MACROECON & INT'L FINANCE WORKSHOP
presented by ***Jennifer La'O***
FRIDAY, Oct. 19, 2012
3:30 pm – 5:00 pm, **Room: HOH-706**

Optimal Monetary Policy with Informational Frictions*

George-Marios Angeletos
MIT and NBER

Jennifer La'O
University of Chicago and NBER

August 11, 2012

Abstract

This paper studies optimal policy in a class of economies in which incomplete information is the source of both nominal and real frictions: firms make both nominal price-setting decisions and real production decisions under imperfect information about the state of the economy. An appropriate notion of constrained efficiency is developed by embedding the informational friction as a measurability constraint on the Ramsey problem. Flexible-price allocations are shown to be constrained efficient, albeit not first-best efficient. The optimal monetary policy replicates flexible-price allocations, but does not target price stability. Rather, it "leans against the wind", targeting a negative correlation between the price level and real economic activity.

JEL codes: E32, E52, D61, D83.

Keywords: business cycles, dispersed information, sticky information, rational inattention, optimal policy, price stability.

*This paper extends, and subsumes, a 2008 working paper that circulated under the title "Dispersed Information over the Business Cycle: Optimal Fiscal and Monetary Policy." We benefited from comments by various colleagues and seminar participants at MIT, Chicago, Harvard, Indiana, John Hopkins, Ohio State, UC Irvine, UCLA, Universidad Carlos III, Wharton, the Federal Reserve Banks of Minneapolis and Philadelphia, the Bank of Spain, the 2008 Hydra Workshop on Dynamic Macroeconomics, the 2008 NBER Monetary Economics Meeting in Boston, and the 2011 SED Meeting in Ghent. We finally owe special thanks to Robert King and Philippe Bacchetta for discussing earlier versions of our paper. *Email:* `angelet@mit.edu`, `jenlao@chicagobooth.edu`.

1 Introduction

Informational frictions are paramount: firms, consumers, and investors alike hold heterogeneous beliefs about the current state of the economy and its likely future path. This could be either because the revelation of information through markets is imperfect (Lucas, 1972) or because people are inattentive (Mankiw and Reis, 2002, Sims, 2003, Woodford, 2003a).

No matter their precise origin, these frictions can have important business cycle implications. For example, Woodford (2003a) and Mackowiak and Wiederholt (2009) argue that rational inattention can justify significant inertia in the response of inflation to nominal shocks; Coibion and Gorodnichenko (2012) provide evidence in support of this class of models; Lorenzoni (2009) uses noisy signals about future TFP to develop a theory of demand shocks; Angeletos and La'O (2012) show how informational frictions can accommodate seemingly self-fulfilling phenomena and forces akin to "animal spirits" within otherwise conventional unique-equilibrium macro models.

Despite these advances on the positive front, it remains unclear as to how informational frictions affect the nature of optimal allocations and the design of optimal policy. The contribution of our paper is to answer this question in the context of a business-cycle economy in which these frictions are the source of both *nominal* and *real* rigidity.

Let us elaborate on what we mean by the latter point. Much of the pertinent literature[1] imposes that firms have imperfect information (or are inattentive) when setting their nominal prices, but allows their employment and output to adjust freely to the true state of nature, as if all production choices were made under *perfect* information. When this is the case, the informational friction is nothing more than a special form of nominal rigidity. By contrast, we are interested in the more realistic scenario in which the informational friction also has a real bite: think of firms making certain production decisions on the basis of the same noisy information (or limited attention) that also applies to their price-setting decisions. Once this is true, the first-best allocation is no more attainable, and the nature of the optimal policy is an open question. To the best of our knowledge, our paper is the first to address this question.

To this goal, we study an an economy in which firms have incomplete information both when setting their prices and when making certain employment, investment, and production choices. The information structure is taken as exogenous, but is otherwise entirely arbitrary. This permits us to identify normative properties that are likely to be robust to multiple interpretations and formalizations of the informational friction.[2] What is more, this permits us to accommodate multiple sources of fluctuations: the business cycle can be driven, not only by shocks to fundamentals such as preferences and technologies, but also by noise in public signals of these fundamentals, as in Jaimovic and Rebelo (2005) and Lorenzoni (2009), or by shocks to higher-order beliefs and forces akin to "animal spirits", as in Angeletos and La'O (2012).

[1] The statements we make in this paragraph apply to, inter alia, Mankiw and Reis (2002), Woodford (2003a), Ball, Mankiw and Reis (2005), Adam (2007), and Mackowiak and Wiederholt (2009).

[2] In particular, our specification directly nests the more special information structures imposed in, inter alia, Mankiw and Reis (2002), Woodford (2003), Ball et all (2005), Adam (2007), and the exogenous-information case of Paciello and Wiederholt (2011). But, in sharp contrast to all these papers, we let the informational friction have a real bite, in the sense explained earlier.

Our analysis then proceeds in three steps.

First, having recognized that the first best is unattainable, we define and study an appropriate notion of *constrained* efficiency, which embeds the real bite of the informational friction as a measurability constraint on the set of feasible allocations. Intuitively, this notion precludes the planner from "curing" a firm's inattentiveness, or from transferring information from one firm to another, but gives him unlimited power in manipulating how firms choose employment, investment and production on the basis of their available information. This notion thus bypasses the details of either the available policy instruments or the underlying market arrangements and, instead, identifies directly the socially optimal utilization of the information that is dispersed among the firms.

Next, we shift focus to the Ramsey problem, in which case the planner's power is restricted by specific policy instruments and specific market arrangements. To isolate the consequences of the real bite of the informational friction, we first study this problem under the assumption that prices are *flexible*, by which we mean the hypothetical situation in which firms are free to post a complete set of state-contingent prices. In this case, we show that the planner can always implement the constrained efficient allocation by setting a sales subsidy that offsets the monopoly distortion. Although this policy does not obtain the first best, it is optimal because it aligns private and social incentives. Once this is true, the available information is utilized in the socially optimal way, no matter how severe the informational friction might be.

Finally, we study the Ramsey problem when prices are *sticky*, by which we mean the situation in which the adjustment of prices to the state of nature is restricted by the presumed informational friction. In this case, we show that the planner can replicate a flexible-price allocation with an appropriate monetary policy if and only if the relative output of any two firms along that allocation satisfies an appropriate measurability condition, which we characterize in due course. While not innocuous, we argue that this restriction can be taken for granted for most practical purposes.[3]

Putting these findings together, we reach our first key result (Theorem 1): no matter the information structure, the optimal monetary policy is one that replicates flexible-price allocations.

The optimality of monetary policies that replicate flexible-price allocations is a cornerstone of modern macroeconomic theory.[4] At face value, our result may therefore appear similar to the standard one. There are, however, two important differences underneath the apparent similarity.

First, the nature of the optimal allocation is different. Unlike the New-Keynesian paradigm, the optimal allocation is no more the first best. Relative to this benchmark, the economy may now respond sluggishly to changes in preferences and technologies, and may also fluctuate in response to shocks that resemble "animal spirits". Seen through the lens of complete-information DSGE models, the economy may thus appear to be ridden with time-varying "labor wedges" or "markup shocks", in which case conventional wisdom might call for monetary policies that seek to stabilize the gap of aggregate output from its first-best level. Our analysis makes clear that such policies are suboptimal even though they may be feasible.

[3]For example, it is automatically satisfied if the technology is Cobb-Douglas and the disutility of labor has a power-form specification, as is often assumed in the literature.

[4]Versions of this result have been established by Goodfriend and King (1997, 2001), Rotemberg and Woodford (1999), Woodford (2003b), and Khan, King and Wolman (2003), among others.

Second, the implementation of the optimal allocation is also different. In the New-Keynesian paradigm, replicating flexible-price allocations is typically synonymous to targeting price stability. The same would have been true here if the informational friction was merely a form of nominal rigidity, as is often assumed in the literature. However, because the informational friction has a real bite, price stability is no more optimal. Instead, the optimal policy is shown to "lean against the wind", targetting a negative relation between the price level and real economic activity.

This result, which is our second key result (Theorem 2), follows directly from the need to make each firm's real choices respond to the firm's information about the state of the economy. Consider, in particular, the empirically plausible case in which constrained efficiency requires that optimistic firms employ, invest, and produce more than pessimistic ones. This can obtain in equilibrium only if optimistic firms face lower relative prices. But since the nominal price set by each firm cannot be a function of the information of another firm, optimistic firms can face lower relative prices only if they themselves set lower nominal prices. It follows that the nominal price of each firm must fall with its own belief, or sentiment, about the state of the economy. But then the aggregate price level must fall with the average sentiment in the economy, and thereby with aggregate economic activity.

Related literature. Our analysis builds on two methodological blocks. First, we follow the more abstract work of Angeletos and Pavan (2007, 2008) in studying an efficiency concept that bypasses the details of policy instruments and identifies the allocation that best utilizes the information that is dispersed in society. Second, we follow the Ramsey literature in taking the primal approach (e.g., Lucas and Stokey, 1983; Chari and Kehoe, 1999; Adao, Correia and Teles, 2003), and in characterizing the optimal monetary policy in relation to the underlying flexible-price allocations (e.g., Woodford, 2003b, Galí, 2008).

Our paper is thus the first to show how the Ramsey methodology can be adapted to environments in which informational frictions are the source of both nominal and real rigidity. This in turn is key to the generality of our results. Whereas most of the pertinent literature is tied to particular specifications of the information structure, the results we obtain in this paper apply to arbitrary information structures, at least in so far as the latter are treated as exogenous to the planner's problem. The spirit of our exercise is thus similar to that of Bergemann and Morris (2011), who also seek to identify predictions that are robust across a large class of information structures. But whereas they focus on the positive predictions of an abstract class of linear-quadratic games, we focus on the normative predictions of micro-founded business-cycle economy.

Finally, it is worth emphasizing how our contribution differs from those of Ball, Mankiw and Reis (2005), Adam (2007), Lorenzoni (2010), and Paciello and Wiederholt (2011). Like our paper, these papers are also concerned with optimal policy in economies with informational frictions. But unlike our paper, they treat the informational friction merely as a nominal rigidity. We elaborate on the importance and the consequences of this difference in due course.

Layout. The rest of the paper is organized as follows. Section 2 introduces the model. Section 3 develops our notion of constrained efficiency. Section 4 studies the sets of flexible- and sticky-price equilibria. Section 5 characterizes the optimal monetary policy. Section 6 discusses the robustness of our findings. Section 7 illustrates our results with a tractable example. Section 8 concludes.

2 The Model

Time is discrete and periods are indexed by $t \in \{0, 1, 2, ...\}$. There is a continuum of monopolistic firms, which produce differentiated goods and are indexed by $i \in I$. Their products are used by a competitive retail sector as intermediate inputs into the production of a final good, which in turn can be either consumed or invested into capital. Finally, there is representative household, which consists of a consumer, a continuum of workers, and a continuum of managers that run the monopolistic intermediate-good firms.[5]

Retailers. The retail sector consists of a representative final-good firm. Its technology is given by a CES aggregator of all the intermediate goods produced by all the monopolistic firms:

$$Y_t = \left[\int_I y_{it}^{\frac{\rho-1}{\rho}} di \right]^{\frac{\rho}{\rho-1}}$$

where Y_t is the quantity of the final good, y_{it} is the quantity of the intermediate good produced by monopolistic firm i, and $\rho > 1$ is the elasticity of substitution across intermediate goods. The firm's objective is to maximize its profit, $P_t Y_t - \int_I p_{it} y_{it} di$, where P_t is the price level (the price of the final good) and p_{it} is the price of the intermediate good i.

Monopolists. Take the firm that produces intermediate good i. Its output is given by

$$y_{it} = A_t F(k_{it}, \ell_{it}),$$

where A_t is an exogenous aggregate productivity shock,[6] F is a CRS production function, k_{it} is the firm's capital stock, and ℓ_{it} is its labor input. Labor input is given by $\ell_{it} = n_{it} h_{it}$, where n_{it} is the level of employment (number of workers) and h_{it} is the level of effort per worker. The firm's capital stock, on the other hand, evolves according to

$$k_{i,t+1} = (1 - \delta) k_{it} + x_{it}$$

where x_{it} is gross investment and $\delta \in [0, 1]$ is the depreciation rate. Finally, the firm's nominal after-tax profit is given by

$$\Pi_{it} = (1 - \tau_t) p_{it} y_{it} - P_t w_{it} n_{it} - P_t x_{it},$$

where τ_t is the tax rate on firm sales and w_{it} is the real wage per worker. The latter is given by $w_{it} = W(h_{it}, s^t)$, for some increasing, convex, and deferentiable function $W : \mathbb{R}_+ \times \mathcal{S}^t \to \mathbb{R}_+$. This simply means that a worker's overall compensation depends on the effort h_{it} the firm requires from him. The firm recognizes this dependence when deciding how many workers to hire or how

[5] Our analysis allows firms to be informationally-constrained but maintains the assumption of a representative, fully-informed, consumer. This is consistent with much of the recent literature on informational frictions, which also maintains this assumption (e.g., Woodford, 2003a, Mankiw and Reis, 2002, Mackowiak and Wiederholt, 2009). Introducing informational frictions on the consumer's side is a natural extension—but it can also be a challenging one, in so far as belief heterogeneity may then be conducive to uninsurable idiosyncratic consumption risk.

[6] We rule out idiosyncratic productivity shocks mostly for expositional reasons; see Appendix B.

much effort to require from each of them, but treats the wage schedule W as exogenous to its own choices—that is, the firm is a price-taker in the labor market.[7]

Households. The preferences of the representative household are given by

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

where C_t is consumption of the final good, ξ_t is an exogenous preference shock, and $\beta \in (0,1)$ is the discount factor. Note that this specification is similar to that in the literature on labor-hoarding, with the functions U and V capturing, respectively, the utility of consumption and the disutility of effort.[8] U is strictly increasing, strictly concave, and differentiable in C, with $U_c(0,\xi) = +\infty$ and $U_c(\infty,\xi) = 0$; and V is strictly increasing, strictly convex, and differentiable in h, with $V(0,\xi) > 0$, $V_h(0,\xi) = 0$, and $V_h(+\infty,\xi) = +\infty$.

The representative household holds two types of assets: nominal bonds and money. The monetary authority is assumed to pay interest on money holdings, making the household indifferent between money and bonds. The budget constraint is therefore given by

$$P_t C_t + B_{t+1} = \int_I \left(\Pi_{it} + P_t W(h_{it}, s^t) n_{it} \right) di + T_t + R_t B_t,$$

where T_t are lump-sum taxes, B_t are nominal financial assets (bonds and money), and R_{t+1} is the nominal interest rate.

Money and the government. We sidestep the micro-foundations of money and, instead, impose the following ad hoc cash-in-advance constraint on total expenditure:

$$P_t Y_t = M_t,$$

where M_t can be interpreted as either money supply or nominal aggregate demand. Nominal aggregate demand is allowed to be contingent on the state of the economy and is assumed to be controlled by the monetary authority. As in Woodford (2003b), this approach abstracts from the utility of money holdings and can be motivated by considering the limit of a "cashless economy" in which the monetary authority pays interest on money and appropriately adjusts the nominal interest rate so as to induce the desired level of nominal spending. Furthermore, by letting the monetary authority control directly the level of nominal spending, we sidestep the ongoing debate on whether interest-rate rules induce a unique equilibrium (Atkeson, Chari and Kehoe, 2010; Cochrane 2011).

While the monetary authority controls M_t, the fiscal authority controls the tax rate τ_t, the lump-sum transfer T_t, and the issue of new nominal assets B_{t+1}. Since there is no revenue from seigniorage (the government pays interest for money), the government budget is given by

$$R_t B_t + T_t = (1 - \tau_t) \int_I p_{it} y_{it} di + B_{t+1}$$

[7] Of course, the wage schedule W is ultimately pinned down in equilibrium by clearing the labor market. But both firms and workers take this schedule as exogenous to their choices.

[8] The particular specification of preferences assumed above simplifies the exposition and is consistent with the literature on labor hoarding, but is not strictly needed for our results. For instance, we could have let the disutility of effort, V, depend on C_t and/or n_{it}.

Furthermore, because Ricardian equivalence holds in our setting as in the vast majority of monetary business-cycle models, we can set $B_t = 0$ without any loss of generality. Finally, we allow the tax rate τ_t to be contingent on the realized state of the economy—although, as it will become clear later, this contingency is not needed for optimality as long as the monopoly distortion is constant, which is the case here.

