Efficiency and Policy with Endogenous Learning*

George-Marios Angeletos MIT

Luigi Iovino Bocconi

Jennifer La'O Columbia

July 27, 2015

Abstract

This paper studies how the endogeneity of information affects the efficiency of the business cycle and the nature of optimal policy. The business cycle is found to be excessively noisy. Optimal taxes and optimal monetary policy are shown to be countercyclical. Such policies incentivize agents to make their choices more sensitive to their idiosyncratic sources of information, which in turn improves the informativeness of market signals and macro statistics.

* *Email:* `angelet@mit.edu`, `luigi.iovino@unibocconi.it`, `jenlao@columbia.edu`. We thank our discussant Christian Hellwig, the organizers Alessandro Pavan and Xavier Vives, and the participants in the workhop on Information, Competition, and Market Frictions at the 2013 Barcelona GSE Summer Forum for their helpful comments and suggestions.

1 Introduction

In any modern economy, the fortunes of each agent depend on the choices of millions of other agents. Market participants therefore anxiously monitor the release of macroeconomic statistics and market signals such as prices for clues about the state of the economy and one another's choices. Even if these signals happen to convey little information about payoff-relevant fundamentals such as preferences and technologies, they help agents coordinate their choices with one another, thus affecting the response of the economy to the underlying business-cycle disturbances. Importantly, the informativeness of these signals is itself a function of how much information agents collect in the first place and how they react to it.

These observations raise the following question: how does the endogenous aggregation 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 prototypical, micro-founded model of the macroeconomy.

Our model shares similar micro-foundations with standard RBC/NK models. More specifically, firms produce differentiated products, implying that individual incentives depend on expectations of aggregate demand. Relative to the standard paradigm, the key added features are the incompleteness and the endogeneity of available information: firms and workers are dispersed in segmented locations ("islands") and make certain employment, production, and pricing choices while lacking common knowledge of one another's fundamentals and choices. As a result of this information incompleteness, business cycles can now originate, not only from shocks to preferences and technologies, but also from various forms of noise.[1] The equilibrium impact of this noise and the associated departure of the economy from the first best crucially depend on how much information agents acquire and how they respond to it.

If all sources of information were exogenous, the economy we study would have belonged, in effect, to the class of incomplete-information games studied in Angeletos and Pavan (2007). In general, the equilibrium use of information in such games need not coincide with the socially efficient one. For the class of micro-founded economies we are interested in, however, this is not the case: if information were exogenous, the equilibrium use of information would have been efficient even without any policy intervention.[2] It follows that our framework

[1] In the present paper, the noise is modeled as a measurement error in the available signals of either the exogenous fundamentals or the endogenous aggregate activity. However, following Angeletos and La'O (2013), noise could also be engineered from shocks to higher-order beliefs, in which case the business cycle looks like the product of extrinsic, sunspot-like, fluctuations.

[2] When we say "without policy intervention" we mean (i) no taxes and (ii) a monetary policy that replicates flexible-price allocations.

isolates the endogeneity of the information as the sole source of inefficiency.[3]

To understand the nature and the policy implications of this inefficiency, we proceed in three steps. First, in the baseline version of our model, we abstract from any nominal rigidities and focus on the endogenous aggregation of information that obtains as agents observe noisy signals of one another's choices and of aggregate economic activity. We then add a nominal friction by requiring that prices cannot be contingent on realized demand; this permits us to study both the robustness of our key insights to the presence of nominal frictions and their translation in the context of monetary policy.

The lessons that emerge can be summarized as follows. First, inefficiency emerges in the endogeneity of the aggregation of information. This inefficiency manifests itself in the form of an excessive contribution of noise to equilibrium fluctuations relative to that of fundamentals: the business cycle is "too noisy". Third, the government can improve upon the laissez-faire equilibrium by following a particular type of countercyclical fiscal and monetary policies.

The rationale behind these policies is quite simple. By making individual incentives relatively less sensitive to realized aggregate economic activity, these policies incentivize the agents to care less about one another's choices and, thereby, to rely relatively more on their idiosyncratic sources of information about the fundamentals. This mechanism increases the informativeness of the available market signals and macro statistics, thereby reducing the business-cycle contribution of noise and increasing welfare.

Related literature. Methodologically, our paper builds on the welfare analysis in Vives (1988) and Angeletos Pavan (2007, 2008). This prior work developed and characterized a notion of constrained efficiency for a certain class of incomplete-information, linear-quadratic games. The notion of constrained efficiency we adopt in this paper is essentially the same. Our marginal contribution rests in characterizing the constrained efficient allocation, and the optimal policy, of a particular micro-founded, business-cycle economy. Vives (2013) studies similar questions, but in a microeconomic market game environment.

Another integral part of our contribution rests in adapting the primal approach from the Ramsey literature (Lucas and Stokey, 1983; Chari and Kehoe, 1999) 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

[3] In the background, there is always the familiar inefficiency associated with monopoly power. This, however, is orthogonal to the type of informational inefficiency that are the focus of our paper and, as usual, can be shut down by considering a subsidy that corrects the monopoly distortion.

agents observe about one another's 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 REE settings, the primal approach we take in this paper could be of broader methodological value.[4]

Finally, the monetary aspect of our analysis is closely related to the following set of papers that have studied optimal monetary policy in the presence of informational frictions: Ball, Mankiw and Reis (2005), Adam (2007), Lorenzoni (2010), Paciello and Wiederholt (2013), and Angeletos and La'O (2012). All of these papers abstract from endogenous aggregation of information. Finally, as emphasized in Angeletos and La'O (2012), the rest of the aforementioned papers 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.

Layout. The rest of the paper is organized as follows. Section 2 introduces the baseline model. Section 3 studies efficiency, implementability, and optimal policy in the baseline model. We then present two extensions in Sections **??** and 4. Section 4 adds a nominal friction and studies the jointly optimal fiscal and monetary policies. Section **??** concludes.

2 The Baseline Model

The framework we use in this paper is similar to the one in Angeletos La'O (2009, 2013b). The key difference is the introduction of endogenous private and public information, which is central to the question of interest for this paper.

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

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

local labor market. At this point, workers and firms in each island have perfect information regarding local productivity, but imperfect information regarding the productivities in 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 is aggregated and 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

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

with

$$U(C) = \frac{C^{1-\gamma}}{1-\gamma} \qquad \text{and} \qquad V(n) = \frac{n^\epsilon}{\epsilon}.$$

where $\gamma \geq 0$ parameterizes the income elasticity of labor supply (also, the coefficient of relative risk aversion), $\epsilon \geq 1$ 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 of all of the commodities that the household purchases and consumes during 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 the 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 period-t price of the commodity produced by the representative firm on island i, $\pi_{i,t}$ is the period-t profit of that firm, $w_{i,t}$ is the period-t wage on island i, and R_t is the period-t nominal gross rate of return on the riskless bond, and B_t is the amount of bonds held in period t.

The objective of each 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 during stage 1 its workers 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 of the income that the worker-members have collected and decides how much to consume in each of the commodities and how much to save (or borrow) in the riskless bond.

Firms. The output of the representative firm on island i during 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, $n_{ij,t}$ is the firm's employment, and $\theta \in (0,1)$ parameterizes 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: $c_{i,t} = q_{i,t} \quad \forall i$. 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 lognormally distributed in the cross-section of islands:

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

where $\bar{a}_t$ denotes the underlying aggregate productivity shock and $\xi_{i,t}$ is a purely idiosyncratic (i.e., island-specific) productivity 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 $\kappa_{A,t} \equiv \sigma_{A,t}^{-2}$ and $\kappa_{\xi,t} \equiv \sigma_{\xi,t}^{-2}$.

Information. In stage 1, when key employment and production choices are made, the firms and worker that are located in any given island face uncertainty about what's going on in other islands. More specifically, these agents get to see 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 informations in other islands.[6]

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

$$x_{i,t}^{a} = \bar{a}_t + \varepsilon_{i,t}^{xa},$$

and the public one by

$$y_t^{a} = \bar{a}_t + \varepsilon_t^{ya},$$

where $\varepsilon_{i,t}^{xa} \sim \mathcal{N}(0, \sigma_{xa}^2)$ and $\varepsilon_t^{ya} \sim \mathcal{N}(0, \sigma_{ya}^2)$ are noise (the former one idiosyncratic, the latter one common). For future reference, we let $\kappa_{xa} \equiv \sigma_{xa}^{-2}$. $\kappa_{ya} \equiv \sigma_{ya}^{-2}$.

Finally, firms and households in each island 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 signal is

$$y_t^{q} = \log Q_t + \varepsilon_t^{yq},$$

while the endogenous private signal is

$$x_{i,t}^{q} = \log Q_t + \varepsilon_{i,t}^{xq},$$

where $\varepsilon_t^{yq} \sim \mathcal{N}(0, \sigma_{yq}^2)$ and $\varepsilon_{i,t}^{xq} \sim \mathcal{N}(0, \sigma_{xq}^2)$ are noises (the first one common, the second one idiosyncratic across islands).

[5] For example, the special case where aggregate productivity follows a random walk could be nested by letting $\mu_t = \bar{a}_{t-1}$ and letting 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 aforementioned public signals are meant to capture the macroeconomic data released by various government agencies.[7] The aforementioned private signals, on the other hand, can be thought of as emerging from firms collecting data about product conditions in particular markets or particular locations; idiosyncratic sampling or measurement error could then justify the idiosyncratic noise in these signals. In a richer version of the model, these private signals could also emerge from localized trading or other interactions, in the same spirit as in Amador and Weill (2009); each island then gets to see the production levels of the islands with which it trades, but each island trades only with a few other islands, so that it is as if each island observes aggregate output with idiosyncratic noise. Putting aside the possible interpretations of these signals, the key modeling feature here is that the informativeness of y_t^q and $x_{i,t}^q$ depend on the choices of other agents. In this regard, these signals are meant to capture various channels of social learning, some of which could be global (public) and some of which could be local (private).

3 Efficiency, Implementability, and Optimal Policy

The characterization of the equilibrium in the absence of policy intervention can be obtained in a similar manner as in Angeletos and La'O (2009). Here, we put aside this issue and focus, instead, on the characterization of efficient and implementable allocations.

Some preliminaries and the efficiency notion

In addition to the usual resource constraint, the fictitious planner of our economy faces an informational constraint: the employment and production choices he can dictate (or incentivize) to any given island i must be measurable in the vector $(a_i, x_i^a, y^a, x_i^q, y^q)$. This measurability constraint epitomizes the informational friction faced by the market mechanism and the planner alike.

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.

Indeed, as long as this is the case, we can guess and verify that the information that

[7] In the extended monetary model (Section 4), we will introduce a signal of the aggregate price level in addition to the signal of aggregate output. Moreover, our analysis can readily accommodate a signal of aggregate employment. See Appendix B for a discussion of alternative signals.

is available to any island i in stage 1 can be summarized by the triplet (a_i, x_i, y), where a_i is the current local productivity; x_i is a Gaussian sufficient statistic for all the private (local) information; and y is a Gaussian sufficient statistic for all the public information. The aforementioned requirement then boils down on restricting attention to the space of allocation in which the output of an island is a log-linear function of the triplet $\omega_i = (a_i, x_i, y)$, namely

$$\log q_i\left(a_i, x_i, y\right) = \varphi_0 + \varphi_a a_i + \varphi_x x_i + \varphi_y y, \tag{1}$$

for arbitrary scalar coefficients $\left(\varphi_0, \varphi_a, \varphi_x, \varphi_y\right) \in \mathbb{R}^4$. For any such allocation, aggregate output is itself a log-linear function of the underlying aggregate shocks:

$$\log Q = \varphi_0' + \left(\varphi_a + \varphi_x\right)\bar{a} + \varphi_y y \tag{2}$$

It follows that the endogenous signals x_i^q and y^q can be transformed into Gaussian signals about the underlying aggregate productivity, thus preserving the Gaussian structure of the information. The aforementioned sufficient statistics can then be constructed by taking, in effect, the projection of the aggregate productivity to the relevant signals.

Let κ_x and κ_y denote the precisions of the statistics x_i and y. Because of the endogeneity of the signals x_i^q and y_i^q, these precisions are endogenous to the production strategy played by all islands.

Lemma 1. *Take any log-linear production strategy of the form (1), for arbitrary* $(\varphi_0, \varphi_a, \varphi_x, \varphi_y)$. *The precisions of the sufficient statistics x and y generated by this strategy are given by*

$$\kappa_x = \sigma_\xi^{-2} + \sigma_{xa}^{-2} + (\varphi_a + \varphi_x)^2\sigma_{xq}^{-2} \qquad \text{and} \qquad \kappa_y = \sigma_{ya}^{-2} + (\varphi_a + \varphi_x)^2\sigma_{yq}^{-2} \tag{3}$$

To understand this result, note that the endogenous signals x_i^q and y^q about aggregate output can be transformed to simple Gaussian signals about the underlying aggregate productivity because aggregate output is itself a log-linear function of $\bar{a}$, as in (2). From this equation (2), note that the sum $\varphi_a + \varphi_x$ determines the sensitivity of aggregate output to the underlying 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 y^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_y of the overall private and public information.

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 $\varphi = (\varphi_0, \varphi_a, \varphi_x, \varphi_y)$ and the information precisions $\kappa = (\kappa_x, \kappa_y)$.

Lemma 2. *There exists a function $\mathcal{W} : \mathbb{R}^6 \to \mathbb{R}$ 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 log-linear production strategy q as in (1) that maximizes welfare subject to condition (3).*

Thus, the notion of constrained efficiency we adopt is similar to that of Angeletos and Pavan (2007, 2008), 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, as it ensures that planner must take into account how different allocations sustain different information structures.[8]

Implementability

Our notion of constrained efficiency allows the planner to choose allocations and associated information structures subject to without any regard to incentive constraints. Following the Ramsey tradition, we nevertheless wish to study the set of allocations that can be implemented with a relatively simple set of tax instruments. More specifically, we consider any combination of the following tax instruments: a linear tax on firm sales, output, employment, or payroll; a linear tax on household labor income; or a linear tax on household consumption. These taxes are collected in stage 2 and can be contingent on the information that is publicly available at that state. Finally, for tractability, we require the tax rates to be log-normal. It is then straightforward to show that any combination of the aforementioned taxes reduces to a single tax wedge between the marginal return and the marginal cost of labor.

Formally, let $\omega_i \equiv (a_i, x_i, y)$ denote the local type of an island and Ω the cross-sectional distribution of such types (also the aggregate state). The key implementability constraint is then identified in the following lemma.

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

$$q(\omega)^{\frac{\epsilon}{\theta}+\frac{1}{\rho}-1} = \left(\tfrac{\rho-1}{\rho}\right)\theta A(\omega)^{\frac{\epsilon}{\theta}}\mathbb{E}\left[(1-\tau(\Omega))Q(\Omega)^{\frac{1}{\rho}-\gamma}|\omega\right]\forall\omega, \tag{4}$$

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

with $Q(\Omega) = \left[\int q(\omega)^{\frac{\rho-1}{\rho}} d\Omega(\omega)\right]^{\frac{\rho}{\rho-1}} \forall\Omega$ *is aggregate output and where* $1-\tau(\Omega)$ *is the total tax wedge induced by the aforementioned taxes.*

The proof of this result is similar to the characterization of equilibrium in Angeletos and La'O (2009) but with the inclusion of a tax wedge. For future reference, let us define

$$\beta \equiv \frac{\frac{\epsilon}{\theta}}{\frac{\epsilon}{\theta}+\frac{1}{\rho}-1} > 1 \qquad \text{and} \qquad \alpha \equiv \frac{\frac{1}{\rho}-\gamma}{\frac{\epsilon}{\theta}+\frac{1}{\rho}-1} < 1, \tag{5}$$

As evident from (4), 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. Finally, to preserve once again the Gaussian structure, we impose that that the tax wedge satisfies

$$-\log(1-\tau(\Omega)) = \tau_0 + \tau_A\bar{a} + \tau_Q \log Q(\Omega) + \tau_y y, \tag{6}$$

for some known coefficients τ_0, τ_A, τ_Q, and τ_y. The coefficients τ_A, τ_Q, and τ_y determine the elasticities of the tax with respect to aggregate productivity, aggregate output, and the common belief about aggregate productivity; the coefficient τ_0 controls the mean level.

