Inflation and Unemployment in the Long Run Revisited*

Zachary Bethune
Rice University

Michael Choi
University of California, Irvine

Sébastien Lotz
Université Panthéon-Assas

Guillaume Rocheteau
University of California, Irvine

June 21, 2024

Abstract

We construct a continuous-time, monetary model with frictional goods and labor markets to revisit the long-run relationship between inflation and unemployment. By endogenizing the value of consumers' outside options and market power in accordance with standard consumer search theory, we generate novel predictions for the slope of the long-run Phillips curve, optimal monetary policy, and outcomes at the frictionless limit. The relationship between inflation and unemployment is non-monotone and, at low inflation rates, an increase in inflation reduces unemployment. The Friedman rule is suboptimal when firms' average bargaining power across markets is low. Markups and markdowns vanish as frictions disappear.

JEL Classification: D82, D83, E40, E50
Keywords: Unemployment, inflation, money, search, bargaining

*We thank seminar participants at the University of Hawaii, Monash University, UC Irvine, University of Wisconsin-Madison, Loyola Marymount University, Université Paris-Panthéon-Assas, the 2023 Dale Mortensen centre conference, the 2023 Search and Matching in Macro and Finance virtual seminar, the 2024 Theories and Methods in Macro conference in Amsterdam, and the 2024 Bristol Macroeconomics Workshop for their comments .

1 Introduction

According to the classical dichotomy, the long-run rate of inflation, perfectly correlated with money growth by virtue of the quantity theory, does not affect the equilibrium unemployment rate, resulting in a vertical long-run Phillips curve. While this view has been embraced by the canonical models of unemployment of Lucas and Prescott (1974) and Mortensen and Pissarides (1994), it has been challenged by Berentsen et al. (2011) – hereafter referred to as BMW. They show that, in theory, anticipated inflation acts as a distortionary tax that penalizes monetary exchange and market activities. In their model, inflation reduces consumers' payment capacity, which has an adverse effect on firms' profits and their incentives to open vacancies in order to hire workers. As a result, the long-run Phillips curve slopes upward, in accordance with Friedman (1977). This logic has been formalized by integrating two workhorse models: the Mortensen and Pissarides (1994) model of unemployment – MP thereafter – and the Lagos and Wright (2005) model of monetary exchange.

While BMW offers an elegant description of goods and labor markets as two frictional markets that open sequentially, they introduce a subtle and overlooked asymmetry regarding the treatment of agents' outside options in each market. In the labor market, a worker who is negotiating her long-term employment contract with a firm can opt out and search for an alternative employer. Hence, by agreeing to an offer, the worker is giving up the opportunity to trade with a different firm, possibly at a different wage. In contrast, in the goods market, matched consumers face no opportunity cost from accepting a trade since their decision does not affect their future trading opportunities. Indeed, by assumption, in every period, consumers want to consume but are matched with at most one producer. Hence, they do not have the option to search for an alternative producer than the one they are matched with to satisfy their *current* desire for consumption. We will show that the lack of meaningful consumer outside options inhibits competitive pressures in the goods market, thereby undermining a core rationale for using search and bargaining models to formalize decentralized markets, namely, that they converge to a rent-free, perfect-competition outcome when trading frictions vanish. Moreover, it shuts down an intuitive market-power channel that operates through the user cost of money, and affects the relationship between inflation and unemployment. Finally, we show it is critical for the design of monetary policy and the optimality of the Friedman rule.

Our paper has two contributions: one is methodological and the other is theoretical. In terms of methodology, we construct a continuous-time monetary model of goods and labor markets with search frictions in both markets that provides a symmetric treatment of workers' and consumers' outside options and the determination of prices and wages. The model

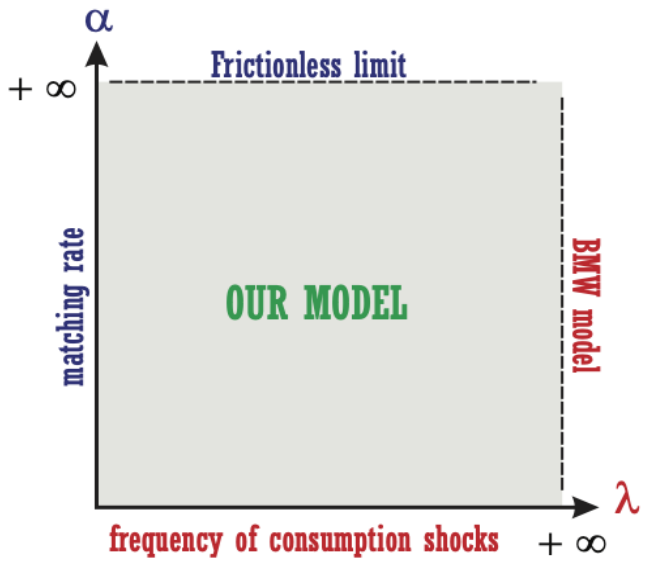


Figure 1: Disentangling search frictions (α) and preference shocks (λ).

has MP and BMW as special cases. Specifically, we endogenize outside options by disentangling preference shocks from random matching shocks in the goods market, as illustrated in Figure 1. Matching in the goods market is represented by $\alpha \in \mathbb{R}_+$, the rate at which consumers find sellers (the vertical axis in Figure 1). The market will be said to be frictionless when $\alpha \to +\infty$. Preference shocks are represented by $\lambda \in \mathbb{R}_+$, the rate at which a buyer who has just consumed receives a new preference shock and wants to consume that same good again (the horizontal axis in Figure 1). BMW is the limit $\lambda = +\infty$, i.e., buyers want to consume all the time. Buyers incur an opportunity cost of trade only when $\lambda < +\infty$. We will show that the economy under $\lambda < +\infty$ differs fundamentally from the one with $\lambda = +\infty$ in terms of its positive implications (e.g., slope of the Phillips curve) as well as its normative implications (e.g., optimality of the Friedman rule).