Stochasticity and information. Our results do not depend on the precise details of how we model either the information structure or the underlying business-cycle shocks. We thus represent the aggregate state of the economy by the history $s^t = (s_0, ..., s_t)$ of an exogenous random variable s_t, which is drawn from a set $\mathcal{S}_t$ according to a conditional probability distribution $\mathcal{F}_t(s_t|s^{t-1})$. Similarly, the information set of firm i is represented by the history $\omega_i^t = (\omega_{i0}, ..., \omega_{it})$ of an arbitrary exogenous signal ω_{it} about the underlying aggregate state: for each i, ω_{it} is drawn from a set Ω_t according to a probability distribution $\mathcal{G}_t(\omega_{it}|s^t, \omega_i^{t-1})$. The sets $\mathcal{S}_t$, Ω_t and the distributions $\mathcal{F}_t$,$\mathcal{G}_t$ are allowed to be arbitrary. To simplify the exposition, we only assume that neither s^t nor any other random variable is ever commonly known: there always exists some heterogeneity in information, although perhaps arbitrarily small.[9]

Remarks. The aggregate preference and technology shocks are measurable in s^t, implying that we can express them as $\xi_t = \xi(s^t)$ and $A_t = A(s^t)$. Nevertheless, the aggregate state s^t may contain not only the preference and technology shocks but also aggregate shocks to the entire hierarchy of beliefs that the firms may hold about these fundamentals and, thereby, about each others' beliefs or actions. The business cycle may thus be driven by "news" and "noise shocks", as in, inter alia, Jaimovich and Rebelo (2009) and Lorenzoni (2009), as well as by "sentiment shocks" and forces akin to "animal spirits", as in Angeletos and La'O (2012). Part of our contribution will be to identify policies that may insulate macroeconomic outcomes from this kind of "non-fundamental" disturbances and, yet, to show that these policies are actually suboptimal.

Furthermore, because the information structure could be arbitrary, our framework can nest or at least proxy for many competing micro-foundations of informational frictions. For example, consider models with "sticky information" as in Mankiw and Reis (2002) and Ball et al (2005). These models are directly nested in our framework by letting $\mathcal{G}_t$ assign probability λ to $\omega_{it} = s^t$ and probability $1 - \lambda$ to $\omega_{it} = \omega_i^{t-1}$, where $\lambda \in (0, 1)$ is the probability with which a firm gets to see the true state and $1 - \lambda$ is the probability with which it learns nothing new. Alternatively, consider models in which firms observe noisy private and, possibly, public signals of the underlying shocks, such as in Morris and Shin (2002), Woodford (2003a), Hellwig (2002), Adam (2007), Nimark (2008), Amador and Weill (2011), and Angeletos and La'O (2009). These works, too, are directly nested in our framework by letting ω_i^t be the collection of such signals observed by firm i up to the beginning of period t. Finally, consider models with endogenous inattention as in Sims (2003), Mackowiak and Wiederholt (2009), and Paciello and Wiederholt (2011). These models are not directly nested

[9]This assumption can be interpreted as a finite bound on the firm' attention capacity. In any event, as it will become clear in due course, this assumption serves only as a minor equilibrium refinement. It rules out situations where all firms make their prices proportionally contingent on a sunspot or another common-knowledge random variable—a possibility that would not affect the set of equilibrium allocations (real variables) but would enrich the set of possible price paths (nominal variables).

in our framework because the signals observed by a firm are endogenous to the firm's information-processing problem and thereby to the planner's problem. Nevertheless, as we discuss in Section 6, some of our results are valid irrespective of this endogeneity.

Nominal versus real frictions. As mentioned in the introduction, a standard practice in the literature is to assume that the informational friction has only a nominal bite: firms are assumed to set prices on the basis of incomplete information, but all quantity margins are left free to adjust to the true state of nature. By contrast, we seek to accommodate the more general, and more realistic, scenario that the informational friction is also a *real* friction.

To this goal, we assume that the firm must not only set its price p_{it} on the basis of ω_{it}, but also choose how many workers to hire, n_{it}, and how much to invest, x_{it}, on the basis of the same information. At the same time, we let worker effort h_{it} adjust to the true state s^t so that supply can always meet demand. The first assumption permits us to introduce a real bite for the informational friction. The second assumption is needed in order to maintain conventional market-clearing equilibrium concepts: if firms fix prices and markets are to clear, there *must* be some margin of adjustment for quantities.[10]

3 Efficient allocations

The conventional approach to policy analysis in business-cycle models is based on first-best efficiency. This approach, however, is not appropriate for environments in which communication and/or cognitive frictions limit the ability of agents to acquire, digest, or respond to information. Surely enough, the planner could achieve first-best outcomes if he had enough instruments to fashion incentives and, in addition, could cure the agents' inattentiveness or otherwise get rid of the informational friction altogether. The question of interest for this paper, however, is what a government can achieve *given* the underlying informational friction.

With this in mind, we start our analysis by revisiting the notions of feasibility and efficiency that must guide policy in the presence of informational frictions.

Definition 1. *A feasible allocation is a collection of contingent plans for aggregate output, consumption, and labor, and firm-level contingent plans for employment, investment, and production choices, that satisfy the following constraints:*

(i) resource feasibility:

$$C_t + \int x_{i,t} di = Y_t = \left[\int y_{it}^{\frac{\rho-1}{\rho}} di \right]^{\frac{\rho}{\rho-1}}$$

$$y_{it} = A_{it} F\left(k_{it}, n_{it} h_{it}\right)$$

$$k_{i,t+1} = (1-\delta) k_{it} + x_{it}$$

(ii) h_{it} and y_{it} are contingent on (ω_i^t, s^t), while n_{it} and x_{it} are contingent only on ω_i^t.

[10]Although make a specific modeling choice regarding which production decisions are restricted to be contingent on ω_i^t and which ones are allowed to adjust to s^t, this modeling choice is not strictly needed for our formal results. All that matters is that some inputs are chosen on the basis of incomplete information, not the precise interpretation of these inputs.

Definition 2. *A constrained efficient allocation is a feasible allocation that maximizes the ex-ante utility of the representative household.*

The first constraint in Definition 1 needs no justification: it is the usual resource feasibility. The second constraint embeds the informational friction as a measurability constraint on allocations: an allocation is (informationally) feasible only if the employment and investment choices it prescribes to any particular firm are contingent at most on its manager's information set, ω_i^t.

If the informational friction represents a geographical segmentation, as, for example, in Lucas (1972) and Angeletos and La'O (2009), this constraint means that the planner cannot transfer information from one "island" to another. If the informational friction represents a cognitive limitation or a certain form of inattentiveness, as, for example, in Mankiw and Reis (2002), Sims (2003), and Woodford (2003a), this constraint simply means that the planner cannot overcome people's inattentiveness. In short, this constraint means that the planner is neither a messenger nor a psychiatrist. The constrained-efficiency concept we propose thus precludes the planner from getting rid of the informational friction, but gives him otherwise complete freedom in manipulating incentives and resources: the planner can simply dictate how firms respond to their available information, whatever that might be.

We henceforth express a feasible allocation with a collection of functions (n, h, x, k, y, C, K, Y) such that $n_{it} = n(\omega_i^t)$, $h_{it} = h(\omega_i^t, s^t)$, $x_{it} = x(\omega_i^t)$, $k_{it} = k(\omega_i^{t-1})$, $y_{it} = y(\omega_i^t, s^t)$, $C_t = C(s^t)$, $N_t = N(s^t)$, $K_t = K(s^{t-1})$, and $Y_t = Y(s^t)$. We can then state the planner's problem as follows:

Planner's problem. *Choose the functions (n, h, x, k, y, C, K, Y) so as to maximize*

$$\max \sum_{t=0}^{\infty} \beta^t \int \left[U\left(C(s^t), \xi(s^t)\right) - \int n(\omega_i^t) V\left(h(\omega_i^t), \xi(s^t)\right) d\mathcal{G}_t(\omega_i^t|s^t) \right] d\mathcal{F}_t(s^t)$$

subject to

$$\begin{aligned} C\left(s^t\right) + K\left(s^t\right) &= Y\left(s^t\right) + (1-\delta)K\left(s^{t-1}\right) \\ Y(s^t) &= \left[\int y(\omega_i^t, s^t)^{\frac{\rho-1}{\rho}} d\mathcal{G}_t(\omega_i^t|s^t) \right]^{\frac{\rho}{\rho-1}} \\ y(\omega_i^t, s^t) &= A(s^t) F\left(k\left(\omega_i^{t-1}\right), n\left(\omega_i^t\right) h\left(\omega_i^t, s^t\right)\right) \\ K\left(s^t\right) &= \int k\left(\omega_i^t\right) d\mathcal{G}_t(\omega_i^t|s^t) \end{aligned}$$

This is akin to the planner's problem in any conventional macroeconomic model. The only difference is that certain choices—here those associated with employment n_{it} and investment x_{it}—are restricted to be contingent on noisy, individual-specific, signals of the underlying state. Notwithstanding this qualification, the characterization of the planner's problem can proceed in a similar fashion.

Because this problem is strictly concave, it has a unique solution, which is pinned down by first-order conditions. To economize on notation, we henceforth let, for any $z \in \{k, \ell\}$,

$$MP_z\left(\omega_i^t, s^t\right) \equiv \left(\frac{y(\omega_i^t, s^t)}{Y(s^t)} \right)^{-\frac{1}{\rho}} A(s^t) \tfrac{\partial}{\partial z} F\left(k(\omega_i^{t-1}), n(\omega_i^t) h(\omega_i^t, s^t)\right)$$

denote firm i's marginal product of input z in terms of the final good. We similarly let

$$U_c(s^t) \equiv \frac{\partial}{\partial C} U\left(C(s^t), \xi(s^t)\right), \; V_h(\omega_i^t, s^t) \equiv \frac{\partial}{\partial h} V\left(h(\omega_i^t, s^t), \xi(s^t)\right), \text{ and } V(\omega_i^t, s^t) \equiv V\left(h(\omega_i^t, s^t), \xi(s^t)\right).$$

We use these short-cuts throughout the paper, with the understanding that these marginal products and marginal utilities depend on the allocation under consideration. We also let $\mathcal{P}(.|\omega_i^t)$ denote the Bayesian posterior conditional on ω_i^t, and $\mathbb{E}[.|\omega_i^t]$ the corresponding conditional expectation. We then reach the following characterization of the efficient allocation.

Proposition 1. *A feasible allocation is constrained efficient if and only if it satisfies the following conditions:*

$$V_h(\omega_i^t, s^t) - U_c(s^t) MP_\ell\left(\omega_i^t, s^t\right) = 0 \quad \forall\, \omega_i^t, s^t \tag{1}$$

$$\mathbb{E}\left[\; V(\omega_i^t, s^t) - U_c(s^t) MP_\ell\left(\omega_i^t, s^t\right) h(\omega_i^t, s^t) \;\middle|\; \omega_i^t \;\right] = 0 \quad \forall\, \omega_i^t \tag{2}$$

$$\mathbb{E}\left[\; U_c(s^t) - \beta U_c(s^{t+1}) \left\{1 - \delta + MP_k(\omega_i^{t+1}, s^{t+1})\right\} \;\middle|\; \omega_i^t \;\right] = 0 \quad \forall\, \omega_i^t \tag{3}$$

The interpretation of this result is simple. Condition (1), which pins down the optimal level of effort, equates the marginal disutility of effort with the social value of the marginal product of labor. Condition (2), which pins down the optimal level of employment, equates the total disutility of working with the social value of the marginal product of labor. Finally, condition (3) is an Euler condition that equates the expected marginal costs and benefits of investment conditional on the relevant information.

Clearly, these conditions are akin to those that characterize the first best. The only essential difference is that the informational friction introduces random wedges between the marginal rates of substitution and marginal rates of transformation of certain real choices. In particular, the marginal social costs and benefits of the employment and investment of any given firm fail to be equated state by state, simply because these choices can be conditioned only on a noisy indicator of the underlying state. Constrained efficiency then requires only that the resulting wedges between marginal costs and benefits are unpredictable conditional on the information of that firm.

These random wedges manifest themselvs as a form of "trembles" that perturb the planner's choice away from the frictionless, first-best allocation. Depending on the information structure, these "trembles" may feature arbitrary correlation across firms as well as with the underlying state s^t. As a result, the positive properties of the constrained efficient allocation may differ drastically from those of the first-best allocation, even though the optimality conditions that characterize these two allocations have a very similar flavor. We revisit this issue later.

4 Flexible-price vs sticky-price allocations

In the preceding section we defined and characterized a welfare benchmark that bypassed the details of available policy instruments: the planner was given direct and unlimited power in dictating how firms respond to their available information. We now shift attention to the conventional Ramsey policy problem, in which the planner's ability to fashion allocations is restricted by specific policy instruments, namely state-contingent rules for the sales tax τ_t and the monetary variable M_t. We thus examine whether these instruments suffice for implementing the constrained-efficient allocation as a market outcome.

We address this question in two steps. First, we study the allocations that can be implemented when nominal prices are "flexible" in the sense that p_{it} is free to adjust to the realized state s^t. Second, we study the allocations that can be implemented when prices are "sticky" in the sense that p_{it} is restricted to be measurable in ω_i^t. The first step therefore isolates the real bite of the informational friction on the set of implementable allocations, while the second step adds the nominal bite. These two steps are instrumental to the characterization of the optimal monetary policy that follows in the subsequent section.

4.1 Flexible-price equilibria

As explained above, the flexible-price concept we adopt removes the bite of informational frictions on nominal prices, while maintaining its hold on real decisions. One can think of this as the firms posting an elastic supply schedule, which lets the price adjust automatically with the level of demand. Alternatively, one can think of the "inattentive" manager delegating the price choice to an "attentive" sales division. One way or another, this means that the informational friction applies only on real allocations, not on the prices that support them.

Definition 3. *Prices are "flexible" or "state-contingent" if and only if p_{it} can be contingent on both ω_i^t and s^t.*

Definition 4. *A flexible-price equilibrium is a feasible allocation along with a collection of prices, wages, and fiscal and monetary policies such that*

(i) the household and all firms are at their respective optima;

(ii) the government budget constraint and the cash-in-advance constraint are satisfied;

(iii) markets clear;

(iv) prices are flexible in the sense of Definition 3.

Fix an arbitrary equilibrium and let $H(s^t)$ be the set of effort levels that obtain in this equilibrium for different realizations of ω_i^t when the aggregate state is s^t. The wage schedule W is part of this equilibrium if and only if it satisfies the following two conditions:

$$U_c(s^t)\ W(h,s^t) \ = \ V(h,\xi(s^t)) \quad \forall h \in H(s^t) \tag{4}$$

$$U_c(s^t)\ W(h,s^t) \ \leq \ V(h,\xi(s^t)) \quad \forall h \notin H(s^t) \tag{5}$$

To understand this, note first that, for any h, the quantity $U_c(s^t)W(h,s^t) - V(h,\xi(s^t))$ gives the marginal contribution to the household's welfare of a worker allocated to a job associated with effort level h. For a worker to prefer such a job to leisure, it had better be the case that the aforementioned difference is non-negative. For labor supply to be finite, it had better be the case that this difference is non-positive. Since labor demand is positive for $h \in H(s^t)$ and zero otherwise, these two properties imply that the labor market can clear if and only if the aforementioned quantity is zero for $h \in H(s^t)$ and non-positive for $h \notin H(s^t)$, which gives the result.[11]

Turning attention to firms, consider first the final-good sector. Its optimal input choices satisfy

$$y_{it} = \left(\frac{p_{it}}{P_t}\right)^{-\rho} Y_t, \tag{6}$$

which gives the demand function faced by the typical intermediate-good monopolistic firm. The latter's objective is to maximize its expectation of the present value of $\mathcal{M}_t \Pi_{it}$, where Π_{it} is its nominal profit and $\mathcal{M}_t \equiv U_c(C_t, \xi_t)/P_t$ is the marginal value of nominal income for the representative household. The firm internalizes the fact that the demand for its product is given by (6), and hence that its real revenue can be expressed as $\frac{p_{it} y_{it}}{P_t} = Y_t^{1/\rho} y_{it}^{1-1/\rho}$. The firm's real profit, net of taxes, labor compensation, and investment costs, is therefore given by

$$\frac{\Pi_{it}}{P_t} = (1-\tau_t) Y_t^{1/\rho} y_{it}^{1-1/\rho} - W(h_{it}, s^t) n_{it} - x_{it}.$$

Recall then that effort h_{it} adjusts to the true state s^t in addition to ω_i^t. On the other hand, employment n_{it} and investment x_{it} are conditioned only on ω_i^t. With this qualification in mind, we conclude that we can write the firm's problem as follows:

$$\max_{n,h,x,k,y} \mathbb{E}\left[\sum_t^\infty \beta^t U_c(s^t) \left\{ (1-\tau(s^t)) Y(s^t)^{1/\rho} y_{it}^{1-1/\rho} - W(h_{it}, s^t) n_{it} - x_{it} \right\} \,\middle|\, \omega_i^t \right]$$

subject to

$$y_{it} = A(s^t) F\left(k_{it}, n_{it} h_{it}\right)$$
$$k_{i,t+1} = (1-\delta) k_{it} + x_{it}$$

Since ρ is greater than 1, F is concave, and W is convex, the above problem is strictly concave. The following first-order conditions therefore pin down the firm's optimal plan:

$$\begin{aligned}
\left(\tfrac{\rho-1}{\rho}\right)\left(1-\tau(s^t)\right) MP_\ell(\omega_i^t, s^t) - W_h(\omega_i^t, s^t) &= 0 \\
\mathbb{E}\left[U_c(s^t) \left\{ \left(\tfrac{\rho-1}{\rho}\right)\left(1-\tau(s^t)\right) MP_\ell(\omega_i^t, s^t) h(\omega_i^t, s^t) - W(\omega_i^t, s^t) \right\} \,\middle|\, \omega_i^t \right] &= 0 \\
\mathbb{E}\left[U_c(s^t) - \beta U_c(s^{t+1}) \left\{ 1 - \delta + \left(\tfrac{\rho-1}{\rho}\right)\left(1-\tau(s^{t+1})\right) MP_k(\omega_i^{t+1}, s^{t+1}) \right\} \,\middle|\, \omega_i^t \right] &= 0
\end{aligned}$$

where $W(\omega_i^t, s^t)$ is a short-cut for $W(h(\omega_i^t, s^t), s^t)$, the compensation paid per worker, and similarly $W_h(\omega_i^t, s^t)$ is a short-cut for $\frac{\partial}{\partial h} W(h(\omega_i^t, s^t), s^t)$, the slope of the wage schedule. The first condition gives the optimal choice of worker effort, which is free to adjust to s^t. The other two conditions give the optimal choices for employment and investment, which are restricted to depend only on ω_i^t.