Definition 2. *The best implementable allocation is production strategy* q *as in (1) that maximizes welfare subject to condition (3) as well as condition (4) for some* $\tau(\Omega)$ *as in* (6).

Because $(\bar{a}, y)$ is a sufficient statistic for the entire exogenous state Ω, we can always write the tax as a function of $(\bar{a}, y)$ alone. However, we favor implementations that express the tax as a function of aggregate productivity $\bar{a}$ and aggregate output Q, rather than as a function of the public signal y for one simple reason: even though in the model y is common knowledge, in practice it seems hard to measure y. After all, the public signal is just a convenient modeling device that permits us to capture a variety of common sources of information that may be available to the market but not necessarily directly observed by the government. It is then reassuring that the contingency on y is strictly needed only in a degenerate case of no practical interest.

Proposition 1. *Consider any strategy as in (1). There exists a state-contingent tax policy as in (6) that implements this strategy as an equilibrium strategy, i.e., that satisfies condition (4), if and only if* $\varphi_a = \beta$*. Moreover, as long as* $(\varphi_x, \varphi_y) \neq 0$*, this can be achieved with a policy that sets* $\tau_y = 0$.

This result highlights the power that the contingencies of taxes on macroeconomic outcomes can have in controlling the decentralized use of information—an insight that extends beyond our model and that was first highlighted by Angeletos and Pavan (2008) within a more abstract framework.[9] The simplest (and more robust) way to understand this result is by ignoring the details of specific implementations and instead looking at the correlations of the tax with underlying productivity and noise shocks. In particular, suppose the tax is negatively correlated with innovations in aggregate productivity and positively correlated with the common noise; the government can induce these correlations by making the tax contingent either on $(\bar{a}, y)$, or on $(\bar{a}, Q)$. Either way, anticipating these correlations, firms 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 simply 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 permit the government to control freely both φ_x and φ_y, the sensitivities to private and public information.[10] The only sensitivity that the policy cannot affect is the one with respect to local productivity, namely the coefficient φ_a; this is simply because the tax is contingent only on macroeconomic outcomes, not on idiosyncratic variables.[11]

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 $\varphi = \left(\varphi_0, \varphi_a, \varphi_x, \varphi_y\right)$ and $\kappa = (\kappa_x, \kappa_y)$; 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 $\kappa_x(\varphi_a + \varphi_x)$ and $\kappa_y(\varphi_a + \varphi_x)$ denote the functions defined by (3). We can then express the planner's problem as follows.

[9] In Section 4, we will show that, to the extent that prices are sticky, monetary policy can play a similar role, although it does not have as much power as state-contingent taxes.

[10] To be precise, the argument made above only explains how taxes impact individual firm incentives holding given the behavior of other firms; that is, they explain they impact of the policy on best responses, not on equilibrium. 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.

[11] We could relax this restriction by letting the marginal tax faced by a firm increase/decrease with its own output. In general, such non-linear taxes could be useful. In our context, however, the omission of such non-linear taxes is without loss of optimality.

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

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 in two steps. We summarize these steps as follows, while we leave the detailed analysis to the appendix.

Auxiliary Problem 1. *Choose* $\varphi = (\varphi_0, \varphi_a, \varphi_x, \varphi_y)$ *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* $W(\bar{\varphi}, \kappa_y(\bar{\varphi}))$

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 (8) and (9) 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 prove the existence of an efficient allocation, to interpret Δ as the shadow value of the informational externality, and to complete the full characterization of the efficient allocation.

Proposition 2. *(i) A constrained efficient strategy always exists. Moreover, there exists a scalar* $\Delta > 0$ *such that the efficient strategy is given by*

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

where the coefficients $(\varphi_a^*, \varphi_x^*, \varphi_y^*)$ *and the associated precisions* (κ_x^*, κ_y^*) *satisfy the following:*

$$\begin{aligned}
\varphi_a^* &= \beta && (7)\\
\varphi_x^* &= \left\{ \frac{(1-\alpha)\kappa_x^*}{(1-\alpha)\kappa_x^* + \kappa_y^* + \kappa_A} \right\} \frac{\alpha}{1-\alpha}\beta + \Delta && (8)\\
\varphi_y^* &= \left\{ \frac{\kappa_y^*}{(1-\alpha)\kappa_x^* + \kappa_y^* + \kappa_A} \right\} \frac{\alpha}{1-\alpha}\beta - \frac{\kappa_y^*}{\kappa_A + \kappa_y^*}\Delta && (9)\\
\kappa_x^* &= \sigma_\xi^{-2} + (\varphi_a^* + \varphi_x^*)^2 \sigma_{xq}^{-2}\\
\kappa_y^* &= \sigma_{ya}^{-2} + (\varphi_a^* + \varphi_x^*)^2 \sigma_{yq}^{-2}
\end{aligned}$$

(ii) The constrained efficient allocation is also the best implementable allocation.

(iii) The equilibrium in the absence of policy (zero taxes) satisfies the same conditions as

the efficient strategy above, replacing Δ *with* 0.

Part (i) of Proposition 2 characterizes the efficient strategy. Part (ii) of Proposition 2 informs us that the efficient allocation is also the best implementable allocation. We can prove part (ii) almost immediately. Note that from Proposition 1, we know that the set of implementable log-linear allocations is strictly smaller than the set of all log-linear allocations. However, the only restriction on implementable allocations is that $\varphi_a = \beta$, a condition which is satisfied for free at the efficient allocation (see equation 7). It follows that the best implementable allocation is simply the efficient one.[12]

The scalar Δ is a wedge that summarizes the impact of the informational externality on the efficient production strategy relative to the one that obtains in equilibrium—or, equivalently, relative to the one that would have maximized welfare for given information. Let us elaborate on this. Without any policy, the equilibrium allocation would be as described in Proposition 2 but with the wedge $\Delta = 0$. Note that the precisions κ_x^* and κ_y^* obtained at the efficient allocation above are higher than those obtained at the equilibrium allocation. If some divine power were to keep the precisions of available information constant at those higher levels and give the planner, then the allocation the planner would choose is the one obtained with $\Delta = 0$ (which in turn coincides with the equilibrium one). Clearly, any deviation from this allocation involves a loss in *allocative* efficiency in the sense that it reduces welfare for given information. But it is only this sacrifice that permits the planner to sustain these higher precisions in the absence of the aforementioned divine power. What then justifies the sacrifice is precisely that these higher precisions contribute to higher welfare. In this sense, 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).

The property that Δ is positive simply means that the efficient allocation features a higher sensitivity to local information that the equilibrium one. This is intuitive: starting from equilibrium, an increase in φ_x raises both κ_x and κ_y. However, the result is not self-evident for three reasons.

First, in general it is not obvious that, starting from equilibrium, a local increase in the precision of information increases welfare. For example, this is not the case in the economy of Morris and Shin (2002). However, the results of the previous section have already ruled out this complication for our class of economies: if information had been exogenous, the equilibrium use of information would have been efficient, guaranteeing that equilibrium welfare would increase with any additional information, whether private or public.

Second, raising the precision of information does not come for free: it comes at the cost of

[12] This will not be true in the extended monetary model; but here it offers a useful benchmark.

raising the cross-sectional dispersion in output, employment and relative prices, which means less allocative efficiency for given information. However, because the equilibrium allocation maximizes allocative efficiency to start with, this cost is only second-order locally, while the welfare benefit of more information is first-order.

Finally, although this last argument establishes the direction of *local* welfare improvements, such local arguments are not necessarily informative about the position of the *global* maximum when the planner's problem fails to be concave, as it is the case here. In particular, a complication is that the planner can induce a high precision of the endogenous signals by picking, not only a sufficiently high $\varphi_a + \varphi_x$, but also a sufficiently negative $\varphi_a + \varphi_x$. 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 negative values for $\varphi_a + \varphi_x$ because they involve unnecessarily high allocative inefficiency: 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. Along with the fact that 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.

Optimal Policy

We can summarize the normative lesson of the preceding analysis in the following.

Corollary 1. *The endogeneity of information renders the business cycle inefficient: the equilibrium degree of social learning is too low, the sensitivity of equilibrium allocations to private information is too low, and their sensitivity to public information is too high.*

We can then deduce the relevant policy implications by characterizing the tax wedge that implements the efficient allocation.

Proposition 3. *The optimal tax is countercyclical in either of the following senses:* $Corr(\tau, \bar{a}) < 0$, $Corr(\tau, \bar{a}|y) < 0$, *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 economic activity to private information, so as to boost social learning; it then also dictates to lower φ_y the sensitivity to public information downwards, 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_y, which is necessary for counterbalancing the allocative inefficiency caused by the higher φ_x but is irrelevant for the efficiency of learning.[13]

A final note on fiscal implementation. There are multiple combinations of the coefficients (τ_a, τ_Q, τ_y) in the baseline model that induce the same incentives and hence implement the efficient allocation. However, if we restrict attention to our preferred class of policies, which allow the tax rate to depend only on aggregate productivity and aggregate output, then there remains a unique policy within that class that implements the efficient allocation. In particular, assuming that good news about productivity raises employment and output ($\varphi_y^* > 0$), our preferred implementation would set $\tau_y = 0$ along with $\tau_Q < 0$ and $\tau_A > 0$. That is, firms should expect their taxes to increase during a boom that is not warranted by aggregate productivity, and to decrease when actual productivity is higher. Moreover, even without this particular selection, the value of the tax as a function of the the aggregate productivity shock $\bar{a}$ and the aggregate noise shock ε (where $\varepsilon = y - \bar{a}$) is always uniquely determined along the equilibrium path. We can thus talk about *the* optimal tax, once again ignoring the details of the particular implementation.

[13] Also, as mentioned above, there are a variety of tax schedules as in (6) that can achieve these cyclical properties. Assuming that good news about productivity raises employment and output ($\varphi_y^* > 0$), our preferred implementation would set $\tau_y = 0$ along with $\tau_Q < 0$ and $\tau_A > 0$. That is, firms should expect their taxes to increase during a boom that is not warranted by aggregate productivity, and to decrease when actual productivity is higher.

4 The Monetary Model

In this section we introduce a New-Keynesian monetary extension of our baseline model, which permits us to study how the dispersion of information interacts with nominal frictions and how it impacts optimal monetary policy.

We modify the baseline model in four dimensions. First, we introduce price rigidities. In particular, we assume that firms set nominal prices at the end of 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.

Second, we let firms make a second employment choice in stage 2 and, accordingly, we let households make a second labor-supply choice in that stage; this permits real output to respond to the new information that becomes available during that stage as well as the monetary policy to have real effects. In particular, the output and consumption level of the typical commodity produced on island ω is now given by

$$c(\omega,\Omega) = A(\omega)n(\omega)^{\theta_1} l(\omega,\Omega)^{\theta_2}$$

where $l(\omega,\Omega)$ denotes the labor employed on island ω during stage 2 and where $\theta_1,\theta_2 > 0, \theta_1+\theta_2 < 1$. Note that second-stage employment can depend on the information that becomes available at that stage, which explains why l is a function of Ω and not just ω. Accordingly, the per-period utility of the representative household is given by

$$U(C(\Omega)) - \int \frac{1}{\epsilon_1} n(\omega)^{\epsilon_1} d\Omega(\omega) - \int \frac{1}{\epsilon_2} l(\omega,\Omega)^{\epsilon_2} d\Omega(\omega).$$

where $\epsilon_1, \epsilon_2 > 1$.[14] Finally, for future reference we let

$$q(\omega) \equiv A(\omega)n(\omega)^{\theta_1} \quad \text{and} \quad Q(\Omega) \equiv \left(\int q(\omega)^{\frac{(\rho-1)\epsilon_2}{\rho\epsilon_2-(\rho-1)\theta_2}} d\Omega(\omega) \right)^{\frac{\rho\epsilon_2-(\rho-1)\theta_2}{(\rho-1)\epsilon_2}}.$$

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) production activity. In particular, we let all agents observe the following two public signals: $y^q = \log Q(\Omega) + \varepsilon^{yq}$ and $y^p = \log P(\Omega) + \varepsilon^{yp}$, where $\varepsilon^{yq} \sim \mathcal{N}(0,\sigma^2_{yq})$ and $\varepsilon^{yp} \sim \mathcal{N}(0,\sigma^2_{yp})$ are the respective noises. Note that the price signal is properly defined only conditional on having first defined

[14] Here, we assume local labor markets for both stages, but our results easily extend if we assume a global labor market during stage 2; the "geographical" boundaries are important only for avoiding perfect information aggregation at stage 1.

what a price is, which is clear on equilibrium but not off equilibrium. We thus extend the definition of the price signal for arbitrary allocations so that it is always a signal of the shadow prices faced by the representative consumer. Finally, it would be straightforward to extend the analysis to situations where some of the output or price signals are private rather than public, as in the baseline model; here, the output and price signals are left as public only for expositional simplicity.

Finally, we let the monetary authority control nominal aggregate demand. In particular, we let

$$M(\Omega) \equiv P(\Omega)C(\Omega)$$

and assume that the monetary authority directly controls $M(\Omega)$. As in Woodford (2003a), this is equivalent to assuming that the monetary authority pays interest on money holdings and appropriately adjusts the nominal interest rate so as to induce the desired level of nominal spending. To simplify the exposition, we specify monetary policy directly in terms of the target level of nominal spending. For tractability, we then restrict attention to the following class of log-normal policy rules:

$$\log M(\Omega) = \lambda_0 + \lambda_A \bar{a} + \lambda_p \log P(\Omega), \tag{10}$$

where the coefficients λ_A and λ_p parameterize the degree to which monetary policy accommodates innovations in aggregate productivity and the aggregate price level.

Note that, once we have identified the optimal monetary policy in terms of a target rule for aggregate nominal demand, it is straightforward to translate that policy it in terms of a target rule for the nominal interest rate—provided, of course, that we take a stand on the behavior of real interest rate under the flexible-price allocation.

Rather than characterizing the equilibrium allocations that obtain under specific monetary policies, we instead move directly to characterizing the set of implementable allocations given arbitrary monetary policy as in (10) and a tax policy as in (6). We then identify the best implementable (or optimal) allocation within this set.

Implementability

In order to preserve tractability, we again guess and verify that the relevant information available to any island in stage 1 can be summarized by $\omega = (a, x, y)$ where a is local productivity; x is a Gaussian sufficient statistic for all private information regarding the aggregate state; and y is a Gaussian sufficient statistic summarizing all public information regarding the aggregate state.