Our theoretical contribution is threefold. The first contribution consists in proving the generic nonmonotonicity of the long-run Phillips curve. At low inflation rates, there is a negative relationship between inflation and unemployment whereas the relationship is positive at high inflation rates. This nonmonotonicity arises from two opposite effects of inflation on equilibrium unemployment. First, there is a negative effect on consumers' real balances, identified in BMW, that raises unemployment. Second, there is a new *market-power effect* according to which inflation reduces the value for consumers of opting out of a trade in order to search for an alternative producer. Indeed, search in monetary economies requires agents to hold real money balances at a cost that increases with inflation. This second effect raises firms' market power and incentivizes them to open more vacancies, thereby reducing

unemployment. The *market-power effect* of inflation dominates at low inflation rates whereas the *real-balance effect* dominates at high inflation rates.

Figure 2 illustrates this finding by plotting the theoretical long-run Phillips curve. There is an inflation rate above the Friedman rule, denoted π^*, that minimizes unemployment. Below that inflation rate, the Phillips curve is downward-sloping, which generates a long-run trade-off for policymakers whose dual mandate is to keep inflation and unemployment low. Above π^*, the long-run Phillips curve is upward-sloping, in which case there is no trade-off between inflation and unemployment. These results rationalize the choice of a positive inflation target between the Friedman rule and π^*.

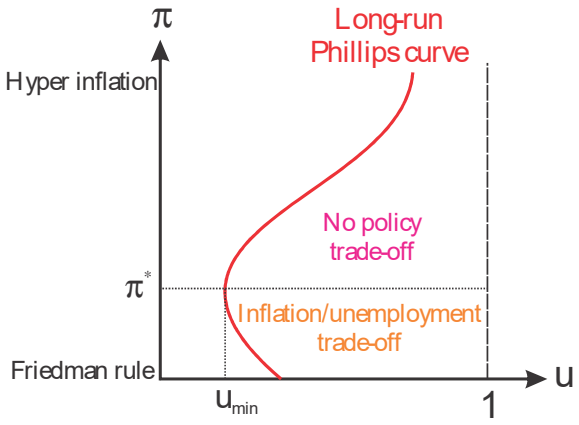


Figure 2: The long-run relation between inflation and unemployment

In addition to showing that the Friedman rule fails to minimize the unemployment rate, the second theoretical contribution establishes that the Friedman rule fails to maximize social welfare when workers and consumers have high enough bargaining power across labor and goods markets. The logic goes as follows. The constrained-efficient allocations are implemented with the Friedman rule and a Hosios condition in each market. When $\lambda < +\infty$, a deviation from the Friedman rule has a first-order, positive effect on firm entry and labor market tightness. There is also a negative effect on consumers' real money balances but the impact on entry is only second order. Hence, it is optimal to raise the money growth rate above the Friedman rule when entry is inefficiently low, e.g., due to high worker's and/or consumer's bargaining powers. Maybe surprisingly, this logic breaks down in the BMW model, when $\lambda = +\infty$. Indeed, a deviation from the Friedman rule only has a second-order effect on firm entry and social welfare and, under bargaining, the Friedman rule is always optimal.

Our third theoretical contribution is related to the limit of equilibrium outcomes as trading frictions vanish. A major result in the literature on decentralized markets is the convergence of equilibrium allocations to a Walrasian, or perfect-competition outcome as

trading frictions vanish (Gale, 1986a,b, 1987). We show that Gale's result on frictionless limits holds in our two-market economies for all $\lambda \in (0, +\infty)$, i.e., markups and markdowns are driven to zero as the process of matching buyers and sellers becomes infinitely efficient. In the goods market, as α increases, the value of searching increases and rents vanish at the limit. In contrast, when $\lambda = +\infty$, i.e. the BMW model, positive rents expand when α increases because consumers hold more cash while there is no competitive pressure to reduce rents since consumers have no outside option. Table 1 provides an overview of the key differences between the predictions of our model and the ones of BMW.

	BMW, $\lambda = +\infty$	Our model, $\lambda \in (0, +\infty)$
Slope of Phillips curve	positive	negative at low inflation rates
Friedman rule	optimal for welfare	suboptimal if low firms' bargaining power
Limit at $\alpha \to +\infty$	rents & markups>0	rents & markup=0

Table 1: Comparison to BMW

In the last section, we quantify the size of the market power and real balance effects of inflation by calibrating a version of our model extended to incorporate realistic heterogeneity in λ across consumer expenditure categories. Using evidence from the Consumer Expenditure Survey, we find that λ averages 0.94 at a monthly frequency with substantial heterogeneity across expenditures categories. We show that the unemployment-minimizing inflation rate, π^*, is slightly less than 0, but can range as high as 9% when the market-power effect is counterfactually strong ($\lambda \to 0$ across goods). For all calibration, the Friedman rule, $\pi = -1.2\%$, maximizes social welfare. An increase in inflation from the Friedman rule to π^* increases markups across all categories, but can either increase or decrease wage markdowns.

In an online appendix, we study two generalizations of our environment to illustrate the robustness of our main insights. In Online Appendix C, we generalize the relation between monetary policy and unemployment by incorporating a short-term nominal interest rate on liquid government bonds and limited participation in the bond market. Monetary policy now has two instruments: the constant money growth rate, and the short-term nominal interest rate. When the inflation rate is sufficiently high, the relation between unemployment and the short-term interest rate is non-monotone. Secondly, in our baseline model, consumer search is a threat that disciplines the demand of firms during their negotiation with consumers, however, on the equilibrium path, consumers do not exercise their option to search once matched with a firm. In Online Appendix D, we consider horizontally differentiated products that induces search along the equilibrium path, and show that our main insight is robust, i.e. the long-run Phillips curve is downward sloping at low inflation rates.