Comparing the above conditions to those in Proposition 1, we see two key differences. First, the marginal product of each input is multiplied by the following wedge, which reflects the combination of the monopoly markup and the sales tax:

$$\phi(s^t) = \left(\frac{\rho-1}{\rho}\right)\left(1-\tau(s^t)\right).$$

[11] Note that, once we fix an equilibrium allocation, the wage schedule that supports it is uniquely determined by condition (4) for all on-equilibrium effort levels, but remains indeterminate for any off-equilibrium effort levels. This residual indeterminacy, however, does not have any allocative consequences: there are just multiple wage schedules that support exactly the same equilibrium allocations.

Second, the wage schedule W shows up in place of the labor disutility V, reflecting the fact that firms care about labor compensation rather than the disutility of effort. However, perfect competition in the labor market guarantees that W and V coincide along the equilibrium allocation; this is shown in conditions (4) and (5).[12] Building on these observations, we reach the following result, which characterizes the set of allocations that can be part of an equilibrium when prices are flexible.

Proposition 2. *A feasible allocation is part of a flexible-price Ramsey equilibrium if and only if there exists a function $\phi : \mathcal{S}^t \to \mathbb{R}_+$ such that the following hold:*

$$V_h(\omega_i^t, s^t) - U_c(s^t)\phi(s^t)MP_\ell\left(\omega_i^t, s^t\right) = 0 \quad \forall\, \omega_i^t, s^t \tag{7}$$

$$\mathbb{E}\left[\ V(\omega_i^t, s^t) - U_c(s^t)\phi(s^t)MP_\ell\left(\omega_i^t, s^t\right) h(\omega_i^t, s^t)\ \middle|\ \omega_i^t\ \right] = 0 \quad \forall\, \omega_i^t \tag{8}$$

$$\mathbb{E}\left[\ U_c(s^t) - \beta U_c(s^{t+1})\left\{1 - \delta + \phi(s^{t+1})MP_k(\omega_i^{t+1}, s^{t+1})\right\}\ \middle|\ \omega_i^t\ \right] = 0 \quad \forall\, \omega_i^t \tag{9}$$

The only difference between the constrained efficient allocation and any flexible-price allocation is therefore the wedge $\phi(s^t)$ that shows up in the above conditions. By appropriately choosing the tax, the Ramsey planner can induce an arbitrary such wedge, and in so doing implement any of the allocations identified by Proposition 2. This explains why ϕ is a "free variable" in this proposition: the set of flexible-price allocations is spanned by varying ϕ. Finally, because the planner can always set $\phi(s^t) = 1$ for all s^t, the following is immediate.

Corollary 1. *The constrained efficient allocation is necessarily contained in the set of flexible-price allocations.*

This result has the same flavor as the optimality of flexible-price allocations in the standard New-Keynesian paradigm. There is, however, an important difference. In the standard paradigm, the set of flexible-price allocations contains the first-best allocation, in which case their optimality is immediate. The same would hold in our setting if we had assumed, like much of the literature, that the informational friction is merely a source of nominal rigidity. But now that the informational friction is also a source of real friction, the first-best allocation is no more contained in the set of flexible-price allocations, and the latter are no more optimal in the *usual* sense.[13] The novelty of our result therefore hinges on formalizing the precise sense in which the optimality of flexible-price allocations is preserved once informational frictions render the first-best allocation unattainable. This is not only a conceptual matter; we elaborate on its policy implications in Section 5.

4.2 Sticky-price equilibria

We now move on to characterize the set of equilibrium allocations when prices are sticky, by which we mean that firms can no more post state-contingent price plans (or supply functions).

[12] In particular, since both W and V are differentiable, (4) and (5) can hold only if $W_h(h, s^t) = V_h(h, \xi(s^t))$ also holds for all $h \in H(s^t)$.

[13] In making this statement, we are ruling out degenerate cases where the first-best allocation is invariant to the state of nature—such as when preferences and technologies are entirely invariant and the state of nature includes only sunspot variables—or is measurable in the information sets of all firms.

Definition 5. *Prices are "sticky" or "non-contingent" if and only if p_{it} is contingent on ω_i^t, but not on s^t.*

Definition 6. *A sticky-price equilibrium is a feasible allocation along with a collection of prices and fiscal and monetary policies such that*

(i) the household and all firms are at their respective optima;

(ii) the government budget constraint and the cash-in-advance constraint are satisfied;

(iii) markets clear.

(iv) prices are sticky.

The notion of nominal "stickiness" we are formalizing above is fully consistent with, inter alia, Mankiw and Reis (2002), Woodford (2003a), Nimark (2008), Mackowiak and Wiederholt (2009), Lorenzoni (2009, 2010), and Paciello and Wiederholt (2011). We are merely making this notion explicit, highlighting how the nominal bite of the informational friction is formally similar to that of the Calvo friction: whether the nominal rigidity originates in incomplete information or Calvo pricing, the essential friction is that prices are not free to adjust to the underlying state.

This form of stickiness affects the behavior of the monopolistic firms, but not that of the workers or of the final-good sector. The wage schedule thus continues to satisfy (4)-(5), and the final-good sector's demand for intermediate inputs continues to satisfy (6). Using this last fact, we can now express the monopolistic firm's real revenue as $\frac{p_{it}y_{it}}{P_t} = (\frac{p_{it}}{P_t})^{1-\rho}Y_t$.

$$\frac{\Pi_{it}}{P_t} = (1-\tau_t)\frac{p_{it}y_{it}}{P_t} - W(h_{it}, s^t)n_{it} - x_{it} = (1-\tau_t)\left(\frac{p_{it}}{P_t}\right)^{1-\rho} Y_t - W(h_t, s^t)n_{it} - x_{it}.$$

Keeping in mind that h_{it} is contingent of both ω_i^t and s^t while n_{it}, x_{it}, and p_{it} are contingent only on ω_{it}, we can thus state the monopolistic firm's optimal pricing and production problem as follows:

$$\max_{n,h,x,k,p} \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t U_c(s^t)\left\{(1-\tau_t)\left(\frac{p_{it}}{P_t}\right)^{1-\rho} Y_t - W(h_t, s^t)n_{it} - x_{it}\right\}\middle|\,\omega_i^t\right]$$

subject to

$$k_{i,t+1} = (1-\delta)\,k_{it} + x_{it} \quad \forall \omega_i^t$$

$$A(s^t)F\left(k_{it}, n_{it}h_{it}\right) = \left(\frac{p_{it}}{P(s^t)}\right)^{-\rho} Y(s^t) \quad \forall \omega_i^t, s^t$$

The first constraint is simply the law of motion for capital. The second constraint, which follows from combining condition (6) with the production function, dictates how labor utilization h_{it} adjusts so as to meet the realized demand, whatever that might be.

Let $\beta^t U_c(s^t)\lambda(\omega_i^t, s^t)$ be the Lagrange multiplier on the second constraint. Next, let $f_\ell(\omega_i^t, s^t)$ be a short-cut for $\frac{\partial}{\partial \ell}F(k(\omega_i^{t-1}), n(\omega_i^t), h(\omega_i^t, s^t))$, the physical marginal product of labor, and similarly let $f_k(\omega_i^t, s^t)$ be a short-cut for the physical marginal product of capital. (These should not be confused with MP_ℓ and MP_k, the marginal products in terms of the final good, which we defined earlier.) The first order conditions with respect to effort h_{it}, employment n_{it}, and investment x_{it} are given by the following:

$$\begin{aligned} \lambda\left(\omega_i^t, s^t\right) A(s^t) f_\ell\left(\omega_i^t, s^t\right) - W_h(\omega_i^t, s^t) &= 0 \qquad (10) \\ \mathbb{E}\left[U_c(s^t)\left\{\lambda(\omega_i^t, s^t) A(s^t) f_\ell\left(\omega_i^t, s^t\right) h(\omega_i^t, s^t) - W(\omega_i^t, s^t)\right\} \mid \omega_i^t\right] &= 0 \qquad (11) \\ \mathbb{E}\left[U_c(s^t) - \beta U_c(s^{t+t})\left\{1 - \delta + \lambda\left(\omega_i^{t+1}, s^{t+1}\right) A(s^{t+1}) f_k\left(\omega_i^{t+1}, s^{t+1}\right)\right\} \mid \omega_i^t\right] &= 0. \qquad (12) \end{aligned}$$

The first-order condition with respect to the price p_{it}, on the other hand, can be stated as follows:

$$\mathbb{E}\left[U_c(s^t) y\left(\omega_i^t, s^t\right)\left\{\left(1 - \tau(s^t)\right)\left(\tfrac{\rho-1}{\rho}\right)\frac{p(\omega_i^t)}{P(s^t)} - \lambda\left(\omega_i^t, s^t\right)\right\}\middle|\, \omega_i^t\right] = 0 \qquad (13)$$

To interpret these conditions, note that $\lambda_{it} = \lambda(\omega_i^t, s^t)$ identifies the real marginal cost the firm has to incur in order to meet the realized demand at the pre-set price. Since the only input that can adjust to meet demand is effort h_{it}, the aforementioned marginal cost is given by the ratio of the price of this input to its marginal product, which gives condition (10). Conditions (11) and (12) then say that employment and capital are chosen so as to minimize the expected cost of servicing demand. Finally, to interpret condition (13), note that $(1 - \tau_t)\,\frac{\rho-1}{\rho}\frac{p_{it}}{P_t}$ is marginal revenue net of taxes. If prices were flexible, marginal revenue and marginal costs would have to be equated state-by-state: $(1 - \tau_t)\,\frac{\rho-1}{\rho}\frac{p_{it}}{P_t}$ would equal λ_{it} for all (ω_i^t, s^t). Now that prices are sticky, the realized marginal revenue and marginal costs will generally fail to coincide as demand varies with information upon which the price cannot be contingent. This is the essence of how informational frictions interact with nominal frictions both in our framework and in all of the related literature: the key is that prices are contingent only on a noisy indicator of the true state. At the same time, optimality in the firm's price-setting behavior requires that the realized pricing "error" (i.e., the difference between marginal revenue and marginal cost) be orthogonal to the variation in output under the appropriate, risk-adjusted, expectation operator of the manager. This is the meaning of condition (13).

To recap, any sticky-price equilibrium is pinned down by the optimality conditions (10)-(13) along with the budgets constraints, the cash-in-advance constraint, and the equilibrium conditions (4)-(5) for the wage schedule. We now move on to characterize these equilibria by a set of necessary and sufficient conditions that solve out for the equilibrium prices and the associated policies and, instead, put restrictions only on the real allocations. Towards this goal, we first introduce the following definition.

Definition 7. *An allocation is log-separable if and only if there exist positive-valued functions Ψ^ω and Ψ^s such that*

$$\log y\left(\omega_i^t, s^t\right) = \log \Psi^\omega(\omega_i^t) + \log \Psi^s(s^t) \qquad (14)$$

Condition (14) requires that the output of a firm can be expressed as the logarithmic sum of two components: one that depends only on the firm's information set ω_i^t (the "ω-component") and another that depends only in the true aggregate state s^t (the "s-component"). We postpone a further discussion of the meaning and the role of this condition for later. After some manipulation of conditions (10)-(13), which we leave for the appendix, we then reach the following result.

Proposition 3. *A feasible allocation can be part of a sticky-price equilibrium if and only if*

(i) there exist functions $\phi : \mathcal{S}^t \to \mathbb{R}_+$ *and* $\chi : \Omega^t \times \mathcal{S}^t \to \mathbb{R}_+$ *such that the following hold:*

$$V_h(\omega_i^t, s^t) - U_c(s^t)\phi(s^t)\chi(\omega_i^t, s^t)MP_\ell\left(\omega_i^t, s^t\right) = 0 \quad \forall\ (\omega_i^t, s^t) \tag{15}$$

$$\mathbb{E}\left[V(\omega_i^t, s^t) - U_c(s^t)\phi(s^t)\chi(\omega_i^t, s^t)MP_\ell\left(\omega_i^t, s^t\right) \;\middle|\; \omega_i^t\right] = 0 \quad \forall\ \omega_i^t \tag{16}$$

$$\mathbb{E}\left[U_c(s^t) - \beta U_c(s^{t+1})\left\{1 - \delta + \phi(s^{t+1})\chi(\omega_i^{t+1}, s^t)MP_k\left(\omega_i^{t+1}, s^{t+1}\right)\right\} \;\middle|\; \omega_i^t\right] = 0 \quad \forall\ \omega_i^t \tag{17}$$

$$\mathbb{E}\left[U_c(s^t)Y\left(s^t\right)^{1/\rho} y\left(\omega_i^t, s^t\right)^{1-1/\rho}\phi(s^t)\left\{\chi(\omega_i^t, s^t) - 1\right\} \;\middle|\; \omega_i^t\right] = 0 \quad \forall\ \omega_i^t \tag{18}$$

(ii) the allocation is log-separable (in the sense of Definition 7).

To understand this result, it is useful to contrast it to Proposition 2, which characterized the set of implementable allocations under flexible-prices.

Consider first part (i). The only difference between this part and Proposition 2 is captured by the additional wedge $\chi_{it} = \chi(\omega_i^t, s^t)$ that shows up in conditions (15), (16) and (17) relative to the corresponding conditions in Proposition 2. This wedge represents random variation in the realized monopoly markup of firm i, which obtains as monetary policy and aggregate demand react to contingencies that were unobserved (or not attended to) at the moment firms set their prices: the realized markup is lower (respectively, higher) than the optimal one if demand turns out to be higher (respectively, lower) than what expected. This wedge therefore encapsulates the extra power that the nominal rigidity gives the planner in terms of manipulating allocations: now that prices are sticky, the planner can induce, not only the wedge ϕ_t by appropriately choosing the tax rate, but also the wedges χ_{it} by appropriately choosing monetary policy.

This extra power, however, is limited. While the planner is entirely free to choose any ϕ_t he wishes, his choice over χ_{it} is restricted by condition (18). The latter requires, in effect, that the stochastic variation in the wedge χ_{it} faced by each firm i be unforecastable on the basis of that firm's information set.[14] Intuitively, this means that monetary policy can impact real allocations only in so far as it is "unanticipated", in the sense of reacting to contingencies that are not measurable in the information sets upon firms set their prices.[15] Furthermore, the ability of monetary policy to "surprise" any given firm hinges on how inattentive, or imperfectly informed, that firm is; this explains why the wedge χ_{it} is firm-specific.

If the above property was the only difference between the set of flexible-price allocations and that of sticky-price allocations, then the latter would necessarily contain the former: by breaking the neutrality of monetary policy, the nominal rigidity would only enlarge the set of implementable allocations. However, this is not the whole story. By restricting the nominal price p_{it} of any given firm i to be contingent on ω_{it} and not on either the true state s^t or the information of any other firm, the nominal rigidity also restricts how the *relative* prices of any two firms can vary with s^t, as well

[14] This statement must be qualified as follows: by "unforecastable" we mean under an appropriate risk-adjusted expectation operator. Formally, (18) can be restated as $\hat{\mathbb{E}}_{it}[\chi_{it}] = 1$, where $\hat{\mathbb{E}}_{it}[X] \equiv \mathbb{E}_{it}[U_c(C_t)\phi_t p_{it} y_{it} \cdot X]$ is the aforementioned risk-adjusted expectation operator.

[15] Note here the implicit analogy: the recent work by Mankiw and Reis (2002), Woodford (2003a) and others has provided novel, and appealing, re-interpretations of the informational friction, but the reason that money is non-neutral in this recent work is formally similar to that in the older work by Lucas (1972).

as with any difference in beliefs between the two firms. In so doing, the nominal rigidity restricts the real allocations that can be implemented by the Ramsey planner relative to the flexible-price case. This extra bite of the nominal rigidity on the set of implementable allocations is captured by the log-separability restriction in part (ii) of the above proposition.

To elaborate on the meaning of this restriction, fix a period t and a state s^t, and take an arbitrary pair of firms i and j that have incomplete and differential information about s^t. For an allocation to be implementable as part of a sticky-price Ramsey equilibrium, it must be that the *nominal* price set by firm i is contingent at most on ω_i^t, and similarly the nominal price set by firm j must be contingent at most on ω_j^t. At the same time, the *relative* price of the two firms is pinned down, from the consumer's side, by their relative output. Putting the two properties together, we infer that any sticky-price allocation must satisfy the following relation between the nominal prices and the relative output of the two firms:

$$\log p(\omega_i^t) - \log p(\omega_j^t) = -\rho \left[\log y(s^t, \omega_i^t) - \log y(s^t, \omega_j^t)\right] \tag{19}$$

Clearly, the above condition can hold for all realizations of ω_i^t, ω_j^t and s^t only if the right-hand side of this condition is independent of s^t, which in turn requires that the dependence of y_{it} on s^t cancels with the corresponding dependence of y_{jt}. This is precisely where the log-separability restriction kicks in: the relative output of any two arbitrary firms is independent of measurable in the joint of their information sets if and only if the allocation is log-separable in the sense of Definition 7.