In order for this conjecture to be true in the presence of endogenous signals, we restrict our attention to allocations that preserve the Gaussian structure of information. We thus restrict ourselves to the space of log-linear production strategies. The employment and production during stages 1 and 2 of any island ω can be reduced to a pair of log-linear strategies as follows:

$$\log q(\omega) = \varphi_0 + \varphi_a a + \varphi_x x + \varphi_y \qquad \text{and} \qquad \log l(\omega, \Omega) = l_0 + l_A \bar{a} + l_a a + l_x x + l_y y, \quad (11)$$

for arbitrary coefficients $\varphi = (\varphi_0, \varphi_a, \varphi_x, \varphi_y) \in \mathbb{R}^4$ and $l = (l_0, l_A, l_a, l_x, l_y) \in \mathbb{R}^5$.[15] When prices are sticky, it is natural to think of firms choosing q and p, rather than q and l; the combination of the first-stage choices with the monetary policy then determines the second-stage employment outcomes. However, for the purpose of studying efficiency, it is more appropriate—or at least more convenient—to reason directly in terms of allocations, independently of whether prices are sticky or not.

Both types of employment are thus restricted to be log-linear in the underlying shocks, which further implies that the log of aggregate output and the log of the aggregate price level are themselves log-Normal. This ensures that the endogenous signals y^q and y^p can be transformed into simple Gaussian signals x and y about the underlying aggregate productivity, thereby verifying our conjecture.

Next, let κ_x and κ_y denote the precisions of the sufficient statistics x and y. Take any pair of strategies as in (11); these will induce certain values for the public precision of κ_y, characterized as follows.

Lemma 4. *Take any pair of strategies as in (11).*

(i) The precision of the available public information is given by

$$\kappa_y = \sigma_{ya}^{-2} + (\varphi_a + \varphi_x)^2 \sigma_{yq}^{-2} + \rho^{-2} (\varphi_a + \varphi_x + \theta_2 l_a + \theta_2 l_x)^2 \sigma_{yp}^{-2} \quad (12)$$

(ii) Provided that $\varphi_a + \varphi_x > 0$ and $\varphi_a + \varphi_x + \theta_2 l_a + \theta_2 l_x > 0$, κ_y increases with both φ_x and l_x.

Lemma 4 is the analog of Lemma 1 in our baseline model. The term $\kappa_q \equiv (\varphi_a + \varphi_x)^2 \sigma_q^{-2}$ in condition (12) captures the learning through the quantity signal y^q, as in the baseline model, while the term $\kappa_p \equiv \rho^{-2}((\varphi_a + \varphi_x) + \theta_2(l_a + l_x))^2 \sigma_{yp}^{-2}$ captures the learning through the price signal y^p. To understand this new term, note that $c_a \equiv \varphi_a + \theta_2 l_a$ and $c_x \equiv \varphi_x + \theta_2 l_x$ measure the sensitivities of the output (and consumption) level of commodity ω to, respectively,

[15] Here it is important to note that the second-stage production allocations can depend on the realized aggregate productivity, simply because the latter has become public information at that stage.

the local productivity and local information. Multiplying those by the elasticity of the demand function, namely $1/\rho$, gives the corresponding sensitivities of the shadow prices faced by the representative consumer. It follows that the precision of the price signal is $\kappa_p = (\frac{1}{\rho}c_a + \frac{1}{\rho}c_x)^2\sigma_{yp}^{-2}$, which explains the new term.

We now identify the set of allocations that can be implemented with a monetary policy rule as in (10) and a simple tax policy as in (6). As mentioned previously, this tax can be any combination of the following tax instruments: a linear tax on firm sales, output, employment, or payroll; a linear tax on household labor income; or a linear tax on household consumption.

However, we restrict the set of tax instruments such that we do not allow for fiscal policies that can tax *differentially* across the two types of labor. With a richer set of fiscal instruments that could levy separate taxes on first and second stage labor, one could enlarge the set of implementable allocations and achieve constrained efficiency. We discuss this possibility below. However, we do not find this richness reasonable. It is unlikely that the government can distinguish between these two stages in practice—after all, these stages are very gross representations for the much richer information and price-setting dynamics that are likely to take place in reality.

For this reason, we find it more appropriate to consider a restricted planning problem, one that identifies the optimal allocation among the ones that can be implemented with the combination of a contingent log-linear tax as in (6) and a contingent monetary policy as in (10).[16] The next lemma characterizes the set of such implementable allocations.

Proposition 4. *(i) When prices are sticky, a pair of strategies as in (11) can be implemented as an equilibrium with an appropriate combination of a contingent linear tax and a monetary policy if and only if the following conditions are satisfied:*[17]

$$\varphi_a = \beta \tag{13}$$

$$l_a = \frac{\epsilon_1}{\epsilon_2\theta_1}(\varphi_a - 1) \tag{14}$$

$$l_x + l_A\frac{\kappa_x}{\kappa_A+\kappa_x+\kappa_y} = \frac{\epsilon_1}{\epsilon_2\theta_1}\varphi_x \tag{15}$$

$$l_y + l_A\frac{\kappa_y}{\kappa_A+\kappa_x+\kappa_y} = \frac{\epsilon_1}{\epsilon_2\theta_1}\varphi_y \tag{16}$$

[16]Keep in mind that, as in the baseline model, a tax on firm output is equivalent to a tax on firm total employment or payroll, a tax on household labor income, or household consumption, or any other tax that has a uniform impact across the two employment choices. Therefore, the results we present here apply more generally to the tax wedge induced by any combination of these taxes.

[17]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 the business-cycle properties of allocations.

(ii) When instead prices are flexible, a pair of strategies as in (11) can be implemented if and only if the following condition is satisfied in addition to conditions (13)-(16):

$$l_A = \frac{\epsilon_1}{\epsilon_2 \theta_1} \frac{\varphi_x}{\beta} \frac{\kappa_A + \kappa_x + \kappa_y}{\kappa_x}. \tag{17}$$

This lemma plays a similar role as 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. Furthermore, it helps understand the extent to which the government can manipulate the decentralized use of information. Indeed, much alike in the baseline model, the government may always induce any φ_x and φ_y it may desire. This is true no matter whether prices are flexible or not: state-contingent taxes alone suffice for this. But when prices are flexible, monetary policy can also help manipulate the decentralized use of information.

The intuition for these results is simple. First, note that, 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 why l_a, l_x and l_y are related to φ_a, φ_x and φ_y as in (13)-(16).[18] Next, note that when prices are are flexible, once the taxes have been set to achieve the desired φ_x and φ_y, the government has no further control on l_A : the sensitivity of second-stage employment and production to the realized aggregate productivity is pinned down by equating the realized marginal returns and costs of stage-2 labor. It is this restriction that gives (17). But when prices are sticky, this restriction is no more present: by designing the extent to which monetary policy accommodates the realized productivity shock, the government can freely choose l_A. Of course, anticipating the monetary policy's response to the realized productivity, firms adjust their own response to the information they have about aggregate productivity when they set their prices—which explains why l_x and l_y are negatively related to l_A in the way defined by conditions (15) and (16).

To see this more clearly, consider the equilibrium prices that are associated with any given implementable allocation. Using the facts that $\log p(\omega) - \log P(\omega) = -\frac{1}{\rho}(\log c(\omega, \Omega) - \log C(\Omega))$ and $\log c(\omega, \Omega) = \log q(\omega) + \theta_2 \log l(\omega, \Omega)$ along with (11), we infer that the equilibrium prices satisfy

$$\log p(\omega) = \psi_0 + \psi_a a + \psi_x x + \psi_y y,$$

[18] In particular, the equality of expected marginal rates of transformation and substitution gives $\mathbb{E}[\log l(\omega, \Omega)] = const + \frac{\epsilon_1 \theta_2}{\epsilon_2 \theta_1} \log n(\omega)$, where *const* includes second-order terms; using then $\log q(\omega) = \log A(\omega) + \theta_1 \log n(\omega)$ and noting that this condition must be satisfied for every $\omega = (a, x, y)$, gives the constraints in part (i) of Proposition 6.

where ψ_a and ψ_x are pinned down by

$$\psi_a = -\frac{1}{\rho}(\varphi_a + \theta_2 l_a) \qquad \text{and} \qquad \psi_x = -\frac{1}{\rho}(\varphi_x + \theta_2 l_x),$$

while ψ_y is indeterminate.[19] Saying that monetary policy can control l_x, the sensitivity of second-stage allocations to local information, is thus synonymous to saying that monetary policy can control ψ_x, the sensitivity of prices to local information. In particular, using (15), we get that 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.

Optimal Allocations

We now proceed to identify the best implementable allocation within our model. We use Proposition 4 to express the planner's problem as follows.

Restricted Planner's Problem. *Choose strategy coefficients* $\varphi = (\varphi_0, \varphi_a, \varphi_x, \varphi_y)$ *and* $l = (l_0, l_A, l_a, l_x, l_y)$ *and a precision* κ_y *so as to maximize* $\mathcal{W}(\varphi, l; \kappa_x, \kappa_y)$ *subject to (12) and (13)-(16).*

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 in conditions (13)-(16). But unlike standard Ramsey exercises, the allocations have also been restricted to respect the dispersion of information that we take as a primitive; this restriction is already embedded in the assumption that the strategy q depends only on locally available information. Finally, the planner takes into account that different allocations induce different information structures according to condition (12).

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

Proposition 5. *There exist scalars* $\Delta_q > 0$ *and* $\Delta_p > 0$ *such that the following are true:*

(i) The first-stage production strategy is given by

$$\log q(\omega) = \varphi_0^* + \varphi_a^* a + \varphi_x^* x + \varphi_y^* y$$

[19] That is, the sensitivities of local prices to local productivity and local information are pinned down by the elasticity of demand and the corresponding sensitivities of local output, while the sensitivity of prices to public information is indeterminate, simply because any component of monetary policy that is public information at the moment prices are set cannot have any real effect.

where

$$\varphi_a^* = \beta \tag{18}$$
$$\varphi_x^* = \left\{\frac{(1-\alpha)\kappa_x^*}{(1-\alpha)\kappa_x^* + \kappa_y^* + \kappa_A}\right\}\frac{\alpha}{1-\alpha}\beta + (\Delta_q + \Delta_p) \tag{19}$$
$$\varphi_y^* = \left\{\frac{\kappa_y^*}{(1-\alpha)\kappa_x^* + \kappa_y^* + \kappa_A}\right\}\frac{\alpha}{1-\alpha}\beta - \frac{\kappa_y^*}{\kappa_A + \kappa_y^*}(\Delta_q + \Delta_p) \tag{20}$$

(ii) The second-stage production strategy is given by

$$\log l\,(\omega, \Omega) = l_0^* + l_A^* \bar{a} + l_a^* a + l_x^* x + l_y^* y$$

where

$$l_a^* = \hat{l}_a \tag{21}$$
$$l_x^* = \hat{l}_x + \left(\frac{\kappa_x}{\kappa_A + \kappa_x + \kappa_y^*}\right)(\lambda_q \Delta_q + \lambda_p \Delta_p) \tag{22}$$
$$l_y^* = \hat{l}_y + \left(\frac{\kappa_y^*}{\kappa_A + \kappa_x + \kappa_y^*}\right)(\lambda_q \Delta_q + \lambda_p \Delta_p) \tag{23}$$
$$l_A^* = \hat{l}_A - (\lambda_q \Delta_q + \lambda_p \Delta_p) \tag{24}$$

where $(\hat{l}_a, \hat{l}_x, \hat{l}_y, \hat{l}_A)$ *are the coefficients that would have obtained in the flexible-price equilibrium when taxes are such that the first-stage production strategy is the optimal one, and where* (λ_q, λ_p) *are positive scalars.*

This result establishes that the impact of learning on the optimal implementable allocations is qualitatively very similar to the one in the baseline model. In particular, the scalars Δ_q and Δ_p are the Lagrange multipliers that measure the social value of increasing the precisions of, respectively, the quantity signal y^q and the price signal y^p. Using similar techniques as in the proof of Proposition 4, we can prove that both Δ_q and Δ_p are positive. Part (i) then shows that it is optimal to increase φ_x relative to the one that would characterize the flexible-price equilibrium allocations in the absence of taxes. By Proposition 6 we know that this increase in φ_x would have been associated with an increase in both l_x and l_A if monetary policy were replicating the flexible-price equilibrium. Part (ii) establishes that it is actually optimal to increase l_x further than that, which in turn is possible only by reducing l_A. We will further discuss how these results translate in terms of the optimal mixture of fiscal and monetary policies in the next subsection. Before doing that, we discuss what are the key forces that drive these results.

If the government could use a differential tax on the two types of labor in addition to the uniform output tax, then the government would have implemented the efficient allocation. The latter would feature the same qualitative properties, except for two differences. First, the precise values of Δ_q and Δ_p would be different. And second, whereas the optimal allocation features $l_A^* + l_x^* < \hat{l}_1 + \hat{l}_x$ and $l_y^* > \hat{l}_y$, the efficient allocation would feature $l_A^* + l_x^* = \hat{l}_A + \hat{l}_x$ and $l_y^* = \hat{l}_y$. To understand this, note that the planner wishes to increase l_x in order to induce more learning. This is true no matter whether a differential tax is available or not. The difference lies on what it takes to achieve this goal. Along the best implementable allocation, by the implementability constraints in Proposition 4 we know that increasing l_x is possible only by reducing l_A more than one-to-one, as well as that this reduction in l_A in turn necessitates an increase in l_y. But when a differential tax is available, these constraints are no longer binding. Rather, the planner can decrease l_A one-to-one with the desired increase in l_x so as to preserve allocative efficiency (in the sense of maintaining the efficient overall sensitivity of second-stage employment, and hence of output and consumption, to aggregate productivity). Moreover, there is clearly no reason to distort l_y.

Finally, if the government could use only output taxes and prices were flexible, then part (i) would continue to hold (with different values for Δ_q and Δ_p), but of course now it would necessarily be the case that $l_A^* = \hat{l}_A, l_x^* = \hat{l}_x, l_y^* = \hat{l}_y$, simply because there would be no instrument that would permit the planner to do otherwise.

We conclude that the nature of the optimal allocation—and hence of the policies that implement it—is driven by, and only by, the combination of three properties of the environment: the presence of informational externalities, which makes it desirable to manipulate the decentralized use of information in a similar manner as in the baseline model; second, the fact the prices are sticky, which gives monetary policy the power to contribute to this goal along with the state-contingent tax; and finally, the absence of a differential tax, which makes it optimal for monetary policy to exercise this power, thus deviating from the principle of replicating the flexible-price equilibrium.

Optimal Policy

Finally we characterize the optimal combination of fiscal and monetary policy. In order to do this, we simply translate the results in Propositions 2 and 5 in terms of the policies that implement the optimal allocation. This yields the following result.

Proposition 6. *(i) The optimal tax is countercyclical, as in the baseline model.*

(ii) The optimal monetary policy is less accommodative of the productivity shock—in the sense that it induces aggregate income to react less to the productivity shock—than the policy

that would have replicated the flexible-price allocations.

Part (i) of Theorem 6 describes optimal fiscal policy while part (ii) characterizes optimal monetary policy.

The intuition stated above for part (i) of Theorem 6 in terms of the baseline model extends to the monetary model. To understand part (ii), recall from Proposition 5 that the optimal allocation satisfies $l_a^* = \hat{l}_a$ and that $l_A^* + l_x^* < \hat{l}_A + \hat{l}_x$ (where the hats indicate the values that would obtain in the flexible-price equilibrium in which taxes are such that the first-state strategy is optimal). Note then that the overall sensitivity of aggregate income to aggregate productivity along the optimal allocation is given by $\varphi_a^* + \varphi_x^* + \theta_2(l_a^* + l_x^* + l_A^*)$, while the corresponding sensitivity of the flexible-price equilibrium in which the first-state strategy coincides with the optimal one is given by $\varphi_a^* + \varphi_x^* + \theta_2(\hat{l}_a + \hat{l}_x + \hat{l}_A)$. It follows that aggregate income responds less to aggregate productivity in the optimal allocation than what would be consistent with the flexible-price equilibrium.

