Learning over the Business Cycle: Policy Implications*

George-Marios Angeletos† Luigi Iovino‡ Jennifer La'O§

September 3, 2020

Abstract

This paper studies the policy implications of the endogeneity of information about the state of the economy. The business cycle can be made less noisy, and more efficient, by incentivizing firms to vary their pricing and production decisions more with their beliefs about the state of the economy. This calls for countercyclical taxes complemented by a monetary policy that "leans against the wind." The optimal policies trade-off allocative efficiency for informational efficiency.

JEL codes: D61, D62, D82, D83, E32, E52

*This paper subsumes an earlier version entitled "Efficiency and Policy with Endogenous Learning," which was prepared for the Workshop on Information, Competition, and Market Frictions at the 2013 Barcelona GSE Summer Forum. We thank our discussant Christian Hellwig, the organizers Alessandro Pavan and Xavier Vives, two anonymous referees, and various seminar participants for helpful comments and suggestions. Angeletos acknowledges the financial support of the National Science Foundation (Award 1757198).

†MIT, 50 Memorial Drive, Cambridge, MA 02142, United States; and NBER. angelet@mit.edu.

‡Bocconi University, Via Roentgen 1, 20136 Milano, Italy; IGIER; and CEPR. luigi.iovino@unibocconi.it.

§Columbia University, 420 W. 118th St., New York, NY 10027, United States; the Federal Reserve Bank of Minneapolis; and NBER. jenlao@columbia.edu.

1 Introduction

Economic agents monitor macroeconomic statistics and market signals such as prices for clues about the state of the economy. But the informativeness of such signals is a function of other agents' behavior. How does this endogeneity of information affect the efficiency of the business cycle and the design of optimal policy?

The contribution of this paper is to address this question within a micro-founded, general-equilibrium, macroeconomic model. Our main lesson is that the optimal policy combines counter-cyclical taxes with a monetary policy that "leans against the wind." A complementary, methodological contribution is to adapt the primal approach of the Ramsey literature to a setting with both incomplete and endogenous information.

Preview. Our model shares the same core micro-foundations as the textbook New Keynesian model, except for three, primary modifications. First, we let the nominal rigidity originate in an informational friction: firms set their prices on the basis of dispersed, noisy information about the state of the economy. Second, we let the informational friction be also a source of real rigidity: firms make certain real production decisions on the basis of the same noisy information. And third, we allow each firm to observe various market signals or macroeconomic statistics whose informational content is endogenous to the choices of other firms.

Although subsets of these three features have appeared in previous work, their *combination* is novel to the literature and essential for our results. We next describe the role played by each of them and explain how they shape the optimal policy.

The first feature, the information-driven *nominal* rigidity, needs no motivation: it is familiar from Woodford (2003a), Mankiw and Reis (2002), Mackowiak and Wiederholt (2009) and a large follow-up literature. But were it not for our setting's two other features, the policy problem would be trivial: the complete-information first best would be implementable with a subsidy that offsets the monopoly distortion and a monetary policy that stabilizes the price level.

The logic is the same as in the textbook New Keynesian model: the optimal subsidy corrects the monopoly distortion and the optimal monetary policy neutralizes the nominal rigidity. Whether the nominal rigidity originates in an informational friction or from Calvo-like sticky prices makes no difference for this logic.[1]

Thus consider our setting's second feature, the information-driven *real* rigidity. This feature, which is borrowed from Angeletos and La'O (2010, 2020), guarantees that the full-information first best is unattainable regardless of the tax and monetary policies: relative to the first best, some misallocation in resources and some dispersion in relative prices is inevitable, and indeed

[1]Like in the textbook New Keynesian model, our model allows for lump-sum taxation. Without it, the requisite subsidy cannot be financed in a non-distortionary way and the first best is not attainable. But as shown in Correia, Nicolini, and Teles (2008), this does not upset the essence of the argument: it only replaces the first best with the kind of flexible-price second best characterized in Lucas and Stokey (1983).

socially desirable, given that firms must base at least some of their production decisions on heterogeneous information about the state of the economy.

This is where our setting's third, and most novel, feature comes into play. The welfare consequences of the real rigidity naturally depend on the precision of the available information. And because this information is endogenous to the choices of others, an informational externality and a new role for policy emerge.

By incentivizing firms to respond more aggressively to variation in their beliefs about the state of the economy, a countercyclical tax on firm revenue or production improves the aggregation of information, thus also reducing the welfare losses due to incomplete information. But because such a tax instrument is too blunt, in a sense we explain below, it has to be complemented by a monetary policy that raises the interest rate above the natural rate during booms (and lowers it during recessions).

Mechanisms at work. To understand the precise role played by monetary policy, it is useful to compare our setting to that of Angeletos and La'O (2020). That paper has shown that the kind of information-driven real rigidity accommodated here redefines the concepts of the "divine coincidence" and the "output gap" that underlie the modern theory of optimal monetary policy. The optimal policy still aims at neutralizing the nominal rigidity, but this no more coincides with minimizing the gap between equilibrium and first-best output, simply because the full-information first best is no longer the right policy benchmark. Instead, the appropriate target level of output is one that displays both less sensitivity, or more inertia, with respect to innovations in underlying fundamentals (total factor productivity) and a positive level of noise- or sentiment-driven fluctuations.

Whereas Angeletos and La'O (2020) treats the information structure as exogenous, here we let it be endogenous—this drives our novel policy conclusions. In their setting, the appropriate gauge for aggregate output is modified for the reason already explained, but the basic policy guidelines remain unchanged: the optimal tax policy serves only the role of correcting the monopoly distortion and the optimal monetary policy serves only the role of neutralizing the nominal rigidity, or replicating the relevant flexible-price outcomes. In our setting, instead, the optimal policies strike a balance between these familiar goals and the novel goal of inducing a better aggregation of information through prices and quantity signals. In other words, the optimal policies trade-off allocative efficiency for informational efficiency.

Our main result is a characterization of this trade-off and its optimal resolution. In particular, we identify two "informational wedges" that serve as sufficient statistics for how the endogeneity of information affects the optimal policy mix. One of them relates to the learning through quantities, the other to the learning through prices. We then show that both wedges enter the determination of optimal taxes, whereas only the second enters the determination of monetary policy. We conclude that in the realistic case in which both kinds of learning are present, the optimal policy mix combines countercyclical taxes with a monetary policy that leans against the wind.

Counter-cyclical taxes serve the goal of improving the informational content of *both* quantity and price signals. By contrast, monetary policy is exclusively connected to the informational content of prices: if all learning were to take place through quantity signals, then optimal monetary policy would only serve the goal of neutralizing the nominal rigidity. Learning from prices is therefore essential for breaking the "divine coincidence" in our setting.

Let us explain the logic. When firms vary their production decisions more aggressively with their private information about the state of the economy, they must also vary their prices more aggressively (and in the direction opposite that of quantities) simply because they face downward-sloping demand curves. This means that when firms respond more aggressively to their own information, *both* the price and the quantity signals become more informative. To induce firms to internalize this information externality, the optimal policy must make firms' expected net returns more sensitive to the state of the economy. Countercyclical taxes serve this goal, whether learning takes place through quantity signals, price signals, or both.

When price signals are absent and the information externality operates entirely through quantity signals, a countercyclical tax on production alone does the job of incentivizing firms to use their information in the socially optimal way. In this case, monetary policy is left with the sole job of neutralizing the nominal rigidity—as desired when information is exogenous. The particular monetary policy that neutralizes nominal rigidities and implements flexible-price allocations is one that targets a negative correlation between the price level and aggregate output; that is, it leans against the wind (Angeletos and La'O, 2020).

When instead learning takes place also through prices, such a tax is not sufficient because it is too blunt. The ideal tax system subsidizes less the production choices that are free to adjust after prices have been set, relative to production choices determined at a similar time as prices; doing so increases the informational content of both quantity and price signals. But note that this would require not only an unrealistic knowledge of which firm choices are set under what information, but also an extremely fine-tuned, differential tax system.

A monetary policy that leans *even more* against the wind relative to the one that implements flexible prices mimics such a differential subsidy. And in contrast to a differential subsidy, this monetary policy implements the socially optimal allocation without requiring the policymaker to know the exact "details" of which production choices and prices are set when and on the basis of what information.

We conclude that a countercyclical tax on production alone improves learning through both quantities and prices, but it does not allow the planner to regulate separately the two kinds of learning. By contrast, combining such a tax with a state-contingent monetary policy facilitates a finer regulation of the two forms of learning.

Related Literature. Methodologically, our paper is at the crossroads of two literatures: the Ramsey literature on optimal taxation and optimal monetary policy (Lucas and Stokey, 1983; Chari, Christiano, and Kehoe, 1994; Correia, Nicolini, and Teles, 2008); and the literature that studies the efficient, decentralized use and aggregation of information (Vives, 1988; Angeletos and Pavan, 2007, 2009).

The latter literature develops and characterizes a notion of constrained efficiency for a certain class of incomplete-information, linear-quadratic games (such as the "beauty contest" games popularized by Morris and Shin, 2002). Relative to those works, our analysis considers different and richer micro-foundations. Furthermore, we are primarily concerned with the characterization of a particular mix of fiscal and monetary policies, as opposed to an abstract notion of constrained efficiency. Nevertheless, the optimal policy problem in our paper can be understood by mapping it to a more abstract constrained-efficiency problem—one that is conceptually the same as those developed in Vives (1988) and Angeletos and Pavan (2007, 2009). We view this translation as an integral part of our contribution.

In so doing, we also offer a concrete example of how the primal approach from the Ramsey literature (Lucas and Stokey, 1983; Chari and Kehoe, 1999) can be adapted to the kind of endogenous-information settings that we are interested in. As in that literature, the analysis becomes much more transparent and straightforward once the policy problem is posed in terms of implementable allocations as opposed to policy instruments. Unlike that literature, however, we must also take into account how different implementable allocations are associated with different information structures, due to the endogeneity of the signals that agents observe about one anothers' choices. As a result, the implementability results that we develop in this paper entail a fixed-point relation between allocations and information structures. Since such a fixed-point relation is endemic to noisy rational expectations equilibrium (REE) settings, the primal approach we take in this paper could be of broader methodological value.[2]

The monetary aspect of our analysis is closely related to the following set of papers that study optimal monetary policy in the presence of informational frictions: Ball, Mankiw, and Reis (2005); Adam (2007); Lorenzoni (2010); Paciello and Wiederholt (2014); Angeletos and La'O (2020). All of these papers abstract from endogenous aggregation of information. With the exception of Angeletos and La'O (2020), they also equate the informational friction with a particular form of nominal friction. By contrast, the lessons we deliver in this paper hinge on the joint property that the informational friction interferes with real allocations even when there is no nominal friction (meaning either that prices are flexible or that monetary policy replicates flexible-price allocations) and that information is endogenously aggregated.

Our focus on learning through prices also brings to mind the voluminous macro and asset-pricing literatures that follow in the traditions of, respectively, Lucas (1972) and Grossman and

[2] Complementary are also Laffont (1985) and Messner and Vives (2001). These papers do not take a Ramsey-like primal approach but share the spirit of studying optimality directly over a particular set of permissible strategies.

Stiglitz (1976, 1980). The following three contributions are worth singling out.

Amador and Weill (2010) show in a macro context how public learning through prices can crowd out valuable private information in the absence of policy. We complement this work by identifying policies that correct the underlying informational externality, thereby also correcting the particular problem emphasized in that paper.

Gaballo (2018) demonstrates how private learning through prices can give rise to multiple equilibria, even when the exogenous noise is small. Such multiplicity is possible in our context as well (at least in principle) and raises the question of whether more "sophisticated" policies may be necessary for unique implementation. This, however, does not affect the essence of our results, as they are derived directly from the solution to the planner's problem (which is, of course, generically unique).

Finally, Vives (2017) studies a class of market games in which firms compete in price-contingent supply schedules and shows how in that context a pecuniary externality may counteract the effects of the learning externality and even push the equilibrium in the opposite direction. Our setting instead features no pecuniary externalities due to complete risk sharing in consumption.[3] This explains both why the informational externality alone pins down the optimal policy and why the direction of the optimal policy intervention is the same regardless of whether the firms' production and pricing choices are strategic complements or substitutes.[4]

Layout. The rest of the paper is organized as follows. Section 2 introduces the baseline model, which abstracts from nominal rigidity and monetary policy. Section 3 studies constrained efficiency, implementability, and optimal policy in the baseline model. These results serve as a stepping stone for the analysis in Section 4. There, we extend the model so that the informational friction becomes the source of both real and nominal rigidity, and we proceed to study the optimal mix of fiscal and monetary policies in Section 5. Section 6 discusses the robustness of these insights to various extensions of the environment and proposes avenues of future research. Section 7 concludes.

2 The Baseline, Non-Monetary Model

We start with a stripped-down version of our framework which abstracts from nominal rigidity and monetary policy. This serves three related goals. First, it permits a cleaner exposition of how the primal approach from the Ramsey literature can be adapted to both dispersed and endogenous information and, by the same token, how the optimal policy problem can be

[3] To be precise, we will assume that all agents belong to the same "big family." As a result, we have complete markets in the sense that all idiosyncratic risk in consumption is insured away. There are, however, missing markets in the sense that there is no "futures market" to perfectly aggregate information at the moment employment and production choices are made. If such a market were to exist, the informational externality would vanish.

[4] For a more abstract analysis of how the interplay of pecuniary and informational externalities may shape policy, see Angeletos and Pavan (2009).

mapped to the more abstract problems studied in Vives (1988) and Angeletos and Pavan (2007, 2009). Second, it sheds light on the role of taxes. And third, it sets the foundations for the analysis of monetary policy in Sections 4-5.

Time and geography. Time is discrete and periods are indexed by $t \in \{0, 1, 2, ...\}$. There is a representative household consisting of a consumer and a continuum of workers. There is a continuum of "islands", indexed by $i \in I = [0, 1]$, which define the boundaries of local labor markets as well as the "geography" of information: information is symmetric within an island, but asymmetric across islands. Each island is inhabited by a representative firm, which specializes in the production of differentiated commodities.

Each period has two stages. In stage 1, the representative household sends a worker to each of the islands. Local labor markets then open, workers decide how much labor to supply, firms decide how much labor to demand, and local wages adjust so as to clear the local labor market. At this point, workers and firms in each island have perfect information regarding local productivity, but imperfect information regarding the productivities of other islands. After employment and production choices are sunk, workers return home and the economy transitions to stage 2. In stage 2, all information that was previously dispersed becomes publicly known, and commodity markets open. Quantities are now pre-determined by the exogenous productivities and the endogenous employment choices made during stage 1, but prices adjust so as to clear product markets.

The representative household. The utility of the representative household is given by

$$\mathcal{U} = \sum_{t=0}^{\infty} \bar{\beta}^t \left[U(C_t) - \int_I V(n_{i,t}) di \right]$$

with $U(C) = \frac{1}{1-\gamma} C^{1-\gamma}$ and $V(n) = \frac{1}{\epsilon} n^{\epsilon}$, where $\gamma \geq 0$ parameterizes the income elasticity of labor supply (also, the reciprocal of the elasticity of intertemporal substitution), $\epsilon \geq 1$ parameterizes the Frisch elasticity of labor supply, $n_{i,t}$ is the labor of the worker who gets located on island i during stage 1 of period t, and C_t is aggregate consumption. The latter is the CES aggregator over all commodities that the household purchases and consumes in stage 2:

$$C_t = \left[\int_I c_{i,t}^{\frac{\rho-1}{\rho}} di \right]^{\frac{\rho}{\rho-1}}$$

where $c_{i,t}$ is the quantity the household consumes in period t of the commodity produced by the representative firm on island i, and $\rho > 1$ is the elasticity of substitution across commodities of different islands.

The representative household receives labor income and profits from all islands in the economy. Its budget constraint is thus given by the following:

$$\int_I p_{i,t} c_{i,t} di + B_{t+1} \leq \int_I \pi_{i,t} di + \int_I w_{i,t} n_{i,t} di + R_t B_t,$$

where $p_{i,t}$ is the nominal price of the commodity produced by the representative firm on island i, $\pi_{i,t}$ is the profit of that firm, $w_{i,t}$ is the nominal wage on island i, and R_t is the nominal gross rate of return on the riskless bond, and B_t is the amount of bonds held in period t.

The objective of the household is simply to maximize expected utility subject to the budget and informational constraints faced by its members. Here, one should think of the worker-members of the household as solving a team problem: they share the same objective (household utility) but have different information sets when making their labor-supply choices. Formally, the household sends off its workers in stage 1 to different islands with instructions on how to supply labor as a function of (i) the information that will be available to them at that stage and (ii) the wage that will prevail in their local labor market. In stage 2, the consumer-member collects all income of its worker-members and decides how much to consume of each commodity and how much to save (or borrow) in the riskless bond.

Firms. The output of the representative firm on island i in period t is given by

$$q_{i,t} = A_{i,t}(n_{i,t})^{\theta}$$

where $A_{i,t}$ is the productivity in island i, $n_{i,t}$ is the firm's employment, and $\theta \in (0, 1]$ is the degree of diminishing returns in production. The firm's realized profit is given by

$$\pi_{i,t} = p_{i,t}q_{i,t} - w_{i,t}n_{i,t}$$

Finally, the objective of the firm is to maximize its expectation of the representative consumer's valuation of its profit, namely, its expectation of $U'(C_t)\pi_{i,t}$.

Aggregates and market clearing. Labor markets operate in stage 1, while product markets operate in stage 2. The wage clears the labor market within each island so that labor supply equals labor demand. For the commodities, market clearing in each product market implies that consumption is equal to output: $c_{i,t} = q_{i,t} \;\; \forall i$. Nominal prices are normalized so that the ideal price index, $P_t \equiv \left[\int p_{it}^{1-\rho} di\right]^{\frac{1}{1-\rho}}$, is fixed at 1. Aggregate output and employment are defined by, respectively,

$$Q_t \equiv \left[\int_I q_{i,t}^{\frac{\rho-1}{\rho}} di\right]^{\frac{\rho}{\rho-1}} \quad \text{and} \quad N_t \equiv \int_I n_{i,t} di.$$

Aggregate and idiosyncratic productivity shocks. We assume that the island-specific productivities $A_{i,t}$ are log-normally distributed in the cross-section of islands:

$$a_{i,t} \equiv \log A_{i,t} = \bar{a}_t + \xi_{i,t}$$

where $\bar{a}_t$ is the aggregate productivity shock and $\xi_{i,t}$ is an idiosyncratic, island-specific, shock. The aggregate shock is drawn from a Normal distribution with mean $\mu_{A,t}$ and variance $\sigma^2_{A,t}$, while the idiosyncratic shock is drawn from a Normal distribution with mean 0 and variance $\sigma^2_{\xi,t}$. The variables $\mu_{A,t}$, $\sigma_{A,t}$ and $\sigma_{\xi,t}$ are common knowledge in period t but need not be deterministic: they could be arbitrary functions of the (public) history of past productivity shocks.[5] For future reference, we let $\kappa_{A,t} \equiv \sigma^{-2}_{A,t}$ and $\kappa_{\xi,t} \equiv \sigma^{-2}_{\xi,t}$.

Information. In stage 1, when key employment and production choices are made, the firms and workers that are located on any given island face uncertainty about what's going on on other islands. More specifically, firms and workers observe the productivity of their own island but not the productivities of other islands. Because local productivities are correlated (through the aggregate productivity shocks), local productivity serves also as a noisy private signal of the distribution of productivities and information of other islands.[6]

In addition to this information, all firms and workers observe exogenous public and private signals about the underlying aggregate productivity. The public signal is given by

$$z^a_t = \bar{a}_t + \varepsilon^{za}_t,$$

and the private signal by

$$x^a_{i,t} = \bar{a}_t + \varepsilon^{xa}_{i,t},$$

where $\varepsilon^{za}_t \sim \mathcal{N}(0, \sigma^2_{za})$ and $\varepsilon^{xa}_{i,t} \sim \mathcal{N}(0, \sigma^2_{xa})$ are noises, the first one common and the second one idiosyncratic across islands. For future reference, we let $\kappa_{xa} \equiv \sigma^{-2}_{xa}$, $\kappa_{za} \equiv \sigma^{-2}_{za}$.

Finally, firms and workers also observe two *endogenous* signals about the production activity that is taking place in other islands, one public and one private. In particular, letting

$$Q_t \equiv \left[\int_I q_{i,t}^{\frac{\rho-1}{\rho}} di \right]^{\frac{\rho-1}{\rho}}$$

measure aggregate output, the endogenous public and private signals are given by, respectively,

$$z^q_t = \log Q_t + \varepsilon^{zq}_t \qquad \text{and} \qquad x^q_{i,t} = \log Q_t + \varepsilon^{xq}_{i,t},$$

where $\varepsilon^{zq}_t \sim \mathcal{N}(0, \sigma^2_{zq})$ and $\varepsilon^{xq}_{i,t} \sim \mathcal{N}(0, \sigma^2_{xq})$ are noises, the first one common and the second one idiosyncratic across islands.

The signal z^q_t is meant to capture macroeconomic data released by various government agencies. For now, this corresponds to a statistic of aggregate GDP. In the extended monetary model (Section 4), we add a signal of the aggregate price level. And although we do not explicitly consider them, signals of aggregate employment or the interest rate introduce the same kind of public learning as that captured by z^q_t here.

[5]For example, the special case in which aggregate productivity follows a random walk can be nested by letting $\mu_t = \bar{a}_{t-1}$ and σ_t be a constant.

[6]The assumption that firms and workers know their own productivities perfectly is inessential; all of our results go through if we allow for uncertainty about local as well as aggregate productivity.

The signal x_t^q, on the other hand, is meant to be a proxy for all kinds of private learning about the state of the economy. For instance, one can think of firms collecting private data about product conditions in particular markets, as in Townsend (1983) and Amador and Weill (2010), or of them extracting information from idiosyncratic market transactions, as in Lucas (1973). The exact modeling of such private sources of information is left outside the analysis, but the essential feature we capture is their dependence on the behavior of others.[7]

3 Efficiency, Implementation, and Optimal Policy

In this section we define and characterize the relevant efficiency benchmark, study the informational externality that underlies it, and show how it can be implemented with taxes.

3.1 Some preliminaries, notation, and the efficiency definition

In addition to the usual resource constraint, the fictitious planner of our economy faces an informational constraint: the employment and production strategies of any given island i must be measurable in the vector $(a_{i,t}, x_{i,t}^a, z_t^a, x_{i,t}^q, z_t^q)$. This measurability constraint encapsulates the informational friction faced by the market mechanism and the planner alike.

Next, note that because there is no capital and all information is revealed at the end of the period, each period is completely separate from one another. For this reason we may drop the time t subscript and the analyze the problem as if it were static. The time dimension will only matter when we study equilibrium implementation and, in particular, the determination of interest rates.

In principle, we could still consider any allocation in which the employment and output of an island are arbitrary functions of the aforementioned signals. For the analysis to remain tractable, however, we must restrict attention to allocations that preserve the Gaussian structure of the information structure. This is true as long as the employment and production choices of an island are log-linear in the private (island-specific) signals.[8]

Indeed, as long as this is the case, we can guess and verify that the information that is available to any island i in stage 1 can be summarized by the triplet (a_i, x_i, z), where a_i is the current local productivity; x_i is a Gaussian sufficient statistic for all the private (local)

[7] The assumed signals can be thought of as special cases of a more general class of signals of the form $\omega_t^q = \log Q_t + \varepsilon_t^{agg} + \varepsilon_t^{idio}$, where ε_t^{agg} is aggregate noise and ε_t^{idio} is idiosyncratic noise. Such signals could be the product of rational inattention over macroeconomic statistics: the measurement error is the source of aggregate noise and rational inattention is the source of idiosyncratic noise. Our results readily extend to such signals, modulo of course that we do not explicitly model the attention choice.

[8] In the absence of endogenous aggregation of information, the planner's optimum and the equilibrium can be characterized under very general assumptions about preferences, technologies, and information structures, as is done in Angeletos and La'O, 2020. Here, we need to assume power forms for preference and technologies, Gaussian shocks and signals, and log-linear policies in order ensure log-linearity in employment and production choices. Log-linearity in employment and production choices, in turn, is needed in order to solve explicitly for the precisions of the endogenous signals, for their dependence on policy, and for their effects on welfare.

information about the aggregate state of the economy; and z is a Gaussian sufficient statistic for all public information.

The precise definitions of these statistics, which are exogenous to the behavior of any single individual or island but are ultimately endogenous to the entire economy, will be provided shortly. For now, it suffices to note that the aforementioned tractability requirement boils down to restricting attention to the space of allocations in which the output of an island is a log-linear function of the triplet $\omega_i \equiv (a_i, x_i, z)$, or

$$q_i = q(\omega_i) = \exp\left\{\varphi_0 + \varphi_a a_i + \varphi_x x_i + \varphi_z z\right\}, \tag{1}$$

for arbitrary scalar coefficients $\varphi = (\varphi_0, \varphi_a, \varphi_x, \varphi_z) \in \mathbb{R}^4$. Let Ω denote the set of all possible ω_i; thus $\omega_i \in \Omega$.

For any such allocation, aggregate output is itself a log-linear function of the aggregate fundamental and the public statistic:

$$Q = Q(\bar{a}, z) = \exp\{\varphi_0' + (\varphi_a + \varphi_x)\,\bar{a} + \varphi_z z\} \tag{2}$$

for some constant φ_0' that differs from φ_0 due to the aggregate effects of dispersion.[9] It follows that the endogenous signals x_i^q and z^q can be transformed into Gaussian signals about the underlying aggregate productivity, thus preserving the Gaussian structure of information. The aforementioned sufficient statistics can then be constructed by taking, in effect, the projection of aggregate productivity onto the relevant signals.