To recap, while part (i) of Proposition 4 encapsulates the extra power that the Ramsey planner enjoys thanks to the non-neutrality of monetary policy, part (ii) identifies the extra constraints that the nominal rigidity imposes on the planner's problem by restricting the stochastic variation in relative prices and real allocations. The following is then immediate.

Corollary 2. *A flexible-price allocation is contained in the set of sticky-price allocations if and only if it is log-separable.*

Combining this result with Corollary 1, we infer that the constrained efficient allocation can be implemented under sticky prices if and only if it is log-separable. This begs the question of what it takes in terms of the primitives of the economy in order to get a log-separable allocation. A sufficient condition is provided in the following.

Proposition 4. *Suppose the production function is Cobb-Douglas and the disutility of effort has a power form. That is,*

$$F(k, nh) = k^{1-\alpha}(nh)^{\alpha} \qquad \textit{and} \qquad V(n) = v_0 + v_1 \frac{n^{1+\epsilon}}{1+\epsilon}, \tag{20}$$

for some $\alpha \in (0,1)$, $\epsilon, v_0, v_1 > 0$.

Then, every flexible-price allocation—including the constrained efficient one—is log-separable, and can therefore be replicated under sticky prices.

Furthermore, the ω*- and* s*-components of output are given by*

$$\Psi^{\omega}(\omega_i^t) = \left[k\left(\omega_i^t\right)^{1-\alpha} n\left(\omega_i^t\right)^{\frac{\alpha\epsilon}{1+\epsilon}}\right]^{\chi} \quad \textit{and} \quad \Psi^{s}(s^t) = \left[A\left(s^t\right)\left(U_c(Y\left(s^t\right))\phi(s^t)Y(s^t)^{\frac{1}{\rho}}\right)^{\frac{\alpha}{1+\epsilon}}\right]^{\chi} \tag{21}$$

where $\chi \equiv \frac{1+\epsilon}{1+\epsilon-\alpha\left(\frac{\rho-1}{\rho}\right)} > 1.$

The aforementioned functional forms are standard in the literature. Furthermore, our subsequent characterization of the cyclical properties of the optimal monetary policy apply even when the constrained efficient allocation is unattainable. We thus see the log-separability condition as relatively innocuous for applied purposes and take it for granted for the remainder of the paper.

5 Optimal monetary policy

We are now ready to state our first key policy result.

Theorem 1. *Whenever the constrained efficient allocation is implementable as a sticky-price equilibrium, it can be attained with a tax that offsets the monopoly distortion and a monetary policy that replicates the corresponding flexible-price allocation.*

This result has the same flavor as the familiar New-Keynesian result that monetary policy must seek to replicate flexible-price allocations once the monopoly distortion is corrected. There are, however, two important differences: first, the nature of the optimal allocation is different; second, the policy targets that help implement the optimal allocation are also different. Let us elaborate on these points.

Consider first the nature of the optimal allocation. As already discussed, the optimal allocation no more coincides with the first best, because (and only because) of the real bite of the informational friction. The optimality of monetary policies that replicate flexible prices therefore has an entirely new meaning—one that relates to whether firms face the right incentives when responding to all sources of variation in their beliefs about the aggregate state of the economy.

To appreciate this point further, in the next section we illustrate how erratic the observable properties of the constrained efficient allocation might look like relative to the first best—and, by implication, how misleading it could be to guide policy on the basis of the conventional goal of minimizing the familiar output gap. More specifically, we consider an example in which the optimal allocation features the following positive properties: (i) sluggish response to the underlying shocks in preferences and technologies; (ii) waves of optimism and pessimism driven by noise in firm expectations of these fundamentals; (iii) "animal spirits" in the sense of random variation in actual and expected economic activity that is no more spanned by variation in either actual or expected fundamentals; and (iv) random variation in measured aggregate labor wedges and/or price markups. Faced with this kind of phenomena, conventional policy analysis would likely recommend policies that seek to stabilize aggregate output around its first-best level. Furthermore, such policies typically exist. Yet, our results make clear that such policies are undesirable, for the appropriate welfare-based policy benchmark is no more the first best.

Consider next the policy targets that help implement the optimal allocation. In the basic New-Keynesian model, replicating flexible-price allocations is usually synonymous to targeting price stability. Furthermore, as long as the first-best allocation is itself contained in the set of flexible-price allocations, the familiar "divine coincidence" holds: stabilizing the price level also stabilizes the output gap (the gap between the equilibrium and the first-best level of output).[16] By contrast, we now show that the optimal monetary policy targets a negative correlation between the price level and real economic activity.

To characterize the cyclical properties of the optimal monetary policy, recall that the constrained efficient allocation is implementable if and only if it is log-separable. It follows that there exist functions Ψ^ω and Ψ^s such that, along that allocation, $y(\omega_i^t, s^t) = \log \Psi^\omega(\omega_i^t) + \log \Psi^s(s^t)$. Aggregating across firms gives $\log Y(s^t) = \log \mathcal{B}(s^t) + \log \Psi^s(s^t)$, where

$$\mathcal{B}(s^t) \equiv \left[\int \Psi^\omega \left(\omega_i^t\right)^{\frac{\rho-1}{\rho}} d\mathcal{G}_t(\omega_i^t | s^t) \right]^{\frac{\rho}{\rho-1}}. \tag{22}$$

Note that $\Psi^\omega(\omega_i^t)$ can be thought as a proxy for the belief, or "sentiment", of firm i: it captures the variation in a firm's output that is driven by the firm's information ω_i^t. Furthermore, from Proposition 2 we know that $\Psi^\omega(\omega_i^t)$ is an increasing function of n_{it} and k_{it}, which in turn implies that $\mathcal{B}(s^t)$ inherits the cyclical properties of aggregate employment and capital. With these observations in mind, we henceforth interpret $\mathcal{B}(s^t)$ interchangably as a proxy for the average belief in the economy or a measure of aggregate economic activity.

Theorem 2. *Along any sticky-price equilibrium that implements the constrained efficient allocation, the price level is inversely related to real economic activity, as proxied by $\mathcal{B}(s)$. In particular,*

$$\log P(s) - \log P(s') = -\tfrac{1}{\rho} \left[\log \mathcal{B}(s) - \log \mathcal{B}(s')\right] \quad \forall s, s' \in \mathcal{S}^t, \forall t \tag{23}$$

To understand this result, recall that, along any sticky price allocation, the relative output of two firms must be independent of the true state s^t conditional on the union of the information sets of the two firms—or else the prices set by the firms would have to be contingent on that state, which would violate the restriction that prices are sticky. It follows that, along any sticky-price allocation, the nominal prices of any two firms i and j must satisfy the following restriction for all realizations of their information sets:

$$\begin{aligned} \log p(\omega_i^t) - \log p(\omega_j^t) &= -\tfrac{1}{\rho} \left[\log y(\omega_i^t, s^t) - \log y(\omega_j^t, s^t)\right] \\ &= -\tfrac{1}{\rho} \left[\log \Psi^\omega(\omega_i^t) - \log \Psi^\omega(\omega_j^t)\right] \end{aligned}$$

The relative price of two firms is therefore inversely related to the relative "sentiment" of these firms. Intuitively, if optimistic firms are to produce more than pessimistic ones, then they must also face lower relative prices. But since i does not observe ω_j^t and, symmetrically, j does not observe ω_i^t, this is possible if and only if the nominal price of each firm is itself negatively related to the firm's own sentiment: it must be that $\log p(\omega_i^t) = -\frac{1}{\rho} \log \Psi^\omega(\omega_i^t)$ up to a constant, and similarly for j.

[16] The optimality of price stability in New-Keynesian settings has to be qualified by a definition of the appropriate price index. For example, if commodity prices are flexible but wages are sticky, then the appropriate price index is a wage index. The arguments we make in this section have nothing to do with this issue: we establish sub-optimality of stabilizing the "right" price index. More specifically, since the prices that are sticky in our framework are those of intermediate-good producers, the appropriate price index is the CES aggregator P_t of the prices of these producers. This keeps our analysis directly comparable to, inter alia, Goodfriend and King (1997, 2001), Rotemberg and Woodford (1999), Woodford (2003), Adao, Coreia and Teles (2003), Khan, King and Wolman (2003), Galí (2008).

Aggregating this across all firms then implies that the price level must move inversely with $\mathcal{B}(s^t)$, which explains the result.

Recall now our earlier observation that $\mathcal{B}(s^t)$ inherits the cyclical properties of employment and capital. As in the standard RBC, these properties ultimately hinge on how various income and substitution effects play against one another—see the example in Section 7 for a concrete illustration. The cyclical properties of $\mathcal{B}(s^t)$ and thereby that of the price level, are therefore ambiguous in general. That being said, the empirically relevant case is one in which employment and capital are procyclical, and thus $\mathcal{B}(s^t)$ is also procyclical. It then follows from Theorem 2 that the optimal monetary policy "leans against the wind" by targeting a negative correlation between the price level and real output.

Corollary 3. *Suppose that capital and employment are procyclical along the constrained efficient allocation. Then, the optimal monetary policy targets a countercyclical price level.*

6 Discussion

The preceding analysis has established two key results: one regarding the optimality of monetary policies that replicate flexible-price allocations, and another regarding the optimality of monetary policies that "lean against the wind".

Not surprisingly, our first result hinges on the availability of sufficiently rich tax instruments. Consider, in particular, an extension of our framework that introduces random disturbances in the underlying monopoly distortion so as to accommodate the conventional formalization of "cost-push" or "mark-up shocks". Such an extension does not affect the characterization of the constrained efficient allocation, but changes in an important way the tax policies that are needed in order to implement that allocation as an equilibrium under either flexible or sticky prices: the tax rate $\tau(s^t)$ must now vary in tandem with the underlying monopoly distortion. To the extent that such state-contingent taxes are available, our preceding results remain intact (modulo, of course, the state-contingency of the optimal tax). Otherwise, flexible-price allocations cease to be constrained efficient and, by direct implication, the optimal monetary policy would no more replicate them.

The above observation qualifies the applicability of our first result in a manner similar to how one must qualify the standard result regarding the optimality of flexible-price allocations in the New-Keynesian paradigm. A related but more distinct qualification emerges if one considers variants of our framework that endogenize the information structure. In this case, state-contingent taxes may help improve welfare by manipulating either the collection or the aggregation of information.[17] To the extent that such taxes are unavailable, the optimal monetary policy may once again deviate from replicating flexible-price allocations in the direction of mimicking the unavailable taxes.

Characterizing the precise nature of the optimal policy under the aforementioned circumstances is beyond the scope of this paper, and is bound to come at the expense of generality.[18] We neverthe-

[17] This possibility has been illustrated in an abstract setting by Angeletos and Pavan (2008). For business-cycle applications, see an earlier version of our paper (Angeletos and La'O, 2008) and the complementary work by Paciello and Wiederholt (2011).

[18] For instance, Ball et al (2005), Adam (2007) and Paciello and Wiederholt (2011) allow for mark-up shocks, but

less expect none of these extensions to seriously affect our second result, namely the one regarding the optimality of policies that "lean against the wind".

Consider, in particular, the possibility of mark-up shocks. In the absence of informational frictions, these shocks provide the conventional rationale for countercyclical monetary policies: such policies mimic the countercyclical subsides that would be needed in order to restore the optimality of flexible-price allocations. We see no reason why this familiar rationale should cease to apply in the presence of informational frictions. Instead, we find it more interesting that our analysis provides a novel rationale for this kind of policies, *without* upsetting the optimality of flexible-price allocations.[19]

Consider, next, the possibility of endogenizing the information structure. In an earlier version of this paper (Angeletos and La'O, 2008), we studied a particular example in which information was allowed to be imperfectly aggregated through certain macroeconomic statistics, such as an index of the price level. We found that allowing for this possibility only reinforced our second result: policies that "lean against the wind" were optimal, not just because of the reasons identified in our preceding analysis, but also because they facilitated the revelation of valuable information through the co-variation of the price level with the underlying aggregate shocks.

To reinforce the case about the likely robustness of our results, we should make the following important observation. Consider condition (23), which establishes a negative relation between the price level $P(s^t)$ and the level of economic activity as measured by $\mathcal{B}(s^t)$. As it can be seen in the proof of Theorem 2 in the Appendix, this condition applies to *every* sticky-price allocation, not just the optimal one. Furthermore, this condition is entirely invariant to either the stochastic structure of the economy or the specification of how information gets collected by firms, aggregated through markets, or communicated by the government.

Different assumptions about the aforementioned primitives therefore map into different predictions about the precise stochastic properties of the optimal allocation and the associated value of $\mathcal{B}(s^t)$. Nonetheless, whatever these primitives and the resulting optimal allocation might be, condition (23) remains valid. It follows that, even if we consider situations where the constrained efficient allocations fails to be implementable or the information structure is endogenous in arbitrary ways, Theorem 2 and Corollary 3 continue to hold as long as we replace the words "constrained efficient allocation" with the words "optimal allocation", whereby we mean the allocation that maximizes welfare within the set of sticky-price allocations.

A similar point applies if we consider variants where the policy maker observes the underlying state of nature and the equilibrium macroeconomic outcomes with measurement error, or is perhaps himself "inattentive" when setting the relevant policy instruments. In this case, the policy maker will of course fail to attain the constrained efficient allocation: the optimal allocation will only be a noised-up proxy of the constrained-efficient one. Nevertheless, condition (23) will continue to hold, guaranteeing that price stability is suboptimal.

only at the expense of assuming away the real bite of informational frictions, which is the central aspect of our contribution. Similarly, the example we consider in Angeletos and La'O (2008) allows for endogenous information aggregation, but restricts attention to a particular Gaussian specification for the information structure.

[19]Plus, recall our earlier point that the fluctuations that obtain along the constrained efficient allocation may be mistaken as the product of mark-up shocks.

These observations underscore, not only the likely robustness of our second result, but also its precise origin: the sub-optimality of price stability we document in this paper derives directly, and exclusively, from the real bite of the informational friction.

To further appreciate this last point, consider for a moment the alternative scenario in which the informational friction has only a nominal bite. This scenario can be nested in our framework by assuming that k_{it} and n_{it} are exogenously fixed. Alternatively, we can consider a variant in which these choices are free to adjust to the true state s^t along with h_{it}. In either case, the optimal allocation is also free to adjust to the true state, which in turn guarantees that there is no reason for the nominal price of each firm to vary with any noisy information that the firm may have about the underlying state at the time it sets its price. It then follows that $\mathcal{B}_t$ is a constant and that price stability is trivially optimal. But as soon as *some* production choices are made on the basis of incomplete and heterogeneous information, the nominal price set by each firm must co-move with the belief of that firm, or else these choices would fail to respond efficiently to the underlying aggregate shocks. It then follows that $\mathcal{B}_t$ is no more a constant and, by direct implication, price stability is no more optimal.

The above discussion also explains the difference between our result and that of Adam (2007), Ball, Mankiw and Reis (2005), and Paciello and Wiederholt (2011). In the absence of markup shocks, these papers find that price stability is optimal. At face value, this seems to contradict our own result. The resolution rests precisely on recognizing that these papers treat the informational friction merely as a nominal friction, while our paper allows it to have a real bite. As soon as this is the case, the optimality of price stability featured in these papers breaks down.[20]

We conclude this section with a brief comment on the robustness of our results to the introduction of idiosyncratic productivity shocks. We consider a particular illustration of this possibility in the next section and a more general extension in Appendix B. Constrained efficiency now requires that the relative output and the relative price of two firms reflect, not only their belief differences, but also their productivity differences. Intuitively, more productive firms must produce more and must therefore also face lower relative prices. When prices are flexible, this requirement poses no problem for implementability. But once prices are sticky, this requirement can be satisfied only in so far as firms have sufficient information about their idiosyncratic shocks when setting their prices. As long as this requirement is satisfied, Theorems 1 and 2 and Corollary 3 remain valid. Otherwise, the aforementioned generalizations of Theorem 2 and Corollary 3, which simply replace "constrained efficient allocation" with "optimal allocation", apply.

7 An illustration

In this section we illustrate our results with the help of an example that abstracts from capital, assumes i.i.d. TFP shocks, and imposes a tractable Gaussian information structure. The goal here is to illustrate in a clean and sharp manner (i) how the constrained efficient allocation can differ from the first-best one and (ii) how the optimal monetary policy targets a negative relation between

[20]To be precise, our statement assumes away non-generic cases in which price stability remains optimal either because all information is perfectly common or because production choices are invariant to the state of nature.

the price level and real economic activity. We do so first for the case with only aggregate TFP shocks and then for a variant that adds idiosyncratic TFP shocks.