1.1 Literature review

Our model builds on Berentsen et al. (2011) that introduces a frictional labor market into the Lagos and Wright (2005) model.[1] Other models with frictional goods and labor markets include Lehmann and Van der Linden (2010), Petrosky-Nadeau and Wasmer (2015, 2017), and Michaillat and Saez (2015). In the first two, there is no money, i.e., agents pay with transferable utility. In the latter, money is introduced in the utility function directly. In contrast, we explicitly formalize the role of money and the associated liquidity constraints that are critical for the formation of the terms of trade. Relative to BMW, we introduce competitive pressures by assuming that consumers have infrequent wishes to consume so that their needs can be fulfilled by different producers across time. This follows the tradition of the consumer search literature where the opportunity cost of trading includes the forgone benefit of searching for other sellers, e.g., McCall (1970), Wolinsky (1986), and Anderson and Renault (1999).[2]

There are alternative approaches to model competition in monetary economies. Head and Kumar (2005); Wang (2016); Wang et al. (2020) introduce noisy search, as in Burdett and Judd (1983), into the Lagos-Wright model.[3] Rocheteau and Wright (2005) and Bethune et al. (2020) introduce directed and partially directed search into monetary economies. A crucial difference is that, in our model, switching across sellers to exercise consumers' market power takes time, i.e., the search process is sequential. Therefore, monetary policies affect the *intertemporal* outside option of searching for alternative sellers by raising the cost of holding money. In the noisy or directed search model above, search is simultaneous and thus the cost of holding money does not affect the outside option of switching to alternative sellers. This difference leads to different predictions regarding the normative and positive impact of monetary policies.

There are alternative explanations for why the long-run Phillips curve could be downward sloping at low inflation rates. In Rocheteau et al. (2007) unemployment emerges due to indivisible labor, the long-run Phillips curve can be downward sloping if leisure and consumption are complements. In Rocheteau and Rodriguez-Lopez (2014), the downward-sloping Phillips curve arises from a Tobin effect that induces agents to substitute real money balances for

[1] Other models of unemployment and inflation based on the Mortensen-Pissarides framework include Shi (1998), Cooley and Quadrini (2004), and Lehmann (2012). A related approach is provided by Williamson (2015). Versions of the model with money and credit include Bethune et al. (2015), Branch et al. (2016), and Gu et al. (2023). A continuous-time version was constructed by Rocheteau and Rodriguez-Lopez (2014).

[2] These models assume that products are horizontally differentiated in order to generate consumer search in equilibrium. We adopt a similar assumption in an extension of our model in Online Appendix D.

[3] Julien et al. (2008), Galenianos and Kircher (2009), and Bajaj and Mangin (2023) also consider imperfect competition by introducing multilateral meetings into search models.

capital and financial assets as inflation increases, thereby promoting job creation and lowering unemployment. In Gu et al. (2023), inflation hurts the unemployed who do not have access to credit, which allows firms to negotiate lower wages. Relative to this literature, we claim that the non-monotone relation between inflation and unemployment is generic provided there is genuine consumer search in the goods market.

There is a literature on search and inflation under menu costs that shows that inflation erodes firms' market power, e.g., Benabou (1988) and Diamond (1993). In those models, the economy is cashless, i.e., money has no transaction role. Inflation reduces market power by preventing firms from maintaining their real price at a monopoly level as in Diamond (1971).[4] In our model, prices are perfectly flexible and, contrary to Benabou (1988) and Diamond (1993), inflation makes search more costly for consumers by raising the cost of carrying cash, which reduces the value of their outside option and raises firms' market power.

The extension in Online Appendix C with liquid bonds and limited participation in the bond market is related to Alvarez et al. (2001), Alvarez et al. (2002), and Williamson (2006). The existence of liquidity-trap equilibria is similar to Williamson (2012) and Rocheteau et al. (2018). Relative to these papers, we add a frictional labor market and endogenous outside options for consumers. Moreover, we provide conditions under which unemployment is minimum at the liquidity trap.

1.2 Empirical evidence

The literature on the empirical relationship among inflation, unemployment and market power is extensive but remains inconclusive, hence there is room to use a micro-founded model to help guide policy. First, there is mixed empirical evidence about the relationship between anticipated inflation and unemployment. King and Watson (1994) use a vector autoregression (VAR) methodology with identified monetary policy shocks, and find evidence of a negative long-run trade-off between inflation and unemployment in the US post war. However, King and Watson (1997) show these results depend on the choice of identifying assumptions. More recently, Benati (2015) adopts both classical and Bayesian structural VARs, and shows that US data is compatible with both positively- and negatively-sloped Phillips curves. Ascari et al. (2022) introduce stochastic trends into a Bayesian VAR. They find that there is a threshold level of inflation, slightly below 4%, below which potential

[4]Relatedly, New Keynesian models can feature a positive correlation between macroeconomic activity and inflation. For instance, in King and Wolman (1996), firms adjust prices infrequently as the inflation rate rises, resulting in a lower markup and a higher output. Devereux and Yetman (2002) show that, when the frequency of price adjustments is endogenized, the correlation between inflation and output becomes non-linear and non-monotone.

output is independent of inflation, and above which potential output and inflation are negatively correlated. This finding is consistent with our theory which predicts that output rises with inflation when inflation is low, and otherwise falls.

In terms of cross-country evidence, Bullard and Keating (1995) study a large sample of postwar economies using a structural VAR approach and find that inflation raised output in low-inflation countries, and either did not affect or reduced output in countries with a higher inflation rate. A meta-study of VAR studies by De Grauwe and Costa Storti (2004) finds that the average effect of a 1% interest rate shock on output is -0.15% after five years, but the standard deviation of the distribution of estimates is 0.27%, so the Phillips curve can have a positive or a negative slope.

