Optimal Monetary Policy with Redistribution*

Jennifer La'O† Wendy Morrison‡

May 5, 2025

Abstract

We study optimal monetary policy in a general equilibrium economy with heterogeneous agents and nominal rigidities. Households differ in type-specific, state-contingent labor productivity and initial firm ownership, yet markets are complete. The fiscal authority has access to a linear tax schedule with non-state-contingent tax rates and uniform, lump-sum transfers. We show that when there are fluctuations in the labor skill distribution, it is optimal for monetary policy to target a state-contingent markup that covaries positively with a sufficient statistic for labor income inequality. In a calibrated version of the model, countercyclical earnings inequality implies countercyclical optimal markups. Therefore, to the extent that redistribution—and not insurance—is the distributional policy goal, optimal monetary policy is output destabilizing.

Keywords: business cycles, inequality, redistribution, monetary policy, fiscal policy, optimal taxation.

JEL codes: E32, E52, E63, D61, D63, H21

---

*We thank Joao Guerreiro, Paolo Martellini, Victor Rios-Rull, and Elisa Rubbo for their insightful and constructive discussions of our paper.

The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

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

‡Federal Reserve Bank of Minneapolis. wendy.morrison@mpls.frb.org.

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 earnings of low-income households exhibit a greater covariance with aggregate fluctuations than those of mid- and high-income households.[1] This heterogeneity in the covariance of individual earnings with aggregate fluctuations generates countercyclical earnings inequality.

Monetary policy is not an ideal tool for achieving distributional objectives—questions of redistribution typically fall within the purview of the fiscal authority. Yet countercyclical earnings inequality calls this division of labor into question: if fiscal instruments cannot respond to short-run movements in the labor earnings distribution, monetary instruments might be a suitable alternative.

How should monetary policy be conducted in light of distributional concerns? Several recent papers tackle this question within the heterogeneous-agent New Keynesian (HANK) framework (Bhandari et al., 2021; Acharya et al., 2023; Nuño and Thomas, 2022; Dávila and Schaab, 2023; McKay and Wolf, 2022; Le Grand et al., 2024). In HANK, households are ex ante identical but face uninsurable, idiosyncratic labor income risk. The lack of insurance and consequent precautionary savings behavior lead to *ex post* differences in wealth across households and consumption inequality. In this context, monetary policy can be used to compensate for missing insurance markets (Bhandari et al., 2021; Acharya et al., 2023). In particular, in response to productivity shocks, optimal monetary policy places a greater weight on output stabilization in HANK than in its representative agent New Keynesian counterpart.

However, empirical studies show that differences in households' earnings exposure to business cycles are, to a large extent, systematic (Guvenen et al., 2014, 2017). That is, *ex ante* differences across households broadly determine how they fare over the business cycle: "fortunes during recessions are predictable by observable characteristics before the recession," according to Guvenen et al. (2014).[2]

Motivated by these findings, we study the optimal conduct of monetary policy in a setting that features systematic, or ex ante, household heterogeneity. In our framework, heterogeneous household labor productivities fluctuate at business cycle frequency, with some households more exposed to aggregate shocks than others—yet markets are complete. Although households can fully insure their consumption against idiosyn-

[1] Parker and Vissing-Jorgensen (2009); Guvenen et al. (2014, 2017); Alves et al. (2020).

[2] Furthermore, a number of studies find that households are able to smooth their consumption in response to income shocks to a considerable degree (Cochrane, 1991; Schulhofer-Wohl, 2011; Guvenen and Smith, 2014; Heathcote et al., 2014).

cratic income risk, short run movements in the labor income distribution introduce a novel redistributive role for monetary policy.

In particular, we show that if a utilitarian planner desires greater redistribution than under laissez faire, then monetary policy should target a countercyclical markup. Specifically, optimal monetary policy raises the markup—equivalently, lowers the real wage—in high income inequality states, and conversely lowers the markup—i.e. raises the real wage—in low income inequality states. If labor earnings inequality is countercyclical, as it is in the data, then the optimal markup is countercyclical.

Optimal policy in our model is directly at odds with that in the HANK model. In HANK, a negative productivity shock decreases household income but leaves nominal debt obligations unchanged, hurting borrowers more than lenders. Bhandari et al. (2021) and Acharya et al. (2023) show that in response to productivity shocks, monetary policy should prevent output from fluctuating one-for-one with its productively-efficient level. In particular, monetary policy can provide greater insurance and improve welfare by lowering the real interest rate and stimulating the economy in downturns, and raising the rate and contracting the economy in upturns. The insurance motive in HANK thereby pushes monetary policy towards output stabilization.

When we remove this insurance motive and isolate a purely redistributive one, we find that optimal monetary policy is instead output destabilizing. This may seem counterintuitive, yet our main result reflects a simple but fundamental insight from public finance: that a linear labor income tax with a positive slope redistributes income from high-skilled, high-earning households to low-skilled, low-earning households. The optimal marginal tax rate trades off equity with efficiency and, in the context of a general equilibrium business cycle model, the labor skill distribution is a key determinant of the optimal tax rate (Werning, 2007).

Our contribution is to apply this simple but powerful insight to monetary policy analysis. By lowering or raising the markup in response to movements in the earnings distribution, monetary policy can essentially manipulate the marginal tax rate and alter the distribution of lifetime income. The surprising implication is that, insofar as greater redistribution is desirable, optimal monetary policy should contract the economy (raise the markup) in downturns when earnings inequality is high, and stimulate the economy (lower the markup) in upturns when earnings inequality is low.

The stark discrepancy between this policy prescription and that in HANK is rooted in the fact that heterogeneity is driven by fundamentally different sources in the two models (ex ante versus ex post), generating different roles for monetary policy (redistribution versus insurance). Given that both ex ante and ex post heterogeneity are present in the data, our paper underscores the potential difficulty of achieving distributional goals with monetary policy and the importance of accounting for the sources of heterogeneity in earnings fluctuations before using monetary policy as a distributional tool.

Framework and Methodology. Our framework is a general equilibrium, heterogeneous agent economy with nominal rigidities. Households are assigned a "type" at birth and remain that type throughout their lifetime. Types map to heterogeneous labor productivities and initial (time-0) firm ownership.

Type-specific labor productivities are state-contingent. We allow these contingencies to be fully general, thereby nesting any exogenous labor income process. We assume that markets are complete: in every period, households can trade a complete set of Arrow securities, a risk-free nominal bond, and firm equity.

A continuum of intermediate-good firms employ workers, produce differentiated goods, and face aggregate productivity shocks. Firms are monopolistically-competitive and set prices subject to a nominal rigidity. In particular, we assume that in each period a fraction of firms must set their prices before observing demand.

A consolidated government sets fiscal and monetary policy under commitment. The government raises tax revenue and issues debt in order to finance lump-sum transfers. We follow the Ramsey approach and allow for linear taxes on consumption, labor income, firm sales, and profits. Monetary policy is fully state-contingent while tax rates are fixed and set optimally at time 0. The lack of fiscal state-contingency, a typical assumption in the New Keynesian literature, can be thought of as a political constraint: the fiscal authority cannot adjust tax rates at business cycle frequency. We furthermore allow for state-contingent lump-sum taxes or transfers, restricted to be uniform across households. Finally, in our baseline model we assume that firm profits are fully taxed—this is not a crucial assumption, and we relax it in an extension.

We adopt a utilitarian welfare function with arbitrary Pareto weights and solve the Ramsey problem using the primal approach (Atkinson and Stiglitz, 1980; Lucas and Stokey, 1983; Chari and Kehoe, 1999). In particular, we adapt the primal approach in Niepelt (2004), Werning (2007), and Bassetto (2014) for flexible-price economies with heterogeneous agents and complete markets, and that in Correia et al. (2008) and Angeletos and La'O (2020) for representative agent economies with nominal rigidities, to our setting that features heterogeneous agents, complete markets, and nominal rigidities. Our main result characterizes fiscal and monetary policy at the Ramsey optimum.

Main Results. We find that when shocks to the labor skill distribution are proportional—that is, when *relative* productivities across types are fixed—the Ramsey optimum can be implemented under flexible prices with the available set of fiscal instruments. Optimal monetary policy replicates flexible-price allocations; it can do so by targeting price stability.

When instead relative productivities vary over the business cycle, it is optimal for monetary policy to deviate from replicating flexible-price allocations and instead target a state-contingent markup. The optimal markup co-varies positively with a sufficient statistic for labor income inequality. Specifically, monetary policy should raise the markup, i.e. lower the real wage, when labor income inequality is high, and conversely lower the markup, i.e. raise the real wage, when income inequality is low.

To understand this result, it is helpful to think of monetary policy as mimicking a state-contingent labor income tax. If the final good price rises above its marginal cost (the wage), it is *as if* households pay an implicit labor income tax; conversely, if the price falls below its marginal cost, it is as if households receive a subsidy. Raising or lowering the markup is thereby similar to manipulating the marginal tax rate of a linear income tax schedule, albeit with an accompanying loss in production efficiency due to price dispersion. A linear income tax is moreover a redistributive tool: although all households face the same marginal tax rate, high income households face higher average tax rates than low income households and therefore a higher tax burden.

The optimal markup (or tax) balances distributional objectives with efficiency. In states with high labor income inequality—states in which high-skilled households earn comparatively more labor income than low-skilled households relative to other states—the marginal redistributional benefit of the markup increases. Conversely, in states with low labor income inequality—states in which the skill distribution is more compressed—the marginal redistributional benefit of the markup falls. Homothetic and separable preferences meanwhile ensure that the distortionary cost of the markup remains constant across states. It follows that the optimal markup rises in high inequality states and falls in low inequality states; in so doing, optimal monetary policy compresses the distribution of lifetime earnings.

Extensions. In our baseline economy, tax rates are non-state-contingent and profits are fully taxed. We consider three extensions in which we relax these restrictions on fiscal policy. The first extension allows tax rates to be partially state-contingent: in particular, we assume taxes are set one period in advance. To the extent that shocks are persistent, fiscal policy can absorb some of the state-contingent pressure placed on monetary policy. In this case we obtain an explicit, closed-form solution for the derivative of the optimal markup with respect to the sufficient statistic for labor income inequality. We show that this slope is strictly positive, and we provide comparative statics for how the slope depends on model parameters.

In a second extension, we assume that profits cannot be fully taxed. With less than full profit taxation, heterogeneity in initial firm ownership introduces an additional channel by which monetary policy affects the distribution of permanent income. Initial equity is a lifetime claim to firm profits. Hence, by shifting firm markups, monetary policy affects equilibrium profits and, by extension, the distribution of financial wealth. Despite this additional channel, our main qualitative result on the optimal conduct of monetary policy remains intact: the optimal markup covaries positively with the same sufficient statistic for labor income inequality.

In a third extension, we allow for tax rates to be set one period in advance and we assume profits are not fully taxed. We find that the derivative of the optimal markup with respect to labor income inequality remains positive. However, the extent to which profits are taxed and the cross-sectional covariance between financial wealth and lifetime labor earnings affect the magnitude of this slope.

Quantitative Illustration. In a simple, quantitative illustration of the model, we calibrate the labor income distribution to match the "worker betas"—the percent change in household labor income growth associated with a percent change in GDP growth—estimated in Guvenen et al. (2017). In this way, business cycle movements in the earnings distribution in the calibrated model directly reflect the unequal incidence of GDP fluctuations documented in the data.

We find that the optimal markup is countercyclical. In our baseline calibration, the elasticity of the optimal markup with respect to real GDP ranges from –.25 to –.39. The countercyclicality of the optimal markup stems from two features: countercyclical earnings inequality as documented in the data, and our main theoretical result that the optimal markup covaries positively with earnings inequality.

The behavior of the optimal markup in our calibrated model—that it rises in recessions and falls in expansions—implies a destabilizing monetary policy. Although a destabilizing stance for monetary policy may be counterintuitive, the behavior of the optimal markup in our calibrated model is consistent with empirical work that documents countercyclical price markups (Bils, 1987; Rotemberg and Woodford, 1999; Bils, Klenow and Malin, 2018) and, more generally, a countercyclical labor wedge (Hall, 1997; Chari, Kehoe and McGrattan, 2007).[3]

Related Literature. The contribution of this paper is to characterize optimal monetary policy when the key motive is redistribution. To this end, we import fundamental

[3]Countercyclical markups are, however, inconsistent with Nekarda and Ramey (2020) who document procyclical markups.

concepts and methods from public finance and adapt these to the study of optimal monetary policy.

The theory of optimal income taxation captures the classic equity-efficiency trade-off. The government must use a distortionary income tax to finance spending and redistribute income. While the typical treatment considers non-linear tax schedules (Mirrlees, 1971), the linear tax case (with an intercept) has received generous attention for its tractability and its connection to the theory of optimal commodity taxation (Ramsey, 1927; Diamond and Mirrlees, 1971; Sheshinski, 1972; Atkinson and Stiglitz, 1980).

We focus on the linear tax case because, as we show in this paper, its lessons carry over to monetary policy. With a linear labor income tax, a positive marginal tax rate implies a higher tax burden on high-income households than on low-income households. When used to finance transfers, a positive tax rate lowers inequality.

In this sense the closest antecedent to our paper is Werning (2007).[4] Werning (2007) studies optimal fiscal policy in a general equilibrium, business cycle model with heterogeneous agents, complete markets, and a linear tax schedule, when a key motive for distortionary taxation is redistribution. Werning (2007) shows that a primary determinant of the optimal tax rate is the labor skill distribution: in response to shocks to the skill distribution, the optimal tax rate rises with labor earnings inequality.

We apply these fundamental insights in public finance to monetary policy. In order to do so, we adopt the primal approach to optimal taxation (Atkinson and Stiglitz, 1980; Chari and Kehoe, 1999) and adapt it to our setting that features nominal rigidites, thereby building on the contributions of Correia et al. (2008), Correia et al. (2013), and Angeletos and La'O (2020).[5]

Using the primal approach, we show how monetary policy resembles a linear tax schedule. This is because monetary policy can generate state-contingent markups which, in allocation space, translate to state-contingent labor wedges with an accompanying loss in production efficiency. The "linearity" comes from the fact that all households face the same markup, i.e. the same marginal tax rate. Optimal monetary policy targets a state-contingent pattern for the optimal markup that resembles the optimal tax rate in Werning (2007). The surprising implication is that, to the extent that labor earnings inequality is countercyclical, optimal monetary policy is destabilizing.

In short, although a vast public finance literature has shown how distortionary taxation can be used to balance redistribution with efficiency, we believe our paper is the first to apply these lessons to monetary policy. If the labor earnings distribution moves

[4] Closely related are Niepelt (2004) and Bassetto (2014). Early contributions on the Ramsey problem with heterogeneous agents include Judd (1985) and Chari and Kehoe (1999).

[5] See also Dávila and Schaab (2023) for a variation of the traditional primal approach: in their primal form, prices and allocations are explicit control variables for the planner.

at business cycle frequency, and if monetary policy can respond more rapidly to short run fluctuations than fiscal policy, then there could be room to use monetary policy as a distributional tool.

A similar justification underlies the distributional role of optimal monetary policy in HANK. Closest to our paper, in this respect, is Bhandari et al. (2021). Bhandari et al. (2021) study optimal policy in a sticky price model with labor income risk and incomplete markets. Households are subject to idiosyncratic and aggregate shocks, can trade only nonstate-contingent nominal debt, and face a linear tax schedule. Like us, Bhandari et al. (2021) solve a Ramsey problem in which monetary policy and transfers can adjust to shocks while tax rates are kept constant (but chosen optimally).[6]

Bhandari et al. (2021) show that an insurance motive shapes the planner's response to shocks. A negative productivity shock decreases household income but leaves nominal debt obligations unchanged, hurting borrowers more than lenders. In response, it is optimal for monetary policy to lower the real interest rate and, in so doing, redistribute wealth from lenders to borrowers. In terms of aggregates, by lowering the nominal rate more than 1-for-1 with the natural rate, the monetary authority tolerates some inflation and prevents output from falling 1-for-1 with its flexible-price counterpart.

A similar finding is present in Acharya et al. (2023). Acharya et al. (2023) analyze a sticky price model with labor income risk and incomplete markets; in order to obtain analytical results, households have constant absolute risk aversion utility and idiosyncratic shocks are normally distributed. Acharya et al. (2023) show that the insurance motive of monetary policy leads to a greater emphasis on output stabilization. In their model, this is driven by two forces. The first is facilitation of greater household self-insurance. The second is an income-risk channel: if the level of idiosyncratic income risk is directly tied to the level of output, and in particular if idiosyncratic risk is countercyclical, then stabilizing output leads to a reduction in consumption risk.

The distributional welfare gains of monetary policy in HANK come from its ability to compensate for missing insurance markets. In contrast, we assume that markets are complete. We thus remove entirely the insurance motive for monetary policy and focus solely on the redistributive motive—this explains the discrepancy in our results.

Using large, panel data on individual earnings from the US Social Security Administration, Guvenen et al. (2014) and Guvenen et al. (2017) find that systematic, forecastable, between-group variation accounts for a large share of the total variation in earnings growth over the business cycle. In fact, earnings profiles are "consistent with countercyclical earnings inequality without needing within-group (idiosyncratic)

[6]Fixed tax rates are a common assumption in New Keynesian models. Bhandari et al. (2021) solve a second Ramsey planner's problem in which the planner has more tools and can adjust tax rates. Similarly in Section 5 we consider an extension of our model in which tax rates are partially state-contingent.

shocks that have countercyclical variances," (Guvenen et al., 2014). While no consensus exists on the exact share of variation in lifetime earnings accounted for by systematic heterogeneity—as well as the "insurable" component of labor income shocks—a number of structural estimations place it well above 50 percent and some as high as 90 percent.[7] This evidence motivates the focus of our paper.

Layout. In Section 2 we describe the economic environment and in Section 3 we characterize equilibrium allocations. In Section 4 we solve the Ramsey problem and characterize optimal monetary policy. In Section 5 we study optimal policy in three extensions of the baseline model. In Section 6 we explore quantitative implications of a calibrated version of the model. Section 7 concludes. All proofs, except for those explicitly provided in the text, are found in the Appendix.

2 The Environment

We study a general equilibrium model with heterogeneous agents and a form of nominal rigidity. Time is discrete, indexed by $t = 0, 1, \ldots, \infty$. 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 s_t. We let $\mu(s^t|s^{t-1})$ denote the probability of history s^t conditional on s^{t-1}, and with slight abuse of notation we let $\mu(s^t)$ denote the unconditional probability of history 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:

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

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 is given by:

[7] See Keane and Wolpin (1997); Storesletten et al. (2004); Huggett et al. (2011); Heathcote et al. (2014); Guvenen and Smith (2014).

$\ell^i(s^t) = \theta^i(s_t)h^i(s^t)$. The household maximizes lifetime expected utility:

$$\sum_t \sum_{s^t} \beta^t \mu(s^t) U(c^i(s^t), \ell^i(s^t)/\theta^i(s_t)). \tag{2}$$

Household i's nominal budget constraint at time t, history s^t is given by:

$$\begin{aligned}(1+\tau_c)P(s^t)c^i(s^t) + b^i(s^t) + \sum_{s^{t+1}|s^t} Q(s^{t+1}|s^t)z^i(s^{t+1}|s^t) + V(s^t)[\sigma^i(s^t) - \sigma^i(s^{t-1})] \\ \leq (1-\tau_\ell)W(s^t)\ell^i(s^t) + (1-\tau_\Pi)(1+\sigma^i(s^{t-1}))\Pi(s^t) + z^i(s^t|s^{t-1}) + (1+i(s^{t-1}))b^i(s^{t-1}) \\ + P(s^t)T(s^t).\end{aligned} \tag{3}$$

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 income tax rates, τ_c and τ_ℓ, and receives a real lump-sum transfer, $T(s^t)$, that can depend on the history s^t. The latter can be positive or negative (a tax), and is uniform across types.

The household can borrow and save via three separate instruments. The first is a one-period, non-state-contingent nominal bond, $b^i(s^t)$, which the household can buy or sell at time t, history s^t, and which pays $(1+i(s^t))b^i(s^t)$ units of money one period later. The second is a complete set of state-contingent Arrow securities, indexed by $s^{t+1}|s^t$. 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 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 nominal bond is a redundant asset but it allows us to represent the one-period interest rate, $i(s^t)$. We assume that initial wealth from bond holdings is zero: $b_0^i = 0$ for all $i \in I$.

The third instrument is equity: the household can buy and sell shares of a fully diversified portfolio of firms. Equity ownership is a claim to aggregate firm profits, denoted in nominal terms by $\Pi(s^t)$ and taxed at a constant rate of $\tau_\Pi \in [0,1]$. If the household enters time t, history s^t, with $1+\sigma^i(s^{t-1})$ shares, it receives dividend $(1-\tau_\Pi)\Pi(s^t)$ per share and it can trade shares at ex-dividend price $V(s^t)$. We assume that the type-i household is endowed with $1+\sigma_0^i$ shares at time 0, with $\sum_{i\in I}\pi^i\sigma_0^i = 0$.

Household's Problem. Given initial bond holdings and initial equity shares, the type-i household chooses a complete contingent plan for consumption, efficiency units of labor, bond holdings, equity holdings, and Arrow security holdings: $\{c^i(s^t), \ell^i(s^t), b^i(s^t), \sigma^i(s^t), (z^i(s^{t+1}|s^t))_{s^{t+1}}\}_{t\geq 0, s^t\in S^t}$, in order to maximize its lifetime expected utility (2) subject to its per-period budget constraint (3) for all $s^t \in S^t$ and no-Ponzi conditions.

Firms. There is a unit mass 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: $y^j(s^t) = A(s_t)n^j(s^t)$, 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 $f^j(s^t) = (1-\tau_r)p_t^j(\cdot)y^j(s^t) - W(s^t)n^j(s^t)$ where τ_r is a constant tax on firm revenue. We postpone for a moment our discussion of the nominal rigidity and the firms' optimization problem—that is, how the price $p_t^j(\cdot)$ is set.

A representative firm produces the final good with a constant elasticity of substitution (CES) production 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 elasticity of substitution $\rho > 1$. 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 and $P(s^t)$ is the price of the final good. The final good producer is perfectly competitive and takes prices as given. Profit maximization implies the standard CES demand function for good j:

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

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. We let:

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

denote nominal tax revenue collected at time t, history s^t, where $C(s^t) \equiv \sum_{i\in I}\pi^i c^i(s^t)$ denotes aggregate consumption, $L(s^t) \equiv \sum_{i\in I}\pi^i\ell^i(s^t)$ denotes aggregate labor in efficiency units, and $\Pi(s^t) \equiv \int_{j\in\mathcal{J}} f^j(s^t)\mathrm{d}j$ denotes aggregate profits. The government can issue both state-contingent and non-state-contingent debt. The government's period-t nominal budget constraint is given by:

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

where $B(s^t)$ is aggregate bond issuance and $Z(s^{t+1}|s^t)$ denotes aggregate Arrow security issuance for each $s^{t+1}|s^t$. Finally, for monetary policy we assume that the monetary authority directly controls nominal aggregate demand according to the following ad hoc, cash-in-advance constraint: $M(s^t) = P(s^t)C(s^t)$. We therefore avoid well-known issues of indeterminacy. Monetary policy is state-contingent: the monetary authority can freely choose $M(s^t) > 0$ in every history.

Market Clearing. Market clearing in the goods and labor markets imply: $C(s^t) = Y(s^t)$ and $L(s^t) = \int_{j\in J} n^j(s^t)\mathrm{d}j$. Market clearing in the financial markets imply: $B(s^t) = \sum_{i\in I}\pi^i b^i(s^t)$, $Z(s^{t+1}|s^t) = \sum_{i\in I}\pi^i z^i(s^{t+1}|s^t)$ for all $s^{t+1}|s^t$, and $\sum_{i\in I}\pi^i\sigma^i(s^t) = 0$.

The Nominal Rigidity. At each date t, Nature draws the state $s_t \in S$ according to 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 $i \in I$, as exogenous mappings from the state space to aggregate productivity and type-specific labor productivities.

Intermediate good firms are price-setters. We equate the nominal rigidity in our model with an informational friction, following e.g. Mankiw and Reis (2002) and Woodford (2003). For tractability we assume a particular specification employed by Correia, Nicolini and Teles (2008): all firms set prices in every period, but only a subset of firms are attentive to the realized, current state.