We thus have that $x_i - \bar{a}$ is idiosyncratic Gaussian noise with variance κ_x^{-1}, and similarly $z - \bar{a}$ is aggregate Gaussian noise with variance κ_z^{-1}, where κ_x and κ_z denote the precisions of these two sufficient statistics. The precisions of these statistics can be expressed as the sum of the precisions of all the underlying, component signals. And because the precision of the information contained in the signals x_i^q and z^q about the underlying fundamental hinges on how strongly economic activity responds to it, κ_x and κ_z are endogenous to any allocation chosen by the planner, in the manner described below.

Lemma 1. *Take any log-linear allocation as in (1), with arbitrary $\varphi = (\varphi_0, \varphi_a, \varphi_x, \varphi_z) \in \mathbb{R}^4$. The precisions of the sufficient statistics x and z generated by this strategy are given by*

$$\kappa_x = \sigma_\xi^{-2} + \sigma_{xa}^{-2} + (\varphi_a + \varphi_x)^2 \sigma_{xq}^{-2} > 0 \qquad \textit{and} \qquad \kappa_z = \sigma_{za}^{-2} + (\varphi_a + \varphi_x)^2 \sigma_{zq}^{-2} > 0. \tag{3}$$

To understand this result, note that the endogenous signals x_i^q and z^q about aggregate output can be transformed into simple Gaussian signals about the underlying aggregate productivity because aggregate output is itself a log-linear function of $\bar{a}$, as in (2). From equation (2), note

[9] See the Appendix for the exact characterization of the gap $\varphi_0' - \varphi_0$ as a function of the CES parameter ρ and the variances of the idiosyncratic productivities and idiosyncratic noises. For all intensive purposes, however, one can safely ignore the "detail" that $\varphi_0' \neq \varphi_0$.

that the sum $\varphi_a + \varphi_x$ determines the sensitivity of aggregate output to aggregate productivity. But this implies that for any given exogenous noises, it is precisely this sensitivity $\varphi_a + \varphi_x$ that determines how much information the endogenous signals x_i^q and z^q contain about aggregate productivity: the more sensitive is aggregate output to aggregate productivity, the more informative these signals. It follows that the sum $\varphi_a + \varphi_x$ thus determines the precisions κ_x and κ_z of the private and public sufficient statistics.

Next, it is straightforward to show that once we restrict attention to the class of allocations that satisfy (1), welfare can be expressed as a function of the strategy coefficients $\varphi = (\varphi_0, \varphi_a, \varphi_x, \varphi_z)$ and the information precisions $\kappa = (\kappa_x, \kappa_z) \in \mathbb{R}_+^2$.

Lemma 2. *There exists a function $\mathcal{W} : \mathbb{R}^4 \times \mathbb{R}_+^2 \to \mathbb{R}$ such that, for any allocation that satisfies (1), welfare (ex ante utility) is given by $W = \mathcal{W}(\varphi; \kappa)$.*

With these observations in mind, we define constrained efficiency (within the log-linear class of strategies) as follows.

Definition 1. *A constrained efficient allocation is a pair (φ, κ), consisting of a log-linear production strategy φ and precisions κ, that maximizes welfare subject to condition (3).*

Thus, the notion of constrained efficiency we adopt is similar to that of Angeletos and Pavan (2007, 2009), appropriately adapted to our micro-founded, business-cycle economy. Condition (1) serves only the need for tractability. Condition (3), on the other hand, is central: it ensures that the planner must take into account how different allocations sustain different information structures.[10]

3.2 Optimal Allocations

We now proceed to characterize the constrained efficient and the best implementable allocations. In general, these allocations could differ. We will show that this is not the case in our baseline model (although it is the case in our monetary extension).

Let us start with the constrained-efficient allocation. Recall that we can express welfare as $W(\varphi; \kappa)$ where $\varphi = (\varphi_0, \varphi_a, \varphi_x, \varphi_z)$ and $\kappa = (\kappa_x, \kappa_z)$; a closed-form expression of this function is provided in the Appendix. Next, recall that the precisions induced by any given strategy are characterized by condition (3) in Lemma 1; let $K : \mathbb{R} \to \mathbb{R}_+^2$ denote the function that maps the sum $\bar{\varphi} \equiv \varphi_a + \varphi_x$ into the values of κ_x and κ_z according to (3). We can then express the planner's problem as follows.

[10]At the same time, it is important to note that our notion of constrained efficiency does not endow the planner with any communication channels in addition to those already available to the market: the planner is prohibited from transferring information from one island to another in any way other than through the endogenous signals (x_i^q, z^q).

Planner's problem. *Choose a strategy* $\varphi = (\varphi_0, \varphi_a, \varphi_x, \varphi_z)$ *and precisions* $\kappa = (\kappa_x, \kappa_z)$ *so as to maximize* $\mathcal{W}(\varphi;\kappa)$ *subject to the constraint that* $\kappa = K(\varphi_a + \varphi_x)$.

The solution to this problem is complicated by the fact that this problem is non-concave and that a closed-form solution for the efficient strategy does not exist. Nevertheless, because the precisions depend on the strategy only through the sum $\varphi_a + \varphi_x$, we can bypass these complications by splitting the planner's problem into two steps. We summarize these steps as follows but leave the detailed derivations to the Appendix.

Auxiliary Problem 1. *Choose* $\varphi = (\varphi_0, \varphi_a, \varphi_x, \varphi_z)$ *so as to maximize* $W(\varphi;\kappa)$ *subject to* $\varphi_a + \varphi_x = \bar{\varphi}$ *and let* Δ *be the Lagrange multiplier on this constraint.*

Auxiliary Problem 2. *Choose* $\bar{\varphi}$ *so as to maximize* $W(\bar{\varphi}; K(\bar{\varphi}))$.

The first step lets the planner optimize over the set of strategies subject to an additional constraint, namely that the sum $\varphi_a + \varphi_x$ equals $\bar{\varphi}$ for some arbitrary $\bar{\varphi}$. The second step then lets the planner optimize over $\bar{\varphi}$ and over the precisions that are induced by it. The first-step problem is strictly concave and, in fact, its first-order conditions can be reduced to a simple linear system. The solution to this problem leads to the conditions (5) and (6) below, which express the efficient strategy as a function of Δ, the Lagrange multiplier on the aforementioned auxiliary constraint. The second step then permits us to interpret the wedge Δ as the shadow value of the informational externality, to prove the existence of a constrained efficient allocation, and to complete its characterization.

Let us define

$$\beta \equiv \frac{\frac{\epsilon}{\theta}}{\frac{\epsilon}{\theta} + \frac{1}{\rho} - 1} > 1 \qquad \text{and} \qquad \alpha \equiv \frac{\frac{1}{\rho} - \gamma}{\frac{\epsilon}{\theta} + \frac{1}{\rho} - 1} < 1.$$

As will become evident below, the coefficient β determines the elasticity of local output to variations in local productivity, while the coefficient α determines the elasticity of local output to variations in (expected) aggregate output.

Proposition 1. *(i) A constrained efficient strategy always exists and is given by*

$$\log q(\omega) = \varphi_0^* + \varphi_a^* a + \varphi_x^* x + \varphi_z^* z,$$

where the coefficients $(\varphi_a^*, \varphi_x^*, \varphi_z^*)$ *and the associated precisions* (κ_x^*, κ_z^*) *are the fixed point to the* *following system:*

$$
\begin{aligned}
\varphi_a^* &= \beta & (4)\\
\varphi_x^* &= \left\{\frac{(1-\alpha)\kappa_x^*}{(1-\alpha)\kappa_x^*+\kappa_z^*+\kappa_A}\right\}\frac{\alpha}{1-\alpha}\beta+\Delta & (5)\\
\varphi_z^* &= \left\{\frac{\kappa_z^*}{(1-\alpha)\kappa_x^*+\kappa_z^*+\kappa_A}\right\}\frac{\alpha}{1-\alpha}\beta-\frac{\kappa_z^*}{\kappa_A+\kappa_z^*}\Delta & (6)\\
\kappa_x^* &= \sigma_\xi^{-2}+(\varphi_a^*+\varphi_x^*)^2\sigma_{xq}^{-2}\\
\kappa_z^* &= \sigma_{za}^{-2}+(\varphi_a^*+\varphi_x^*)^2\sigma_{zq}^{-2}
\end{aligned}
$$

for some Δ, which itself is proportional to the sum $\frac{\partial W}{\partial \kappa_z}\frac{\partial \kappa_z}{\partial \bar{\varphi}}+\frac{\partial W}{\partial \kappa_x}\frac{\partial \kappa_x}{\partial \bar{\varphi}}$ and is strictly positive whenever $\alpha \neq 0$.

(ii) An equilibrium in the absence of policy (zero taxes) also exists and satisfies the same conditions as the efficient strategy above, replacing Δ with 0.

Part (i) of Proposition 1 characterizes the efficient strategy. Part (ii) contrasts it to the equilibrium in the absence of policy intervention ("laissez faire"). Together, these properties establish that policy intervention is warranted whenever $\Delta > 0$, which in turn is true whenever $\alpha \neq 0$.

Before elaborating on the economic meaning of this result, let us note that the result is intentionally silent about the constant φ_0. This differs between the laissez-faire equilibrium and the constrained efficient allocation for a familiar reason that is of no interest here: the monopoly distortion. This distortion is orthogonal to the informational friction and can be corrected with an *acyclical*, or constant, subsidy on production. In what follows, we disregard the monopoly distortion and concentrate on the informational friction and more specifically on the scalar Δ.

This scalar is a wedge that summarizes the impact of the informational externality on the efficient production strategy relative to the laissez-faire equilibrium—or, equivalently, relative to the allocation that would have maximized welfare had information been exogenous.

In the absence of taxes, the equilibrium allocation is described in Proposition 1 but without the wedge Δ. Note that the precisions κ_x^* and κ_z^* obtained at the efficient allocation above are higher than those obtained at the equilibrium, precisely because the planner induces a higher $\bar{\varphi}$. Had information been exogenous, Δ would have been zero and the planner would have chosen the same allocation as the equilibrium without policy intervention. This is because in the economy under consideration, barring any informational externalities, there is no misalignment between the privately and socially optimal use of information.

Any deviation from the equilibrium allocation thus involves a loss in terms of *allocative* efficiency: it reduces welfare for given information. But it is only this sacrifice that permits the planner to engineer an increase in the precision of the available information. Furthermore, such an increase is welfare improving precisely because the equilibrium use of information is efficient to start with. If the latter were not true, as for example in the case of Morris and Shin (2002), an improvement in the precision of the available information could map to a deterioration of welfare.

What then justifies the aforementioned sacrifice is precisely that these higher precisions contribute to higher welfare. In short, the planner trades off less *allocative* efficiency (i.e., less welfare for given information) for more *informational* efficiency (i.e., higher welfare via better information).

To be precise, the above argument establishes the direction of a *local* welfare improvement starting from the equilibrium without policy intervention. But such local arguments are not necessarily informative about the position of the *global* maximum when the planner's problem fails to be concave as is the case here because, and only because, of the endogeneity of information.[11]

A concrete example of how the local argument could fail is that the planner can induce a high precision for the endogenous signals by choosing, not only a high enough positive value for the sum $\varphi_a + \varphi_x$, but also a sufficiently *negative* value for it. This is simply because the informativeness of the endogenous signals depends only on the absolute value of the sensitivity of aggregate output to aggregate productivity, not on the sign of this sensitivity. We can nevertheless rule out this possibility with a different, non-local argument: the planner can always achieve the same precision along with higher allocative efficiency by choosing the symmetrically positive value for $\varphi_a + \varphi_x$. This is because the value of $\varphi_a + \varphi_x$ that maximizes allocative efficiency—i.e., the equilibrium one—is positive to start with and the welfare function is symmetric around this point. Moreover any positive value for $\varphi_a + \varphi_x$ that is lower than the equilibrium one is clearly suboptimal—for locally raising this value would have improved *both* allocative and informational efficiency.

The above example also hints at the possible non-convexity of the planner's problem. By the same token, we cannot rule out the possibility that there are multiple values of Δ solving the system of equations given in Proposition 1. When such a possibility occurs, it means that there are multiple solutions to the FOCs of the planner's problem. Only one of them, of course, is the optimum, for the planner's optimum is (generically) unique regardless of whether the planner's problem is concave or not; all our subsequent results pertain to that solution.

Finally, the discussion above presumes that $\Delta > 0$, which is the case as long as $\alpha \neq 0$. When instead α is zero, Δ is also zero, and the need for policy intervention vanishes, for the simple reason that there is no value for knowing the aggregate state of the economy in the knife-edge case in which $\alpha = 0$. We expand on this point shortly.

[11] In the absence of endogenous information, the concavity of the planner's problem follows directly from the concavity of preferences and of the production function.

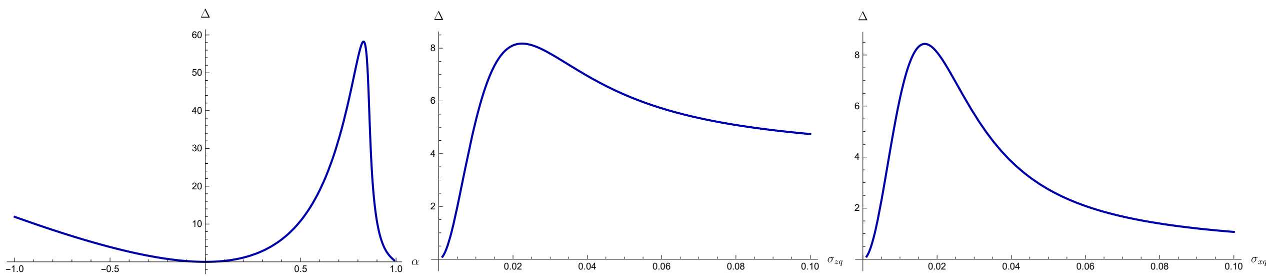

Figure 1. The wedge Δ for different values of α (Panel 1) and for different levels of noise in the endogenous public signal (Panel 2) and the endogenous private signal (Panel 3).

3.3 The Informational Wedge

In Proposition 1, the informational externality, as measured by the sum $\frac{\partial W}{\partial \kappa_z}\frac{\partial \kappa_z}{\partial \bar{\varphi}} + \frac{\partial W}{\partial \kappa_x}\frac{\partial \kappa_x}{\partial \bar{\varphi}}$, is itself evaluated at the constrained efficient strategy. By the same token, the wedge Δ is jointly determined with the coefficients φ^* and the precisions κ^*. The details can be found in the Appendix. This complication prevents a closed-form characterization of Δ.

This complication does not matter for the *qualitative* properties of the efficient strategy discussed above, nor for the qualitative cyclical properties of the optimal policy that will be discussed shortly. For these it suffices to know that Δ is strictly positive as long as $\alpha \neq 0$.

However, in order to study the comparative statics of Δ with respect to the economy's deep parameters, we must resort to numerical simulations. The model is sufficiently parsimonious that we may conduct a rather extensive exploration of its parameter space. In Figure 1, we use a particular parameterization to illustrate a few key findings which are qualitatively robust across a wide range of parameter values. For each of these simulations, we compute and plot the Δ that maximizes welfare. Although the model is too stylized to permit a serious quantitative evaluation, the particular parameterization used in this figure is reasonably realistic: we use conventional values for the underlying preference and technology parameters and empirically plausible values for the measurement errors in the endogenous signals.[12]

Consider first Panel A of Figure 1, which plots Δ as a function of the degree of strategic complementarity, or the GE feedback, as measured by α. When $\alpha = 0$, the GE feedback is muted and there is no informational externality: $\Delta = 0$. This coincides with our results stated in Proposition 1. In this case, firms care only about their own productivity, which they already know; they do not care about the state of the economy, because the positive demand effect of others' production on their revenue is exactly offset by the negative effect on labor costs. As a result, the social value of learning about the state of economy is zero, and there is no informational externality. There is therefore no reason for policy to intervene when $\alpha = 0$.

Away from $\alpha = 0$, the value of learning about the state of the economy is instead strictly

[12]The parameterization used here is the same as that used later in the monetary extension, modulo the exclusion of nominal rigidity and price signals. See Section 5.4 for a detailed description.

positive, and so is Δ. This is true whether firms' actions are strategic complements ($\alpha > 0$) or strategic substitutes ($\alpha < 0$). In the former case, the positive effect of others' production on a firm's demand and revenue outweighs the negative effect on labor costs; firms thereby have an incentive to produce more when other firms produce more. In the latter case, the opposite properties are true. In both cases, however, firms care to know what others do and either do the same or the opposite. This explains why Δ is strictly positive on both sides of $\alpha = 0$.

Interestingly, though, Panel A of Figure 1 shows that even if we restrict attention to the positive domain for α, i.e. the region of strategic complementarity, the wedge Δ is non-monotonic in α. As α increases from 0 to 1, the wedge Δ initially increases but eventually starts falling, converging to zero as α approaches 1.

To understand the downward-sloping region, note that when strategic complementarities are sufficiently strong, firms care a lot about coordinating their behavior. They thereby find it optimal to largely disregard any private information about the state of the economy and instead condition their behavior heavily on the public signal, even if the latter is a rather poor signal about underlying aggregate fundamentals. In the limit as $\alpha \to 1$, the public signal serves in equilibrium as a nearly perfect coordinating device, or signal of the behavior of others. This removes the planner's desire to distort equilibrium allocations in favor of improving learning. As for the endogenous private signals, these are optimally disregarded in favor of public signals in this limit, hence the value of improving their quality vanishes, too. As a result, Δ approaches zero as α approaches 1. This limit thereby has the same policy implication as when $\alpha = 0$, but the rationale is quite different.

Panels B and C of Figure 1 plot Δ for different levels of noise in the public and private endogenous signals, respectively. The relationship is again non-monotonic. When noise in either of these signals is zero, all firms observe the economy's fundamentals perfectly; as a result, there is no need to distort allocations to improve learning and $\Delta = 0$. As the noise in either of these signals moves away from zero, learning becomes imperfect and Δ increases.

To understand the downward-sloping region of Panel B, note that when the noise in the public signal becomes sufficiently high, holding the private noise fixed, firms endogenously disregard the public signal in favor of the private signal. As a result, Δ decreases. However, in the limit as the noise of the public signal tends to infinity, it is *as if* there were no public signal and only a private signal. In this case, firms learn purely from the private signal, but noise in the private signal implies a strictly positive informational externality. Therefore, as public signal noise approaches infinity, Δ falls and converges from above to a strictly positive constant. The same, but reverse, intuition holds true in Panel C as the noise in the private signal tends to infinity, holding the public noise fixed.

To sum up, this example highlights why the informational wedge is likely to be highest when both the degree of strategic complementarity and measurement errors are "moderate." A quantitative translation of this statement, however, is beyond the scope of this paper. That said, we will offer a back-of-the-envelope calculation of the wedge and of its footprint on both taxes and monetary policy within a calibrated version of the monetary extension in Section 4.

Finally, let us emphasize that although α matters for the *magnitude* of Δ, it does not matter for its *sign*: $\Delta > 0$ for both $\alpha > 0$ and $\alpha < 0$. This is because the nature of the informational externality is invariant to whether firm choices are strategic complements or substitutes, and because there is no inefficiency other than the informational externality.[13]

3.4 Implementation and Optimal Policy

Our notion of constrained efficiency allows the planner to choose allocations and associated information structures without any consideration of whether and how these outcomes can be implemented in a market-based equilibrium. We now show how such an allocation can in fact be achieved with countercyclical taxes.

Consider any combination of the following tax instruments: a linear tax $\tau^R(\bar{a}, z)$ on firm revenue or sales, a linear tax $\tau^L(\bar{a}, z)$ on household labor income, and a linear tax $\tau^C(\bar{a}, z)$ on household consumption (a sales tax that is uniform across commodities). These taxes are collected in stage 2 and can be contingent on the information that is publicly available at that time. To maintain tractability and guarantee that equilibrium allocations are log-Normal, these taxes are assumed to be log-linear functions of $(\bar{a}, z)$. Finally, we assume that in order for the government to balance its budget, the government has access to additional lump-sum taxes or transfers to the household.

With $\omega \in \Omega$ once again summarizing the information available to an island in stage 1 (inclusive of its productivity), we define an equilibrium as follows.

Definition 2. *A (log-linear) equilibrium is the combination of signal precisions* $(\kappa_x, \kappa_z) \in {\mathbb{R}_+}^2$ *and a production strategy* $q : \Omega \to \mathbb{R}_+$ *as in (1), along with an employment strategy* $n : \Omega \to \mathbb{R}_+$*, tax rate functions* $\tau^R, \tau^L, \tau^C : \mathbb{R}^2 \to \mathbb{R}$*, a wage function* $w : \Omega \to \mathbb{R}_+$*, a price function* $p : \Omega \times \mathbb{R}^2 \to \mathbb{R}$*, and household consumption demand and labor supply functions, such that:*

(i) Given the signal precisions, the remaining elements constitute a competitive equilibrium in the sense that the production and employment choices are optimal for the firms and the households, the wages and the prices clear the labor and goods markets, and the government's budget constraint is satisfied in all states.

(ii) Given the production strategy, the signal precisions are generated according to (3).

Next, we show that any such tax combination reduces to a single tax wedge between the marginal return and the marginal cost of labor. The key implementability constraint is then identified in the following lemma.

[13] Had there been pecuniary externalities or other sources of inefficiency, as for example in Angeletos and Pavan (2007, 2009) and Vives (2017), we would expect the wedge Δ between the equilibrium and the optimal allocations to combine the informational externality with these other sources of inefficiency. To the extent that the sign of the latter depends on the sign of the firms' strategic interaction, the sign of Δ could also depend on it.

Lemma 3. *A production strategy is implementable with the aforementioned tax instruments if and only it solves the following fixed-point problem:*

$$q(\omega)^{\frac{\epsilon}{\theta}+\frac{1}{\rho}-1} = \left(\frac{\rho-1}{\rho}\right)\theta A(\omega)^{\frac{\epsilon}{\theta}}\mathbb{E}\left[(1-\tau(\bar{a},z))Q(\bar{a},z)^{\frac{1}{\rho}-\gamma}|\omega\right], \forall\omega, \tag{7}$$

where $q(\omega)$ is the production of an island of type ω, $A(\omega)$ is its productivity, $Q(\bar{a},z)$ is aggregate output, and $1-\tau(\bar{a},z)$ is the combined tax wedge induced by the aforementioned tax instruments.

The proof of this result is similar to the characterization of equilibrium in Angeletos and La'O (2010) but with the inclusion of a tax wedge. Basically, condition (7) means that the marginal cost of production in each island is equated to the local expectation of the marginal revenue product, net of taxes.

Let us express the tax wedge as follows

$$-\log(1-\tau(\bar{a},z)) = \tau_0 + \tau_A\bar{a} + \tau_z z, \tag{8}$$

for some known coefficients $(\tau_0, \tau_A, \tau_z) \in \mathbb{R}^3$. The coefficient τ_0 parameterizes the mean value of the tax and helps correct the monopoly distortion—a familiar function that, as already mentioned, is of no interest to us. Central for our purposes are instead the coefficients τ_A and τ_z. These coefficients determine the elasticities of the tax with respect to, respectively, the underlying fundamental and the public statistic; as we explain next, their ultimate function is to regulate how information is used and aggregated in equilibrium.[14]

To see how such contingent taxes can impact the decentralized use and aggregation of information, suppose the tax is negatively correlated with innovations in aggregate productivity and positively correlated with the common noise. Anticipating these correlations, firms choosing output according to equation (7) will have an incentive during stage 1 to react more strongly to any information they may have about aggregate productivity and less strongly to any information they may have about the common noise. It follows that firms will unambiguously increase their response to their private sources of information; whether they will at the same time reduce their response to common information then depends on whether the positive correlation of the tax with the underlying common noise is sufficiently strong relative to its negative correlation with respect to aggregate productivity. This explains why state-contingent taxes can separately regulate both φ_x and φ_z, the sensitivities to private and public information.[15]

[14] Here we have chosen to express the taxes as a function of $\bar{a}$ and z. But since $\log Q$ and $\log N$ are both linear (and indeed non-colinear) combinations of $\bar{a}$ and z, we could equivalently condition taxes on any pair among the set $\{\log Q(\bar{a},z), \log N(\bar{a},z), \bar{a}, z\}$.

[15] To be precise, the argument made above only explains how taxes impact individual firm incentives holding the behavior of other firms constant; that is, it explains the impact of the policy on best responses, but not on equilibrium behavior. However, the equilibrium could fail to inherit the comparative-static properties of best responses only when the degree of complementarity is too strong (namely $\alpha > 1$), which is never the case here due to the assumed micro-foundations.

We can then deduce the relevant policy implications by characterizing the tax wedge that implements the efficient allocation. Let ε denote the noise in the public signal z.

Proposition 2. *There exists a state-contingent tax policy as in (8) that implements the efficient allocation. The optimal tax is countercyclical in either of the following senses:* $Corr(\tau, \bar{a}) < 0$, $Corr(\tau, \bar{a}|z) < 0$, *and* $Corr(\tau, Q) < 0$. *Moreover, the tax is positively correlated with the noise:* $Corr(\tau, \varepsilon) > 0$.

The countercyclicality of optimal taxes follows directly from comparing equilibrium and efficient allocations. Recall that, when information is endogenous, efficiency dictates the government to raise φ_x, the sensitivity of production to private information, so as to boost social learning. At the same time, it also dictates to lower φ_z, the sensitivity to public information, so as to preserve allocative efficiency. How can the tax system provide the agents with the right incentives for these goals to materialize in equilibrium? For the agents to find it optimal to raise their response to their private information about aggregate productivity, it better be that they expect the tax to fall—and hence their net-of-tax returns to increase—with any positive innovations in aggregate productivity. And for them to find it optimal to decrease their response to public information, it better be that they expect the tax to increase with the public signal or, equivalently, with the noise in it. This explains why the optimal tax must be negatively correlated with $\bar{a}$ and positively correlated with ε along the equilibrium.