Preferences take a conventional power-form specification:

$$U(C) = \frac{C^{1-\gamma}}{1-\gamma} \qquad \text{and} \qquad V(h) = 1 + \frac{1}{1+\epsilon} h^{1+\epsilon} \tag{24}$$

with $\gamma, \epsilon \geq 0$. Capital is fixed at $k_{it} = 1$ and the technology is Cobb-Douglas:

$$y_{it} = A_t \ell_{it}^{\alpha}, \tag{25}$$

with $\alpha \in (0,1)$. The TFP shock is log-normal and i.i.d. over time: $a_t \equiv \log A_t \sim \mathcal{N}(0, \sigma_A^2)$, with $\sigma_A > 0$. Finally, in each period t, each firm i observes $\omega_{it} = (x_{it}, z_{it})$, where

$$x_{it} = a_t + v_{it} \qquad \text{and} \qquad z_{it} = a_t + \varepsilon_t + u_{it},$$

and where $v_{it} \sim \mathcal{N}(0, \sigma_v^2)$ and $u_{it} \sim \mathcal{N}(0, \sigma_u^2)$ are idiosyncratic noises, $\varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)$ is a common noise, and $\sigma_v, \sigma_u, \sigma_\epsilon > 0$. The first signal can be interpreted as any private information that each firm might have about its own productivity, which here coincides with the aggregate one—a variant of the present example that delivers similar results is one in which the log-productivity of firm i is itself given by $x_{it} = a_t + v_{it}$, in which case v_{it} represents an idiosyncratic productivity shock. The second signal can be interpreted as public sources of information about the state of the economy, in which case the correlated error ε_t may reflect, for example, measurement error in macroeconomic statistics while the idiosyncratic error u_{it} may be interpreted as the product of limited attention. Alternatively, ε_t can be thought of as a proxy for either the kind of noise shocks studied in Lorenzoni (2009) or the kind of "animal spirits" studied in Angeletos and La'O (2012).

Let N_t be aggregate employment, defined as the average of n_{it} across the firms. As a reference point, we first characterize the complete-information first-best allocation.[21]

Proposition 5. *There exist scalars Λ_a^* and Φ_a^* such that the first best satisfies*

$$\log n_{it} = \log N_t = \Lambda_a^* a_t \qquad \textit{and} \qquad \log y_{it} = \log Y_t = \Phi_a^* a_t$$

Furthermore, $\Phi_a^* > 0$ necessarily, while $\Lambda_a^* > 0$ if and only if $\gamma < 1$.

In the first best, the levels of employment and output are identical across firms, since there is no heterogeneity, and the business cycle is driven merely by the aggregate TFP shock. As already explained, if the informational friction had only a nominal bite (as in much of the literature), the first-best allocation would be implementable with a monetary policy that targets price stability. But now that the informational friction has a real bite, the first-best is no more attainable. Instead, policy must be guided by the notion of constrained efficiency we have developed in this paper. For the example under consideration, this means the following.

[21] In all the results that follow, we omit the intercepts of the various policy rules and report only their stochastic components. For instance, the first condition in Proposition ?? must be read as $\log n_{it} = \log N_t = const + \Lambda_a^* a_t$ for some constant *const*, and similarly for the second condition.

Proposition 6. *There exist scalars $\lambda_x, \lambda_z, \phi_x, \phi_z, \phi_a \in \mathbb{R}$ and $\Lambda_a, \Lambda_\varepsilon, \Phi_a, \Phi_\varepsilon \in \mathbb{R}$ such that the following properties hold along the constrained efficient allocation:*

(i) individual employment and output are given by

$$\begin{aligned} \log n_{it} &= \lambda_x x_{it} + \lambda_z z_{it} \\ \log y_{it} &= \phi_x x_{it} + \phi_z z_{it} + \phi_a a_t \end{aligned}$$

(ii) aggregate employment and output are given by

$$\begin{aligned} \log N_t &= \Lambda_a a_t + \Lambda_\varepsilon \varepsilon_t \\ \log Y_t &= \Phi_a a_t + \Phi_\varepsilon \varepsilon_t \end{aligned}$$

Furthermore, $\Lambda_a > 0$ if and only if $\Lambda_a^* > 0$, and either one implies all of the following: $0 < \Lambda_a < \Lambda_a^*$, $0 < \Phi_a < \Phi_a^*$, $\Lambda_\varepsilon > 0$, and $\Phi_\varepsilon > 0$.

To understand part (i), recall that the number of workers hired by a firm is contingent only on that firm's information, which explains why n_{it} is a function of only x_{it} and z_{it}. But as the level of effort h_{it} adjusts to the true productivity shock, so does the firm's output, which explains why y_{it} is also a function of a_t. Part (ii) then follows from aggregating across firms and noting that, as the idiosyncratic noises v_{it} and u_{it} wash out at the aggregate, macroeconomic outcomes depend only on the TFP shock a_t and the correlated noise shock ε_t.

Comparing Proposition 5 and Proposition 6, we see that the first best and the constrained efficient allocations differ in a number of ways. First, under the constrained efficient allocation, a firm's employment and output respond to the idiosyncratic noise in the firm's information. Thus, unlike the first best, the constrained efficient allocation features a certain level of noise-driven heterogeneity in real outcomes. Second, firms also respond to the correlated error ε_t. Thus, while the business cycle is driven only by TFP shocks along the first best, now it is also driven by a certain type of noise shocks. Finally, for the natural case where $\Lambda_a^* > 0$, we have that employment and output exhibit a dampened response to TFP innovations in the constrained efficient allocation relative to the first best.

The above results verify our earlier discussion regarding how different the positive properties of the constrained efficient allocation might be relative to the first best. We now proceed to illustrate how the optimal monetary policy "leans against the wind". To this goal, recall from Theorem 2 that the price level hinges on the aggregate belief proxy $\mathcal{B}(s^t)$. In the present example, it is easy to show that there exists a scalar $\vartheta > 0$ such that $\Psi^\omega(\omega_{it}) = \vartheta \log n(\omega_{it})$ and therefore $\mathcal{B}(s^t) = \vartheta \log N(s^t)$. The following is then immediate.

Proposition 7. *There exists a scalar $\psi > 0$ such that, in any sticky-price equilibrium that implements the constrained efficient allocation, prices satisfy the following properties:*

$$\log p_{it} = -\psi \log n_{it} = -\psi \left(\lambda_x x_{it} + \lambda_z z_{it}\right)$$

$$\log P_t = -\psi \log N_t = -\psi \left(\Lambda_a a_t + \Lambda_\varepsilon \varepsilon_t\right)$$

That is, for the constrained efficient allocation to be implemented, the price set by a firm must be inversely related to its efficient level of employment, and the aggregate price level must itself be inversely related to aggregate employment. It follows that the price level is counter-cyclical if and only if aggregate employment is pro-cyclical along the constrained efficient allocation.

This is of course merely a clean illustration of the more general result in Corollary 3. The tractability of the present example, however, permit us to characterize the cyclical properties of the efficient level of employment, and thereby of the price level, in terms of the primitive parameters. In particular, we can show that the noise shock always induces positive co-movement between employment and output. The TFP shock, on the other hand, can induce either positive or negative co-movement, depending on whether it is income or substitution effects that dominate. As in the case of complete information, the substitution effect turns out to dominate if and only if $\gamma < 1$. It follows that $\gamma < 1$ suffices for the overall co-movement of employment and output to be positive, and therefore for the price level to be counter-cyclical. Finally, Λ_a and Λ_ε are zero if and only if $\gamma = 1$, which proves that price stability is optimal only in the degenerate case in which the efficient level of employment is entirely invariant to the underlying state of nature and, therefore, the real bite of the informational friction is no more binding.

We now consider variant that adds idiosyncratic TFP shocks. In particular, we continue to assume that i's signal in period t is given by $\omega_{it} = (x_{it}, z_{it})$, but re-reinterpret x_{it} as firm-specific TFP. Individual output is thus given by $y_{it} = A_{it}\ell_{it}^{\alpha}$, with

$$\log A_{it} = x_{it} = \log A_t + v_{it}.$$

We like this variant for four reasons. First, it illustrates how our results can be robust to the introduction of idiosyncratic productivity shocks. Second, this variant seems consistent with the spirit of rational inattention, which predicts that agents should pay more attention to the variables that matter most for their decisions.[22] Finally, this variant permits us to disentangle the information a firm has about its own technology and marginal costs from the information it has about "aggregate demand". To see what we mean by this, note that, since the idiosyncratic shock v_{it} is much more volatile than the aggregate shock a_t, we can think of x_{it} as a signal that is nearly perfectly informative of the firm's idiosyncratic conditions but is almost entirely uninformative about macroeconomic conditions. At the same time, we can control how much information firms have about the exogenous aggregate state, and thereby about all endogenous macroeconomic outcomes, through the signal z_{it}. In line with Lorenzoni (2009) and Angeletos and La'O (2009, 2012), this variant thus also permits us to interpret a_t as a "supply" shock and ε_t as an "expectational" or "demand" shock.[23]

[22] The assumption that firms observe a signal of A_{it} rather of η_{it} is squarely consistent with rational inattention, since for any given aggregate output Y_t a firm's profits hinge only on A_{it}, not on its underlying components. At the same time, the assumption that this signal is perfect is only for simplicity. If we relax this assumption, the constrained efficient allocation is no more be implementable, but the optimal implementable allocation inherits the qualitative properties of the constrained one.

[23] To see this more clearly, note that the following properties hold whenever each firm knows its own TFP. When a_t increases, each firm recognizes that its own productivity has increased, and therefore that its own marginal cost function has fallen—which one can think of as a supply shock. When instead ε_t increases, each firm recognizes that its

Putting aside these motivating factors, this variant is also interesting because it gives us a particularly simple characterization for the optimal monetary policy: for the efficient allocation to be implemented, the price level must satisfy

$$\log P(s^t) \;=\; -\,\Gamma \cdot \log Y(s^t) \tag{26}$$

for some constant $\Gamma > 0$ that depends only on the parameters ρ, α, ϵ, and γ (see Appendix). The price level is therefore necessarily counter-cylical, and the policy target is a simple linear rule between the price level and aggregate output.

In fact, as long as individual firms know their own productivity, the above result continues to hold even if we allow for persistence in the underlying aggregate and/or idiosyncratic TFP shocks, or if we let the firms receive arbitrary Gaussian signals about the aggregate shocks. This variant thus illustrates that the monetary policy might be able to implement, or at least proxy, the optimal allocation with a simple target rule that does not depend on the details of either the stochastic structure of the productivity shocks or the information structure. Furthermore, to the extent that Γ is close enough to 1, one can also proxy this with a policy that targets nominal GDP stabilization—a target that has been proposed by some economists, albeit without clear micro-foundations.

8 Concluding remarks

In the past decade, a rapidly growing literature has renewed interest on the macroeconomic implications of informational frictions. Nevertheless, this literature has largely shied away from studying how informational frictions impact the normative properties of the business cycle and the design of optimal monetary policy in so far as these frictions constrain, not only the price-setting decisions of firms, but also their employment and production responses.

This paper contributes towards filling this gap and, in so doing, it also revisits the familiar Ramsey (or primal) approach in the presence of informational frictions. In this regard, the definition and characterization of the appropriate notions of constrained efficient, flexible-price, and sticky-price allocations are integral parts of our own contribution. What is more, by showing how the primal approach—which has proved so fruitful in the past—can be adapted to models with informational frictions, we hope to make a broader methodological contribution that may prove useful beyond the boundaries of the particular policy questions we address in this paper.

The framework within which we conduct our analysis bypasses any specific "micro-foundation" of the informational friction and, instead, captures the latter by an arbitrary measurability constraint on the set of feasible allocations and prices. Most importantly, it allows the informational friction to have a bite, not only on the firms' price-setting decisions, but also on some of their real production choices. In these respects, our analysis provides a clean and quite general benchmark for understanding how the real bite of informational frictions impact the optimality of flexible-price allocations and the nature of optimal monetary policy.

productivity and its marginal cost function have stayed the same, yet each firm expects aggregate economic activity to expand and thereby the demand for its product to increase—which one can think of as a demand shock.

As with any other benchmark, our analysis is of course subject to limitations. The one that seems most intriguing from a conceptual perspective regards the potential endogeneity of the information structure, while the one that seems of first order from a practical perspective regards the fact that we ruled out conventional, Calvo-like, price stickiness. Our analysis has also focused on how informational frictions impact firm decisions; adding such frictions on the consumer side need not upset the key insights of this paper but is certainly a natural direction for future research.

References

[1] Adao, Bernardino, Isabel Correia, and Pedro Teles (2003), "Gaps and Triangles," *Review of Economic Studies* 70(4), 699-713.

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

[3] Amador, Manuel, and Pierre-Olivier Weill (2011), "Learning from Private and Public Observations of Others Actions,"*Journal of Economic Theory*, forthcoming.

[4] Angeletos, George-Marios, and Jennifer La'O (2008), "Dispersed Information over the Business Cycle: Optimal Fiscal and Monetary Policy," MIT mimeo.

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

[6] Angeletos, George-Marios, and Jennifer La'O (2012), "Sentiments," MIT/Chicago Booth mimeo.

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

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

[9] Atkeson, Andrew, Varadarajan V. Chari, and Patrick J. Kehoe (2010), "Sophisticated Monetary Policies,"*Quarterly Journal of Economics* 125, 47-89.

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

[11] Bergemann, Dirk, and Stephen Morris, (2011), "Robust Predictions in Games of Incomplete Information," Princeton ETC working paper.

[12] Cochrane, John (2011), "Determinacy and Identification with Taylor Rules, " *Journal of Political Economy* 119, 565-615.

[13] Coibion, Olivier, and Yuriy Gorodnichenko (2012), "What can survey forecasts tell us about informational rigidities?", *Journal of Political Economy* 120, 116-159.

[14] 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.

[15] Galí, Jordi (2008), *Monetary Policy, Inflation and the Business Cycle: An Introduction to the New Keynesian Framework,* Princeton: Princeton University Press.

[16] Goodfriend, Marvin, and Robert King (1997), "The New Neoclassical Synthesis and the Role of Monetary Policy," *NBER Macroeconomics Annual 1997*, Cambridge: MIT Press.

[17] Goodfriend, Marvin, and Robert King (2001), “The Case for Price Stability,” NBER Working Paper 8423.

[18] Hellwig, Christian (2002), “Public Announcements, Adjustment Delays and the Business Cycle,” UCLA mimeo.

[19] Khan, Aubhik, Robert King, and Alexander Wolman (2003), “Optimal Monetary Policy,” *Review of Economic Studies* 70, 825-860.

[20] Jaimovich, Nir, and Sergio Rebelo (2009), “Can News About the Future Drive the Business Cycle?,” *American Economic Review* 99, 1097-1118.

[21] Lorenzoni, Guido (2009), “A Theory of Demand Shocks,” *American Economic Review* 99(5), 2050-2084.

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

[23] Lucas, Robert E., Jr. (1972), “Expectations and the Neutrality of Money,” *Journal of Economic Theory* 4, 103-124.

[24] 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.

[25] Mackowiak, Bartosz, and Mirko Wiederholt (2009), “Optimal Sticky Prices under Rational Inattention,” *American Economic Review*, 99, 769-803.

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

[27] Morris, Stephen, and Hyun Song Shin (2002), “The Social Value of Public Information,” *American Economic Review* 92, 1521-1534.

[28] Nimark, Kristoffer (2008), “Dynamic Pricing and Imperfect Common Knowledge,” *Journal of Monetary Economics* 55, 365-382.

[29] Paciello, Luigi, and Mirko Wiederholt (2011), “Exogenous Information, Endogenous Information, and Optimal Monetary Policy,” Northwestern/EIEF Mimeo.

[30] Rotemberg, Julio, and Michael Woodford (1999), “The Cyclical Behavior of Prices and Costs.,” in J. B. Taylor & M. Woodford (eds.), *Handbook of Macroeconomics*, Elsevier.

[31] Sims, Christopher (2003), “Implications of Rational Inattention,” *Journal of Monetary Economics* 50, 665-690.

[32] Woodford, Michael (2003a), “Imperfect Common Knowledge and the Effects of Monetary Policy,” in P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford, eds., *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, Princeton University Press.

[33] Woodford, Michael (2003b), *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton: Princeton University Press.

Appendix A: Proofs

Proof of Proposition 1. This follows from taking the first-order conditions of the planner's problem defined in page 3 and noting that this problem is strictly convex, guaranteeing that these conditions are both necessary and sufficient for efficiency. **QED.**

Proof of Proposition 2. To prove necessity, note first that the orthogonality conditions (7)-(9) follow directly from the analysis in the main text. All that remains is therefore to show that, as usual, resource feasibility follows from the combination of budgets and market clearing. Thus consider the nominal budget constraint of the household, which is given by

$$P\left(s^t\right)C\left(s^t\right) = \int \left(\Pi\left(\omega_i^t, s^t\right) + P\left(s^t\right)W(\omega_i^t, s^t)n\left(\omega_i^t\right)\right)d\mathcal{G}_t(\omega_i^t|s^t) + T\left(s^t\right).$$

Firm profits, in turn, are given by

$$\Pi\left(\omega_i^t, s^t\right) = \left(1-\tau_t\right)p(\omega_i^t, s^t)y\left(\omega_i^t, s^t\right) - P\left(s^t\right)W(\omega_i^t, s^t)n\left(\omega_i^t\right) - P\left(s^t\right)x\left(\omega_i^t\right),$$

Aggregating over profits, using the budget constraint of the government, and capital accumulation equation, we get that

$$P\left(s^t\right)C\left(s^t\right) = \int p\left(\omega_i^t\right)y\left(\omega_i^t, s^t\right)d\mathcal{G}_t(\omega_i^t|s^t) - P\left(s^t\right)\left(K\left(s^t\right) - \left(1-\delta\right)K\left(s^{t-1}\right)\right)$$

Using then the fact that $\int p\left(\omega_i^t\right)y\left(\omega_i^t, s^t\right)d\mathcal{G}_t(\omega_i^t|s^t) = P_tY_t$, we arrive at the resource constraint, as expected.