Formally, we assume that in every period a mass $\kappa \in [0, 1)$ of randomly-selected firms are inattentive, or "sticky." All other firms, of mass $1 - \kappa$, are attentive, or "flexible." We let $\mathcal{J}^s \subset \mathcal{J}$ denote the set of "sticky-price" firms and $\mathcal{J}^f \subset \mathcal{J}$ denote the set of "flexible-price" firms, with $\mathcal{J}^f = (\mathcal{J}^s)'$.

Sticky-price firms at time t are inattentive to the current state, s_t, and hence set their price based solely on their knowledge of the history of past states, s^{t-1}. We denote the price they set by $p^s_t(s^{t-1})$. The subscript t indicates that this is the price set *at time* t by the sticky-price firm, even though the price function itself is measurable in s^{t-1}.

Flexible-price firms at time t are attentive to the current state, s_t, as well as the history of past states, s^{t-1}. It follows that these firms can set their price based on knowledge of the entire history, s^t. We denote the price they set by $p^f_t(s^t)$. The subscript t similarly indicates that this is the nominal price set at time t by the flexible-price firm. However, unlike the sticky-price function, the flexible-price function is measurable in s^t.

Implicit in these 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, $j \in \mathcal{J}^s$, and which firms are flexible, $j \in \mathcal{J}^f$. Intermediate good firms make their nominal pricing decisions given their information sets: s^{t-1}if sticky, s^t if flexible. Once nominal prices are set, the aggregate state becomes common knowledge. Given intermediate good prices, the representative final good firm purchases inputs and produces the final good, and households make their consumption, savings, and effort choices. All allocations adjust so that supply equals demand and markets clear.

Given this description of the nominal rigidity, we state the problems of the two types of firms as follows. At time t, a flexible-price firm $j \in \mathcal{J}^f$ solves:

$$p^f_t(s^t) \in \arg\max_{p^j_t} \left\{ (1 - \tau_r) p^j_t y^j(s^t) - \frac{W(s^t)}{A(s_t)} y^j(s^t) \right\}$$

subject to (4). At time t, a sticky-price firm $j \in \mathcal{J}^s$ solves:

$$p_t^s(s^{t-1}) \in \arg\max_{p_t^j} \sum_{s^t|s^{t-1}} Q(s^t|s^{t-1}) \left\{ (1-\tau_r) p_t^j y^j(s^t) - \frac{W(s^t)}{A(s_t)} y^j(s^t) \right\}$$

subject to (4).

A flexible-price firm sets its nominal price so as to maximize firm profits, and it does so state-by-state. On the other hand, a sticky-price firm sets its nominal price so as to maximize its expectation, conditional on s^{t-1}, of the investors' valuation of firm profits.

The firm is owned by its investors. For this reason it weighs profits across states not only by their conditional probabilities, $\mu(s^t|s^{t-1})$, but also by the stochastic discount factor of the marginal investor. Given that markets are complete, it does not matter which household's stochastic discount factor we use—in equilibrium, marginal rates of substitution are equated across all households. By no arbitrage we can equivalently value firm profits using the Arrow prices $Q(s^t|s^{t-1})$. While this may appear obvious, later on in our analysis we verify that the equilibrium Arrow prices indeed reflect the true conditional probabilites and the appropriate pricing kernel.

Finally, note that all firms—sticky and flexible—solve a static problem. This is because every firm is free to adjust its price in every period; it follows that no firm need take into account future periods or states when setting its current price.

2.1 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 $y^j(s^t) = A(s_t) n^j(s^t)$ for all $j \in \mathcal{J}$;

$$\sum_{i \in I} \pi^i c^i(s^t) = C(s^t) = Y(s^t) = \left[\int_{j \in \mathcal{J}} y^j(s^t)^{\frac{\rho-1}{\rho}} \mathrm{d}j \right]^{\frac{\rho}{\rho-1}}; \qquad \text{and} \tag{6}$$

$$\sum_{i \in I} \pi^i \ell^i(s^t) = L(s^t) = \int_{j \in \mathcal{J}} n^j(s^t) \mathrm{d}j \tag{7}$$

for all $s^t \in S^t$.

Let $\mathcal{X}$ denote the set of feasible allocations. We are interested in the subset of feasible allocations that can be supported in equilibrium in this economy. We define an equilibrium as follows.

Definition 2. An equilibrium is an allocation x, a price system

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

a policy

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

and a set of financial market positions

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

such that: (i) in every t, s^t: $p_t^s(s^{t-1})$ solves the sticky-price firm's problem, for all $j \in \mathcal{J}^s$; the price $p_t^f(s^t)$ solves the flexible-price firm's problem, for all $j \in \mathcal{J}^f$; (ii) the aggregate price level is given by:

$$P(s^t) = \left[\kappa p_t^s(s^{t-1})^{1-\rho} + (1-\kappa) p_t^f(s^t)^{1-\rho}\right]^{\frac{1}{1-\rho}}; \tag{8}$$

(iii) for all t, s^t: prices and allocations satisfy (4) for all $j \in \mathcal{J}$; (iv) given the price system and the policy, $\{c^i(s^t), \ell^i(s^t), b^i(s^t), \sigma^i(s^t), (z^i(s^{t+1}|s^t))_{s^{t+1}}\}_{t\geq 0, s^t\in S^t}$ solves the household's problem of type i, for every $i \in I$; (v) for all t, s^t: the government budget constraint is satisfied and $M(s^t) = P(s^t)C(s^t)$; and (vi) markets clear.

Remarks on the model. *Heterogeneity with market completeness.* Household types are fixed, however household labor productivity can vary over states in a general and flexible manner characterized by the arbitrary functions $\theta^i : S \to \mathbb{R}_+$. This formulation nests all exogenous labor income processes, including those that feature a high degree of heterogeneity in the covariance of individual labor earnings with aggregate shocks. In the proceeding analysis we show that the completeness of markets implies that households fully insure themselves against idiosyncratic income risk: equilibrium household consumption varies only with aggregate consumption. Heterogeneity in consumption is therefore determined entirely "ex ante" rather than "ex post."

Lump-sum taxes and transfers. In the standard, single-agent Ramsey framework, only distorting taxes are available. Lump-sum taxes—or any combination of taxes that may replicate them—are a priori ruled out; otherwise, the first best would be attainable. When instead households are heterogeneous, Werning (2007) shows how one can incorporate a lump-sum tax without sacrificing earlier lessons from the Ramsey literature on optimal taxation.

We follow Werning (2007) and allow for a uniform lump-sum tax or transfer. Uniformity of this tax across types ensures that the first best is unattainable. One can think of the uniformity restriction as an informational constraint: the fiscal authority cannot distinguish among household types.

Lack of fiscal state-contingency. The nature of optimal monetary policy depends on the set of available fiscal instruments. We assume distortionary, linear taxation as in Ramsey. In particular, we allow for taxes on consumption, labor income, firm sales, and firm profits; we therefore do not artificially restrict *what* can be taxed.

However, in our baseline model we assume tax rates are fixed at time 0 (but can be chosen optimally). State-contingency of monetary policy but nonstate-contingency of tax rates is the typical assumption made in New Keynesian models; it is motivated by the notion that monetary policy can respond more rapidly to short-run fluctuations than fiscal policy, perhaps due to political constraints. In any case, we relax this assumption in Section 5: we solve a planner's problem in which tax rates can adjust to past states.

Throughout we assume full state-contingency the lump-sum tax or transfer. It turns out that this state-contingency is without loss of generality—markets are complete, the government can issue state-contingent debt, and the infinitely-lived households have rational expectations. It follows that households are Ricardian as in Barro (1974).

3 Equilibrium Characterization

Consider first the individual household's problem. 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 are equated across households. The Negishi (1960) characterization of competitive equilibria follows.

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

$$U^m(C(s^t), L(s^t), s_t, \varphi) \equiv \max_{(c^i(s^t), \ell^i(s^t))_{i \in I}} \sum_{i \in I} \varphi^i \pi^i U(c^i(s^t), \ell^i(s^t)/\theta^i(s_t)) \tag{9}$$

$$\textit{s.t.} \quad C(s^t) = \sum_{i \in I} \pi^i c^i(s^t), \quad \textit{and} \quad L(s^t) = \sum_{i \in I} \pi^i \ell^i(s^t). \tag{10}$$

Proof. See Appendix A.2. □

In any equilibrium there is 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

$U^m(C, L; \varphi)$.[8] Relative prices satisfy the representative household's intratemporal and intertemporal conditions:

$$-\frac{U_L^m(s^t)}{U_C^m(s^t)} = \left[\frac{1-\tau_\ell}{1+\tau_c}\right]\frac{W(s^t)}{P(s^t)}, \tag{11}$$

$$\frac{U_C^m(s^t)}{P(s^t)} = \beta(1+i(s^t))\sum_{s^{t+1}|s^t}\mu(s^{t+1}|s^t)\frac{U_C^m(s^{t+1})}{P(s^{t+1})}, \tag{12}$$

for all $s^t \in S^t$, 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 labor. Condition (11) indicates that the marginal rate of substitution between aggregate consumption and aggregate labor is equal to the after-tax real wage; condition (12) is the Euler equation corresponding to the one-period nominal bond. Furthermore, the set of Arrow prices and the ex-dividend share price satisfy, respectively:

$$Q(s^{t+1}|s^t) = \frac{\beta U_C^m(s^{t+1})}{U_C^m(s^t)}\frac{P(s^t)}{P(s^{t+1})}\mu(s^{t+1}|s^t), \quad \forall s^{t+1}|s^t; \tag{13}$$

$$V(s^t) = \sum_{s^{t+1}|s^t} Q(s^{t+1}|s^t)[(1-\tau_\Pi)\Pi(s^{t+1}) + V(s^{t+1})]. \tag{14}$$

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. Therefore equations (11)-(14) hold with U^i in place of U^m. This verifies our earlier claim that the Arrow prices appropriately reflect the investors' conditional probabilities and stochastic discount factor.

With general preferences, the unique solution to the static sub-problem in (9) 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. With the separable and iso-elastic preferences assumed in (2), the solution can be written in closed form:

$$c^i(s^t) = \omega_C^i(\varphi)C(s^t) \qquad \text{and} \qquad \ell^i(s^t) = \omega_L^i(\varphi, s_t)L(s^t), \tag{15}$$

$$\text{where} \quad \omega_C^i(\varphi) \equiv \frac{(\varphi^i)^{1/\gamma}}{\sum_{k\in I}\pi^k(\varphi^k)^{1/\gamma}} \quad \text{and} \quad \omega_L^i(\varphi, s_t) \equiv \frac{(\varphi^i)^{-1/\eta}\theta^i(s_t)^{\frac{1+\eta}{\eta}}}{\sum_{k\in I}\pi^k(\varphi^k)^{-1/\eta}\theta^k(s_t)^{\frac{1+\eta}{\eta}}}. \tag{16}$$

Individual consumption and labor are thereby proportional to their aggregates.

Household i's shares of aggregate consumption and aggregate labor are given by $\omega_C^i(\varphi)$ and $\omega_L^i(\varphi, s_t)$, respectively. Its consumption share is fixed and depends only on the

[8] We follow the notation in Werning (2007) and let the superscript m stand for "market."

Negishi weights, φ, and the coefficient of relative risk aversion. Markets are complete—as a result, households insure all idiosyncratic risk and face only aggregate risk in consumption. In contrast, its share of labor is a function of the Negishi weights, φ, the Frisch elasticity of labor supply, 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 labor productivity—but it is allocated efficiently.

At the household's optimum, its lifetime budget constraint holds with equality. Using equations (11)-(14) to substitute out after-tax prices, we obtain the following conditions corresponding to the time-0 budget constraint for each type $i \in I$:

$$\sum_t \sum_{s^t} \beta^t \mu(s^t) \left[U_C^m(s^t)\omega_C^i(\varphi)C(s^t) + U_L^m(s^t)\omega_L^i(\varphi, s_t)L(s^t)\right] \tag{17}$$

$$= U_C^m(s_0)\bar{T} + \sigma_0^i \frac{1-\tau_\Pi}{1+\tau_c} \sum_t \sum_{s^t} \beta^t \mu(s^t) U_C^m(s^t) \frac{\Pi(s^t)}{P(s^t)},$$

$$\text{where} \quad \bar{T} \equiv \frac{1}{1+\tau_c} \sum_t \sum_{s^t} \mu(s^t) \frac{\beta^t U_C^m(s^t)}{U_C^m(s_0)} \left[T(s^t) + (1-\tau_\Pi)\frac{\Pi(s^t)}{P(s^t)}\right]. \tag{18}$$

see Appendix A.3 for its derivation. The conditions in (17) indicate that for any household, its lifetime expenditure on consumption is equal to its lifetime wealth.

With multiple household types, there is a *set* of implementability conditions: one corresponding to the budget constraint for each type $i \in I$. The combination of linear taxes and uniform lump-sum transfers give the planner some ability to redistribute. The quantity $\bar{T}$ on the right hand side of equation (17) represents, in part, the time-0 value of lifetime transfers. Redistribution, however, is limited. The planner cannot achieve *any* desired distribution across households: transfers are non-targeted, hence $\bar{T}$ is uniform across types $i \in I$. It follows that the conditions in (17) are joint restrictions on equilibrium allocations.

Furthermore, it is clear from (17) and (18) that the assumed state-contingency of the lump-sum transfer is without loss of generality. As in Barro (1974), changes in the timing of lump-sum taxes or transfers are irrelevant for the decisions of the individual households; what matters is the value of $\bar{T}$. Ricardian equivalence holds.

The implementability conditions in (17) resemble the implementability conditions in Niepelt (2004) and Werning (2007).[9] A key characteristic, however, that distinguishes our conditions from the previous literature is the presence of profits. In our economy, monopolistic competition among intermediate-good firms results in equilibrium profits. Real profits, $\Pi(s^t)/P(s^t)$, enter the household's budget constraints as dividend payouts. We subsume the "common" component of these dividend payouts into $\bar{T}$.

[9] See also Bassetto (2014).

However, there is an "uncommon" component due to heterogeneity in initial endowments of equity. The final term on the right-hand side of equation (17) represents the household's heterogeneous exposure, σ_0^i, to the time-0 value of lifetime after-tax real profits. Complete markets imply that it is irrelevant whether the household holds on to its initial shares from time-0 onward, or trades them away—it is only the initial claim on firm profits that matters for its lifetime budget constraint. This final term disappears when there is either no heterogeneity in initial equity ($\sigma_0^i = 0$ for all $i \in I$) or when profits are fully taxed ($\tau_\Pi = 1$).

Firm optimality. Turning now to the firms, the unique solution to the flexible-price firm's problem is given by:

$$p_t^f(s^t) = \left[(1-\tau_r)\left(\frac{\rho-1}{\rho}\right)\right]^{-1} \frac{W(s^t)}{A(s_t)}, \qquad \forall s^t \in S^t. \tag{19}$$

Firm optimality equates marginal cost with after-tax marginal revenue. The firm's optimal price, therefore, is a constant markup over its nominal marginal cost $W(s^t)/A(s_t)$.

The unique solution to the sticky-price firm's problem is similarly given by:

$$p_t^s(s^{t-1}) = \left[(1-\tau_r)\left(\frac{\rho-1}{\rho}\right)\right]^{-1} \sum_{s^t|s^{t-1}} \frac{W(s^t)}{A(s_t)} q(s^t|s^{t-1}) \tag{20}$$

$$\text{where} \quad q(s^t|s^{t-1}) \equiv \frac{\mu(s^t|s^{t-1}) U_C^m(s^t) C(s^t) P(s^t)^{\rho-1}}{\sum_{s^t|s^{t-1}} \mu(s^t|s^{t-1}) U_C^m(s^t) C(s^t) P(s^t)^{\rho-1}}. \tag{21}$$

denote risk-adjusted conditional probabilities, conditional on s^{t-1}. By construction, these satisfy $\sum_{s^t|s^{t-1}} q(s^t|s^{t-1}) = 1$. Equation (20) states that the firm's optimal price is equal to a markup over its risk-weighted, conditional expectation of its nominal marginal cost, $W(s^t)/A(s_t)$.

Comparing (19) and (20), we infer that: $p_t^s(s^{t-1}) = \sum_{s^t|s^{t-1}} q(s^t|s^{t-1}) p_t^f(s^t)$. That is, the optimal price of the sticky-price firm is equal to its risk-adjusted expectation of the optimal price of the flexible-price firm.

3.1 Equilibrium Allocations

The following proposition characterizes the set of allocations that can be implemented as part of an equilibrium in this economy. In any equilibrium, all sticky-price firms set their prices according to (20) and all flexible-price firms set their prices according to (19). It follows 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 the output, labor, and profits of the flexible-price firms by $y^f(s^t)$, $n^f(s^t)$, and $\pi^f(s^t)$, respectively.

Proposition 1. *A feasible allocation $x \in \mathcal{X}$ can be implemented in equilibrium if and only if there exist a vector $\varphi \equiv (\varphi^i)$, and scalars $\bar{T} \in \mathbb{R}$, $\chi \in \mathbb{R}_+$, $\hat{\vartheta} \in \mathbb{R}_{\geq 0}$, 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 $s^t \in S^t$,*

$$\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho} + \frac{U_L^m(s^t)}{U_C^m(s^t)}\frac{1}{\chi A(s_t)} = 0; \tag{22}$$

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

$$\sum_{s^t|s^{t-1}} U_C^m(s^t)y^s(s^t)\left\{\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho} + \frac{U_L^m(s^t)}{U_C^m(s^t)}\frac{1}{\chi A(s_t)}\right\}\mu(s^t|s^{t-1}) = 0; \tag{23}$$

and (iii) for all $i \in I$:

$$\sum_t\sum_{s^t}\beta^t\mu(s^t)\left[U_C^m(s^t)\omega_C^i(\varphi)C(s^t) + U_L^m(s^t)\omega_L^i(\varphi,s_t)L(s^t)\right] \tag{24}$$
$$= U_C^m(s_0)\bar{T} + \sigma_0^i\hat{\vartheta}\sum_t\sum_{s^t}\beta^t\mu(s^t)\left[\chi\frac{\rho}{\rho-1}U_C^m(s^t)C(s^t) + U_L^m(s^t)L(s^t)\right]$$

Proof. See Appendix A.4. □

Proposition 1 characterizes the set of allocations that can be supported as part of an equilibrium. Aside from resource and technology constraints, any equilibrium allocation satisfies three additional sets of constraints.

Part (i) indicates that in any equilibrium, there is no output dispersion within the set of sticky-price firms nor within the set of flexible-price firms. This follows from the optimal pricing decisions of both sets of firms and the demand functions. However, there can be differences in production across the two sets of firms.

Part (ii) states that in any equilibrium, condition (22) 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 (11). The resulting condition simply states that the marginal cost to the flexible-price firm of producing an extra unit of output is equated with its marginal revenue.

Condition (23) 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 (22) corresponding to flexible-price firm optimality, the only difference being that in (23), the marginal cost and marginal revenue of the sticky-price firm are equated in risk-weighted expectation, conditional on s^{t-1}.

Finally, part (iii) ensures that all budget constraints are satisfied in equilibrium. These conditions—one per household type—correspond to the lifetime budget constraints in (17), but with real profits, $\Pi(s^t)/P(s^t)$, expressed in terms of the allocation. The government's budget constraint holds by Walras's law.

3.2 The power of monetary and fiscal policy

The primary channel by which monetary and fiscal policy influence equilibrium allocations is through the labor wedge. To see this explicitly, we rewrite our equilibrium conditions as follows. First, the optimal price of the sticky-price firm in (20) can be expressed as:

$$p_t^s(s^{t-1}) = \epsilon(s^t)\left[(1-\tau_r)\left(\frac{\rho-1}{\rho}\right)\right]^{-1}\frac{W(s^t)}{A(s_t)}, \tag{25}$$

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

$$\epsilon(s^t) \equiv \frac{\sum_{s^t|s^{t-1}}[q(s^t|s^{t-1})W(s^t)/A(s_t)]}{W(s^t)/A(s_t)}. \tag{26}$$

Formally, $\epsilon(s^t)$ is defined as the firm's optimal forecast error of $W(s^t)/A(s_t)$, conditional on information set s^{t-1}. Therefore, $\epsilon(s^t)$ acts as a stochastic wedge between the firm's price and its ex-post optimal price, i.e. the markup over nominal marginal cost. Because the sticky-price firm has incomplete information, it cannot perfectly forecast its nominal marginal cost, and as a result, a state-contingent wedge emerges that can be interpreted as the firm's "pricing mistake."

Aggregating over the sticky- and flexible-price firms' prices according to (8) and combining the aggregate price level with the representative household's intratemporal optimality condition in (11), we obtain the following equilibrium condition:

$$-\frac{U_L^m(s^t)}{U_C^m(s^t)} = \chi\left[\kappa\epsilon(s^t)^{1-\rho}+(1-\kappa)\right]^{-\frac{1}{1-\rho}}A(s_t), \quad \text{where} \quad \chi \equiv \left(\frac{\rho-1}{\rho}\right)\frac{(1-\tau_\ell)(1-\tau_r)}{1+\tau_c}. \tag{27}$$

This condition indicates that the marginal rate of substitution between labor and consumption, $-U_L^m(s^t)/U_C^m(s^t)$, is equated with the marginal rate of transformation, $A(s_t)$, modulo a labor wedge. The labor wedge is the product of two components.

The first is a fiscal component, denoted by χ. This wedge is the product of multiple terms: the consumption, sales, and labor income taxes, and the intermediate-good markup. It follows that χ is a lever of fiscal policy that is both time- and state-invariant.

The second component of the labor wedge, given by $[\kappa\epsilon(s^t)^{1-\rho} + (1-\kappa)]^{-\frac{1}{1-\rho}}$, contains the state-contingent "pricing errors," $\epsilon(s^t)$, made by the fraction κ of inattentive firms. The nominal rigidity thereby gives rise to a state-contingent component of the labor wedge; this component represents the lever of monetary policy. By shifting $\epsilon(s^t)$, monetary policy can shift allocations in ways fiscal policy cannot.

The power of monetary policy, however, is limited in two ways, corresponding to parts (i) and (ii) of Proposition 1. First, $\epsilon(s^t)$ introduces a wedge between the prices of the sticky- and flexible-price firms: $p_t^s(s^{t-1}) = \epsilon(s^t)p_t^f(s^t)$. This in turn drives a wedge between the sticky-price and flexible-price firms' output, implying a loss in production efficiency. Second, by construction, the forecast errors $\epsilon(s^t)$ must "average out" to 1. This is the interpretation of the implementability condition in (23): monetary policy cannot surprise firms "on average." This constraint on equilibrium allocations follows directly from the optimal price-setting behavior of sticky-price firms; it is therefore a natural consequence of rational expectations.

Untaxed profits. The final term in equation (24) corresponds to household-i's heterogeneous exposure, σ_0^i, to the lifetime value of after-tax real profits. The weakly-positive scalar $\hat{\vartheta}$, given by $\hat{\vartheta} \equiv (1-\tau_\Pi)/(1-\tau_\ell)$, parameterizes an additional lever of fiscal policy aside from χ.

The profit tax itself does not distort households' intratemporal or intertemporal decisions; taxing profits and distributing them uniformly thereby constitutes a non-distortionary channel of redistribution available to the government. For this reason it will never be optimal to set this tax within its interior. If, for example, equity shares covary positively with lifetime labor earnings and there is a desire to redistribute from households with high lifetime earnings to those with low lifetime earnings, it will be optimal to fully tax profits: $\tau_\Pi = 1$. If, however, the government is unable to fully tax profits, then monetary and fiscal policy would affect the distribution of dividend income through changes in real profits. In order to let this channel play a role, we make the following ad-hoc assumption about the government's ability to tax profits.

Assumption 1. Let $\vartheta \geq 0$. The tax rates τ_Π and τ_ℓ are such that $\hat{\vartheta} = \vartheta$.

If profits are fully taxed, then $\vartheta = 0$. In this case, heterogeneity in firm ownership plays no role. If instead $\vartheta > 0$, then the government cannot fully tax profits or drive τ_ℓ to negative infinity. In this case, heterogeneity in firm ownership affords a second channel by which monetary and fiscal policy affect the equilibrium distribution of income.

For the remainder of our analysis, we index economies by $\vartheta \geq 0$. For a given ϑ, we let $\mathcal{X}^e(\vartheta) \subset \mathcal{X}$ denote the set of equilibrium allocations in economy-ϑ: all allocations that satisfy Proposition 1 with $\hat{\vartheta} = \vartheta$.