Note that it is the combination of the two cyclical properties of the tax—its negative correlation with $\bar{a}$ and its positive correlation with ε—that achieves full efficiency. However, it is only the negative correlation with $\bar{a}$ that is the key instrument for increasing φ_x and thereby for boosting the aggregation of information over the business cycle. The positive correlation with the noise is instrumental only for reducing φ_z, which is necessary for preserving some allocative efficiency, but is irrelevant for the efficiency of learning.

Finally, let us clarify the following subtlety. The taxes identified above guarantee the *existence* of an equilibrium that implements the planner's optimum but does not rule out the existence of other equilibria that are worse in terms of information aggregation and welfare. Whether these taxes are present or not, multiplicity of equilibria is possible solely because of the endogeneity of information, and for reasons similar to those articulated in Amador and Weill (2010) and Gaballo (2018).[16] If multiplicity happens to occur with the particular linear taxes we have described, and if the planner wants to guarantee uniqueness of the equilibrium that implements the optimum, then the planner may have to resort to more sophisticated, non-linear taxes. But since our policy lessons are derived directly from the optimal allocation itself, they directly extend to such

[16] In particular, this possibility appears to hinge on whether learning is private or public. When learning is private, there is strategic complementarity in the use of private information: the more other agents rely on private information, the more precise the information contained in endogenous private signals, and the higher the individual willingness to rely on them. This opens the door to multiple equilibria. By contrast, when learning is public, there is strategic substitutability, contributing to uniqueness.

policies: *any* taxes that implement the optimum have to be countercyclical.

4 The Monetary Model

In this section we consider the full version of our model which allows the informational friction to be the source of both real and nominal rigidity. This model sheds light on the joint determination of optimal fiscal and monetary policies.

Set up. We modify the baseline model in three dimensions. First, we introduce price rigidities. In particular, we assume that firms set nominal prices in stage 1 while information is still dispersed, and cannot adjust them in stage 2 in response to the new information that becomes available at that stage. We refer to this scenario as "sticky prices" and to the alternative in which prices are free to adjust in stage 2 as "flexible prices."[17]

Second, we allow firms to make an additional labor-demand choice in stage 2 and, accordingly, we let households make a second labor-supply choice in that stage. In particular, we assume the production of the typical commodity be given by

$$y(\omega, \bar{a}, z) = A(\omega) n(\omega)^{\theta} l(\omega, \bar{a}, z)^{1-\theta},$$

where $y(\omega, \bar{a}, z)$ denotes output, $n(\omega)$ denotes the labor input in stage 1, $l(\omega, \bar{a}, z)$ denotes the labor input in stage 2, and $\theta \in (0, 1)$; accordingly, we assume the per-period utility of the representative household be given by

$$U(C(\bar{a}, z)) - \int \frac{1}{\epsilon} n(\omega)^{\epsilon} dF(\omega|\bar{a}, z) - \int \frac{1}{\epsilon} l(\omega, \bar{a}, z)^{\epsilon} dF(\omega|\bar{a}, z),$$

where $C(\bar{a}, z)$ is the same CES aggregator as that in the baseline model, $F(\cdot|\bar{a}, z)$ is the cdf of ω conditional on the aggregate state $(\bar{a}, z)$, and $\epsilon > 1$.

Third, we let firms and workers in each island observe signals of the (nominal) prices set by firms in other islands in addition to signals of the (real) quantities. In particular, we denote by $Y(\bar{a}, z)$ and $P(\bar{a}, z)$ the real aggregate output and the nominal price level[18] and let each firm

[17] This terminology is borrowed from Angeletos and La'O (2020) but is non-standard. What that work and our paper alike call "sticky prices" is often referred to in the related literature as "flexible prices with information constraints" in order to emphasize that that there is no ad hoc nominal rigidity of the Calvo type. Nevertheless, the terminology proposed in Angeletos and La'O (2020) and adopted here is most appropriate for understanding the mapping between models in which the nominal rigidity originates from an informational friction, as the model used here and in related works (e.g., Woodford, 2003b), and the textbook New Keynesian framework, in which the nominal rigidity takes the Calvo-like form. The adopted terminology is also consistent with that in Correia, Nicolini, and Teles (2008) and Correia et al. (2013); in those works, "sticky firms" refer to a group of firms whose prices cannot be measurable in the concurrent aggregate productivity. Clearly, the essence is the same whether this restriction is interpreted as "pre-determined prices" or as "informationally-constrained prices."

[18] $Y(\bar{a}, z)$ is defined by the same CES aggregator as $C(\bar{a}, z)$, replacing $c(\omega, \bar{a}, z)$ with $y(\omega, \bar{a}, z)$, and $P(\bar{a}, z)$ is defined as the corresponding ideal price index. Clearly, $Y(\bar{a}, z) = C(\bar{a}, z)$ by market clearing in the goods market (or, equivalently, by resource feasibility).

observe a total of four endogenous signals about these objects. Two of these signals are public and are given by

$$z^y = \log Y(\bar{a}, z) + \varepsilon^{zy} \quad \text{and} \quad z^p = \log P(\bar{a}, z) + \varepsilon^{zp},$$

where $\varepsilon^{zy} \sim \mathcal{N}(0, \sigma_{zy}^2)$ and $\varepsilon^{zp} \sim \mathcal{N}(0, \sigma_{zp}^2)$ are their respective noises. The remaining two are private and are given by

$$x_i^y = \log Y(\bar{a}, z) + \varepsilon_i^{xy} \quad \text{and} \quad x_i^p = \log P(\bar{a}, z) + \varepsilon_i^{xp},$$

where $\varepsilon_i^{xy} \sim \mathcal{N}(0, \sigma_{xy}^2)$ and $\varepsilon_i^{xp} \sim \mathcal{N}(0, \sigma_{xp}^2)$ are their respective noises. We interpret the former two signals as readily accessible and commonly known macroeconomic statistics and the latter two as proxies for private learning about economic activity.

Remark. Relative to the baseline model, the only new terms in both technology and preferences are those pertaining to $l(\omega, \bar{a}, z)$, the second-stage labor input. The inclusion of this input, and the fact that it can adjust freely to the realized state of nature, ensures that markets clear: firms can adjust their supply to meet demand at the pre-determined (or information-constrained) prices. This type of assumption is standard in the New Keynesian literature, including the strand that emphasizes informational frictions.

Consider, for instance, the models studied in Mankiw and Reis (2002), Woodford (2003a) and Mackowiak and Wiederholt (2009). The *unique* type of labor featured in these models maps exactly to the second-stage labor input in our framework. What is truly new in our framework is therefore the first-stage labor input, or the inclusion of a production choice that is subject to the same informational friction as the firms' pricing choice. This circles back to our earlier discussion of how the "real" rigidity (the property that at least some production decisions are made on the basis of incomplete and endogenous information) is essential for our results.

At the same time, let us emphasize that the inclusion of the second-stage labor input per se does not drive our results. As will become clear from Proposition 4 below, if we shut down the learning from prices, the following two properties hold regardless of θ, the relative importance of the second-stage input. First, the rationale for a monetary policy that departs from replicating flexible price allocations disappears; and second, the optimal taxes are still driven by the same logic (and indeed the same math) as in our baseline model.

Substance and solution strategy. As in the baseline model, the essential issue here is the endogeneity of the information contained in these signals. If, instead, information had been exogenous, our framework would have been nested in that of Angeletos and La'O (2020). Their results would have guaranteed that the following key lessons from the New Keynesian framework (Correia, Nicolini, and Teles, 2008) extend to the presence of informational frictions. First, that optimal taxes are acyclical; and, second, that optimal monetary policy merely neutralizes the nominal rigidity, or replicates flexible prices. From this perspective, our key contribution here is to show how the endogeneity of information upends these lessons.

There is a subtle technical difference from our baseline model due to the introduction of price signals. In the baseline model, *all* endogenous signals are well defined for arbitrary allocations and regardless of whether such allocations were chosen directly by the planner or implemented via decentralized markets. The same is true here for the quantity signals, but not for the price signals: the latter are well defined *only* in the context of market-based implementations.

This precludes the solution strategy taken in the baseline model. There, we could define and characterize a natural efficiency benchmark without explicit consideration of implementability. Here, we must first characterize the set of the combinations of quantities, prices, and information structures that can be implemented as a market outcome given the available policy instruments; and only then can we proceed to identify the best allocation within this set.

Taxes and monetary policy. The above discussion brings to the forefront the question of what the available policy instruments are. We maintain the taxes from the baseline model but add monetary policy as a new instrument.

The following property extends from our baseline model to the present setting: a tax on firm revenue is equivalent to a tax on firm output, total employment, or payroll; a tax on household labor income or consumption; or any other tax that has a uniform impact across the two labor inputs. With this in mind, we think of the tax on firm revenue as a proxy for all of the above and specify it as as follows:

$$-\log(1-\tau(\bar{a},z)) = \tau_0 + \tau_A\bar{a} + \tau_z z, \tag{9}$$

for some scalars (τ_0, τ_A, τ_z) and for some random variable z that is a sufficient statistic for all available public information (similarly defined as in the baseline model). We describe monetary policy with the following policy rule for the nominal interest rate:

$$\log\left(1 + R(\bar{a},z)\right) = \rho_0 + \rho_A\bar{a} + \rho_z z, \tag{10}$$

for some scalars (ρ_0, ρ_A, ρ_z). Condition (9) is the same as (8) in the baseline model; condition (10) is the analogue for monetary policy. We impose a log-linear specification on both policies in order to preserve the Gaussian nature of the information structure.

In the following section, we first characterize the combinations of allocations, prices, and information structures that can be implemented with arbitrary policies of the aforementioned kind. We then identify the best allocation within this set. Finally, we characterize the policies that support it as an equilibrium. Before jumping into the formal arguments, it is useful to anticipate the following two properties, which help clarify the role played by monetary policy.

First, note that the coefficients ρ_z and ρ_A parameterize the response of monetary policy to the public signal and the innovations in aggregate output, respectively. Since firms set prices in stage 1, when they know z but not $\bar{a}$, monetary policy can have real effects insofar as it responds to the variation in $\bar{a}$ that is not spanned by z. By the same token, ρ_z has no impact on the real allocations and the information structure; only ρ_A matters. This conclusion contrasts with taxation, in which case both τ_z and τ_A have real effects.

Second, if prices had been flexible (i.e., free to adjust in stage 2), taxes would still have real effects, but monetary policy would not. Therefore, by letting prices be sticky (i.e., determined in stage 1), we effectively *expand* the set of allocations and information structures that can be attained by the planner. In particular, an important intermediate step of the subsequent analysis will be to show that, for any given tax policy, there always exists a monetary policy that replicates the corresponding flexible-price outcomes, as well as monetary policies that induce different allocations and information structures. A key final result will be that the former kind of monetary policy is optimal when there are only quantity signals, whereas the latter kind becomes optimal once there are price signals.

5 Implementability and Optimality

This section contains our main results. We first characterize the allocations and prices that can be supported by the available policy instruments. This gives us the "primal" representation of the policy problem. We use this to identify the optimal allocation and the optimal information structure. We proceed to recover the policies that support this allocation and information structure in an equilibrium.

5.1 Implementability in the Monetary Model

As in the baseline model, we guess (and subsequently verify) that the relevant information available to any island in stage 1 can be summarized by the triplet $\omega = (a, x, z)$ with $\omega \in \Omega$, where a is local productivity; x is a Gaussian sufficient statistic for all private information regarding the aggregate state; and z is a Gaussian sufficient statistic summarizing all public information regarding the aggregate state. We similarly guess (and subsequently verify) that the employment and production levels of any island ω can be expressed as follows:

$$\log q(\omega) = \varphi_0 + \varphi_a a + \varphi_x x + \varphi_z z \qquad \text{and} \qquad \log l(\omega, \bar{a}, z) = l_0 + l_A \bar{a} + l_a a + l_x x + l_z z, \tag{11}$$

where $q(\omega)$ is now defined as $q(\omega) \equiv A(\omega) n(\omega)^\theta$, a composite of productivity and the first-stage labor choice, and where $(\varphi_0, \varphi_a, \varphi_x, \varphi_z) \in \mathbb{R}^4$ and $(l_0, l_A, l_a, l_x, l_z) \in \mathbb{R}^5$ are coefficients that are indirectly under the control of the planner.

Three clarifications are needed here. First, although $q(\omega)$ plays a similar technical role as in the baseline model, it now has a more subtle interpretation: instead of being the entire output of a firm or island, it is the component of it that is determined in stage 1, that is, it excludes the input that is free to adjust to monetary policy. Second, although the planner has control over the nine coefficients $(\varphi_0, \varphi_a, \varphi_x, \varphi_z; l_0, l_A, l_a, l_x, l_z) \in \mathbb{R}^9$ that parameterize the state dependence of the real allocations, this control is limited by certain implementability restrictions. These restrictions will be derived shortly.

The third clarification is the following. Since firms set prices in stage 1 along with first-stage labor (equivalently q), it would seem most natural to specify firms' behavior in terms of a pair of strategies for q and p. However, one can always recast a firm's pricing-setting choice as a choice of a state-contingent plan for how its flexible stage-2 input, l, will adjust to realized demand. For this reason we may specify firms' behavior as a pair of strategies for q and l. This recasting, which shifts focus away from price-setting choices towards implementable allocations, works best for our purposes and is a defining feature of the primal approach to optimal policy.[19]

In fact, this equivalence can be inferred from our following equilibrium definition.

Definition 3. *A (log-linear) equilibrium is a combination of signal precisions* $(\kappa_x, \kappa_z) \in \mathbb{R}^2$*, a "sticky" price function* $p : \Omega \to \mathbb{R}$*, production & employment strategies* $q : \Omega \to \mathbb{R}_+$*,* $n : \Omega \to \mathbb{R}_+$*,* $l : \Omega \times \mathbb{R}^2 \to \mathbb{R}$ *as in (11), policies* $\tau : \mathbb{R}^2 \to \mathbb{R}$ *and* $R : \mathbb{R}^2 \to \mathbb{R}$ *as in (9) and (10), a wage function* $w : \Omega \to \mathbb{R}_+$*, and household consumption demand and labor supply functions, such that:*

(i) Given the signal precisions and the tax and monetary policies, employment and prices are optimal for the firms, quantities are optimal for the household, the labor and goods markets clear, the government's budget constraint is satisfied, and the nominal interest rate satisfies

$$U'\left(C(\bar{a}, z)\right) = \bar{\beta}\mathbb{E}\left[\left(1 + R(\bar{a}, z)\right) \frac{P(\bar{a}, z)}{P(\bar{a}_{+1}, z_{+1})} U'\left(C(\bar{a}_{+1}, z_{+1})\right) \middle| \bar{a}, z\right] \tag{12}$$

where $(\bar{a}_{+1}, z_{+1})$ *corresponds to the aggregate state in the following period.*

(ii) Given the production and pricing strategies, the signal precisions are generated accordingly.

Our equilibrium definition is analogous to our earlier equilibrium definition for the baseline flexible-price model (Definition 2); the only modifications here are the following. First, because prices are now sticky, i.e. measurable in ω, second-stage labor must adjust in order for goods markets to clear. Second, the nominal interest rate is such that it satisfies the intertemporal Euler equation for the household (2), written here in recursive form. Third, we again impose that the precisions of the private and public signal are generated endogenously by the allocation; however, we have yet to show how they are generated. We turn to this consideration next.

Implementability results. The rest of this subsection is organized in three results. The first result (Lemma 4) characterizes the information structures that are induced by the aforementioned strategies under arbitrary coefficients $(\varphi_0, \varphi_a, \varphi_x, \varphi_z; l_0, l_A, l_a, l_x, l_z) \in \mathbb{R}^9$. The second result (Proposition 3) works out the implementability restrictions on these coefficients; together with the first result, this amounts to finding the set of the implementable combinations

[19] For examples of the primal approach applied to models with nominal rigidity, see Correia, Nicolini, and Teles (2008) and Correia, Farhi, Nicolini, and Teles (2013), as well as the related work of Angeletos and La'O (2020).

of allocations and information structures. The last result (Lemma 5) characterizes the associated prices.

We start with the first result, which is the analogue of Lemma 1 in the baseline model.

Lemma 4. *Take any pair of strategies as in (11). Let κ_x and κ_z denote the precisions of the sufficient statistics of, respectively, the private and the public information that obtain in equilibrium when all firms follow the aforementioned strategies. Then,*

$$\kappa_x = \sigma_{xa}^{-2} + \Phi^2\sigma_{xy}^{-2} + \Psi^2\sigma_{xp}^{-2} \qquad \textit{and} \qquad \kappa_z = \sigma_{za}^{-2} + \Phi^2\sigma_{zy}^{-2} + \Psi^2\sigma_{zp}^{-2} \tag{13}$$

where

$$\Phi = \varphi_a + \varphi_x + (1-\theta)(l_a + l_x + l_A) \quad \textit{and} \quad \Psi = \frac{1}{\rho}\left(\varphi_a + \varphi_x + (1-\theta)(l_a + l_x)\right) \tag{14}$$

measure the elasticities of, respectively, the aggregate level of output and the aggregate price level with respect to the underlying fundamental, conditional on the public information.

The interpretation of (13) is straightforward: the terms $\kappa_{xy} \equiv \Phi^2\sigma_{xy}^{-2}$ and $\kappa_{xp} \equiv \Psi^2\sigma_{xp}^{-2}$ that show up in κ_x capture the precision of the *private* learning that obtains through the quantity and price signals, respectively; the corresponding terms in κ_z capture the corresponding *public* learning. Only (14) deserves some explanation. To this goal, note that (11) implies that aggregate output can be expressed, up to a constant, as

$$\log Y(\bar{a}, z) = (\varphi_a + \varphi_x + (1-\theta)(l_a + l_x + l_A))\,\bar{a} + (\varphi_z + (1-\theta)l_z)\,z.$$

Because z and, hence, also the second term above is known, observing the available quantity signal is akin to observing the first term above plus measurement error. This explains the formula for and interpretation of Φ. The logic for Ψ is similar, except that it requires the characterization of the equilibrium price level. This missing piece follows from the last result of this subsection (Lemma 5) and from the more detailed analysis in the Appendix.

Let us move on to the second result, which identifies the relevant implementability restrictions. The exact definitions of α and β for the monetary model are in the proof of Proposition 3 in the Appendix. We also assume that the parameters are such that $\alpha > 0$. Although this assumption is required only for the characterization of optimal fiscal and monetary policy in Section 5.3, this is the most realistic case; we thus assume it from the outset.

Proposition 3. *(i) A pair of strategies as in (11) can be implemented as an equilibrium with an appropriate combination of a linear tax and monetary policy as in (9) and (10) if and only if the* *following conditions are satisfied:*[20]

$$\begin{aligned} \varphi_a &= \beta & (15) \\ l_a &= \frac{1}{\theta}(\varphi_a - 1) & (16) \\ l_x + l_A \frac{\kappa_x}{\kappa_A + \kappa_x + \kappa_z} &= \frac{1}{\theta}\varphi_x & (17) \\ l_z + l_A \frac{\kappa_z}{\kappa_A + \kappa_x + \kappa_z} &= \frac{1}{\theta}\varphi_z & (18) \end{aligned}$$

(ii) Had prices been flexible (i.e., free to adjust to realized demand), a pair of strategies as in (11) would have been implemented if and only if the following condition was satisfied in addition to conditions (15)-(18):

$$l_A = \frac{1}{\theta}\frac{\varphi_x}{\beta}\frac{\kappa_A + \kappa_x + \kappa_z}{\kappa_x}. \tag{19}$$

Proposition 3 plays a role similar to the familiar "implementability" results in the Ramsey literature: it represents the optimal policy problem in terms of the allocations that are induced by the policy rather than the policy instruments themselves. In our context, Proposition 3 sheds light on how the planner can regulate the decentralized use of information and thereby also its aggregation through the available price and quantity signals.

As in the baseline model, the planner can induce any φ_x and φ_z she may desire by appropriately choosing the contingencies of the taxes. However, conditional on picking these coefficients, her control over the remaining coefficients is limited either entirely (under flexible prices) or partially (under sticky prices). In particular, if prices were flexible, the planner would have no further control: all remaining coefficients would be fixed functions of the chosen pair (φ_x, φ_z) and deep parameters. Instead, since prices are sticky, the planner has an extra degree of freedom: by appropriately choosing monetary policy, it can induce any l_A it wishes and thereby also influence the pair (l_x, l_z), over and above its control of the pair (φ_x, φ_z).

More specifically, whether prices are flexible or sticky, the absence of a differential tax on the two types of labor implies that the equilibrium necessarily equates the (expected) marginal rate of transformation between these two types of labor with the corresponding marginal rate of substitution in preferences. This explains how and why (l_a, l_x, l_z) are related to $(\varphi_a, \varphi_x, \varphi_z)$ in (15)-(18).[21] Furthermore, if prices were flexible, once taxes had been set to achieve the desired φ_x and φ_z, the sensitivity of second-stage employment and production to the realized aggregate productivity would be pinned down by equating the realized marginal returns and costs of stage-2 labor. It is this restriction that gives (19). But since prices are sticky, this restriction is no more present: by designing the extent to which monetary policy accommodates the realized

[20]There is also a restriction between φ_0 and l_0 which we omit because it is of no interest: φ_0 and l_0 are irrelevant for both the endogenous precisions and business-cycle properties of the real allocations.

[21]In particular, the equality of expected marginal rates of transformation and substitution gives $\mathbb{E}[\log l(\omega, \bar{a}, z)] = const + \log n(\omega)$, where $const$ includes second-order terms; using then $\log q(\omega) = \log A(\omega) + \theta \log n(\omega)$ and noting that this condition must be satisfied for every $\omega = (a, x, z)$, gives the constraints in part (i) of Proposition 3.

productivity shock, the planner can effectively choose any l_A she wishes, in addition to the free choice of the pair (φ_x, φ_z). In other words, sticky prices amount to an extra degree of freedom.

Finally, note that as the planner chooses a higher l_A, which amounts to a more accommodative monetary policy, firms respond optimally by raising the sensitivity of their own prices to the information they have about aggregate productivity when they set their prices, thus, partly offsetting the monetary policy. This explains why l_x and l_z are negatively related to l_A in the way defined by conditions (17) and (18), and why the planner has no control over l_x and l_z other than that afforded through the free choice of l_A. In short, when prices are sticky, monetary policy affords exactly one extra degree of freedom over choosing allocations.

So far we have shown that the set of implementable allocations and information structures is characterized by the combination of condition (11), Lemma 4, and Proposition 3. We conclude this subsection with the characterization of the prices that are associated with any element of this set.

By consumer optimality and market clearing in the goods markets, we have

$$\frac{p(\omega)}{P(\bar{a}, z)} = \left(\frac{c(\omega, \bar{a}, z)}{C(\bar{a}, z)}\right)^{-\frac{1}{\rho}} \quad \text{and} \quad c(\omega, \bar{a}, z) = y(\omega, \bar{a}, z) = q(\omega) l(\omega, \bar{a}, z)^{1-\theta}.$$

Combining these equilibrium conditions with (11), we reach the following result.

Lemma 5. *Pick any allocation and information structure that satisfy the combination of condition (11), Lemma 4, and Proposition 3. The associated equilibrium prices satisfy*

$$\log p(\omega) = \psi_0 + \psi_a a + \psi_x x + \psi_z z,$$

where ψ_a and ψ_x are pinned down by

$$\psi_a = -\frac{1}{\rho}(\varphi_a + (1-\theta) l_a) \qquad \textit{and} \qquad \psi_x = -\frac{1}{\rho}(\varphi_x + (1-\theta) l_x),$$

while ψ_z is indeterminate.

Because any component of monetary policy that is public information at the moment prices are set cannot have any real effect, the dependence of monetary policy and prices to z is indeterminate. By contrast, the dependence of a firm's price on its own productivity and on its private signal are uniquely determined. To understand why, note first that, once the planner has picked a real allocation, there is a unique collection of relative prices that support it in equilibrium. Note next that the relative price of two firms, A and B, can move in a particular manner with the private information of firm A only if the nominal price of firm A moves in the exact same manner, simply because the nominal price of firm B cannot possibly be contingent on the private information of another firm. Furthermore, this is true whether the relevant private information is a private signal about the aggregate state of the economy or merely the firm's own productivity. This provides the intuition for why ψ_a and ψ_x are uniquely pinned down once the corresponding production coefficients are fixed. The specific formulas given in the lemma above follow from the optimality conditions of the household or, equivalently, the inverse demand functions faced by the firms.

A direct corollary of Proposition 3 and Lemma 5 is that by controlling l_A, and thereby l_x, monetary policy also influences ψ_x, the sensitivity of prices to local information. In particular, by (17), ψ_x is decreasing in l_A: the more firms expect monetary policy to accommodate the realized productivity shock, the higher their incentive to raise their prices in response to any private information they may have about aggregate productivity. This observation anticipates the role of monetary policy in regulating the aggregation of information through prices.

5.2 Optimal Allocations

We now proceed to identify the best implementable allocation and information structure. First, we express welfare as a function of the strategy coefficients presented in condition (11) and the precisions of the private and public sufficient statistics provided in (13).

Lemma 6. *There exists a function $\mathcal{W} : \mathbb{R}^{11} \to R$ such that the expected utility of the representative household can be expressed as $\mathcal{W}(\varphi, l; \kappa_x, \kappa_z)$.*

Next, we use Lemma 4 and Proposition 3 to express the planner's problem as follows.