Finally, again consider part (i) and suppose for a moment that monetary policy were replicating the flexible-price equilibrium. From our discussion above, we know that in this case the optimal fiscal policy is a countercyclical tax in order to induce a higher φ_x. But now note that if the optimal monetary policy is less accommodative than implementing flexible prices, this induces the opposite effect: other things equal, a less accommodative policy implies a lower incentive to react to local information about aggregate productivity. It follows that the optimal tax necessarily becomes even more countercyclical; if it were not, the monetary policy would have induced a lower φ_x, which is a contradiction.

We conclude that the basic intuition for the optimality of making monetary policy less accommodative of the productivity shock is that this induces more learning through price signals.

References

[1] Adam, Klaus (2007), "Optimal Monetary Policy with Imperfect Common Knowledge," *Journal of Monetary Economics* 54(2), 276-301.

[2] Angeletos, George-Marios, and Jennifer La'O (2009), "Noisy Business Cycles," *NBER Macroeconomics Annual 2009*, 24, 319-378.

[3] Angeletos, George-Marios, and Jennifer La'O (2012), "Optimal Monetary Policy with Informational Frictions." MIT/Columbia Mimeo.

[4] Angeletos, George-Marios, and Alessandro Pavan (2007), "Efficient Use of Information and Social Value of Information," *Econometrica* 75(4), 1103-1142.

[5] Angeletos, George-Marios, and Alessandro Pavan (2008), "Policy with Dispersed Information," *Journal of the European Economic Association* 7, 11-60.

[6] Ball, Laurence, Gregory N. Mankiw, and Ricardo Reis (2005) "Monetary policy for inattentive economies," *Journal of Monetary Economics* 52(4), 703-725.

[7] Chari, Varadarajan V., and Patrick J. Kehoe (1999), "Optimal fiscal and monetary policy," in J. B. Taylor & M. Woodford, eds., *Handbook of Macroeconomics*, vol. 1, chapter 26, 1671-1745, Elsevier.

[8] Colombo, Luca, Gianluca Femminis, and Alessandro Pavan. 2012. "Information Acquisition and Welfare." Mimeo, Northwestern University.

[9] Laffont (1985), "On the Welfare Analysis of Rational Expectations Equilibria with Asymmetric Information," *Econometrica*, 53, 1–30.

[10] Llosa, Luis and Venky Venkateswaran (2012), "Efficiency with Endogenous Information Choice," NYU Stern mimeo.

[11] Lorenzoni, Guido (2010), "Optimal Monetary Policy with Uncertain Fundamentals and Dispersed Information." *Review of Economic Studies* 77, 305-38.

[12] Lucas, Robert E., Jr., and Nancy Stokey (1983), "Optimal fiscal and monetary policy in an economy without capital," *Journal of Monetary Economics* 12(1), 55-93.

[13] Mackowiak, Bartosz, and Mirko Wiederholt (2011), "Inattention to Rare Events." Mimeo, European Central Bank/Nothwestern University.

[14] Messner, S., and Xavier Vives (2001), "Informational and Economic Efficiency in REE with Asymmetric Information," Working Paper 2678, CEPR.

[15] Paciello, Luigi, and Mirko Wiederholt (2013), "Exogenous Information, Endogenous Information, and Optimal Monetary Policy," *Review of Economic Studies*, forthcoming.

[16] Vives, Xavier (1988): "Aggregation of Information in Large Cournot Markets," *Econometrica*, 56, 851–876.

[17] Vives, Xavier (2013), "Endogenous Public Information and Welfare in Market Games," IESE Business School mimeo.

Appendix

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

$$y^q = \log Q(\Omega) + \varepsilon^{yq}$$

where

$$\log Q(\Omega) = \varphi_0' + \varphi_a a + \varphi_x x + \varphi_y y$$

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

$$\tilde{y}^q \equiv \frac{y^q - \varphi_0' - \varphi_y y}{\varphi_a + \varphi_x} = \bar{a} + \tilde{\varepsilon}^{yq} \tag{25}$$

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

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

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

$$y = \frac{\kappa_{ya}}{\kappa_y} y^a + \frac{\kappa_{yq}}{\kappa_y} \tilde{y}^q,$$

where κ_{ya} and κ_{yq} are the precisions of these two signals, while $\kappa_y = \kappa_{ya} + \kappa_{yq}$ is the overall precision of the sufficient statistic y.

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

Proof of Lemma 2 Take an arbitrary log-linear strategy of the form $q(a, x, y) = \varphi_0 + \varphi_a a + \varphi_x x + \varphi_y y$. For any coefficients $\varphi = (\varphi_0, \varphi_a, \varphi_x, \varphi_y)$ and any precisions $\kappa = (\kappa_x, \kappa_y)$, the implied level of welfare (ex-ante utility) can be expressed as follows:

$$\mathbb{E}u = \mathcal{W}(\varphi; \kappa) \equiv \frac{1}{1-\gamma} \exp V_c(\varphi; \kappa) - \frac{1}{\epsilon} \exp V_n(\varphi; \kappa), \tag{26}$$

where

$$\begin{aligned} V_c(\varphi; \kappa) \;&\equiv\; (1-\gamma)\left(\varphi_0 + (\varphi_a + \varphi_x + \varphi_y)\mu\right) \\ &+ \frac{1}{2}(1-\gamma)\left(\frac{\rho - 1}{\rho}\right)\left[\frac{\varphi_a^2}{\kappa_\xi} + \frac{\varphi_x^2}{\kappa_x} + 2\frac{\varphi_a \varphi_x}{\kappa_x}\right] \\ &+ \frac{1}{2}(1-\gamma)^2 \left[\frac{\varphi_y^2}{\kappa_y} + \frac{(\varphi_a + \varphi_x + \varphi_y)^2}{\kappa_A}\right] \end{aligned} \tag{27}$$

$$\begin{aligned} V_n(\varphi; \kappa) \;&\equiv\; \frac{\epsilon}{\theta}\left(\varphi_0 + (\varphi_a + \varphi_x + \varphi_y - 1)\mu\right) \\ &+ \frac{1}{2}\frac{\epsilon^2}{\theta^2}\left[\frac{(\varphi_a - 1)^2}{\kappa_\xi} + \frac{\varphi_x^2}{\kappa_x} + 2\frac{(\varphi_a - 1)\varphi_x}{\kappa_x} + \frac{\varphi_y^2}{\kappa_y} + \frac{(\varphi_a + \varphi_x + \varphi_y - 1)^2}{\kappa_A}\right] \end{aligned} \tag{28}$$

Proof of Lemma 3 We consider a combination of the following tax instruments: a linear tax $\tau^R(\Omega)$ on firm revenue, a linear tax $\tau^L(\Omega)$ on household labor income, and a linear tax $\tau^C(\Omega)$ 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 the $(\bar{a}, Q, y)$:

$$\begin{aligned} -\log(1 - \tau^R(\Omega)) &= \tau_0^R + \tau_A^R \bar{a} + \tau_Q^R \log Q(\Omega) + \tau_y^R y, \\ -\log(1 - \tau^L(\Omega)) &= \tau_0^L + \tau_A^L \bar{a} + \tau_Q^L \log Q(\Omega) + \tau_y^L y, \\ \log(1 + \tau^C(\Omega)) &= \tau_0^C + \tau_A^C \bar{a} + \tau_Q^C \log Q(\Omega) + \tau_y^C y. \end{aligned}$$

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

$$\pi(\omega, \Omega) = \left(1 - \tau^R(\Omega)\right) p(\omega, \Omega) q(\omega) - w(\omega) n(\omega),$$

while the budget constraint of the household is given by

$$(1+\tau^C(\Omega))\int p(\omega,\Omega)c(\omega)d\Omega(\omega)=\int \pi(\omega,\Omega)d\Omega(\omega)+(1-\tau^L(\Omega))\int w(\omega)n(\omega)d\Omega(\omega)+T(\Omega)$$

where $T(\Omega)$ 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

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

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

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

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:

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

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

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

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

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

Equivalently, the tax wedge is given by (6) with $\tau_0\equiv-\log\chi+\tau_0^R+\tau_0^C+\tau_0^L$, $\tau_A\equiv\tau_A^R+\tau_A^C+\tau_A^L$, $\tau_Q\equiv\tau_Q^R+\tau_Q^C+\tau_Q^L$, and $\tau_y\equiv\tau_y^R+\tau_y^C+\tau_y^L$.

Proof of Proposition 1 As stated in Lemma 3, the equilibrium strategy must solve the following fixed point:

$$q(\omega)^{\frac{\epsilon}{\theta}+\frac{1}{\rho}-1} = \left(\frac{\rho-1}{\rho}\right)\theta A(\omega)^{\frac{\epsilon}{\theta}}\mathbb{E}\left[\exp(-\tau_0-\tau_A\bar{a}-\tau_y y)Q(\Omega)^{\frac{1}{\rho}-\gamma-\tau_Q}|\omega\right] \tag{29}$$

with $Q(\Omega)=\left[\int q(\omega)^{\frac{\rho-1}{\rho}}d\Omega(\omega)\right]^{\frac{\rho}{\rho-1}}$. It follows that the equilibrium strategy is given by

$$\log q(\omega) = \hat{\varphi}_0(\tau)+\hat{\varphi}_a(\tau)\,a+\hat{\varphi}_x(\tau)\,x+\hat{\varphi}_y(\tau)\,y \tag{30}$$

where

$$\hat{\varphi}_a(\tau) = \beta \tag{31}$$

$$\hat{\varphi}_x(\tau) = \left(1-\frac{\theta}{\epsilon\hat{\alpha}}\tau_A\right)\left(\frac{(1-\hat{\alpha})\kappa_x}{(1-\hat{\alpha})\kappa_x+\kappa_y+\kappa_A}\right)\frac{\hat{\alpha}}{1-\hat{\alpha}}\beta \tag{32}$$

$$\hat{\varphi}_y(\tau) = \frac{1}{1-\hat{\alpha}}\left(\frac{\kappa_y}{\kappa_x}\hat{\varphi}_x(\tau)-\frac{\beta\theta}{\epsilon}\tau_y\right) \tag{33}$$

$$\hat{\varphi}_0(\tau) = \frac{1}{\frac{\epsilon}{\theta}+\gamma-1+\tau_Q}\left[-\tau_0+\left(-\tau_A+\left(\frac{1}{\rho}-\gamma-\tau_Q\right)(\hat{\varphi}_a+\hat{\varphi}_x)\right)\frac{\kappa_A}{\kappa_A+\kappa_x+\kappa_y}\mu\right. \tag{34}$$

$$+\left(\frac{1}{\rho}-\gamma-\tau_Q\right)\left(\frac{\rho-1}{\rho}\right)\frac{(\hat{\varphi}_a+\hat{\varphi}_x)^2}{2}\sigma_x^2+\frac{1}{2}\left(\frac{1}{\rho}-\gamma-\tau_Q\right)^2(\hat{\varphi}_a+\hat{\varphi}_x)^2\sigma_0^2 \tag{35}$$

$$\left.+\frac{1}{2}\tau_A^2\sigma_0^2-\tau_A\left(\frac{1}{\rho}-\gamma-\tau_Q\right)(\hat{\varphi}_a+\hat{\varphi}_x)\sigma_0^2+\log\left(\theta\frac{\rho-1}{\rho}\right)\right] \tag{36}$$

and where

$$\hat{\alpha}=\hat{\alpha}(\tau_Q)\equiv\frac{\frac{1}{\rho}-\gamma-\tau_Q}{\frac{\epsilon}{\theta}+\frac{1}{\rho}-1}=\alpha-\frac{\tau_Q}{\frac{\epsilon}{\theta}+\frac{1}{\rho}-1} \tag{37}$$

represents the equilibrium degree of complementarity induced by the policy.

We now prove the claims in the proposition. Pick an arbitrary strategy for which $\varphi_a=\beta$ and let $(\varphi_0^{\#},\varphi_x^{\#},\varphi_y^{\#})$ denote the remaining coefficients. Set $\tau_Q=0$, in which case $\hat{\alpha}=\alpha$. From condition (32), there is a unique value for τ_A that induces $\hat{\varphi}_x(\tau)=\varphi_x^{\#}$; this is given by

$$\tau_A=\frac{\epsilon}{\theta}\left\{\alpha-\varphi_x^{\#}\frac{(1-\alpha)\kappa_x+\kappa_y+\kappa_A}{\beta\kappa_x}\right\}. \tag{38}$$

From (33), there is then a unique value for τ_y that induces $\hat{\varphi}_y(\tau)=\varphi_y^{\#}$; this is given by

$$\tau_y=\frac{\epsilon}{\beta\theta}\left\{\frac{\kappa_y}{\kappa_x}\varphi_x^{\#}-(1-\alpha)\varphi_y^{\#}\right\}. \tag{39}$$

Combining these two results, we have a unique pair (τ_A, τ_y) that induces the desired $(\varphi_x^{\#}, \varphi_y^{\#})$. But then from (34) there is also a unique τ_0 that induces $\hat{\varphi}_0(\tau) = \varphi_0^{\#}$. This proves the desired strategy can always be implemented with a policy that allows $\tau_y \neq 0$.

Now restrict $\tau_y = 0$ and suppose $\varphi_x^{\#}, \varphi_y^{\#} \neq 0$. From condition (37), there always exists a unique τ_Q such that

$$(1 - \hat{\alpha}(\tau_Q)) \frac{\kappa_x}{\kappa_y} = \frac{\varphi_x^{\#}}{\varphi_y^{\#}}$$

and hence such that $\hat{\varphi}_x(\tau) / \hat{\varphi}_y(\tau) = \varphi_x^{\#}/\varphi_y^{\#}$. Given this τ_Q, condition (32) gives a unique τ_A such that $\hat{\varphi}_x(\tau) = \varphi_x^{\#}$, which along with previous result gives also $\hat{\varphi}_y(\tau) = \varphi_y^{\#}$. Finally, given these values for (τ_Q, τ_A), there is then a unique τ_0 that ensures $\hat{\varphi}_0(\tau) = \varphi_0^{\#}$. This proves that, as long as $\varphi_x^{\#}, \varphi_y^{\#} \neq 0$, the policy can be contingent only on aggregate productivity and aggregate output.

Proof of Proposition 2 *Part (i).* Recall from Lemma 1 that any given strategy induces a κ_x and a κ_y as functions of $\varphi_a + \varphi_x$; let $\kappa_x(\varphi_a + \varphi_x)$ and $\kappa_y(\varphi_a + \varphi_x)$ denotes these functions. We can then express the planner's problem as follows:

Planner's problem. *Choose* $\varphi = (\varphi_0, \varphi_a, \varphi_x, \varphi_y)$ *and* (κ_x, κ_y) *so as to maximize* $W(\varphi, \kappa)$ *subject to* $\kappa_x = \kappa_x(\varphi_a + \varphi_x)$ *and* $\kappa_y = \kappa_y(\varphi_a + \varphi_x)$.

