} \tag{129}
$$

We can also rewrite (128) as:

$$
0 = (1-\kappa) \varsigma^Y(s^t) Y(s^t)^{1/\rho} y^f(s^t)^{\frac{\rho-1}{\rho}} + (1-\kappa) \varsigma^L(s^t) \frac{y^f(s^t)}{A(s_t)} - \frac{1}{\rho} \xi(s^t) \chi(s^{t-1}) U_C^m(s^t) \left[\frac{y^f(s^t)}{Y(s^t)} \right]^{-1/\rho} \tag{130}
$$

Summing (129) and (130) yields:

$$
0 = \varsigma^Y(s^t) Y(s^t) + \varsigma^L(s^t) L(s^t) - \frac{1}{\rho} \xi(s^t) \chi(s^{t-1}) U_C^m(s^t) \left[\frac{y^f(s^t)}{Y(s^t)} \right]^{-1/\rho}, \tag{131}
$$

which can be rewritten as follows:

$$
-\frac{\varsigma^L(s^t)}{\varsigma^Y(s^t) - \frac{1}{\rho} \xi(s^t) \chi(s^{t-1}) U_C^m(s^t) \left[\frac{y^f(s^t)}{Y(s^t)} \right]^{-1/\rho} \frac{1}{Y(s^t)}} = \frac{Y(s^t)}{L(s^t)}. \tag{132}
$$

Next, the FOC with respect to $C(s^t)$ is given by:

$$0 = \mathcal{W}_C(s^t) - \varsigma^C(s^t) + \xi(s^t)\chi(s^{t-1})U^m_{CC}(s^t)\left(\frac{y^f(s^t)}{Y(s^t)}\right)^{-1/\rho}, \tag{133}$$

and the FOC with respect to $Y(s^t)$ is given by:

$$0 = -\varsigma^Y(s^t) + \varsigma^C(s^t) + \frac{1}{\rho}\xi(s^t)\chi(s^{t-1})U^m_C(s^t)\left(\frac{y^f(s^t)}{Y(s^t)}\right)^{-1/\rho}\frac{1}{Y(s^t)}. \tag{134}$$

Combining (133) and (134) yields:

$$\varsigma^Y(s^t) = \mathcal{W}_C(s^t) + \xi(s^t)\chi(s^{t-1})U^m_{CC}(s^t)\left(\frac{y^f(s^t)}{Y(s^t)}\right)^{-1/\rho} + \frac{1}{\rho}\xi(s^t)\chi(s^{t-1})U^m_C(s^t)\left(\frac{y^f(s^t)}{Y(s^t)}\right)^{-1/\rho}\frac{1}{Y(s^t)} \tag{135}$$

The FOC with respect to $L(s^t)$ implies:

$$\varsigma^L(s^t) = \mathcal{W}_L(s^t) + \xi(s^t)U^m_{LL}(s^t)\frac{1}{A(s_t)}. \tag{136}$$

Finally, we use (135) and (136) to substitute for $\varsigma^Y(s^t)$ and $\varsigma^L(s^t)$ in (132) and obtain:

$$-\frac{\mathcal{W}_L(s^t) + \xi(s^t)U^m_{LL}(s^t)\frac{1}{A(s_t)}}{\mathcal{W}_C(s^t) + \xi(s^t)\chi(s^{t-1})U^m_{CC}(s^t)\left(\frac{y^f(s^t)}{Y(s^t)}\right)^{-1/\rho}} = \frac{Y(s^t)}{L(s^t)}$$

as in (119).

B.3 Proof of Theorem 3

We substitute the expressions for $\mathcal{W}_C(s^t)$ and $\mathcal{W}_L(s^t)$ from (96) and (97) into (119) and obtain the following Ramsey optimality condition:

$$-\frac{U^m_L(s^t)}{U^m_C(s^t)}\left\{\frac{\sum_{i\in I}\pi^i\omega^i_L(\varphi,s_t)\left[\frac{\lambda^i}{\varphi^i}+\nu^i(1+\eta)\right] + \xi(s^t)\frac{U^m_{LL}(s^t)}{U^m_L(s^t)}\frac{1}{A(s_t)}}{\sum_{i\in I}\pi^i\omega^i_C(\varphi)\left[\frac{\lambda^i}{\varphi^i}+\nu^i(1-\gamma)\right] + \xi(s^t)\chi(s^{t-1})\frac{U^m_{CC}(s^t)}{U^m_C(s^t)}\left(\frac{y^f(s^t)}{Y(s^t)}\right)^{-1/\rho}}\right\} = \frac{Y(s^t)}{L(s^t)}.$$

Therefore the optimal monetary wedge, as defined in (38), satisfies:

$$1 - \tau^*_M(s^t) = \frac{\mathcal{H}(s^{t-1}) + \xi(s^t)\frac{U^m_{CC}(s^t)}{U^m_C(s^t)}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho}}{\mathcal{I}(s_t) + \xi(s^t)\frac{U^m_{LL}(s^t)}{U^m_L(s^t)}\frac{1}{A(s_t)}}. \tag{137}$$

where we let $\mathcal{H}(s^{t-1}) \equiv \chi(s^{t-1})^{-1}\Omega_C > 0$.

First, note that when $\xi(s^t) = 0$, the constraint is slack. Therefore, it is clear that $\tau^*_M(s^t) = 0$ if and only if

$$\mathcal{I}(s_t) = \bar{\mathcal{I}}(s^{t-1}) \equiv \mathcal{H}(s^{t-1}).$$

Next we use the representation of the monetary tax found in (104) and repeated here:

$$A(s_t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho} = (1-\tau_M^*(s^t))\frac{Y(s^t)}{L(s^t)} \tag{138}$$

Substituting the optimal monetary wedge from (137) into (138) we obtain:

$$A(s_t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho} = \left\{\frac{\mathcal{H}(s^{t-1})+\xi(s^t)\frac{U^m_{CC}(s^t)}{U^m_C(s^t)}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho}}{\mathcal{I}(s_t)+\xi(s^t)\frac{U^m_{LL}(s^t)}{U^m_L(s^t)}\frac{1}{A(s_t)}}\right\}\frac{Y(s^t)}{L(s^t)}.$$

Rearrangement, yields:

$$1 = \frac{\mathcal{H}(s^{t-1})Y(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1/\rho}+\xi(s^t)\frac{U^m_{CC}(s^t)C(s^t)}{U^m_C(s^t)}}{\mathcal{I}(s_t)A(s_t)L(s^t)+\xi(s^t)\frac{U^m_{LL}(s^t)L(s^t)}{U^m_L(s^t)}},$$

which reduces to:

$$\mathcal{I}(s_t)+(\eta+\gamma)\frac{\xi(s^t)}{A(s_t)L(s^t)}-\mathcal{H}(s^{t-1})\frac{Y(s^t)}{A(s_t)L(s^t)}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1/\rho}=0$$

We define

$$\hat{\xi}(s^t)\equiv\frac{\xi(s^t)}{A(s_t)L(s^t)}\mathcal{H}(s^{t-1})^{-1}$$

We have that:

$$\mathcal{I}(s_t)+(\eta+\gamma)\mathcal{H}(s^{t-1})\hat{\xi}(s^t)-\mathcal{H}(s^{t-1})\frac{Y(s^t)}{A(s_t)L(s^t)}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1/\rho}=0$$

Next, using condition (138) we have the following optimality condition:

$$\mathcal{I}(s_t)+(\eta+\gamma)\mathcal{H}(s^{t-1})\hat{\xi}(s^t)-\mathcal{H}(s^{t-1})(1-\tau_M(s^t))^{-1}=0$$