The empirical evidence on inflation and firms' market power is also mixed. For instance, Chirinko and Fazzari (2000) and Neiss (2001) find a positive relationship between inflation and firm-level markups, while Banerjee and Russell (2001, 2005) find a negative relationship. The topic has gained renewed interest following the 2021-22 inflation surge and discussion of "greed-flation" as a source of rising prices (see, e.g., DePillis, 2022). Glover et al. (2023) document that firm-level markups in the US increased by 3.4 percentage points in 2021 while inflation increased by around 3 percentage points, and Hansen et al. (2023) find that increases in corporate profits contributed up to 45% of the Euro area inflation.

2 Environment

The benchmark environment, which builds on Choi and Rocheteau (2021), can be interpreted as a version of Lagos and Wright (2005), and Rocheteau and Wright (2005) where centralized and decentralized markets co-exist in continuous time.

Time, agents, and commodities Time is continuous and indexed by $t \in \mathbb{R}_+$. The economy is composed of three types of infinitely-lived agents: a unit measure of workers, a measure ω of consumers, and an endogenous measure of firms, n. There are two perishable goods, $y \in \mathbb{R}_+$ and $c \in \mathbb{R}$. Good c is taken as the numéraire, can be consumed and produced by all agents, and is traded competitively and continuously over time. Good y is produced by worker-firm pairs and valued by consumers only.

Preferences Preferences of consumers, workers, and firms are represented by the following lifetime expected utility functions:

$$\mathcal{U}^b = \mathbb{E}\left[\int_0^\infty e^{-\rho t} dC(t) + \sum_{n=1}^{+\infty} e^{-\rho T_n} \varepsilon_{T_n} \upsilon\left[y(T_n)\right]\right], \tag{1}$$

$$\mathcal{U}^w = \mathbb{E}\left[\int_0^\infty e^{-\rho t} dC(t)\right], \text{ and} \tag{2}$$

$$\mathcal{U}^f = \mathbb{E}\left[\int_0^\infty e^{-\rho t} dC(t)\right], \tag{3}$$

where $C(t)$ measures the cumulative net consumption of the numéraire good.[5] Negative consumption of the numéraire good is interpreted as production. So preferences of all agents are linear in the numéraire.[6] The second term on the right side of (1) represents the consumption of the y good produced by firms for consumers. In (1) the idiosyncratic stochastic process $\{T_n\}$ indicates the times at which the consumer gets to consume good y. The utility function, $\upsilon(y)$, is such that $\upsilon(0) = 0$, $\upsilon' > 0$, and $\upsilon'' < 0$.

We introduce a meaningful outside option for consumers by assuming that the desire to consume is infrequent. At any point in time, consumers can be in one of two states, idle or active, captured by $\varepsilon_t \in \{0, 1\}$.[7] An idle consumer has no desire to consume ($\varepsilon_t = 0$). The desire to consume arrives at Poisson rate $\lambda > 0$, in which case the consumer becomes active ($\varepsilon_t = 1$). This desire is fulfilled after the consumption of any quantity $y > 0$ (e.g., consumers are satiated) or it disappears at Poisson rate $\gamma \geq 0$. In both events, the active consumer becomes idle. The measure of active consumers is denoted ω_1 and the measure of idle consumers is ω_0. The assumption that the consumer is satiated after having consumed any positive y is made for tractability. Alternatively, one could assume that the consumer wishes to purchase one unit of a good and negotiates with the firm about the quality, y, to be produced. In that case, satiation occurs after the consumption of one unit of any positive quality.

As an example, assume consumers wish to consume ice cream at some Poisson rate λ. Once they have consumed it, or if they have waited too long, they no longer desire ice cream for a while. (In the calibrated version of the model, agents buy different goods, possibly characterized by different λ's.) The assumption $\lambda < +\infty$, according to which consumers

[5] A similar cumulative consumption process is assumed in the continuous-time model of OTC trades of Duffie et al. (2005). If consumption (or production) of the numéraire happens in flows, then $C(t)$ admits a density, $dC(t) = c(t)dt$. If the buyer consumes or produces a discrete quantity of the numéraire good at some instant t, then $C(t^+) - C(t^-) \neq 0$.

[6] In the formulation of this environment, we separate agents according to their role (consumer, worker, firm) relative to the consumption or production of good y. It would be equivalent to consider a household composed of a unit measure of workers and a measure ω of buyers with a single integrated budget constraint that incorporates firms' profits, as in, e.g., Shi (1998).

[7] In Online Appendix D, we assume preferences are $\varepsilon\upsilon(y)$ and that ε is drawn from a continuous distribution with a positive density on an interval of $\mathbb{R}_+$.

remain temporarily idle (satiated) following consumption, $y_t > 0$, generates an opportunity cost from accepting a trade. This cost, together with bargaining shares, determines sellers' market power. In BMW, $\lambda = +\infty$, in which case consumers are always active.

Technology Each worker-firm pair produces both types of output: a constant flow, $x > 0$, of the numéraire good and endogenous quantities, y, of the decentralized-market good in bilateral meetings with consumers. The production of y units of the decentralized-market good requires $\varphi(y)$ units of numéraire.[8] It is such that $\varphi(0) = \varphi'(0) = 0$, $\varphi'(y) > 0$, and $\varphi''(y) > 0$. We denote y^* such that $\varphi'(y^*) = \upsilon'(y^*)$.

Frictions in the goods market We denote $q \equiv n/\omega_1$ as the market tightness in the decentralized goods market. It is defined as the ratio of the measure of firms to the measure of active consumers. The matching rate of a consumer is $\alpha \equiv \alpha(q)$ where $\alpha(0) = 0$, $\alpha' > 0$, $\alpha'(0) = +\infty$, and $\alpha'' < 0$. The matching rate of a firm is $\alpha^s \equiv \alpha(q)/q$.