To prove sufficiency, we need to find prices and policies that sustain the candidate allocation as an equilibrium. Thus, take any allocation that satisfies conditions (7)-(9) and let us propose the following contingent nominal prices:

$$p(\omega_i^t, s^t) = \left(\frac{y(\omega_i^t, s^t)}{Y(s^t)}\right)^{-\frac{1}{\rho}}$$

It follows that price level is constant:

$$P(s^t) = 1$$

Clearly, these nominal prices implement the right relative prices. Next, let the real wage schedule, the tax rate and the monetary policy be given by

$$W(h, s^t) = \frac{V(h, \xi(s^t))}{U_c(s^t)}, \qquad \tfrac{\rho-1}{\rho}(1-\tau_t(s^t)) = \phi(s^t), \qquad \text{and} \qquad M(s^t) = Y(s^t),$$

Finally, let the nominal interest rate—also the real one since the price level is constant—be

$$R_t = \frac{U_c(s^t)}{\mathbb{E}\left[U_c(s^{t+1})\right| s^t]}$$

and set $B(s^t) = 0$ and $T(s^t) = \tau(s^t) \int p(\omega_i t) y(\omega_i^t, s^t) d\mathcal{F}(\omega_i^t | s^t)$.

With the prices and the policies defined as above, the following are true. First, the optimality conditions of the household and the final-good sector, the cash-in-advance constraint, and the government budget are satisfied automatically. Second, the budget constraint of the household and market clearing then follow from the resource constraints. Finally, the optimality conditions of the monopolistic firms follow from conditions (7)-(9). We conclude that the aforementioned prices and policies sustain that candidate allocation as part of a flexible-price equilibrium, which completes the proof. **QED.**

Proof of Proposition 3. To prove necessity, note that feasibility follows again from the combination of budgets and market clearing. Next, from the analysis in the main text, we have that any sticky-price allocation satisfies the following conditions:

$$\begin{aligned}
\mathbb{E}\left[U_c(s^t) y\left(\omega_i^t, s^t\right) \left\{ \phi(s^t) \left(\frac{y\left(\omega_i^t, s^t\right)}{Y(s^t)} \right)^{-\frac{1}{\rho}} - \lambda\left(\omega_i^t, s^t\right) \right\} \middle| \omega_i^t \right] &= 0 \qquad \forall\, \omega_i^t \\
\mathbb{E}\left[U_c(s^t) - \beta U_c(s^t) \left\{ 1 - \delta + \lambda\left(\omega_i^{t+1}, s^{t+1}\right) A(\omega_i^{t+1}) f_k\left(\omega_i^{t+1}, s^{t+1}\right) \right\} \middle| \omega_i^t \right] &= 0 \qquad \forall\, \omega_i^t \\
\mathbb{E}\left[V(\omega_i^t, s^t) - U_c(s^t) \lambda(\omega_i^t, s^t) A(s^t) f_\ell\left(\omega_i^t, s^t\right) h(\omega_i^t, s^t) \middle| \omega_i^t \right] &= 0 \qquad \forall\, \omega_i^t \\
V_h(\omega_i^t, s^t) - U_c(s^t) \lambda\left(\omega_i^t, s^t\right) A(s^t) f_\ell\left(\omega_i^t, s^t\right) &= 0 \qquad \forall\, \omega_i^t, s^t
\end{aligned}$$

where, for any $z \in \{\ell, k\}$, $f_z\left(\omega_i^t, s^t\right) \equiv \frac{\partial}{\partial z} F\left(k\left(\omega_i^{t-1}\right), \ell\left(\omega_i^t, s^t\right)\right)$. Conditions (10)-(13) in part (i) follow directly from the above once we let

$$\chi(\omega_i^t, s^t) = \frac{\lambda(\omega_i^t, s^t)}{\phi(s^t) \left(\frac{y\left(\omega_i^t, s^t\right)}{Y(s^t)} \right)^{-\frac{1}{\rho}}}$$

Finally, part (ii), namely the fact that any equilibrium allocation is log-separable, follows directly from condition (6) if we let $\Psi^\omega\left(\omega_i^t\right) \equiv p\left(\omega_i^t\right)^{-\rho}$ and $\Psi^s\left(s^t\right) \equiv Y(s^t) P(s^t)^\rho$.

Consider now sufficiency. Take any allocation that satisfies properties (i) and (ii) in the proposition; we need to find prices and policies that sustain the candidate allocation as an equilibrium. Because the allocation is separable, we have that

$$y\left(\omega_i^t, s^t\right) = \Psi^\omega(\omega_i^t) \Psi^s(s^t)$$

for some functions Ψ^ω and Ψ^s. Let us then propose the following nominal prices:

$$p(\omega_i^t) = \Psi^\omega\left(\omega_i^t\right)^{-\frac{1}{\rho}},$$

which are clearly measurable in ω_i^t. It follows that the price level satisfies

$$P(s^t) = \left[\int p\left(\omega_i^t\right)^{1-\rho} d\mathcal{G}_t(\omega_i^t|s^t)\right]^{\frac{1}{1-\rho}} = \left[\int \Psi^\omega\left(\omega_i^t\right)^{\frac{\rho-1}{\rho}} d\mathcal{G}_t(\omega_i^t|s^t)\right]^{\frac{1}{1-\rho}},$$

while aggregate output satisfies

$$Y\left(s^t\right) = \Psi^s\left(s^t\right)\left[\int \Psi^\omega\left(\omega_i^t\right)^{\frac{\rho-1}{\rho}} d\mathcal{G}_t(\omega_i^t|s^t)\right]^{\frac{\rho}{\rho-1}},$$

and therefore relative prices satisfy

$$\frac{p(\omega_i^t)}{P(s^t)} = \frac{\Psi^\omega\left(\omega_i^t\right)^{-\frac{1}{\rho}}}{\left[\int \Psi^\omega\left(\omega_i^t\right)^{\frac{\rho-1}{\rho}} d\mathcal{G}_t(\omega_i^t|s^t)\right]^{\frac{1}{1-\rho}}} = \left(\frac{y(\omega_i^t, s^t)}{Y(s^t)}\right)^{-\frac{1}{\rho}}$$

That is, we can find nominal prices that implement the right relative prices while being measurable in ω_i^t. Next, let the real wage, the tax rate and the monetary policy be given by

$$W(h, s^t) = \frac{V(h, \xi(s^t))}{U_c(s^t)}, \qquad \tfrac{\rho-1}{\rho}(1 - \tau_t(s^t)) = \phi(s^t), \qquad \text{and} \qquad M(s^t) = P(s^t)Y(s^t).$$

Finally, let the nominal interest rate be

$$R_t = \frac{U_c(s^t)/P(s^t)}{\mathbb{E}\left[U_c(s^{t+1})/P(s^{t+1})|\, s^t\right]}$$

and set $B(s^t) = 0$ and $T(s^t) = \tau(s^t)\int p(\omega_i^t)y(\omega_i^t, s^t)d\mathcal{F}(\omega_i^t|s^t)$. With the prices and policies defined as above, the optimality conditions of the household and the final-good sector, the cash-in-advance constraint, and the government budget are satisfied automatically; the budget constraint of the household and market clearing then follow from the resource constraints; and the optimality conditions of the monopolistic firms follow from conditions (10)-(13). We conclude that the aforementioned prices and policies sustain the candidate allocation as part of a sticky-price equilibrium, which completes the proof. **QED.**

Proof of Proposition 4. Take any flexible-price equilibrium. From Proposition 2 we know that $h(\omega_i^t, s^t)$ is pinned down by condition (7). Since the technology satisfies (20), this condition can be expressed as

$$V_h(\omega_i^t, s^t) = U_c(Y\left(s^t\right)\phi(s^t)\left(\frac{y\left(\omega_i^t, s^t\right)}{Y(s^t)}\right)^{-\frac{1}{\rho}}\alpha\frac{y\left(\omega_i^t, s^t\right)}{n\left(\omega_i^t\right)h\left(\omega_i^t, s^t\right)}$$

Thus, with the assumed specification for $V\left(h\right)$ in (20), for any realization of $\left(\omega_i^t, s^t\right)$ the following two equations must hold:

$$v_1 h\left(\omega_i^t, s^t\right)^\epsilon = U_c(Y\left(s^t\right)\phi(s^t)\left(\frac{y\left(\omega_i^t, s^t\right)}{Y(s^t)}\right)^{-\frac{1}{\rho}}\alpha\frac{y\left(\omega_i^t, s^t\right)}{n\left(\omega_i^t\right)h\left(\omega_i^t, s^t\right)} \tag{27}$$

$$y\left(\omega_i^t, s^t\right) = A\left(s^t\right)k\left(\omega_i^t\right)^{1-\alpha}\left(n\left(\omega_i^t\right)h\left(\omega_i^t, s^t\right)\right)^\alpha \tag{28}$$

We can solve (27) and (28) simultaneously for $y\left(\omega_i^t, s^t\right)$ and $h(\omega_i^t, s^t)$. We thereby get that equilibrium output is given by

$$y\left(\omega_i^t, s^t\right) = \left[A\left(s^t\right) k\left(\omega_i^t\right)^{1-\alpha} n\left(\omega_i^t\right)^{\alpha\left(1-\frac{1}{1+\epsilon}\right)} \left[U_c(Y\left(s^t\right))\phi(s^t)Y(s^t)^{\frac{1}{\rho}} \frac{\alpha}{v_1}\right]^{\frac{\alpha}{1+\epsilon}}\right]^{1/\left(1-\frac{\alpha}{1+\epsilon}\left(\frac{\rho-1}{\rho}\right)\right)}$$

Thus, output $y\left(\omega_{it}, s^t\right)$ is log-separable in ω_i^t and s^t

$$y\left(\omega_{it}, s^t\right) = \Psi^\omega(\omega_i^t)\Psi^s(s^t)$$

with

$$\Psi^\omega(\omega_i^t) = \left[k\left(\omega_i^t\right)^{1-\alpha} n\left(\omega_i^t\right)^{\alpha\left(\frac{\epsilon}{1+\epsilon}\right)}\right]^{\frac{1+\epsilon}{1+\epsilon-\alpha\left(\frac{\rho-1}{\rho}\right)}} \quad (29)$$

$$\Psi^s(s^t) = \left[A\left(s^t\right)\left(U_c(Y\left(s^t\right))\phi(s^t)Y(s^t)^{\frac{1}{\rho}}\right)^{\frac{\alpha}{1+\epsilon}}\right]^{\frac{1+\epsilon}{1+\epsilon-\alpha\left(\frac{\rho-1}{\rho}\right)}} \quad (30)$$

where we abstract from the constant scalar $(\alpha/v_1)^{\frac{\alpha}{1+\epsilon-\alpha\left(\frac{\rho-1}{\rho}\right)}}$.

This confirms that, with the Cobb-Douglas production function and power form for the disutility of effort as in (20), *every* flexible-price equilibrium allocation is log-separable, and can therefore be replicated under sticky prices. **QED.**

Proof of Theorem 1. First, comparing the conditions in Proposition 2 to those in Proposition 1, we see that the only difference between the set of flexible-price allocations and the constrained efficient allocation is the wedge $\phi(s^t)$ appearing in conditions (7)-(9), relative to the corresponding conditions for the constrained efficient allocation, namely conditions (1)-(3). This wedge is simply the product of the monopoly markup and the tax rate:

$$\phi(s^t) = \left(\frac{\rho-1}{\rho}\right)\left(1-\tau(s^t)\right).$$

Now, take the constrained efficient allocation. By Proposition 1, this allocation necessarily satisfies conditions (1)-(3). Conditions (7)-(9) of Proposition 2 are then trivially satisfied once we let $\phi(s^t) = 1$ for all s^t. Therefore, the constrained efficient allocation can always be implemented as a flexible-price allocation with a non-contingent subsidy that merely offsets the monopoly distortion, namely $\tau(s^t) = \tau^* \equiv -\frac{1}{\rho-1} \quad \forall\ s^t$.

Second, comparing the conditions in Proposition 3 to those in Proposition 2, we see that the only difference between the set of sticky-price allocations and the set of flexible-price allocations is captured by the additional wedge $\chi_{it} = \chi(\omega_{it}, s^t)$ appearing in conditions (15)-(17) relative to the corresponding conditions for flexible-price allocations, namely conditions (7)-(9). Now, take any flexible-price allocation. By Proposition 2, this allocation necessarily satisfies conditions (7)-(9). Conditions (15)-(18) are then trivially satisfied once we let $\chi(\omega_{it}, s^t) = 1$ for all (ω_{it}, s^t). That is, any flexible-price allocation necessarily satisfies part (i) of Proposition 2. It follows that a flexible-price allocation can be replicated under sticky prices if and only if it satisfies part (ii) of that proposition, that is, if and only if it is log-separable.

Together, this implies that whenever the constrained efficient allocation is implementable as a sticky-price equilibrium (that is, whenever the constrained efficient allocation is log-separable), it is implemented with a tax that offsets the monopoly distortion and a monetary policy that replicates the corresponding flexible-price allocation. **QED.**

Proof of Theorem 2. Following the proof of Proposition 3, for any arbitrary common-knowledge process J_t', nominal prices are given by

$$p(\omega_i^t) = e^{J_t'} \Psi^\omega \left(\omega_i^t\right)^{-\frac{1}{\rho}}$$

It follows that the aggregate price level is given by

$$P(s^t) = \left[\int p\left(\omega_i^t\right)^{1-\rho} d\mathcal{G}_t(\omega_i^t|s^t)\right]^{\frac{1}{1-\rho}} = e^{J_t'} \left[\int \Psi^\omega \left(\omega_i^t\right)^{\frac{\rho-1}{\rho}} d\mathcal{G}_t(\omega_i^t|s^t)\right]^{\frac{1}{1-\rho}},$$

We may thus express the aggregate price level in terms of $\mathcal{B}(s^t) \equiv \left[\int \Psi^\omega \left(\omega_i^t\right)^{\frac{\rho-1}{\rho}} d\mathcal{G}_t(\omega_i^t|s^t)\right]^{\frac{\rho}{\rho-1}}$ as follows

$$P(s^t) = e^{J_t'} \mathcal{B}(s^t)^{-\frac{1}{\rho}}$$

Condition (23) then follows. **QED.**

Proof of Proposition 5. Given the Cobb-Douglas production function (25), the following equations are necessary and sufficient for the complete-information first-best allocation

$$v'\left(h(s^t)\right) = u'\left(C(s^t)\right)\alpha\frac{y\left(s^t\right)}{n\left(s^t\right)h\left(s^t\right)} = 0 \quad \forall\ s^t \tag{31}$$

$$\left(1+v\left(h(s^t)\right)\right) = u'\left(C(s^t)\right)\alpha\frac{y\left(s^t\right)}{n\left(s^t\right)h\left(s^t\right)}h\left(s^t\right) = 0 \quad \forall\ s^t \tag{32}$$

$$y\left(s^t\right) = A\left(s^t\right)\left(n\left(s^t\right)h\left(s^t\right)\right)^\alpha \quad \forall\ s^t \tag{33}$$

$$C\left(s^t\right) = y\left(s^t\right) \tag{34}$$

Combining (31) and (32), we get that

$$1 + v\left(h(\omega_i^t, s^t)\right) = v'\left(h(s^t)\right)h\left(s^t\right)$$

Next, using the specification for v in (24), we may rewrite this as

$$\frac{1+\epsilon}{\epsilon} = h\left(s^t\right)^{1+\epsilon} \tag{35}$$

or

$$h\left(s^t\right) = \bar{h} = \left(\frac{1+\epsilon}{\epsilon}\right)^{\frac{1}{1+\epsilon}}$$

That is, the first-best level of effort is constant. We can now also solve for employment. From (31) we have that

$$v'\left(h(s^t)\right) = u'\left(C(s^t)\right)\alpha\frac{y\left(s^t\right)}{n\left(s^t\right)\bar{h}}$$

Using the specification for u and v in (24), we may rewrite this as

$$\bar{h}^{1+\epsilon} = y(s^t)^{1-\gamma}\alpha\frac{1}{n\left(s^t\right)}$$

Plugging in for $y\left(s^t\right)$ from the production function and solving for $n\left(s^t\right)$, we get that

$$n\left(s^t\right) = A\left(s^t\right)^{\frac{(1-\gamma)}{1-\alpha(1-\gamma)}}\left[\bar{h}^{\alpha(1-\gamma)-(1+\epsilon)}\alpha\right]^{\frac{1}{1-\alpha(1-\gamma)}}$$