We let g be the function defined by:

$$g(\mathcal{I}(s_t),\tau_M(s^t))\equiv\mathcal{I}(s_t)+\mathcal{H}(s^{t-1})\left[(\eta+\gamma)\hat{\xi}(s^t)-(1-\tau_M(s^t))^{-1}\right].$$

Therefore, the optimal monetary tax satisfies: $g(\mathcal{I}(s_t),\tau_M^*(s^t)) = 0$. By the implicit function theorem:

$$\frac{d\tau_M^*(s^t)}{d\mathcal{I}(s_t)} = -\frac{dg/d\mathcal{I}(s_t)}{dg/d\tau_M^*(s^t)} = -\frac{1}{\mathcal{H}(s^{t-1})\left\{(\eta+\gamma)\frac{d\hat{\xi}(s^t)}{d\tau_M(s^t)}-(1-\tau_M^*(s^t))^{-2}\right\}}$$

The derivative of the optimal monetary tax satisfies:

$$\frac{d\tau_M^*(s^t)}{d\mathcal{I}(s_t)} = \frac{1}{\mathcal{H}(s^{t-1})}\left\{(1-\tau_M^*(s^t))^{-2}-(\eta+\gamma)\frac{d\hat{\xi}(s^t)}{d\tau_M^*(s^t)}\right\}^{-1} \tag{139}$$

An expression for $\hat{\xi}(s^t)$. The planner's optimality condition in (129) implies:

$$\varsigma^L(s^t) = -\varsigma^Y(s^t)A(s_t)\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho}$$

Substituting this into (131) we obtain:

$$0 = \varsigma^Y(s^t)Y(s^t) - \varsigma^Y(s^t)\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho} A(s_t)L(s^t) - \frac{1}{\rho}\xi(s^t)\chi(s^{t-1})U_C^m(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho},$$

which can be rewritten as:

$$0 = \frac{\varsigma^Y(s^t)}{\chi(s^{t-1})U_C^m(s^t)}\left[\frac{Y(s^t)}{A(s_t)L(s^t)} - \left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho}\right] - \frac{1}{\rho}\frac{\xi(s^t)}{A(s_t)L(s^t)}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho}$$

Rearranging, we get that:

$$\frac{\xi(s^t)}{A(s_t)L(s^t)} = \rho\frac{\varsigma^Y(s^t)}{\chi(s^{t-1})U_C^m(s^t)}\left[\frac{Y(s^t)}{A(s_t)L(s^t)}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1/\rho} - \left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1/\rho}\right] \tag{140}$$

Therefore

$$\hat{\xi}(s^t) = \rho\frac{\mathcal{H}(s^{t-1})^{-1}}{\chi(s^{t-1})U_C^m(s^t)}\varsigma^Y(s^t)\left[\frac{Y(s^t)}{A(s_t)L(s^t)}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1/\rho} - \left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1/\rho}\right] \tag{141}$$

In what follows, we turn back to the equivalent equilibrium representation used in Section (A.14) of the Appendix. Recall the forecast errors $\epsilon(s^t)$ defined in (29). From equation (114) we have that the monetary tax satisfies:

$$1 - \tau_M(s^t) = \frac{\kappa\epsilon(s^t)^{-\rho} + (1-\kappa)}{\kappa\epsilon(s^t)^{1-\rho} + (1-\kappa)}.$$

Combining this with (138), we obtain:

$$\frac{Y(s^t)}{A(s_t)L(s^t)}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1/\rho} = \frac{\kappa\epsilon(s^t)^{1-\rho} + (1-\kappa)}{\kappa\epsilon(s^t)^{-\rho} + (1-\kappa)}$$

Furthermore, using (108), the following is also true:

$$\left[\frac{y^s s^t)}{Y(s^t)}\right]^{-1/\rho}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1/\rho} = \left[\epsilon(s^t)^{-\rho}\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1/\rho} = \epsilon(s^t)$$

Therefore, we may rewrite (141) as follows:

$$\hat{\xi}(s^t) = \rho\frac{\mathcal{H}(s^{t-1})^{-1}}{\chi(s^{t-1})U_C^m(s^t)}\varsigma^Y(s^t)\left[\frac{(1-\kappa)(1-\epsilon(s^t))}{\kappa\epsilon(s^t)^{-\rho} + (1-\kappa)}\right] \tag{142}$$

Next we use the planner's optimality condition in (135):

$$\varsigma^Y(s^t) = \mathcal{W}_C(s^t) + \xi(s^t)\chi(s^{t-1}) \left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho} \frac{U_C^m(s^t)}{Y(s^t)} \left[\frac{U_{CC}^m(s^t)C(s^t)}{U_C^m(s^t)} + \frac{1}{\rho}\right]$$

Substituting in for $\xi(s^t)$ from (140) we have that:

$$\varsigma^Y(s^t) = \mathcal{W}_C(s^t) + \varsigma^Y(s^t)(1-\gamma\rho) \left[1 - \left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho} \frac{A(s_t)L(s^t)}{Y(s^t)}\right]$$

Therefore

$$\varsigma^Y(s^t) = \mathcal{W}_C(s^t) \left[1 - (1-\gamma\rho) \left[1 - \left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho} \frac{A(s_t)L(s^t)}{Y(s^t)}\right]\right]^{-1}$$

We have that:

$$\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho} \frac{A(s_t)L(s^t)}{Y(s^t)} = \epsilon(s^t) \left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho} \frac{A(s_t)L(s^t)}{Y(s^t)} = \epsilon(s^t) \frac{\kappa\epsilon(s^t)^{-\rho} + (1-\kappa)}{\kappa\epsilon(s^t)^{1-\rho} + (1-\kappa)}$$

Therefore

$$\varsigma^Y(s^t) = \mathcal{W}_C(s^t) \left[1 - (1-\gamma\rho) \left[1 - \epsilon(s^t) \frac{\kappa\epsilon(s^t)^{-\rho} + (1-\kappa)}{\kappa\epsilon(s^t)^{1-\rho} + (1-\kappa)}\right]\right]^{-1}$$

which reduces to

$$\varsigma^Y(s^t) = \mathcal{W}_C(s^t) \left[1 - (1-\gamma\rho)(1-\kappa) \left[\frac{1-\epsilon(s^t)}{\kappa\epsilon(s^t)^{1-\rho} + (1-\kappa)}\right]\right]^{-1}$$

Therefore

$$\varsigma^Y(s^t) = \mathcal{W}_C(s^t) \frac{\kappa\epsilon(s^t)^{1-\rho} + (1-\kappa)}{\kappa\epsilon(s^t)^{1-\rho} + (1-\gamma\rho)(1-\kappa)\epsilon(s^t) + \gamma\rho(1-\kappa)}$$

Substituting this expression for $\varsigma^Y(s^t)$ into (142), we obtain the following expression for $\hat{\xi}(s^t)$ as a function of $\epsilon(s^t)$:

$$\hat{\xi}(s^t) = \rho \left[\frac{(1-\kappa)(1-\epsilon(s^t))}{\kappa\epsilon(s^t)^{1-\rho} + (1-\gamma\rho)(1-\kappa)\epsilon(s^t) + \gamma\rho(1-\kappa)}\right] \left[\frac{\kappa\epsilon(s^t)^{1-\rho} + (1-\kappa)}{\kappa\epsilon(s^t)^{-\rho} + (1-\kappa)}\right].$$

Next, for shorthand we let

$$\Sigma(\epsilon) \equiv \frac{(1-\kappa)(1-\epsilon)}{\kappa\epsilon^{1-\rho} + (1-\gamma\rho)(1-\kappa)\epsilon + \gamma\rho(1-\kappa)} \tag{143}$$