Payments We distinguish meetings according to the method of payment. A fraction χ^d of meetings are such that the consumer can produce the numéraire to pay for his consumption. We interpret these meetings as credit meetings. With complement probability, χ^m, the consumer cannot produce the numéraire in the meeting and is not trusted to repay his debt in the future. In such meetings, the consumer pays with fiat money, an intrinsically useless object that is perfectly storable and durable. The quantity of money at time t is denoted M_t. The constant money growth rate is $\pi \equiv \dot{M}_t/M_t$ and new money is injected in the economy through lump-sum transfers (or taxes if $\pi < 0$) to consumers. The price of money in terms of the numéraire is denoted ϕ_t and the lump-sum transfer is denoted $\tau_t = \phi_t \dot{M}_t$.

In pairwise meetings in the goods market, the quantities produced and consumed, and payments, are determined according to the proportional solution of Kalai (1977), where the share of the surplus received by sellers is $\mu \in [0, 1]$. The reasons for using the Kalai solution are explained in Aruoba et al. (2007) and Hu and Rocheteau (2020), who also provide strategic foundations based on a variant of the Rubinstein game.[9]

Frictions in the labor market Labor market tightness is defined as the ratio of vacancies per unemployed worker, $\theta \equiv \nu/u$. The job finding rate of a worker is $f(\theta)$ with $f(0) = 0$,

[8] We also worked out a version of the model where the variable cost is in terms of workers' labor or disutility. It does not affect the allocations or results, except for the expression of the worker's compensation. See Online Appendix E.

[9] The monotonicity property of the Kalai solution guarantees that it implements efficient quantities at the Friedman rule (Aruoba et al., 2007). This result is instrumental for some of our proofs.

$f' > 0$, $f'(0) = +\infty$, and $f'' < 0$. The vacancy filling rate is $f(\theta)/\theta$. The flow cost of opening a vacancy is $k > 0$. We restrict our attention to contracts where workers are paid a constant wage, w, that is negotiated between the worker and the firm according to the Nash/Kalai solution. Unemployed workers receive a flow income b.

3 Equilibrium

We focus on steady-state equilibria where the distribution of agents across states and their value functions are constant.

3.1 Goods market

The value function of an active consumer with a units of real balances is $V^b(a)$ and the value function of an idle consumer is $W^b(a)$. Given the linearity of preferences with respect to the numéraire good, both value functions are linear with $V^b(a) = a + V^b$ and $W^b(a) = a + W^b$. (See Choi and Rocheteau, 2021, for details.) We denote $Z \equiv V^b - W^b$ as the loss in terms of lifetime expected utility from transitioning from being active to being idle. It represents an opportunity cost from accepting a trade. We interpret Z as the value of the consumer's outside option: it is the value of continuing searching normalized by the value of being idle.

The outcome of the negotiation between a consumer and a firm is a pair, $(p, y) \in \mathbb{R}^2_+$, where p is the payment by the consumer expressed in the numéraire and y is the output produced by the firm. The surplus of the firm is the difference between the firm's revenue expressed in terms of the numéraire and the variable cost to produce y, $p - \varphi(y)$. The firm does not incur an opportunity cost since producing for its current consumer does not affect its ability to serve its future consumers. The surplus of the consumer is the difference between the utility from consuming y net of the payment and the opportunity cost of accepting a trade, $\upsilon(y) - p - Z$. There are gains from trade if

$$\max_y \{\upsilon(y) - \varphi(y) : \varphi(y) \leq a\} > Z. \tag{4}$$

For a monetary trade to be incentive feasible, the payment, which is bounded above by a, must at least cover the firm's variable cost, $\varphi(y)$. Moreover, the utility of consumption net of the variable cost must be greater than the consumer's opportunity cost.

A necessary condition for (4) to hold is that $Z \leq \upsilon(y^*) - \varphi(y^*)$. So, the opportunity cost of the consumer is bounded above by the first-best surplus. If $a \geq \varphi(y^*)$, the condition is also sufficient. If $a < \varphi(y^*)$, then the liquidity constraint binds and (4) can be reexpressed

as $\upsilon \circ \varphi^{-1}(a) - a > Z$. The existence of gains from trade requires that the consumer holds enough real balances.

The terms of trade, which are determined according to the Kalai proportional solution, solve

$$\max_{p,y} \{p - \varphi(y)\} \quad \text{s.t. } p - \varphi(y) = \mu \left[\upsilon(y) - \varphi(y) - Z\right] \quad \text{and} \quad p \leq a. \tag{5}$$

Equation (5) can be understood as the firm choosing (p, y) to maximize its profits subject to the conditions that: (i) the profits are a fraction $\mu \in [0, 1]$ of the whole gains from trade and (ii) the consumer's payment does not exceed her real balances. Since the firm's and the consumer's surpluses are proportional to each other, the problem can be reduced to the maximization of the joint profits subject to a liquidity constraint:

$$\max_{p,y} \{\upsilon(y) - \varphi(y)\} \quad \text{s.t. } p = \varphi(y) + \mu \left[\upsilon(y) - \varphi(y) - Z\right] \leq a. \tag{6}$$

If the liquidity constraint does not bind, then $y = y^*$ and $p = (1 - \mu)\varphi(y^*) + \mu\upsilon(y^*) - \mu Z$. Otherwise, $p = a$ and $a = (1 - \mu)\varphi(y) + \mu\upsilon(y) - \mu Z$. The solution to the bargaining problem, (6), is represented graphically in Figure 3 in the utility space (U^b, U^s), where U^b is the buyer's utility of consumption net of the payment and U^s are the profits of the seller. The bargaining outcome is obtained at the intersection of the downward-sloping Pareto frontier and the upward-sloping proportional sharing rule. We denote $y\,(a, Z)$ as the outcome of the negotiation. It is nondecreasing in a and Z. In a credit trade, $y = y^*$.