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_y are kept constant at $\kappa_x = \kappa_x(\bar{\varphi})$ and $\kappa_y = \kappa_y(\bar{\varphi})$. The second step is to optimize over the sum $\bar{\varphi}$ and the precision κ_x and κ_y subject to the constraint that $\kappa_x = \kappa_x(\bar{\varphi})$ and $\kappa_y = \kappa_y(\bar{\varphi})$. 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, as well as 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 $\kappa = (\kappa_x(\bar{\varphi}), \kappa_y(\bar{\varphi}))$, 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:

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

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

$$\varphi_0 \;:\; 0 = \exp V_c(\varphi, \kappa) - \frac{1}{\theta} \exp V_n(\varphi, \kappa). \tag{40}$$

This guarantees that $V_c = V_n - \log\theta$ at the efficient allocation and gives φ_0 as a function of $\varphi_a, \varphi_x, \varphi_y, \kappa_x, \kappa_y$ 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:

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

For fixed η, this is a linear system of three equations in the three coefficients φ_a, φ_x and φ_y. 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_y as follows:

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

Letting

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

gives conditions (8) and (9). Finally, note that Δ is just a 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:

$$\Delta=\bar{\varphi}-\left\{\frac{\kappa_x+\kappa_y+\kappa_A}{(1-\alpha)\kappa_x+\kappa_y+\kappa_A}\right\}\beta. \tag{41}$$

Using this into conditions (8) and (9), we can obtain the optimal coefficients as functions of the sum $\bar{\varphi}$ and the precisions κ_x and κ_y. 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 also as function of the sum $\bar{\varphi}$ and the precisions κ_x and κ_y. In particular, using the FOC with respect to φ_0, we get that

$$\mathcal{W}(\varphi,\kappa)=\left(\frac{\frac{\epsilon}{\theta}-1+\gamma}{1-\gamma}\frac{\theta}{\epsilon}\right)\exp V_c(\varphi,\kappa) \tag{42}$$

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

$$\mathcal{TW}(\varphi(\bar{\varphi},\kappa),\kappa)=W(\bar{\varphi},\kappa)\equiv A(\kappa)-B(\kappa)\left(\bar{\varphi}-f(\kappa)\right)^2 \tag{43}$$

where

$$B(\kappa)\equiv\frac{\epsilon}{2\theta(1-\alpha)}\frac{\kappa_0+(1-\alpha)\kappa_x+\kappa_y}{\kappa_x(\kappa_A+\kappa_y)}>0$$

and

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

(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, (43) 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 $\kappa_x = \kappa_x(\bar{\varphi})$ and $\kappa_y = \kappa_y(\bar{\varphi})$. Letting

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

the planner's problem reduces to the following unidimensional 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

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

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

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

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

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

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 $\kappa_y \to \infty$, 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

$$\frac{\partial W}{\partial \bar{\varphi}} + \frac{\partial W}{\partial \kappa_y}\frac{\partial \kappa_y}{\partial \bar{\varphi}} + \frac{\partial W}{\partial \kappa_x}\frac{\partial \kappa_x}{\partial \bar{\varphi}} = 0. \tag{44}$$

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 (41) 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 (44). Indeed, combining these two conditions gives

$$\eta = \frac{\partial W}{\partial \kappa_y}\frac{\partial \kappa_y}{\partial \bar{\varphi}} + \frac{\partial W}{\partial \kappa_x}\frac{\partial \kappa_x}{\partial \bar{\varphi}},$$

which means that the Lagrange multiplier measures the social value of increasing the precision of available public information by increasing the sensitivity of allocations to local information.

Proof of Proposition 1 This follows directly from Proposition 2 and from the property that $\Delta > 0$ in the optimal allocation but $\Delta = 0$ in the equilibrium allocation without policy.

Proof of Proposition 3 Without any loss of generality, set $\tau_Q = 0$ and express the tax as a function of $(\bar{a}, y)$ alone. From conditions (38) and (39) in the proof of Proposition 1, we have that the optimal tax satisfies

$$\begin{aligned}
\tau_A^* &= \frac{\epsilon}{\theta}\left(\alpha - \varphi_x^* \frac{(1-\alpha)\kappa_x + \kappa_y + \kappa_A}{\beta\kappa_x}\right) \\
\tau_y^* &= \frac{\epsilon}{\beta\theta}\left(\frac{\kappa_y}{\kappa_x}\varphi_x^* - (1-\alpha)\varphi_y^*\right)
\end{aligned}$$

Using the characterization of φ_x^* and φ_y^* from Proposition 2, we get

$$\tau_A^* = -\lambda\Delta \qquad \text{and} \qquad \tau_y^* = \frac{\kappa_y}{\kappa_A + \kappa_y}\lambda\Delta, \tag{45}$$

where

$$\lambda \equiv \frac{\varepsilon}{\beta\theta}\frac{(1-\alpha)\kappa_x + \kappa_y + \kappa_A}{\kappa_x} > 0.$$

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

To interpret this result, note first that $y = \bar{a} + \varepsilon$, where ε is noise. The property that $\tau_A < 0$ means the that tax is negatively correlated with aggregate productivity for given common belief y; that is, it is negatively correlated with the surprise component in realized aggregate productivity. At the same time, the property that $\tau_a + \tau_y < 0$ 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. Next, the property that $\tau_y > 0$ 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

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

their covariance is given by

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

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

$$Cov(-\log(1-\tau), \log Q) = (\tau_A^* + \tau_y^*)(\varphi_a^* + \varphi_x^*)\frac{1}{\kappa_A} + \left\{\tau_A^*\frac{1}{\kappa_A} + \tau_y^*\frac{\kappa_0 + \kappa_y}{\kappa_A\kappa_y}\right\}\varphi_y^*.$$

By (45), the last term is necessarily zero. Next, note that $\varphi_a^* + \varphi_x^*$ is necessarily positive, while $\tau_A^* + \tau_y^*$ is necessarily negative. We conclude that the tax is negatively correlated with aggregate output.

Proof of Proposition ?? We first solve for the planner's choice of κ_x. As in the proof of Proposition 2, we can write welfare as

$$\mathcal{W}(\varphi;\kappa) \equiv \frac{1}{1-\gamma}\exp V_c(\varphi;\kappa) - \frac{1}{\epsilon}\exp V_n(\varphi;\kappa),$$

where $V_c(\varphi;\kappa)$ and $V_n(\varphi;\kappa)$ are functions defined in (27) and (28). Taking the derivative of welfare with respect to κ_x yields the following expression

$$\frac{\partial \mathcal{W}}{\partial \kappa_x} = \frac{1}{1-\gamma}\frac{\partial \exp V_c}{\partial \kappa_x} - \frac{1}{\epsilon}\frac{\partial \exp V_n}{\partial \kappa_x}$$

where

$$\begin{aligned}
\frac{\partial \exp V_c}{\partial \kappa_x} &= \exp V_c\left[\frac{1}{2}(1-\gamma)\left(\frac{\rho-1}{\rho}\right)\left(-\frac{\varphi_x^2}{\kappa_x^2} - 2\frac{\varphi_a\varphi_x}{\kappa_x^2}\right)\right] \\
\frac{\partial \exp V_n}{\partial \kappa_x} &= \exp V_n\left[\frac{1}{2}\frac{\epsilon^2}{\theta^2}\left(-\frac{\varphi_x^2}{\kappa_x^2} - 2\frac{(\varphi_a-1)\varphi_x}{\kappa_x^2}\right)\right]
\end{aligned}$$

We can thus write the partial derivative of welfare as follows.

$$\frac{\partial \mathcal{W}}{\partial \kappa_x} = \frac{1}{1-\gamma}\exp V_c\left[\frac{1}{2}(1-\gamma)\left(\frac{\rho-1}{\rho}\right)\left(-\frac{\varphi_x^2}{\kappa_x^2} - 2\frac{\varphi_a\varphi_x}{\kappa_x^2}\right)\right] - \frac{1}{\epsilon}\exp V_n\left[\frac{1}{2}\frac{\epsilon^2}{\theta^2}\left(-\frac{\varphi_x^2}{\kappa_x^2} - 2\frac{(\varphi_a-1)\varphi_x}{\kappa_x^2}\right)\right] \tag{46}$$

Next, from the FOC condition of the planner's problem with respect to φ_0 (40), we have that

$$\exp V_c(\varphi,\kappa) - \frac{1}{\theta}\exp V_n(\varphi,\kappa) = 0$$

Substituting this into (46), we obtain the following simplified expression for the partial derivative of welfare

$$\frac{\partial \mathcal{W}}{\partial \kappa_x} = \left\{\frac{\epsilon}{\theta}\frac{1}{2}\left(\frac{\rho-1}{\rho}\right)\left[-\frac{\varphi_x^2}{\kappa_x^2} - 2\frac{\varphi_a\varphi_x}{\kappa_x^2}\right] - \frac{1}{2}\frac{\epsilon^2}{\theta^2}\left[-\frac{\varphi_x^2}{\kappa_x^2} - 2\frac{(\varphi_a-1)\varphi_x}{\kappa_x^2}\right]\right\}\frac{1}{\epsilon}\exp V_n \tag{47}$$

The planner's optimal choice of κ_x occurs where $\partial\mathcal{W}/\partial\kappa_x = \partial\mathcal{C}/\partial\kappa_x$.

Next, we can do a similar exercise to characterize the firm's optimal choice of κ_x. The firm's objective is given by

$$\mathcal{U} = \frac{U'\left(Q(\Omega)\right)}{P(\Omega)} \Pi\left(\omega, \Omega\right)$$

We can rewrite this as

$$\mathcal{U} = U'\left(Q(\Omega)\right) Q(\Omega)^{\frac{1}{\rho}} q(\omega)^{1-\frac{1}{\rho}} - n(\omega)^{\epsilon} = Q(\Omega)^{1-\gamma}\left(\frac{q(\omega)}{Q(\Omega)}\right)^{\frac{\rho-1}{\rho}} - \left(\frac{q(\omega)}{A(\omega)}\right)^{\frac{\epsilon}{\theta}}$$

Again using the log-linear strategies, we have that

$$\mathcal{U}\left(\phi, \kappa_i, \kappa_{-i}\right) \equiv \exp \tilde{V}_c\left(\phi, \kappa_i, \kappa_{-i}\right) - \exp \tilde{V}_n\left(\phi, \kappa_i, \kappa_{-i}\right),$$

where[20]

$$\begin{aligned}
\tilde{V}_c(\varphi;\kappa) \quad &\equiv \quad (1-\gamma)\left(\varphi_0' + (\varphi_a + \varphi_x + \varphi_y)\mu\right) + \frac{1}{2}(1-\gamma)^2\left[\frac{\varphi_y^2}{\kappa_y} + \frac{(\varphi_a + \varphi_x + \varphi_y)^2}{\kappa_A}\right] \\
&\quad + \frac{1}{2}\left(\frac{\rho-1}{\rho}\right)^2\left[\frac{\varphi_a^2}{\kappa_\xi} + \frac{\varphi_x^2}{\kappa_x} + 2\frac{\varphi_a\varphi_x}{\kappa_x}\right] + \left(\frac{\rho-1}{\rho}\right)(\varphi_0 - \varphi_0')
\end{aligned}$$

$$\begin{aligned}
\tilde{V}_n(\varphi;\kappa) \quad &\equiv \quad \frac{\epsilon}{\theta}\left(\varphi_0 + (\varphi_a + \varphi_x + \varphi_y - 1)\mu\right) \\
&\quad + \frac{1}{2}\frac{\epsilon^2}{\theta^2}\left[\frac{(\varphi_a - 1)^2}{\kappa_\xi} + \frac{\varphi_x^2}{\kappa_x} + 2\frac{(\varphi_a - 1)\varphi_x}{\kappa_x} + \frac{\varphi_y^2}{\kappa_y} + \frac{(\varphi_a + \varphi_x + \varphi_y - 1)^2}{\kappa_A}\right]
\end{aligned}$$

Taking the derivative of $\mathcal{U}\left(\phi, \kappa_i, \kappa_{-i}\right)$ with respect to κ_x yields the following expression

$$\frac{\partial \mathcal{U}}{\partial \kappa_{x,i}} = \frac{\partial \exp \tilde{V}_c}{\partial \kappa_{x,i}} - \frac{\partial \exp \tilde{V}_n}{\partial \kappa_{x,i}}$$

where

$$\begin{aligned}
\frac{\partial \exp \tilde{V}_c}{\partial \kappa_{x,i}} \quad &= \quad \exp \tilde{V}_c\left[\frac{1}{2}\left(\frac{\rho-1}{\rho}\right)^2\left(-\frac{\varphi_x^2}{\kappa_x^2} - 2\frac{\varphi_a\varphi_x}{\kappa_x^2}\right)\right] \\
\frac{\partial \exp \tilde{V}_n}{\partial \kappa_{x,i}} \quad &= \quad \exp \tilde{V}_n\left[\frac{1}{2}\frac{\epsilon^2}{\theta^2}\left(-\frac{\varphi_x^2}{\kappa_x^2} - 2\frac{(\varphi_a - 1)\varphi_x}{\kappa_x^2}\right)\right]
\end{aligned}$$

[20] Note that $\varphi_0' \equiv \varphi_0 + \left(\frac{\rho-1}{\rho}\right)\frac{(\varphi_a+\varphi_x)^2}{2}\left[\frac{\varphi_a^2}{\kappa_\xi} + \frac{\varphi_x^2}{\kappa_x} + 2\frac{\varphi_a\varphi_x}{\kappa_x}\right]$, but that the κ_x that shows up here is $\kappa_{x,-i}$ i.e. it is the dispersion in the aggregate output of all other agents, which agent i takes as given. Thus when we take the FOC with respect to $\kappa_{x,i}$ this term doesn't show up.

We can thus write the partial derivative of the firm's objective function as follows.

$$\frac{\partial \mathcal{U}}{\partial \kappa_{x,i}} = \exp \tilde{V}_c \left[\frac{1}{2} \left(\frac{\rho - 1}{\rho} \right)^2 \left[-\frac{\varphi_x^2}{\kappa_x^2} - 2\frac{\varphi_a \varphi_x}{\kappa_x^2} \right] \right] - \exp \tilde{V}_n \left[\frac{1}{2} \frac{\epsilon^2}{\theta^2} \left[-\frac{\varphi_x^2}{\kappa_x^2} - 2\frac{(\varphi_a - 1)\varphi_x}{\kappa_x^2} \right] \right] \quad (48)$$

Next we have from the firm's FOC with respect to φ_0, we have that

$$0 = \left(\frac{\rho - 1}{\rho} \right) \exp \tilde{V}_c - \frac{\epsilon}{\theta} \exp \tilde{V}_n$$

Substituting this into (48), we obtain the following simplified expression

$$\frac{\partial \mathcal{U}}{\partial \kappa_{x,i}} = \left\{ \frac{\epsilon}{\theta} \frac{1}{2} \left(\frac{\rho - 1}{\rho} \right) \left[-\frac{\varphi_x^2}{\kappa_x^2} - 2\frac{\varphi_a \varphi_x}{\kappa_x^2} \right] - \frac{1}{2} \frac{\epsilon^2}{\theta^2} \left[-\frac{\varphi_x^2}{\kappa_x^2} - 2\frac{(\varphi_a - 1)\varphi_x}{\kappa_x^2} \right] \right\} \exp \tilde{V}_n \quad (49)$$

The firm's optimal choice of $\kappa_{x,i}$ occurs where $\partial \mathcal{U} / \partial \kappa_{x,i} = \partial \mathcal{C} / \partial \kappa_{x,i}$.