Therefore:

$$\hat{\xi}(s^t) = \rho\Sigma(\epsilon(s^t))(1-\tau_M(s^t))^{-1} \tag{144}$$

Derivative of $\tau_M(s^t)$. The derivative of the optimal monetary "tax" satisfies (139). Evaluating this derivative at the benchmark in which $\tau_M(s^t) = 0$, we have:

$$\left.\frac{d\tau_M(s^t)}{d\mathcal{I}(s_t)}\right|_{\tau_M(s^t)=0} = \frac{1}{\mathcal{H}(s^{t-1})}\left\{1-(\eta+\gamma)\left.\frac{d\hat{\xi}(s^t)}{d\tau_M(s^t)}\right|_{\tau_M(s^t)=0}\right\}^{-1} \tag{145}$$

where $\hat{\xi}(s^t)$ satisfies (144). Taking the first derivative of the expression in (144), we get:

$$\frac{d\hat{\xi}}{d\tau_M} = \rho\Sigma(\epsilon)(1-\tau_M)^{-2} + \rho(1-\tau_M)^{-1}\frac{d\Sigma}{d\epsilon}\frac{d\epsilon}{d\tau_M}.$$

The derivative $d\Sigma/d\epsilon$ satisfies:

$$\frac{d\Sigma(\epsilon)}{d\epsilon} = -(1-\kappa)\frac{\kappa\epsilon^{1-\rho}+(1-\rho)\kappa\epsilon^{-\rho}(1-\epsilon)+(1-\kappa)}{(\kappa\epsilon^{1-\rho}+(1-\gamma\rho)(1-\kappa)\epsilon+\gamma\rho(1-\kappa))^2}$$

Next, we obtain $d\epsilon/d\tau_M$ as follows. The monetary tax satisfies equation (114):

$$1-\tau_M(\epsilon) = \frac{\kappa\epsilon^{-\rho}+(1-\kappa)}{\kappa\epsilon^{1-\rho}+(1-\kappa)}$$

Rearranging, we obtain the following expression:

$$\kappa\epsilon^{1-\rho} - \kappa\epsilon^{-\rho} - \tau_M\kappa\epsilon^{1-\rho} - \tau_M(1-\kappa) = 0$$

We let $\hat{g}$ be the function defined by:

$$\hat{g}(\tau_M,\epsilon) \equiv (1-\tau_M)\kappa\epsilon^{1-\rho} - \kappa\epsilon^{-\rho} - \tau_M(1-\kappa)$$

Therefore, $\hat{g}(\tau_M,\epsilon) = 0$. By the implicit function theorem:

$$\frac{d\epsilon}{d\tau_M} = -\frac{d\hat{g}/d\tau_M}{dg/d\epsilon} = \frac{\kappa\epsilon^{1-\rho}+(1-\kappa)}{(1-\rho)(1-\tau_M)\kappa\epsilon^{-\rho}+\rho\kappa\epsilon^{-\rho-1}}$$

Therefore

$$\begin{aligned}\frac{d\hat{\xi}}{d\tau_M} =& \rho\frac{(1-\kappa)(1-\epsilon)}{\kappa\epsilon^{1-\rho}+(1-\gamma\rho)(1-\kappa)\epsilon+\gamma\rho(1-\kappa)}(1-\tau_M)^{-2} \\ &-\rho(1-\tau_M)^{-1}(1-\kappa)\left[\frac{\kappa\epsilon^{1-\rho}+(1-\rho)\kappa\epsilon^{-\rho}(1-\epsilon)+(1-\kappa)}{(\kappa\epsilon^{1-\rho}+(1-\gamma\rho)(1-\kappa)\epsilon+\gamma\rho(1-\kappa))^2}\right]\left[\frac{\kappa\epsilon^{1-\rho}+(1-\kappa)}{(1-\rho)(1-\tau_M)\kappa\epsilon^{-\rho}+\rho\kappa\epsilon^{-\rho-1}}\right]\end{aligned}$$

Evaluating this derivative at $\tau_M = 0$, or equivalently at $\epsilon = 1$, we have:

$$\left.\frac{d\hat{\xi}}{d\tau_M}\right|_{\tau_M=0} = -\rho(1-\kappa)\left[\frac{\kappa+(1-\kappa)}{(\kappa+(1-\gamma\rho)(1-\kappa)+\gamma\rho(1-\kappa))^2}\right]\left[\frac{\kappa+(1-\kappa)}{(1-\rho)\kappa\frac{\kappa+(1-\kappa)}{\kappa+(1-\kappa)}+\rho\kappa}\right]$$

which reduces to

$$\left.\frac{d\hat{\xi}}{d\tau_M}\right|_{\tau_M=0} = -\rho(1-\kappa)\left[\frac{1}{(1-\rho)\kappa+\rho\kappa}\right] = -\rho\frac{1-\kappa}{\kappa}$$

Substituting this into (145), we obtain:

$$\left.\frac{d\tau_M(s^t)}{d\mathcal{I}(s_t)}\right|_{\tau_M(s^t)=0} = \frac{1}{\bar{\mathcal{I}}(s^{t-1})\left\{1+\rho(\eta+\gamma)\frac{1-\kappa}{\kappa}\right\}} > 0.$$

where $\bar{\mathcal{I}}(s^{t-1}) \equiv \mathcal{H}(s^{t-1})$. Therefore, taking a first-order Taylor approximation of $\tau_M^*(s^t)$ around the point $\mathcal{I}(s_t) = \bar{\mathcal{I}}(s^{t-1})$, we have:

$$\tau_M^*(s^t) \approx 0 + \frac{1}{\bar{\mathcal{I}}(s^{t-1})\left\{1+\rho(\eta+\gamma)\frac{1-\kappa}{\kappa}\right\}}[\mathcal{I}(s_t) - \bar{\mathcal{I}}(s^{t-1})],$$

which coincides with the expression in (40).

C Appendix: Proofs for Heterogeneous Equity Shares Economy

In this section of the appendix we provide the proofs for the economy with heterogeneous equity shares presented in Section 7.

C.1 Derivation of Implementability Conditions (45)

We derive condition (45). We take the household's budget constraint in (44) for type $i \in I$, multiply both sides by $\Lambda^i(s^t)$, and use the household's FOCs in (52) and (53) to substitute out consumption and labor prices. Doing so, we obtain:

$$\begin{aligned} U_c^i(s^t)c^i(s^t) + \frac{1}{\theta^i(s_t)}U_\ell^i(s^t)\ell^i(s^t) =& \Lambda^i(s^t)z^i(s^t|s^{t-1}) - \Lambda^i(s^t)\sum_{s^{t+1}|s^t} Q(s^{t+1}|s^t)z^i(s^{t+1}|s^t) - \Lambda^i(s^t)b^i(s^t) \\ &+ \Lambda^i(s^t)(1+i(s^{t-1}))b^i(s^{t-1}) + \Lambda^i(s^t)P(s^t)\bar{T}(s^t) + \Lambda^i(s^t)\sigma^i(1-\tau_\Pi)\Pi(s^t) \end{aligned}$$

where $\bar{T}(s^t) = T(s^t) + (1-\tau_\Pi)\Pi(s^t)/P(s^t)$ as before. Multiplying both sides by $\beta^t\mu(s^t)$, summing over t and s^t, and using the household's intertemporal optimality conditions (57)-(56) to cancel terms, we obtain:

$$\sum_t\sum_{s^t}\beta^t\mu(s^t)\left[U_c^i(s^t)c^i(s^t) + \frac{1}{\theta^i(s_t)}U_\ell^i(s^t)\ell^i(s^t)\right] \leq U_c^i(s_0)\bar{T} + \sigma^i\frac{1-\tau_\Pi}{1+\tau_c}\sum_t\sum_{s^t}\beta^t\mu(s^t)U_c^i(s^t)\frac{\Pi(s^t)}{P(s^t)},$$

where

$$\bar{T} \equiv \frac{1}{U_c^i(s_0)(1+\tau_c)}\sum_t\sum_{s^t}\beta^t\mu(s^t)U_c^i(s^t)\bar{T}(s^t)$$

for all $i \in I$ as before. Finally, using the solution and envelope conditions for the static sub-problem described in Lemma 1, as well as the fact that individual allocations satisfy (16), we can rewrite the above conditions as:

$$\sum_t \sum_{s^t} \beta^t \mu(s^t) \left[U_C^m(s^t) \omega_C^i(\varphi) C(s^t) + U_L^m(s^t) \omega_L^i(\varphi, s_t) L(s^t) \right] = U_c^i(s_0) \bar{T} + \sigma^i (1 - \tau_\Pi) \sum_t \sum_{s^t} \beta^t \mu(s^t) \frac{U_C^m(s^t)}{1 + \tau_c} \frac{\Pi(s^t)}{P(s^t)},$$

where $\bar{T}$ is as in (19), as was to be shown.