Planner's Problem. *Choose strategy coefficients $\varphi = (\varphi_0, \varphi_a, \varphi_x, \varphi_z)$ and $l = (l_0, l_A, l_a, l_x, l_z)$ and precisions κ_x and κ_z so as to maximize $\mathcal{W}(\varphi, l; \kappa_x, \kappa_z)$ subject to (13) and (15)-(18).*

Like standard Ramsey policy problems, this problem imposes certain implementability constraints on the set of allocations that the planner can choose; as explained before, these constraints are summarized by conditions (15)-(18). In our problem the planner must also take into account that different allocations induce different information structures; this explains why the planner controls κ_x and κ_z, subject to condition (13).

The solution to this problem is characterized in the following result.

Proposition 4. *There exist scalars $\Delta_Y > 0$ and $\Delta_p > 0$, which depend on the information parameters, such that the optimal allocation satisfies*

$$\log q\left(\omega\right) = \varphi_0^* + \varphi_a^* a + \varphi_x^* x + \varphi_z^* z$$

$$\log l\left(\omega, \bar{a}, z\right) = l_0^* + l_A^* \bar{a} + l_a^* a + l_x^* x + l_z^* z$$

with the following coefficients:

$$\varphi_a^* = \beta \tag{20}$$

$$\varphi_x^* = \left\{ \frac{(1-\alpha)\kappa_x^*}{(1-\alpha)\kappa_x^* + \kappa_z^* + \kappa_A} \right\} \frac{\alpha}{1-\alpha}\beta + (\Delta_y + \Delta_p) \tag{21}$$

$$\varphi_z^* = \left\{ \frac{\kappa_z^*}{(1-\alpha)\kappa_x^* + \kappa_z^* + \kappa_A} \right\} \frac{\alpha}{1-\alpha}\beta - \frac{\kappa_z^*}{\kappa_A + \kappa_z^*}(\Delta_y + \Delta_p) \tag{22}$$

$$l_a^* = \hat{l}_a \tag{23}$$

$$l_x^* = \hat{l}_x + \left(\frac{\kappa_x^*}{\kappa_A + \kappa_x^* + \kappa_z^*} \right) \lambda \Delta_p \tag{24}$$

$$l_z^* = \hat{l}_z + \left(\frac{\kappa_z^*}{\kappa_A + \kappa_x^* + \kappa_z^*} \right) \lambda \Delta_p \tag{25}$$

$$l_A^* = \hat{l}_A - \lambda \Delta_p \tag{26}$$

where $(\hat{l}_a, \hat{l}_x, \hat{l}_z, \hat{l}_A)$ *are the coefficients that would have obtained if prices were flexible (but taxes were fixed at their sticky-price optimal level) and* λ *is a positive scalar that is invariant to information parameters (it only depends only on preferences and technologies).*

This result establishes that the impact of learning on the optimal implementable allocations resembles qualitatively that in the baseline model. In particular, the scalars Δ_y and Δ_p are the Lagrange multipliers that measure the social value of increasing the endogenous precisions of the quantity and price signals, respectively.

Consider first the sensitivity of first-stage decisions to information. Modulo the new definitions of α and β, the coefficients given by the conditions (21) and (22) coincide with their counterparts in Proposition 1 if we let the total information wedge be $\Delta \equiv \Delta_y + \Delta_p$. The intuition is of course the same. The planner corrects the information externality and boosts social learning by raising the sensitivity of first-stage decisions to private information. At the same time, it preserves allocative efficiency by lowering the sensitivity to public information.

Let us now turn to second-stage decisions. Suppose for a moment that agents did not learn from price signals (formally, let $\sigma_{zp}^2 = \sigma_{xp}^2 = \infty$). As a result, $\Delta_p = 0$, and by conditions (24)-(26) second-stage coefficients would coincide with their counterparts under flexible prices. Likewise, Proposition 4 would coincide exactly with Proposition 1 for the baseline model with condition (26) requiring that monetary policy be set so as to replicate flexible prices. Therefore, the novelty here relative to our baseline analysis is to allow for learning through prices and to show how this requires optimal monetary policy to depart from the benchmark of replicating flexible-price allocations without changing the essence of the optimal taxes.

To gain intuition for this result, recall that sticky prices provide the planner with an extra degree of freedom. Whether the planner uses this extra lever and, thus, moves away from the allocations that would obtain under flexible prices depends on whether the distorted allocations come with a further boost to social learning. This is the case when, and only when, prices serve as a signal of the state of the economy.

The argument above suggests that in the general case in which agents learn also from prices, optimal implementable allocations differ from those that would be consistent with flexible prices. More specifically, note that, as the planner increases φ_x to render social learning more effective, if monetary policy were to replicate the flexible-price outcomes, by Proposition 3 this increase in φ_x would have been associated with an increase in l_A. Condition (26) establishes that it is actually optimal to reduce l_A relative to this benchmark. That is, the optimal level of output moves less with productivity than its flexible-price counterpart (with taxes).

There are two complementary ways to understand this result. One is in terms of monetary policy mimicking a missing tax instrument. Another is in terms of the optimal covariation between nominal prices and real economic activity that monetary policy should aim to implement. We discuss the first perspective here and the second in the next subsection.

To understand the first perspective, let us momentarily abstract from sticky prices and monetary policy. Let us also put aside how Proposition 4 was obtained in the first place, where the wedges Δ_y and Δ_p originate, and what they capture. Instead, let us treat those wedges as exogenous and take as given that the planner wishes to implement the real allocations characterized in Proposition 4. What is the most direct way to achieve this goal if the planner has access to a completely unrestricted set of tax instruments?

The answer is that the planner should apply a differential subsidy on the early and the late production choices. In particular, the subsidy on early production choices should be higher than the subsidy on later production choices. To read this property from Proposition 4, note first that what we have called "flexible-price allocations" corresponds, under the present perspective, to the allocations implemented with a uniform subsidy across all production choices, i.e. to both n and l (equivalently q and l). It follows from condition (26) that relative to the case where all production choices are equally subsidized, late production choices ought to be subsidized *less* by an amount proportional to Δ_p. Equivalently, n has to be subsidized in proportion to the sum $\Delta_y + \Delta_p$, but l has to be subsidized only in proportion to Δ_y.

From this perspective, a monetary policy that leans further against the wind relative to the one that implements flexible-price outcomes assumes the role of the missing differential subsidy. Such a monetary policy, by contracting during a boom and expanding during a recession, mimics the effect of introducing a procyclical subsidy that has a bigger footprint on the late production choices than on the earlier ones. This as-if differential tax partly offsets the uniform subsidy, so that in the end the effective subsidy on l is proportional only to Δ_y, whereas that on n is proportional to the sum $\Delta_y + \Delta_p$, as desired.

This echoes how monetary policy works in the textbook New Keynesian model: in that model, too, whenever monetary policy departs from flexible-price allocations, it does so only to mimic a missing tax instrument (Correia, Nicolini, and Teles, 2008). In particular, in the oftconsidered case of a monetary policy that "leans against cost-push shocks," the missing tax instrument is the uniform (across inputs) state-contingent subsidy needed to offset a time-varying monopoly. Here, this kind of tax instrument is allowed, and thereby the rationale for this policy is different. Monetary policy still mimics *some* type of missing tax instrument, but in this case it is a subsidy that can differentiate between early and late choices.

Clearly, such a differential subsidy is hard to envision in the real world, especially given the difficulty of figuring out the exact timing of the different production choices and the information upon which they are based. This explains why we view a "direct" implementation implausible in practice. An "indirect" implementation via monetary policy helps bypass this problem for the following basic reason: insofar as certain production decisions take place at a similar time or on the basis of similar information as price-setting decisions, monetary policy will naturally have less control over these real decisions relative to those that *must* adjust after prices are fixed. In other words, monetary policy naturally has a differential real effect, akin to that of a differential tax. Importantly, monetary policy allows the planner to mimic the needed differential tax without explicit knowledge of which real decisions are made on the basis of restricted information, or which ones are most directly regulated by monetary policy.

This discussion explains the *mechanics* of the optimal monetary policy, but not the *function* served by it. To understand this, and to pave the way for the second perspective on what monetary policy accomplishes (which we turn to in the next subsection), we must ask why Proposition 4 holds in the first place, and in particular why optimality calls for a larger subsidy on n than on l. The abstract math is that n contributes to learning through both quantities and prices, wheres l contributes only through quantities. The logic is that, because firms choose n at the same time and on the basis of the same information as p, a more information-sensitive n translates to more information revelation through both quantities and prices. By contrast, because l adjusts after prices have been fixed, a more information-sensitive l contributes to more learning only through quantities. It follows that internalizing the social learning through quantities calls for a procyclical subsidy on both n and l, whereas internalizing the social learning through prices calls for a subsidy only on l, or equivalently for a monetary policy that mimics the aforementioned differential tax.

We expand on this logic and formally characterize the policy mix that supports the optimal allocation in the next subsection.

5.3 Optimal Policy Mix

To characterize the optimal combination of fiscal and monetary policy, we simply combine the the results in Propositions 3 and 4 to obtain the following result.

Proposition 5. *(i) The optimal tax is countercyclical, as in the baseline model. In particular,*

$$\tau_A^* + \tau_z^* = -\chi_1 \Delta_y - \chi_2 \Delta_p, \tag{27}$$

for some positive scalars χ_1, χ_2.

(ii) The optimal monetary policy is less accommodative of the productivity shock than the policy that replicates flexible-price allocations. In particular,

$$\rho_A^* = \hat{\rho}_A + \chi_3 \Delta_p, \tag{28}$$

where χ_3 *is a positive scalar and* $\hat{\rho}_A$ *is the coefficient in the interest-rate rule that replicates flexible-price outcomes.*

This result translates the informational wedges from Proposition 4 into policy prescriptions. Part (i) says that optimal taxes are countercyclical, part (ii) says that monetary policy deviates from the benchmark of replicating flexible prices towards further "leaning against the wind," or raising interest rates more aggressively during booms.

The intuition behind part (i) is similar to that in our baseline model. The novelty is that there are now two informational wedges driving optimal taxes, one related to the learning through quantities and the other related to the learning through prices. This is because countercyclical taxes incentivize firms to react more aggressively to their information when making their production and pricing choices, thereby improving the learning through both channels.

The logic for taxes holds true even if monetary policy were restricted to replicate flexible-price outcomes. For any given information structure, replication of flexible-price outcomes maximizes allocative efficiency by minimizing relative-price distortions. But part (ii) of Proposition 4 establishes that when and only when there is learning through prices, optimal monetary policy deviates from this benchmark: it trades-off less allocative efficiency, or more relative-price distortions, for the sake of inducing more learning through prices.

In the previous subsection, we have explained how the counter-cyclicality of the optimal monetary policy can be understood in terms of mimicking a missing tax instrument. We now expand on a second, complementary perspective, which relates to why the optimal monetary policy "leans against the wind" in the sense of inducing a negative relation between prices and real economic activity and, in particular, why this helps improve social learning.

To this goal, it is important to understand first the reference point from which the optimal monetary policy departs. For any given information structure, had monetary policy replicated flexible-price outcomes, it would have maximized allocative efficiency and minimized relative-price distortions. But it would *not* have stabilized the nominal price level. Instead, it would have allowed the nominal price level to move in the direction opposite that of real output, for the reason first explained in Angeletos and La'O (2020).

Let us first review this reason. Due to the real rigidity, it is efficient for the relative production of any two firms to vary with their relative information, or beliefs, about the state of the economy. Next, because the demand curve faced by any given firm is downward-sloping, in order to preserve efficient movements in relative production, a firm's nominal price must move in the direction opposite its real quantity in response to any variation in its private information. But aggregate booms tend to correlate with a wave of optimism across a majority of firms, while efficiency dictates that the prices of these firms move in the direction opposite their quantities. As a result, the aggregate price level must be negatively correlated with aggregate output in order to preserve allocative efficiency.

Let us now relate this benchmark to our own result. If taxes were set to zero and monetary policy were to replicate flexible-price outcomes, nominal prices would have moved in the direction opposite aggregate output and aggregate productivity. In our context, this would have *already* allowed prices to aggregate and reveal information about the state of economy. Starting from this benchmark, the cost of a slightly different monetary policy in terms of allocative efficiency is only second order. But the benefit of inducing more learning is first order—and this is precisely what the Δ_y and Δ_p wedges capture. Furthermore, because nominal prices are negatively related to real economic activity in the benchmark, the direction of the optimal deviation thereby is clear: it is optimal to raise *both* the positive correlation of the real quantities with the underlying fundamental and the corresponding negative correlation with prices.

For the reasons already explained, both goals can be accomplished in large part by a pro-cyclical subsidy, or counter-cyclical tax, on production. But insofar as there is learning through prices, monetary policy should lean even more against the wind relative to the flexible-price benchmark. If instead monetary policy were to move in the other direction, that of stabilizing the price level, it would both impede learning and reduce welfare.

Finally, were we to consider a version of this model with fully flexible prices, i.e. zero nominal rigidity, but with learning from both quantities and prices, then the optimal allocation described above would be unattainable. Under flexible prices we would no more have the extra lever afforded by monetary policy, and consequently we would no more be able to use that lever to mitigate the distortion between first and second-stage input choices: equilibrium would necessarily equate their marginal rate of transformation with their corresponding marginal rate of substitution (as in condition 19). In this restricted case, there would still be value for countercyclical taxes, but the optimal tax would be less countercyclical than the one characterized under sticky prices. By the same token, there would be less learning.

5.4 A Numerical Illustration

Our model is too stylized to permit a quantitative evaluation of the theoretical results obtained above. In our model all the learning takes place "within a period," whereas in a more realistic model it would take place across periods. Notwithstanding this limitation, it is instructive to illustrate our results within a "calibrated" version of our model.

We focus on a case where learning is all private. This can be motivated in at least one of two ways. First, in the tradition of Lucas (1972), one may have a prior that most learning happens in a decentralized fashion. Second, in the tradition of Sims (2003), Woodford (2003b) and Mackowiak and Wiederholt (2009), one may argue that even if aggregate statistics are readily available, firms may pay little attention to them, effectively observing them with idiosyncratic noise. Formally, we set $\sigma_{zy} = \sigma_{zp} = \infty$ and let all information be private.[22]

Consider first ϵ, the inverse of the Frisch elasticity of labor supply, and γ, which in models without capital captures mainly the income elasticity of labor supply. We follow Woodford (2003b) and set $\epsilon = 1.3$ and $\gamma = 0.2$. These values imply that the complete-information version is consistent with the empirical regularity that output and employment have the same cyclical behavior over the business cycle and real wages are relatively flat. We next set $\theta = 0.5$, which means that roughly half of production is fixed on the basis of incomplete information and half adjusts freely to realized aggregate demand. To calibrate the idiosyncratic risk in TFP, we refer to the NBER-CES Manufacturing Database, which computes a measure of TFP for all 6-digit NAICS manufacturing industries. This suggests a value for σ_ξ equal to 0.08, or an 8% standard deviation in firm-level TFP from year to year. For aggregate productivity, on the other hand, we set σ_A equal to 0.02, or a 2% standard deviation in aggregate TFP.

We do not have a strong prior on what is the appropriate parameterization of the degree of strategic complementarity. This naturally depends on whether one interprets α narrowly, as a measure of the aggregate demand externalities allowed in our simple model, or more broadly as a proxy for all other additional sources of complementarity left outside our model, such as those originating from financial frictions. Under the narrow interpretation, Angeletos and La'O (2010) argue that a value of α around or above 0.5 can be justified for an "elastic" neoclassical economy of the kind argued for in the RBC literature (King and Rebelo, 1999). We adopt this as our baseline value but explore how the results vary as we vary α in the entire $[0, 1]$ range.

What about the noise in the available private signals? For our baseline, we set $\sigma_{xy} = \sigma_{xp} = 0.03$ on the basis of the following rationale. These values, along with our value for idiosyncratic TFP, imply that the overall signal-to-noise ratio in a firm's overall information about aggregate output is equal to 1.[23] This means that whenever aggregate output goes up by one unit, the average forecast of it goes up by half a unit. In other words, the size of the forecast error is commensurate to the size of the innovation in the forecast. Such a pattern is broadly consistent with the evidence from surveys of macroeconomic forecasts documented in Coibion and Gorodnichenko (2012, 2015). We use these values as our baseline but explore how the results vary as we vary the noise in either signal.

Figure 2 computes the wedges Δ_Y and Δ_p as a function of α and the levels of noise in the two endogenous signals. Both wedges in the monetary model behave qualitatively the same as

[22]The assumption $\sigma_{zp} = \infty$ can also be motivated on purely empirical grounds, that the signal contained in inflation about real economic activity is almost negligible (e.g., Angeletos, Collard, and Dellas, 2018).

[23]The variance of a firm's posterior about aggregate TFP is $(\sigma_\xi^{-2} + \sigma_{xy}^{-2} + \sigma_{xp}^{-2})^{-2}$, which under our parameterization is approximately $(0.02)^2$, or the same as σ_A^2. And since aggregate output is proportional to aggregate TFP, this verifies the claim made above.

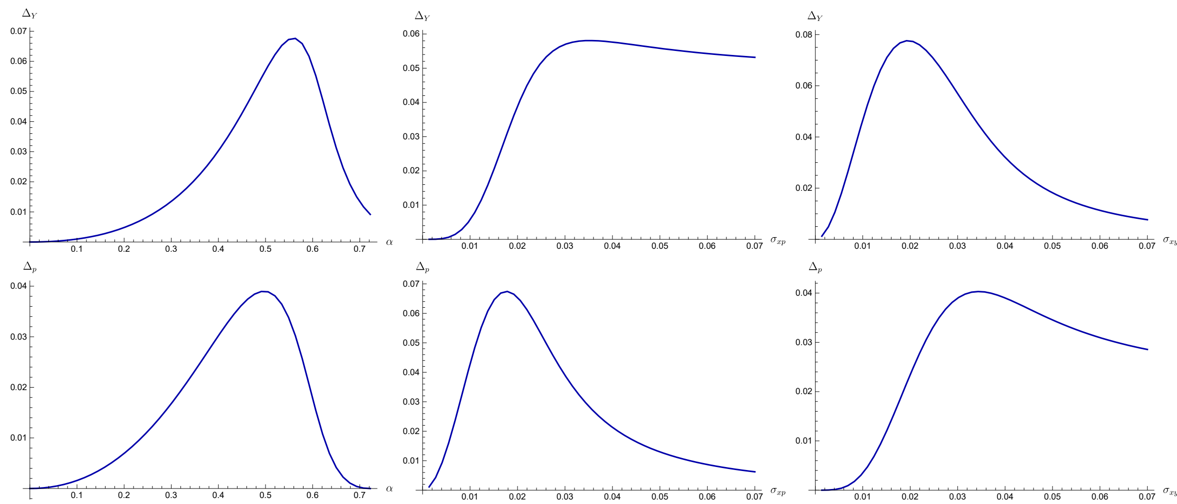

Figure 2. The two wedges (Δ_Y in the first row, Δ_p in the second) for different values of the strategic complementarity (first column), the noise in the price signal (second column) and the noise in the quantity signal (third column).

the wedge in the baseline model.[24]

We now turn to optimal policy. Because aggregate output is a log-linear function of aggregate TFP, we can readily re-express the tax and the nominal interest rate that support the optimal allocation as follows:

$$-\log(1-\tau) = \tau_Y^* \log Y \qquad \text{and} \qquad \log(1+R) = \rho_Y^* \log Y.$$

Under this representation, τ_Y^* measures the optimal cyclical elasticity of the tax and ρ_Y^* measures the optimal cyclical elasticity of the nominal interest rate.[25] Similarly, let $\hat{\rho}_Y$ denote the cyclical elasticity of the nominal interest rate required to replicate the corresponding flexible-price allocation. The difference $\rho_Y^* - \hat{\rho}_Y$ thereby provides a simple measure of the countercyclicality of monetary policy: the more positive this quantity is, the less accommodative the optimal monetary policy is over the business cycle. Finally, to assess the overall impact of the optimal fiscal-and-monetary policy mix, we compute the quantity $y_A^*/y_A^\circ - 1$, where y_A^* and y_A° denote the elasticities of aggregate output to TFP under, respectively, the planner's optimum and the "laissez faire" equilibrium in which taxes are non-contingent and monetary policy replicates flexible prices.

Figure 3 illustrates how the optimal countercyclicality of fiscal and monetary policy depends on the degree of strategic complementarity and the level of noise in the available price and quantity signals. Let us first comment on the signs of the measures seen in this figure. The negative value for τ_Y^* means that the optimal taxes are countercyclical. Similarly, the positive

[24]Also note that when $\theta < 1$, there is an upper bound on α, which is obtained by taking the highest admissible value of ρ; this is given by 0.72 under our parameterization. This explains the domain of α in the figure.

[25]These elasticities are given simply by $\tau_Y^* = \tau_A^*/y_A^*$ and $\rho_Y^* = \rho_A^*/y_A^*$, where τ_A^* and ρ_A^* are the corresponding elasticities in terms of productivity, as characterized in Proposition 5, and y_A^* is the elasticity of aggregate output to aggregate productivity along the optimal allocation.

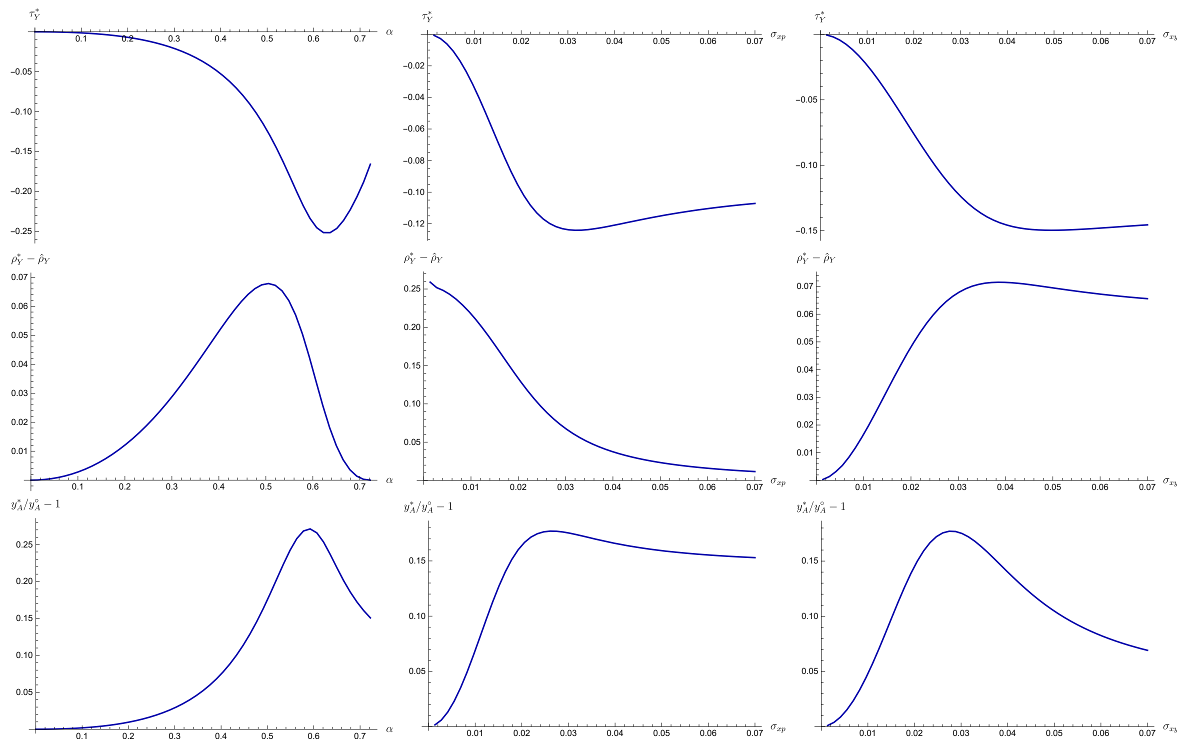

Figure 3. The countercyclicality of the tax (first row) and the countercyclicality of monetary policy, as manifested in nominal interest rates (second row), and their overall effect on output (third row), for different values of the strategic complementarity (first column), the level of noise in the price signals (second column), and the level of noise in the quantity signals (last column).

value for $\rho_Y^* - \hat{\rho}_Y$ indicates that optimal monetary policy "leans against the wind:" interest rates are higher than their flexible-price counterparts. Together, fiscal and monetary policy render output more responsive to the underlying TFP shock than in the laissez faire benchmark with non-contingent taxes and a monetary policy that replicates flexible prices.

Let us next turn to the effect of α, the degree of strategic complementarity. Here, we see the policy translation of the non-monotonic pattern we documented earlier for the wedges: the value of both countercylical taxes and countercyclical monetary policy is highest when α is neither too high nor too low. For extreme values of α, there is little such value either because firm decisions are nearly independent (for low α) or because the complementarity is so strong that firms optimally disregard any private information and hence there is little scope for social learning to begin with (for high α).

The effect of the two levels of noises on the optimal policies also mirror their effects on the wedges. The only subtlety here is the following. As explained earlier on, the optimal monetary policy deviates from the benchmark of replicating flexible prices only insofar as there is learning through prices. It follows that when the noise in the price signals is sufficiently large, such a deviation is not worthwhile, which in turn explains why $\rho_Y^* - \hat{\rho}_Y$ converges to zero as σ_{xp} alone goes to infinity. By contrast, because countercyclical taxes serve the dual role of improving the aggregation of information through both quantities and price signals, τ_Y^* stays bounded away from zero as either σ_{xp} or σ_{xy} become larger and larger. For τ_Y^* to vanish, both sources of learning have to be muted.