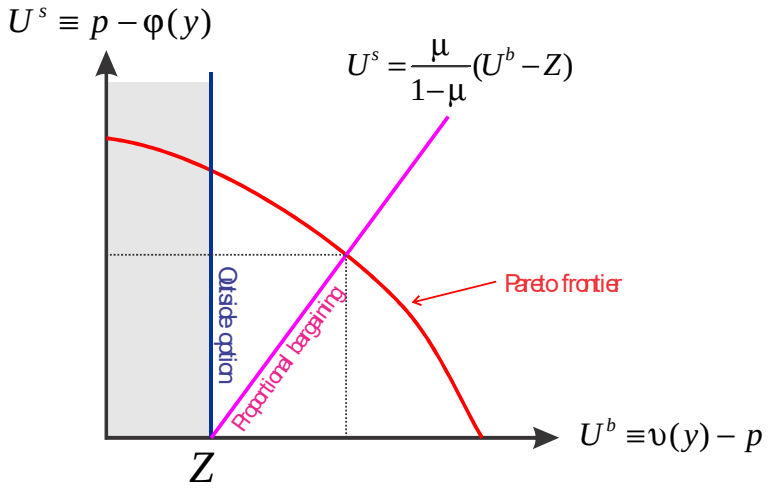


Figure 3: Bargaining in the goods market

In order to establish the connection between Z and firms' market power, we define the

markup associated with a transaction as

$$MKUP \equiv \frac{p - \varphi(y)}{\varphi(y)} = \frac{\mu \left[\upsilon(y) - \varphi(y) - Z \right]}{\varphi(y)}. \tag{7}$$

The markup is the difference between the revenue from a sale and the variable production cost expressed in percentage terms of the latter.[10] For given y, the markup increases with μ but decreases with Z. Moreover, when $y = y(a, Z)$, the markup tends to 0 as Z tends to its upper bound, $\upsilon(y^*) - \varphi(y^*)$, irrespective of μ.

We define the surpluses in monetary and credit matches as

$$S^m(a, Z) \equiv \upsilon[y(a, Z)] - \varphi[y(a, Z)] - Z \text{ and} \tag{8}$$

$$S^d(Z) \equiv \upsilon(y^*) - \varphi(y^*) - Z. \tag{9}$$

The surplus in monetary matches, given by (8), is increasing in a, and decreasing in Z. The surplus in credit matches, given by (9), is decreasing in Z.

We now turn to the value functions of the consumer. The HJB equation for the value function of an active buyer, V^b, in a steady state (i.e. $\dot{V}^b = 0$) is

$$\rho V^b = \max_{a \geq 0} \left\{ -ia + \tau + \alpha(1 - \mu) \left[\chi^m S^m(a, Z) + \chi^d S^d(Z) \right] - \gamma Z \right\}, \tag{10}$$

where $i \equiv \rho + \pi$ can be interpreted as the nominal interest rate on an illiquid bond (i.e., a bond that cannot serve as means of payment in the decentralized goods market). An active buyer incurs the cost of holding real balances, ia, receives a lump-sum transfer, τ, enters a match with a firm at Poisson rate α, and receives $1 - \mu$ share of the trade surplus. With probability χ^m, the match is monetary and the surplus is S^m. With complement probability, χ^d, the consumer can pay with the numéraire and the surplus is S^d. According to the last term on the right side of (10), if the preference shock is reversed and the consumer no longer wants to consume, at Poisson rate γ, he incurs a lifetime utility loss of $Z \equiv V^b - W^b$.

From the right side of (10), the optimal choice of real balances is given by

$$a \in \arg\max_{\hat{a} \geq 0} \left\{ -i\hat{a} + \alpha(1 - \mu)\chi^m S^m(\hat{a}, Z) \right\}. \tag{11}$$

The optimal a maximizes the consumer's expected surplus in monetary matches net of the

[10] Our measure of the markup is the average price of producing y goods, p/y, over average costs, $\varphi(y)/y$. This is related but different from defining the markup as price over marginal cost, $\varphi'(y)$, and more directly comparable to measures of gross sales margins that we use in the calibration.

cost of holding real balances. From the first-order condition,

$$a = (1-\mu)\varphi(y) + \mu\upsilon(y) - \mu Z \quad \text{where } \frac{\alpha(1-\mu)\chi^m\left[\upsilon'(y) - \varphi'(y)\right]}{\mu\upsilon'(y) + (1-\mu)\varphi'(y)} = i, \tag{12}$$

if $-ia + \alpha(1-\mu)\chi^m S^m(a, Z) \geq 0$. Otherwise, $a = 0$. Fixing the matching rate, α, from (12), the output in monetary matches does not depend on the value of the consumer's outside option but the payment does. As Z increases, consumers purchase the same amount of goods but reduce their payment.

The HJB equation for the value function of an idle buyer, W^b, is

$$\rho W^b = \tau + \lambda Z. \tag{13}$$

The idle buyer receives a preference shock with Poisson arrival rate λ, in which case she becomes active and enjoys a lifetime utility gain of $Z \equiv V^b - W^b$.

Substituting (13) from (10), the outside option of the consumer, Z, solves

$$(\rho + \lambda + \gamma) Z = \max_{a \geq 0} \left\{ -ia + \alpha(1-\mu)\left[\chi^m S^m(a, Z) + \chi^d S^d(Z)\right] \right\}. \tag{14}$$

The left side of (14) can be interpreted as the opportunity cost of consumer search – the effective discount rate multiplied by the value of consumers' outside options – while the right side is the expected return from the search activity. The effective discount rate is composed of the rate of time preference, ρ, the rate at which idle consumers become active, λ, and the rate at which active consumers become idle, γ. The right side is decreasing in Z, is positive when $Z = 0$ provided that $\chi^d > 0$, and it approaches 0 as $Z \to \upsilon(y^*) - \varphi(y^*)$. Hence, as shown in Figure 4, there is a unique $Z \in (0, \upsilon(y^*) - \varphi(y^*))$ solution to (14). As the nominal interest rate increases, the curve representing the right side of (14) moves downward and Z decreases. Intuitively, since the cost of holding real balances increases, searching for trading opportunities in the goods market becomes more costly, and thus the value of consumers' outside options decrease.