C.2 Proof of Proposition 7

Necessity. Necessity of parts (i) and (ii) of the proposition follow from the same steps as those used to prove Proposition 2. What remains to be shown is necessity of the budget implementability conditions in 46.

First, we derive an expression for real profits, $\Pi(s^t)/P(s^t)$, in terms of allocations alone. To do so, we first write aggregate profits in the following way:

$$\Pi(s^t) = (1 - \kappa) \Pi^f(s^t) + \kappa \Pi^s(s^t)$$

where $\Pi^f(s^t)$ denotes profits of the flexible-price firms and $\Pi^s(s^t)$ denotes profits of the sticky price firms in history s^t. Profits of these firms are given by:

$$\Pi^f(s^t) = \left[(1 - \tau_r) p_t^f(s^t) - \frac{W(s^t)}{A(s^t)} \right] y^f(s^t) \qquad \text{and} \qquad \Pi^s(s^t) = \left[(1 - \tau_r) p_t^s(s^{t-1}) - \frac{W(s^t)}{A(s^t)} \right] y^s(s^t).$$

Combining these expressions with the demand functions,

$$\left[\frac{y^f(s^t)}{Y(s^t)} \right]^{-1/\rho} = \frac{p_t^f(s^t)}{P(s^t)} \qquad \text{and} \qquad \left[\frac{y^s(s^t)}{Y(s^t)} \right]^{-1/\rho} = \frac{p_t^s(s^{t-1})}{P(s^t)}, \tag{146}$$

we find that real profits of these firms are given by:

$$\frac{\Pi^f(s^t)}{P(s^t)} = \left[(1 - \tau_r) \left[\frac{y^f(s^t)}{Y(s^t)} \right]^{-1/\rho} - \frac{W(s^t)}{P(s^t) A(s^t)} \right] y^f(s^t) \qquad \text{and} \qquad \frac{\Pi^s(s^t)}{P(s^t)} = \left[(1 - \tau_r) \left[\frac{y^s(s^t)}{Y(s^t)} \right]^{-1/\rho} - \frac{W(s^t)}{P(s^t) A(s^t)}$$

This implies aggregate real profits can be written as:

$$\frac{\Pi(s^t)}{P(s^t)} = (1 - \tau_r) \left[\kappa \left[\frac{y^s(s^t)}{Y(s^t)} \right]^{1-1/\rho} + (1 - \kappa) \left[\frac{y^f(s^t)}{Y(s^t)} \right]^{1-1/\rho} \right] Y(s^t) - \frac{W(s^t)}{P(s^t)} A(s^t)^{-1} \left[(1 - \kappa) y^f(s^t) + \kappa y^s(s^t) \right]$$

Next resource constraints satisfy (36), which implies that real profits are given by:

$$\frac{\Pi(s^t)}{P(s^t)} = (1 - \tau_r) Y(s^t) - \frac{W(s^t)}{P(s^t)} L(s^t)$$

Next, we replace the real wage $W(s^t)/P(s^t)$ in the above expression using the representative household's intratemporal condition, (13). This gives us the following expression for real profits:

$$\frac{\Pi(s^t)}{P(s^t)} = (1-\tau_r)C(s^t) + \frac{1+\tau_c}{1-\tau_\ell}\frac{U_L^m(s^t)}{U_C^m(s^t)}L(s^t) \tag{147}$$

Multiplying both sides of this by $U_C^m(s^t)/1+\tau_c$, we get:

$$\frac{U_C^m(s^t)}{1+\tau_c}\frac{\Pi(s^t)}{P(s^t)} = (1-\tau_r)\frac{U_C^m(s^t)}{1+\tau_c}C(s^t) + \frac{U_L^m(s^t)}{1-\tau_\ell}L(s^t),$$

and as a result, the implementability condition in (45) becomes:

$$\sum_t\sum_{s^t}\beta^t\mu(s^t)\left[U_C^m(s^t)\omega_C^i(\varphi)C(s^t) + U_L^m(s^t)\omega_L^i(\varphi,s_t)L(s^t)\right] = U_c^i(s_0)\bar{T}$$
$$+\sigma^i\frac{1-\tau_\Pi}{1-\tau_\ell}\sum_t\sum_{s^t}\beta^t\mu(s^t)\left[(1-\tau_r)\frac{1-\tau_\ell}{1+\tau_c}U_C^m(s^t)C\right.$$

Finally, using the definition of χ in (25), we obtain:

$$\sum_t\sum_{s^t}\beta^t\mu(s^t)\left[U_C^m(s^t)\omega_C^i(\varphi)C(s^t) + U_L^m(s^t)\omega_L^i(\varphi,s_t)L(s^t)\right] = U_c^i(s_0)\bar{T}.$$
$$+\sigma^i\frac{1-\tau_\Pi}{1-\tau_\ell}\sum_t\sum_{s^t}\beta^t\mu(s^t)\left[\chi\frac{\rho}{\rho-1}U_C^m(s^t)C(s^t) + U\right.$$

Finally, we define

$$\vartheta \equiv \frac{1-\tau_\Pi}{1-\tau_\ell},$$

and obtain the expression in (46).

Sufficiency. Follows the same argument as in the sufficiency proof of Proposition 2.

C.3 The Ramsey Problem

In this section of the appendix, we state and solve the Ramsey problem. We assume the same utilitarian social welfare function (30) as before. Again we let $\pi^i\nu^i$ denote the Lagrange multiplier on the implementability condition (46) of type $i \in I$ and subsume these into the maximand. Provided that $\vartheta = \bar{\vartheta}$, we can define a new pseudo-welfare function $\hat{\mathcal{W}}(\cdot)$ as follows:

$$\begin{aligned}\hat{\mathcal{W}}(C(s^t),L(s^t),s_t;\varphi,\nu,\lambda,\sigma,\chi) \equiv\ & \mathcal{W}(C(s^t),L(s^t),s_t;\varphi,\nu,\lambda)\\ & -\bar{\vartheta}\left(\sum_{i\in I}\pi^i\nu^i\sigma^i\right)\left[\chi\frac{\rho}{\rho-1}U_C^m(s^t)C(s^t) + U_L^m(s^t)L(s^t)\right]\end{aligned}$$

where $\mathcal{W}$ is as in (31). With this, we can recast the Ramsey planning problem as follows.

Ramsey Planner's Problem. *The Ramsey planner chooses an allocation* $x \equiv \{y^s(s^t), y^f(s^t), C(s^t), Y(s^t), L(s^t)\}_{t\geq 0, s^t \in S^t}$, *market weights* $\varphi \equiv (\varphi^i)$, *and constants* $\bar{T} \in \mathbb{R}$ *and* $\chi \in \mathbb{R}_+$, *in order to maximize:*

$$\sum_t \sum_{s^t} \beta^t \mu(s^t) \hat{\mathcal{W}}(C(s^t), L(s^t), s_t; \varphi, \nu, \lambda, \sigma, \chi) - U_C^m(s_0)\bar{T} \sum_{i\in I} \pi^i \nu^i \tag{148}$$

subject to (36), (26), *and* (27).