Finally, let us note that although our model is far too stylized to allow for a serious quantitative evaluation, Figure 3 indicates that the documented effects could be non-trivial. Depending on the exact parameterization (within the range considered), the optimal policy may have taxes decrease up to 25 bps and the nominal rate increase up to 25 bps for every percentage point increase in output over the business cycle; the optimal output response is greater than the laissez-faire counterpart by a comparable amount.

6 Discussion

In this section, we discuss the robustness of our insights to various perturbations of the environment and hint at possible directions for future work.

6.1 Imperfectly informed policy

Our analysis has focused on the imperfection of information within the private sector, but has assumed that policy itself can be contingent on the true state. Let us first clarify that this assumption does not require an informational asymmetry between the policymaker and the private sector at any given point of time.

To see this, consider our extended model. The private decisions made in stage 2 are conditioned on exactly the same information as that upon which the policymaker sets taxes and monetary policy. Still, policy "works" because it affects incentives faced in stage 1, when that information is coarser. This highlights that our basic insights hold even if policy can depend only on public information, provided that: (i) such information arrives gradually; (ii) policy can react to public information that was not available at the time certain private decisions were made; and (iii) private decisions are influenced by the anticipation of such future policy reaction. Such *anticipatory* effects of policy are common place in richer macroeconomic models, where private decisions are forward looking. The novelty here is to illustrate how they can regulate the aggregation of information.

In this light, it is essential that policy responds to information that was not originally available to private agents. This is what the contingency of the taxes or the interest rate on the realized productivity captures. But such future information of the policymaker need not be perfect.

To see this, consider our baseline model and let taxes depend on two noisy public signals: the original one available in stage 1, and an additional one which becomes available at stage 2. Denote the former by z_t, as before, and the latter by z'_t. Let the latter be given by $z'_t = a_t + \epsilon'_t$, where ϵ'_t is the measurement error in the policymaker's observation of the fundamental in stage 2. The original analysis is nested by letting the variance of ϵ'_t be zero. More generally, as long as this measurement error is unpredictable, every agent's stage-1 expectation of the policy satisfies

$$\mathbb{E}_i\left[\tau(z'_t, z_t)\right] = \mathbb{E}_i\left[\tau(a_t, z_t)\right].$$

It follows that the set of stage-1 decisions implemented by policies contingent on (z'_t, z_t) is the same as that implemented by policies contingent on (a_t, z_t). And because, at least in our baseline model, only stage-1 decisions enter the precisions of the endogenous signals, it follows that the two types of policies are equally effective in regulating the aggregation of information.

This of course does not mean that such measurement error is of no consequence. To the extent that the variation in taxes (or monetary policy) induced by such measurement error necessarily distorts private decisions made in stage 2, this may naturally limit the desirability of the state-contingent policies we have identified. Intuitively, we expect the optimal cyclicality of taxes and of monetary policy to attenuate with the level of such measurement error. Clearly, this consideration could be important quantitatively, but it does not change the qualitative lesson.

In the above argument we let z'_t be an *exogenous* signal of the fundamental and concluded that the noise in it is likely to reduce (in absolute value) the optimal τ_A. But now consider the case in which z'_t is an *endogenous* signal, given for example by $z'_t = \log Y_t + \epsilon'_t$. In this case, there is a force contributing in the opposite direction: a higher absolute value for τ_A reveals valuable information to the policymaker itself. This appears to reinforce our message in the following sense: the kind of policies we characterize in this paper may facilitate not only more learning within the private sector but also more information transmission from the private sector to the policymaker and therefore to better, more informed, policies. We leave the investigation of this idea for future work.

6.2 Learning from policy

Throughout our analysis we have assumed that agents observe taxes and the interest rate only in the second stage, after the fundamental is revealed and all actions have been taken. But what if agents can observe a signal of the policy itself—the tax rate or the interest rate—at the moment they make their decisions? We now sketch why this possibility does not change the essence of the policy problem. In fact, it only reinforces our policy lessons.

To start with, consider our baseline real model and let agents observe, in addition to the previously introduced signals, the following private signal about the tax rate:[26]

$$x^\tau = -\log(1 - \tau(\bar{a}, z)) + \epsilon^{x\tau},$$

[26] The focus on private signals facilitates the two interpretations discussed next. But the formal argument would not change if the signal were public.

where $\varepsilon^{x\tau}$ is Gaussian idiosyncratic noise, with fixed variance. This noise could be interpreted as a proxy for rational inattention.[27] Alternatively, x^τ could be the *actual* tax faced by an agent, in which case $\epsilon^{x\tau}$ corresponds to an idiosyncratic tax shock. Regardless of the interpretation, using our log-linear specification of τ, we have that x^τ contains the same information as that of the following, "normalized" signal:

$$\bar{x}^\tau \equiv \frac{x^\tau - \tau_0 - \tau_z z}{\tau_A} = \bar{a} + \widetilde{\varepsilon}^{x\tau},$$

where $\widetilde{\varepsilon}^{x\tau} \equiv \varepsilon^{x\tau}/\tau_A$. Thus, by choosing a larger absolute value for τ_A, the planner can now induce more learning via the above signal. Intuitively, the more counter-cyclical taxes are, the more informative is any signal about them.

How does this modify the optimal policy problem? In our baseline analysis, the planner chooses a negative value for τ_A so as to induce more cyclicality in aggregate output and thereby more learning through signals of aggregate output. This was an *indirect* way of manipulating the information available to agents. Now, there is also a *direct* way of doing so, calling for an even more negative value for τ_A.[28]

A similar logic applies to monetary policy. Consider our extended monetary model and let agents observe a noisy private signal of the interest rate. Again, the noise can be interpreted as either a proxy for rational inattention or as an idiosyncratic interest-rate shock. In either case, a monetary policy that induces a more countercyclical aggregate interest rate is beneficial not only for the reasons explained in our main analysis, but also because it directly boosts agents' learning from interest rates.

[27]Such an interpretation stretches the assumption that the variance of $\varepsilon^{x\tau}$ is exogenously fixed. See Subsection 6.4 for why this issue may not be essential.

[28]Let us sketch the formal argument. In the original model, welfare was expressed as $W(\varphi, \kappa)$, where φ is the elasticity of output with respect to productivity and κ is the vector of precisions of the two sufficient statistics. Furthermore, κ was a function of only φ, which was itself regulated by τ_A. It followed that the optimal choice of τ_A was solving the following FOC:

$$\left\{\frac{\partial W}{\partial \varphi} + \frac{\partial W}{\partial \kappa}\frac{\partial \kappa}{\partial \varphi}\right\}\frac{\partial \varphi}{\partial \tau_A}\bigg|_{\tau_A^*} = 0,$$

where $\frac{\partial \varphi}{\partial \tau_A}$ was encapsulating implementability. Now, welfare is still given by the same function of φ and κ, but now the latter depends not only on φ (in essentially the same way as before) but also on τ_A (because of the direct effect of τ_A on the precision of $\bar{z}^\tau$). As a result, the total derivative of welfare with respect to τ_A is now given by

$$\frac{dW}{d\tau_A} = \left\{\frac{\partial W}{\partial \varphi} + \frac{\partial W}{\partial \kappa}\frac{\partial \kappa}{\partial \varphi}\right\}\frac{\partial \varphi}{\partial \tau_A} + \frac{\partial W}{\partial \kappa}\frac{\partial \kappa}{\partial \tau_A}.$$

Note that $\frac{\partial W}{\partial \kappa} > 0$ (more information is socially valuable, for the reasons already explained) and $\frac{\partial \kappa}{\partial |\tau_A|} > 0$ (precision increases with the absolute value of τ_A). Evaluating the above at τ_A^*, i.e. the tax that is optimal in the absence of the policy signal, gives (remember that $\tau_A^* < 0$)

$$\frac{dW}{d\tau_A} = \frac{\partial W}{\partial \kappa}\frac{\partial \kappa}{\partial \tau_A} = -\frac{\partial W}{\partial \kappa}\frac{\partial \kappa}{\partial |\tau_A|} < 0,$$

which proves that it is optimal, at least locally, to reduce τ_A below τ_A^*. To complete the argument, one only has to verify that this local-deviation logic extends globally. We suspect that this can be done in a similar way as in the proof of Proposition 1.

6.3 Removing state-contingent taxes

Throughout we have allowed the planner to vary taxes with the business cycle, thus providing economic agents with the incentive to respond more strongly to their private information about the state of the economy. We have shown that this leads monetary policy to deviate from the benchmark of replicating flexible-price allocations towards inducing a more counter-cyclical price level. But what if such state-contingent taxes were unavailable?

In this case, monetary policy must substitute not only for the *differential* tax missing in our main analysis, but now also for the *uniform* tax that was previously used to induce efficient first-stage production choices. In more practical terms, this means that monetary policy must balance its previously-articulated role of inducing a more counter-cyclical price level with its new role of inducing more pro-cyclical aggregate output. Depending on which role is more important, which in turn depends on which form of learning (from prices or quantities) is stronger, monetary policy could now be either more or less counter-cyclical than the one that replicates flexible prices.

To illustrate this point, the left panel of Figure 4 plots the wedge between the optimal sensitivity of the interest rate to aggregate output when taxes are constrained to be zero (denoted by $\rho^*_{Y,\tau=0}$) and its counterpart in the benchmark case where the planner ignores the informational externality and replicates flexible prices (denoted by $\hat{\rho}_{Y,\tau=0}$); this wedge is plotted for different parameterizations of the precisions of the endogenous private signals. More specifically, we shut down the public signals, we fix the sum of the precisions of the two endogenous private signals (i.e., the sum $1/\sigma^2_{xy} + 1/\sigma^2_{xp}$), and we vary their ratio. As we move from the left to the right on the x axis, we therefore let the output signal become more precise at the expense of the price signal. In the right panel, we report the corresponding optimal counter-cyclicality of the aggregate price level.

The qualitative message of the figure is clear. When learning occurs mostly through prices, the role articulated in our main analysis dominates and monetary policy is more "hawkish" than the one that implements flexible prices (i.e., $\rho^*_{Y,\tau=0} > \hat{\rho}_{Y,\tau=0}$). But once learning through quantities becomes sufficiently important, the need to substitute for the missing counter-cyclical taxes takes over and monetary policy turns "dovish" (i.e., $\rho^*_{Y,\tau=0} < \hat{\rho}_{Y,\tau=0}$).

Which of these two scenarios is more relevant in practice, and how they interact with our earlier point about policy signals, is an empirical/quantitative question beyond the scope of our paper. We would thus like to close this subsection with the following two remarks.

First, when learning from quantities is relatively more important, there may be a rationale for "dovish" monetary policy in the sense described above, but this does not necessarily translate to targeting a less counter-cyclical price level. This is evident in the right panel of Figure 4, where the optimal sensitivity of the price level is shown to be relatively flat in the relative importance

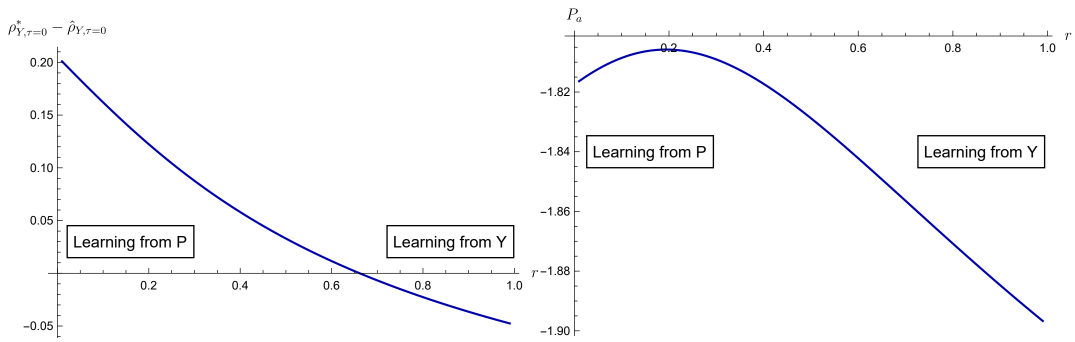

Figure 4. The wedge $\rho^*_{Y,\tau=0} - \hat{\rho}_{Y,\tau=0}$ (left panel a) and the sensitivity P_a of the price level to aggregate TFP (right panel) for different values of $r \in (0,1)$, where $1/\sigma^2_{xy} = r\bar{\kappa}$ and $1/\sigma^2_{xp} = (1-r)\,\bar{\kappa}$, for some positive constant $\bar{\kappa}$.

of two kinds of learning.[29]

Second, in the US data all relevant macroeconomic outcomes (aggregate employment, output, the price level, the interest rate) are nearly orthogonal to utilization-adjusted TFP at business-cycle frequencies. This does not mean that our theory is inapplicable. The productivity shock in our model is merely a proxy for all shocks that drive efficient business cycles. It follows that the resolution to the aforementioned question hinges on the resolution to the more challenging, long-standing question of what fraction of the business cycle is efficient.

6.4 Rational inattention

Like Woodford (2003a) and others, we have invited the interpretation of the idiosyncratic noise in the observation of the available quantity, price, and policy signals as a proxy for rational inattention à la Sims (2003). We have not, however, taken this interpretation "seriously" in the sense of endogenizing the agents' inattention or information-acquisition choice. Had we done so, policy could regulate the information structure not only via the channel studied so far (the equilibrium aggregation of information) but also via an additional channel: the equilibrium collection of information.

This, however, need not change our main conclusions. To understand why, note first that there is no deep conceptual difference between the *use* of information and the *collection* of information: both notions ultimately relate to how closely actions track the underlying fundamental. Efficiency in the *use* of information thereby naturally extends to efficiency in the *collection* of information, at least insofar as we model the cost of information as an arbitrary but fixed function of the joint distribution between the agents' actions and the underlying fundamentals. This point has been formalized in Angeletos and La'O (2020) for an environment similar to the one considered here but excludes the endogenous *aggregation* of information.[30] It follows that the informational externality originating in the aggregation of information remains

[29] Of course, this figure serves only as proof of concept.

[30] See in particular Online Appendix B of Angeletos and La'O (2020).

the sole source of inefficiency, and its correction remains the sole goal for policy. In a nutshell, the lessons developed here are robust to endogenizing the acquisition of information.

There are, however, two possible exceptions to this statement (and to the result of Angeletos and La'O (2020) upon which the above argument rests). The first regards the aforementioned assumption that the cost of information can be expressed as a fixed, policy-invariant function of the joint distribution between the agents' actions and the *exogenous* state of nature. This assumption is commonplace in the macroeconomics literature on rational inattention, (e.g., Mackowiak and Wiederholt, 2009, 2015; Myatt and Wallace, 2012; Sims, 2010). But it may be violated if the cost of information depends on the joint distribution between the agents' actions and *endogenous* economic outcomes: think of agents tracking the relevant quantities and prices directly, as opposed to tracking the underlying, possibly much richer, state of nature. As explained in Angeletos and Sastry (2019) and Hébert and La'O (2020), this possibility may or may not give rise to an additional externality, depending on certain "details" of the cost functional, such as whether the cost functional takes the Sims-Shannon mutual-information form or other forms proposed in recent decision-theoretic literature. How this subtle issue interacts with our paper's policy lesson is an open question.

The second exception regards risk sharing. Our paper and Angeletos and La'O (2020) alike have abstracted from incomplete risk sharing by assuming that all agents belong to a single "big family." This guarantees that there is no social value from reducing the inequality caused by dispersed information or rational inattention. If, instead, one allows for such a social value, and if there is also a limit to how much redistribution can be obtained via taxes, then there could be a reason for regulating the use of information away from the benchmarks we have characterized here. Colombo, Femminis, and Pavan (2014) provide an example of this sort, albeit in a model that abstracts from endogenous information aggregation.

6.5 Inefficient business-cycle shocks

Throughout the analysis, we have focused on productivity shocks. These can be thought more generally as proxies for shocks that trigger *efficient* business cycles. What if we were instead to consider shocks that trigger socially inefficient business cycles, such as shocks to monopoly markups or other distortions?

When the planner can undo the impact of such shocks on allocations with appropriate state-contingent taxes, the planner can also guarantee that welfare is invariant to the precision of information about them. In this "ideal" case, the optimal taxes and the optimal monetary policy are determined *as if* information were exogenous. But if policy is sufficiently constrained so that equilibrium allocations are sensitive to such shocks, welfare is likely to decrease with the precision of information about them.[31] It follows that, with inefficient fluctuations, the

[31] See Angeletos, Iovino, and La'O (2016) for the articulation of this point in a model that is similar to the present

basic logic behind our results works in reverse: the policymaker now wishes to minimize the information revealed through quantities and prices about these shocks. Paciello and Wiederholt (2014) and Baeriswyl and Cornand (2010) study models featuring inefficient fluctuations but abstract from the endogenous aggregation of information. The former focuses on the firms' collection of information, the latter on the signaling role of policy. A bridge between their analyses and ours is an interesting direction for future research.

7 Conclusion

In this paper we have shown that the endogeneity of information contained in macroeconomic statistics and market outcomes about the state of the economy calls for counter-cyclical taxes coupled with a monetary policy that leans against the wind. We have explained how such policies incentivize firms to act more aggressively on their private information about the economy, thus improving the information revealed to other firms, boosting allocative efficiency, and reducing the "noise" in business cycles. We have distinguished two main channels of such learning, one through real quantities and another through nominal prices, and have explained how each of them contributes to shaping optimal policy.

Our model contains three key features: (i) a real rigidity due to informational friction; (ii) a nominal rigidity due to informational friction; and (iii) the endogenous aggregation of information. Although subsets of these features can be found in previous work, their *combination* is novel to the literature and critical to our policy lessons.

The real rigidity, isolated in our baseline analysis, was essential for understanding why there is social value in improving the aggregation of information in the first place: were it not for the real aspect of the informational friction, a monetary policy that replicates flexible prices would have implemented the *complete-information* first-best outcomes, negating the value of any intervention. The nominal rigidity, on the other hand, was key to letting monetary policy be non-neutral and, hence, be able to assist in the aforementioned goal. And finally, were it not for the endogeneity of information aggregation, the state-contingency of taxes would be unnecessary and there would be no reason for monetary policy to depart from replicating the flexible-price benchmark.

In deriving these lessons, we have allowed the planner to vary taxes with the business cycle, thus providing economic agents with an incentive to respond more strongly to their private information about the state of the economy. In the New Keynesian literature, such state-contingent taxes are typically assumed away so as to open the door for monetary policy to stabilize the economy against inefficient fluctuations, such as those triggered by shocks to monopoly power or other distortions. Such shocks are absent here and so, too, is the

one but abstracts from endogenous information.

standard rationale for state-contingent taxes and monetary policy. Instead, state-contingent taxes and monetary policy are useful because they serve a novel function: they help internalize informational externalities and boost social learning. The quantitative evaluation of this function is an open question.

References

Adam, Klaus. 2007. “Optimal Monetary Policy with Imperfect Common Knowledge.” *Journal of Monetary Economics* 54 (2):267–301.

Amador, Manuel and Pierre-Olivier Weill. 2010. “Learning from Prices: Public Communication and Welfare.” *Journal of Political Economy* 118 (5):866–907.

Angeletos, George-Marios, Fabrice Collard, and Harris Dellas. 2018. “Business Cycle Anatomy.” *NBER Working Paper No. 24875* .

Angeletos, George-Marios, Luigi Iovino, and Jennifer La’O. 2016. “Real Rigidity, Nominal Rigidity, and the Social Value of Information.” *American Economic Review* 106 (1):200–227.

Angeletos, George-Marios and Jennifer La’O. 2010. “Noisy Business Cycles.” In *NBER Macroeconomics Annual 2009, Volume 24*. University of Chicago Press, 319–378.

———. 2020. “Optimal Monetary Policy with Informational Frictions.” *Journal of Political Economy* 128 (3):1027–1064.

Angeletos, George-Marios and Alessandro Pavan. 2007. “Efficient Use of Information and Social Value of Information.” *Econometrica* 75 (4):1103–1142.

———. 2009. “Policy with Dispersed Information.” *Journal of the European Economic Association* 7 (1):11–60.

Angeletos, George-Marios and Karthik Sastry. 2019. “Inattentive Economies: General Equilibrium and Welfare Theorems.” *MIT mimeo* .

Baeriswyl, Romain and Camille Cornand. 2010. “The Signaling Role of Policy Actions.” *Journal of Monetary Economics* 57 (6):682 – 695.

Ball, Laurence, N. Gregory Mankiw, and Ricardo Reis. 2005. “Monetary Policy for Inattentive Economies.” *Journal of Monetary Economics* 52 (4):703–725.

Chari, Varadarajan V and Patrick J Kehoe. 1999. “Optimal Fiscal and Monetary Policy.” In *Handbook of Macroeconomics*, vol. 1. Elsevier, 1671–1745.

Chari, V.V., Lawrence Christiano, and Patrick Kehoe. 1994. “Optimal Fiscal Policy in a Business Cycle Model.” *Journal of Political Economy* 102 (4):617–52.

Coibion, Olivier and Yuriy Gorodnichenko. 2012. “What Can Survey Forecasts Tell Us about Information Rigidities?” *Journal of Political Economy* 120 (1):116–159.

———. 2015. “Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts.” *American Economic Review* 105 (8):2644–78.

Colombo, Luca, Gianluca Femminis, and Alessandro Pavan. 2014. “Information acquisition and welfare.” *The Review of Economic Studies* 81 (4):1438–1483.

Correia, Isabel, Emmanuel Farhi, Juan Pablo Nicolini, and Pedro Teles. 2013. “Unconventional Fiscal Policy at the Zero Bound.” *American Economic Review* 103 (4):1172–1211.

Correia, Isabel, Juan Pablo Nicolini, and Pedro Teles. 2008. “Optimal Fiscal and Monetary Policy: Equivalence Results.” *Journal of Political Economy* 116 (1):141–170.

Gaballo, Gaetano. 2018. “Price dispersion, private uncertainty, and endogenous nominal rigidities.” *The Review of Economic Studies* 85 (2):1070–1110.

Grossman, Sanford J. and Joseph E. Stiglitz. 1976. “Information and Competitive Price Systems.” *The American Economic Review Papers and Proceedings* 66 (2):246–253.

———. 1980. “On the Impossibility of Informationally Efficient Markets.” *The American Economic Review* 70 (3):393–408.

Hébert, Benjamin and Jennifer La’O. 2020. “Information Acquisition, Efficiency, and Non-Fundamental Volatility.” *Columbia/Stanford Mimeo* .

King, Robert G and Sergio T Rebelo. 1999. “Resuscitating Real Business Cycles.” In *Handbook of macroeconomics*, vol. 1. Elsevier, 927–1007.

Laffont, Jean-Jacques. 1985. “On the Welfare Analysis of Rational Expectations Equilibria with Asymmetric Information.” *Econometrica: Journal of the Econometric Society* :1–29.

Lorenzoni, G. 2010. “Optimal Monetary Policy with Uncertain Fundamentals and Dispersed Information.” *The Review of Economic Studies* 77 (1):305–338.

Lucas, Robert E. Jr. 1972. “Expectations and the Neutrality of Money.” *Journal of Economic Theory* 4 (2):103–124.

———. 1973. “Some International Evidence on Output-Inflation Tradeoffs.” *American Economic Review* 63 (3):326–334.

Lucas, Robert Jr. and Nancy L. Stokey. 1983. “Optimal Fiscal and Monetary policy in an Economy without Capital.” *Journal of Monetary Economics* 12 (1):55–93.

Mackowiak, Bartosz and Mirko Wiederholt. 2009. “Optimal Sticky Prices under Rational Inattention.” *American Economic Review* 99 (3):769–803.

———. 2015. “Business Cycle Dynamics under Rational Inattention.” *Review of Economic Studies* 82 (4):1502–1532.

Mankiw, N. Gregory and Ricardo Reis. 2002. “Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve.” *Quarterly Journal of Economics* 117 (4):1295–1328.

Messner, Simon and Xavier Vives. 2001. “Informational and Economic Efficiency in REE with Asymetric Information.” *INSEAD Working Paper Series* .

Morris, Stephen and Hyun Song Shin. 2002. “Social Value of Public Information.” *American Economic Review* 92 (5):1521–1534.

Myatt, David P and Chris Wallace. 2012. “Endogenous Information Acquisition in Coordination Games.” *Review of Economic Studies* 79 (1):340–374.

Paciello, Luigi and Mirko Wiederholt. 2014. “Exogenous Information, Endogenous Information, and Optimal Monetary Policy.” *Review of Economic Studies* 81 (1):356–388.

Sims, Christopher A. 2003. “Implications of Rational Inattention.” *Journal of Monetary Economics* 50 (3):665–690.

———. 2010. “Rational Inattention and Monetary Economics.” *Handbook of Monetary Economics* 3:155–181.

Townsend, Robert M. 1983. “Forecasting the forecasts of others.” *The Journal of Political Economy* :546–588.

Vives, Xavier. 1988. “Aggregation of Information in Large Cournot Markets.” *Econometrica* :851–876.

———. 2017. “Endogenous Public Information and Welfare in Market Games.” *The Review of Economic Studies* 84 (2):935–963.

Woodford, Michael. 2003a. “Imperfect Common Knowledge and the Effects of Monetary Policy.” *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps* .

———. 2003b. *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press.