4 The Ramsey Problem and Optimal Monetary Policy

The goal of this paper is to solve the Ramsey problem and study its implications for optimal policy. We assume a utilitarian social welfare function given by:

$$\mathcal{U} \equiv \sum_{i \in I} \lambda^i \pi^i \sum_t \sum_{s^t} \beta^t \mu(s^t) U(c^i(s^t), \ell^i(s^t)/\theta^i(s_t)) \tag{28}$$

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

Definition 3. Given $\vartheta \geq 0$, a Ramsey optimum is an allocation x that maximizes (28) subject to $x \in \mathcal{X}^e(\vartheta)$.

There are two sources of heterogeneity in our model: labor income heterogeneity and heterogeneity in initial firm ownership. For the remainder of this section we isolate the former and abstract from the latter. Specifically, we solve the Ramsey problem in a baseline economy in which profits are fully taxed: $\vartheta = 0$. Full profit taxation renders heterogeneity in firm ownership irrelevant. In Section 5, we relax this assumption and consider the case in which $\vartheta > 0$.

With $\vartheta = 0$, the implementability conditions in (24) reduce to:

$$\sum_t \sum_{s^t} \beta^t \mu(s^t) \left[U_C^m(s^t) \omega_C^i(\varphi) C(s^t) + U_L^m(s^t) \omega_L^i(\varphi, s_t) L(s^t) \right] = U_C^m(s_0) \bar{T}, \quad \forall i \in I. \tag{29}$$

Following the standard procedure, we write the Ramsey problem in Lagrangian form. Let $\pi^i \nu^i$ denote the Lagrange multiplier on the implementability condition (29) for type $i \in I$, and let $\nu \equiv (\nu^i)_{i \in I}$ denote the set of multipliers. 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

$$\sum_t \sum_{s^t} \beta^t \mu(s^t) \mathcal{W}(C(s^t), L(s^t), s_t; \varphi, \nu, \lambda) - U_C^m(s_0) \bar{T} \sum_{i \in I} \pi^i \nu^i, \tag{30}$$

subject to resource constraints,

$$Y(s^t) = \left[\kappa y^s(s^t)^{\frac{\rho-1}{\rho}} + (1-\kappa) y^f(s^t)^{\frac{\rho-1}{\rho}} \right]^{\frac{\rho}{\rho-1}}, \qquad L(s^t) = \kappa \frac{y^s(s^t)}{A(s_t)} + (1-\kappa) \frac{y^f(s^t)}{A(s_t)}, \tag{31}$$

$C(s^t) \leq Y(s^t)$, and conditions (22) and (23), where the function $\mathcal{W}$ incorporates the implementability conditions in (29) into the planner's maximand as follows:

$$\begin{aligned} \mathcal{W}(C(s^t), L(s^t), s_t; \varphi, \nu, \lambda) \equiv & \sum_{i \in I} \pi^i \lambda^i U(\omega_C^i(\varphi) C(s^t), \omega_L^i(\varphi, s_t) L(s^t)/\theta^i(s_t)) \\ & + \sum_{i \in I} \pi^i \nu^i \left[U_C^m(s^t) \omega_C^i(\varphi) C(s^t) + U_L^m(s^t) \omega_L^i(\varphi, s_t) L(s^t) \right]. \end{aligned} \tag{32}$$

In Appendix A.5 we characterize the Ramsey optimum. We are primarily interested in what Ramsey optimality implies for optimal monetary policy, which we turn to next.

4.1 Optimal Monetary Policy

The following theorem provides a characterization of optimal monetary and fiscal policy. In particular, we express monetary policy in terms of the aggregate markup, $\mathcal{M}(s^t)$, defined as the aggregate price level divided by nominal marginal cost:

$$\mathcal{M}(s^t) \equiv \frac{P(s^t)}{W(s^t)/A(s^t)}. \tag{33}$$

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

Theorem 1. *Let $\vartheta = 0$ and let x^* be a Ramsey optimum in this economy. Let χ^* be the fiscal wedge at x^*, let $\bar{\mathcal{M}} > 0$ be an arbitrary positive constant, and let $\mathcal{I} : S \to \mathbb{R}_+$ be a positively-valued function defined by:*

$$\mathcal{I}(s_t) \equiv \frac{\sum_{i\in I} \tilde{\pi}^i (\varphi^i)^{-1/\eta} (\theta^i(s_t))^{\frac{1+\eta}{\eta}}}{\sum_{i\in I} \pi^i (\varphi^i)^{-1/\eta} (\theta^i(s_t))^{\frac{1+\eta}{\eta}}} > 0, \qquad \textit{where} \qquad \tilde{\pi}^i \equiv \pi^i \left[\frac{\lambda^i}{\varphi^i} + \nu^i(1+\eta)\right]. \tag{34}$$

There exists a threshold, $\bar{\mathcal{I}}(s^{t-1}) > 0$, such that x^* can be implemented with tax rates that satisfy:

$$(1-\tau_r)^{-1}\frac{\rho}{\rho-1} = \bar{\mathcal{M}}, \qquad \textit{and} \qquad \frac{1-\tau_\ell}{1+\tau_c}(\bar{\mathcal{M}})^{-1} = \chi^*. \tag{35}$$

and a monetary policy that targets a state-contingent markup $\mathcal{M}^*(s^t)$ that satisfies:

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

Proof. See Appendix A.6. □

The function $\mathcal{I}(s_t)$ can be interpreted as a sufficient statistic for the level of labor income inequality in state s_t. Recall that λ^i are the Pareto weights, φ^i are the Negishi weights, and ν^i are the multipliers on the implementability conditions in (29). Notably all are scalars—they do not depend on the realized state or history. As a result, $\mathcal{I}(s_t)$ is history-independent and in particular depends only on the current realization of the labor skill distribution, $(\theta^i(s_t))_{i\in I}$.

As we show in an example below, the weights $\lambda^i/\varphi^i + \nu^i(1+\eta)$ are increasing in human wealth: types with high lifetime labor earnings have larger weights at the Ramsey optimum than types with low lifetime labor earnings.[10] As the labor productivities $\theta^i(s_t)$

[10] While high-types 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.

of the high-type households increase relative to those of lower-types, the numerator of $\mathcal{I}(s_t)$ grows relative to its denominator. As a result, $\mathcal{I}(s_t)$ is higher in states in which high lifetime labor earnings households are relatively more productive than the low lifetime labor earnings households. Furthermore, the extent to which $\mathcal{I}(s_t)$ responds to relative movements in the labor skill distribution depends on the Frisch elasticity of labor supply, $1/\eta$.

Theorem 1 characterizes fiscal and monetary policy at the Ramsey optimum. Clearly there is no unique fiscal implementation of the optimal fiscal wedge χ^*, and any implementation of χ^* results in the same behavior for optimal monetary policy. We thus index implementations by $\bar{\mathcal{M}}$, the aggregate markup under flexible-prices. Note that targeting a constant markup is equivalent to targeting price stability, i.e. zero inflation.

We find that optimal monetary policy targets a state-contingent markup: this markup depends on the level of labor income inequality, as proxied for by $\mathcal{I}(s_t)$. There exists a critical threshold $\bar{\mathcal{I}}(s^{t-1})$ such that when labor income inequality is strictly greater than this threshold, the optimal markup is greater than $\bar{\mathcal{M}}$. When labor income inequality is below this threshold, the optimal markup is lower than $\bar{\mathcal{M}}$. When $\mathcal{I}(s_t)$ is exactly equal to the threshold, the optimal markup is equal to $\bar{\mathcal{M}}$. We discuss equilibrium interest rates and prices shortly.

Proportional Shocks to the Labor Skill Distribution. To build intuition, we consider a special case in which the skill distribution exhibits proportional shocks.

Corollary 1. *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:*

$$\theta^i(s_t) = \vartheta^i \Theta(s_t), \quad \forall s_t \in S, \tag{36}$$

then $\mathcal{M}^*(s^t) = \bar{\mathcal{M}}$ *for all* $s^t \in S^t$.

As a corollary to Theorem 1, Corollary 1 provides sufficient conditions under which it is optimal for monetary policy to target a constant markup. These conditions are: separable and homothetic preferences and proportional shocks to the labor productivity distribution. When these conditions are met, the function $\mathcal{I}(s_t)$ reduces to a constant: $\mathcal{I}(s_t) \equiv \bar{\mathcal{I}}$, for all $s_t \in S$. The optimal level of redistribution can be achieved with a constant fiscal wedge and a monetary policy that replicates flexible prices.

Disproportional Shocks to the Labor Skill Distribution. Moving away from this special case, we illustrate the effect of disproportional shocks on optimal policy with a simple numerical example. In this example there are two household types—a high-type and

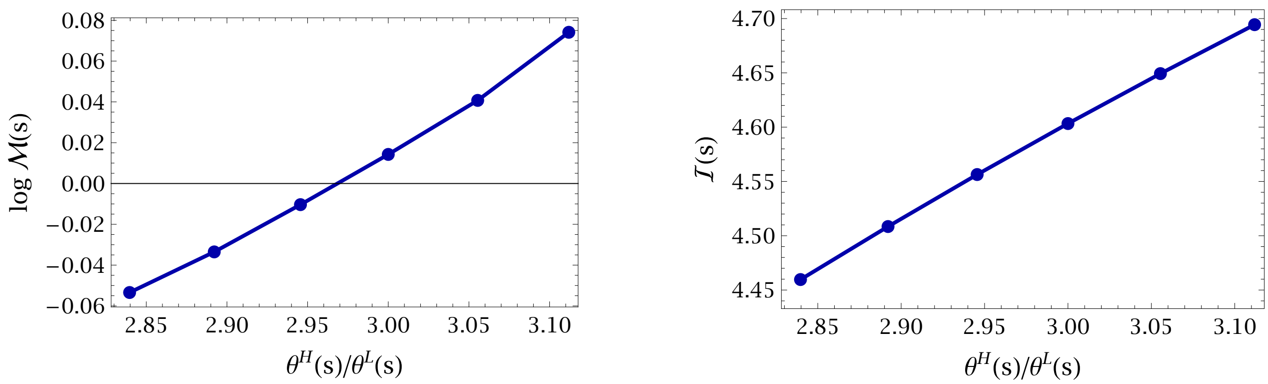

Figure 1. The optimal markup, $\log \mathcal{M}^*(s^t)$, as a function of $\theta^H(s_t)/\theta^L(s_t)$ (left panel). $\mathcal{I}(s_t)$ as a function of $\theta^H(s_t)/\theta^L(s_t)$ (right panel).

a low-type—indexed by $i \in \{H, L\}$, of equal sizes ($\pi^H = \pi^L = 1/2$). We consider a labor skill distribution in which the high-type is always more productive than the low-type, but we let the ratio $\theta^H(s_t)/\theta^L(s_t)$ fluctuate across 6 possible states. We assume states are uniformly distributed and i.i.d.: $\mu(s'|s) = 1/6$ for all $s, s' \in S$, and that $A(s_t) = 1$ in all states. Finally, we set $\beta = .98$, $\eta = 1$, $\gamma = 2$, $\kappa = .25$, $\rho = 2$, and τ_r such that $\bar{\mathcal{M}} = 1$.

We numerically solve for the Ramsey optimum with equal Pareto weights: $\lambda^H = \lambda^L = 1$. The left panel of Figure 1 plots the optimal markup for different values of $\theta^H(s_t)/\theta^L(s_t)$. As this ratio increases, i.e. as the high-type becomes more productive relative to the low-type, the optimal markup increases. To check that this is in line with the predictions of our theory, in the right panel of Figure 1 we plot $\mathcal{I}(s_t)$ as a function of $\theta^H(s_t)/\theta^L(s_t)$.[11]

Intuition. Recall that monetary and fiscal policy together act as a labor wedge or a tax on labor income. In fact, we can rewrite Theorem 1 in terms of the labor wedge that supports the Ramsey optimum in equilibrium. Specifically, if we define the optimal implicit "monetary tax," $\tau_M^*(s^t)$, as the state-contingent component of the labor wedge at the Ramsey optimum as follows:

$$-\frac{U_L^m(s^t)}{U_C^m(s^t)} = \chi^*(1 - \tau_M^*(s^t))\frac{Y(s^t)}{L(s^t)}, \tag{37}$$

then the optimal implicit tax satisfies:

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

[11] In this example, the weight $\lambda^i/\varphi^i + \nu^i(1+\eta)$ of the high-type is greater than that of the low-type.

The markup is equivalent to the implicit monetary tax, $\tau_M^*(s^t)$. If the final good price rises above its marginal cost, it is *as if* households pay an implicit income tax; conversely, if the price falls below its marginal cost, it is as if households receive a subsidy. Therefore, by raising or lowering the markup, monetary policy essentially manipulates the marginal tax rate of a linear income tax, albeit with an accompanying loss in production efficiency due to price dispersion. Theorem 1 indicates that when $\mathcal{I}(s_t)$ rises above the critical threshold, the optimal monetary tax is positive; when $\mathcal{I}(s_t)$ falls below the threshold, the optimal monetary tax is negative; and when $\mathcal{I}(s_t)$ is exactly equal to the threshold, the optimal monetary tax is zero.

To understand this pattern, consider the distributional nature of a linear labor income tax. Given a linear income tax that finances lump-sum transfers, all households face the same marginal tax rate, but high-skilled, high-earning households face higher average tax rates than low-skilled, low-earning households. A positive tax rate therefore reduces inequality. It follows that if the planner desires a more equal distribution of lifetime earnings across households than under laissez-faire (zero taxes and replication of flexible prices), then the optimal implicit tax rate is, on average, positive.

In the case of proportional shocks to the labor income distribution, the optimal implicit tax rate can be implemented with fiscal policy alone (Corollary 1). This is because the optimal tax rate in any state is the rate at which the marginal cost of distorting the intratemporal margin is equated with its marginal distributional benefit. When preferences are separable and homothetic, the marginal cost of distorting the intratemporal margin is constant across states.[12] Meanwhile, when shocks to the labor skill distribution are proportional, the marginal distributional benefit of this distortion is also invariant across states. In this case, the ratio of labor productivity across any two types, $i, j \in I$, is constant:

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

With no movement in the *relative* skill distribution, there are no states in which labor income taxation is more beneficial than others. It follows that a constant fiscal wedge equal to χ^* is sufficient to support the Ramsey optimum, and state-contingent monetary policy is unnecessary. Optimal monetary policy replicates flexible prices.

On the other hand, when labor productivities vary disproportionately across states, the marginal benefit of labor income taxation also varies—rendering state-contingent monetary policy useful. In our simple numerical example, high-type households are

[12] Under such preferences, perfect "tax smoothing" arises in Lucas and Stokey (1983), partially affirming arguments made in Barro (1979). One can understand this as an application of the classic uniform commodity taxation result: under homothetic and separable preferences, it is optimal to tax goods at a uniform rate (Atkinson and Stiglitz, 1980; Chari and Kehoe, 1999).

always more productive than low-type households and hence have higher lifetime earnings. However, there are states in which the productivity of the high-type is especially high relative to that of the low-type, and states in which the productivities of the two types are more compressed. In the former case, the marginal distributional benefit of the tax rate increases: high-types always face a higher tax burden than low-types but more so in these states. Conversely, in the latter case, the marginal distributional benefit of the tax rate falls. It follows that the optimal implicit tax rises in the "high inequality" states and falls in the "low inequality" states.

Optimal monetary policy mimics this state-contingent pattern: it targets a high markup in high inequality states and a low markup in low inequality states. Another way to understand this is through movements in the real wage: optimal monetary policy lowers the real wage in high inequality states and raises the real wage in low inequality states. By distorting the economy in this state-contingent manner, optimal monetary policy compresses the *lifetime* labor earnings distribution across households.

Note that monetary policy cannot perfectly replicate a true labor income tax: abandoning replication of flexible prices results in price dispersion and, hence, misallocation. Nevertheless, starting from the flexible-price benchmark where production efficiency is maximized, any loss in production efficiency due to monetary policy is, to a first-order, zero. It follows that if there is any incentive for monetary policy to move away from replicating flexible prices—here, when relative productivities fluctuate—the Ramsey planner finds it optimal to do so.

One concern with optimal monetary policy may be that it depends critically on the assumption that profits are fully taxed. We show in the following section, Section 5, that this is not the case: the qualitative behavior of the optimal markup is robust to less than full profit taxation and heterogeneity in firm ownership.

4.2 Nominal Implementation: Price Levels and Interest Rates

We next use Theorem 1 to characterize equilibrium price levels and nominal interest rates consistent with the Ramsey optimum. As is standard in the New Keynesian literature, nominal implementation of a specific allocation is not unique. We therefore choose to focus on a particular implementation in which the sticky-price firms optimally set their prices equal to the aggregate price level in the previous period—this is consistent with one-period Calvo as in, e.g., Correia, Nicolini and Teles (2008). In Appendix C we provide a more general characterization of nominal implementation.

Specifically, let the optimal price of the sticky-price firms at time t be equal to the aggregate price level in the previous period: $p^s_t(s^{t-1}) = P(s^{t-1}) > 0$. Equivalently, the common belief of the aggregate price level at time t, based only on past events s^{t-1}, is the previous period's price level.

Given this belief, the economy replicates the flexible price outcome in period t, history s^t, if $P(s^t) = P(s^{t-1})$. This is the case in which both flexible-price firms and sticky-price firms set their price equal to $P(s^{t-1})$. We call this price level stabilization. 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 2. *Let $\vartheta = 0$. A Ramsey optimum x^* can be implemented with tax rates that satisfy* (35)*, and a state-contingent markup described in Theorem 1. An aggregate price level, $P(s^t)$, and nominal interest rate, $i(s^t)$, consistent with these policies satisfy:*

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

Proof. See Appendix C for a more general version of this proposition and its proof. □

The behavior of the aggregate price level in Proposition 2 is consistent with the optimal markup in Theorem 1. To understand this, consider first the flexible price outcome. In order to replicate flexible prices in any particular history s^t, monetary policy must target price stability: $P(s^t) = P(s^{t-1})$. In this case the wage moves one-for-one with aggregate productivity, the flexible-price firms find it optimal to set their prices equal to $P(s^{t-1})$, and the markup satisfies $\mathcal{M}(s^t) = \bar{\mathcal{M}}$. Replicating the flexible price outcome is optimal when $\mathcal{I}(s_t) = \bar{\mathcal{I}}(s^{t-1})$.

Consider next the case in which $\mathcal{I}(s_t)$ rises above the threshold and it is optimal for the markup to increase. In order to generate an increase in the markup, the nominal wage must fall unexpectedly relative to aggregate productivity. 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" at the past price level. As a result, the aggregate price level falls, but less so than the nominal wage, so that the markup increases.

Conversely, when $\mathcal{I}(s_t)$ falls below the threshold, it is optimal for the markup to fall. In order to generate a fall in the markup, the nominal wage must rise unexpectedly relative to aggregate productivity. In this case, the aggregate price level also rises, but less so than wages, so that the markup falls.

The nominal interest rate must be consistent with these price movements and the allocation. In particular, the equilibrium nominal interest rate satisfies the bond Euler equation in (12). A rise in the markup is consistent with a tightening of the nominal interest rate relative to the natural rate: this contracts the economy so that the optimal level of output is lower than its flexible-price counterpart. Conversely, a fall in the markup is consistent with a loosening of the nominal interest rate relative to the natural rate: this expands the economy so that the optimal level of output is higher than under flexible prices. By contracting the economy when labor income inequality is high and expanding the economy when labor income inequality is low, optimal monetary policy compresses the lifetime labor earnings distribution at the expense of some efficiency.

Remarks on Multiplicity and Robustness. We have presented a particular implementation in which the common belief of the aggregate price at time t, based on s^{t-1}, is the previous period's price level. This is a natural implementation, however it is not unique. In Appendix C we characterize the set of all possible nominal implementations of which Proposition 2 is a particular case. All implementations, however, feature the same general behavior of the price level and nominal interest rate described in Proposition 2. In this sense, relative movements in prices and interest rates are robust across nominal implementations.

The behavior of nominal variables at the Ramsey optimum, however, would depend on the relative stickiness of prices versus wages. In our model, prices are "sticky" in the sense that price-setting firms face informational constraints, while nominal wages are fully flexible. 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. An increase in the price level with a less than commensurate increase in the nominal wage is equivalent to a lower real wage, or, a higher markup. Conversely a fall in the markup would require an unexpected fall in the price level.

For this reason, in terms of understanding the optimal conduct of monetary policy, we wish to put less emphasis on the nominal implementation in Proposition 2 and more emphasis on the behavior of the optimal markup in Theorem 1. In particular, what is robust to the relative stickiness of prices versus wages is the labor wedge that supports the Ramsey optimum. This is one advantage of using the primal approach.

5 Extensions

Thus far in our analysis we have made two stark assumptions on fiscal policy. One is that tax rates are nonstate-contingent, and the other is that profits are fully taxed, $\tau_{\Pi} = 1$. In this section we relax both assumptions: first individually, then jointly.

5.1 Partially State-Contingent Taxes

We first relax our restriction on the nonstate-contingency of fiscal tools. We allow tax rates to be partially state-contingent: they can 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}. We maintain our assumption that profits are fully taxed: $\tau_\Pi = 1$.

With tax rates that can adjust to past states, the fiscal authority has the flexibility to respond to shocks with a one-period lag. To the extent that shocks are persistent, fiscal policy can absorb some of the state-contingent pressure placed on monetary policy. In this case we obtain a sharper characterization of the behavior of the optimal markup near the flexible price benchmark.

Theorem 2. *Let tax rates be set one period in advance. There exists a threshold $\bar{\mathcal{I}}(s^{t-1}) > 0$ such that, to a first-order approximation around $\mathcal{I}(s_t) = \bar{\mathcal{I}}(s^{t-1})$, the optimal state-contingent markup $\mathcal{M}^*(s^t)$ satisfies:*

$$\log \mathcal{M}^*(s^t) - \log \bar{\mathcal{M}} \approx \delta_0[\mathcal{I}(s_t)/\bar{\mathcal{I}}(s^{t-1}) - 1] \quad (38)$$

where

$$\delta_0 = \frac{1}{1 + \rho(\eta + \gamma)\frac{1-\kappa}{\kappa}} \in (0, 1). \quad (39)$$

Proof. See Appendix E.2. □

When we allow tax rates to be set one-period in advance, our main result on the optimal conduct of monetary policy remains intact and we obtain a sharper characterization of the optimal markup near its flexible price benchmark level. In particular, we show that to a first-order at $\mathcal{I}(s_t) = \bar{\mathcal{I}}(s^{t-1})$, the optimal markup 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 markup responds to an increase in $\mathcal{I}(s_t)$: a larger value for δ_0 indicates a more aggressive response of monetary policy, whereas a lower value indicates a less aggressive response. An explicit, closed-form expression for this derivative is given in (39). We find that δ_0 is strictly positive, strictly less than 1, and a function of the primitives ρ, γ, η, and κ.

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

Second, δ_0 is increasing in $\kappa/(1-\kappa)$, the mass of sticky-price firms relative to the mass of flexible-price firms. Consider the limit in which $\kappa \to 1$. In this case δ_0 approaches one. When nearly all firms in the economy are sticky, movements in monetary policy away from flexible-rice allocations result in near zero losses in production efficiency; monetary policy therefore approximates a labor income tax. In this limit, monetary policy perfectly mimics a state-contingent tax rate which responds one-for-one with changes in $\mathcal{I}(s_t)$. In the opposite limit in which $\kappa \to 0$, δ_0 approaches zero. That is, when nearly all firms in the economy are flexible, monetary policy has no power.

5.2 Constrained Profit Taxation

We next relax the assumption that profits are fully taxed. In what follows we constrain profit taxation to be only partial, so that $\vartheta > 0$. We maintain our assumption that all tax rates are constant.

Recall that $\mathcal{X}^e(\vartheta) \subset \mathcal{X}$ denotes the set of equilibrium allocations in economy-ϑ. This set is characterized in Proposition 1. We now consider the Ramsey problem for economy-ϑ with $\vartheta > 0$. A Ramsey optimum for this economy is characterized in Appendix D; its implications for optimal monetary policy are as follows.