We let $\beta^t \mu(s^t)(1-\kappa)\xi(s^t)$ and $\beta^t \mu(s^{t-1})\kappa \upsilon(s^{t-1})$ denote the Lagrange multipliers on the implementability conditions (26) and (27), respectively. We obtain the following Ramsey optimality condition.

Proposition 11. *A Ramsey optimum* x^* *satisfies, for all* $s^t \in S^t$,

$$-\frac{\hat{\mathcal{W}}_L(s^t) + (U_L^m(s^t) + U_{LL}^m(s^t)L(s^t))\left[\kappa \upsilon(s^{t-1})\frac{y^s(s^t)}{A(s_t)L(s^t)} + (1-\kappa)\xi(s^t)\frac{y^f(s^t)}{A(s_t)L(s^t)}\right]}{\hat{\mathcal{W}}_C(s^t) + \chi(U_C^m(s^t) + U_{CC}^m(s^t)C(s^t))\left[\kappa \upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{1-1/\rho} + (1-\kappa)\xi(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1-1/\rho}\right]} = \frac{Y(s^t)}{L(s^t)}. \tag{149}$$

where

$$\hat{\mathcal{W}}_C(s^t) = \mathcal{W}_C(s^t) - \bar{\vartheta}\left(\sum_{i\in I}\pi^i\nu^i\sigma^i\right)\chi\frac{\rho}{\rho-1}[U_C^m(s^t) + U_{CC}^m(s^t)C(s^t)] \tag{150}$$

$$\hat{\mathcal{W}}_L(s^t) = \mathcal{W}_L(s^t) - \bar{\vartheta}\left(\sum_{i\in I}\pi^i\nu^i\sigma^i\right)[U_L^m(s^t) + U_{LL}^m(s^t)L(s^t)] \tag{151}$$

Proof. The planner's problem is the same as in the proof of Proposition 8, with the exception of maximizing $\hat{\mathcal{W}}(s^t)$ instead of $\mathcal{W}(s^t)$. Therefore, following the same procedure as the proof of Proposition 8, we obtain the expression in (149).

Furthermore, taking the first derivatives of $\hat{\mathcal{W}}(C(s^t), L(s^t), s_t; \varphi, \nu, \lambda, \sigma, \chi)$, defined in (149). Taking the derivatives of this, we get:

$$\hat{\mathcal{W}}_C(s^t) = \mathcal{W}_C(s^t) - \bar{\vartheta}\chi\frac{\rho}{\rho-1}\left(\sum_{i\in I}\pi^i\nu^i\sigma^i\right)[U_C^m(s^t) + U_{CC}^m(s^t)C(s^t)]$$

$$\hat{\mathcal{W}}_L(s^t) = \mathcal{W}_L(s^t) - \bar{\vartheta}\left(\sum_{i\in I}\pi^i\nu^i\sigma^i\right)[U_L^m(s^t) + U_{LL}^m(s^t)L(s^t)]$$

□

C.4 Proof of Theorem 4

At the Ramsey optimum, we have:

$$-\frac{\mathcal{W}_L(s^t) + (U^m_L(s^t) + U^m_{LL}(s^t)L(s^t))\left[\kappa\upsilon(s^{t-1})\frac{y^s(s^t)}{A(s_t)L(s^t)} + (1-\kappa)\xi(s^t)\frac{y^f(s^t)}{A(s_t)L(s^t)} - \bar{\vartheta}\left(\sum_{i\in I}\pi^i\nu^i\sigma^i\right)\right]}{\mathcal{W}_C(s^t) + \chi(U^m_C(s^t) + U^m_{CC}(s^t)C(s^t))\left[\kappa\upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{1-1/\rho} + (1-\kappa)\xi(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1-1/\rho} - \bar{\vartheta}\frac{\rho}{\rho-1}\left(\sum_{i\in I}\pi^i\nu^i\sigma^i\right)\right]} \quad (152)$$

$$= \frac{Y(s^t)}{L(s^t)}$$

With separable and iso-elastic utility, $\mathcal{W}_C(s^t)$ and $\mathcal{W}_L(s^t)$ satisfy (96) and (97). Substituting these expressions for $\mathcal{W}_C(s^t)$ and $\mathcal{W}_L(s^t)$ into (152), we obtain:

$$-\frac{U^m_L(s^t)\left\{\sum_{i\in I}\pi^i\omega^i_L(\varphi,s_t)\left[\frac{\lambda^i}{\varphi^i} + \nu^i(1+\eta)\right] + (1+\eta)\left[\kappa\upsilon(s^{t-1})\frac{y^s(s^t)}{A(s_t)L(s^t)} + (1-\kappa)\xi(s^t)\frac{y^f(s^t)}{A(s_t)L(s^t)} - \bar{\vartheta}\left(\sum_{i\in I}\pi\right.\right.\right.}{U^m_C(s^t)\sum_{i\in I}\pi^i\omega^i_C(\varphi)\left[\frac{\lambda^i}{\varphi^i} + \nu^i(1-\gamma)\right] + \chi(1-\gamma)\left[\kappa\upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{1-1/\rho} + (1-\kappa)\xi(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1-1/\rho} - \bar{\vartheta}\frac{\rho}{\rho-1}\left(\sum\right.\right.}$$

$$= \frac{Y(s^t)}{L(s^t)}$$

Therefore the optimal monetary wedge satisfies:

$$1-\tau^*_M(s^t) = \frac{(\chi^*)^{-1}\sum_{i\in I}\pi^i\omega^i_C(\varphi)\left[\frac{\lambda^i}{\varphi^i} + \nu^i(1-\gamma)\right] + (1-\gamma)\left[\kappa\upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{1-1/\rho} + (1-\kappa)\xi(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1-1/\rho} - \right.}{\sum_{i\in I}\pi^i\omega^i_L(\varphi,s_t)\left[\frac{\lambda^i}{\varphi^i} + \nu^i(1+\eta)\right] + (1+\eta)\left[\kappa\upsilon(s^{t-1})\frac{y^s(s^t)}{A(s_t)L(s^t)} + (1-\kappa)\xi(s^t)\frac{y^f(s^t)}{A(s_t)L(s^t)} - \bar{\vartheta}\left(\sum\right.\right.}$$

Next we define a function $\mathcal{I}(s_t)$ and a scalar $\mathcal{H}$ as follows:

$$\mathcal{I}(s_t) \equiv \sum_{i\in I}\pi^i\omega^i_L(s^t)\left[\frac{\lambda^i}{\varphi^i} + \nu^i(1+\eta)\right], \qquad \text{and} \qquad \mathcal{H} \equiv (\chi^*)^{-1}\Omega_C.$$

The optimal monetary wedge can then be written as follows:

$$1-\tau^*_M(s^t) = \frac{\mathcal{H} + (1-\gamma)\left[\kappa\upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{1-1/\rho} + (1-\kappa)\xi(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1-1/\rho} - \bar{\vartheta}\frac{\rho}{\rho-1}\left(\sum_{i\in I}\pi^i\nu^i\sigma^i\right)\right]}{\mathcal{I}(s_t) + (1+\eta)\left[\kappa\upsilon(s^{t-1})\frac{y^s(s^t)}{A(s_t)L(s^t)} + (1-\kappa)\xi(s^t)\frac{y^f(s^t)}{A(s_t)L(s^t)} - \bar{\vartheta}\left(\sum_{i\in I}\pi^i\nu^i\sigma^i\right)\right]}. \quad (153)$$