Thus

$$\log n\left(s^t\right) = \frac{(1-\gamma)}{1-\alpha\left(1-\gamma\right)}\log A\left(s^t\right)$$

Finally, using the fact that $y\left(s^t\right) = A\left(s^t\right)\left(n\left(s^t\right)\bar{h}\right)^{\alpha}$, we get that

$$\begin{aligned}
\log y\left(s^t\right) &= \log A\left(s^t\right) + \alpha\log n\left(s^t\right) \\
&= \log A\left(s^t\right) + \frac{\alpha\left(1-\gamma\right)}{1-\alpha\left(1-\gamma\right)}\log A\left(s^t\right) \\
&= \frac{1}{1-\alpha\left(1-\gamma\right)}\log A\left(s^t\right)
\end{aligned}$$

This verifies the statement in the Proposition with

$$\begin{aligned}
\Lambda_a^* &= \frac{(1-\gamma)}{1-\alpha\left(1-\gamma\right)} \\
\Phi_a^* &= \frac{1}{1-\alpha\left(1-\gamma\right)}
\end{aligned}$$

Finally, note that $\Lambda_a^* > 0$ if and only if $\gamma < 1$. **QED.**

Proof of Proposition 6. Given the Cobb-Douglas production function (25), the following equations are necessary and sufficient for efficiency

$$v'\left(h(\omega_i^t, s^t)\right) - u'\left(C(s^t)\right)\left(\frac{y\left(\omega_i^t, s^t\right)}{Y(s^t)}\right)^{-\frac{1}{\rho}}\alpha\frac{y\left(\omega_i^t, s^t\right)}{n\left(\omega_i^t\right)h\left(\omega_i^t, s^t\right)} = 0 \quad \forall\, \omega_i^t, s^t \tag{36}$$

$$\mathbb{E}\left[\left(1 + v\left(h(\omega_i^t, s^t)\right)\right) - u'\left(C(s^t)\right)\left(\frac{y\left(\omega_i^t, s^t\right)}{Y(s^t)}\right)^{-\frac{1}{\rho}}\alpha\frac{y\left(\omega_i^t, s^t\right)}{n\left(\omega_i^t\right)h\left(\omega_i^t, s^t\right)}h\left(\omega_i^t, s^t\right)\,\middle|\,\omega_i^t\right] = 0 \quad \forall\, \omega_i^t \tag{37}$$

$$y\left(\omega_i^t, s^t\right) = A\left(s^t\right)\left(n\left(\omega_i^t\right)h\left(\omega_i^t, s^t\right)\right)^{\alpha} \quad \forall\, \omega_i^t, s^t \tag{38}$$

$$C(s^t) = Y\left(s^t\right) = \left[\int y\left(\omega_{it}, s^t\right)^{\frac{\rho-1}{\rho}}d\mathcal{G}_t(\omega_{it}|s^t)\right]^{\frac{\rho}{\rho-1}} \quad \forall\, s^t \tag{39}$$

Thus, for any realization of $\left(\omega_i^t, s^t\right)$, the following two equations must hold:

$$v'\left(h(\omega_i^t, s^t)\right) = u'\left(Y(s^t)\right)\left(\frac{y\left(\omega_i^t, s^t\right)}{Y(s^t)}\right)^{-\frac{1}{\rho}}\alpha\frac{y\left(\omega_i^t, s^t\right)}{n\left(\omega_i^t\right)h\left(\omega_i^t, s^t\right)} \tag{40}$$

$$y\left(\omega_i^t, s^t\right) = A\left(s^t\right)\left(n\left(\omega_i^t\right)h\left(\omega_i^t, s^t\right)\right)^{\alpha} \tag{41}$$

Using the assumed specification for u and v in (24), we can solve equations (40) and (41) simultaneously for $y\left(\omega_i^t, s^t\right)$ and $h(\omega_i^t, s^t)$. We thereby get that the efficient firm output level is given by

$$y\left(\omega_{it}, s^t\right) = \Psi^\omega(\omega_i^t)\Psi^s(s^t) \tag{42}$$

where

$$\begin{aligned}\log \Psi^\omega(\omega_i^t) &= \frac{\alpha\epsilon}{1+\epsilon-\alpha\left(\frac{\rho-1}{\rho}\right)}\log n\left(\omega_i^t\right) \\ \log \Psi^s(s^t) &= \frac{1+\epsilon}{1+\epsilon-\alpha\left(\frac{\rho-1}{\rho}\right)}\log\left(A\left(s^t\right)Y(s^t)^{\frac{\alpha}{1+\epsilon}\left(\frac{1}{\rho}-\gamma\right)}\right)\end{aligned}$$

where we abstract from the scalar constant $\alpha^{\frac{\alpha}{1+\epsilon}/\left(1-\frac{\alpha}{1+\epsilon}\left(\frac{\rho-1}{\rho}\right)\right)}$.

The constrained efficient level of effort is given by

$$h\left(\omega_i^t, s^t\right) = n\left(\omega_i^t\right)^{-\eta_n} A\left(s^t\right)^{\eta_A} Y(s^t)^{\eta_Y} \tag{43}$$

where

$$\eta_n = \frac{1-\alpha\left(\frac{\rho-1}{\rho}\right)}{(1+\epsilon)-\alpha\left(\frac{\rho-1}{\rho}\right)}, \qquad \eta_A = \frac{\left(\frac{\rho-1}{\rho}\right)}{(1+\epsilon)-\alpha\left(\frac{\rho-1}{\rho}\right)}, \text{ and } \qquad \eta_Y = \frac{\left(\frac{1}{\rho}-\gamma\right)}{(1+\epsilon)-\alpha\left(\frac{\rho-1}{\rho}\right)}$$

For any realization of s^t, we can now solve for the aggregates. Let $N\left(s^t\right)$ denote the simple average of employment

$$N\left(s^t\right) = \int n\left(\omega_{it}, s^t\right) d\mathcal{G}_t(\omega_{it}|s^t) \tag{44}$$

Aggregate output is given by

$$Y\left(s^t\right) = \left[N\left(s^t\right)^{\alpha\left(1-\frac{1}{1+\epsilon}\right)} A\left(s^t\right) Y(s^t)^{\frac{\alpha}{1+\epsilon}\left(\frac{1}{\rho}-\gamma\right)}\right]^{1/\left(1-\frac{\alpha}{1+\epsilon}\left(\frac{\rho-1}{\rho}\right)\right)}$$

where again we abstract from any constant scalar component. Solving this expression for $Y\left(s^t\right)$, we get that

$$Y\left(s^t\right) = A\left(s^t\right)^{\Upsilon_A} N\left(s^t\right)^{\Upsilon_N} \tag{45}$$

where

$$\Upsilon_A = \frac{1+\epsilon}{1+\epsilon-\alpha\left(1-\gamma\right)} \qquad \text{and} \qquad \Upsilon_N = \frac{\alpha\epsilon}{1+\epsilon-\alpha\left(1-\gamma\right)}$$

Therefore, given $n\left(\omega_i^t\right)$ and $N\left(s^t\right)$, the behavior of $y\left(\omega_i^t, s^t\right)$ and $h(\omega_i^t, s^t)$ are pinned down by (42), (43), and (45). What then remains to be characterized is the behavior of $n\left(\omega_i^t\right)$ and $N\left(s^t\right)$. We now show that there exists a fixed point in $n\left(\omega_{i,t}\right)$ and $N\left(s^t\right)$ which pins down their joint solution. The optimality condition for labor is given by (37).

$$\mathbb{E}\left[\left(1+v\left(h(\omega_i^t, s^t)\right)\right) - u'\left(C(s^t)\right)\left(\frac{y\left(\omega_i^t, s^t\right)}{Y(s^t)}\right)^{-\frac{1}{\rho}} \alpha \frac{y\left(\omega_i^t, s^t\right)}{n\left(\omega_i^t\right) h\left(\omega_i^t, s^t\right)} h\left(\omega_i^t, s^t\right) \;\middle|\; \omega_i^t\right] = 0$$

Combining this with (36), gives us that

$$\mathbb{E}\left[\ \left(1+v\left(h(\omega_i^t,s^t)\right)\right)-v'\left(h(\omega_i^t,s^t)\right)h\left(\omega_i^t,s^t\right)\ \middle|\ \omega_i^t\ \right]=0$$

Next, using the specification for v in (24), we may rewrite this as

$$\frac{1+\epsilon}{\epsilon}=\mathbb{E}\left[\ h(\omega_i^t,s^t)^{1+\epsilon}\ \middle|\ \omega_i^t\ \right] \tag{46}$$

Using the fact that $h\left(\omega_i^t,s^t\right)=n\left(\omega_i^t\right)^{-\eta_n}A\left(s^t\right)^{\eta_A}Y(s^t)^{\eta_Y}$, we may therefore express (46) in terms of $n\left(\omega_i^t\right)$ as follows

$$n\left(\omega_i^t\right)^{\eta_n(1+\epsilon)}=\frac{\epsilon}{1+\epsilon}\mathbb{E}\left[\ \left(A\left(s^t\right)^{\eta_A}Y(s^t)^{\eta_Y}\right)^{1+\epsilon}\ \middle|\ \omega_i^t\ \right]$$

Finally, using the fact that aggregate output satisfies $Y\left(s^t\right)=A\left(s^t\right)^{\Upsilon_A}N\left(s^t\right)^{\Upsilon_N}$, we then have that

$$n\left(\omega_i^t\right)^{\eta_n(1+\epsilon)}=\frac{\epsilon}{1+\epsilon}\mathbb{E}\left[\ \left(A\left(s^t\right)^{\eta_A+\Upsilon_A\eta_Y}N\left(s^t\right)^{\Upsilon_N\eta_Y}\right)^{(1+\epsilon)}\ \middle|\ \omega_i^t\ \right]$$

This fixed-point representation pins down the constrained efficient $n\left(\omega_i^t\right)$ for any information structure. However, given the Gaussian information structure we have imposed, we propose that the constrained efficient levels of $n\left(\omega_i^t\right)$ and $N\left(s^t\right)$ are jointly log-normal. This implies that the aforementioned functional equation can be restated in the following log-linear form:

$$\log n\left(\omega_i^t\right)=(1-\chi)\Lambda_a^*\mathbb{E}\left[a\left(s^t\right)\middle|\omega_{i,t}\right]+\chi\mathbb{E}\left[\log N\left(s^t\right)\middle|\omega_{i,t}\right] \tag{47a}$$

with

$$\chi\equiv\left(\frac{\alpha\epsilon}{1+\epsilon-\alpha(1-\gamma)}\right)\frac{\frac{1}{\rho}-\gamma}{1-\alpha\left(\frac{\rho-1}{\rho}\right)}<1\qquad\text{and}\qquad\Lambda_a^*=\frac{1-\gamma}{1-\alpha(1-\gamma)}$$

We now proceed to solve for the fixed point to the above condition. We guess and verify the log-linear fixed point. Suppose that the constrained efficient production strategy takes a log-linear form given by $\log n\left(\omega_i^t\right)=\lambda_x x_{it}+\lambda_z z_t$ for some coefficients λ_x,λ_z. It follows that $N\left(s^t\right)$ is indeed log-normal and given by $\log N\left(s^t\right)=\lambda_x a_t+\lambda_z z_t$. This implies

$$\mathbb{E}\left[\log N\left(s^t\right)\middle|\omega_{i,t}\right]=\lambda_x\mathbb{E}\left[a\left(s^t\right)\middle|\omega_{i,t}\right]+\lambda_z z_t$$

where $\mathbb{E}\left[a\left(s^t\right)\middle|\omega_{i,t}\right]=\frac{\kappa_x}{\kappa_x+\kappa_z+\kappa_A}x_{i,t}+\frac{\kappa_z}{\kappa_x+\kappa_z+\kappa_A}z_t$. Substituting this expression into (47a) gives us

$$\log n\left(\omega_i^t\right)=\left((1-\chi)\Lambda_a^*+\chi\lambda_x\right)\left[\frac{\kappa_x}{\kappa_x+\kappa_z+\kappa_A}x_{i,t}+\frac{\kappa_z}{\kappa_x+\kappa_z+\kappa_A}z_t\right]+\chi\lambda_z z_t$$

For this to coincide with $\log n\left(\omega_i^t\right)=\lambda_x x_{it}+\lambda_z z_t$ for every (x_{it},z_t) it is necessary and sufficient that the coefficients (λ_x,λ_z) satisfy

$$\begin{aligned}\lambda_x &= \left((1-\chi)\Lambda_a^*+\chi\lambda_x\right)\frac{\kappa_x}{\kappa_x+\kappa_z+\kappa_A}\\ \lambda_z &= \left((1-\chi)\Lambda_a^*+\chi\lambda_x\right)\frac{\kappa_z}{\kappa_x+\kappa_z+\kappa_A}+\chi\lambda_z\end{aligned}$$

The unique solution to this system is given below.

Lemma 1. *Given the Gaussian information structure, constrained efficient labor satisfies*

$$\log n\left(\omega_i^t\right) = \lambda_x x_{it} + \lambda_z z_t \tag{48}$$

with

$$\lambda_x = \frac{(1-\chi)\,\kappa_x}{(1-\chi)\,\kappa_x + \kappa_z + \kappa_A}\Lambda_a^* \tag{49}$$

$$\lambda_z = \frac{\kappa_z}{(1-\chi)\,\kappa_x + \kappa_z + \kappa_A}\Lambda_a^* \tag{50}$$

This implies that aggregate $N\left(s^t\right)$ is given by

$$\log N\left(s^t\right) = \lambda_x a_t + \lambda_z z_t \tag{51}$$

Next, from (45) we have that aggregate output is given by $Y\left(s^t\right) = A\left(s^t\right)^{\Upsilon_A} N\left(s^t\right)^{\Upsilon_N}$. Combining this with the aggregate labor condition above (51), we get that equilibrium aggregate output satisfies

$$\log Y\left(s^t\right) = \Upsilon_A a_t + \Upsilon_N\left(\lambda_x a_t + \lambda_z z_t\right) \tag{52}$$

This verifies part (ii) of Proposition (6), with

$$\begin{aligned}
\Lambda_a &= \lambda_x + \lambda_z, \quad \text{and} \quad \Lambda_\varepsilon = \lambda_z, \\
\Phi_a &= \Upsilon_A + \Upsilon_N \Lambda_a \quad \text{and} \quad \Phi_\varepsilon = \Upsilon_N \Lambda_\varepsilon
\end{aligned}$$

Finally, we have that individual firm output is given by

$$y\left(\omega_{it}, s^t\right) = \Psi^\omega(\omega_i^t)\Psi^s(s^t)$$

where

$$\begin{aligned}
\log \Psi^\omega(\omega_i^t) &= \frac{\alpha\epsilon}{1+\epsilon-\alpha\left(\frac{\rho-1}{\rho}\right)} \log n\left(\omega_i^t\right) \\
\log \Psi^s(s^t) &= \frac{1+\epsilon}{1+\epsilon-\alpha\left(\frac{\rho-1}{\rho}\right)} \log\left(A\left(s^t\right) Y(s^t)^{\frac{\alpha}{1+\epsilon}\left(\frac{1}{\rho}-\gamma\right)}\right)
\end{aligned}$$

then

$$y\left(\omega_{it}, s^t\right) = \frac{\alpha\epsilon}{1+\epsilon-\alpha\left(\frac{\rho-1}{\rho}\right)} \log n\left(\omega_i^t\right) + \frac{1+\epsilon}{1+\epsilon-\alpha\left(\frac{\rho-1}{\rho}\right)} \log A\left(s^t\right) + \frac{\alpha\left(\frac{1}{\rho}-\gamma\right)}{1+\epsilon-\alpha\left(\frac{\rho-1}{\rho}\right)} \log Y(s^t)$$

Substituting for $\log n\left(\omega_i^t\right)$ from (48) and for $\log Y\left(s^t\right)$ from (52), we rewrite this in terms of (x_{it}, z_t, a_t)

$$\begin{aligned}
y\left(\omega_{it}, s^t\right) &= \frac{\alpha\epsilon}{1+\epsilon-\alpha\left(\frac{\rho-1}{\rho}\right)}\left(\lambda_x x_{it} + \lambda_z z_t\right) + \frac{1+\epsilon}{1+\epsilon-\alpha\left(\frac{\rho-1}{\rho}\right)} a_t \\
&\quad + \frac{\alpha\left(\frac{1}{\rho}-\gamma\right)}{1+\epsilon-\alpha\left(\frac{\rho-1}{\rho}\right)}\left(\Upsilon_A a_t + \Upsilon_N\left(\lambda_x a_t + \lambda_z z_t\right)\right)
\end{aligned}$$

This verifies the statement in the Proposition, with

$$\begin{aligned}
\phi_x &= \frac{\alpha\epsilon}{1+\epsilon-\alpha\left(\frac{\rho-1}{\rho}\right)}\lambda_x \\
\phi_z &= \frac{\alpha\epsilon}{1+\epsilon-\alpha\left(\frac{\rho-1}{\rho}\right)}\lambda_z + \frac{\alpha\left(\frac{1}{\rho}-\gamma\right)}{1+\epsilon-\alpha\left(\frac{\rho-1}{\rho}\right)}\Upsilon_N\lambda_z \\
\phi_a &= \frac{1+\epsilon}{1+\epsilon-\alpha\left(\frac{\rho-1}{\rho}\right)} + \frac{\alpha\left(\frac{1}{\rho}-\gamma\right)}{1+\epsilon-\alpha\left(\frac{\rho-1}{\rho}\right)}\left(\Upsilon_A+\Upsilon_N\lambda_x\right)
\end{aligned}$$