Theorem 3. *Let $\vartheta > 0$. There exists a threshold $\bar{\mathcal{I}}_\vartheta(s^{t-1}) > 0$, such that the optimal state-contingent markup $\mathcal{M}^*(s^t)$ satisfies:*

$$\begin{array}{ll} \mathcal{M}^*(s^t) > \bar{\mathcal{M}} & \textit{if and only if} \;\; \mathcal{I}(s_t) > \bar{\mathcal{I}}_\vartheta(s^{t-1}); \\ \mathcal{M}^*(s^t) = \bar{\mathcal{M}} & \textit{if and only if} \;\; \mathcal{I}(s_t) = \bar{\mathcal{I}}_\vartheta(s^{t-1}); \\ \mathcal{M}^*(s^t) < \bar{\mathcal{M}} & \textit{if and only if} \;\; \mathcal{I}(s_t) < \bar{\mathcal{I}}_\vartheta(s^{t-1}). \end{array}$$

Proof. See Appendix E.3. □

When profits are only partially taxed and initial equity shares are heterogeneous, the behavior of the optimal markup resembles that in the baseline economy with full profit taxation. In particular, it remains optimal for the markup to rise in states with high income inequality and to fall in states in which the productivity distribution is more compressed.

Initially this result may seem surprising. When monetary policy abandons the flexible price benchmark and unexpectedly increases or decreases the aggregate markup, firm profits are altered. Although monetary policy can only raise the markup in one state if it lowers it in another, on average these movements in profits may not fully cancel each other out. Therefore, abandoning flexible-price allocations generally implies some amount of redistribution of lifetime financial wealth.

However, this redistribution of financial wealth occurs at time 0: it is a change in households' lifetime budget sets. In this sense, while it may affect the *average* distributional benefit of fiscal policy, it does not affect the optimal timing of state-contingent monetary policy.[13] The marginal benefit of manipulating the markup varies across states in the same manner as when profits are fully taxed. By lowering the real wage when labor income inequality is high and raising the real wage when labor income inequality is low, state-contingent monetary policy optimally compresses the lifetime labor earnings distribution in the desired direction.

5.3 Constrained Profit Taxation and Partially State-Contingent Taxes

We now jointly relax both restrictions on fiscal policy. We assume $\vartheta > 0$ and we allow tax rates to be set one period in advance. Specifically, we let τ_c and τ_r at time t be contingent on s^{t-1}. Optimal monetary policy in this economy can be characterized as follows.

Theorem 4. *Let $\vartheta > 0$ and tax rates be set one period in advance. There exists a threshold $\bar{\mathcal{I}}_\vartheta(s^{t-1}) > 0$ such that, to a first-order approximation around $\mathcal{I}(s_t) = \bar{\mathcal{I}}_\vartheta(s^{t-1})$, the optimal state-contingent markup $\mathcal{M}^*(s^t)$ satisfies:*

$$\log \mathcal{M}^*(s^t) - \log \bar{\mathcal{M}} \approx \delta_0 \frac{1}{\mathcal{H}_\vartheta(s^{t-1}) + \frac{\rho}{\rho-1}(\gamma-1)\vartheta \sum_{i\in I} \pi^i \nu^i \sigma_0^i} [\mathcal{I}(s_t) - \bar{\mathcal{I}}_\vartheta(s^{t-1})], \tag{40}$$

where $\mathcal{H}_\vartheta(s^{t-1}) \equiv \chi^*(s^{t-1})^{-1} \frac{\sum_{i\in I} \tilde{\pi}^i (\varphi^i)^{1/\gamma}}{\sum_{i\in I} \pi^i (\varphi^i)^{1/\gamma}} > 0$.

Proof. See Appendix E.4. □

When both restrictions on fiscal policy are relaxed, our main result on the optimal conduct of monetary policy remains intact. We again obtain a sharp characterization of the optimal markup near its flexible price benchmark level: in particular, to a first-order around $\mathcal{I}(s_t) = \bar{\mathcal{I}}_\vartheta(s^{t-1})$, the optimal markup is strictly increasing in $\mathcal{I}(s_t)$.

We can compare this behavior to the economy with full profit taxation, $\vartheta = 0$, but with tax rates set one period in advance, as presented in Theorem 2. In this economy,

$$\log \mathcal{M}^*(s^t) - \log \bar{\mathcal{M}} \approx \delta_0 \frac{1}{\mathcal{H}_0(s^{t-1})} [\mathcal{I}(s_t) - \bar{\mathcal{I}}_0(s^{t-1})], \tag{41}$$

where $\mathcal{H}_0(s^{t-1}) > 0$ takes the same form as $\mathcal{H}_\vartheta(s^{t-1})$ but evaluated at the Ramsey optimum for economy-0. Equation (41) directly corresponds to equation (38), noting that $\bar{\mathcal{I}}_0(s^{t-1}) = \mathcal{H}_0(s^{t-1})$.

We make the following heuristic argument comparing the slope in (40) to that in (41). The terms $\mathcal{H}_\vartheta(s^{t-1})$ and $\mathcal{H}_0(s^{t-1})$, although they take the same functional form, are difficult to compute and compare. We therefore focus on the other term determining

[13] Fiscal policy, χ, can be adjusted until its marginal distortionary cost is in line with its marginal redistributional benefit.

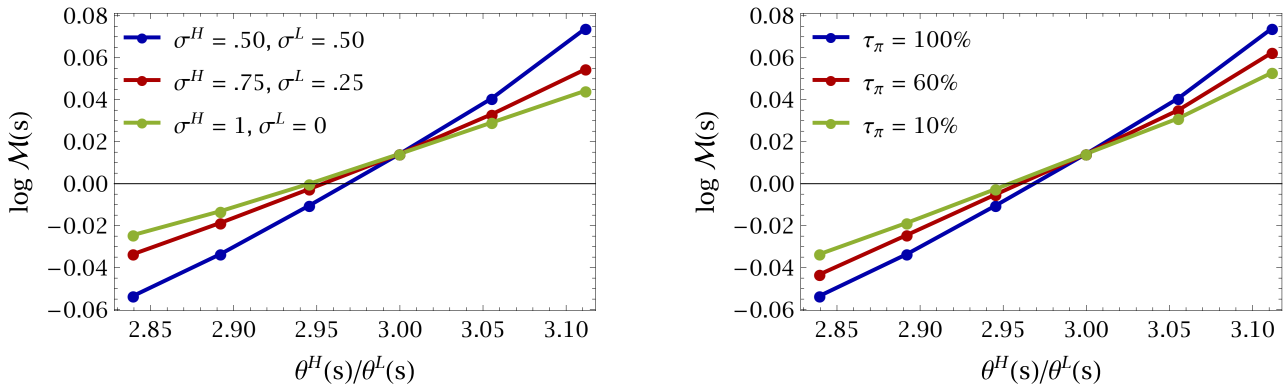

Figure 2. The optimal markup as a function of $\theta^H(s_t)/\theta^L(s_t)$ for different initial distributions of equity (left panel) and for different levels of the profit tax (right panel).

the difference in slopes: $\vartheta \sum_{i\in I} \pi^i \nu^i \sigma_0^i$. This key term is the product of two components: ϑ, which parameterizes the extent to which profits are left untaxed, and $\sum_{i\in I} \pi^i \nu^i \sigma_0^i$, the cross-sectional covariance between σ_0^i, the initial equity shares, and ν^i, the multipliers on the implementability conditions in (24).

Recall that households with higher lifetime labor earnings have higher values of ν^i. A positive value for $\sum_{i\in I} \pi^i \nu^i \sigma_0^i$ thereby indicates a positive cross-sectional covariance between initial equity and lifetime labor earnings: high human wealth households own greater shares of the firm.

When $\vartheta \sum_{i\in I} \pi^i \nu^i \sigma_0^i$ is strictly positive—when profits are not fully taxed and when initial equity and lifetime labor earnings are positively correlated—this term contributes to a "lower" slope in economy-ϑ than in economy-0. This is consistent with the intuition that when profits are not fully taxed and high human wealth households own greater shares of the firm, an increase in the markup in high-inequality states reduces overall labor income inequality but could increase profits in those states. As a result, optimal monetary policy responds less aggressively to movements in $\mathcal{I}(s_t)$. While this is only a heuristic argument, we show that it holds in the following numerical example.

Numerical Illustration. To illustrate this result, we return to our simple example of two household types, a high-type and a low-type, of equal sizes described in Section 4. We again let the the ratio $\theta^H(s_t)/\theta^L(s_t)$ fluctuate across 6 possible states and we fix all parameter values as in our previous example. In this exercise we vary the distribution of initial equity and the profit tax itself.

In the left panel of Figure 2 we plot the optimal markup as a function of $\theta^H(s_t)/\theta^L(s_t)$ for three economies: one with equal initial firm ownership, one in which the low-type owns 25% and the high-type owns 75%, and one in which the low-type owns zero shares and the high-type owns 100% of the firm. We keep the profit tax constant, set at $\tau_\Pi = 10\%$. Our baseline ($\vartheta = 0$) is nested by the economy with equal firm ownership. We find that in all three economies, the optimal markup is increasing in the ratio $\theta^H(s_t)/\theta^L(s_t)$. As the distribution of initial shares becomes more unequal, the slope of the optimal markup with respect to the ratio of productivities falls but remains positive.

In the right panel we plot the optimal markup as a function of $\theta^H(s_t)/\theta^L(s_t)$ for three economies: one with $\tau_\Pi = 100\%$ as in our baseline, another with $\tau_\Pi = 60\%$, and a third with $\tau_\Pi = 10\%$. We keep constant the initial distribution of equity: the low-type owns 20% of the firm and the high type owns 80%. We find that in all three economies the optimal markup is increasing in the ratio $\theta^H(s_t)/\theta^L(s_t)$. As the profit tax falls, the slope falls but remains positive.

6 Quantitative Illustration

We next consider a simple, calibrated version of the model and compute the model-implied elasticity of the optimal markup with respect to aggregate output.

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 et al. (2017) to construct the functions $\theta^i : S \to \mathbb{R}_+$.[14] We assume 5 household types: the first four types capture the bottom 90 percent of the labor income distribution, while the last type captures the top decile. We partition the type space in this way to capture the non-monoticity of worker betas over the income distribution observed in the data. Worker-betas are monotonically falling in earnings levels throughout most of the income distribution then rise again at the very top of the distribution. Our type partition is able to capture this U-shape of household earnings exposure.

We use the U.S. Department of the Treasury estimates of the labor income distribution in 2019 to construct the long-run labor productivity distribution. We interpret a period in our model as a year and use annual data on real GDP from the Bureau of Economic Analysis from 1948 to 2023. We detrend the data and, using annual growth rates, we calculate unconditional probabilities for 4 distinct economic states: a severe recession (growth lower than 4.5 percent below trend), a mild recession (growth between 4.5 and 1.9 percent below trend), normal times (between 1.9 percent below trend and 1.7 percent above trend), and high growth (greater than 1.7 percent above trend). Equating the average labor income growth rate with the average growth rate of GDP, we use the worker betas to translate the percent change in GDP growth in each state into

[14] Guvenen et al. (2017) use a large, panel data set on individual earnings from the US Social Security Administration in which the same individuals can be tracked over time.

Table 1. Model-implied optimal markups and elasticities

State	$d\log Y$	μ	$\kappa = .25$		$\kappa = .06$	
			$d\log \mathcal{M}$	Elasticity	$d\log \mathcal{M}$	Elasticity
severe recession	–5.23	.09	1.32	–.25	.22	–.04
mild recession	–3.30	.09	1.30	–.39	.20	–.06
normal times	0	.49	0	—	0	—
high growth	3.21	.33	–.8	–.25	–.15	–.05

Notes. This table reports model-implied optimal markups and elasticities when price change frequencies are calibrated to match Nakamura and Steinsson (2008) (κ = .25) and Bils and Klenow (2004) (κ = .06), respectively.

percent changes in labor income growth for each household. We then translate these changes in labor income into changes in labor productivity.

We set $\beta = .98$. We use an elasticity of intertemporal substitution of .5 ($\gamma = 2$), following Hall (2009). We set the Frisch elasticity of labor supply to 2 ($\eta = .5$), in line with "macro" elasticities (Hall, 2009). The elasticity of substitution across goods, ρ, is set to 6, a value used commonly throughout the New Keynesian literature (McKay, Nakamura and Steinsson, 2016). For aggregate productivity, we use the annual series on total factor productivity (TFP) growth from Fernald (2014) to calculate average TFP growth in the years corresponding to each state.

We calibrate the share of sticky-price firms, κ, by converting estimates of the monthly frequency of price changes from Nakamura and Steinsson (2008) and Bils and Klenow (2004) into annual probabilities of a price remaining unchanged. Nakamura and Steinsson (2008) report that roughly 11 percent of prices change per month; the corresponding number from Bils and Klenow (2004) is 21 percent. Assuming that price changes are i.i.d. across months, these estimates imply κ =.25 and $\kappa = .06$, respectively.

With these parameter values, we solve numerically for the Ramsey optimum with equal Pareto weights: $\lambda_i = 1$ for all $i \in I$. In terms of fiscal policy, we assume profits are fully taxed and follow the implementation in (35) with $\bar{\mathcal{M}} = 1$.

Results. The first two columns of Table 1 report percent deviations of real GDP from trend and unconditional probabilities for each state. Model-implied optimal markups expressed in percent deviations from normal times, and their elasticities with respect to real GDP, are reported for the two specifications of κ.

We find that in either specification, the optimal markup is counter-cyclical. In our preferred specification of κ = .25, the optimal markup grows by 1.32 percent in severe recessions and falls by .8 percent in periods of high growth. These numbers imply a range for the elasticity of the optimal markup with respect to real GDP of –.25 to –.39. In the specification with more flexible prices (κ = .06), the optimal markup grows by .22 percent in severe recessions and falls by .15 percent in high growth periods; implied elasticities range from –.04 to –.06.

The countercyclicality of the optimal markup is the natural consequence of two features: countercyclical earnings inequality (in the data) and optimal monetary policy as prescribed by the model. As noted above, estimates of worker betas feature a striking pattern: earnings exposure to GDP growth is monotonically falling in income throughout the majority of the distribution. As output falls in a recession, the labor income of low-skilled workers declines disproportionately, resulting in an increase in earnings inequality. Countercyclical earnings inequality, coupled with the positive covariance between earnings inequality and the optimal markup, together imply countercyclical optimal markups.

The behavior of the optimal markup in our model is thereby consistent with work that documents countercyclical price markups and, more generally, a countercyclical labor wedge. It is firmly established that the labor wedge, defined as the ratio of the marginal product of labor to the marginal rate of substitution between consumption and leisure, is countercyclical (Hall, 1997; Chari, Kehoe and McGrattan, 2007). While the labor wedge can arise from both product and labor market distortions, a number of studies find evidence of countercyclical price markups: Bils (1987); Rotemberg and Woodford (1999); Bils, Klenow and Malin (2018). Given the countercyclicality of earnings inequality in the data, optimal monetary policy in our model is broadly consistent with these findings.

In terms of magnitudes, Bils, Klenow and Malin (2018) show that, depending on the wage measure used, the price markup elasticity with respect to real GDP can range from –.32 to –2.17. When using the frequency of price changes estimated by Nakamura and Steinsson (2008), the elasticities for the optimal markup implied by our model are consistent with the lower end (in absolute value) of this range.

To conclude, given the countercyclicality of earnings inequality present in the data, according to our calibrated model monetary policy should raise the nominal interest rate above the natural rate and contract the economy in recessions, and it should lower the nominal interest rate below the natural rate and stimulate the economy in expansions. This implies a monetary policy stance that is output destabilizing: GDP at the Ramsey optimum fluctuates more than one-for-one with its productively-efficient, flexible-price level. Yet by distorting the economy in this state-contingent pattern—by raising the real wage when earnings inequality is low, and lowering the real wage when earnings inequality is high—optimal monetary policy compresses the lifetime labor earnings distribution and achieves greater redistribution.

7 Conclusion

The variation in lifetime earnings accounted for by ex ante heterogeneity is often estimated to be above 50 percent and sometimes as high as 90 percent (Keane and Wolpin, 1997; Storesletten et al., 2004; Huggett et al., 2011).[15] Moreover, systematic exposure to short run fluctuations is unequal: Guvenen et al. (2014) and Guvenen et al. (2017) find that forecastable, between-group variation accounts for a large share of the total variation in earnings growth over the business cycle. Finally, a number of studies find substantial evidence of consumption smoothing in response to income shocks (Cochrane, 1991; Schulhofer-Wohl, 2011), and that a considerable fraction of individual income risk is insurable (Guvenen and Smith, 2014; Heathcote et al., 2014).

Motivated by these findings, in this paper we study optimal monetary policy in an economy with heterogeneous agents in which the key distributional motive for monetary policy is redistribution rather than insurance. We find that when household labor productivities fluctuate disproportionately over the business cycle, it is optimal for monetary policy to target a markup that co-varies positively with a sufficient statistic for labor income inequality. When we calibrate our model to reflect the unequal incidence of GDP fluctuations across households in the data, we find that countercyclical earnings inequality implies countercyclical optimal markups. In our baseline calibration, the elasticity of the optimal markup with respect to real GDP ranges from –.25 to –.39. This pattern is consistent with work that documents countercyclical price markups and, more broadly, a countercyclical labor wedge.

Our model shows that, insofar as earnings inequality is systematic and countercyclical, optimal monetary policy is output destabilizing. This policy prescription is at odds with optimal monetary policy in HANK, where missing insurance markets leads monetary policy to partially stabilize output in response to TFP shocks. Because both ex ante and ex post heterogeneity are present in the data, our paper highlights the potential difficulty of achieving distributional aims with monetary policy and the importance of accounting for the underlying sources of heterogeneity in earnings fluctuations before using monetary policy as a distributional tool.

[15] Keane and Wolpin (1997) find that between-type variation accounts for 90 percent of the total variance in lifetime utility. Storesletten, Telmer and Yaron (2004) conclude that roughly half of the variance of lifetime earnings is attributable to initial heterogeneity, while Huggett, Ventura and Yaron (2011) instead estimate that fraction at 62 percent.

References

Acharya, S., Challe, E., Dogra, K., 2023. Optimal monetary policy according to hank. American Economic Review 113, 1741–1782.

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.

Atkinson, A.B., Stiglitz, J.E., 1980. Lectures on Public Economics. McGraw-Hill, New York.

Barro, R.J., 1974. Are government bonds net wealth? Journal of Political Economy 82, 1095–1117.

Barro, R.J., 1979. On the determination of the public debt. Journal of Political Economy 87, 940–971.

Bassetto, M., 2014. Optimal fiscal policy with heterogeneous agents. Quantitative Economics 5, 675–704.

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

Bils, M., 1987. The cyclical behavior of marginal cost and price. The American Economic Review 77, 838–855.

Bils, M., Klenow, P.J., 2004. Some evidence on the importance of sticky prices. Journal of Political Economy 112, 947–985.

Bils, M., Klenow, P.J., Malin, B.A., 2018. Resurrecting the role of the product market wedge in recessions. American Economic Review 108, 1118–46.

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

Chari, V.V., Kehoe, P.J., McGrattan, E.R., 2007. Business cycle accounting. Econometrica 75, 781–836.

Cochrane, J.H., 1991. A simple test of consumption insurance. Journal of Political Economy 99, 957–976.

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., 2023. Optimal monetary policy with heterogeneous agents: Discretion, commitment, and timeless policy. NBER Working Paper 30961 .

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

Fernald, J.G., 2014. A Quarterly, Utilization-Adjusted Series on Total Factor Productivity. Working Paper Series 2012-19. Federal Reserve Bank of San Francisco.

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: Papers and Proceedings 107, 398–403.

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

Hall, R., 2009. Reconciling cyclical movements in the marginal value of time and the marginal product of labor. Journal of Political Economy 117, 281–323.

Hall, R.E., 1997. Macroeconomic fluctuations and the allocation of time. Journal of Labor Economics 15, pp. S223–S250.

Heathcote, J., Storesletten, K., Violante, G.L., 2014. Consumption and labor supply with partial insurance: An analytical framework. American Economic Review 104, 2075–2126.

Huggett, M., Ventura, G., Yaron, A., 2011. Sources of lifetime inequality. The American Economic Review 101, 2923–2954.

Judd, K., 1985. Redistributive taxation in a simple perfect foresight model. Journal of Public Economics 28, 59–83.

Keane, M.P., Wolpin, K.I., 1997. The career decisions of young men. Journal of Political Economy 105, 473–522.

Le Grand, F., Martin-Baillon, A., Ragot, X., 2024. Should monetary policy care about redistribution? optimal monetary and fiscal policy with heterogeneous agents. Working Paper .

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

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., Nakamura, E., Steinsson, J., 2016. The power of forward guidance revisited. American Economic Review 106, 3133–3158.

McKay, A., Wolf, C., 2022. Optimal policy rules in hank. Working Paper .

Mirrlees, J.A., 1971. An exploration in the theory of optimum income taxation12. The Review of Economic Studies 38, 175–208.

Nakamura, E., Steinsson, J., 2008. Five facts about prices: A reevaluation of menu cost models. The Quarterly Journal of Economics 123, 1415–1464.

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

Nekarda, C.J., Ramey, V.A., 2020. The cyclical behavior of the price-cost markup. Journal of Money, Credit and Banking 52, 319–353.

Niepelt, D., 2004. Tax smoothing versus tax shifting. Review of Economic Dynamics 7, 27–51.

Nuño, G., Thomas, C., 2022. Optimal redistributive inflation. Annals of Economics and Statistics 146, 3–64.

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

Ramsey, F.P., 1927. A contribution to the theory of taxation. The Economic Journal 37, 47–61.

Rotemberg, J.J., Woodford, M., 1999. The cyclical behavior of prices and costs, in: Taylor, J.B., Woodford, M. (Eds.), Handbook of Macroeconomics. Elsevier. volume 1 of *Handbook of Macroeconomics*. chapter 16, pp. 1051–1135.

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

Sheshinski, E., 1972. The optimal linear income tax. The Review of Economic Studies 39, 297–302.

Storesletten, K., Telmer, C.I., Yaron, A., 2004. Consumption and risk sharing over the life cycle. Journal of Monetary Economics 51, 609–633.

U.S. Department of the Treasury, 2019. Distribution of Income by Source (2019 Income Levels). Https://home.treasury.gov/system/files/131/Distribution-of-Income-by-Source-2019.pdf.

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

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

A Proofs for the Baseline Model

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:

$$\mu(s^t)U_c^i(s^t) - \mu(s^t)\Lambda^i(s^t)(1+\tau_c)P(s^t) = 0, \tag{42}$$

$$\mu(s^t)\frac{1}{\theta^i(s_t)}U_\ell^i(s^t) + \mu(s^t)\Lambda^i(s^t)(1-\tau_\ell)W(s^t) = 0, \tag{43}$$

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. Combining (42) and (43), we obtain the household's intratemporal condition:

$$-\frac{1}{\theta^i(s_t)}\frac{U_\ell^i(s^t)}{U_c^i(s^t)} = \frac{(1-\tau_\ell)W(s^t)}{(1+\tau_c)P(s^t)} \tag{44}$$

The first-order condition with respect to nominal bonds $b^i(s^t)$ is given by:

$$-\beta^t\mu(s^t)\Lambda^i(s^t) + \beta^{t+1}\sum_{s^{t+1}|s^t}\mu(s^{t+1})\Lambda^i(s^{t+1})(1+i(s^t)) = 0. \tag{45}$$

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

$$-\beta^t\mu(s^t)\Lambda^i(s^t)Q(s^{t+1}|s^t) + \beta^{t+1}\mu(s^{t+1})\Lambda^i(s^{t+1}) = 0. \tag{46}$$

The first-order condition with respect to equity shares $\sigma^i(s^t)$ is given by:

$$-\beta^t\mu(s^t)\Lambda^i(s^t)V(s^t) + \beta^{t+1}\sum_{s^{t+1}|s^t}\mu(s^{t+1})\Lambda^i(s^{t+1})[(1-\tau_\Pi)\Pi(s^{t+1}) + V(s^{t+1})] = 0 \tag{47}$$

The household's transversality conditions are given by: $\lim_{t\to\infty}\sum_{s^t}\mu(s^t)\Lambda^i(s^t)b^i(s^t) = 0$, $\lim_{t\to\infty}\sum_{s^t}\mu(s^t)\Lambda^i(s^t)V(s^t)\sigma^i(s^t) = 0$, and $\lim_{t\to\infty}\sum_{s^t}\mu(s^t)\Lambda^i(s^t)Q(s^{t+1}|s^t)z^i(s^{t+1}) = 0$.