Threshold. We first consider the conditions under which $\tau^*_M(s^t) = 0$. In this state: $y^s(s^t) = y^f(s^t) = Y(s^t) = A(s_t)L(s^t)$. Condition (153) reduces to:

$$1 = \frac{\mathcal{H} + (1-\gamma)\left[\kappa\upsilon(s^{t-1}) + (1-\kappa)\xi(s^t) - \bar{\vartheta}\frac{\rho}{\rho-1}\left(\sum_{i\in I}\pi^i\nu^i\sigma^i\right)\right]}{\mathcal{I}(s_t) + (1+\eta)\left[\kappa\upsilon(s^{t-1}) + (1-\kappa)\xi(s^t) - \bar{\vartheta}\left(\sum_{i\in I}\pi^i\nu^i\sigma^i\right)\right]}$$

Furthermore, conditions (77) and (79) imply that $\xi(s^t) = \upsilon(s^{t-1})$ in this state. Therefore:

$$1 = \frac{\mathcal{H} + (1-\gamma)\upsilon(s^{t-1}) - \bar{\vartheta}(1-\gamma)\frac{\rho}{\rho-1}\left(\sum_{i\in I}\pi^i\nu^i\sigma^i\right)}{\mathcal{I}(s_t) + (1+\eta)\upsilon(s^{t-1}) - \bar{\vartheta}(1+\eta)\left(\sum_{i\in I}\pi^i\nu^i\sigma^i\right)}$$

Solving this for $\mathcal{I}(s_t)$ we obtain the following threshold:

$$\hat{\mathcal{I}}(s^{t-1}) = \mathcal{H} - (\eta+\gamma)\upsilon(s^{t-1}) + \bar{\vartheta}\left[(1+\eta) - (1-\gamma)\frac{\rho}{\rho-1}\right]\left(\sum_{i\in I}\pi^i\nu^i\sigma^i\right)$$

When $\mathcal{I}(s_t) = \bar{\mathcal{I}}(s^{t-1})$ the optimal monetary tax is equal to zero: $\tau_M^*(s^t) = 0$.

Finally, using the fact that $\bar{\mathcal{I}}(s^{t-1}) = \mathcal{H} - (\eta+\gamma)\upsilon(s^{t-1})$, we have that:

$$\hat{\mathcal{I}}(s^{t-1}) = \bar{\mathcal{I}}(s^{t-1}) + \bar{\vartheta}\left[(1+\eta) - (1-\gamma)\frac{\rho}{\rho-1}\right]\left(\sum_{i\in I}\pi^i\nu^i\sigma^i\right).$$

The fictitious tax wedge. We next define a fictitious tax wedge as follows:

$$1 - \hat{\tau}(s^t) \equiv \frac{\mathcal{H} + (1-\gamma)\upsilon(s^{t-1}) - \bar{\vartheta}(1-\gamma)\frac{\rho}{\rho-1}\left(\sum_{i\in I}\pi^i\nu^i\sigma^i\right)}{\mathcal{I}(s_t) + (1+\eta)\upsilon(s^{t-1}) - \bar{\vartheta}(1+\eta)\left(\sum_{i\in I}\pi^i\nu^i\sigma^i\right)} \tag{154}$$

This wedge is unambiguously falling in $\mathcal{I}(s_t)$, as all other terms are constants. Furthermore, note that when $\mathcal{I}(s_t) = \hat{\mathcal{I}}(s^{t-1})$, this wedge is equal to one. As a result, the fictitious tax $\hat{\tau}(s^t)$ trivially satisfies:

$$\begin{aligned} \hat{\tau}(s^t) > 0 \quad & \text{if and only if} \quad \mathcal{I}(s_t) > \hat{\mathcal{I}}(s^{t-1}), \\ \hat{\tau}(s^t) = 0 \quad & \text{if and only if} \quad \mathcal{I}(s_t) = \hat{\mathcal{I}}(s^{t-1}), \\ \hat{\tau}(s^t) < 0 \quad & \text{if and only if} \quad \mathcal{I}(s_t) < \hat{\mathcal{I}}(s^{t-1}). \end{aligned}$$

The remainder of the proof follows the exact same steps and logic as in the proof of Theorem 2.

C.5 One-Period-Ahead Tax Rates

In this section of the appendix, we characterize the heterogeneous equity share economy with one-period-ahead tax rates. We state the planner's problem as follows

Ramsey Planner's Problem. *The Ramsey planner chooses an allocation,*

$$x \equiv \{y^s(s^t), y^f(s^t), C(s^t), Y(s^t), L(s^t)\}_{t\geq 0, s^t\in S^t},$$

market weights $\varphi \equiv (\varphi^i)$*, and scalar* $\bar{T} \in \mathbb{R}$*, in order to maximize* (148) *subject to* (118) *and* (117).

We let $\beta^t\mu(s^t)(1-\kappa)\xi(s^t)$ denote the Lagrange multiplier on the implementability condition (117). The Ramsey optimum can be characterized as follows.

Proposition 12. *A Ramsey optimum x^* satisfies*

$$-\frac{\hat{\mathcal{W}}_L(s^t)+\xi(s^t)U^m_{LL}(s^t)\frac{1}{A(s_t)}}{\hat{\mathcal{W}}_C(s^t)+\xi(s^t)\chi(s^{t-1})U^m_{CC}(s^t)\left(\frac{y^f(s^t)}{Y(s^t)}\right)^{-1/\rho}}=\frac{Y(s^t)}{L(s^t)},\qquad \forall s^t\in S^t. \tag{155}$$

Proof. The planner's problem is the same as in the proof of Proposition 10, with the exception of maximizing $\hat{\mathcal{W}}(s^t)$ instead of $\mathcal{W}(s^t)$. Therefore, following the same procedure as in the proof of Proposition 10, we obtain the expression in (155). □

C.6 Proof of Theorem 5

The Ramsey optimum satisfies (155). Substituting in our expressions for $\hat{\mathcal{W}}_C(s^t)$ and $\hat{\mathcal{W}}_L(s^t)$ we get:

$$-\frac{\mathcal{W}_L(s^t)-\bar{\vartheta}\left(\sum_{i\in I}\pi^i\nu^i\sigma^i\right)\left[U^m_L(s^t)+U^m_{LL}(s^t)L(s^t)\right]+\xi(s^t)U^m_{LL}(s^t)\frac{1}{A(s_t)}}{\mathcal{W}_C(s^t)-\bar{\vartheta}\chi(s^{t-1})\frac{\rho}{\rho-1}\left(\sum_{i\in I}\pi^i\nu^i\sigma^i\right)\left[U^m_C(s^t)+U^m_{CC}(s^t)C(s^t)\right]+\xi(s^t)\chi(s^{t-1})U^m_{CC}(s^t)\left(\frac{y^f(s^t)}{Y(s^t)}\right)^{-1/\rho}}=\frac{Y(s^t)}{L(s^t)}$$

Next, substituting in the expressions for $\mathcal{W}_C(s^t)$ and $\mathcal{W}_L(s^t)$ from (96) and (97), we obtain the following Ramsey optimality condition:

$$-\frac{U^m_L(s^t)}{U^m_C(s^t)}\left\{\frac{\sum_{i\in I}\pi^i\omega^i_L(\varphi,s_t)\left[\frac{\lambda^i}{\varphi^i}+\nu^i(1+\eta)\right]+\xi(s^t)\frac{U^m_{LL}(s^t)}{U^m_L(s^t)}\frac{1}{A(s_t)}-(1+\eta)\bar{\vartheta}\left(\sum_{i\in I}\pi^i\nu^i\sigma^i\right)}{\sum_{i\in I}\pi^i\omega^i_C(\varphi)\left[\frac{\lambda^i}{\varphi^i}+\nu^i(1-\gamma)\right]+\xi(s^t)\chi(s^{t-1})\frac{U^m_{CC}(s^t)}{U^m_C(s^t)}\left(\frac{y^f(s^t)}{Y(s^t)}\right)^{-1/\rho}-\chi(s^{t-1})(1-\gamma)\bar{\vartheta}\frac{\rho}{\rho-1}\left(\sum_{i\in I}\pi^i\nu^i\sigma^i\right)}\right.$$

Therefore the optimal monetary wedge, as defined in (38), satisfies:

$$1-\tau^*_M(s^t)=\frac{\mathcal{H}(s^{t-1})+\xi(s^t)\frac{U^m_{CC}(s^t)}{U^m_C(s^t)}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho}-(1-\gamma)\bar{\vartheta}\frac{\rho}{\rho-1}\sum_{i\in I}\pi^i\nu^i\sigma^i}{\mathcal{I}(s_t)+\xi(s^t)\frac{U^m_{LL}(s^t)}{U^m_L(s^t)}\frac{1}{A(s_t)}-(1+\eta)\bar{\vartheta}\sum_{i\in I}\pi^i\nu^i\sigma^i}. \tag{156}$$

where we let $\mathcal{H}(s^{t-1})\equiv\chi(s^{t-1})^{-1}\Omega_C>0$.