Finally, we compute the measure of idle and active buyers. The steady-state measure of active consumers solves $(\gamma + \alpha)\omega_1 = \lambda(\omega - \omega_1)$, or rewritten,

$$\omega_1 = \frac{\lambda}{\gamma + \alpha(q) + \lambda}\omega. \tag{15}$$

The measure of active consumers decreases with tightness in the goods market, q.

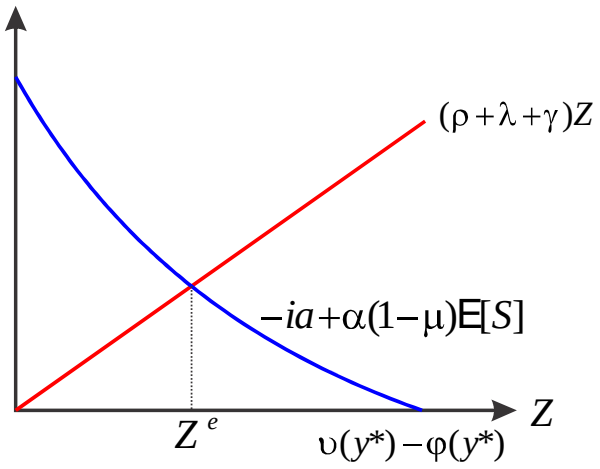


Figure 4: Determination of the value of consumers' outside options

3.2 Labor market

The HJB equation for the value of a match composed of a worker and a firm is J that solves

$$(\rho + \delta)\, J = \alpha^s \mu \left[\chi^m S^m(a, Z) + \chi^d S^d(Z)\right] + x - \rho U, \tag{16}$$

where δ is the job destruction rate, $\alpha^s \equiv \alpha(q)/q$ is the arrival rate of consumers, and U is the value function of an unemployed worker. The firm receives a share, μ, of the surplus generated by each trade. The second term on the right side is the flow of numéraire good produced by the worker-firm pair. The last term is the reservation wage of an unemployed worker. It solves

$$\rho U = b + f(\theta)\beta J, \tag{17}$$

where b is the income when unemployed, $\beta \in [0, 1]$ is the worker's bargaining share, and $f(\theta)$ is the job finding rate.

The HJB equation for an employed worker is

$$\rho E = w - \delta\beta J. \tag{18}$$

On the right side of (18), the worker receives a wage w, and, at rate δ, the match with the firm is destroyed, which generates a capital loss equal to βJ. We subtract ρU from both sides and use that $E - U = \beta J$ to obtain

$$w = (\rho + \delta)\, \beta J + \rho U. \tag{19}$$

The first term on the right side is the fraction β of the value of the match that the worker

captures. The last term corresponds to the reservation wage of the worker. We substitute $(\rho+\delta)\beta J$ by its expression given by (16) to simplify the wage equation as follows:

$$w = \beta\left\{\alpha^s\mu\left[\chi^m S^m(a,Z) + \chi^d S^d(Z)\right] + x\right\} + (1-\beta)\rho U. \tag{20}$$

Free entry of firms in the labor market implies

$$k = \frac{f(\theta)}{\theta}(1-\beta)J. \tag{21}$$

The flow cost of posting a vacancy is equal to the vacancy filling rate multiplied by the value of a filled job, $(1-\beta)J$. Substituting $f(\theta)$ from (21) into (17), $\rho U = b + \beta k\theta/(1-\beta)$. We substitute this expression into (16) and replace J from (21) to obtain the equilibrium condition for market tightness at the steady state,

$$(\rho+\delta)\frac{k\theta}{f(\theta)} = (1-\beta)\left\{\alpha^s\mu\left[\chi^m S^m(a,Z) + \chi^d S^d(Z)\right] + x - b\right\} - \beta k\theta. \tag{22}$$

Relative to the DMP model, the novelty in (22) is the term $\alpha^s\mu\left[\chi^m S^m(a,Z) + \chi^d S^d(Z)\right]$ that represents sales in the frictional goods market. There is a positive θ solution to (22). It is increasing in a and decreasing in q and Z – firms are more profitable when consumers have a higher payment capacity but they are less profitable when consumers have better outside options.

We use $\rho U = b + \beta k\theta/(1-\beta)$ to rewrite the wage in (20) as

$$w = \beta\left\{\alpha^s\mu\left[\chi^m S^m(a,Z) + \chi^d S^d(Z)\right] + x\right\} + (1-\beta)b + \beta k\theta. \tag{23}$$

Relative to both MP and BMW, the value of consumers' outside options in the goods market affects the wage. As Z increases, the market power of the firm in the goods market decreases, which reduces profits and wages.

We define the wage markdown in a symmetric fashion as the markup,

$$MKDOWN \equiv \frac{\hat{x} - w}{\hat{x}},$$

where $\hat{x}$ is the net expected revenue generated by a filled job, $\hat{x} = \mathbb{E}[p - \varphi(y)] + x$. It is also equal to

$$\hat{x} = \alpha^s\mu\left[\chi^m S^m(a,Z) + \chi^d S^d(Z)\right] + x, \tag{24}$$

which allows us to rewrite the wage markdown as

$$MKDOWN \equiv \frac{(1-\beta)\left\{\alpha^s \mu \left[\chi^m S^m(a,Z) + \chi^d S^d(Z)\right] + x - b\right\} - \beta k\theta}{\alpha^s \mu \left[\chi^m S^m(a,Z) + \chi^d S^d(Z)\right] + x}. \tag{25}$$

The markdown depends on bargaining powers in labor and goods markets as well as consumers' outside options in the goods market.