A.2 Proof of Lemma 1

Markets are complete. The single, lifetime budget constraint of the household of type i can be represented as:

$$\sum_t\sum_{s^t}\hat{q}(s^t)\left[(1+\tau_c)c^i(s^t) - (1-\tau_\ell)\frac{W(s^t)}{P(s^t)}\ell^i(s^t)\right] = \sum_t\sum_{s^t}\hat{q}(s^t)\left[T(s^t) + (1+\sigma_0^i)(1-\tau_\Pi)\frac{\Pi(s^t)}{P(s^t)}\right]$$

where $\hat{q}(s^t)$ represents the Arrow-Debreu price of one unit of consumption in period t, history s^t, normalized so that $\hat{q}(s_0) = 1$, $W(s^t)/P(s^t)$ is the real wage, and $\Pi(s^t)/P(s^t)$ are real profits. Let $1/\varphi^i$ denote the Lagrange multiplier on this budget constraint. The household's first-order conditions w.r.t. $c^i(s^t)$ and $\ell^i(s^t)$ are given by:

$$\varphi^i U_c^i(s^t) - (1+\tau_c)\frac{\hat{q}(s^t)}{\beta^t \mu(s^t)} = 0 \qquad \text{and} \qquad \varphi^i \frac{1}{\theta^i(s_t)} U_\ell^i(s^t) + (1-\tau_\ell)\frac{\hat{q}(s^t)}{\beta^t \mu(s^t)}\frac{W(s^t)}{P(s^t)} = 0.$$

These conditions hold for all t, s^t and for all types $i \in I$ and imply that in equilibrium:

$$\varphi^i U_c^i(s^t) = \varphi^j U_c^j(s^t), \qquad \forall i, j \in I; \tag{48}$$

$$\varphi^i \frac{1}{\theta^i(s_t)} U_\ell^i(s^t) = \varphi^j \frac{1}{\theta^j(s_t)} U_\ell^j(s^t), \qquad \forall i, j \in I; \tag{49}$$

$$-\frac{1}{\theta^i(s_t)}\frac{U_\ell^i(s^t)}{U_c^i(s^t)} = -\frac{1}{\theta^j(s_t)}\frac{U_\ell^j(s^t)}{U_c^j(s^t)}, \qquad \forall i, j \in I; \tag{50}$$

for all t, s^t. These conditions and the resource constraints in (10) pin down the equilibrium allocation.

Consider now the static subproblem. Take any t, s^t and let $\rho_C(s^t)$ and $\rho_L(s^t)$ be the Lagrange multipliers on the constraints in (10). The first-order conditions w.r.t. $c^i(s^t)$ and $\ell^i(s^t)$ are given by:

$$\varphi^i U_c^i(s^t) - \rho_C(s^t) = 0 \qquad \text{and} \qquad \varphi^i \frac{1}{\theta^i(s_t)} U_\ell^i(s^t) + \rho_L(s^t) = 0.$$

These conditions imply (48)-(50). Again, these conditions, along with the resource constraints in (10), pin down the allocation. It follows that the solution to the sub-problem coincides with the equilibrium allocation. The envelope conditions for the static subproblem are given by: $U_C^m(s^t) = \varphi^i U_c^i(s^t)$ and $U_L^m(s^t) = \varphi^i \frac{1}{\theta^i(s_t)} U_\ell^i(s^t)$, for all $i \in I$. Next, with the separable and iso-elastic preferences assumed, conditions (48) and (49) can be written as:

$$\varphi^i c^i(s^t)^{-\gamma} = \varphi^j c^j(s^t)^{-\gamma} \quad \text{and} \quad \varphi^i \frac{1}{\theta^i(s_t)} \left[\frac{\ell^i(s^t)}{\theta^i(s_t)}\right]^\eta = \varphi^j \frac{1}{\theta^j(s_t)} \left[\frac{\ell^j(s^t)}{\theta^j(s_t)}\right]^\eta,$$

for all $\forall i, j \in I$. Combining these conditions with the resource constraints in (10), we obtain the linear expressions in (15) for individual consumption and labor with shares given by (16). Finally, note that conditions (44)-(47) imply conditions (11)-(14).

A.3 Derivation of Budget Implementability Conditions

We take the household's budget constraint for type $i \in I$, multiply both sides by $\Lambda^i(s^t)$, and use the household's FOCs in (42) and (43) to substitute out consumption and labor prices. Doing so, we obtain:

$$
\begin{aligned}
U_c^i(s^t)c^i(s^t) + \frac{1}{\theta^i(s_t)}U_\ell^i(s^t)\ell^i(s^t) =& \Lambda^i(s^t)z^i(s^t|s^{t-1}) - \Lambda^i(s^t)\sum_{s^{t+1}|s^t} Q(s^{t+1}|s^t)z^i(s^{t+1}|s^t) \\
&+ \Lambda^i(s^t)(1+i(s^{t-1}))b^i(s^{t-1}) - \Lambda^i(s^t)b^i(s^t) + \Lambda^i(s^t)P(s^t)\bar{T}(s^t) \\
&- \Lambda^i(s^t)V(s^t)(\sigma^i(s^t) - \sigma^i(s^{t-1})) + \Lambda^i(s^t)(1-\tau_\Pi)\Pi(s^t)\sigma^i(s^{t-1})
\end{aligned}
$$

where we let $\bar{T}(s^t) \equiv T(s^t) + (1-\tau_\Pi)\frac{\Pi(s^t)}{P(s^t)}$. Multiplying both sides by $\beta^t\mu(s^t)$, summing over t and s^t, and using the household's FOCs (45)-(47) to cancel terms, we obtain:

$$
\begin{aligned}
\sum_t \sum_{s^t} \beta^t \mu(s^t)\left[U_c^i(s^t)c^i(s^t) + \frac{1}{\theta^i(s_t)}U_\ell^i(s^t)\ell^i(s^t)\right] =& U_c^i(s_0)\bar{T} \\
&+ \sigma_0^i \Lambda^i(s_0)[(1-\tau_\Pi)\Pi((s_0) + V(s_0)]
\end{aligned} \tag{51}
$$

where $\bar{T} \equiv \frac{1}{U_C^m(s_0)}\sum_t \sum_{s^t} \beta^t\mu(s^t)\frac{U_C^m(s^t)}{(1+\tau_c)}\bar{T}(s^t)$. Note that for this latter term we have used the envelope condition $U_C^m(s^t) = \varphi^i U_c^i(s^t)$.

Next, we use the household FOC (47) to write the equity share price as follows:

$$
V(s^t) = \beta \sum_{s^{t+1}|s^t} \mu(s^{t+1}|s^t)\frac{\Lambda^i(s^{t+1})}{\Lambda^i(s^t)}[(1-\tau_\Pi)\Pi(s^{t+1}) + V(s^{t+1})].
$$

Iterating this forward and evaluating at time 0,

$$
V(s_0) = \sum_{t=1}^{\infty}\sum_{s^t} \beta^t \mu(s^t)\frac{\Lambda^i(s^t)}{\Lambda^i(s_0)}(1-\tau_\Pi)\Pi(s^t).
$$

Substituting this price per share at time 0 into 51, using the household FOC $U_c^i(s^t) = \Lambda^i(s^t)(1+\tau_c)P(s^t)$ as well as the solution to the static sub-problem described in Lemma 1 and its envelope conditions, we obtain the following implementability conditions:

$$
\begin{aligned}
&\sum_t \sum_{s^t} \beta^t \mu(s^t)\left[U_C^m(s^t)\omega_C^i(\varphi)C(s^t) + U_L^m(s^t)\omega_L^i(\varphi, s_t)L(s^t)\right] \\
&= U_C^m(s_0)\bar{T} + \sigma_0^i \frac{1-\tau_\Pi}{1+\tau_c}\sum_t \sum_{s^t} \beta^t \mu(s^t) U_C^m(s^t)\frac{\Pi(s^t)}{P(s^t)},
\end{aligned}
$$

where

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

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

A.4 Proof of Proposition 1

Necessity. All flexible-price firms set the same nominal price; similarly all sticky-price firms set the same nominal price. Combining this observation with the demand functions,

$$\frac{y^f(s^t)}{Y(s^t)} = \left[\frac{p_t^f(s^t)}{P(s^t)}\right]^{-\rho} \quad \text{and} \quad \frac{y^s(s^t)}{Y(s^t)} = \left[\frac{p_t^s(s^{t-1})}{P(s^t)}\right]^{-\rho}. \tag{52}$$

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 (19). Rearranging, dividing through by $P(s^t)$, and using the demand functions in (52), yields:

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

Combining the above condition with the household's intratemporal optimality condition (11), we obtain the equilibrium necessary condition

$$\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho} + \frac{U_L^m(s^t)}{U_C^m(s^t)}\frac{1}{\chi A(s_t)} = 0; \tag{53}$$

with χ defined by

$$\chi \equiv \left(\frac{\rho-1}{\rho}\right)\frac{(1-\tau_\ell)(1-\tau_r)}{1+\tau_c}. \tag{54}$$

The first-order condition for the sticky-price firm's problem is given by:

$$\sum_{s^t|s^{t-1}} Q(s^t|s^{t-1})Y(s^t)\left(\frac{p_t^s(s^{t-1})}{P(s^t)}\right)^{-\rho}\left\{\frac{p_t^s(s^{t-1})}{P(s^t)} - \left[(1-\tau_r)\left(\frac{\rho-1}{\rho}\right)\right]^{-1}\frac{W(s^t)}{P(s^t)A(s_t)}\right\} = 0. \tag{55}$$

Substituting in the equilibrium Arrow prices (13), and solving for $p_t^s(s^{t-1})$ yields:

$$p_t^s(s^{t-1}) = \left[(1-\tau_r)\left(\frac{\rho-1}{\rho}\right)\right]^{-1}\sum_{s^t|s^{t-1}}\frac{W(s^t)}{A(s_t)}q(s^t|s^{t-1}) \tag{56}$$

as stated in the main text. Next, using the CES demand function (52) and the equilibrium Arrow prices, the sticky price firm's FOC in (55) 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 (11), we obtain the equilibrium necessary condition

$$\sum_{s^t|s^{t-1}} U_C^m(s^t) y^s(s^t) \left\{ \left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho} + \frac{U_L^m(s^t)}{U_C^m(s^t)} \frac{1}{\chi A(s_t)} \right\} \mu(s^t|s^{t-1}) = 0; \tag{57}$$

with χ defined in (54).

What remains to be shown is necessity of the budget implementability conditions in (24). First, we derive an expression for real profits, $\Pi(s^t)/P(s^t)$, in terms of allocations alone. We can 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:

$$\Pi^f(s^t) = \left[(1-\tau_r)p_t^f(s^t) - \frac{W(s^t)}{A(s^t)}\right] y^f(s^t) \quad \text{and} \quad \Pi^s(s^t) = \left[(1-\tau_r)p_t^s(s^{t-1}) - \frac{W(s^t)}{A(s^t)}\right] y^s(s^t).$$

Combining these expressions with the demand functions,

$$\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho} = \frac{p_t^f(s^t)}{P(s^t)} \quad \text{and} \quad \left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho} = \frac{p_t^s(s^{t-1})}{P(s^t)}, \tag{58}$$

we find that real profits of these firms are given by:

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

Together, these imply that aggregate real profits can be written as:

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

Finally, the resource constraints in (31) imply 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, (11). This gives us the following expression for real profits:

$$\frac{\Pi(s^t)}{P(s^t)} = (1-\tau_r)C(s^t) + \frac{1+\tau_c}{1-\tau_\ell}\frac{U_L^m(s^t)}{U_C^m(s^t)}L(s^t) \tag{59}$$

Multiplying both sides of this by $U^m_C(s^t)/1+\tau_c$, we get:

$$\frac{U^m_C(s^t)}{1+\tau_c}\frac{\Pi(s^t)}{P(s^t)} = (1-\tau_r)\frac{U^m_C(s^t)}{1+\tau_c}C(s^t) + \frac{U^m_L(s^t)}{1-\tau_\ell}L(s^t),$$

Substituting this expression for real profits into (17), we obtain the following condition:

$$\sum_t \sum_{s^t} \beta^t \mu(s^t)\left[U^m_C(s^t)\omega^i_C(\varphi)C(s^t) + U^m_L(s^t)\omega^i_L(\varphi, s_t)L(s^t)\right]$$
$$= U^m_C(s_0)\bar{T} + \sigma^i_0 \frac{1-\tau_\Pi}{1-\tau_\ell}\sum_t \sum_{s^t}\beta^t \mu(s^t)\left[\frac{(1-\tau_\ell)(1-\tau_r)}{1+\tau_c}U^m_C(s^t)C(s^t) + U^m_L(s^t)L(s^t)\right].$$

Using the definition of χ, these conditions can be written as:

$$\sum_t \sum_{s^t} \beta^t \mu(s^t)\left[U^m_C(s^t)\omega^i_C(\varphi)C(s^t) + U^m_L(s^t)\omega^i_L(\varphi, s_t)L(s^t)\right]$$
$$= U^m_C(s_0)\bar{T} + \sigma^i_0 \frac{1-\tau_\Pi}{1-\tau_\ell}\sum_t \sum_{s^t}\beta^t \mu(s^t)\left[\chi\frac{\rho}{\rho-1}U^m_C(s^t)C(s^t) + U^m_L(s^t)L(s^t)\right]$$

Finally, we define $\vartheta \equiv \frac{1-\tau_\Pi}{1-\tau_\ell}$ and obtain the condition in (24).

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 $\mathcal{R}$, a policy $\mathcal{P}$, and a set of financial market positions $\mathcal{A}$, that support x as a sticky-price equilibrium; we construct these as follows.

First, we construct nominal prices as follows. 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. We set $p^s_t(s^{t-1}) = \mathcal{B}_t(s^{t-1})$. Next, we can decompose the sticky- and flexible-price firm output each into two components:

$$y^s(s^t) = \phi^s(s^{t-1})\Phi(s^t) \qquad \text{and} \qquad y^f(s^t) = \phi^f(s^t)\Phi(s^t). \tag{60}$$

where we set $\phi^s(s^{t-1}) \equiv \mathcal{B}_t(s^{t-1})^{-\rho}$. Therefore, $p^s_t(s^{t-1}) = \phi^s(s^{t-1})^{-1/\rho}$. The output decomposition in (60) implies $\Phi(s^t) = y^s(s^t)/\mathcal{B}_t(s^{t-1})^{-\rho}$ and $\phi^f(s^t) = y^f(s^t)/\Phi(s^t)$. Finally, we set the price of the flexible-price firm as follows: $p^f_t(s^t) = \phi^f(s^t)^{-1/\rho}$. These prices, along with the production function $Y(s^t) = \left[\int_{j\in\mathcal{J}} y^j(s^t)^{\frac{\rho-1}{\rho}}\mathrm{d}j\right]^{\frac{\rho}{\rho-1}}$, imply that the aggregate price level is given by:

$$P(s^t) = \left[\frac{Y(s^t)}{\Phi(s^t)}\right]^{-1/\rho}. \tag{61}$$

These prices furthermore ensure that the CES demand curves in (52) are satisfied. We set the money supply such that $M(s^t) = P(s^t)Y(s^t)$.

Next, we set the tax rates $(\tau_\ell, \tau_c, \tau_r)$ such that they jointly satisfy (54). For any strictly positive χ and $\rho > 1$, such tax rates exist. Combining this with conditions (53) and (57), we obtain the following two conditions:

$$\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho} + \frac{U_L^m(s^t)}{U_C^m(s^t)}\frac{1+\tau_c}{1-\tau_\ell}\left[(1-\tau_r)\left(\frac{\rho-1}{\rho}\right)\right]^{-1}\frac{1}{A(s_t)} = 0; \tag{62}$$

$$\sum_{s^t|s^{t-1}} \mu(s^t|s^{t-1})U_C^m(s^t)y^s(s^t)\left\{\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho} + \frac{U_L^m(s^t)}{U_C^m(s^t)}\frac{1+\tau_c}{1-\tau_\ell}\left[(1-\tau_r)\left(\frac{\rho-1}{\rho}\right)\right]^{-1}\frac{1}{A(s_t)}\right\} = 0. \tag{63}$$

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

$$W(s^t) = -\frac{U_L^m(s^t)}{U_C^m(s^t)}\left(\frac{1+\tau_c}{1-\tau_\ell}\right)P(s^t), \tag{64}$$

and therefore satisfy the household's intratemporal condition in (11). Substituting the above expression for the real wage into (62) and (63), combining these with the CES demand functions in (52), and rearranging gives us the following two conditions:

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

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

These conditions imply that both the flexible-price and the sticky-price firm's optimality conditions, (19) and (56), are satisfied. Finally, we set the Arrow prices, the nominal interest rate, and the ex-dividend share price so that they satisfy equilibrium conditions (12)-(14) for all $s^t \in S^t$.

What remains to be shown is that we can construct financial asset holdings such that the household's budget constraints are satisfied at this allocation in every history. Given $\bar{T}$, we first construct a sequence $\{\bar{T}(s^t)\}_{t\geq 0, s^t\in S^t}$ that satisfies the following condition:

$$\bar{T} = \frac{1}{U_c^m(s_0)(1+\tau_c)}\sum_t\sum_{s^t}\beta^t\mu(s^t)U_c^m(s^t)\bar{T}(s^t). \tag{65}$$

Given such a sequence $\{\bar{T}(s^t)\}$, we set transfers such that $T(s^t) = \bar{T}(s^t) - (1-\tau_\Pi)\Pi(s^t)/P(s^t)$ for all $s^t \in S^t$. Next, we take the household's budget constraint for type $i \in I$ for all periods and states following and including period r, history s^r; we multiply these budget constraints by $\beta^{t-r}\mu(s^t|s^r)\Lambda^i(s^t)$ and sum over all periods and states following and including period r, history s^r. Doing so, we get:

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

Using the household's FOCs (45)-(47) the above equation reduces to:

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

where $a^i(s^r)$ is defined as:

$$a^i(s^r)\equiv(1+i(s^{r-1}))b^i(s^{r-1})+z^i(s^r|s^{r-1})+\sigma^i(s^{r-1})\left[(1-\tau_\Pi)\Pi(s^r)+V(s^r)\right] \tag{66}$$

Therefore $a^i(s^r)$ represents the total nominal financial assets (cash-on-hand) that household i carries into period r, history s^r. Next, using conditions (42) and (43), we obtain the following expression for the total *real* financial assets that household i carries into period r, history s^r:

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

This can equivalently be written as follows:

$$\frac{a^i(s^r)}{P(s^t)}=\left(\frac{U_C^m(s^r)}{1+\tau_c}\right)^{-1}\sum_{t=r+1}^{\infty}\sum_{s^t}\beta^{t-r}\mu(s^t|s^r)\left[U_C^m(s^t)\omega_C^i(\varphi)C(s^t)+U_L^m(s^t)\omega_L^i(\varphi,s_t)L(s^t)\right]$$
$$-\sum_{t=r+1}^{\infty}\sum_{s^t}\beta^{t-r}\mu(s^t|s^r)\frac{U_C^m(s^t)}{U_C^m(s^r)}\bar{T}(s^t).$$

Finally, the government's budget constraint holds by Walras's law.

A.5 The Ramsey Optimum

Ramsey Planning 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 (30), subject to resource constraints: (31), $C(s^t)\leq Y(s^t)$, (22) and (23).

We allow for the inequality constraint: $C(s^t) \leq Y(s^t)$; that is, the planner has free disposal of the final good. 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 (22) and (23), respectively.

Proposition 3. *A Ramsey optimum x^* satisfies, for all $s^t \in S^t$,*

$$-\frac{\mathcal{W}_L(s^t) + (U_L^m(s^t) + U_{LL}^m(s^t)L(s^t))\left[\kappa\upsilon(s^{t-1})\frac{y^s(s^t)}{A(s_t)L(s^t)} + (1-\kappa)\xi(s^t)\frac{y^f(s^t)}{A(s_t)L(s^t)}\right]}{\mathcal{W}_C(s^t) + \chi(U_C^m(s^t) + U_{CC}^m(s^t)C(s^t))\left[\kappa\upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{1-1/\rho} + (1-\kappa)\xi(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1-1/\rho}\right]} = \frac{Y(s^t)}{L(s^t)}. \tag{67}$$

Proof. We write the planner's Lagrangian as follows:

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

along with Karush-Kuhn-Tucker conditions:

$$Y(s^t) - C(s^t) \geq 0, \quad \varsigma^C(s^t) \geq 0, \quad \text{and} \quad \varsigma^C(s^t)[Y(s^t) - C(s^t)] = 0, \quad \forall s^t \in S^t.$$

The FOCs with respect to $y^s(s^t)$ and $y^f(s^t)$ satisfy, respectively:

$$\begin{aligned}
0 =& \kappa\varsigma^Y(s^t)\left[\kappa y^s(s^t)^{\frac{\rho-1}{\rho}} + (1-\kappa)y^f(s^t)^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}-1} y^s(s^t)^{\frac{\rho-1}{\rho}} + \kappa\varsigma^L(s^t)\frac{y^s(s^t)}{A(s_t)} \\
&+ \kappa\upsilon(s^{t-1})y^s(s^t)\left\{\chi U_C^m(s^t)\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho}\frac{\rho-1}{\rho} + U_L^m(s^t)\frac{1}{A(s_t)}\right\},
\end{aligned} \tag{68}$$

$$\begin{aligned}
0 =& (1-\kappa)\varsigma^Y(s^t)\left[\kappa y^s(s^t)^{\frac{\rho-1}{\rho}} + (1-\kappa)y^f(s^t)^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}-1} y^f(s^t)^{\frac{\rho-1}{\rho}} \\
&+ (1-\kappa)\varsigma^L(s^t)\frac{y^f(s^t)}{A(s_t)} + (1-\kappa)\xi(s^t)y^f(s^t)\left\{\chi U_C^m(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho}\frac{\rho-1}{\rho} + U_L^m(s^t)\frac{1}{A(s_t)}\right\}.
\end{aligned} \tag{69}$$

Adding (68) to (69) and rearranging gives us the following condition:

$$-\frac{\varsigma^L(s^t)+U_L^m(s^t)\frac{1}{A(s_t)L(s^t)}\left[\kappa\upsilon(s^{t-1})y^s(s^t)+(1-\kappa)\xi(s^t)y^f(s^t)\right]}{\varsigma^Y(s^t)+\chi\left(1-\frac{1}{\rho}\right)U_C^m(s^t)\left[\kappa\upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{1-1/\rho}+(1-\kappa)\xi(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1-1/\rho}\right]}=\frac{Y(s^t)}{L(s^t)} \tag{70}$$

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

$$\varsigma^C(s^t)=\mathcal{W}_C(s^t)+\chi U_{CC}^m(s^t)Y(s^t)\left[\kappa\upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{1-1/\rho}+(1-\kappa)\xi(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1-1/\rho}\right], \tag{71}$$

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

$$\varsigma^Y(s^t)=\varsigma^C(s^t)+\chi\frac{1}{\rho}U_C^m(s^t)\left[\kappa\upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{1-1/\rho}+(1-\kappa)\xi(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1-1/\rho}\right] \tag{72}$$

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

$$\varsigma^L(s^t)=\mathcal{W}_L(s^t)+\kappa\upsilon(s^{t-1})y^s(s^t)U_{LL}^m(s^t)\frac{1}{A(s_t)}+(1-\kappa)\xi(s^t)y^f(s^t)U_{LL}^m(s^t)\frac{1}{A(s_t)}, \tag{73}$$

Combining (71) and (72) we get:

$$\begin{aligned}\varsigma^Y(s^t)=&\mathcal{W}_C(s^t)+\chi\left\{\left[U_{CC}^m(s^t)Y(s^t)+\frac{1}{\rho}U_C^m(s^t)\right]\right. \\ &\left.\times\left[\kappa\upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{1-1/\rho}+(1-\kappa)\xi(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1-1/\rho}\right]\right\}\end{aligned} \tag{74}$$

Substituting the above expression for $\varsigma^Y(s^t)$ and the expression for $\varsigma^L(s^t)$ in (73) into (70), and noting that $Y(s^t)=C(s^t)$, we obtain the optimality condition in (67). Furthermore, $\varsigma^C(s^t)>0$ for all $s^t\in S^t$. □

A.6 Proof of Theorem 1

From Proposition (3), at the Ramsey optimum:

$$-\frac{\mathcal{W}_L(s^t)+(1+\eta)U_L^m(s^t)\frac{Y(s^t)}{A(s_t)L(s^t)}\left[\kappa\upsilon(s^{t-1})\frac{y^s(s^t)}{Y(s^t)}+(1-\kappa)\xi(s^t)\frac{y^f(s^t)}{Y(s^t)}\right]}{\mathcal{W}_C(s^t)+\chi(1-\gamma)U_C^m(s^t)\left[\kappa\upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}}+(1-\kappa)\xi(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}}\right]}=\frac{Y(s^t)}{L(s^t)}. \tag{75}$$

where we have used the fact that with separable and iso-elastic utility, $\frac{U^m_{CC}(s^t)C(s^t)}{U^m_C(s^t)} = -\gamma$ and $\frac{U^m_{LL}(s^t)L(s^t)}{U^m_L(s^t)} = \eta$. Furthermore, note that $\mathcal{W}_C(s^t)$ and $\mathcal{W}_L(s^t)$ satisfy:

$$\mathcal{W}_C(s^t) = U^m_C(s^t) \sum_{i \in I} \pi^i \omega^i_C(\varphi) \left[\frac{\lambda^i}{\varphi^i} + \nu^i(1-\gamma) \right] \tag{76}$$

$$\mathcal{W}_L(s^t) = U^m_L(s^t) \sum_{i \in I} \pi^i \omega^i_L(\varphi, s_t) \left[\frac{\lambda^i}{\varphi^i} + \nu^i(1+\eta) \right]. \tag{77}$$

We proceed by characterizing the implicit monetary tax defined in (37). We define a function $\mathcal{I}(s_t)$ and a scalar $\mathcal{H}$ as follows:

$$\mathcal{I}(s_t) \equiv \sum_{i \in I} \pi^i \omega^i_L(\varphi, s_t) \left[\frac{\lambda^i}{\varphi^i} + \nu^i(1+\eta) \right] \qquad \text{and} \qquad \mathcal{H} \equiv (\chi^*)^{-1} \Omega_C, \tag{78}$$

Then (75), (37) (76), (77) imply that the optimal monetary wedge satisfies:

$$1 - \tau^*_M(s^t) = \frac{\mathcal{H} + (1-\gamma)\left[\kappa \upsilon(s^{t-1}) \left[\frac{y^s(s^t)}{Y(s^t)} \right]^{\frac{\rho-1}{\rho}} + (1-\kappa)\xi(s^t) \left[\frac{y^f(s^t)}{Y(s^t)} \right]^{\frac{\rho-1}{\rho}} \right]}{\mathcal{I}(s_t) + (1+\eta)\frac{Y(s^t)}{A(s_t)L(s^t)} \left[\kappa \upsilon(s^{t-1}) \frac{y^s(s^t)}{Y(s^t)} + (1-\kappa)\xi(s^t) \frac{y^f(s^t)}{Y(s^t)} \right]} \tag{79}$$

Threshold. We first consider the conditions under which $\tau^*_M(s^t) = 0$. In this state: $y^s(s^t) = y^f(s^t) = Y(s^t) = A(s_t)L(s^t)$. Furthermore, by the planner FOCs (68) and (69), $y^s(s^t) = y^f(s^t) = Y(s^t)$ if and only if $\xi(s^t) = \upsilon(s^{t-1})$; we state and prove this formally in Lemma 2. Therefore, in this case condition (79) reduces to:

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

Solving this for $\mathcal{I}(s_t)$ we obtain the following threshold: $\bar{\mathcal{I}}(s^{t-1}) = \mathcal{H} - (\eta+\gamma)\upsilon(s^{t-1})$. Therefore if $\mathcal{I}(s_t) = \bar{\mathcal{I}}(s^{t-1})$, the optimal monetary tax is equal to zero: $\tau^*_M(s^t) = 0$.