A Appendix

A.1 Proofs for Section 3

Proof of Lemma 1. Take any log-linear strategy of the form $q(a,x,z) = \varphi_0 + \varphi_a a + \varphi_x x + \varphi_z z$, for arbitrary coefficients $(\varphi_0, \varphi_a, \varphi_x, \varphi_z)$. The endogenous public signal is then given by

$$z^q = \log Q(\bar{a}, z) + \varepsilon^{zq}$$

where

$$\log Q(\bar{a}, z) = \varphi_0' + \varphi_a a + \varphi_x x + \varphi_z z$$

is the log of aggregate output. It follows that the public signal z^q can be transformed into an unbiased Gaussian signal $\tilde{z}^q$ about aggregate productivity, defined as follows:

$$\tilde{z}^q \equiv \frac{z^q - \varphi_0' - \varphi_z z}{\varphi_a + \varphi_x} = \bar{a} + \tilde{\varepsilon}^{zq}$$

where $\tilde{\varepsilon}^{zq} \equiv \varepsilon^{zq}/(\varphi_a + \varphi_x)$. The precision of this signal is

$$\kappa_{zq} \equiv \frac{1}{Var(\tilde{\varepsilon}^{zq})} = (\varphi_a + \varphi_x)^2 \sigma_{zq}^{-2}.$$

Standard Bayesian updating then implies that the sufficient statistic z of available public information is given by a weighted average of the exogenous productivity signal z^a and the (normalized) endogenous output signal $\tilde{z}^q$:

$$z = \frac{\kappa_{za}}{\kappa_z} z^a + \frac{\kappa_{zq}}{\kappa_z} \tilde{z}^q,$$

where κ_{za} and κ_{zq} are the precisions of these two signals, while $\kappa_z = \kappa_{za} + \kappa_{zq}$ is the overall precision of the sufficient statistic z.

The analysis of the private signal x_i^q is similar: it can be transformed into an unbiased signal with precision $\kappa_{xq} = (\varphi_a + \varphi_x)^2 \sigma_{xq}^{-2}$. QED.

Proof of Lemma 2. Take any log-linear strategy of the form (1). For arbitrary coefficients $\varphi = (\varphi_0, \varphi_a, \varphi_x, \varphi_z)$ and arbitrary precisions $\kappa = (\kappa_x, \kappa_z)$, the implied level of welfare (ex-ante utility) can be expressed as:

$$\mathbb{E}u = \mathcal{W}(\varphi;\kappa) \equiv \frac{1}{1-\gamma} \exp V_c(\varphi;\kappa) - \frac{1}{\epsilon} \exp V_n(\varphi;\kappa), \tag{A.1}$$

where

$$\begin{aligned} V_c(\varphi;\kappa) \equiv\ & (1-\gamma)\left(\varphi_0 + (\varphi_a + \varphi_x + \varphi_z)\mu\right) \\ & + \frac{1}{2}(1-\gamma)\left(\frac{\rho-1}{\rho}\right)\left[\frac{\varphi_a^2}{\kappa_\xi} + \frac{\varphi_x^2}{\kappa_x} + 2\frac{\varphi_a\varphi_x}{\kappa_x}\right] + \frac{1}{2}(1-\gamma)^2\left[\frac{\varphi_z^2}{\kappa_z} + \frac{(\varphi_a+\varphi_x+\varphi_z)^2}{\kappa_A}\right], \end{aligned} \tag{A.2}$$

$$\begin{aligned} V_n(\varphi;\kappa) \equiv\ & \frac{\epsilon}{\theta}\left(\varphi_0 + (\varphi_a + \varphi_x + \varphi_z - 1)\mu\right) \\ & + \frac{1}{2}\frac{\epsilon^2}{\theta^2}\left[\frac{(\varphi_a-1)^2}{\kappa_\xi} + \frac{\varphi_x^2}{\kappa_x} + 2\frac{(\varphi_a-1)\varphi_x}{\kappa_x} + \frac{\varphi_z^2}{\kappa_z} + \frac{(\varphi_a+\varphi_x+\varphi_z-1)^2}{\kappa_A}\right]. \end{aligned} \tag{A.3}$$

QED.

Proof of Proposition 1. We prove the two parts of the proposition together. Recall from Lemma 1 that any given strategy induces a κ_x and a κ_z as functions of $\varphi_a + \varphi_x$; let $K_1(\varphi_a + \varphi_x)$ and $K_2(\varphi_a + \varphi_x)$ denote the first and second element of the vector $K(\varphi_a + \varphi_x)$ as defined in (3). We can then express the planner's problem as follows:

Planner's problem. Choose $\varphi = (\varphi_0, \varphi_a, \varphi_x, \varphi_z)$ and $\kappa = (\kappa_x, \kappa_z)$ so as to maximize $W(\varphi; \kappa)$ subject to $\kappa = K(\varphi_a + \varphi_x)$.

To solve this problem, we proceed in two steps. The first step is to characterize the strategy that is optimal subject to the constraint that the sum $\varphi_a + \varphi_x$ is kept constant at some $\bar{\varphi} \in \mathbb{R}$ and accordingly the precisions κ_x and κ_z are kept constant at $\kappa_x = K_1(\bar{\varphi})$ and $\kappa_z = K_2(\bar{\varphi})$. The second step is to optimize over the sum $\bar{\varphi}$ and the precision κ_x and κ_z subject to the constraint that $\kappa_x = K_1(\bar{\varphi})$ and $\kappa_z = K_2(\bar{\varphi})$. The first step permits us to characterize the efficient allocation as a function of the Lagrange multiplier associated with the constraint $\varphi_a + \varphi_x = \bar{\varphi}$. The second step permits us to interpret this Lagrange multiplier as the shadow value of the informational externality, to prove the existence of an efficient allocation, and to complete its characterization by showing that this multiplier is strictly positive.

Thus consider the first step. Fix some $\bar{\varphi} \in \mathbb{R}$, let $\kappa = K(\bar{\varphi})$, and consider the following constrained problem:

Auxiliary problem 1. *Choose φ so as to maximize $\mathcal{W}(\varphi; \kappa)$ subject to $\varphi_a + \varphi_x = \bar{\varphi}$.*

Note that $\mathcal{W}$ is differentiable in φ for fixed κ. Let $\tilde{\eta}$ denote the Lagrange multiplier for the constraint $\varphi_a + \varphi_x = \bar{\varphi}$. The first-order conditions for this problem are then the following:

$$\begin{aligned}
\varphi_0 &\;:\; 0 = \frac{\partial \mathcal{W}}{\partial \varphi_0}, \\
\varphi_a &\;:\; 0 = \frac{\partial \mathcal{W}}{\partial \varphi_a} + \tilde{\eta}, \\
\varphi_x &\;:\; 0 = \frac{\partial \mathcal{W}}{\partial \varphi_x} + \tilde{\eta}, \\
\varphi_z &\;:\; 0 = \frac{\partial \mathcal{W}}{\partial \varphi_z}.
\end{aligned}$$

Using the characterization of $\mathcal{W}$, the first of these conditions reduces to the following:

$$\varphi_0 \;:\; 0 = \exp V_c(\varphi; \kappa) - \frac{1}{\theta} \exp V_n(\varphi; \kappa). \tag{A.4}$$

This guarantees that $V_c = V_n - \log\theta$ at the efficient allocation and gives φ_0 as a function of $\varphi_a, \varphi_x, \varphi_z, \kappa_x, \kappa_z$ and exogenous parameters. Let $V \equiv V_c = V_n - \log\theta$ and let $\eta \equiv e^{-V}\tilde{\eta}$. The rest of the first-order conditions reduce to the following:

$$
\begin{aligned}
\varphi_a \ : \ 0 \ &= \ \left(\frac{\rho-1}{\rho}\right)\frac{\varphi_a}{\kappa_\xi} + \left(\frac{\rho-1}{\rho}\right)\frac{\varphi_x}{\kappa_x} + (1-\gamma)\frac{(\varphi_a+\varphi_x+\varphi_z)}{\kappa_A} \\
&\quad -\frac{\epsilon}{\theta}\frac{(\varphi_a-1)}{\kappa_\xi} - \frac{\epsilon}{\theta}\frac{\varphi_x}{\kappa_x} - \frac{\epsilon}{\theta}\frac{(\varphi_a+\varphi_x+\varphi_z-1)}{\kappa_A} + \eta, \\
\varphi_x \ : \ 0 \ &= \ \left(\frac{\rho-1}{\rho}\right)\frac{\varphi_x}{\kappa_x} + \left(\frac{\rho-1}{\rho}\right)\frac{\varphi_a}{\kappa_x} + (1-\gamma)\frac{(\varphi_a+\varphi_x+\varphi_z)}{\kappa_A} \\
&\quad -\frac{\epsilon}{\theta}\frac{\varphi_x}{\kappa_x} - \frac{\epsilon}{\theta}\frac{(\varphi_a-1)}{\kappa_x} - \frac{\epsilon}{\theta}\frac{(\varphi_a+\varphi_x+\varphi_z-1)}{\kappa_A} + \eta, \\
\varphi_z \ : \ 0 \ &= \ (1-\gamma)\frac{\varphi_z}{\kappa_z} + (1-\gamma)\frac{(\varphi_a+\varphi_x+\varphi_z)}{\kappa_A} - \frac{\epsilon}{\theta}\frac{\varphi_z}{\kappa_z} - \frac{\epsilon}{\theta}\frac{(\varphi_a+\varphi_x+\varphi_z-1)}{\kappa_A}.
\end{aligned}
$$

For fixed η, this is a linear system of three equations in the three coefficients φ_a, φ_x and φ_z. Subtracting the first equation from the second, we obtain

$$\varphi_a^{**} = \frac{\frac{\epsilon}{\theta}}{\frac{\epsilon}{\theta}+\frac{1}{\rho}-1} \equiv \beta.$$

We can then solve the remaining two equations for φ_x and φ_z as follows:

$$
\begin{aligned}
\varphi_x^{**} \ &= \ \left\{\frac{(1-\alpha)\kappa_x}{(1-\alpha)\kappa_x+\kappa_z+\kappa_A}\right\}\frac{\alpha}{1-\alpha}\beta + \frac{1}{\frac{\epsilon}{\theta}+\frac{1}{\rho}-1}\left\{\frac{\kappa_x(\kappa_z+\kappa_A)}{(1-\alpha)\kappa_x+\kappa_z+\kappa_A}\right\}\eta \\
\varphi_z^{**} \ &= \ \left\{\frac{\kappa_z}{(1-\alpha)\kappa_x+\kappa_z+\kappa_A}\right\}\frac{\alpha}{1-\alpha}\beta - \frac{1}{\frac{\epsilon}{\theta}+\frac{1}{\rho}-1}\left\{\frac{\kappa_x\kappa_z}{(1-\alpha)\kappa_x+\kappa_z+\kappa_A}\right\}\eta
\end{aligned}
$$

Letting

$$\Delta \equiv \frac{1}{\frac{\epsilon}{\theta}+\frac{1}{\rho}-1}\left(\frac{\kappa_x(\kappa_z+\kappa_A)}{(1-\alpha)\kappa_x+\kappa_z+\kappa_A}\right)\eta,$$

gives conditions (5) and (6). Finally, note that Δ is a simple rescaling of the Lagrange multiplier η, so we can think of Δ itself as the relevant Lagrange multiplier. Using then the above results along with the constraint $\varphi_a + \varphi_x = \bar{\varphi}$, we can express Δ (or equivalently η) as follows:

$$\Delta = \bar{\varphi} - \left\{\frac{\kappa_x+\kappa_z+\kappa_A}{(1-\alpha)\kappa_x+\kappa_z+\kappa_A}\right\}\beta. \tag{A.5}$$

Substituting this into conditions (5) and (6), we can obtain the optimal coefficients as functions of the sum $\bar{\varphi}$ and the precisions κ_x and κ_z. Let $\varphi(\bar{\varphi};\kappa)$ denote this solution; for the rest of this proof, whenever we write φ, we mean $\varphi = \varphi(\bar{\varphi};\kappa)$.

We can then express the level of welfare obtained at this solution as a function of the sum $\bar{\varphi}$ and the precisions κ_x and κ_z. In particular, using the FOC with respect to φ_0, we get that

$$\mathcal{W}(\varphi;\kappa) = \left(\frac{\frac{\epsilon}{\theta}-1+\gamma}{1-\gamma}\frac{\theta}{\epsilon}\right)\exp V_c(\varphi;\kappa). \tag{A.6}$$

Since $\frac{\epsilon}{\theta} - 1 + \gamma > 0$ and $\frac{\epsilon}{\theta} > 0$, we can consider the following monotone transformation of welfare:

$$\mathcal{TW}(\varphi; \kappa) \equiv \frac{1}{1-\gamma} V_c(\varphi; \kappa).$$

Using then the characterization of the efficient coefficients, we conclude that

$$\mathcal{TW}(\varphi(\bar{\varphi}; \kappa), \kappa) = W(\bar{\varphi}; \kappa) \equiv A(\kappa) - B(\kappa) \left(\bar{\varphi} - f(\kappa)\right)^2, \tag{A.7}$$

where

$$B(\kappa) \equiv \frac{\epsilon}{2\theta(1-\alpha)} \frac{\kappa_A + (1-\alpha)\kappa_x + \kappa_z}{\kappa_x(\kappa_A + \kappa_z)} > 0,$$

and

$$f(\kappa) \equiv \frac{\kappa_A + \kappa_x + \kappa_z}{(1-\alpha)\kappa_x + \kappa_z + \kappa_A} \beta = \arg\max_{\bar{\varphi}} W(\bar{\varphi}; \kappa) = \arg\max_{\bar{\varphi}} \mathcal{W}(\varphi(\bar{\varphi}; \kappa), \kappa).$$

(The precise value of $A(\kappa)$ has no particular interest, so it is omitted.) This result has a simple interpretation. Note that $f(\kappa)$ identifies the sum $\bar{\varphi} = \varphi_a + \varphi_x$ that would have been efficient had information been exogenous (equivalently, $\varphi(f(\kappa); \kappa)$ are simply the coefficients of the efficient allocation when $\Delta = 0$). Hence, (A.7) expresses welfare as a monotone transformation of the quadratic distance between any value $\bar{\varphi}$ that the planner may choose and the one that would have been optimal from a purely allocative perspective. Clearly, the only reason that the efficient $\bar{\varphi}$ may differ from $f(\kappa)$ is the informational externality.

We now proceed to the second step, namely that of optimizing over the sum $\bar{\varphi} = \varphi_a + \varphi_x$ and the induced precisions $\kappa_x = K_1(\bar{\varphi})$ and $\kappa_z = K_2(\bar{\varphi})$. Letting

$$\bar{W}(\bar{\varphi}) \equiv W\left(\bar{\varphi}; \kappa(\bar{\varphi})\right),$$

the planner's problem reduces to the following uni-dimensional problem:

Auxiliary problem 2. *Choose $\bar{\varphi} \in \mathbb{R}$ so as to maximize $\bar{W}(\bar{\varphi})$.*

First, note that, because $f(\kappa) > 0$, it is necessarily the case that, for any given κ, $W(\bar{\varphi}; \kappa) > W(-\bar{\varphi}; \kappa)$ whenever $\bar{\varphi} > 0$. And because $\kappa(\bar{\varphi}) = \kappa(-\bar{\varphi})$, it is immediate that $\bar{W}(\bar{\varphi}) > \bar{W}(-\bar{\varphi})$ whenever $\bar{\varphi} > 0$, which means that it is never optimal to choose $\bar{\varphi} < 0$.

Next, we can show that

$$\frac{\partial W}{\partial \kappa_z} = \frac{\epsilon}{2\theta\kappa_z^2} \varphi_2(\bar{\varphi}; \kappa)^2 = \frac{\epsilon(\beta - (1-\alpha)\bar{\varphi})^2}{2\theta(1-\alpha)^2(\kappa_A + \kappa_z)^2}.$$

Along with the fact that κ_z is a quadratic function of $\bar{\varphi}$, this guarantees that

$$\frac{\partial W}{\partial \kappa_z} \frac{\partial \kappa_z}{\partial \bar{\varphi}} \to 0 \quad \text{as} \quad \bar{\varphi} \to \infty.$$

In words, the social value of a marginal increase in the precision κ_z of public information vanishes as this precision goes to infinity. A similar result holds for private information:

$$\frac{\partial W}{\partial \kappa_x} = \frac{\epsilon}{2\theta(1-\alpha)\kappa_x^2} \varphi_1(\bar{\varphi}; \kappa)^2 = \frac{\epsilon(\beta - \bar{\varphi})^2}{2\theta(1-\alpha)\kappa_x^2},$$

and hence

$$\frac{\partial W}{\partial \kappa_x}\frac{\partial \kappa_x}{\partial \bar{\varphi}} \to 0 \quad \text{as} \quad \bar{\varphi} \to \infty.$$

At the same time, because

$$\frac{\partial W}{\partial \bar{\varphi}} = -2B(\kappa)\left(\bar{\varphi} - f(\kappa)\right),$$

and because $B(\kappa) \to \frac{\epsilon}{2\theta(1-\alpha)\kappa_x} > 0$ and $f(\kappa) \to \beta$ as $\kappa_z \to \infty$, we have that

$$\frac{\partial W}{\partial \bar{\varphi}} \to -\infty \quad \text{as} \quad \bar{\varphi} \to \infty.$$

Combining, we conclude that

$$\frac{\partial \bar{W}(\bar{\varphi})}{\partial \bar{\varphi}} \to -\infty \quad \text{as} \quad \bar{\varphi} \to \infty.$$

Along with the facts that $\bar{W}(\bar{\varphi})$ is continuous in $\bar{\varphi}$ and that it is without loss of optimality to restrict $\bar{\varphi} \in [0, \infty)$, this guarantees the existence of a solution to auxiliary problem 2 (and hence the existence of an efficient allocation).

Let $\bar{\varphi}^* \geq 0$ denote any such a solution. Since $\bar{W}$ is differentiable, this solution must satisfy $\frac{\partial \bar{W}}{\partial \bar{\varphi}} = 0$. Using the definition of $\bar{W}$, this is equivalent to

$$\frac{\partial W}{\partial \bar{\varphi}} + \frac{\partial W}{\partial \kappa_z}\frac{\partial \kappa_z}{\partial \bar{\varphi}} + \frac{\partial W}{\partial \kappa_x}\frac{\partial \kappa_x}{\partial \bar{\varphi}} = 0. \tag{A.8}$$

Note that the second and the third term are always non-negative. Whenever $0 \leq \bar{\varphi} < f(\kappa)$, the first term is strictly positive, so that the sum is also strictly positive; this rules out $\bar{\varphi}^* \in [0, f(\kappa))$. Moreover, when $\bar{\varphi} = f(\kappa)$, the first term is zero, but now the other two terms are strictly positive, so that the sum is also strictly positive; this rules out $\bar{\varphi}^* = f(\kappa)$. It follows that $\bar{\varphi}^* > f(\kappa)$ necessarily. From (A.5) and the definition of $f(\kappa)$, we have that, at the efficient allocation, $\Delta = \bar{\varphi}^* - f(\kappa)$. It follows that $\Delta > 0$, as claimed in the proposition.

Finally, that Δ (or equivalently η) represents the shadow value of the informational externality follows directly from the envelope condition of auxiliary problem 1, namely $\frac{\partial W}{\partial \bar{\varphi}} = -\eta$, along with the first-order condition of auxiliary problem 2, namely condition (A.8). Indeed, combining these two conditions gives

$$\eta = \frac{\partial W}{\partial \kappa_z}\frac{\partial \kappa_z}{\partial \bar{\varphi}} + \frac{\partial W}{\partial \kappa_x}\frac{\partial \kappa_x}{\partial \bar{\varphi}}.$$

The Lagrange multiplier thereby measures the social value of increasing the precision of available information by increasing the sensitivity of allocations to local information. QED.

Proof of Lemma 3. We consider a combination of the following tax instruments: a linear tax $\tau^R(\bar{a}, z)$ on firm revenue, a linear tax $\tau^L(\bar{a}, z)$ on household labor income, and a linear tax $\tau^C(\bar{a}, z)$ on household consumption (a sales tax that is uniform across commodities). To guarantee the existence of an equilibrium where the allocations are log-normal, these taxes are assumed to be log-linear functions of $(\bar{a}, z)$:

$$\begin{aligned}
-\log(1-\tau^R(\bar{a}, z)) &= \tau_0^R + \tau_A^R \bar{a} + \tau_z^R z, \\
-\log(1-\tau^L(\bar{a}, z)) &= \tau_0^L + \tau_A^L \bar{a} + \tau_z^L z, \\
\log(1+\tau^C(\bar{a}, z)) &= \tau_0^C + \tau_A^C \bar{a} + \tau_z^C z.
\end{aligned}$$

Given these taxes, the firm's realized net-of-tax profits are given by

$$\pi(\omega, \bar{a}, z) = \left(1-\tau^R(\bar{a}, z)\right) p(\omega, \bar{a}, z) q(\omega) - w(\omega) n(\omega),$$

while the budget constraint of the household is given by

$$\begin{aligned}
&(1+\tau^C(\bar{a}, z)) \int p(\omega, \bar{a}, z) c(\omega, \bar{a}, z) dF(\omega|\bar{a}, z) \\
&= \int \pi(\omega, \bar{a}, z) dF(\omega|\bar{a}, z) + (1-\tau^L(\bar{a}, z)) \int w(\omega) n(\omega) dF(\omega|\bar{a}, z) + T(\bar{a}, z)
\end{aligned}$$

where $T(\bar{a}, z)$ is a lump-sum transfer or tax. (By the government budget, the latter is equal to the revenue from all the taxes.) It follows that the optimal labor supply of the typical worker on island ω is given by

$$n(\omega)^{\epsilon-1} = w(\omega) \mathbb{E}\left[(1-\tau^L(\bar{a}, z)) \frac{U'(C(\bar{a}, z))}{(1+\tau^C(\bar{a}, z)) P(\bar{a}, z)} \middle| \omega \right],$$

while the consumer's stochastic discount factor is given by $\frac{U'(Q(\bar{a},z))}{(1+\tau^C(\bar{a},z))P(\bar{a},z)}$. The firm's objective is thus given by

$$\mathbb{E}\left[\frac{U'\left(Q(\bar{a}, z)\right)}{(1+\tau^C(\bar{a}, z)) P(\bar{a}, z)} \left(\left(1-\tau^R(\bar{a}, z)\right) P(\bar{a}, z) Q(\bar{a}, z)^{1/\rho} q(\omega)^{1-1/\rho} - w(\omega) n(\omega) \right) \middle| \omega \right].$$

Taking the FOC for the firm's problem, substituting the equilibrium wage, and guessing that the taxes and the allocations are jointly log-normal (which they are in the equilibrium we construct in the main text), we conclude that the equilibrium level of employment is pinned down by the following condition:

$$n(\omega)^{\epsilon-1} = \left(\frac{\rho-1}{\rho}\right) \mathbb{E}\left[\frac{\chi\left(1-\tau^R(\bar{a}, z)\right)(1-\tau^L(\bar{a}, z))}{1+\tau^C(\bar{a}, z)} U'\left(Q(\bar{a}, z)\right) \left(\frac{q(\omega)}{Q(\bar{a}, z)}\right)^{-\frac{1}{\rho}} \left(\theta A(\omega) n(\omega)^{\theta-1}\right) \middle| \omega \right].$$

where χ is a constant that depends on second-order terms. The result then follows by defining the tax wedge as

$$1-\tau(\bar{a}, z) \equiv \frac{\chi\left(1-\tau^R(\bar{a}, z)\right)(1-\tau^L(\bar{a}, z))}{1+\tau^C(\bar{a}, z)}.$$

and substituting in for $n(\omega)$,

$$n(\omega) = \left(\frac{q(\omega)}{A(\omega)}\right)^{\frac{1}{\theta}}$$

Equivalently, the tax wedge is given by (8) with $\tau_0 \equiv -\log \chi + \tau_0^R + \tau_0^C + \tau_0^L$, $\tau_A \equiv \tau_A^R + \tau_A^C + \tau_A^L$, and $\tau_z \equiv \tau_z^R + \tau_z^C + \tau_z^L$. QED.

The following lemma describes the set of strategies as in (1) that can be implemented by a tax policy as in (8). This is a key step to prove that the optimal allocations defined by (4)-(6) are implementable.

Lemma 7. *Consider any log-linear strategy as in (1). There exists a state-contingent tax policy as in (8) that implements this strategy as an equilibrium strategy, i.e., that satisfies condition (7), if and only if* $\varphi_a = \beta$.