We can thus use this expression for $\partial \mathcal{U} / \partial \kappa_{x,i}$ and compare it to our expression for $\partial \mathcal{W} / \partial \kappa_x$ in (47) . One can easily verify that $V_n = \tilde{V}_n$. Using this, we have that

$$\frac{\partial \mathcal{W}}{\partial \kappa_x} = \frac{1}{\epsilon} \frac{\partial \mathcal{U}}{\partial \kappa_{x,i}} > 0$$

Thus, these objects only differ by the Frisch elasticity constant $\epsilon \in [1, \infty)$. Using this expression, we can now prove parts (i) and (ii) of the proposition. For this, let $g_p(\kappa_x) \equiv \partial \mathcal{W} / \partial \kappa_x$, $g_{eq}(\kappa_x) \equiv \partial \mathcal{U} / \partial \kappa_x$ and $f(\kappa_x) \equiv \partial \mathcal{C} / \partial \kappa_x$. Next, let κ_x^* denote the value of κ_x for which $g_p(\kappa_x) = f(\kappa_x)$ and let κ_x^{eq} denote the value of κ_x for which $g_{eq}(\kappa_x) = f(\kappa_x)$; κ_x^* and κ_x^{eq} thus respectively denote the efficient and equilibrium levels of private information. Due to convexity of $\mathcal{C}$, we have that f is a monotonically increasing function of κ_x. Finally, in order to ensure that this problem is well-behaved and that such levels κ_x^* and κ_x^{eq} exist, we assume that $\mathcal{C}$ is sufficiently convex. Specifically we assume that $f' > g_p'$ and $f' > g_{eq}'$.[21]

Part (i). If $\epsilon = 1$, then for any value of κ_x, $g_p(\kappa_x) = g_{eq}(\kappa_x)$. By definition, κ_x^* is the value of κ_x for which $g_p(\kappa_x^*) = f(\kappa_x^*)$. But this implies that κ_x^* is also the value of κ_x for which $g_{eq}(\kappa_x) = f(\kappa_x)$. It is thus immediate that $\kappa_x^* = \kappa_x^{eq}$, i.e. the equilibrium collection of information is efficient.

Part (ii). If $\epsilon > 1$, then for any value of κ_x, $g_p(\kappa_x) < g_{eq}(\kappa_x)$. We now show that in this case, $\kappa_x^* < \kappa_x^{eq}$. We can prove this by contraction. Suppose it were instead the case that $\kappa_x^{eq} \leq \kappa_x^*$. Since f is monotonically increasing, we have that $f(\kappa_x^{eq}) \leq f(\kappa_x^*)$. This implies

[21] Note that $\epsilon g_p' = g_{eq}'$. In principle, g_p' and g_{eq}' can be either negative or positive (but both must be of the same sign).

that

$$g_{eq}\left(\kappa_x^{eq}\right) = f\left(\kappa_x^{eq}\right) \leq f\left(\kappa_x^*\right) = g_p\left(\kappa_x^*\right)$$

Combining this with the fact that $g_p\left(\kappa_x\right) < g_{eq}\left(\kappa_x\right)\ \forall\ \kappa_x$, we obtain the following set of inequalities: $g_{eq}\left(\kappa_x^{eq}\right) \leq g_p\left(\kappa_x^*\right) < g_{eq}\left(\kappa_x^*\right)$. But this implies that

$$\frac{g_{eq}\left(\kappa_x^*\right) - g_{eq}\left(\kappa_x^{eq}\right)}{\kappa_x^* - \kappa_x^{eq}} > \frac{g_p\left(\kappa_x^*\right) - g_{eq}\left(\kappa_x^{eq}\right)}{\kappa_x^* - \kappa_x^{eq}} = \frac{f\left(\kappa_x^*\right) - f\left(\kappa_x^{eq}\right)}{\kappa_x^* - \kappa_x^{eq}}$$

But this contradicts the assumption that $g_{eq}' < f'$. Thus, $\kappa_x^* \geq \kappa_x^{eq}$ cannot be possible, and hence it must be the case that $\kappa_x^* < \kappa_x^{eq}$, i.e. firms collect too much private information in equilibrium relative to the efficient level.

Proof of Proposition ?? This follows directly from immediately from Proposition **??**.

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

$$-\rho\left(\log p(\omega) - \log P(\Omega)\right) = \left(\log c(\omega,\Omega) - \log C(\Omega)\right)$$

where

$$\begin{aligned}
\log p(\omega) &= const + \psi_a a + \psi_x x + \psi_y y \\
\log P(\Omega) &= const + \psi_a \bar{a} + \psi_x \bar{a} + \psi_y y \\
\log c(\omega,\Omega) &= const + (\varphi_a + \theta_2 l_a)a + (\varphi_x + \theta_2 l_x)x + (\varphi_y + \theta_2 l_y)y + \theta_2 l_A \bar{a} \\
\log C(\Omega) &= const + (\varphi_a + \theta_2 l_a)\bar{a} + (\varphi_x + \theta_2 l_x)\bar{a} + (\varphi_y + \theta_2 l_y)y + \theta_2 l_A \bar{a}
\end{aligned}$$

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

$$-\rho(\psi_a a + \psi_x x - (\psi_a + \psi_x)\bar{a}) = (\varphi_a + \theta_2 l_a)a + (\varphi_x + \theta_2 l_x)x - (\varphi_a + \theta_2 l_a + \varphi_x + \theta_2 l_x)\bar{a},$$

which is true if and only if

$$\psi_a = -\frac{1}{\rho}(\varphi_a + \theta_2 l_a) \tag{50}$$

$$\psi_x = -\frac{1}{\rho}(\varphi_x + \theta_2 l_x) \tag{51}$$

Finally, note that the observation of $y^p = \log P(\Omega) + \varepsilon^p$ is equivalent to the observation of the unbiased Gaussian signal

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

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

$$\kappa_{yp} = (\psi_a + \psi_x)^2 \sigma_p^{-2} = \frac{1}{\rho^2}(\varphi_a + \varphi_x + \theta_2(l_a + l_x))^2 \sigma_p^{-2},$$

which together with the fact that $\kappa_y = \sigma_a^{-2} + \kappa_{yq} + \kappa_{yp}$ and the characterization of κ_{yq} from Lemma 1 gives the result.

Part (ii). This is immediate from condition (12).

Proof of Proposition 4 In equilibrium, $l(\omega, \Omega)$ adjusts in stage 2 so as to satisfy the the consumer's demand:

$$\frac{p(\omega)}{P(\Omega)} = \left(\frac{q(\omega)\, l(\omega, \Omega)^{\theta_2}}{C(\Omega)} \right)^{-\frac{1}{\rho}} \tag{52}$$

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

$$\mathbb{E}\left[\frac{U'(C(\Omega))}{P(\Omega)} \left((1-\tau(\Omega)) C(\Omega) p^{1-\rho} P(\Omega)^{\rho} - w_2(\omega) \left(\frac{p}{P(\Omega)} \right)^{-\frac{\rho}{\theta_2}} \left(\frac{C(\Omega)}{q} \right)^{\frac{1}{\theta_2}} - w_1(\omega) \left(\frac{q}{e^a} \right)^{\frac{1}{\theta_1}} \right) \middle| \omega \right].$$

Note that this objective is strictly concave in $p^{1-\rho}$ and q^{1/θ_1}, 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

$$n(\omega)^{\epsilon_1 - 1} = w_1(\omega) \mathbb{E}\left[\frac{U'(C(\Omega))}{P(\Omega)} \middle| \omega \right] \quad \text{and} \quad l(\omega)^{\epsilon_2 - 1} = w_2(\omega) \mathbb{E}\left[\frac{U'(C(\Omega))}{P(\Omega)} \middle| \omega \right]$$

Solving these conditions for $w_1(\omega)$ 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:

$$p(\omega)^{1-\rho+\frac{\rho\epsilon_2}{\theta_2}} = \frac{\mathbb{E}\left[P(\Omega)^{\frac{\rho\epsilon_2}{\theta_2}}C(\Omega)^{\frac{\epsilon_2}{\theta_2}}q(\omega)^{-\frac{\epsilon_2}{\theta_2}}\middle|\,\omega\right]}{\theta_2\mathbb{E}\left[\left(\frac{\rho-1}{\rho}\right)(1-\tau(\Omega))C(\Omega)^{1-\gamma}P(\Omega)^{\rho-1}\middle|\,\omega\right]} \tag{53}$$

$$q(\omega)^{\frac{\epsilon_1}{\theta_1}} = p(\omega)^{1-\rho}e^{\frac{\epsilon_1}{\theta_1}a}\theta_1\mathbb{E}\left[\left(\frac{\rho-1}{\rho}\right)(1-\tau(\Omega))C(\Omega)^{1-\gamma}P(\Omega)^{\rho-1}\middle|\,\omega\right] \tag{54}$$

Using (52), we can restate these conditions in terms of allocations alone as follows:

$$0 = n(\omega)^{\epsilon_1-1} - \mathbb{E}\left[\left(\frac{\rho-1}{\rho}\right)(1-\tau(\Omega))U'(C(\Omega))\left(\frac{c(\omega,\Omega)}{C(\Omega)}\right)^{-\frac{1}{\rho}}\left(\theta_1\frac{c(\omega,\Omega)}{n(\omega)}\right)\middle|\,\omega\right]$$

$$0 = \mathbb{E}\left[l(\omega,\Omega)\left\{l(\omega,\Omega)^{\epsilon_2-1} - \left(\frac{\rho-1}{\rho}\right)(1-\tau(\Omega))U'(C(\Omega))\left(\frac{c(\omega,\Omega)}{C(\Omega)}\right)^{-\frac{1}{\rho}}\left(\theta_2\frac{c(\omega,\Omega)}{l(\omega,\Omega)}\right)\right\}\middle|\,\omega\right]$$

Rearranging these conditions, we get

$$\mathbb{E}\left[l(\omega,\Omega)^{\epsilon_2}\middle|\,\omega\right] = \frac{\theta_2}{\theta_1}e^{-\frac{\epsilon_1}{\theta_1}a}q(\omega)^{\frac{\epsilon_1}{\theta_1}} \tag{55}$$

$$q(\omega)^{\frac{\epsilon_1}{\theta_1}+\frac{1}{\rho}-1} = e^{\frac{\epsilon_1}{\theta_1}a}\theta_1\left(\frac{\rho-1}{\rho}\right)\mathbb{E}\left[(1-\tau)C(\Omega)^{\frac{1}{\rho}-\gamma}l(\omega,\Omega)^{\theta_2\left(\frac{\rho-1}{\rho}\right)}\middle|\,\omega\right] \tag{56}$$

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 (55) and (56) along with the resource constraint

$$C(\Omega) = \left[\int\left(q(\omega)l(\omega)^{\theta_2}\right)^{\frac{\rho-1}{\rho}}d\Omega(\omega)\right]^{\frac{\rho}{\rho-1}}; \tag{57}$$

(ii) the nominal prices satisfy condition (52); and (iii) the monetary policy satisfies

$$M(\Omega) = P(\Omega)C(\Omega). \tag{58}$$

We now seek to translate conditions (52)-(58) in terms of the relevant coefficients that parameterize the allocations, prices and policy under a log-normal specification. Thus let

$$\begin{aligned}
\log q(\omega) &= const + \varphi_a a + \varphi_x x + \varphi_y y \\
\log l\,(\omega, \Omega) &= const + l_A \bar{a} + l_a a + l_x x + l_y y \\
\log C(\Omega) &= const + c_A \bar{a} + c_y y \\
\log p(\omega) &= const + \psi_a a + \psi_x x + \psi_y y \\
\log(1 - \tau(\Omega)) &= const - \tau_A \bar{a} - \tau_y y \\
\log M(\Omega) &= const + \lambda_A \bar{a} + \lambda_y y
\end{aligned}$$

for some coefficients $(\varphi_a, \varphi_x, ..., \lambda_A, \lambda_y)$. (To simplify the derivations, we have expressed the policies as functions of $\bar{a}$ and y; except for degenerate cases, can always translate these contingencies in terms of contingencies on $\bar{a}$ and C or P.) Note that the resource constraint (57) is satisfied if and only if

$$c_A = (\varphi_a + \varphi_x) + \theta_2(l_a + l_x + l_A) \tag{59}$$

$$c_y = \varphi_y + \theta_2 l_y \tag{60}$$

while the nominal-demand condition (58) is satisfied if and only if

$$\lambda_A = c_A + (\psi_a + \psi_x) \tag{61}$$

$$\lambda_y = c_y + \psi_y \tag{62}$$

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

$$-\rho\,(\log p(\omega) - \log P(\Omega)) = (\log c(\omega, \Omega) - \log C(\Omega))$$

where

$$\log c(\omega) = \log q(\omega) + \theta_2 \log l(\omega) = const + (\varphi_a + \theta_2 l_a)a + (\varphi_x + \theta_2 l_x)x + (\varphi_y + \theta_2 l_y)y + \theta_2 l_A \bar{a}$$

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

$$-\rho(\psi_a a + \psi_x x - (\psi_a + \psi_x)\bar{a}) = (\varphi_a + \theta_2 l_a)a + (\varphi_x + \theta_2 l_x)x - (\varphi_a + \theta_2 l_a + \varphi_x + \theta_2 l_x)\bar{a}.$$

This is true if and only if

$$\psi_a = -\frac{1}{\rho}(\varphi_a + \theta_2 l_a) \qquad \text{and} \qquad \psi_x = -\frac{1}{\rho}(\varphi_x + \theta_2 l_x)$$

Finally, note that conditions (55) and (56) may be rewritten as follows:

$$\mathbb{E}[\log l\,(\omega,\Omega)\,|\omega] = const + \frac{\epsilon_1}{\epsilon_2\theta_1}\left(\log q(\omega) - a\right) \tag{63}$$

$$\log q(\omega) = const + \beta a - k(\tau_A \mathbb{E}[\bar{a}|\omega] + \tau_y y) + \frac{\alpha}{\chi}\mathbb{E}[\log C(\Omega)|\omega] \tag{64}$$

where

$$\beta \equiv \frac{\frac{\epsilon_1}{\theta_1}}{\frac{\epsilon_1}{\theta_1} - (\rho-1)\nu} > 1, \quad \alpha \equiv \left(\tfrac{1}{\rho} - \gamma\right)\frac{\rho\nu\chi}{\frac{\epsilon_1}{\theta_1} - (\rho-1)\nu},$$

$$\nu \equiv \frac{\epsilon_2}{\rho(\epsilon_2-\theta_2)+\theta_2} > \frac{1}{\rho}, \qquad \chi \equiv \frac{\epsilon_2}{(\epsilon_2-\theta_2)+\gamma\theta_2} > 0, \qquad k \equiv \nu\rho\frac{\theta_1}{\epsilon_1}\beta > 0.$$

Clearly, condition (63) holds for all ω if and only if

$$l_a = \frac{\epsilon_1}{\epsilon_2\theta_1}(\varphi_a - 1) \tag{65}$$

$$l_x = \frac{\epsilon_1}{\epsilon_2\theta_1}\varphi_x - l_A\frac{\kappa_x}{\kappa} \tag{66}$$

$$l_y = \frac{\epsilon_1}{\epsilon_2\theta_1}\varphi_y - l_A\frac{\kappa_y}{\kappa} \tag{67}$$

while condition (64) holds for all ω if and only if

$$\varphi_a = \beta \tag{68}$$

$$\varphi_x = -k\tau_A\frac{\kappa_x}{\kappa} + \frac{\alpha}{\chi}c_A\frac{\kappa_x}{\kappa} \tag{69}$$

$$\varphi_y = -k\left(\tau_A\frac{\kappa_y}{\kappa} + \tau_y\right) + \frac{\alpha}{\chi}\left(c_A\frac{\kappa_y}{\kappa} + c_y\right) \tag{70}$$

where c_A and c_y are given by (59)-(60).