A fictitious tax wedge. We next define a fictitious tax wedge:

$$1 - \hat{\tau}(s^t) \equiv \frac{\mathcal{H} + (1-\gamma)\upsilon(s^{t-1})}{\mathcal{I}(s_t) + (1+\eta)\upsilon(s^{t-1})} \tag{80}$$

The wedge $1 - \hat{\tau}(s^t)$ is continuous and strictly decreasing in $\mathcal{I}(s_t)$, as all other terms are constants (conditional on s^{t-1}). Furthermore, note that $\hat{\tau}(s^t) = 0$ if and only if $\mathcal{I}(s_t) = \bar{\mathcal{I}}(s^{t-1})$. As a result, the fictitious tax $\hat{\tau}(s^t)$ trivially satisfies: $\hat{\tau}(s^t) > 0$ if and only if $\mathcal{I}(s_t) > \bar{\mathcal{I}}(s^{t-1})$, $\hat{\tau}(s^t) = 0$ if and only if $\mathcal{I}(s_t) = \bar{\mathcal{I}}(s^{t-1})$, and $\hat{\tau}(s^t) < 0$ if and only if $\mathcal{I}(s_t) < \bar{\mathcal{I}}(s^{t-1})$.

The optimal monetary wedge. We next characterize $y^f(s^t)$, $y^s(s^t)$, and the multipliers $\xi(s^t)$ and $\upsilon(s^{t-1})$ at the optimal allocation.

Lemma 2. *At the optimal allocation:*

(i) $y^f(s^t) > y^s(s^t)$ *if and only if* $\tau_M^*(s^t) > 0$, $y^f(s^t) = y^s(s^t)$ *if and only if* $\tau_M^*(s^t) = 0$, *and* $y^f(s^t) < y^s(s^t)$ *if and only if* $\tau_M^*(s^t) < 0$*; and*

(ii) $\xi(s^t) < \upsilon(s^{t-1})$ *if and only if* $\tau_M^*(s^t) > 0$, $\xi(s^t) = \upsilon(s^{t-1})$ *if and only if* $\tau_M^*(s^t) = 0$, *and* $\xi(s^t) > \upsilon(s^{t-1})$ *if and only if* $\tau_M^*(s^t) < 0$.

Proof. See Section A.7 of this appendix. □

We use this lemma in what follows. From our resource constraints, note that aggregate labor and aggregate output satisfy:

$$1 = \kappa \frac{y^s(s^t)}{A(s_t)L(s^t)} + (1-\kappa)\frac{y^f(s^t)}{A(s_t)L(s^t)}, \quad \text{and} \quad 1 = \kappa \left[\frac{y^s(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}} + (1-\kappa)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}}. \tag{81}$$

Combining (81) with (79) and rearranging, we obtain the following expression for the optimal monetary tax:

$$1 - \tau_M^*(s^t) = \frac{\mathcal{H} + (1-\gamma)\upsilon(s^{t-1}) + (1-\kappa)(1-\gamma)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}}(\xi(s^t) - \upsilon(s^{t-1}))}{\mathcal{I}(s_t) + (1+\eta)\upsilon(s^{t-1}) + (1-\kappa)(1+\eta)\frac{y^f(s^t)}{Y(s^t)}\frac{Y(s^t)}{A(s_t)L(s^t)}(\xi(s^t) - \upsilon(s^{t-1}))} \tag{82}$$

We want to compare this to the fictitious tax wedge defined in (80). In order to do so, we take the inverse wedges from (82) and (80), and combining the two we obtain:

$$\begin{aligned}\frac{1}{1-\tau_M^*(s^t)} =& \frac{1}{1-\hat{\tau}(s^t)} + \frac{(1-\kappa)(1+\eta)}{\mathcal{H} + (1-\gamma)\upsilon(s^{t-1})}\left[\frac{y^f(s^t)}{Y(s^t)}\frac{Y(s^t)}{A(s_t)L(s^t)}\right](\xi(s^t) - \upsilon(s^{t-1})) \\ &- \frac{(1-\kappa)(1-\gamma)}{\mathcal{H} + (1-\gamma)\upsilon(s^{t-1})}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{\frac{\rho-1}{\rho}}\frac{1}{1-\tau_M^*(s^t)}(\xi(s^t) - \upsilon(s^{t-1}))\end{aligned} \tag{83}$$

Next, if we combine (37) with the implementability condition (22), we obtain the following expression for the optimal monetary tax:

$$\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho} = (1 - \tau_M^*(s^t))\frac{Y(s^t)}{A(s_t)L(s^t)}. \tag{84}$$

Combining (84) with (83), we obtain the following condition:

$$\frac{1}{1-\tau_M^*(s^t)} = \frac{1}{1-\hat{\tau}(s^t)} + \frac{(1-\kappa)(\eta+\gamma)}{\mathcal{H} + (1-\gamma)\upsilon(s^{t-1})}\frac{y^f(s^t)}{Y(s^t)}\frac{Y(s^t)}{A(s_t)L(s^t)}(\xi(s^t) - \upsilon(s^{t-1})) \tag{85}$$

It follows that $\xi(s^t) < \upsilon(s^{t-1})$ if and only if $\tau_M^*(s^t) < \hat{\tau}(s^t)$, $\xi(s^t) = \upsilon(s^{t-1})$ if and only if $\tau_M^*(s^t) = \hat{\tau}(s^t)$, and $\xi(s^t) > \upsilon(s^{t-1})$ if and only if $\tau_M^*(s^t) > \hat{\tau}(s^t)$.

Consider again the case in which: $y^s(s^t) = y^f(s^t) = Y(s^t) = A(s_t)L(s^t)$. From Lemma 2 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 $\hat{\tau}(s^t) = 0$. This implies $\tau_M^*(s^t) = 0$ if and only if $\mathcal{I}(s_t) = \bar{\mathcal{I}}(s^{t-1})$.

Consider second the case in which $y^f(s^t) > y^s(s^t)$. From Lemma 2 we have that in this state, $\tau_M^*(s^t) > 0$ and $\xi(s^t) < \upsilon(s^{t-1})$. From the expression above, the latter implies $\tau_M^*(s^t) < \hat{\tau}(s^t)$. Therefore $0 < \tau_M^*(s^t) < \hat{\tau}(s^t)$. Finally, $\hat{\tau}(s^t) > 0$ implies $\mathcal{I}(s_t) > \bar{\mathcal{I}}(s^{t-1})$.

Consider third the case in which $y^f(s^t) < y^s(s^t)$. From Lemma 2 we have that in this state, $\tau_M^*(s^t) < 0$ and $\xi(s^t) > \upsilon(s^{t-1})$. From the expression above, the latter implies $\tau_M^*(s^t) > \hat{\tau}(s^t)$. Therefore $\hat{\tau}(s^t) < \tau_M^*(s^t) < 0$. Finally, $\hat{\tau}(s^t) < 0$ implies $\mathcal{I}(s_t) < \bar{\mathcal{I}}(s^{t-1})$.

We prove the converse statements by contradiction. Let $\mathcal{I}(s_t) > \bar{\mathcal{I}}(s^{t-1})$ and suppose that $\tau_M^*(s^t) < 0$. From Lemma 2, this implies $\xi(s^t) > \upsilon(s^{t-1})$, which further implies $\tau_M^*(s^t) > \hat{\tau}(s^t)$. It follows that $\hat{\tau}(s^t) < \tau_M^*(s^t) < 0$. But $\hat{\tau}(s^t) < 0$ is a contradiction of $\mathcal{I}(s_t) > \bar{\mathcal{I}}(s^{t-1})$. Therefore $\tau_M^*(s^t) > 0$. Similarly let $\mathcal{I}(s_t) < \bar{\mathcal{I}}(s^{t-1})$ and suppose that $\tau_M^*(s^t) > 0$. From Lemma 2, this implies $\xi(s^t) < \upsilon(s^{t-1})$, which further implies $\tau_M^*(s^t) < \hat{\tau}(s^t)$. It follows that $0 < \tau_M^*(s^t) < \hat{\tau}(s^t)$. But $\hat{\tau}(s^t) > 0$ is a contradiction of $\mathcal{I}(s_t) < \bar{\mathcal{I}}(s^{t-1})$. Therefore $\tau_M^*(s^t) < 0$.

We have thus proved the following statement:

Lemma 3. *At the optimal allocation: $\tau_M^*(s^t) > 0$ if and only if $\mathcal{I}(s_t) > \bar{\mathcal{I}}(s^{t-1})$; $\tau_M^*(s^t) = 0$ if and only if $\mathcal{I}(s_t) = \bar{\mathcal{I}}(s^{t-1})$; and $\tau_M^*(s^t) < 0$ if and only if $\mathcal{I}(s_t) < \bar{\mathcal{I}}(s^{t-1})$.*

Note that by combining this result with Lemma 2, it follows that $y^f(s^t) > y^s(s^t)$ if and only if $\mathcal{I}(s_t) > \bar{\mathcal{I}}(s^{t-1})$, $y^f(s^t) = y^s(s^t)$ if and only if $\mathcal{I}(s_t) = \bar{\mathcal{I}}(s^{t-1})$, and $y^f(s^t) < y^s(s^t)$ if and only if $\mathcal{I}(s_t) < \bar{\mathcal{I}}(s^{t-1})$;

Next, we combine (37) with the intratemporal condition in (11) and obtain the following condition:

$$\frac{W(s^t)}{P(s^t)} = (1-\tau_r)\left(\frac{\rho-1}{\rho}\right)(1-\tau_M^*(s^t))\frac{Y(s^t)}{L(s^t)} \tag{86}$$

Consider the fiscal implementation that sets $(1-\tau_r)^{-1}\frac{\rho}{\rho-1} = \bar{\mathcal{M}}$ for an arbitrary $\bar{\mathcal{M}} > 0$. Using this fiscal implementation, we combine the above expression with the implementability condition in (84) and infer that the aggregate markup satisfies:

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

with $\rho > 1$. It follows that: $\mathcal{M}^*(s^t) > \bar{\mathcal{M}}$ if and only if $y^f(s^t) > y^s(s^t)$, $\mathcal{M}^*(s^t) = \bar{\mathcal{M}}$ if and only if $y^f(s^t) = y^s(s^t)$, and $\mathcal{M}^*(s^t) < \bar{\mathcal{M}}$ if and only if $y^f(s^t) < y^s(s^t)$. Combining this result with Lemma 3, we obtain the result stated in the theorem.

A.7 Proof of Lemma 2

Part (i). Combining our expression for $L(s^t)$ from (81) with (84) and rearranging,

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

Combining this with the second expression in (81) we obtain:

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

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

Part (ii). The planner's first-order conditions (68) to (69) can be rewritten as follows:

$$0=\varsigma^Y(s^t)\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho}+\varsigma^L(s^t)\frac{1}{A(s_t)}+\upsilon(s^{t-1})\left\{\chi U_C^m(s^t)\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho}\frac{\rho-1}{\rho}+U_L^m(s^t)\frac{1}{A(s_t)}\right\}, \tag{87}$$

$$0=\varsigma^Y(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho}+\varsigma^L(s^t)\frac{1}{A(s_t)}+\xi(s^t)\left\{\chi U_C^m(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho}\frac{\rho-1}{\rho}+U_L^m(s^t)\frac{1}{A(s_t)}\right\}. \tag{88}$$

Combining these, we obtain the following condition which must hold at the Ramsey 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 implementability condition (22) 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:

$$\upsilon(s^{t-1})\left(\frac{\rho-1}{\rho}-\left[\frac{y^s(s^t)}{y^f(s^t)}\right]^{1/\rho}\right)-\xi(s^t)\left(\frac{\rho-1}{\rho}-1\right)\left[\frac{y^s(s^t)}{y^f(s^t)}\right]^{1/\rho}=\frac{\varsigma^Y(s^t)}{\chi U_C^m(s^t)}\left(\left[\frac{y^s(s^t)}{y^f(s^t)}\right]^{1/\rho}-1\right) \tag{89}$$

Consider the term $\varsigma^Y(s^t)/\chi U_C^m(s^t)$ on the right-hand side of condition (89). The scalar χ is strictly positive by definition. Next, because individual utility is strictly increasing in consumption, it is clear that $U_C^m(s^t)$ is strictly positive for all $s^t \in S^t$. Finally, $\varsigma^Y(s^t)$ is strictly positive for all $s^t \in S^t$; we verify this statement later on in the proof.

We first consider the case in which $y^s(s^t) = y^f(s^t)$. In this case, condition (89) reduces to:

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

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

We now prove the converse. Consider the case in which $\xi(s^t) = \upsilon(s^{t-1})$. In this case, condition (89) reduces to:

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

which implies: $y^s(s^t) = y^f(s^t)$.

Consider next the case in which $y^s(s^t) > y^f(s^t)$. Then condition (89) implies

$$\xi(s^t)\left(\frac{\rho-1}{\rho}-1\right)\left[\frac{y^s(s^t)}{y^f(s^t)}\right]^{1/\rho}<\upsilon(s^{t-1})\left(\frac{\rho-1}{\rho}-\left[\frac{y^s(s^t)}{y^f(s^t)}\right]^{1/\rho}\right),$$

which furthermore implies

$$\xi(s^t)\left(\frac{\rho-1}{\rho}-1\right)<\upsilon(s^{t-1})\left(\frac{\rho-1}{\rho}\left[\frac{y^s(s^t)}{y^f(s^t)}\right]^{-1/\rho}-1\right). \tag{90}$$

Furthermore note that

$$\upsilon(s^{t-1})\left(\frac{\rho-1}{\rho}\left[\frac{y^s(s^t)}{y^f(s^t)}\right]^{-1/\rho}-1\right)<\upsilon(s^{t-1})\left(\frac{\rho-1}{\rho}-1\right). \tag{91}$$

Combining (90) and (91), we infer:

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

which finally implies: $\xi(s^t) > \upsilon(s^{t-1})$. We can prove the converse statement by contradiction.

Consider next the case in which $y^s(s^t) < y^f(s^t)$. Then condition (89) implies

$$\upsilon(s^{t-1})\left(\frac{\rho-1}{\rho}-\left[\frac{y^s(s^t)}{y^f(s^t)}\right]^{1/\rho}\right) < \xi(s^t)\left(\frac{\rho-1}{\rho}-1\right)\left[\frac{y^s(s^t)}{y^f(s^t)}\right]^{1/\rho},$$

which furthermore implies

$$\upsilon(s^{t-1})\left(\frac{\rho-1}{\rho}\left[\frac{y^f(s^t)}{y^s(s^t)}\right]^{1/\rho}-1\right) < \xi(s^t)\left(\frac{\rho-1}{\rho}-1\right). \tag{92}$$

Furthermore note that

$$\upsilon(s^{t-1})\left(\frac{\rho-1}{\rho}-1\right) < \upsilon(s^{t-1})\left(\frac{\rho-1}{\rho}\left[\frac{y^f(s^t)}{y^s(s^t)}\right]^{1/\rho}-1\right). \tag{93}$$

Combining (92) and (93), we infer:

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

which finally implies: $\xi(s^t) < \upsilon(s^{t-1})$. We can prove the converse statement by contradiction.

We have thus shown that:

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

Combining this result with the result stated in part (i) of Lemma 2, we obtain the result stated in part (ii).

What remains to be shown is that the multiplier $\varsigma^Y(s^t)$ is strictly positive for all $s^t \in S^t$. We combine the planner's optimality conditions (71) and (72) and obtain the following expression for $\varsigma^Y(s^t)$:

$$\varsigma^Y(s^t) = \varsigma^C(s^t) + \frac{1}{\rho}\frac{U_C^m(s^t)}{U_{CC}^m(s^t)C(s^t)}(\varsigma^C(s^t) - \mathcal{W}_C(s^t))$$

Next, iso-elastic utility implies:

$$\varsigma^Y(s^t) = \left[1 - \frac{1}{\gamma\rho}\right]\varsigma^C(s^t) + \frac{1}{\gamma\rho}\mathcal{W}_C(s^t)$$

with $\rho > 1$ and $\gamma > 1$. It is clear from the definition of $\mathcal{W}(\cdot)$ that its first derivative with respect to aggregate consumption, denoted by $\mathcal{W}_C(s^t)$, is strictly positive. Furthermore, recall that in our proof of Proposition 3 [in Appendix A.5], we show that the Karush–Kuhn–Tucker multiplier $\varsigma^C(s^t)$ is strictly positive for all $s^t \in S^t$. Therefore $\varsigma^Y(s^t)$ is strictly positive for all $s^t \in S^t$.

B Equivalent Equilibrium Representation

In this section of the appendix we provide an equivalent characterization of the equilibrium in our economy using the forecast errors $\epsilon(s^t)$ defined in (26). Such a representation gives rise to a few auxiliary results, Lemmas (4), (5), and (6), that we find useful in later proofs.

First, note that equation (25) implies $p^s_t(s^{t-1}) = \epsilon(s^t)p^f_t(s^t)$. It follows from the CES demand equations in ()

$$\frac{y^f(s^t)}{Y(s^t)} = \left[\frac{p^f_t(s^t)}{P(s^t)}\right]^{-\rho} \qquad \text{and} \qquad \frac{y^s(s^t)}{Y(s^t)} = \left[\frac{p^s_t(s^{t-1})}{P(s^t)}\right]^{-\rho}. \tag{94}$$

that relative quantities across the two types of firms satisfy:

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

Therefore:

$$y^s(s^t) = \epsilon(s^t)^{-\rho} y^f(s^t) \tag{95}$$

Note that the flexible-price allocation coincides with $\epsilon(s^t) = 1$ for all $s^t \in S^t$.

Lemma 4. *For any $s^t \in S^t$, equilibrium aggregate output satisfies:*

$$Y(s^t) = A(s_t)\Delta(\epsilon(s^t))L(s^t) \tag{96}$$

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

$$\Delta(\epsilon) \equiv \left\{\frac{\left[\kappa\epsilon^{1-\rho} + (1-\kappa)\right]^{-\frac{1}{1-\rho}}}{\left[\kappa\epsilon^{-\rho} + (1-\kappa)\right]^{1/\rho}}\right\}^{\rho} > 0. \tag{97}$$

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

Proof. See Appendix B.1. □

Proposition 4 provides a succinct characterization of the efficiency wedge in this economy. When monetary policy implements flexible-price allocations—that is, when it sets $\epsilon(s^t) = 1$ 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 $\epsilon(s^t)$ deviates from one 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 leads to misallocation of inputs. This manifests as an efficiency wedge, or TFP loss. The term $\Delta(\epsilon)$ represents this efficiency wedge.

The following proposition provides a similar result for the equilibrium labor wedge in this economy.

Lemma 5. *For any $s^t \in S^t$, aggregate output and labor joint satisfy:*

$$-\frac{U_L^m(s^t)}{U_C^m(s^t)} = \chi\Gamma(\epsilon(s^t))A(s_t) \tag{98}$$

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

$$\Gamma(\epsilon) \equiv \left[\kappa\epsilon^{1-\rho} + (1-\kappa)\right]^{\frac{1}{\rho-1}} > 0. \tag{99}$$

The function Γ 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. See Appendix B.2. □

Finally, we can relate implicit monetary tax to the forecast error. In this representation we can think of the monetary tax as a function of ϵ, that is, $\tau_M : \mathbb{R}_+ \to (-\infty, 1)$.

Lemma 6. *The monetary tax satisfies:*

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

The function τ_M is continuous, differentiable, and satisfies: sign$(\tau_M(\epsilon))$ = sign$(\epsilon - 1)$ for all $\epsilon > 0$. It follows that:

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

Proof. See Appendix B.3. □

B.1 Proof of Lemma 4

We combine (95) with the resource constraints in (31) and obtain the following expressions for aggregate output and labor:

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

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 (96) 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 continuous and differentiable. 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:

$$\frac{d\Delta(\epsilon)}{d\epsilon} = \kappa\rho\Delta(\epsilon)\epsilon^{-\rho-1}\left\{\left[\kappa\epsilon^{-\rho} + (1-\kappa)\right]^{-1} - \left[\kappa\epsilon^{-\rho+1} + (1-\kappa)\right]^{-1}\epsilon\right\} \tag{100}$$

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 (100), $d\Delta(\epsilon)/d\epsilon > 0$ 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:

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

as was to be shown.

B.2 Proof of Lemma 5

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 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, 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 (98) 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 continuous and differentiable. Furthermore, $\Gamma(1) = 1$. Finally, the first derivative of $\Gamma(\epsilon)$ 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 $\epsilon > 0$.

B.3 Proof of Lemma 6

Combining the definition of the implicit monetary tax in 37 with (98) yields:

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

Combining the above expression with (96) 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 the functions for Δ and Γ from (97) and (99), we find that the implicit monetary tax satisfies:

$$1-\tau_M(\epsilon) = \frac{\kappa\epsilon^{-\rho} + (1-\kappa)}{\kappa\epsilon^{-(\rho-1)} + (1-\kappa)}. \tag{101}$$

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

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

Note that $\tau_M(\epsilon)$ is continuous and differentiable on the domain $\epsilon > 0$. With this expression we can prove the last part of Proposition 6. Note that the denominator is strictly positive for all $\epsilon > 0$. Furthermore $\kappa\epsilon^{-\rho} > 0$ for all $\epsilon > 0$. Therefore $\text{sign}(\tau_M(\epsilon)) = \text{sign}(\epsilon - 1)$ for all $\epsilon > 0$.