Proof of Lemma 7. By Lemma 3, the equilibrium strategy must solve the following fixed point:

$$q(\omega)^{\frac{\epsilon}{\theta}+\frac{1}{\rho}-1} = \left(\frac{\rho-1}{\rho}\right)\theta A(\omega)^{\frac{\epsilon}{\theta}}\mathbb{E}\left[\exp(-\tau_0 - \tau_A\bar{a} - \tau_z z)Q(\bar{a},z)^{\frac{1}{\rho}-\gamma}|\omega\right] \tag{A.9}$$

with $Q(\bar{a},z)$ given by (2). It follows that the equilibrium strategy is given by

$$\log q(\omega) = \hat{\varphi}_0(\tau) + \hat{\varphi}_a(\tau)\,a + \hat{\varphi}_x(\tau)\,x + \hat{\varphi}_z(\tau)\,z,$$

where

$$\hat{\varphi}_a(\tau) = \beta$$

$$\hat{\varphi}_x(\tau) = \left(1 - \frac{\theta}{\epsilon\alpha}\tau_A\right)\left(\frac{(1-\alpha)\kappa_x}{(1-\alpha)\kappa_x + \kappa_z + \kappa_A}\right)\frac{\alpha}{1-\alpha}\beta \tag{A.10}$$

$$\hat{\varphi}_z(\tau) = \frac{1}{1-\alpha}\left(\frac{\kappa_z}{\kappa_x}\hat{\varphi}_x(\tau) - \frac{\beta\theta}{\epsilon}\tau_z\right) \tag{A.11}$$

$$\begin{aligned}\hat{\varphi}_0(\tau) = \frac{1}{\frac{\epsilon}{\theta}+\gamma-1}\Bigg[&-\tau_0 + \left(-\tau_A + \left(\frac{1}{\rho}-\gamma\right)(\hat{\varphi}_a(\tau)+\hat{\varphi}_x(\tau))\right)\frac{\kappa_A}{\kappa_A+\kappa_x+\kappa_z}\mu \\ &+\left(\frac{1}{\rho}-\gamma\right)\left(\frac{\rho-1}{\rho}\right)\frac{(\hat{\varphi}_a(\tau)+\hat{\varphi}_x(\tau))^2}{2}\sigma_x^2 + \frac{1}{2}\left(\frac{1}{\rho}-\gamma\right)^2(\hat{\varphi}_a(\tau)+\hat{\varphi}_x(\tau))^2\sigma_0^2 \\ &+\frac{1}{2}\tau_A^2\sigma_0^2 - \tau_A\left(\frac{1}{\rho}-\gamma\right)(\hat{\varphi}_a(\tau)+\hat{\varphi}_x(\tau))\sigma_0^2 + \log\left(\theta\frac{\rho-1}{\rho}\right)\Bigg].\end{aligned} \tag{A.12}$$

We now prove the claim in the lemma. Pick any log-linear strategy as in (1) for which $\varphi_a = \beta$ and let $(\varphi_0^{\#}, \varphi_x^{\#}, \varphi_z^{\#})$ denote the remaining coefficients. From condition (A.10), there is a unique value for τ_A that induces $\hat{\varphi}_x(\tau) = \varphi_x^{\#}$; this is given by

$$\tau_A = \frac{\epsilon}{\theta}\left\{\alpha - \varphi_x^{\#}\frac{(1-\alpha)\kappa_x + \kappa_z + \kappa_A}{\beta\kappa_x}\right\}. \tag{A.13}$$

From (A.11), there is then a unique value for τ_z that induces $\hat{\varphi}_z(\tau) = \varphi_z^{\#}$; this is given by

$$\tau_z = \frac{\epsilon}{\beta\theta}\left\{\frac{\kappa_z}{\kappa_x}\varphi_x^{\#} - (1-\alpha)\varphi_z^{\#}\right\}. \tag{A.14}$$

Combining these two results, we have a unique pair (τ_A, τ_z) that induces the desired $(\varphi_x^{\#}, \varphi_z^{\#})$. Finally, from (A.12) there is also a unique τ_0 that induces $\hat{\varphi}_0(\tau) = \varphi_0^{\#}$. QED.

Proof of Proposition 2. From conditions (A.13) and (A.14) in the proof of Lemma 7, the optimal tax satisfies

$$\begin{aligned}\tau_A^* &= \frac{\epsilon}{\theta}\left(\alpha - \varphi_x^* \frac{(1-\alpha)\kappa_x^* + \kappa_z^* + \kappa_A}{\beta\kappa_x^*}\right),\\ \tau_z^* &= \frac{\epsilon}{\beta\theta}\left(\frac{\kappa_z^*}{\kappa_x^*}\varphi_x^* - (1-\alpha)\varphi_z^*\right).\end{aligned}$$

Using the characterization of φ_x^* and φ_z^* from Proposition 1, we get

$$\tau_A^* = -\lambda\Delta \qquad \text{and} \qquad \tau_z^* = \frac{\kappa_z^*}{\kappa_A + \kappa_z^*}\lambda\Delta, \tag{A.15}$$

where

$$\lambda \equiv \frac{\epsilon}{\beta\theta}\frac{(1-\alpha)\kappa_x^* + \kappa_z^* + \kappa_A}{\kappa_x^*} > 0.$$

It follows that $\Delta > 0$ is both necessary and sufficient for each of the following properties: $\tau_A^* < 0$, $\tau_A^* + \tau_z^* < 0$, and $\tau_z^* > 0$.

To interpret this result, note first that $z = \bar{a} + \varepsilon$, where ε is noise. The property that $\tau_A^* < 0$ means the that tax is negatively correlated with aggregate productivity for given common belief z; that is, it is negatively correlated with the surprise component in realized aggregate productivity. At the same time, the property that $\tau_A^* + \tau_z^* < 0$ means that the tax is negatively correlated with aggregate productivity for given noise ε; that is, the overall effect of the productivity shock is also negative. Third, the property that $\tau_z^* > 0$ means that the tax is positively correlated with the noise. Finally, to understand the overall cyclical behavior of the optimal tax, consider the covariance between the (log) tax and (log) output. Since

$$-\log(1-\tau(\bar{a}, z)) = \tau_0^* + (\tau_A^* + \tau_z^*)\bar{a} + \tau_z^*\varepsilon \quad \text{and} \quad \log Q(\bar{a}, z) = \varphi_0^* + (\varphi_a^* + \varphi_x^* + \varphi_z^*)\bar{a} + \varphi_z^*\varepsilon,$$

their covariance is given by

$$Cov(-\log(1-\tau), \log Q) = (\tau_A^* + \tau_z^*)(\varphi_a^* + \varphi_x^* + \varphi_z^*)Var(\bar{a}) + \tau_z^*\varphi_z^* Var(\varepsilon)$$

Using the fact that $Var(\bar{a}) = 1/\kappa_A$ and $Var(\varepsilon) = 1/\kappa_z^*$ and rearranging, we get

$$Cov(-\log(1-\tau), \log Q) = (\tau_A^* + \tau_z^*)(\varphi_a^* + \varphi_x^*)\frac{1}{\kappa_A} + \left\{\tau_A^*\frac{1}{\kappa_A} + \tau_z^*\frac{\kappa_A + \kappa_z^*}{\kappa_A\kappa_z^*}\right\}\varphi_z^*.$$

By (A.15), the last term is necessarily zero. Next, note that $\varphi_a^* + \varphi_x^*$ is necessarily positive, while $\tau_A^* + \tau_z^*$ is necessarily negative. We conclude that the tax is negatively correlated with aggregate output. QED.

A.2 Proofs for Section 4

Proof of Lemma 4. From the consumer's optimal demand, we have that the (shadow) prices must satisfy

$$-\rho\left(\log p(\omega)-\log P(\bar{a},z)\right)=\left(\log c(\omega,\bar{a},z)-\log C(\bar{a},z)\right),$$

where

$$\begin{aligned}
\log p(\omega) &= const+\psi_a a+\psi_x x+\psi_z z,\\
\log P(\bar{a},z) &= const+\psi_a \bar{a}+\psi_x \bar{a}+\psi_z z,\\
\log c(\omega,\bar{a},z) &= const+(\varphi_a+(1-\theta)l_a)a+(\varphi_x+(1-\theta)l_x)x+(\varphi_z+(1-\theta)l_z)z+(1-\theta)l_A\bar{a},\\
\log C(\bar{a},z) &= const+(\varphi_a+(1-\theta)l_a)\bar{a}+(\varphi_x+(1-\theta)l_x)\bar{a}+(\varphi_z+(1-\theta)l_z)z+(1-\theta)l_A\bar{a}.
\end{aligned}$$

It follows that the following must hold for all $(a,x,z,\bar{a})$:

$$-\rho(\psi_a a+\psi_x x-(\psi_a+\psi_x)\bar{a})=(\varphi_a+(1-\theta)l_a)a+(\varphi_x+(1-\theta)l_x)x-(\varphi_a+(1-\theta)l_a+\varphi_x+(1-\theta)l_x)\bar{a},$$

which is true if and only if

$$\psi_a=-\frac{1}{\rho}(\varphi_a+(1-\theta)l_a), \tag{A.16}$$

$$\psi_x=-\frac{1}{\rho}(\varphi_x+(1-\theta)l_x). \tag{A.17}$$

Finally, note that the observation of $z^p=\log P(\bar{a},z)+\varepsilon^p$ is equivalent to the observation of the unbiased Gaussian signal

$$\tilde{z}^p\equiv\frac{z^p-const-\psi_z z}{\psi_a+\psi_x}=\bar{a}+\tilde{\varepsilon}_p,$$

where $\tilde{\varepsilon}_p=\varepsilon_p/(\psi_a+\psi_x)$. We conclude that the precision of the public price signal is given by

$$\kappa_{zp}\equiv(\psi_a+\psi_x)^2\sigma_{zp}^{-2}=\frac{1}{\rho^2}(\varphi_a+\varphi_x+(1-\theta)(l_a+l_x))^2\sigma_{zp}^{-2}.$$

The proof for the public signal on $\log Y(\bar{a},z)$ is given in the main text. We thus have

$$\kappa_{zy}\equiv(\varphi_a+\varphi_x+(1-\theta)(l_a+l_x+l_A))^2\sigma_{zy}^{-2},$$

which, together with $\kappa_z=\sigma_{za}^{-2}+\kappa_{zy}+\kappa_{zp}$, gives the expression for κ_z in (13). Analogous arguments give the expression for κ_x. QED.

Proof of Proposition 3. In equilibrium, $l(\omega,\bar{a},z)$ adjusts in stage 2 so as to satisfy the the consumer's demand:

$$\frac{p\left(\omega\right)}{P(\bar{a},z)}=\left(\frac{q\left(\omega\right)l\left(\omega,\bar{a},z\right)^{1-\theta}}{C(\bar{a},z)}\right)^{-\frac{1}{\rho}} \tag{A.18}$$

Solving for $l(\omega,\bar a,z)$ and substituting into the firm's objective, the latter reduces to the following:

$$\mathbb{E}\left[\frac{U'\left(C(\bar a,z)\right)}{P(\bar a,z)}\left((1-\tau(\bar a,z))C(\bar a,z)p^{1-\rho}P(\bar a,z)^{\rho}-w_2(\omega)\left(\frac{p}{P(\bar a,z)}\right)^{-\frac{\rho}{1-\theta}}\left(\frac{C(\bar a,z)}{q}\right)^{\frac{1}{1-\theta}}-w_1(\omega)\left(\frac{q}{e^{a}}\right)^{\frac{1}{\theta}}\right)\middle|\,\omega\right].$$

Note that this objective is strictly concave in p and q which guarantees that the FOCs are both necessary and sufficient and that they uniquely pin down the solution to the firm's problem for given wages. Next, note that the equilibrium wages satisfy

$$n(\omega)^{\epsilon-1}=w_1(\omega)\mathbb{E}\left[\frac{U'(C(\bar a,z))}{P(\bar a,z)}\middle|\,\omega\right]\quad\text{and}\quad l(\omega)^{\epsilon-1}=w_2(\omega)\mathbb{E}\left[\frac{U'(C(\bar a,z))}{P(\bar a,z)}\middle|\,\omega\right].$$

Solving these conditions for wages $w_1(\omega)$ and $w_2(\omega)$ and substituting the solutions into the first-order conditions for the firm's problem gives us the following two conditions for the equilibrium price and production choices taken in stage 1:

$$p\,(\omega)^{1-\rho+\frac{\rho\epsilon}{1-\theta}}=\frac{\mathbb{E}\left[P(\bar a,z)^{\frac{\rho\epsilon}{1-\theta}}C(\bar a,z)^{\frac{\epsilon}{1-\theta}}q(\omega)^{-\frac{\epsilon}{1-\theta}}\middle|\,\omega\right]}{(1-\theta)\,\mathbb{E}\left[\left(\frac{\rho-1}{\rho}\right)(1-\tau(\bar a,z))C(\bar a,z)^{1-\gamma}P(\bar a,z)^{\rho-1}\middle|\,\omega\right]},\tag{A.19}$$

$$q(\omega)^{\frac{\epsilon}{\theta}}=p\,(\omega)^{1-\rho}\,e^{\frac{\epsilon}{\theta}a}\theta\mathbb{E}\left[\left(\frac{\rho-1}{\rho}\right)(1-\tau(\bar a,z))C(\bar a,z)^{1-\gamma}P(\bar a,z)^{\rho-1}\middle|\,\omega\right].\tag{A.20}$$

Using (A.18), we can restate these conditions in terms of allocations alone as follows:

$$0=n(\omega)^{\epsilon-1}-\mathbb{E}\left[\left(\frac{\rho-1}{\rho}\right)(1-\tau(\bar a,z))U'\left(C(\bar a,z)\right)\left(\frac{c(\omega,\bar a,z)}{C(\bar a,z)}\right)^{-\frac{1}{\rho}}\left(\theta\frac{c(\omega,\bar a,z)}{n(\omega)}\right)\middle|\,\omega\right],$$

$$0=\mathbb{E}\left[l(\omega,\bar a,z)\left\{l(\omega,\bar a,z)^{\epsilon-1}-\left(\frac{\rho-1}{\rho}\right)(1-\tau(\bar a,z))U'\left(C(\bar a,z)\right)\left(\frac{c(\omega,\bar a,z)}{C(\bar a,z)}\right)^{-\frac{1}{\rho}}\left((1-\theta)\frac{c(\omega,\bar a,z)}{l(\omega,\bar a,z)}\right)\right\}\middle|\,\omega\right]$$

Rearranging these conditions, we get

$$\mathbb{E}\left[l\left(\omega,\bar a,z\right)^{\epsilon}\middle|\,\omega\right]=\frac{1-\theta}{\theta}e^{-\frac{\epsilon}{\theta}a}q(\omega)^{\frac{\epsilon}{\theta}},\tag{A.21}$$

$$q(\omega)^{\frac{\epsilon}{\theta}+\frac{1}{\rho}-1}=e^{\frac{\epsilon}{\theta}a}\theta\left(\frac{\rho-1}{\rho}\right)\mathbb{E}\left[(1-\tau(\bar a,z))C(\bar a,z)^{\frac{1}{\rho}-\gamma}l\left(\omega,\bar a,z\right)^{(1-\theta)\left(\frac{\rho-1}{\rho}\right)}\middle|\,\omega\right].\tag{A.22}$$

The first condition equates the (expected) marginal rates of transformation and substitution between l and n. We conclude that a set of allocations, prices and policies constitute an equilibrium if and only if the following hold: (i) the allocations and the tax policy satisfy conditions (A.21) and (A.22), along with the resource constraint

$$C(\bar a,z)=\left[\int\left(q(\omega)l(\omega,\bar a,z)^{1-\theta}\right)^{\frac{\rho-1}{\rho}}dF(\omega|\bar a,z)\right]^{\frac{\rho}{\rho-1}};\tag{A.23}$$

(ii) the nominal prices satisfy condition (A.18); and (iii) the interest-rate rule satisfies the Euler condition

$$C(\bar a,z)^{-\gamma}=\bar\beta\left(1+R(\bar a,z)\right)P(\bar a,z)\mathbb{E}\left[\frac{C(\bar a_{+1},z_{+1})^{-\gamma}}{P(\bar a_{+1},z_{+1})}\middle|\,\bar a,z\right].\tag{A.24}$$

We now seek to translate conditions (A.18)-(A.24) in terms of the relevant coefficients that parameterize the allocations, prices and policy under a log-normal specification. Thus let:

$$\begin{aligned}
\log q(\omega) &= const + \varphi_a a + \varphi_x x + \varphi_z z, \\
\log l\left(\omega, \bar{a}, z\right) &= const + l_A \bar{a} + l_a a + l_x x + l_z z, \\
\log C(\bar{a}, z) = \log Y(\bar{a}, z) &= const + c_A \bar{a} + c_z z, \\
\log p(\omega) &= const + \psi_a a + \psi_x x + \psi_z z, \\
\log(1 - \tau(\bar{a}, z)) &= const - \tau_A \bar{a} - \tau_z z, \\
\log\left(1 + R(\bar{a}, z)\right) &= const + \rho_A \bar{a} + \rho_z z,
\end{aligned}$$

for some coefficients $(\varphi_a, \varphi_x, ..., \rho_A, \rho_z)$.

The resource constraint (A.23) is satisfied if and only if

$$c_A = (\varphi_a + \varphi_x) + (1 - \theta)(l_a + l_x + l_A), \tag{A.25}$$

$$c_z = \varphi_z + (1 - \theta) l_z. \tag{A.26}$$

Also, the interest rate is pinned down by the Euler equation (A.24). Taking the logs of both sides and using the fact that expectations about future variables are constant by the assumption of i.i.d. shocks, we can restate condition (A.24) as

$$\log\left(1 + R(\bar{a}, z)\right) = const - \gamma \log Y(\bar{a}, z) - \log P(\bar{a}, z).$$

Using the log-normal specification above, the latter condition is satisfied if and only if

$$\rho_A = -\gamma c_A - (\psi_a + \psi_x), \quad \text{and} \tag{A.27}$$

$$\rho_z = -\gamma c_z - \psi_z. \tag{A.28}$$

Next, we can rewrite the consumer's demand function as

$$-\rho\left(\log p(\omega) - \log P(\bar{a}, z)\right) = \left(\log c(\omega, \bar{a}, z) - \log C(\bar{a}, z)\right),$$

where

$$\begin{aligned}
\log c(\omega, \bar{a}, z) &= \log q(\omega) + (1 - \theta) \log l(\omega, \bar{a}, z) \\
&= const + (\varphi_a + (1 - \theta) l_a) a + (\varphi_x + (1 - \theta) l_x) x + (\varphi_z + (1 - \theta) l_z) z + (1 - \theta) l_A \bar{a}.
\end{aligned}$$

It follows that the following must hold for all $(a, x, z, \bar{a})$:

$$-\rho(\psi_a a + \psi_x x - (\psi_a + \psi_x)\bar{a}) = (\varphi_a + (1-\theta) l_a) a + (\varphi_x + (1-\theta) l_x) x - (\varphi_a + (1-\theta) l_a + \varphi_x + (1-\theta) l_x)\bar{a}.$$

This is true if and only if

$$\psi_a = -\frac{1}{\rho}(\varphi_a + (1 - \theta) l_a) \qquad \text{and} \qquad \psi_x = -\frac{1}{\rho}(\varphi_x + (1 - \theta) l_x).$$

Finally, note that conditions (A.21) and (A.22) may be rewritten as follows:

$$\mathbb{E}[\log l\,(\omega,\bar{a},z)\,|\omega] = const + \frac{1}{\theta}(\log q(\omega) - a), \qquad \text{and} \tag{A.29}$$

$$\log q(\omega) = const + \beta a - k(\tau_A \mathbb{E}[\bar{a}|\omega] + \tau_z z) + \frac{\alpha}{\chi}\mathbb{E}[\log C(\bar{a},z)|\omega], \tag{A.30}$$

where

$$\beta \equiv \frac{\frac{\epsilon}{\theta}}{\frac{\epsilon}{\theta} - (\rho-1)\nu} > 1, \quad \alpha \equiv \left(\tfrac{1}{\rho} - \gamma\right)\frac{\rho\nu\chi}{\frac{\epsilon}{\theta} - (\rho-1)\nu},$$

$$\nu \equiv \frac{\epsilon}{\rho(\epsilon - 1 + \theta) + 1 - \theta} > \frac{1}{\rho}, \qquad \chi \equiv \frac{\epsilon}{(\epsilon - 1 + \theta) + \gamma(1-\theta)} > 0, \qquad k \equiv \nu\rho\frac{\theta}{\epsilon}\beta > 0.$$

Clearly, condition (A.29) holds for all ω if and only if

$$l_a = \frac{1}{\theta}(\varphi_a - 1), \tag{A.31}$$

$$l_x = \frac{1}{\theta}\varphi_x - l_A\frac{\kappa_x}{\kappa_A + \kappa_x + \kappa_z}, \tag{A.32}$$

$$l_z = \frac{1}{\theta}\varphi_z - l_A\frac{\kappa_z}{\kappa_A + \kappa_x + \kappa_z}, \tag{A.33}$$

while condition (A.30) holds for all ω if and only if

$$\varphi_a = \beta, \tag{A.34}$$

$$\varphi_x = -k\tau_A\frac{\kappa_x}{\kappa_A + \kappa_x + \kappa_z} + \frac{\alpha}{\chi}c_A\frac{\kappa_x}{\kappa_A + \kappa_x + \kappa_z}, \tag{A.35}$$

$$\varphi_z = -k\left(\tau_A\frac{\kappa_z}{\kappa_A + \kappa_x + \kappa_z} + \tau_z\right) + \frac{\alpha}{\chi}\left(c_A\frac{\kappa_z}{\kappa_A + \kappa_x + \kappa_z} + c_z\right), \tag{A.36}$$

where c_A and c_z are given by (A.25)-(A.26).

Note that conditions (A.31) through (A.34) give the implementability constraints stated in the proposition, completing the proof of the necessity of these conditions for an allocation to be part of an equilibrium. We next prove sufficiency.

Pick arbitrary $(\varphi_x, \varphi_z, l_A)$ and let $(\varphi_a, l_a, l_x, l_z)$ satisfy conditions (A.31) through (A.34). Note that there is a unique $(\varphi_a, l_a, l_x, l_z)$ that has this property for any given $(\varphi_x, \varphi_z, l_A)$. Next, pick an arbitrary ψ_z and let $(c_A, c_z, \psi_a, \psi_x)$ be determined as in (A.25)-(A.17). Next, let (τ_A, τ_z) be the unique solution to (A.35)-(A.36); for future reference, this solution is given by

$$\tau_A = \frac{1}{\chi k}\left\{\alpha c_A - \chi\frac{\kappa}{\kappa_x}\varphi_x\right\} \tag{A.37}$$

$$\tau_z = \frac{1}{\chi k}\left\{\alpha c_z - \chi\left(\varphi_z - \varphi_x\frac{\kappa_z}{\kappa_x}\right)\right\} \tag{A.38}$$

where $\chi k > 0$. Finally, set (ρ_A, ρ_z) as in (A.27)-(A.28). By construction, the allocations, prices and policies defined in this way constitute an equilibrium, which completes the sufficiency argument.

Part (ii). The proof of this part is similar to that of part (i), except for one key difference: now the marginal costs and returns of stage-2 employment must be equated state-by-state, not just in expectation. In particular, we would have

$$n(\omega)^{\epsilon-1} = \left(\frac{\rho-1}{\rho}\right)\mathbb{E}\left[(1-\tau(\bar{a},z))U'\left(C(\bar{a},z)\right)\left(\frac{c(\omega,\bar{a},z)}{C(\bar{a},z)}\right)^{-\frac{1}{\rho}}\left(\theta\frac{c(\omega,\bar{a},z)}{n(\omega)}\right)\middle|\,\omega\right] \quad \text{(A.39)}$$

$$l(\omega,\bar{a},z)^{\epsilon-1} = \left(\frac{\rho-1}{\rho}\right)(1-\tau(\bar{a},z))U'\left(C(\bar{a},z)\right)\left(\frac{c(\omega,\bar{a},z)}{C(\bar{a},z)}\right)^{-\frac{1}{\rho}}\left((1-\theta)\frac{c(\omega,\bar{a},z)}{l(\omega,\bar{a},z)}\right) \quad \text{(A.40)}$$

It is this additional restriction that pins down l_A. (A detailed derivation is available upon request.) QED.

Proof of Lemma 5. The proof is contained in the proof of Proposition 3. In particular, note that the argument is silent about the value for ψ_z, which is thereby undetermined. QED.

Proof of Lemma 6. Take any pair of strategies as in (11). Welfare (ex-ante utility) may be written as

$$\mathcal{W}(\varphi,l;\kappa_x,\kappa_z) = \frac{1}{1-\gamma}\exp V_c(\varphi,l,\kappa) - \frac{1}{\epsilon}\exp V_l(\varphi,l,\kappa) - \frac{1}{\epsilon}\exp V_n(\varphi,l,\kappa),$$

where $\varphi = (\varphi_a,\varphi_x,\varphi_z)$, $l = (l_a,l_x,l_z,l_A)$, $\kappa = (\kappa_x,\kappa_z)$, and where

$$\begin{aligned}
V_c(\varphi,l,\kappa) \equiv\ & (1-\gamma)\left(\varphi_0 + (1-\theta)l_0 + \left[(1-\theta)l_A + (\varphi_a+(1-\theta)l_a) + (\varphi_x+(1-\theta)l_x) + (\varphi_z+(1-\theta)l_z)\right]\mu\right) \\
& + \frac{1}{2}(1-\gamma)\left(\frac{\rho-1}{\rho}\right)\left[\frac{(\varphi_a+(1-\theta)l_a)^2}{\kappa_\xi} + \frac{(\varphi_x+(1-\theta)l_x)^2}{\kappa_x} + 2\frac{(\varphi_a+(1-\theta)l_a)(\varphi_x+(1-\theta)l_x)}{\kappa_x}\right] \\
& + \frac{1}{2}(1-\gamma)^2\left[\frac{(\varphi_z+(1-\theta)l_z)^2}{\kappa_z} + \frac{((1-\theta)l_A + (\varphi_a+(1-\theta)l_a) + (\varphi_x+(1-\theta)l_x) + (\varphi_z+(1-\theta)l_z))^2}{\kappa_A}\right], \\
V_l(\varphi,l,\kappa) \equiv\ & \epsilon\left(l_0 + (l_A+l_a+l_x+l_z)\mu\right) + \frac{1}{2}\epsilon^2\left[\left(\frac{l_a^2}{\kappa_\xi} + \frac{l_x^2}{\kappa_x} + 2\frac{l_a l_x}{\kappa_x}\right) + \frac{l_z^2}{\kappa_z} + \frac{(l_A+l_a+l_x+l_z)^2}{\kappa_A}\right], \\
V_n(\varphi,l,\kappa) \equiv\ & \frac{\epsilon}{\theta}\left(\varphi_0 + (\varphi_a+\varphi_x+\varphi_z-1)\mu\right) \\
& + \frac{1}{2}\frac{\epsilon^2}{\theta^2}\left[\frac{(\varphi_a-1)^2}{\kappa_\xi} + \frac{\varphi_x^2}{\kappa_x} + 2\frac{(\varphi_a-1)\varphi_x}{\kappa_x} + \frac{\varphi_z^2}{\kappa_z} + \frac{(\varphi_a+\varphi_x+\varphi_z-1)^2}{\kappa_A}\right].
\end{aligned}$$

QED.