Note that conditions (65) through (68) 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 $(\varphi_x, \varphi_y, l_A)$ and let $(\varphi_a, l_a, l_x, l_y)$ satisfy conditions (65) through (68). Note that there is a unique $(\varphi_a, l_a, l_x, l_y)$ that has this property for any given $(\varphi_x, \varphi_y, l_A)$. Next, pick an arbitrary ψ_y and let $(c_A, c_y, \psi_a, \psi_x)$ be determined as in (59)-(51). Next, let (τ_A, τ_y) be the unique solution to (69)-(70); for future reference, this solution is given by

$$\tau_A = \frac{1}{\chi k}\left\{\alpha c_A - \chi\frac{\kappa}{\kappa_x}\varphi_x\right\} \tag{71}$$

$$\tau_y = \frac{1}{\chi k}\left\{\alpha c_y - \chi\left(\varphi_y - \varphi_x\frac{\kappa_y}{\kappa_x}\right)\right\} \tag{72}$$

where $\chi k > 0$. Finally, set (λ_A, λ_y) as in (61)-(62). 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

$$n(\omega)^{\epsilon_1 - 1} = \left(\frac{\rho-1}{\rho}\right)\mathbb{E}\left[(1-\tau(\Omega))U'\left(C(\Omega)\right)\left(\frac{c(\omega,\Omega)}{C(\Omega)}\right)^{-\frac{1}{\rho}}\left(\theta_1\frac{c(\omega,\Omega)}{n(\omega)}\right)\Bigg|\,\omega\right] \quad (73)$$

$$l(\omega,\Omega)^{\epsilon_2 - 1} = \left(\frac{\rho-1}{\rho}\right)(1-\tau(\Omega))U'\left(C(\Omega)\right)\left(\frac{c(\omega,\Omega)}{C(\Omega)}\right)^{-\frac{1}{\rho}}\left(\theta_2\frac{c(\omega,\Omega)}{l(\omega,\Omega)}\right) \quad (74)$$

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

Proof of Proposition 5 *Parts (i) and (ii).* Take any allocation in which $\log q(\omega) = \varphi_0 + \varphi_a a + \varphi_x x + \varphi_y y$ and $\log l(\omega) = l_0 + l_A \bar{a} + l_a a + l_x x + l_y y$. Welfare (ex-ante utility) is then given by

$$\mathcal{W}(\varphi, l, \kappa) = \frac{1}{1-\gamma}\exp V_c(\varphi, l, \kappa) - \frac{1}{\epsilon_2}\exp V_l(\varphi, l, \kappa) - \frac{1}{\epsilon_1}\exp V_n(\varphi, l, \kappa)$$

where $\varphi = (\varphi_a, \varphi_x, \varphi_y)$, $l = (l_a, l_x, l_y, l_A)$, and $\kappa = (\kappa_x, \kappa_y)$, and where

$$\begin{aligned}
V_c(\varphi, l, \kappa) \equiv\ & (1-\gamma)\left(\varphi_0 + \theta_2 l_0 + \left[\theta_2 l_A + (\varphi_a + \theta_2 l_a) + (\varphi_x + \theta_2 l_x) + \left(\varphi_y + \theta_2 l_y\right)\right]\mu\right) \\
& + \frac{1}{2}(1-\gamma)\left(\frac{\rho-1}{\rho}\right)\left[\frac{(\varphi_a + \theta_2 l_a)^2}{\kappa_\xi} + \frac{(\varphi_x + \theta_2 l_x)^2}{\kappa_x} + 2\frac{(\varphi_a + \theta_2 l_a)(\varphi_x + \theta_2 l_x)}{\kappa_x}\right] \\
& + \frac{1}{2}(1-\gamma)^2\left[\frac{\left(\varphi_y + \theta_2 l_y\right)^2}{\kappa_y} + \frac{\left(\theta_2 l_A + (\varphi_a + \theta_2 l_a) + (\varphi_x + \theta_2 l_x) + \left(\varphi_y + \theta_2 l_y\right)\right)^2}{\kappa_A}\right] \\
V_l(\varphi, l, \kappa) \equiv\ & \epsilon_2\left(l_0 + (l_A + l_a + l_x + l_y)\mu\right) + \frac{1}{2}\epsilon_2^2\left[\left(\frac{l_a^2}{\kappa_\xi} + \frac{l_x^2}{\kappa_x} + 2\frac{l_a l_x}{\kappa_x}\right) + \frac{l_y^2}{\kappa_y} + \frac{(l_A + l_a + l_x + l_y)^2}{\kappa_A}\right] \\
V_n(\varphi, l, \kappa) \equiv\ & \frac{\epsilon_1}{\theta_1}\left(\varphi_0 + \left(\varphi_a + \varphi_x + \varphi_y - 1\right)\mu\right) \\
& + \frac{1}{2}\frac{\epsilon_1^2}{\theta_1^2}\left[\frac{(\varphi_a - 1)^2}{\kappa_\xi} + \frac{\varphi_x^2}{\kappa_x} + 2\frac{(\varphi_a - 1)\varphi_x}{\kappa_x} + \frac{\varphi_y^2}{\kappa_y} + \frac{\left(\varphi_a + \varphi_x + \varphi_y - 1\right)^2}{\kappa_A}\right]
\end{aligned}$$

We henceforth consider a relaxed problem, where we ignore the constraint on φ_a imposed by (13); 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

$$\begin{aligned}\varphi_0 &: \quad 0 = \exp V_c - \frac{1}{\theta_1}\exp V_n \\ l_0 &: \quad 0 = \theta_2 \exp V_c - \exp V_l\end{aligned}$$

Hence, at the optimal allocation, $\exp V \equiv \exp V_c = \frac{1}{\theta_1}\exp V_n = \frac{1}{\theta_2}\exp V_l > 0$. Let the Lagrange multipliers on the implementability constraints (13)-(16) be, respectively, $e^V\mu_a, e^V\mu_x$, and $e^V\mu_y$. Next, as in the proof of Proposition 2, we can represent the informational externalities by two Lagrange multipliers, one for the sum $\varphi_a+\varphi_x$, which determines κ_q, the precision of the output signal; and another for the sum $\varphi_a+\varphi_x+\theta_2(l_a+l_x)$, which determines κ_p, the precision of the price signal. Let these multipliers be, respectively, $e^V\eta_q$ and $e^V\eta_p$. We can then state the rest of the first-order conditions of the optimal policy problem as follows.

First, the conditions for the stage-1 strategy are the following:

$$\begin{aligned}
\varphi_a \; : \quad 0 \; = \; & \left(\frac{\rho-1}{\rho}\right)\left(\frac{(\varphi_a+\theta_2 l_a)}{\kappa_\xi}+\frac{(\varphi_x+\theta_2 l_x)}{\kappa_x}\right) \\
& +(1-\gamma)\frac{\left(\theta_2 l_A+(\varphi_a+\theta_2 l_a)+(\varphi_x+\theta_2 l_x)+\left(\varphi_y+\theta_2 l_y\right)\right)}{\kappa_A} \\
& -\frac{\epsilon_1}{\theta_1}\left[\frac{(\varphi_a-1)}{\kappa_\xi}+\frac{\varphi_x}{\kappa_x}+\frac{\left(\varphi_a+\varphi_x+\varphi_y-1\right)}{\kappa_A}\right]+\eta_q+\eta_p-\frac{\varepsilon_1}{\varepsilon_2\theta_1}\mu_a \\
\varphi_x \; : \quad 0 \; = \; & \left(\frac{\rho-1}{\rho}\right)\left(\frac{(\varphi_x+\theta_2 l_x)}{\kappa_x}+\frac{(\varphi_a+\theta_2 l_a)}{\kappa_x}\right) \\
& +(1-\gamma)\frac{\left(\theta_2 l_A+(\varphi_a+\theta_2 l_a)+(\varphi_x+\theta_2 l_x)+\left(\varphi_y+\theta_2 l_y\right)\right)}{\kappa_A} \\
& -\frac{\epsilon_1}{\theta_1}\left[\frac{\varphi_x}{\kappa_x}+\frac{(\varphi_a-1)}{\kappa_x}+\frac{\left(\varphi_a+\varphi_x+\varphi_y-1\right)}{\kappa_A}\right]+\eta_q+\eta_p-\frac{\varepsilon_1}{\varepsilon_2\theta_1}\mu_x \\
\varphi_y \; : \quad 0 \; = \; & (1-\gamma)\left[\frac{\left(\varphi_y+\theta_2 l_y\right)}{\kappa_y}+\frac{\left(\theta_2 l_A+(\varphi_a+\theta_2 l_a)+(\varphi_x+\theta_2 l_x)+\left(\varphi_y+\theta_2 l_y\right)\right)}{\kappa_A}\right] \\
& -\frac{\epsilon_1}{\theta_1}\left[\frac{\varphi_y}{\kappa_y}+\frac{\left(\varphi_a+\varphi_x+\varphi_y-1\right)}{\kappa_A}\right]-\frac{\varepsilon_1}{\varepsilon_2\theta_1}\mu_y
\end{aligned}$$

And second, the conditions for the stage-2 strategy are the following:

$$
\begin{aligned}
l_a \ : \quad 0 \ &= \ \left(\frac{\rho-1}{\rho}\right)\left(\frac{(\varphi_a+\theta_2 l_a)}{\kappa_\xi}+\frac{(\varphi_x+\theta_2 l_x)}{\kappa_x}\right) \\
&+(1-\gamma)\left[\frac{\left(\theta_2 l_A+(\varphi_a+\theta_2 l_a)+(\varphi_x+\theta_2 l_x)+\left(\varphi_y+\theta_2 l_y\right)\right)}{\kappa_A}\right] \\
&-\epsilon_2\left[\frac{l_a}{\kappa_\xi}+\frac{l_x}{\kappa_x}+\frac{(l_A+l_a+l_x+l_y)}{\kappa_A}\right]+\eta_p+\frac{\mu_a}{\theta_2} \\
l_x \ : \quad 0 \ &= \ \left(\frac{\rho-1}{\rho}\right)\left(\frac{(\varphi_x+\theta_2 l_x)}{\kappa_x}+\frac{(\varphi_a+\theta_2 l_a)}{\kappa_x}\right) \\
&+(1-\gamma)\left[\frac{\left(\theta_2 l_A+(\varphi_a+\theta_2 l_a)+(\varphi_x+\theta_2 l_x)+\left(\varphi_y+\theta_2 l_y\right)\right)}{\kappa_A}\right] \\
&-\epsilon_2\left[\frac{l_x}{\kappa_x}+\frac{l_a}{\kappa_x}+\frac{(l_A+l_a+l_x+l_y)}{\kappa_A}\right]+\eta_p+\frac{\mu_x}{\theta_2} \\
l_y \ : \quad 0 \ &= \ (1-\gamma)\left[\frac{\left(\varphi_y+\theta_2 l_y\right)}{\kappa_y}+\frac{\left(\theta_2 l_A+(\varphi_a+\theta_2 l_a)+(\varphi_x+\theta_2 l_x)+\left(\varphi_y+\theta_2 l_y\right)\right)}{\kappa_A}\right] \\
&-\epsilon_2\left[\frac{l_y}{\kappa_y}+\frac{(l_A+l_a+l_x+l_y)}{\kappa_A}\right]+\frac{\mu_y}{\theta_2} \\
l_A \ : \quad 0 \ &= \ (1-\gamma)\left[\frac{\left(\theta_2 l_A+(\varphi_a+\theta_2 l_a)+(\varphi_x+\theta_2 l_x)+\left(\varphi_y+\theta_2 l_y\right)\right)}{\kappa_A}\right]-\epsilon_2\left[\frac{(l_A+l_a+l_x+l_y)}{\kappa_A}\right] \\
&+\frac{\kappa_x}{\kappa}\frac{\mu_x}{\theta_2}+\frac{\kappa_y}{\kappa}\frac{\mu_y}{\theta_2}
\end{aligned}
$$

For any given η_q and η_p, the combination of these seven FOCs with the three implementability constraints (14)-(16) 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_y) and the implementability multipliers (μ_a, μ_x, μ_y). The solution to this system gives the following results: For the stage-1 allocation, we get

$$
\varphi_a^* \ = \ \beta \tag{75}
$$

$$
\varphi_x^* \ = \ \left\{\frac{(1-\alpha)\kappa_x^*}{(1-\alpha)\kappa_x^*+\kappa_y^*+\kappa_A}\right\}\frac{\alpha}{1-\alpha}\beta+\delta_q\eta_q+\delta_p\eta_p \tag{76}
$$

$$
\varphi_y^* \ = \ \left\{\frac{\kappa_y^*}{(1-\alpha)\kappa_x^*+\kappa_y^*+\kappa_A}\right\}\frac{\alpha}{1-\alpha}\beta-\frac{\kappa_y^*}{\kappa_A+\kappa_y^*}\left(\delta_q\eta_q+\delta_p\eta_p\right) \tag{77}
$$

where

$$\delta_q \equiv \frac{\theta_1\kappa_x\left(\kappa_A+\kappa_y\right)\left(\beta\left(\varepsilon_2\theta_1+\varepsilon_1\theta_2\right)\left(\kappa_A+\kappa_x+\kappa_y\right)-\varepsilon_1\theta_2\left(\kappa_A+\left(1-\alpha\right)\kappa_x+\kappa_y\right)\right)}{\varepsilon_1\left(\varepsilon_2\theta_1+\varepsilon_1\theta_2\right)\left(\kappa_A+\kappa_x+\kappa_y\right)\left(\kappa_A+\left(1-\alpha\right)\kappa_x+\kappa_y\right)} > 0$$

$$\delta_p \equiv \frac{\left(\beta\varepsilon_2\theta_1+\varepsilon_1\theta_2+\beta\varepsilon_1\theta_2\right)\kappa_x\left(\kappa_A+\kappa_y\right)}{\varepsilon_1\varepsilon_2\left(\kappa_A+\left(1-\alpha\right)\kappa_x+\kappa_y\right)} > 0$$

For the stage-2 allocation, we get

$$\begin{aligned}
l_A^* &= \hat{l}_A - \left(\zeta_q\eta_q+\zeta_p\eta_p\right)\\
l_a^* &= \hat{l}_a\\
l_x^* &= \hat{l}_x + \left(\frac{\kappa_x}{\kappa_A+\kappa_x+\kappa_y^*}\right)\left(\lambda_q\delta_q\eta_q+\lambda_p\delta_p\eta_p\right)\\
l_y^* &= \hat{l}_y + \left(\frac{\kappa_y^*}{\kappa_0+\kappa_x+\kappa_y^*}\right)\left(\lambda_q\delta_q\eta_q+\lambda_p\delta_p\eta_p\right)
\end{aligned}$$

where

$$\begin{aligned}
\hat{l}_A &\equiv \frac{(\kappa_A+\kappa_x+\kappa_y)\varphi_x^*\varepsilon_1}{\beta\kappa_x\varepsilon_2\theta_1}\\
\hat{l}_a &\equiv \frac{\epsilon_1}{\epsilon_2\theta_1}\left(\varphi_a^*-1\right)\\
\hat{l}_x &\equiv \frac{\epsilon_1}{\epsilon_2\theta_1}\varphi_x^* - \hat{l}_A\frac{\kappa_x}{\kappa}\\
\hat{l}_y &\equiv \frac{\epsilon_1}{\epsilon_2\theta_1}\varphi_y^* - \hat{l}_A\frac{\kappa_y}{\kappa}
\end{aligned}$$

and where

$$\begin{aligned}
\lambda_q &\equiv \frac{\left(\beta\varepsilon_2\theta_1+(\beta-1)\varepsilon_1\theta_2\right)\left(\kappa_A+\kappa_y\right)}{\beta\varepsilon_2\left(\varepsilon_2\theta_1+\varepsilon_1\theta_2\right)} > 0\\
\lambda_p &\equiv \frac{\left(\beta\varepsilon_2\theta_1+(\beta-1)\varepsilon_1\theta_2\right)\left(\kappa_A+\kappa_x+\kappa_y\right)}{\beta\varepsilon_2^2\theta_1} > 0
\end{aligned}$$

Note that, by Proposition 4, $(\hat{l}_A, \hat{l}_a, \hat{l}_x, \hat{l}_y)$ identifies the stage-2 allocation that would obtain in the (unique) flexible-price equilibrium in which the stage-1 allocation is given by (75)-(77). Finally, for the implementability multipliers, we get

$$\mu_a = \mu_x = \tfrac{\varepsilon_2\theta_1}{\varepsilon_1+\varepsilon_2\theta_1}\eta_q + \tfrac{\varepsilon_2\theta_1(1-\theta_2)}{\varepsilon_1+\varepsilon_2\theta_1}\eta_p \qquad \text{and} \qquad \mu_y = 0$$

Letting

$$\Delta_q \equiv \eta_q\delta_q \qquad \text{and} \qquad \Delta_p \equiv \eta_p\delta_p$$

completes the proof of all the conditions in the proposition.