C Implementation

In the main text, Proposition 2 provides a particular implementation of the Ramsey optimum in terms of the behavior of aggregate price levels and the nominal interest rate. We now provide a more general version of this proposition and its proof.

First, from the sufficiency portion of our proof of Proposition 1 in Appendix A.4, we 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. We can thus index implementations by $\mathcal{B}_t(s^{t-1})$.

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 4. *Given $\mathcal{B}_t(s^{t-1}) > 0$, 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:*

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

Proof. See Appendix C.1. □

Proposition 2 presented in the main text is a special case of Proposition 4 with $\mathcal{B}_t(s^{t-1}) = P(s^{t-1})$.

C.1 Proof of Proposition 4

From the sufficiency portion of our proof of Proposition 1 in Appendix A.4, recall that for any equilibrium allocation, we can construct nominal prices as follows. Given $\mathcal{B}_t(s^{t-1}) > 0$, we set: $p_t^s(s^{t-1}) = \mathcal{B}_t(s^{t-1})$. This implies in terms of our output decomposition in 60 that $\phi^s(s^{t-1})^{-1/\rho} = \mathcal{B}_t(s^{t-1})$. It follows that:

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

The aggregate price level thereby satisfies:

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

We can write the deviation of the price level from the expected price as follows:

$$\log P(s^t) - \log \mathcal{B}_t(s^{t-1}) = -\frac{1}{\rho}(\log Y(s^t) - \log y^s(s^t)) \tag{102}$$

with $\rho > 1$. Lemma 2 states that $Y(s^t) > y^s(s^t)$ if and only if $\tau_M^*(s^t) > 0$, $Y(s^t) = y^s(s^t)$ if and only if $\tau_M^*(s^t) = 0$, and $Y(s^t) < y^s(s^t)$ if and only if $\tau_M^*(s^t) < 0$. It follows that:

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

Finally, combining this with the result in Lemma 3, the behavior of the aggregate price level stated in the proposition follows.

Next we turn to the nominal interest rate. The equilibrium nominal interest rate satisfies the Euler equation in 12:

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

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 102 into the above expression. Doing so, we obtain:

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

The next step in our proof requires the following lemma:

Lemma 7. *Let $\hat{Y}(s^t)$ denote the level of output in history s^t under flexible prices. At the optimal allocation, $Y(s^t) < \hat{Y}(s^t)$ if and only if $\tau_M^*(s^t) > 0$, $Y(s^t) = \hat{Y}(s^t)$ if and only if $\tau_M^*(s^t) = 0$, and $Y(s^t) > \hat{Y}(s^t)$ if and only if $\tau_M^*(s^t) < 0$.*

Proof. See Section C.2 of this appendix. □

Recall that Lemma 2 states that $Y(s^t) > y^s(s^t)$ if and only if $\tau_M^*(s^t) > 0$, $Y(s^t) = y^s(s^t)$ if and only if $\tau_M^*(s^t) = 0$, and $Y(s^t) < y^s(s^t)$ if and only if $\tau_M^*(s^t) < 0$. Therefore, combining the results of Lemmas 2 and 7, it follows that: $i(s^t) > \hat{\imath}(s^t)$ if and only if $\tau_M^*(s^t) > 0$, $i(s^t) = \hat{\imath}(s^t)$ if and only if $\tau_M^*(s^t) = 0$, and $i(s^t) < \hat{\imath}(s^t)$ if and only if $\tau_M^*(s^t) < 0$. The result stated in Proposition 2 follows by combining this with Lemma 3.

C.2 Proof of Lemma 7

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

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

These are 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 equivalent equilibrium representation articulated in Appendix B, 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 a system of 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 (4) and (5) 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:

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

since:

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

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

$$\log Y(s^t) - \log \hat{Y}(s^t) \approx \frac{1}{\eta + \gamma} \log \Gamma(\epsilon(s^t)) \tag{103}$$

From Lemma (5), we have that the function $\Gamma : \mathbb{R}_+ \to \mathbb{R}_+$ satisfies the following:

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

Finally, using this in equation (103) and the result of Lemma (6), it follows that:

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

as was to be shown.

D Extensions

In this section of the appendix, we characterize the three economies presented in Section 5. These three economies are extensions of our baseline model.

D.1 Partially State-Contingent Taxes

We consider the economy with partially state-contingent tax rates corresponding to Section 5.1. We begin by characterizing the set of equilibrium allocations, $\mathcal{X}^e$.

Proposition 5. *A feasible allocation $x \in \mathcal{X}$ can be implemented as an equilibrium allocation 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 $s^{t-1} \in S^{t-1}$,

$$\left[\frac{y^f(s|s^{t-1})}{Y(s|s^{t-1})}\right]^{-1/\rho} \frac{A(s)U_C^m(s|s^{t-1})}{-U_L^m(s|s^{t-1})} = \left[\frac{y^f(s'|s^{t-1})}{Y(s'|s^{t-1})}\right]^{-1/\rho} \frac{A(s')U_C^m(s'|s^{t-1})}{-U_L^m(s'|s^{t-1})}, \qquad \forall s, s'|s^{t-1}; \tag{104}$$

and (iii) condition (29) holds for every $i \in I$.

Proof. See Appendix E.1. □

Proposition 5 characterizes the set $\mathcal{X}^e$ when tax rates can be set one period in advance. Note that the conditions stated in part (ii) of the proposition 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:

$$\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho} \frac{A(s_t)U_C^m(s^t)}{-U_L^m(s^t)} = \frac{1}{\chi(s^{t-1})}, \qquad \forall s^t|s^{t-1}. \tag{105}$$

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

Ramsey Planning 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 (30) subject to (31) and (105).

We let $\beta^t \mu(s^t)(1-\kappa)\xi(s^t)$ denote the Lagrange multiplier on the implementability condition (105). The Ramsey optimum can be characterized as follows.

Proposition 6. *A Ramsey optimum x^* satisfies*

$$-\frac{\mathcal{W}_L(s^t) + \xi(s^t)U_{LL}^m(s^t)\frac{1}{A(s_t)}}{\mathcal{W}_C(s^t) + \xi(s^t)\chi(s^{t-1})U_{CC}^m(s^t)\left(\frac{y^f(s^t)}{Y(s^t)}\right)^{-1/\rho}} = \frac{Y(s^t)}{L(s^t)}, \qquad \forall s^t \in S^t. \tag{106}$$

Proof. The planner's FOCs with respect to $y^s(s^t)$ and $y^f(s^t)$ satisfy, respectively:

$$0 = \kappa\varsigma^Y(s^t)Y(s^t)^{1/\rho}y^s(s^t)^{\frac{\rho-1}{\rho}} + \kappa\varsigma^L(s^t)\frac{y^s(s^t)}{A(s_t)} \tag{107}$$

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

Summing these two conditions yields:

$$0 = \varsigma^Y(s^t)Y(s^t) + \varsigma^L(s^t)L(s^t) - \frac{1}{\rho}\xi(s^t)\chi(s^{t-1})U_C^m(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho}, \tag{108}$$

which can be rewritten as follows:

$$-\frac{\varsigma^L(s^t)}{\varsigma^Y(s^t) - \frac{1}{\rho}\xi(s^t)\chi(s^{t-1})U_C^m(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho}\frac{1}{Y(s^t)}} = \frac{Y(s^t)}{L(s^t)}. \tag{109}$$

Next, the FOCs with respect to $C(s^t)$ and $Y(s^t)$ combined imply:

$$\varsigma^Y(s^t) = \mathcal{W}_C(s^t) + \xi(s^t)\chi(s^{t-1})\left(\frac{y^f(s^t)}{Y(s^t)}\right)^{-1/\rho}\left[U_{CC}^m(s^t) + \frac{1}{\rho}U_C^m(s^t)\frac{1}{Y(s^t)}\right] \tag{110}$$

The FOC with respect to $L(s^t)$ satisfies:

$$\varsigma^L(s^t) = \mathcal{W}_L(s^t) + \xi(s^t)U_{LL}^m(s^t)\frac{1}{A(s_t)}. \tag{111}$$

Substituting the expression for $\varsigma^Y(s^t)$ in (110) and the expression for $\varsigma^L(s^t)$ in (111) into (109) we obtain the optimality condition in (106). □

D.2 Constrained Profit Taxation

We consider the economy with constrained profit taxation corresponding to Section 5.2. We state and solve the Ramsey problem. Again we let $\pi^i\nu^i$ denote the Lagrange multiplier on the implementability condition (24) of type $i \in I$ and subsume these into the maximand. Given ϑ, we can define a new pseudo-welfare function $\hat{\mathcal{W}}(\cdot)$ as follows:

$$\begin{aligned}\hat{\mathcal{W}}(C(s^t), L(s^t), s_t; \varphi, \nu, \lambda, \sigma, \chi, \vartheta) \equiv\ & \mathcal{W}(C(s^t), L(s^t), s_t; \varphi, \nu, \lambda) \\ & - \left(\vartheta\sum_{i\in I}\pi^i\nu^i\sigma_0^i\right)\left[\chi\frac{\rho}{\rho-1}U_C^m(s^t)C(s^t) + U_L^m(s^t)L(s^t)\right]\end{aligned} \tag{112}$$

where $\mathcal{W}(\cdot)$ is defined in (32). With this, we recast the Ramsey planning problem as follows.

Ramsey Planning 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:

$$\sum_t\sum_{s^t}\beta^t\mu(s^t)\hat{\mathcal{W}}(C(s^t), L(s^t), s_t; \varphi, \nu, \lambda, \sigma, \chi, \vartheta) - U_C^m(s_0)\bar{T}\sum_{i\in I}\pi^i\nu^i \tag{113}$$

subject to (31), (22), and (23).

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 (22) and (23), respectively. We obtain the following Ramsey optimality condition.

Proposition 7. *A Ramsey optimum x^* satisfies, for all $s^t \in S^t$,*

$$-\frac{\hat{\mathcal{W}}_L(s^t) + (U^m_L(s^t) + U^m_{LL}(s^t)L(s^t))\left[\kappa\upsilon(s^{t-1})\frac{y^s(s^t)}{A(s_t)L(s^t)} + (1-\kappa)\xi(s^t)\frac{y^f(s^t)}{A(s_t)L(s^t)}\right]}{\hat{\mathcal{W}}_C(s^t) + \chi(U^m_C(s^t) + U^m_{CC}(s^t)C(s^t))\left[\kappa\upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{1-1/\rho} + (1-\kappa)\xi(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1-1/\rho}\right]} = \frac{Y(s^t)}{L(s^t)}. \tag{114}$$

where

$$\hat{\mathcal{W}}_C(s^t) = \mathcal{W}_C(s^t) - \left(\vartheta\sum_{i\in I}\pi^i\nu^i\sigma^i_0\right)\chi\frac{\rho}{\rho-1}[U^m_C(s^t) + U^m_{CC}(s^t)C(s^t)] \tag{115}$$

$$\hat{\mathcal{W}}_L(s^t) = \mathcal{W}_L(s^t) - \left(\vartheta\sum_{i\in I}\pi^i\nu^i\sigma^i_0\right)[U^m_L(s^t) + U^m_{LL}(s^t)L(s^t)] \tag{116}$$

Proof. The planner's problem is the same as Ramsey problem in the baseline economy, 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 3, we obtain the expression in (114). The first derivatives of $\hat{\mathcal{W}}(\cdot)$ with respect to $C(s^t)$ and $L(s^t)$ are stated in (115) and (116). □

D.3 Constrained Profit Taxation and Partially State-Contingent Taxes

We consider the economy with constrained profit taxation *and* one-period-ahead tax rates corresponding to Section 5.3. Specifically, we let τ_c and τ_r at time t be contingent on s^{t-1}. We begin by characterizing the set of equilibrium allocations, $\mathcal{X}^e$.

Lemma 8. *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 $\varphi \equiv (\varphi^i)$, a scalar $\bar{T} \in \mathbb{R}$, and a weakly positive scalar $\vartheta \in \mathbb{R}_{\geq 0}$, such that parts (i)-(ii) of Proposition 5 are satisfied, and condition (24) holds for every $i \in I$.*

Proof. The proof is analogous to the proof of Proposition 5; see Appendix E.1. □

Next, we state the planner's problem as follows.

Ramsey Planning 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 (113) subject to (31) and (105).

We let $\beta^t\mu(s^t)(1-\kappa)\xi(s^t)$ denote the Lagrange multiplier on the implementability condition (105). The Ramsey optimum can be characterized as follows.

Lemma 9. *A Ramsey optimum x^* satisfies*

$$-\frac{\hat{\mathcal{W}}_L(s^t) + \xi(s^t)U^m_{LL}(s^t)\frac{1}{A(s_t)}}{\hat{\mathcal{W}}_C(s^t) + \xi(s^t)\chi(s^{t-1})U^m_{CC}(s^t)\left(\frac{y^f(s^t)}{Y(s^t)}\right)^{-1/\rho}} = \frac{Y(s^t)}{L(s^t)}, \qquad \forall s^t \in S^t. \tag{117}$$

Proof. The planner's problem is the same as in the proof of Proposition 6, with the exception of maximizing $\hat{\mathcal{W}}(s^t)$ instead of $\mathcal{W}(s^t)$. Therefore, we follow the same procedure as in the proof of Proposition 6, and we obtain the expression in (117). This is identical to the Ramsey optimality condition (106) but with $\mathcal{W}_C(s^t)$ and $\mathcal{W}_L(s^t)$ replaced by $\hat{\mathcal{W}}_C(s^t)$ and $\hat{\mathcal{W}}_L(s^t)$. □

E Proofs for Extensions

E.1 Proof of Proposition 5

Necessity. The necessity argument follows similar steps as the proof of Proposition 1. In particular, we combine the flexible-price firm's optimality condition (19) with the CES demand function (52) and the household's intratemporal optimality condition (11) 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 markup and taxes, and rewrite the previous condition as follows:

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

This is a necessary condition for an allocation to be supported in equilibrium. Note that the above is equivalent to the conditions stated in (104).

We can similarly combine the sticky-price firm optimality condition with the CES demand function (52) and the household's intratemporal optimality condition (11) and obtain the following equilibrium condition:

$$\sum_{s^t|s^{t-1}} U_C^m(s^t)y^s(s^t)\left\{\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho} + \frac{U_L^m(s^t)}{U_C^m(s^t)}\frac{1}{\chi(s^{t-1})A(s_t)}\right\}\mu(s^t|s^{t-1}) = 0; \tag{118}$$

for all $s^{t-1} \in S^{t-1}$. Therefore (118) 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 1.

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 5. We show that there exists a price system

$\mathcal{R}$, a policy $\mathcal{P}$, and a set of financial market positions $\mathcal{A}$, that support x as a sticky-price equilibrium; we construct these as follows.

First, we construct nominal prices as in the sufficiency portion of the proof of Proposition 1. We decompose output as in (60). Given $\mathcal{B}_t(s^{t-1}) > 0$, we set $\phi^s(s^{t-1}) \equiv \mathcal{B}_t(s^{t-1})^{-\rho}$ and prices as follows:

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

This implies $\Phi(s^t) = y^s(s^t)/\mathcal{B}_t(s^{t-1})^{-\rho}$ and $\phi^f(s^t) = y^f(s^t)/\Phi(s^t)$. These prices, along with the feasibility constraint (6), imply that the aggregate price level is given by (61). These prices furthermore ensure that the CES demand curves in (52) are satisfied. We set the money supply such that $M(s^t) = P(s^t)Y(s^t)$.

Next, note that the conditions stated in (104) 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:

$$\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho} + \frac{1}{\chi(s^{t-1})}\frac{U_L^m(s^t)}{U_C^m(s^t)}\frac{1}{A(s_t)} = 0. \tag{119}$$

These conditions 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:

$$\chi(s^{t-1}) = -\frac{\sum_{s^t|s^{t-1}} y^s(s^t)U_L^m(s^t)\frac{1}{A(s_t)}\mu(s^t|s^{t-1})}{\sum_{s^t|s^{t-1}} y^s(s^t)U_C^m(s^t)\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho}\mu(s^t|s^{t-1})} > 0. \tag{120}$$

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:

$$\frac{(1-\tau_\ell(s^{t-1}))(1-\tau_r(s^{t-1}))}{1+\tau_c(s^{t-1})} = \left(\frac{\rho-1}{\rho}\right)^{-1}\chi(s^{t-1}). \tag{121}$$

For any strictly positive $\chi(s^{t-1})$ and $\rho > 1$, such tax rates exist.

Combining (121) with condition (119), we obtain:

$$0 = \left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho} + \left(\frac{\rho-1}{\rho}\right)^{-1}\left[\frac{1+\tau_c(s^{t-1})}{(1-\tau_r(s^{t-1}))(1-\tau_\ell(s^{t-1}))}\right]\frac{U_L^m(s^t)}{U_C^m(s^t)}\frac{1}{A(s_t)}. \tag{122}$$

Furthermore, combining (121) with condition (120) and rearranging, we obtain:

$$0 = \sum_{s^t|s^{t-1}} \mu(s^t|s^{t-1})U_C^m(s^t)y^s(s^t)\left\{\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho} + \left(\frac{\rho-1}{\rho}\right)^{-1}\left[\frac{1+\tau_c(s^{t-1})}{(1-\tau_r(s^{t-1}))(1-\tau_\ell(s^{t-1}))}\right]\frac{U_L^m(s^t)}{U_C^m(s^t)}\frac{1}{A(s_t)}\right\}. \tag{123}$$

Next, we set the real wage as follows:

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

and therefore satisfy the household's intratemporal condition in (11). Substituting the above expression for the real wage into (122) and (123), we obtain:

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

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

Combining these with the CES demand functions in (52), 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 1.

E.2 Proof of Theorem 2

We substitute the expressions for $\mathcal{W}_C(s^t)$ and $\mathcal{W}_L(s^t)$ from (76) and (77) into (106) and obtain the following Ramsey optimality condition:

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

Therefore the optimal monetary wedge, as defined in (37), satisfies:

$$1-\tau_M^*(s^t) = \frac{\mathcal{H}(s^{t-1}) + \xi(s^t) \frac{U_{CC}^m(s^t)}{U_C^m(s^t)} \left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho}}{\mathcal{I}(s_t) + \xi(s^t) \frac{U_{LL}^m(s^t)}{U_L^m(s^t)} \frac{1}{A(s_t)}}. \tag{124}$$

where $\mathcal{I}(s_t)$ is defined in (78) and 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 in (84). Substituting the optimal monetary tax from (124) into (84) we obtain:

$$A(s_t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho} = \left\{\frac{\mathcal{H}(s^{t-1}) + \xi(s^t)\frac{U^m_{CC}(s^t)}{U^m_C(s^t)}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho}}{\mathcal{I}(s_t) + \xi(s^t)\frac{U^m_{LL}(s^t)}{U^m_L(s^t)}\frac{1}{A(s_t)}}\right\}\frac{Y(s^t)}{L(s^t)}.$$

With the iso-elastic preferences, this reduces to:

$$\mathcal{I}(s_t) + (\eta+\gamma)\frac{\xi(s^t)}{A(s_t)L(s^t)} - \mathcal{H}(s^{t-1})\frac{Y(s^t)}{A(s_t)L(s^t)}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1/\rho} = 0$$

Again using condition (84), we have that that optimal monetary tax satisfies:

$$\mathcal{I}(s_t) + (\eta+\gamma)\mathcal{H}(s^{t-1})\hat{\xi}(s^t) - \mathcal{H}(s^{t-1})(1-\tau_M(s^t))^{-1} = 0$$

where $\hat{\xi}(s^t) \equiv \frac{\xi(s^t)}{A(s_t)L(s^t)}\mathcal{H}(s^{t-1})^{-1}$. We define a function $g(\mathcal{I}(s_t), \tau_M(s^t)) \equiv \mathcal{I}(s_t) + \mathcal{H}(s^{t-1})[(\eta+\gamma)\hat{\xi}(s^t) - (1-\tau_M(s^t))^{-1}]$. The optimal monetary tax satisfies: $g(\mathcal{I}(s_t), \tau_M^*(s^t)) = 0$. By the implicit function theorem:

$$\frac{d\tau_M^*(s^t)}{d\mathcal{I}(s_t)} = -\frac{dg/d\mathcal{I}(s_t)}{dg/d\tau_M^*(s^t)} = -\frac{1}{\mathcal{H}(s^{t-1})\left\{(\eta+\gamma)\frac{d\hat{\xi}(s^t)}{d\tau_M(s^t)} - (1-\tau_M^*(s^t))^{-2}\right\}}$$

Therefore the derivative of the optimal monetary tax satisfies:

$$\frac{d\tau_M^*(s^t)}{d\mathcal{I}(s_t)} = \frac{1}{\mathcal{H}(s^{t-1})}\left\{(1-\tau_M^*(s^t))^{-2} - (\eta+\gamma)\frac{d\hat{\xi}(s^t)}{d\tau_M^*(s^t)}\right\}^{-1} \tag{125}$$

Our next goal is to obtain an expression for $\hat{\xi}(s^t)$. The planner's optimality condition in (107) can be written as: $\varsigma^L(s^t) = -\varsigma^Y(s^t)A(s_t)\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho}$. Substituting this into (108) and rearranging, we obtain:

$$\frac{\xi(s^t)}{A(s_t)L(s^t)} = \rho\frac{\varsigma^Y(s^t)}{\chi(s^{t-1})U^m_C(s^t)}\left[\frac{Y(s^t)}{A(s_t)L(s^t)}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1/\rho} - \left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1/\rho}\right]. \tag{126}$$

Therefore

$$\hat{\xi}(s^t) = \rho\frac{\mathcal{H}(s^{t-1})^{-1}}{\chi(s^{t-1})U^m_C(s^t)}\varsigma^Y(s^t)\left[\frac{Y(s^t)}{A(s_t)L(s^t)}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1/\rho} - \left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1/\rho}\right]. \tag{127}$$

In what follows, we use the equivalent equilibrium representation in Section (B). Combining (101). with (84), 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, $y^s(s^t) = \epsilon(s^t)^{-\rho}y^f(s^t)$ implies that 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 (127) as follows:

$$\hat{\xi}(s^t) = \rho\frac{\mathcal{H}(s^{t-1})^{-1}}{\chi(s^{t-1})U^m_C(s^t)}\varsigma^Y(s^t)\left[\frac{(1-\kappa)(1-\epsilon(s^t))}{\kappa\epsilon(s^t)^{-\rho}+(1-\kappa)}\right] \tag{128}$$

Next we use the planner's optimality condition in (110). Substituting in for $\xi(s^t)$ from (126) into (110), and solving for $\varsigma^Y(s^t)$ we get:

$$\varsigma^Y(s^t) = \mathcal{W}_C(s^t)\left[1-(1-\gamma\rho)\left[1-\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{-1/\rho}\frac{A(s_t)L(s^t)}{Y(s^t)}\right]\right]^{-1}. \tag{129}$$

Furthermore, note that

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

Substituting this into (129) gives us:

$$\varsigma^Y(s^t) = \mathcal{W}_C(s^t)\frac{\kappa\epsilon(s^t)^{1-\rho}+(1-\kappa)}{\kappa\epsilon(s^t)^{1-\rho}+(1-\gamma\rho)(1-\kappa)\epsilon(s^t)+\gamma\rho(1-\kappa)}$$

Finally, substituting the above expression for $\varsigma^Y(s^t)$ into (128), we obtain the following expression for $\hat{\xi}(s^t)$ as a function of $\epsilon(s^t)$ and $\tau_M(s^t)$:

$$\hat{\xi}(s^t) = \rho\Sigma(\epsilon(s^t))(1-\tau_M(s^t))^{-1} \quad \text{where} \quad \Sigma(\epsilon) \equiv \frac{(1-\kappa)(1-\epsilon)}{\kappa\epsilon^{1-\rho}+(1-\gamma\rho)(1-\kappa)\epsilon+\gamma\rho(1-\kappa)}$$

Derivative of $\tau_M(s^t)$. The derivative of the optimal monetary "tax" satisfies (125). Evaluating this derivative at the benchmark in which $\tau_M(s^t) = 0$, we have:

$$\left.\frac{d\tau_M(s^t)}{d\mathcal{I}(s_t)}\right|_{\tau_M(s^t)=0} = \frac{1}{\mathcal{H}(s^{t-1})}\left\{1-(\eta+\gamma)\left.\frac{d\hat{\xi}(s^t)}{d\tau_M(s^t)}\right|_{\tau_M(s^t)=0}\right\}^{-1} \tag{130}$$

The derivative of $\hat{\xi}$ with respect to τ_M is given by:

$$\frac{d\hat{\xi}}{d\tau_M} = \rho\left\{\Sigma(\epsilon)(1-\tau_M)^{-2} + (1-\tau_M)^{-1}\frac{d\Sigma}{d\epsilon}\frac{d\epsilon}{d\tau_M}\right\}. \tag{131}$$

where $\frac{d\Sigma(\epsilon)}{d\epsilon} = -(1-\kappa)\frac{\kappa\epsilon^{1-\rho}+(1-\rho)\kappa\epsilon^{-\rho}(1-\epsilon)+(1-\kappa)}{(\kappa\epsilon^{1-\rho}+(1-\gamma\rho)(1-\kappa)\epsilon+\gamma\rho(1-\kappa))^2}$.