Finally, note that

$$\begin{aligned}
\Lambda_a &= \frac{(1-\chi)\,\kappa_x+\kappa_z}{(1-\chi)\,\kappa_x+\kappa_z+\kappa_A}\Lambda_a^* \\
\Lambda_\varepsilon &= \frac{\kappa_z}{(1-\chi)\,\kappa_x+\kappa_z+\kappa_A}\Lambda_a^*
\end{aligned}$$

Therefore, $\Lambda_a > 0$ and $\Lambda_\varepsilon > 0$ if and only if $\Lambda_a^* > 0$. Furthermore, $|\Lambda_a| < |\Lambda_a^*|$. **QED.**

Proof of Proposition 7. From (42), we may write $y\left(\omega_{it}, s^t\right) = \Psi^\omega(\omega_i^t)\Psi^s(s^t)$ with

$$\log\Psi^\omega(\omega_i^t) = \frac{\alpha\epsilon}{1+\epsilon-\alpha\left(\frac{\rho-1}{\rho}\right)}\log n\left(\omega_i^t\right).$$

Following the proof of Proposition 3, nominal prices are given by

$$\log p\left(\omega_{it}\right) = -\frac{1}{\rho}\log\Psi^\omega(\omega_i^t)$$

Combining the above two equations, we conclude that

$$\log p\left(\omega_{it}\right) = -\psi\log n(\omega_{it}) \tag{53}$$

where

$$\psi \equiv \frac{\alpha\epsilon}{\rho\,(1+\epsilon)-\alpha\,(\rho-1)} > 0$$

Finally, aggregating (53) across firms gives the aggregate price level as a decreasing function of aggregate employment: $\log P_t = -\psi\log N_t$. Combining this with the closed-form results in Proposition 6 immediately verifies the statements in Proposition 7. **QED.**

Proof of Condition (26). With this variant, for any realization of $\left(\omega_i^t, s^t\right)$, the following two equations must hold:

$$v'\left(h(\omega_i^t, s^t)\right) = u'\left(Y(s^t)\right)\left(\frac{y\left(\omega_i^t, s^t\right)}{Y(s^t)}\right)^{-\frac{1}{\rho}}\alpha\frac{y\left(\omega_i^t, s^t\right)}{n\left(\omega_i^t\right)h\left(\omega_i^t, s^t\right)} \tag{54}$$

$$y\left(\omega_i^t, s^t\right) = A\left(\omega_i^t\right)\left(n\left(\omega_i^t\right)h\left(\omega_i^t, s^t\right)\right)^\alpha \tag{55}$$

Note that the difference between these equations and (40) and (41) is that A is now measurable in ω_i^t. Using the assumed specification for u and v in (24), we can solve these simultaneously for $y\left(\omega_i^t, s^t\right)$ and $h(\omega_i^t, s^t)$. We find that the efficient firm output level is given by

$$y\left(\omega_{it}, s^t\right) = \Psi^\omega(\omega_i^t)\Psi^s(s^t) \tag{56}$$

with

$$\Psi^\omega(\omega_i^t) = \left[A(\omega_i^t)n\left(\omega_i^t\right)^{\alpha\left(\frac{\epsilon}{1+\epsilon}\right)}\right]^{\frac{1+\epsilon}{1+\epsilon-\alpha\left(\frac{\rho-1}{\rho}\right)}} \tag{57}$$

$$\Psi^s(s^t) = Y(s^t)^{\frac{\alpha\left(\frac{1}{\rho}-\gamma\right)}{1+\epsilon-\alpha\left(\frac{\rho-1}{\rho}\right)}} \tag{58}$$

and where we abstract from the scalar constant $\alpha^{\frac{\alpha}{1+\epsilon}/\left(1-\frac{\alpha}{1+\epsilon}\left(\frac{\rho-1}{\rho}\right)\right)}$.

Using (56), we may then write aggregate output as follows

$$Y\left(s^t\right) = \left[\int y\left(\omega_{it}, s^t\right)^{\frac{\rho-1}{\rho}} d\mathcal{G}_t(\omega_{it}|s^t)\right]^{\frac{\rho}{\rho-1}} = \Psi^s(s^t)\mathcal{B}(s^t) \tag{59}$$

where we recall that $\mathcal{B}(s^t)$ is defined by

$$\mathcal{B}(s^t) \equiv \left[\int \Psi^\omega\left(\omega_i^t\right)^{\frac{\rho-1}{\rho}} d\mathcal{G}_t(\omega_i^t|s^t)\right]^{\frac{\rho}{\rho-1}}$$

Substituting our expression for $\Psi^s(s^t)$ from (58) into (59), we get that

$$Y\left(s^t\right) = Y(s^t)^{\frac{\alpha\left(\frac{1}{\rho}-\gamma\right)}{1+\epsilon-\alpha\left(\frac{\rho-1}{\rho}\right)}}\mathcal{B}(s^t)$$

Solving this for $Y\left(s^t\right)$, we find that we may write aggregate output as

$$Y\left(s^t\right) = \mathcal{B}(s^t)^{\chi_Y} \tag{60}$$

with

$$\chi_Y \equiv \frac{1+\epsilon-\alpha\left(\frac{\rho-1}{\rho}\right)}{1+\epsilon-\alpha\left(1-\gamma\right)} > 0$$

We now consider prices. Following the proof of Theorem 2, for any arbitrary common-knowledge process J_t', nominal prices are given by $p(\omega_i^t) = e^{J_t'}\Psi^\omega\left(\omega_i^t\right)^{-\frac{1}{\rho}}$. It follows that the aggregate price level is given by

$$P(s^t) = \mathcal{B}(s^t)^{-\frac{1}{\rho}} \tag{61}$$

where we abstract from the common-knowledge process J_t'.

Aggregate output and the price level are then determined by equations (60), and (61), written here in logs as

$$\log Y(s^t) = \chi_Y \log Q_t$$
$$\log P(s^t) = -\frac{1}{\rho}\log Q_t$$

This verifies that we may write the aggregate price level as log-linear function of aggregate output:

$$\log P(s^t) = -\Gamma \log Y(s^t)$$

with

$$\Gamma \equiv \frac{1}{\rho}\frac{1}{\chi_Y} = \frac{1}{\rho}\left(\frac{1+\epsilon-\alpha\left(1-\gamma\right)}{1+\epsilon-\alpha\left(\frac{\rho-1}{\rho}\right)}\right) > 0$$

QED.

Appendix B: Extension with Idiosyncratic Shocks

In this appendix we extend the analysis in the presence of idiosyncratic productivity shocks.

Setup. We modify the model by letting firm-level productivity vary with both aggregate and idiosyncratic shocks. In particular, we let the output of firm i be

$$y_{it} = A_{it} F\left(k_{it}, \ell_{it}\right),$$

where $\ell_{it} = n_{it} h_{it}$ and were firm-level productivity A_{it} is given by

$$A_{it} = A_t \eta_{it},$$

with A_t being the aggregate shock (common across firms) and η_{it} being the idiosyncratic shock (i.i.d. across firms).

We next modify the information structure so that firms can have arbitrary information for either the aggregate state (which contains the aggregate productivity shock) or their idiosyncratic productivity shock. Thus, the aggregate state continues to be given by the history $s^t = (s_0, ..., s_t)$ of an exogenous random variable s_t as in the baseline model. The idiosyncratic productivity shocks, on the other hand, are drawn identically and independently across the firms, but can be correlated over time and their distribution may vary with the aggregate state: for each i, η_{it} is drawn from a set $\mathcal{H}_t \subset \mathbb{R}_+$ according to a probability distribution $\mathcal{F}_t^{\eta}(\eta_{it}|\eta_i^{t-1}, s^t)$, where $\eta_i^{t-1} \equiv (\eta_{i0}, ..., \eta_{it-1})$. With this notation at hand, we can express aggregate productivity as $A_t = \bar{A}(s^t)$ and firm-level productivity as $A_{it} = A(s^t, \eta_i^t) \equiv \bar{A}(s^t)\eta_{it}$. Finally, the information set of firm i is represented by the history $\omega_i^t = (\omega_{i0}, ..., \omega_{it})$ of an exogenous signal ω_{it} about the underlying aggregate state and the firm's idiosyncratic productivity shocks: for each i, ω_{it} is drawn from a set Ω_t according to a probability distribution $\mathcal{G}_t(\omega_{it}|s^t, \eta_i^t, \omega_i^{t-1})$. The latter means that ω_{it} may contain information, not only about s^t, but also about η_i^t.

This formalization is very flexible, allowing for arbitrary persistence and heteroscedasticity in the idiosyncratic shocks as well as for firms to have incomplete information for either the aggregate state of the economy or their idiosyncratic productivity shocks. It also allows for arbitrary dynamics in how this information may be updated over time. Two special cases that we will consider in the sequel is when ω_i^t contains either the idiosyncratic shock η_{it} or the total firm-level productivity A_{it}.

Results. As in the main text, employment n_{it} and investment x_{it} must be measurable in ω_i^t. But now that there are idiosyncratic shocks, market clearing can obtain only if we allow firm output to adjust, not only to the aggregate state s^t, but also to the idiosyncratic productivity shocks. The analysis in the main text then remains nearly intact. In particular, all we have to do is to replace $A(s^t)$, $h(\omega_i^t, s^t)$, and $y(\omega_i^t, s^t)$ with, respectively, $A(z_i^t)$, $h(\omega_i^t, z_i^t)$ and $y(\omega_i^t, z_i^t)$, where $z_t^i \equiv (s^t, \eta_i^t)$.

Notwithstanding this notational adjustment, Propositions 1, 2 and 3, Theorems 1 and 2, and Corollaries 1, 2, and 3 continue to hold in exactly the same form as in the main text. The only difference emerges in Proposition 4, regarding the conditions that suffice for the constrained efficient allocation to be log-separable (and hence implementable under sticky prices).

In particular, the economic meaning of the log-separability restriction remains the same: for the relative output of any two firms i and j to be spanned by their information sets ω_i^t and ω_j^t, it must be that the dependence of i's output to the true state cancels with that of j's output. But now that there is productivity heterogeneity, this can happen only if that heterogeneity is also spanned by the two firms' information sets. The following variant of Proposition 4 thus applies.

Proposition 4b 1. *A flexible-price allocation is log-separable if the following two properties hold:*

(i) The technology is Cobb-Douglas and the disutility of effort is isoelastic (as in Proposition 4).

(ii) Firm-level productivity is itself log-separable: there exist functions A^ω and A^s such that

$$\log A(z_{it}) = \log A^\omega(\omega_i^t) + \log A^s(s^t)$$

The only difference from Proposition 4 is therefore property (ii), which is a joint restriction on the productivity process and the information structure. This property is trivially satisfied in each of the following four cases. (i) Suppose there are only unknown aggregate productivity shocks (as in the main text); then, $A^\omega(\omega_i^t) = 1$ and $A^s(s^t) = A_t$. (ii) Suppose firms know perfectly their own idiosyncratic shock but not the aggregate shock; then $A^\omega(\omega_i^t) = \eta_{it}$ and $A^s(s^t) = A_t$. (iii) Suppose that there are both aggregate and idiosyncratic shocks to productivity, and firms know perfectly their overall productivity A_{it}; then, $A^\omega(\omega_i^t) = A_{it} = A_t\eta_{it}$ and $A^s(s^t) = 1$. (iv) Suppose that aggregate productivity has two components, $A_t = A_{1t}A_{2t}$, and that firms observe $A_{2t}\eta_{it}$, i.e., the product of the idiosyncratic shock and one component of the aggregate shock; then,$A^\omega(\omega_i^t) = A_{2t}\eta_{it}$ and $A^s(s^t) = A_{1t}$.

Finally, to see when log-separability may fail, consider the case where there are only idiosyncratic productivity shocks and firms do not know these shocks when setting their prices. Because effort can still adjust to the idiosyncratic shock, more productive firms will produce more along the constrained efficient allocation. For this to be implemented in equilibrium, more productive firms must face lower relative prices. This is no issue if prices are flexible. But once prices are sticky, this can have obtained only if firms knew their idiosyncratic productivity when setting prices. It follows that the constrained efficient allocation will not be implementable under sticky prices when firms do not know their productivity.

If we define a "generic" information structure as one where firms face uncertainty about all exogenous random variables, then the constrained efficient allocation will fail to be implementable under sticky prices whenever the information structure is generic in the aforementioned sense.

Nevertheless, a variant of Theorem 2 continues to hold in no so far we measure the aggregate belief proxy $\mathcal{B}(s^t)$ along that optimal implementable allocation, whatever that might be. By direct implication, Corollary 3 also continues to hold.

To recap, the introduction of idiosyncratic productivity shocks makes it less likely that the constrained efficient allocation can be implemented under sticky prices, but does not affect the essence of our result regarding the cyclical nature of the optimal monetary policy. The latter must "lean against the wind" irrespectively of whether this attains, or only proxies, the constrained efficient allocation.

Proof of Proposition 4b. Suppose properties (i) and (ii) of Proposition 4b hold for some flexible-price allocation. From Proposition 2 we know that, in any flexible price equilibrium allocation, effort $h(\omega_i^t, s^t)$ is pinned down by the following condition:

$$V_h(\omega_i^t, s^t) = U_c(s^t)\phi(s^t)MP_\ell\left(\omega_i^t, s^t\right)$$

Property (i) implies that the above condition can be reduced to the following:

$$V_h(\omega_i^t, s^t) = U_c(Y\left(s^t\right)\phi(s^t)\left(\frac{y\left(\omega_i^t, s^t\right)}{Y(s^t)}\right)^{-\frac{1}{\rho}}\alpha\frac{y\left(\omega_i^t, s^t\right)}{n\left(\omega_i^t\right)h\left(\omega_i^t, s^t\right)}$$

Furthermore, with the assumed specification for $V(h)$ in property (i), for any realization of $\left(\omega_i^t, s^t\right)$ the following two equations must hold:

$$v_1 h\left(\omega_i^t, s^t\right)^\epsilon = U_c(Y\left(s^t\right)\phi(s^t)\left(\frac{y\left(\omega_i^t, s^t\right)}{Y(s^t)}\right)^{-\frac{1}{\rho}}\alpha\frac{y\left(\omega_i^t, s^t\right)}{n\left(\omega_i^t\right)h\left(\omega_i^t, s^t\right)} \tag{62}$$

$$y\left(\omega_i^t, s^t\right) = A_{it}k\left(\omega_i^t\right)^{1-\alpha}\left(n\left(\omega_i^t\right)h\left(\omega_i^t, s^t\right)\right)^\alpha \tag{63}$$

We can solve (62) and (63) simultaneously for $y\left(\omega_i^t, s^t\right)$ and $h(\omega_i^t, s^t)$. We thereby get that equilibrium output is given by

$$y\left(\omega_i^t, s^t\right) = \left[A_{it}k\left(\omega_i^t\right)^{1-\alpha}n\left(\omega_i^t\right)^{\alpha\left(1-\frac{1}{1+\epsilon}\right)}\left[U_c(Y\left(s^t\right))\phi(s^t)Y(s^t)^{\frac{1}{\rho}}\frac{\alpha}{v_1}\right]^{\frac{\alpha}{1+\epsilon}}\right]^{1/\left(1-\frac{\alpha}{1+\epsilon}\left(\frac{\rho-1}{\rho}\right)\right)}$$

Next, using property (ii), namely, that productivity is log-separable: $A_{it} = A^\omega(\omega_i^t)A^s(s^t)$, it is immediate that output $y\left(\omega_{it}, s^t\right)$ is log-separable in ω_i^t and s^t:

$$y\left(\omega_{it}, s^t\right) = \Psi^\omega(\omega_i^t)\Psi^s(s^t)$$

with

$$\Psi^\omega(\omega_i^t) = \left[A^\omega(\omega_i^t)k\left(\omega_i^t\right)^{1-\alpha}n\left(\omega_i^t\right)^{\alpha\left(\frac{\epsilon}{1+\epsilon}\right)}\right]^{\frac{1+\epsilon}{1+\epsilon-\alpha\left(\frac{\rho-1}{\rho}\right)}} \tag{64}$$

$$\Psi^s(s^t) = \left[A^s(s^t)\left(U_c(Y\left(s^t\right))\phi(s^t)Y(s^t)^{\frac{1}{\rho}}\right)^{\frac{\alpha}{1+\epsilon}}\right]^{\frac{1+\epsilon}{1+\epsilon-\alpha\left(\frac{\rho-1}{\rho}\right)}} \tag{65}$$

where we abstract from the constant scalar $(\alpha/v_1)^{\frac{\alpha}{1+\epsilon-\alpha\left(\frac{\rho-1}{\rho}\right)}}$. This confirms that a flexible-price allocation in which properties (i) and (ii) hold is log-separable. **QED.**