First, note that when $\xi(s^t)=0$, the constraint is slack. Therefore, it is clear that $\tau^*_M(s^t)=0$ if and only if

$$\mathcal{I}(s_t)=\hat{\mathcal{I}}(s^{t-1})\equiv\mathcal{H}(s^{t-1})+\bar{\vartheta}\left[(1+\eta)-(1-\gamma)\frac{\rho}{\rho-1}\right]\sum_{i\in I}\pi^i\nu^i\sigma^i$$

Next we use the representation of the monetary tax found in (104) and repeated here:

$$A(s_t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho}=(1-\tau^*_M(s^t))\frac{Y(s^t)}{L(s^t)}$$

Substituting the optimal monetary wedge from (156) into (104) we obtain:

$$A(s_t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho}=\left\{\frac{\mathcal{H}(s^{t-1})+\xi(s^t)\frac{U^m_{CC}(s^t)}{U^m_C(s^t)}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho}-(1-\gamma)\bar{\vartheta}\frac{\rho}{\rho-1}\sum_{i\in I}\pi^i\nu^i\sigma^i}{\mathcal{I}(s_t)+\xi(s^t)\frac{U^m_{LL}(s^t)}{U^m_L(s^t)}\frac{1}{A(s_t)}-(1+\eta)\bar{\vartheta}\sum_{i\in I}\pi^i\nu^i\sigma^i}\right\}\frac{Y(s^t)}{L(s^t)}.$$

Rearrangement, yields:

$$1 = \frac{\mathcal{H}(s^{t-1})Y(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1/\rho} + \xi(s^t)\frac{U^m_{CC}(s^t)C(s^t)}{U^m_C(s^t)} - (1-\gamma)\bar{\vartheta}\frac{\rho}{\rho-1}\sum_{i\in I}\pi^i\nu^i\sigma^i Y(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1/\rho}}{\mathcal{I}(s_t)A(s_t)L(s^t) + \xi(s^t)\frac{U^m_{LL}(s^t)L(s^t)}{U^m_L(s^t)} - (1+\eta)\bar{\vartheta}\sum_{i\in I}\pi^i\nu^i\sigma^i A(s_t)L(s^t)}.$$

which reduces to:

$$\begin{aligned} 0 =& \mathcal{I}(s_t) + (\eta+\gamma)\frac{\xi(s^t)}{A(s_t)L(s^t)} - \mathcal{H}(s^{t-1})\frac{Y(s^t)}{A(s_t)L(s^t)}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1/\rho} \\ & - \left\{(1+\eta) - (1-\gamma)\frac{\rho}{\rho-1}\frac{Y(s^t)}{A(s_t)L(s^t)}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1/\rho}\right\}\bar{\vartheta}\sum_{i\in I}\pi^i\nu^i\sigma^i \end{aligned}$$

As before, we define

$$\hat{\xi}(s^t) \equiv \frac{\xi(s^t)}{A(s_t)L(s^t)}\mathcal{H}(s^{t-1})^{-1}$$

We have that:

$$\mathcal{I}(s_t)+(\eta+\gamma)\mathcal{H}(s^{t-1})\hat{\xi}(s^t)-\mathcal{H}(s^{t-1})\frac{Y(s^t)}{A(s_t)L(s^t)}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1/\rho}-\left\{(1+\eta)-(1-\gamma)\frac{\rho}{\rho-1}\frac{Y(s^t)}{A(s_t)L(s^t)}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1/\rho}\right\}\bar{\vartheta}\sum_{i\in}$$

Next, using condition (104) we have the following optimality condition:

$$\mathcal{I}(s_t)+(\eta+\gamma)\mathcal{H}(s^{t-1})\hat{\xi}(s^t)-\mathcal{H}(s^{t-1})(1-\tau_M(s^t))^{-1}-\left[(1+\eta)-(1-\gamma)\frac{\rho}{\rho-1}(1-\tau_M(s^t))^{-1}\right]\bar{\vartheta}\sum_{i\in I}\pi^i\nu^i\sigma^i = 0$$

We let $\hat{g}$ be the function defined by:

$$\begin{aligned} \hat{g}(\mathcal{I}(s_t),\tau_M(s^t)) \equiv& \mathcal{I}(s_t) + \mathcal{H}(s^{t-1})\left[(\eta+\gamma)\hat{\xi}(s^t) - (1-\tau_M(s^t))^{-1}\right] \\ & - (1+\eta)\bar{\vartheta}\sum_{i\in I}\pi^i\nu^i\sigma^i + (1-\gamma)\frac{\rho}{\rho-1}\bar{\vartheta}\sum_{i\in I}\pi^i\nu^i\sigma^i(1-\tau_M(s^t))^{-1}. \end{aligned}$$

Therefore, the optimal monetary tax satisfies: $g(\mathcal{I}(s_t),\tau^*_M(s^t)) = 0$. By the implicit function theorem:

$$\frac{d\tau^*_M(s^t)}{d\mathcal{I}(s_t)} = -\frac{d\hat{g}/d\mathcal{I}(s_t)}{d\hat{g}/d\tau^*_M(s^t)} = -\frac{1}{\mathcal{H}(s^{t-1})\left\{(\eta+\gamma)\frac{d\hat{\xi}(s^t)}{d\tau_M(s^t)} - (1-\tau^*_M(s^t))^{-2}\right\} + (1-\gamma)\frac{\rho}{\rho-1}\bar{\vartheta}\sum_{i\in I}\pi^i\nu^i\sigma^i(1-\tau^*_M(s^t))^{-}}$$

Therefore the derivative of the optimal monetary tax satisfies:

$$\frac{d\tau^*_M(s^t)}{d\mathcal{I}(s_t)} = \frac{1}{\mathcal{H}(s^{t-1})}\left\{\left[1-(1-\gamma)\mathcal{H}(s^{t-1})^{-1}\frac{\rho}{\rho-1}\bar{\vartheta}\sum_{i\in I}\pi^i\nu^i\sigma^i\right](1-\tau^*_M(s^t))^{-2} - (\eta+\gamma)\frac{d\hat{\xi}(s^t)}{d\tau_M(s^t)}\right\}^{-1} \tag{157}$$

An expression for $\hat{\xi}(s^t)$. Following the same steps as before, we get the following expression for $\hat{\xi}(s^t)$:

$$\hat{\xi}(s^t) = \rho \frac{\mathcal{H}(s^{t-1})^{-1}}{\chi(s^{t-1})U_C^m(s^t)} \varsigma^Y(s^t) \left[\frac{(1-\kappa)(1-\epsilon(s^t))}{\kappa\epsilon(s^t)^{-\rho} + (1-\kappa)} \right] \tag{158}$$