Proof of Proposition 4. We henceforth consider a relaxed problem, where we ignore the constraint on φ_a imposed by (15); it will turn out that the solution to this relaxed problem satisfies this constraint, which means that the solution to the relaxed problem is also the solution to our initial problem.

The first-order conditions of the (relaxed) problem with respect to φ_0 and l_0 give

$$\begin{aligned}
\varphi_0 &: \quad 0 = \exp V_c - \frac{1}{\theta}\exp V_n \\
l_0 &: \quad 0 = (1-\theta)\exp V_c - \exp V_l.
\end{aligned}$$

Hence, at the optimal allocation, $\exp V \equiv \exp V_c = \frac{1}{\theta}\exp V_n = \frac{1}{1-\theta}\exp V_l > 0$. Let the Lagrange multipliers on the implementability constraints (16)-(18) be, respectively, $e^V\mu_a$, $e^V\mu_x$, and $e^V\mu_z$.

Next, as in the proof of Proposition 1, we can represent the informational externalities by two Lagrange multipliers, one for the sum $\varphi_a+\varphi_x+(1-\theta)(l_a+l_x+l_A)$ which determines the precision of the output signals, κ_{zy} and κ_{xy}; and another for the sum $\varphi_a + \varphi_x + (1-\theta)(l_a + l_x)$ which determines the precision of the price signals, κ_{zp} and κ_{xp}. Let these multipliers be, respectively, $e^V\eta_Y$ and $e^V\eta_p$. We can then state the rest of the first-order conditions of the optimal policy problem as follows.

The first-order conditions for the stage-1 strategy are the following:

$$
\begin{aligned}
\varphi_a:\quad 0 \quad &= \quad \left(\frac{\rho-1}{\rho}\right)\left(\frac{(\varphi_a+(1-\theta)l_a)}{\kappa_\xi}+\frac{(\varphi_x+(1-\theta)l_x)}{\kappa_x}\right)\\
&\quad +(1-\gamma)\frac{((1-\theta)l_A+(\varphi_a+(1-\theta)l_a)+(\varphi_x+(1-\theta)l_x)+(\varphi_z+(1-\theta)l_z))}{\kappa_A}\\
&\quad -\frac{\epsilon}{\theta}\left[\frac{(\varphi_a-1)}{\kappa_\xi}+\frac{\varphi_x}{\kappa_x}+\frac{(\varphi_a+\varphi_x+\varphi_z-1)}{\kappa_A}\right]+\eta_Y+\eta_p-\frac{1}{\theta}\mu_a,\\
\varphi_x:\quad 0 \quad &= \quad \left(\frac{\rho-1}{\rho}\right)\left(\frac{(\varphi_x+(1-\theta)l_x)}{\kappa_x}+\frac{(\varphi_a+(1-\theta)l_a)}{\kappa_x}\right)\\
&\quad +(1-\gamma)\frac{((1-\theta)l_A+(\varphi_a+(1-\theta)l_a)+(\varphi_x+(1-\theta)l_x)+(\varphi_z+(1-\theta)l_z))}{\kappa_A}\\
&\quad -\frac{\epsilon}{\theta}\left[\frac{\varphi_x}{\kappa_x}+\frac{(\varphi_a-1)}{\kappa_x}+\frac{(\varphi_a+\varphi_x+\varphi_z-1)}{\kappa_A}\right]+\eta_Y+\eta_p-\frac{1}{\theta}\mu_x,\\
\varphi_z:\quad 0 \quad &= \quad (1-\gamma)\left[\frac{(\varphi_z+(1-\theta)l_z)}{\kappa_z}+\frac{((1-\theta)l_A+(\varphi_a+(1-\theta)l_a)+(\varphi_x+(1-\theta)l_x)+(\varphi_z+(1-\theta)l_z))}{\kappa_A}\right]\\
&\quad -\frac{\epsilon}{\theta}\left[\frac{\varphi_z}{\kappa_z}+\frac{(\varphi_a+\varphi_x+\varphi_z-1)}{\kappa_A}\right]-\frac{1}{\theta}\mu_z;
\end{aligned}
$$

and the first-order conditions for the stage-2 strategy are the following:

$$
\begin{aligned}
l_a:\quad 0 \quad &= \quad \left(\frac{\rho-1}{\rho}\right)\left(\frac{(\varphi_a+(1-\theta)l_a)}{\kappa_\xi}+\frac{(\varphi_x+(1-\theta)l_x)}{\kappa_x}\right)\\
&\quad +(1-\gamma)\left[\frac{((1-\theta)l_A+(\varphi_a+(1-\theta)l_a)+(\varphi_x+(1-\theta)l_x)+(\varphi_z+(1-\theta)l_z))}{\kappa_A}\right]\\
&\quad -\epsilon\left[\frac{l_a}{\kappa_\xi}+\frac{l_x}{\kappa_x}+\frac{(l_A+l_a+l_x+l_z)}{\kappa_A}\right]+\eta_Y+\eta_p+\frac{\mu_a}{1-\theta},\\
l_x:\quad 0 \quad &= \quad \left(\frac{\rho-1}{\rho}\right)\left(\frac{(\varphi_x+(1-\theta)l_x)}{\kappa_x}+\frac{(\varphi_a+(1-\theta)l_a)}{\kappa_x}\right)\\
&\quad +(1-\gamma)\left[\frac{((1-\theta)l_A+(\varphi_a+(1-\theta)l_a)+(\varphi_x+(1-\theta)l_x)+(\varphi_z+(1-\theta)l_z))}{\kappa_A}\right]\\
&\quad -\epsilon\left[\frac{l_x}{\kappa_x}+\frac{l_a}{\kappa_x}+\frac{(l_A+l_a+l_x+l_z)}{\kappa_A}\right]+\eta_Y+\eta_p+\frac{\mu_x}{1-\theta},\\
l_z:\quad 0 \quad &= \quad (1-\gamma)\left[\frac{(\varphi_z+(1-\theta)l_z)}{\kappa_z}+\frac{((1-\theta)l_A+(\varphi_a+(1-\theta)l_a)+(\varphi_x+(1-\theta)l_x)+(\varphi_z+(1-\theta)l_z))}{\kappa_A}\right]\\
&\quad -\epsilon\left[\frac{l_z}{\kappa_z}+\frac{(l_A+l_a+l_x+l_z)}{\kappa_A}\right]+\frac{\mu_z}{1-\theta},\\
l_A:\quad 0 \quad &= \quad (1-\gamma)\left[\frac{((1-\theta)l_A+(\varphi_a+(1-\theta)l_a)+(\varphi_x+(1-\theta)l_x)+(\varphi_z+(1-\theta)l_z))}{\kappa_A}\right]-\epsilon\left[\frac{(l_A+l_a+l_x+l_z)}{\kappa_A}\right]\\
&\quad +\eta_Y+\frac{\kappa_x}{\kappa_A+\kappa_x+\kappa_z}\frac{\mu_x}{1-\theta}+\frac{\kappa_z}{\kappa_A+\kappa_x+\kappa_z}\frac{\mu_z}{1-\theta}.
\end{aligned}
$$

For any given η_Y and η_p, the combination of these seven FOCs with the three implementability constraints (16)-(18) defines a linear system of 10 equations in 10 unknowns, the allocation coefficients $(\varphi_a, \varphi_x, \varphi_x)$ and (l_A, l_a, l_x, l_z) and the implementability multipliers (μ_a, μ_x, μ_z). The solution to this system gives the following results. For the stage-1 allocation,

$$\varphi_a^* = \beta, \tag{A.41}$$

$$\varphi_x^* = \left\{\frac{(1-\alpha)\kappa_x^*}{(1-\alpha)\kappa_x^* + \kappa_z^* + \kappa_A}\right\}\frac{\alpha}{1-\alpha}\beta + \delta_Y\eta_Y + \delta_p\eta_p, \tag{A.42}$$

$$\varphi_z^* = \left\{\frac{\kappa_z^*}{(1-\alpha)\kappa_x^* + \kappa_z^* + \kappa_A}\right\}\frac{\alpha}{1-\alpha}\beta - \frac{\kappa_z^*}{\kappa_A + \kappa_z^*}\left(\delta_Y\eta_Y + \delta_p\eta_p\right), \tag{A.43}$$

where

$$\delta_Y \equiv \frac{(1-\alpha)\theta\kappa_x^*(\kappa_A + \kappa_z^*)}{(\gamma+\epsilon-1)((1-\alpha)\kappa_x^* + \kappa_z^* + \kappa_A)} > 0, \quad \text{and}$$

$$\delta_p \equiv \frac{(\epsilon\theta - \alpha(\gamma+\epsilon-1+\theta(1-\gamma)))\kappa_x^*(\kappa_A + \kappa_z^*)}{\epsilon(\gamma+\epsilon-1)((1-\alpha)\kappa_x^* + \kappa_z^* + \kappa_A)} > 0.$$

For the stage-2 allocation,

$$\begin{aligned} l_A^* &= \hat{l}_A - \lambda\delta_p\eta_p, \\ l_a^* &= \hat{l}_a, \\ l_x^* &= \hat{l}_x + \left(\frac{\kappa_x^*}{\kappa_A + \kappa_x^* + \kappa_z^*}\right)\lambda\delta_p\eta_p, \\ l_z^* &= \hat{l}_z + \left(\frac{\kappa_z^*}{\kappa_A + \kappa_x^* + \kappa_z^*}\right)\lambda\delta_p\eta_p, \end{aligned}$$

where

$$\begin{aligned} \hat{l}_A &\equiv \frac{(\kappa_A + \kappa_x^* + \kappa_z^*)\varphi_x^*}{\beta\kappa_x^*\theta}, \\ \hat{l}_a &\equiv \frac{1}{\theta}(\varphi_a^* - 1), \\ \hat{l}_x &\equiv \frac{1}{\theta}\varphi_x^* - \hat{l}_A\frac{\kappa_x^*}{\kappa_A + \kappa_x^* + \kappa_z^*}, \\ \hat{l}_z &\equiv \frac{1}{\theta}\varphi_z^* - \hat{l}_A\frac{\kappa_z^*}{\kappa_A + \kappa_x^* + \kappa_z^*}, \end{aligned}$$

where

$$\lambda \equiv \frac{(\gamma+\epsilon-1)(\kappa_A + \kappa_x^* + \kappa_z^*)(\kappa_A + (1-\alpha)\kappa_x^* + \kappa_z^*)}{(1-\alpha)\theta(\gamma+\epsilon-1+\theta(1-\gamma))\kappa_x^*(\kappa_A + \kappa_z^*)} > 0.$$

Note that, by Proposition 3, $(\hat{l}_A, \hat{l}_a, \hat{l}_x, \hat{l}_z)$ identifies the stage-2 allocation that would obtain in the (unique) flexible-price equilibrium in which the stage-1 allocation is given by (A.41)-(A.43). Finally, for the implementability multipliers,

$$\mu_a = \mu_x = \mu_z = 0.$$

Letting

$$\Delta_Y \equiv \eta_Y\delta_Y \qquad \text{and} \qquad \Delta_p \equiv \eta_p\delta_p$$

completes the proof of all the conditions in the proposition.

What remains to be shown is that Δ_Y and Δ_p (or, equivalently, η_Y and η_p) are strictly positive. In the remainder of the proof, to simplify expressions we focus on the case in which endogenous information is only public, that is, $\kappa_{xy} = \kappa_{xp} = 0$. Note that

$$\begin{aligned}\eta_Y &= e^{-V} \frac{\partial \mathcal{W}}{\partial \kappa_z} \frac{\partial \kappa_z}{\partial \kappa_{zy}} \frac{\partial \kappa_{zy}}{\partial \varphi_x}, \\ \eta_p &= e^{-V} \frac{\partial \mathcal{W}}{\partial \kappa_z} \frac{\partial \kappa_z}{\partial \kappa_{zp}} \frac{\partial \kappa_{zp}}{\partial \varphi_x}.\end{aligned}$$

Next, note that $\frac{\partial \kappa_z}{\partial \kappa_{zy}} = \frac{\partial \kappa_z}{\partial \kappa_{zp}} = 1$. Also,

$$\frac{\partial \mathcal{W}}{\partial \kappa_z} = -(1-\gamma) \exp V_c \frac{(\varphi_z + (1-\theta) l_z)^2}{\kappa_z^2} + \frac{\epsilon}{(1-\theta)^2} \exp V_l \frac{((1-\theta) l_z)^2}{\kappa_z^2} + \frac{\epsilon}{\theta^2} \exp V_n \frac{\varphi_z^2}{\kappa_z^2}. \tag{A.44}$$

At the optimal allocation, we know that $\exp V \equiv \exp V_c = \frac{1}{\theta} \exp V_n = \frac{1}{1-\theta} \exp V_l$, as well as that $\mu_a = \mu_x = \mu_z = 0$. Using the first set of equalities, we obtain:

$$\frac{\partial \mathcal{W}}{\partial \kappa_z} = \frac{e^V}{\kappa_z^2} \left\{ (\gamma - 1) (\varphi_z + (1-\theta) l_z)^2 + \frac{\epsilon}{1-\theta} ((1-\theta) l_z)^2 + \frac{\epsilon}{\theta} \varphi_z^2 \right\}.$$

Using the second set of equalities, along with the FOCs with respect to (l_a, l_x, l_z, l_A), we can express l_z as a function of φ_z and η_Y as follows:

$$l_z = \frac{(1-\gamma)\varphi_z - \kappa_z \eta_Y}{(\epsilon - 1 + \theta) + \gamma(1-\theta)}.$$

It follows that

$$\frac{\partial \mathcal{W}}{\partial \kappa_z} = e^V \frac{1}{\beta} (1-\alpha) \left\{ \frac{\epsilon}{\theta} \frac{\varphi_z^2}{\kappa_z^2} + \frac{1-\theta}{\epsilon - 1 + \gamma} \eta_Y^2 \right\} > 0.$$

Finally, recall that $\frac{\partial \kappa_{zy}}{\partial \varphi_x} > 0$ if and only if $\varphi_a + \varphi_x + (1-\theta)(l_a + l_x + l_A) > 0$, while $\frac{\partial \kappa_{zp}}{\partial \varphi_x} > 0$ if and only if $\varphi_a + \varphi_x + (1-\theta)(l_a + l_x) > 0$. Combining these results, we conclude that $\Delta_Y > 0$ and $\Delta_p > 0$ if and only if the optimal allocation satisfies $\varphi_a + \varphi_x + (1-\theta)(l_a + l_x + l_A) > 0$ and $\varphi_a + \varphi_x + (1-\theta)(l_a + l_x) > 0$.

To prove that these inequalities are satisfied, we proceed in a fashion similar to the proof of Proposition 1. Let us define the 2-by-1 vectors

$$\bar{v} \equiv \begin{bmatrix} \frac{1}{\rho} [\varphi_a + \varphi_x + (1-\theta)(l_a + l_x)] \\ \varphi_a + \varphi_x + (1-\theta)(l_a + l_x + l_A) \end{bmatrix} \qquad \text{and} \qquad v(\kappa) \equiv \begin{bmatrix} \frac{1}{\theta\rho} \frac{(\beta - 1 + \theta)(\kappa_A + \kappa_x + \kappa_z)}{(1-\alpha)\kappa_x + \kappa_z + \kappa_A} \\ \frac{1}{\theta} \frac{(\beta - (1-\theta)(1-\alpha))(\kappa_A + \kappa_x + \kappa_z)}{(1-\alpha)\kappa_x + \kappa_z + \kappa_A} \end{bmatrix}.$$

It is immediate to show that $v_1(\kappa), v_2(\kappa) > 0$. As in Proposition 1, $\bar{v} - v(\kappa)$ is the distance between any value $\bar{v}$ that the planner may choose and the one that would have been optimal from a purely allocative perspective. Moreover, welfare can be expressed as (a monotone transformation of) a quadratic form of this distance. In particular, using the FOCs with respect to φ_0 and l_0, we get that welfare is given by

$$\mathcal{W}(\varphi, l, \kappa) = \frac{\epsilon - 1 + \gamma}{1-\gamma}\frac{1}{\epsilon} \exp V_c(\varphi, l, \kappa).$$

Since $\epsilon - 1 + \gamma > 0$ and $\frac{1}{\epsilon} > 0$, we can once again consider the following monotone transformation of welfare:

$$\mathcal{TW}(\varphi, l, \kappa) \equiv \frac{1}{1-\gamma} V_c(\varphi, l, , \kappa).$$

Using the characterization of the optimal coefficients (φ, l) as functions of $\bar{v}$, we conclude that

$$\mathcal{TW}(\varphi, l, \kappa) = W(\bar{v}, \kappa) \equiv A(\kappa) + (\bar{v} - v(\kappa))' B(\kappa) (\bar{v} - v(\kappa)), \tag{A.45}$$

where $A(\kappa)$ is scalar that identifies the level welfare attained when $\bar{v} = v(\kappa)$, while $B(\kappa)$ is a 2-by-2 matrix that identifies the Hessian of the (transformed) welfare function W. The latter is given by the following:

$$\begin{aligned}
B(\kappa) &\equiv \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}, \qquad \text{where} \\
b_{11} &\equiv -\frac{\epsilon\rho(1-\theta+\rho(-1+\epsilon+\theta))\kappa_x + \epsilon\rho(1-\theta)(1-\rho+\rho\epsilon)(\kappa_A+\kappa_z)}{2(-1+\gamma+\epsilon)(1-\theta)\kappa_x(\kappa_A+\kappa_x+\kappa_z)}, \\
b_{22} &\equiv -\frac{\epsilon}{2(\kappa_A+\kappa_z)} - \frac{\epsilon^2\theta}{2(-1+\gamma+\epsilon)(1-\theta)(\kappa_A+\kappa_x+\kappa_z)}, \\
b_{12} &\equiv b_{21} \equiv \frac{\epsilon^2\theta\rho}{2(-1+\gamma+\epsilon)(1-\theta)(\kappa_A+\kappa_x+\kappa_z)}.
\end{aligned}$$

Note that $b_{11} < 0$ and that the determinant of $B(\kappa)$ is positive:

$$\begin{aligned}
det(B(\kappa)) &= b_{11}b_{22} - b_{12}b_{21} \\
&= \frac{\epsilon^3\theta\rho(1-\rho+\rho\epsilon)}{4(-1+\gamma+\epsilon)^2(1-\theta)\kappa_x(\kappa_A+\kappa_x+\kappa_z)} \\
&\quad + \frac{\epsilon^2\rho(1-\theta+\rho(-1+\epsilon+\theta))\kappa_x + \epsilon^2\rho(1-\theta)(1-\rho+\rho\epsilon)(\kappa_A+\kappa_z)}{4(-1+\gamma+\epsilon)(1-\theta)\kappa_x(\kappa_A+\kappa_z)(\kappa_A+\kappa_x+\kappa_z)} \\
&> 0.
\end{aligned}$$

It follows that the matrix $B(\kappa)$ is negative definite and, hence, the aforementioned quadratic form for welfare in (A.45) is also negative definite. The same type of arguments as in Proposition 1 then imply that the optimal $\bar{v}$ is positive and indeed greater than $v(\kappa)$, which in turn guarantees that $\Delta_Y > 0$ and $\Delta_p > 0$. QED.

Finally, note that none of the above derivations require the assumption that $\alpha > 0$. That restriction will be used only in the proof of Proposition 5.

Proof of Proposition 5. We prove the proposition in reverse order.

Part (ii). At the optimal allocation, aggregate output is given by $\log Y(\bar{a}, z) = const + c_A^* \bar{a} + c_z^* z$, where $c_A^* = (\varphi_a^* + \varphi_x^*) + (1-\theta)(l_a^* + l_x^* + l_A^*)$ and $c_z^* = \varphi_z^* + (1-\theta) l_z^*$. If monetary policy were replicating the flexible-price allocations, then aggregate output would be given by $\log Y(\bar{a}, z) = const + \hat{c}_A \bar{a} + \hat{c}_z z$, where $\hat{c}_A = (\varphi_a^* + \varphi_x^*) + (1-\theta)(\hat{l}_a + \hat{l}_x + \hat{l}_A)$ and $\hat{c}_z = \varphi_z^* + (1-\theta)\hat{l}_z$. Analogous expressions hold for the price level.

Using condition (A.27) we have

$$\rho_A^* - \hat{\rho}_A = \left(\frac{1}{\rho} - \gamma\right)(1-\theta)(l_x^* - \hat{l}_x) - \gamma(1-\theta)(l_A^* - \hat{l}_A).$$

From Proposition 4, $l_x^* - \hat{l}_x = \frac{\kappa_x^*}{\kappa_A + \kappa_x^* + \kappa_z^*}(\hat{l}_A - l_A^*)$; thus,

$$\begin{aligned}
\rho_A^* - \hat{\rho}_A &= \left[\left(\frac{1}{\rho} - \gamma\right)\frac{\kappa_x^*}{\kappa_A + \kappa_x^* + \kappa_z^*} + \gamma\right](1-\theta)\left(\hat{l}_A - l_A^*\right) \\
&= \left[\frac{1}{\rho}\frac{\kappa_x^*}{\kappa_A + \kappa_x^* + \kappa_z^*} + \gamma\frac{\kappa_A + \kappa_z^*}{\kappa_A + \kappa_x^* + \kappa_z^*}\right](1-\theta)\lambda\Delta_p \\
&\equiv \chi_3 \Delta_p.
\end{aligned}$$

Part (i). Consider now the optimal tax. From conditions (A.37) and (A.38), we know that the optimal tax satisfies

$$\begin{aligned}
\tau_A^* &= \frac{1}{\chi k}\left\{\alpha c_A^* - \chi\frac{\kappa_A + \kappa_x^* + \kappa_z^*}{\kappa_x^*}\varphi_x^*\right\}, \quad \text{and} \\
\tau_z^* &= \frac{1}{\chi k}\left\{\alpha c_z^* - \chi\left(\varphi_z^* - \varphi_x^*\frac{\kappa_z^*}{\kappa_x^*}\right)\right\}.
\end{aligned}$$

Let

$$\begin{aligned}
\hat{\tau}_A &\equiv \frac{1}{\chi k}\left\{\alpha \hat{c}_A - \chi\frac{\kappa_A + \kappa_x^* + \kappa_z^*}{\kappa_x^*}\varphi_x^*\right\}, \quad \text{and} \\
\hat{\tau}_z &\equiv \frac{1}{\chi k}\left\{\alpha \hat{c}_z - \chi\left(\varphi_z^* - \varphi_x^*\frac{\kappa_z^*}{\kappa_x^*}\right)\right\};
\end{aligned}$$

these coefficients identify the tax policy that would be required in order to implement the optimal stage-1 sensitivities, φ_x^* and φ_z^*, if monetary policy were replicating the flexible-price allocations associated with these stage-1 sensitivities. It is easy to verify that

$$\hat{\tau}_A = -\bar{\lambda}\Delta \qquad \text{and} \qquad \hat{\tau}_z = -\bar{\lambda}\frac{\kappa_z^*}{\kappa_A + \kappa_z^*}\Delta,$$

where $\Delta \equiv \Delta_Y + \Delta_p$ and $\bar{\lambda} \equiv \frac{(\epsilon\theta - \alpha(-1+\gamma+\epsilon+\theta-\gamma\theta))^2(\kappa_A + (1-\alpha)\kappa_x^* + \kappa_z^*)}{(1-\alpha)\epsilon^2(-1+\gamma+\epsilon)\theta\kappa_x^*} > 0$. Note that the optimal tax satisfies

$$\tau_A^* = \hat{\tau}_A + \frac{1}{\chi k}\alpha(c_A^* - \hat{c}_A) \qquad \text{and} \qquad \tau_z^* = \hat{\tau}_z + \frac{1}{\chi k}\alpha(c_z^* - \hat{c}_z).$$

By Proposition 4,

$$c_A^* - \hat{c}_A = (1-\theta)(l_x^* - \hat{l}_x + l_A^* - \hat{l}_A) = -(1-\theta)\frac{\kappa_A + \kappa_z^*}{\kappa_A + \kappa_x^* + \kappa_z^*}\lambda\Delta_p$$

and

$$c_z^* - \hat{c}_z = (1-\theta)(l_z^* - \hat{l}_z) = (1-\theta)\frac{\kappa_z^*}{\kappa_A + \kappa_x^* + \kappa_z^*}\lambda\Delta_p.$$

Therefore,

$$\begin{aligned}
\tau_A^* + \tau_z^* &= \hat{\tau}_A + \hat{\tau}_z - \frac{1}{\chi k}\alpha\frac{\kappa_A}{\kappa_A + \kappa_x^* + \kappa_z^*}\lambda\Delta_p \\
&= -\bar{\lambda}\left(1 + \frac{\kappa_z^*}{\kappa_A + \kappa_z^*}\right)(\Delta_Y + \Delta_p) - \frac{1}{\chi k}\alpha\frac{\kappa_A}{\kappa_A + \kappa_x^* + \kappa_z^*}\lambda\Delta_p.
\end{aligned}$$

The statement of the proposition follows from assuming that $\alpha > 0$ and letting $\chi_1 \equiv \bar{\lambda}\left(1 + \frac{\kappa_z^*}{\kappa_A + \kappa_z^*}\right) > 0$ and $\chi_2 \equiv \bar{\lambda}\left(1 + \frac{\kappa_z^*}{\kappa_A + \kappa_z^*}\right) + \frac{1}{\chi k}\alpha\frac{\kappa_A}{\kappa_A + \kappa_x^* + \kappa_z^*}\lambda > 0$. QED.