Next, we obtain an expression for $d\epsilon/d\tau_M$ as follows. The monetary tax satisfies equation (101). Rearranging, $\kappa\epsilon^{1-\rho} - \kappa\epsilon^{-\rho} - \tau_M\kappa\epsilon^{1-\rho} - \tau_M(1-\kappa) = 0$. We define a function $\varrho(\tau_M,\epsilon) \equiv (1-\tau_M)\kappa\epsilon^{1-\rho} - \kappa\epsilon^{-\rho} - \tau_M(1-\kappa)$. By the implicit function theorem:

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

Therefore the last term in (131) satisfies:

$$\begin{aligned}\Sigma(\epsilon)(1-\tau_M)^{-2} + (1-\tau_M)^{-1}\frac{d\Sigma}{d\epsilon}\frac{d\epsilon}{d\tau_M} =& \frac{(1-\kappa)(1-\epsilon)}{\kappa\epsilon^{1-\rho}+(1-\gamma\rho)(1-\kappa)\epsilon+\gamma\rho(1-\kappa)}(1-\tau_M)^{-2} \\ &- \left\{\frac{1-\kappa}{1-\tau_M}\left[\frac{\kappa\epsilon^{1-\rho}+(1-\rho)\kappa\epsilon^{-\rho}(1-\epsilon)+(1-\kappa)}{(\kappa\epsilon^{1-\rho}+(1-\gamma\rho)(1-\kappa)\epsilon+\gamma\rho(1-\kappa))^2}\right]\right. \\ &\left.\times\left[\frac{\kappa\epsilon^{1-\rho}+(1-\kappa)}{(1-\rho)(1-\tau_M)\kappa\epsilon^{-\rho}+\rho\kappa\epsilon^{-\rho-1}}\right]\right\}\end{aligned} \tag{132}$$

Evaluating this term at $\tau_M = 0$, or equivalently at $\epsilon = 1$, we have:

$$\left[\Sigma(\epsilon)(1-\tau_M)^{-2} + (1-\tau_M)^{-1}\frac{d\Sigma}{d\epsilon}\frac{d\epsilon}{d\tau_M}\right]_{\tau_M=0} = -\left(\frac{1-\kappa}{\kappa}\right)$$

Therefore $d\hat{\xi}/d\tau_M\Big|_{\tau_M=0} = -\rho\frac{1-\kappa}{\kappa}$. Substituting this into (130) yields

$$\left.\frac{d\tau_M(s^t)}{d\mathcal{I}(s_t)}\right|_{\tau_M(s^t)=0} = \frac{1}{\bar{\mathcal{I}}(s^{t-1})\left[1+\rho(\eta+\gamma)\frac{1-\kappa}{\kappa}\right]}[> 0.$$

where $\bar{\mathcal{I}}(s^{t-1}) \equiv \mathcal{H}(s^{t-1})$.

Consider now the markup. The equilibrium condition in (86) implies that the optimal markup satisfies:

$$1 = \frac{\mathcal{M}^*(s^t)}{\bar{\mathcal{M}}}(1-\tau_M^*(s^t))\frac{Y(s^t)}{A(s^t)L(s^t)}$$

Taking logs and rearranging:

$$\log\mathcal{M}^*(s^t) - \log\bar{\mathcal{M}} = -\log(1-\tau_M^*(s^t)) - \log\frac{Y(s^t)}{A(s^t)L(s^t)}$$

Therefore, to a first-order Taylor approximation around the point $\mathcal{I}(s_t) = \bar{\mathcal{I}}(s^{t-1})$,

$$\log\mathcal{M}^*(s^t) - \log\bar{\mathcal{M}} \approx \tau_M^*(s^t) \approx \frac{1}{\bar{\mathcal{I}}(s^{t-1})\left[1+\rho(\eta+\gamma)\frac{1-\kappa}{\kappa}\right]}[\mathcal{I}(s_t) - \bar{\mathcal{I}}(s^{t-1})],$$

as was to be shown.

E.3 Proof of Theorem 3

The Ramsey optimum satisfies (114). Substituting into (114) our expressions for $\hat{\mathcal{W}}_C(s^t)$ and $\hat{\mathcal{W}}_L(s^t)$ from (115) and (116), as well as our expressions for $\mathcal{W}_C(s^t)$ and $\mathcal{W}_L(s^t)$ from (76) and (77), and solving for the implicit optimal monetary wedge, we get:

$$
\begin{aligned}
1-\tau^*_M(s^t) = &\left\{ (\chi^*)^{-1}\sum_{i\in I}\pi^i\omega^i_C(\varphi)\left[\frac{\lambda^i}{\varphi^i}+\nu^i(1-\gamma)\right]\right.\\
&\left. +(1-\gamma)\left[\kappa\upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{1-1/\rho}+(1-\kappa)\xi(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1-1/\rho}-\frac{\rho}{\rho-1}\vartheta\sum_{i\in I}\pi^i\nu^i\sigma^i\right]\right\}\\
&\times\left\{\sum_{i\in I}\pi^i\omega^i_L(\varphi,s_t)\left[\frac{\lambda^i}{\varphi^i}+\nu^i(1+\eta)\right]\right.\\
&\left. +(1+\eta)\left[\kappa\upsilon(s^{t-1})\frac{y^s(s^t)}{A(s_t)L(s^t)}+(1-\kappa)\xi(s^t)\frac{y^f(s^t)}{A(s_t)L(s^t)}-\vartheta\sum_{i\in I}\pi^i\nu^i\sigma^i\right]\right\}^{-1}
\end{aligned}
$$

Therefore the optimal monetary wedge satisfies:

$$
1-\tau^*_M(s^t) = \frac{\mathcal{H}+(1-\gamma)\left[\kappa\upsilon(s^{t-1})\left[\frac{y^s(s^t)}{Y(s^t)}\right]^{1-1/\rho}+(1-\kappa)\xi(s^t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1-1/\rho}-\frac{\rho}{\rho-1}\vartheta\sum_{i\in I}\pi^i\nu^i\sigma^i_0\right]}{\mathcal{I}(s_t)+(1+\eta)\left[\kappa\upsilon(s^{t-1})\frac{y^s(s^t)}{A(s_t)L(s^t)}+(1-\kappa)\xi(s^t)\frac{y^f(s^t)}{A(s_t)L(s^t)}-\vartheta\sum_{i\in I}\pi^i\nu^i\sigma^i_0\right]}. \tag{133}
$$

where $\mathcal{I}(s_t)$ and $\mathcal{H}$ are defined in (78).

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 (133) reduces to:

$$
1=\frac{\mathcal{H}+(1-\gamma)\left[\kappa\upsilon(s^{t-1})+(1-\kappa)\xi(s^t)-\frac{\rho}{\rho-1}\vartheta\sum_{i\in I}\pi^i\nu^i\sigma^i_0\right]}{\mathcal{I}(s_t)+(1+\eta)\left[\kappa\upsilon(s^{t-1})+(1-\kappa)\xi(s^t)-\vartheta\sum_{i\in I}\pi^i\nu^i\sigma^i_0\right]}
$$

Furthermore, conditions (87) and (88) imply that $\xi(s^t)=\upsilon(s^{t-1})$ in this state. Therefore:

$$
1=\frac{\mathcal{H}+(1-\gamma)\upsilon(s^{t-1})-(1-\gamma)\frac{\rho}{\rho-1}\vartheta\sum_{i\in I}\pi^i\nu^i\sigma^i_0}{\mathcal{I}(s_t)+(1+\eta)\upsilon(s^{t-1})-(1+\eta)\vartheta\sum_{i\in I}\pi^i\nu^i\sigma^i_0}
$$

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

$$
\bar{\mathcal{I}}_\vartheta(s^{t-1})=\mathcal{H}-(\eta+\gamma)\upsilon(s^{t-1})+\left[(1+\eta)-(1-\gamma)\frac{\rho}{\rho-1}\right]\vartheta\sum_{i\in I}\pi^i\nu^i\sigma^i_0.
$$

Therefore if $\mathcal{I}(s_t)=\bar{\mathcal{I}}_\vartheta(s^{t-1})$, the optimal monetary tax is equal to zero: $\tau^*_M(s^t)=0$. Finally, letting$\bar{\mathcal{I}}_0(s^{t-1})=\mathcal{H}-(\eta+\gamma)\upsilon(s^{t-1})$, we obtain the following expression:

$$
\bar{\mathcal{I}}_\vartheta(s^{t-1})=\bar{\mathcal{I}}_0(s^{t-1})+\left[(1+\eta)+\frac{\rho}{\rho-1}(\gamma-1)\right]\vartheta\sum_{i\in I}\pi^i\nu^i\sigma^i_0. \tag{134}
$$

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

$$1-\hat{\tau}_{\vartheta}(s^t) \equiv \frac{\mathcal{H}+(1-\gamma)\upsilon(s^{t-1})-(1-\gamma)\frac{\rho}{\rho-1}\vartheta\sum_{i\in I}\pi^i\nu^i\sigma_0^i}{\mathcal{I}(s_t)+(1+\eta)\upsilon(s^{t-1})-(1+\eta)\vartheta\sum_{i\in I}\pi^i\nu^i\sigma_0^i} \tag{135}$$

The wedge $1-\hat{\tau}(s^t)$ is continuous and strictly decreasing in $\mathcal{I}(s_t)$, as all other terms are constants (conditional on s^{t-1}). Furthermore, note that $\hat{\tau}(s^t)=0$ if and only if $\mathcal{I}(s_t)=\bar{\mathcal{I}}_{\vartheta}(s^{t-1})$. As a result, the fictitious tax $\hat{\tau}(s^t)$ trivially satisfies:

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

The remainder of the proof follows the exact same steps as in the proof of Theorem 1.

E.4 Proof of Theorem 4

The Ramsey optimum satisfies (117). Substituting in our expressions for $\hat{\mathcal{W}}_C(s^t)$ and $\hat{\mathcal{W}}_L(s^t)$ from (115) and (116):

$$-\frac{\mathcal{W}_L(s^t)-\left(\vartheta\sum_{i\in I}\pi^i\nu^i\sigma^i\right)\left[U_L^m(s^t)+U_{LL}^m(s^t)L(s^t)\right]+\xi(s^t)U_{LL}^m(s^t)\frac{1}{A(s_t)}}{\mathcal{W}_C(s^t)-\chi(s^{t-1})\frac{\rho}{\rho-1}\left(\vartheta\sum_{i\in I}\pi^i\nu^i\sigma^i\right)\left[U_C^m(s^t)+U_{CC}^m(s^t)C(s^t)\right]+\xi(s^t)\chi(s^{t-1})U_{CC}^m(s^t)\left(\frac{y^f(s^t)}{Y(s^t)}\right)^{-1/\rho}}$$
$$=\frac{Y(s^t)}{L(s^t)}$$

This optimality condition, along with our expressions for $\mathcal{W}_C(s^t)$ and $\mathcal{W}_L(s^t)$ in (76) and (77), imply that the optimal monetary wedge, defined in (37), satisfies:

$$1-\tau_M^*(s^t)=\frac{\mathcal{H}_{\vartheta}(s^{t-1})+\xi(s^t)\frac{U_{CC}^m(s^t)}{U_C^m(s^t)}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho}-(1-\gamma)\frac{\rho}{\rho-1}\vartheta\sum_{i\in I}\pi^i\nu^i\sigma_0^i}{\mathcal{I}(s_t)+\xi(s^t)\frac{U_{LL}^m(s^t)}{U_L^m(s^t)}\frac{1}{A(s_t)}-(1+\eta)\vartheta\sum_{i\in I}\pi^i\nu^i\sigma_0^i}. \tag{136}$$

where $\mathcal{I}(s_t)$ is defined in (78) and we let $\mathcal{H}_{\vartheta}(s^{t-1})\equiv\chi(s^{t-1})^{-1}\Omega_C(\varphi)>0$.

First, note that when $\xi(s^t)=0$, the constraint is slack. Therefore, $\tau_M^*(s^t)=0$ if and only if

$$\mathcal{I}(s_t)=\bar{\mathcal{I}}_{\vartheta}(s^{t-1})\equiv\mathcal{H}_{\vartheta}(s^{t-1})+\left[(1+\eta)-(1-\gamma)\frac{\rho}{\rho-1}\right]\vartheta\sum_{i\in I}\pi^i\nu^i\sigma_0^i$$

Note that $\bar{\mathcal{I}}_0(s^{t-1})\equiv\mathcal{H}_0(s^{t-1})$. Substituting the optimal monetary wedge from (136) into (84) we obtain:

$$A(s_t)\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho}=\left\{\frac{\mathcal{H}_{\vartheta}(s^{t-1})+\xi(s^t)\frac{U_{CC}^m(s^t)}{U_C^m(s^t)}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{-1/\rho}-(1-\gamma)\frac{\rho}{\rho-1}\vartheta\sum_{i\in I}\pi^i\nu^i\sigma^i}{\mathcal{I}(s_t)+\xi(s^t)\frac{U_{LL}^m(s^t)}{U_L^m(s^t)}\frac{1}{A(s_t)}-(1+\eta)\vartheta\sum_{i\in I}\pi^i\nu^i\sigma^i}\right\}\frac{Y(s^t)}{L(s^t)}.$$

With the iso-elastic preferences, this reduces to:

$$0 = \mathcal{I}(s_t) + (\eta+\gamma)\frac{\xi(s^t)}{A(s_t)L(s^t)} - \mathcal{H}_\vartheta(s^{t-1})\frac{Y(s^t)}{A(s_t)L(s^t)}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1/\rho}$$
$$- \left\{(1+\eta) - (1-\gamma)\frac{\rho}{\rho-1}\frac{Y(s^t)}{A(s_t)L(s^t)}\left[\frac{y^f(s^t)}{Y(s^t)}\right]^{1/\rho}\right\}\left(\vartheta\sum_{i\in I}\pi^i\nu^i\sigma_0^i\right)$$

Again using condition (84), we have that that optimal monetary tax satisfies:

$$0 = \mathcal{I}(s_t) + (\eta+\gamma)\mathcal{H}_\vartheta(s^{t-1})\hat{\xi}(s^t) - \mathcal{H}_\vartheta(s^{t-1})(1-\tau_M(s^t))^{-1}$$
$$- \left[(1+\eta) - (1-\gamma)\frac{\rho}{\rho-1}(1-\tau_M(s^t))^{-1}\right]\left(\vartheta\sum_{i\in I}\pi^i\nu^i\sigma_0^i\right)$$

where $\hat{\xi}(s^t)$ is defined according to:

$$\hat{\xi}(s^t) \equiv \frac{\xi(s^t)}{A(s_t)L(s^t)}\mathcal{H}(s^{t-1})^{-1}. \tag{137}$$

We define a function

$$\hat{g}(\mathcal{I}(s_t),\tau_M(s^t)) \equiv \mathcal{I}(s_t) + \mathcal{H}_\vartheta(s^{t-1})\left[(\eta+\gamma)\hat{\xi}(s^t) - (1-\tau_M(s^t))^{-1}\right]$$
$$- (1+\eta)\left(\vartheta\sum_{i\in I}\pi^i\nu^i\sigma_0^i\right) + (1-\gamma)\frac{\rho}{\rho-1}\left(\vartheta\sum_{i\in I}\pi^i\nu^i\sigma_0^i\right)(1-\tau_M(s^t))^{-1}.$$

The optimal monetary tax satisfies: $\hat{g}(\mathcal{I}(s_t),\tau_M^*(s^t)) = 0$. By the implicit function theorem:

$$\frac{d\tau_M^*(s^t)}{d\mathcal{I}(s_t)} = -\frac{d\hat{g}/d\mathcal{I}(s_t)}{d\hat{g}/d\tau_M^*(s^t)}$$
$$= \frac{1}{\mathcal{H}_\vartheta(s^{t-1})\left\{(\eta+\gamma)\frac{d\hat{\xi}(s^t)}{d\tau_M(s^t)} - (1-\tau_M^*(s^t))^{-2}\right\} + (1-\gamma)\frac{\rho}{\rho-1}\vartheta\sum_{i\in I}\pi^i\nu^i\sigma^i(1-\tau_M^*(s^t))^{-2}}$$

Therefore the derivative of the optimal monetary tax satisfies:

$$\frac{d\tau_M^*(s^t)}{d\mathcal{I}(s_t)} = \frac{\mathcal{H}_\vartheta(s^{t-1})^{-1}}{\left[1-(1-\gamma)\frac{\rho}{\rho-1}\left(\vartheta\sum_{i\in I}\pi^i\nu^i\sigma_0^i\right)\mathcal{H}_\vartheta(s^{t-1})^{-1}\right](1-\tau_M^*(s^t))^{-2} - (\eta+\gamma)\frac{d\hat{\xi}(s^t)}{d\tau_M(s^t)}} \tag{138}$$

Our next goal is to obtain an expression for $\hat{\xi}(s^t)$. Following the same steps as in the proof of Theorem 2, we obtain the following expression for $\hat{\xi}(s^t)$:

$$\hat{\xi}(s^t) = \rho\frac{\mathcal{H}_\vartheta(s^{t-1})^{-1}}{\chi(s^{t-1})U_C^m(s^t)}\varsigma^Y(s^t)\left[\frac{(1-\kappa)(1-\epsilon(s^t))}{\kappa\epsilon(s^t)^{-\rho}+(1-\kappa)}\right] \tag{139}$$

Next we use the planner's optimality condition in (110). Following the same steps as in the proof of Theorem 4, we find that $\varsigma^Y(s^t)$ satisfies:

$$\varsigma^Y(s^t) = \hat{\mathcal{W}}_C(s^t)\frac{\kappa\epsilon(s^t)^{1-\rho} + (1-\kappa)}{\kappa\epsilon(s^t)^{1-\rho} + (1-\gamma\rho)(1-\kappa)\epsilon(s^t) + \gamma\rho(1-\kappa)}.$$

Substituting this expression for $\varsigma^Y(s^t)$ into (139), we obtain the following expression for $\hat{\xi}(s^t)$ as a function of $\epsilon(s^t)$:

$$\hat{\xi}(s^t) = \rho\frac{\mathcal{H}_\vartheta(s^{t-1})^{-1}}{\chi(s^{t-1})U_C^m(s^t)}\hat{\mathcal{W}}_C(s^t)\Sigma(\epsilon(s^t))(1-\tau_M(s^t))^{-1}.$$

where $\Sigma(\epsilon)$ is defined by

$$\Sigma(\epsilon) \equiv \frac{(1-\kappa)(1-\epsilon)}{\kappa\epsilon^{1-\rho} + (1-\gamma\rho)(1-\kappa)\epsilon + \gamma\rho(1-\kappa)}$$

as in the proof of Theorem 2. Furthermore, substituting in our expression for $\hat{\mathcal{W}}_C(s^t)$ from (115), we obtain:

$$\hat{\xi}(s^t) = \rho\frac{\mathcal{H}_\vartheta(s^{t-1})^{-1}}{\chi(s^{t-1})}\left[\frac{\mathcal{W}_C(s^t)}{U_C^m(s^t)} - \chi(s^{t-1})(1-\gamma)\frac{\rho}{\rho-1}\left(\vartheta\sum_{i\in I}\pi^i\nu^i\sigma_0^i\right)\right]\Sigma(\epsilon)(1-\tau_M(s^t))^{-1}$$

Using the fact that $\mathcal{W}_C(s^t) = U_C^m(s^t)\Omega_C(\varphi)$, the above expression reduces to:

$$\hat{\xi}(s^t) = \rho\left[1-(1-\gamma)\frac{\rho}{\rho-1}\left(\vartheta\sum_{i\in I}\pi^i\nu^i\sigma_0^i\right)\mathcal{H}_\vartheta(s^{t-1})^{-1}\right]\Sigma(\epsilon)(1-\tau_M(s^t))^{-1}. \tag{140}$$

Derivative of $\tau_M(s^t)$. The derivative of the optimal monetary tax satisfies (138). Evaluating this derivative at the benchmark in which $\tau_M(s^t) = 0$, we have:

$$\left.\frac{d\tau_M^*(s^t)}{d\mathcal{I}(s_t)}\right|_{\tau_M(s^t)=0} = \frac{\mathcal{H}_\vartheta(s^{t-1})^{-1}}{\left[1-(1-\gamma)\frac{\rho}{\rho-1}\left(\vartheta\sum_{i\in I}\pi^i\nu^i\sigma_0^i\right)\mathcal{H}_\vartheta(s^{t-1})^{-1}\right] - (\eta+\gamma)\left.\frac{d\hat{\xi}(s^t)}{d\tau_M(s^t)}\right|_{\tau_M(s^t)=0}} \tag{141}$$

where $\hat{\xi}(s^t)$ satisfies (140). Taking the first derivative of the expression in (140), we get:

$$\frac{d\hat{\xi}}{d\tau_M} = \rho\left[1-(1-\gamma)\frac{\rho}{\rho-1}\left(\vartheta\sum_{i\in I}\pi^i\nu^i\sigma_0^i\right)\mathcal{H}_\vartheta(s^{t-1})^{-1}\right]\left[\Sigma(\epsilon)(1-\tau_M)^{-2} + (1-\tau_M)^{-1}\frac{d\Sigma}{d\epsilon}\frac{d\epsilon}{d\tau_M}\right] \tag{142}$$

Note that the last term in (142) coincides with the last term in (131). Therefore, as in the proof of Theorem 2, we have that the last term in (142) satisfies (132). Evaluating this term at $\tau_M = 0$, or equivalently at $\epsilon = 1$, we have:

$$\left[\Sigma(\epsilon)(1-\tau_M)^{-2} + (1-\tau_M)^{-1}\frac{d\Sigma}{d\epsilon}\frac{d\epsilon}{d\tau_M}\right]_{\tau_M=0} = -\left(\frac{1-\kappa}{\kappa}\right)$$

And furthermore evaluating (142) at $\tau_M = 0$, we get:

$$\left.\frac{d\hat{\xi}}{d\tau_M}\right|_{\tau_M=0} = -\rho\frac{1-\kappa}{\kappa}\left[1-(1-\gamma)\frac{\rho}{\rho-1}\left(\vartheta\sum_{i\in I}\pi^i\nu^i\sigma_0^i\right)\mathcal{H}(s^{t-1})^{-1}\right]$$

Substituting this expression into (141) yields:

$$\left.\frac{d\tau_M^*(s^t)}{d\mathcal{I}(s_t)}\right|_{\tau_M(s^t)=0} = \frac{1}{(1+\rho(\eta+\gamma)\frac{1-\kappa}{\kappa})\left[\mathcal{H}_\vartheta(s^{t-1})+(\gamma-1)\frac{\rho}{\rho-1}\vartheta\sum_{i\in I}\pi^i\nu^i\sigma_0^i\right]}$$

Therefore, to a first-order Taylor approximation around $\mathcal{I}(s_t) = \hat{\mathcal{I}}(s^{t-1})$,

$$\begin{aligned}\log\mathcal{M}^*(s^t) - \log\bar{\mathcal{M}} &\approx \tau_M^*(s^t)\\ &\approx \frac{1}{(1+\rho(\eta+\gamma)\frac{1-\kappa}{\kappa})\left[\mathcal{H}_\vartheta(s^{t-1})+(\gamma-1)\frac{\rho}{\rho-1}\vartheta\sum_{i\in I}\pi^i\nu^i\sigma_0^i\right]}[\mathcal{I}(s_t)-\bar{\mathcal{I}}_\vartheta(s^{t-1})]\end{aligned}$$

as was to be shown.