What remains is to show that Δ_q and Δ_p (or, equivalently, η_q and η_p) are positive. Towards this goal, first note that

$$\begin{aligned}\eta_q &= e^{-V}\frac{\partial\mathcal{W}}{\partial\kappa_y}\frac{\partial\kappa_y}{\partial\kappa_q}\frac{\partial\kappa_q}{\partial\varphi_x}\\ \eta_p &= e^{-V}\frac{\partial\mathcal{W}}{\partial\kappa_y}\frac{\partial\kappa_y}{\partial\kappa_p}\frac{\partial\kappa_p}{\partial\varphi_x}\end{aligned}$$

Next, note that $\frac{\partial\kappa_y}{\partial\kappa_q} = \frac{\partial\kappa_y}{\partial\kappa_p} = 1$ and

$$\frac{\partial\mathcal{W}}{\partial\kappa_y} = -(1-\gamma)\exp V_c\frac{\left(\varphi_y+\theta_2 l_y\right)^2}{\kappa_y^2} + \frac{\epsilon_2}{\theta_2^2}\exp V_l\frac{(\theta_2 l_y)^2}{\kappa_y^2} + \frac{\epsilon_1}{\theta_1^2}\exp V_n\frac{\varphi_y^2}{\kappa_y}. \tag{78}$$

At the optimal allocation, we know that $\exp V \equiv \exp V_c = \frac{1}{\theta_1}\exp V_n = \frac{1}{\theta_2}\exp V_l$, as well as that $\mu_y = 0$ and $\mu_a = \mu_x$. Using the first fact, we get

$$\frac{\partial\mathcal{W}}{\partial\kappa_y} = \frac{e^V}{\kappa_y^2}\left\{(\gamma-1)\left(\varphi_y+\theta_2 l_y\right)^2 + \frac{\epsilon_2}{\theta_2}(\theta_2 l_y)^2 + \frac{\epsilon_1}{\theta_1}\varphi_y^2\right\}.$$

Using the second fact along with the FOCs with respect to (l_a, l_x, l_y, l_A) and the implementability constraint for φ_x, we can express l_y as a function of φ_y and μ_x :

$$l_y = \frac{1}{(\epsilon_2-\theta_2)+\gamma\theta_2}\left\{(1-\gamma)\varphi_y - \frac{\kappa_x\kappa_y}{\kappa_A+\kappa_x+\kappa_y}\frac{\mu_x}{\theta_2}\right\}$$

It follows that

$$\frac{\partial\mathcal{W}}{\partial\kappa_y} = \frac{e^V}{\beta}\left\{(1-\alpha)\frac{\epsilon_1}{\theta_1}\frac{\varphi_y^2}{\kappa_y^2} + \frac{(\beta-1+\alpha)\epsilon_1\theta_2+\beta\epsilon_2\theta_1}{\epsilon_2^2\theta_1\theta_2(\kappa_0+\kappa_x+\kappa_y)^2}\mu_x^2\kappa_x^2\right\}.$$

Since $\beta > 1$ necessarily, it is immediate that $\alpha \geq 0$ (which we have assumed) suffices for $\frac{\partial\mathcal{W}}{\partial\kappa_y} > 0$. Finally, recall that $\frac{\partial\kappa_q}{\partial\varphi_x} > 0$ if and only $\varphi_a+\varphi_x > 0$, while $\frac{\partial\kappa_p}{\partial\varphi_x} > 0$ if and only if $\varphi_a+\varphi_x+\theta_2 l_a+\theta_2 l_x > 0$. Combining these result, we conclude that $\Delta_q > 0$ and $\Delta_p > 0$ if and only if the optimal allocation satisfies $\varphi_a+\varphi_x > 0$ and $\varphi_a+\varphi_x+\theta_2 l_a+\theta_2 l_x > 0$.

To prove this, we proceed in a similar fashion as in Proposition 4. Let

$$\bar{v} \equiv \begin{bmatrix}\varphi_a+\varphi_x\\ \varphi_a+\varphi_x+\theta_2 l_a+\theta_2 l_x\end{bmatrix} \quad\text{and}\quad v(\kappa) \equiv \begin{bmatrix}\frac{\kappa_A+\kappa_x+\kappa_y}{(1-\alpha)\kappa_x+\kappa_y+\kappa_A}\beta\\ \frac{\kappa_A+\kappa_x+\kappa_y}{(1-\alpha)\kappa_x+\kappa_y+\kappa_A}\beta\left(1+\frac{\varepsilon_1\theta_2}{\varepsilon_2\theta_1}\right)\end{bmatrix}$$

As in Proposition 4, $\bar{v}-v(\kappa)$ 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

$$\mathcal{W}(\varphi, l, \kappa) = \frac{\left(\frac{\theta_1}{\epsilon_1} + \frac{\theta_2}{\epsilon_2}\right)^{-1} - 1 + \gamma}{1-\gamma} \left(\frac{\theta_1}{\epsilon_1} + \frac{\theta_2}{\epsilon_2}\right) \exp V_c(\varphi, l, \kappa)$$

Since $\left(\frac{\theta_1}{\epsilon_1} + \frac{\theta_2}{\epsilon_2}\right)^{-1} - 1 + \gamma > 0$ and $\frac{\theta_1}{\epsilon_1} + \frac{\theta_2}{\epsilon_2} > 0$, 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 then the characterization of the optimal coefficients (φ, l) as functions of $\bar{v}$, we conclude that

$$\mathcal{TW}(\varphi, l, \kappa) = W(\bar{v}, \kappa) \equiv A(\kappa) - (\bar{v} - v(\kappa))' B(\kappa) (\bar{v} - v(\kappa)), \tag{79}$$

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:

$$\begin{aligned}
B(\kappa) &\equiv \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \\
b_{11} &\equiv -\frac{(\varepsilon_2\theta_1 + \varepsilon_1\theta_2)(\kappa_A + \kappa_x + \kappa_y)((1-\alpha)\varepsilon_1\theta_2\kappa_x + \beta(\varepsilon_2\theta_1 + \varepsilon_1\theta_2)(\kappa_A + \kappa_y))}{(1-\alpha)\varepsilon_2\theta_1^2\theta_2\kappa_x^2(\kappa_A + \kappa_y)} \\
b_{22} &\equiv -\frac{\beta\varepsilon_2(\beta(\varepsilon_2\theta_1 + \varepsilon_1\theta_2)(\kappa_A + \kappa_x + \kappa_y) - \varepsilon_1\theta_2(\kappa_0 + (1-\alpha)\kappa_x + \kappa_y))}{(1-\alpha)\theta_2\kappa_x^2(\beta\varepsilon_2\theta_1 + (\beta-1)\varepsilon_1\theta_2)} \\
b_{12} \equiv b_{21} &\equiv \frac{\beta(\varepsilon_2\theta_1 + \varepsilon_1\theta_2)(\kappa_A + \kappa_x + \kappa_y)}{2(1-\alpha)\theta_1\theta_2\kappa_x^2}
\end{aligned}$$

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

$$\begin{aligned} det(B) &= b_{11}b_{22} - b_{12}b_{21} \\ &= \frac{\beta\left(\varepsilon_2\theta_1 + \varepsilon_1\theta_2\right)\left(\kappa_A + \kappa_x + \kappa_y\right)}{4\left(1-\alpha\right)^2\theta_1^2\theta_2^2\kappa_x^4\left(\beta\varepsilon_2\theta_1 + (\beta-1)\varepsilon_1\theta_2\right)\left(\kappa_A + \kappa_y\right)} \times \\ &\quad \times \Big[4\alpha\varepsilon_1\theta_2\kappa_x\left(\left(1-\alpha\right)\varepsilon_1\theta_2\kappa_x + \beta\left(\varepsilon_2\theta_1 + \varepsilon_1\theta_2\right)\left(\kappa_A + \kappa_y\right)\right) + \\ &\quad + \left(\beta\varepsilon_2\theta_1 + (\beta-1)\varepsilon_1\theta_2\right)\left(\kappa_A + \kappa_x + \kappa_y\right)\left(4\left(1-\alpha\right)\varepsilon_1\theta_2\kappa_x + 3\beta\left(\varepsilon_2\theta_1 + \varepsilon_1\theta_2\right)\left(\kappa_A + \kappa_y\right)\right)\Big] \\ &> 0. \end{aligned}$$

It follows that the matrix $B(\kappa)$ is negative definite and, hence, the aforementioned quadratic form for welfare in (79) is also negative definite. The same type of arguments as in Proposition 2 then imply that the optimal $\bar{v}$ is positive, and indeed higher than $v(\kappa)$, which in turn guarantees that $\Delta_q > 0$ and $\Delta_p > 0$.

Proof of Proposition 6 We prove the result in reverse order.

Part (ii). At the optimal allocation, aggregate consumption is given by

$$\log C(\Omega) = const + c_A^*\bar{a} + c_y^* y = const + (c_A^* + c_y^*)\bar{a} + c_y^*\varepsilon$$

where $c_A^* = (\varphi_a^* + \varphi_x^*) + \theta_2(l_a^* + l_x^* + l_A^*)$ and $c_y^* = \varphi_y^* + \theta_2 l_y^*$. If monetary policy were replicating the flexible-price allocations, then aggregate consumption would be given by

$$\log C(\Omega) = const + \hat{c}_A\bar{a} + \hat{c}_y y = const + (\hat{c}_A + \hat{c}_y)\bar{a} + \hat{c}_y\varepsilon$$

where $\hat{c}_A = (\varphi_a^* + \varphi_x^*) + \theta_2(\hat{l}_a + \hat{l}_x + \hat{l}_A)$ and $\hat{c}_y = \varphi_y^* + \theta_2\hat{l}_y$. From part (ii) of Proposition 5, it is immediate that $c_A^* < \hat{c}_A$ and $c_A^* + c_y^* < \hat{c}_A + \hat{c}_y$. The first inequality means that if one fixes the common belief about aggregate productivity (namely y) and considers the surprise component in the realization of aggregate productivity (namely $\bar{a} - y$), then the response of aggregate consumption to this surprise component is lower in the optimal allocation than in the flexible-price allocation. The second inequality means that if one considers the overall response of consumption to the aggregate productivity keeping the noise (ε), then this is also lower. So, no matter which way one sees it, the optimal monetary policy is less accommodative of the productivity shock than what would have been consistent with replicating the flexible-price allocations.

Part (i). Write the tax in terms of the productivity and the noise shocks as

$$-\log(1-\tau) = \tau_0^* + \tau_A^* \bar{a} + \tau_\varepsilon^* \varepsilon$$

We seek to prove that $\tau_A^* < 0$. Re-expressing the tax as a function of the productivity shock and the sufficient statistic of the public information gives

$$-\log(1-\tau) = \tau_0 + \tau_A \bar{a} + \tau_y y$$

where $\tau_A \equiv \tau_A^* - \tau_\varepsilon^*$ and $\tau_y = \tau_\varepsilon^*$. From conditions (71) and (72) in the proof of Proposition **??**, we know that the optimal tax satisfies

$$\begin{aligned}
\tau_A^* &= \frac{1}{\chi k}\left\{\alpha c_A^* - \chi \frac{\kappa}{\kappa_x}\varphi_x^*\right\} \\
\tau_y^* &= \frac{1}{\chi k}\left\{\alpha c_y^* - \chi \left(\varphi_y^* - \varphi_x^* \frac{\kappa_y}{\kappa_x}\right)\right\}
\end{aligned}$$

where $c_A^* = (\varphi_a^* + \varphi_x^*) + \theta_2(l_a^* + l_x^* + l_A^*)$ and $c_y^* = \varphi_y^* + \theta_2 l_y^*$. Let

$$\begin{aligned}
\hat{\tau}_A &\equiv \frac{1}{\chi k}\left\{\alpha \hat{c}_A - \chi \frac{\kappa}{\kappa_x}\varphi_x^*\right\} \\
\hat{\tau}_y &\equiv \frac{1}{\chi k}\left\{\alpha \hat{c}_y - \chi \left(\varphi_y^* - \varphi_x^* \frac{\kappa_y}{\kappa_x}\right)\right\}
\end{aligned}$$

where $\hat{c}_A \equiv (\varphi_a^* + \varphi_x^*) + \theta_2(\hat{l}_a + \hat{l}_x + \hat{l}_A)$ and $\hat{c}_y \equiv \varphi_y^* + \theta_2 \hat{l}_y$; this identifies the tax policy that would be required in order to implemented the optimal stage-1 sensitivities, φ_x^* and φ_y^*, if monetary policy were replicating the flexible-price allocations associated with these stage-1 sensitivities. It is easy to verify that

$$\hat{\tau}_A = -\lambda\Delta \qquad \text{and} \qquad \hat{\tau}_y = -\lambda \frac{\kappa_y}{\kappa_A + \kappa_y}\Delta$$

where $\Delta \equiv \Delta_q + \Delta_p$ and $\lambda \equiv \frac{(\kappa_A + (1-\alpha)\kappa_x + \kappa_y)\epsilon_1\epsilon_2}{\kappa_x(\beta\epsilon_2\theta_1 + (\beta - 1 + \alpha)\epsilon_1\theta_2)} > 0$. This result is identical to the result in the baseline model, except for the different value of the coefficient λ. Once again, $\Delta > 0$ is necessary and sufficient for $\hat{\tau}_A < 0$ and $\hat{\tau}_A + \hat{\tau}_y < 0$; that is, the tax that would have implemented the optimal stage-1 choices if monetary policy replicated the flexible-price allocation is countercyclical, much alike in the baseline model. But now note that the optimal tax satisfies

$$\begin{aligned}
\tau_A^* &= \hat{\tau}_A + \frac{1}{\chi k}\alpha(c_A^* - \hat{c}_1) \\
\tau_y^* &= \hat{\tau}_y + \frac{1}{\chi k}\alpha(c_y^* - \hat{c}_y)
\end{aligned}$$

By assumption, $\alpha > 0$, while by part (ii) of Proposition 5, $c_A^* < \hat{c}_A$ and $c_A^* + c_y^* < \hat{c}_A + \hat{c}_y$. It follows that $\tau_A^* < \hat{\tau}_A$ and $\tau_A^* + \tau_y^* < \hat{\tau}_A + \hat{\tau}_y$, which means that the optimal tax is even more countercyclical.