Next we use the planner's optimality condition:

$$\varsigma^Y(s^t) = \hat{\mathcal{W}}_C(s^t) + \xi(s^t)\chi(s^{t-1}) \left[\frac{y^f(s^t)}{Y(s^t)} \right]^{-1/\rho} \frac{U_C^m(s^t)}{Y(s^t)} \left[\frac{U_{CC}^m(s^t)C(s^t)}{U_C^m(s^t)} + \frac{1}{\rho} \right]$$

Therefore

$$\varsigma^Y(s^t) = \hat{\mathcal{W}}_C(s^t) \frac{\kappa\epsilon(s^t)^{1-\rho} + (1-\kappa)}{\kappa\epsilon(s^t)^{1-\rho} + (1-\gamma\rho)(1-\kappa)\epsilon(s^t) + \gamma\rho(1-\kappa)}$$

Substituting this expression for $\varsigma^Y(s^t)$ into (156), we obtain the following expression for $\hat{\xi}(s^t)$ as a function of $\epsilon(s^t)$:

$$\hat{\xi}(s^t) = \rho \frac{\mathcal{H}(s^{t-1})^{-1}}{\chi(s^{t-1})U_C^m(s^t)} \hat{\mathcal{W}}_C(s^t)\Sigma(\epsilon)(1-\tau_M(s^t))^{-1}$$

where $\Sigma(\epsilon)$ is defined in (143) and $\tau_M(s^t)$ satisfies equation (114). Furthermore:

$$\hat{\mathcal{W}}_C(s^t) = \mathcal{W}_C(s^t) - \bar{\vartheta}\chi(s^{t-1})\frac{\rho}{\rho-1}\left(\sum_{i\in I}\pi^i\nu^i\sigma^i\right)[U_C^m(s^t) + U_{CC}^m(s^t)C(s^t)]$$

Thus

$$\hat{\xi}(s^t) = \rho \frac{\mathcal{H}(s^{t-1})^{-1}}{\chi(s^{t-1})} \left[\frac{\mathcal{W}_C(s^t)}{U_C^m(s^t)} - \bar{\vartheta}\chi(s^{t-1})\frac{\rho}{\rho-1}\left(\sum_{i\in I}\pi^i\nu^i\sigma^i\right) \frac{U_C^m(s^t) + U_{CC}^m(s^t)C(s^t)}{U_C^m(s^t)} \right] \Sigma(\epsilon)(1-\tau_M(s^t))^{-1}$$

Using the fact that $\mathcal{W}_C(s^t) = U_C^m(s^t)\Omega_C(\varphi)$ and $\mathcal{H}(s^{t-1}) \equiv \chi(s^{t-1})^{-1}\Omega_C$ the above reduces to:

$$\hat{\xi}(s^t) = \rho \left[1 - (1-\gamma)\mathcal{H}(s^{t-1})^{-1}\bar{\vartheta}\frac{\rho}{\rho-1}\sum_{i\in I}\pi^i\nu^i\sigma^i \right] \Sigma(\epsilon)(1-\tau_M(s^t))^{-1} \tag{159}$$

Derivative of $\tau_M(s^t)$. The derivative of the optimal monetary "tax" satisfies (157). Evaluating this derivative at the benchmark in which $\tau_M(s^t) = 0$, we have:

$$\left.\frac{d\tau_M^*(s^t)}{d\mathcal{I}(s_t)}\right|_{\tau_M(s^t)=0} = \frac{1}{\mathcal{H}(s^{t-1})} \left\{ \left[1 - (1-\gamma)\mathcal{H}(s^{t-1})^{-1}\bar{\vartheta}\frac{\rho}{\rho-1}\sum_{i\in I}\pi^i\nu^i\sigma^i \right] - (\eta+\gamma) \left.\frac{d\hat{\xi}(s^t)}{d\tau_M(s^t)}\right|_{\tau_M(s^t)=0} \right\}^{-1} \tag{160}$$

where $\hat{\xi}(s^t)$ satisfies (159). Taking the first derivative of the expression in (159), we get:

$$\frac{d\hat{\xi}}{d\tau_M} = \rho \left[1 - (1-\gamma)\mathcal{H}(s^{t-1})^{-1}\bar{\vartheta}\frac{\rho}{\rho-1}\sum_{i\in I}\pi^i\nu^i\sigma^i \right] \left\{ \Sigma(\epsilon)(1-\tau_M)^{-2} + (1-\tau_M)^{-1}\frac{d\Sigma}{d\epsilon}\frac{d\epsilon}{d\tau_M} \right\}$$

Following the same procedure as the proof of Theorem 3 we have that:

$$\Sigma(\epsilon)(1-\tau_M)^{-2} + (1-\tau_M)^{-1}\frac{d\Sigma}{d\epsilon}\frac{d\epsilon}{d\tau_M} = \frac{(1-\kappa)(1-\epsilon)}{\kappa\epsilon^{1-\rho} + (1-\gamma\rho)(1-\kappa)\epsilon + \gamma\rho(1-\kappa)}(1-\tau_M)^{-2}$$
$$- (1-\tau_M)^{-1}(1-\kappa)\left[\frac{\kappa\epsilon^{1-\rho} + (1-\rho)\kappa\epsilon^{-\rho}(1-\epsilon) + (1-\kappa)}{(\kappa\epsilon^{1-\rho} + (1-\gamma\rho)(1-\kappa)\epsilon + \gamma\rho(1-\kappa))^2}\right]\left[\frac{\quad}{(1-\rho)(}\right.$$

Evaluating this derivative at $\tau_M = 0$, or equivalently at $\epsilon = 1$, we get:

$$\left.\frac{d\hat{\xi}}{d\tau_M}\right|_{\tau_M=0} = -\rho\frac{1-\kappa}{\kappa}\left[1 - (1-\gamma)\mathcal{H}(s^{t-1})^{-1}\bar{\vartheta}\frac{\rho}{\rho-1}\sum_{i\in I}\pi^i\nu^i\sigma^i\right]$$

Substituting this expression into (160), we obtain:

$$\left.\frac{d\tau_M^*(s^t)}{d\mathcal{I}(s_t)}\right|_{\tau_M(s^t)=0} = \frac{1}{\mathcal{H}(s^{t-1})}\left\{(1+\rho(\eta+\gamma)\frac{1-\kappa}{\kappa})\left[1-(1-\gamma)\mathcal{H}(s^{t-1})^{-1}\bar{\vartheta}\frac{\rho}{\rho-1}\sum_{i\in I}\pi^i\nu^i\sigma^i\right]\right\}^{-1}$$

which can be rewritten as:

$$\left.\frac{d\tau_M^*(s^t)}{d\mathcal{I}(s_t)}\right|_{\tau_M(s^t)=0} = \frac{1}{(1+\rho(\eta+\gamma)\frac{1-\kappa}{\kappa})\left[\mathcal{H}(s^{t-1}) + (\gamma-1)\bar{\vartheta}\frac{\rho}{\rho-1}\sum_{i\in I}\pi^i\nu^i\sigma^i\right]}$$

Therefore, taking a first-order Taylor approximation of $\tau_M^*(s^t)$ around the point $\mathcal{I}(s_t) = \hat{\mathcal{I}}(s^{t-1})$, we have:

$$\tau_M^*(s^t) \approx 0 + \frac{1}{(1+\rho(\eta+\gamma)\frac{1-\kappa}{\kappa})\left[\mathcal{H}(s^{t-1}) + (\gamma-1)\bar{\vartheta}\frac{\rho}{\rho-1}\sum_{i\in I}\pi^i\nu^i\sigma^i\right]}[\mathcal{I}(s_t) - \hat{\mathcal{I}}(s^{t-1})],$$

which coincides with the expression in (49).