The measures of employed and unemployed workers, at the steady state, are

$$n = \frac{f(\theta)}{\delta + f(\theta)} \quad \text{and} \quad u = \frac{\delta}{\delta + f(\theta)}, \tag{26}$$

respectively. Equation (26), which is analogous to the Beveridge curve, gives a positive, steady-state relationship between employment and market tightness.

Finally, we denote the value of a firm as $\Pi \equiv (1-\beta)J$. From (21), $\Pi = \frac{k\theta}{f(\theta)}$. There is a monotone relationship between the value of a firm and labor market tightness. Market capitalization is $K \equiv n\Pi$. It can be written as $K = \frac{k\theta}{\delta + f(\theta)}$. There is also a monotone relationship between K and θ.

3.3 Definition of equilibrium

In order to simplify the definition of an equilibrium, we derive a relationship between the tightness of the goods market, q, and the tightness of the labor market, θ. From the definition $q\omega_1 \equiv n$, (15), and (26), the relationship between q and θ is given by

$$\frac{\lambda \omega q}{\gamma + \lambda + \alpha(q)} = \frac{f(\theta)}{\delta + f(\theta)}. \tag{27}$$

The implicit solution, $q = Q(\theta)$, from (27) is an increasing function of θ with $Q(0) = 0$ and $Q'(\theta) > 0$. Using $Q(\theta)$, we rewrite the matching rates for firms and consumers in the goods market as

$$\alpha^s(\theta) \equiv \frac{\alpha\left[Q(\theta)\right]}{Q(\theta)} \quad \text{and} \quad \alpha^b(\theta) \equiv \alpha\left[Q(\theta)\right],$$

respectively. The rate at which firms sell their output decreases with θ. Indeed, as θ increases, the number of active firms increases, which raises tightness in the goods market and reduces firms' matching rate with consumers.

Using the definitions of α^s and α^b above, from (11), (14), (22), and (23), an equilibrium

is a 4-tuple, (θ, a, Z, w), solution to

$$(\rho+\delta)\frac{k\theta}{f(\theta)} = (1-\beta)\left\{\alpha^s(\theta)\mu\left[\chi^m S^m(a,Z)+\chi^d S^d(Z)\right]+x-b\right\}-\beta k\theta, \quad (28)$$

$$a \in \arg\max_{\hat{a}\geq 0}\left\{-i\hat{a}+\alpha^b(\theta)(1-\mu)\chi^m S^m(\hat{a},Z)\right\}, \quad (29)$$

$$(\rho+\lambda+\gamma)Z = -ia+\alpha^b(\theta)(1-\mu)\left[\chi^m S^m(a,Z)+\chi^d S^d(Z)\right], \text{ and} \quad (30)$$

$$w = \beta\left\{\alpha^s(\theta)\mu\left[\chi^m S^m(a,Z)+\chi^d S^d(Z)\right]+x\right\}+(1-\beta)b+\beta k\theta. \quad (31)$$

Proposition 1 *If $\chi^d > 0$, then there exists an active steady-state equilibrium.*

The logic for the existence result goes as follows.[11] From (30) we express the value of the consumer's outside option, Z, as a function of labor market tightness, θ. It is an increasing function – as labor market tightness increases, more firms are created, and hence the measure of producers per consumer in the goods market increases, which improves consumers' outside options. From (29), we express consumers' real balances, a, as a function of θ. We obtain an increasing function, because as θ increases, the average time for an active consumer to find a producer decreases, which in turn reduces the average holding cost of real balances. We then substitute $Z(\theta)$ and $a(\theta)$ into (28) to obtain a single equilibrium condition in θ. We use the continuity properties of this equation and its values at $\theta = 0$ and $\theta = +\infty$ to establish that a positive solution exists. A sufficient condition for the existence of an active equilibrium is that a positive measure of transactions is conducted with credit, $\chi^d > 0$. Indeed, if θ becomes very small, the expected revenue of firms becomes very large because they can serve a large measure of consumers per firm. Hence, irrespective of b or x, there is always a sufficiently low θ so that firms' profits are positive and entry is profitable. Before we get to our main result, we consider some special cases.

3.4 Cashless economies

Suppose that all meetings in the decentralized goods market are credit meetings, $\chi^d = 1$, as in, e.g., Petrosky-Nadeau and Wasmer (2015, 2017). From (28)-(31), an equilibrium can then be reduced to a triple (θ, Z, w) that solves

$$(\rho+\delta)\frac{k\theta}{f(\theta)} = (1-\beta)\left\{\alpha^s(\theta)\mu\left[\upsilon(y^*)-\varphi(y^*)-Z\right]+x-b\right\}-\beta k\theta, \quad (32)$$

$$(\rho+\lambda+\gamma)Z = \alpha^b(\theta)(1-\mu)\left[\upsilon(y^*)-\varphi(y^*)-Z\right], \text{ and} \quad (33)$$

$$w = \beta\left\{\alpha^s(\theta)\mu\left[\upsilon(y^*)-\varphi(y^*)-Z\right]+x\right\}+(1-\beta)b+\beta k\theta. \quad (34)$$

[11] The proof is available in an online appendix.

We start by considering two polar cases where Z has no effect on labor market outcomes. If firms have no bargaining power in the decentralized goods market, $\mu = 0$, then θ, which is determined from (32), is independent of Z and identical to the equilibrium condition of the MP model. The other polar case is when $\mu = 1$. From (33), $Z = 0$ and from (32) θ is uniquely determined by

$$(\rho + \delta) \frac{k\theta}{f(\theta)} = (1 - \beta) \{\alpha^s(\theta) [\upsilon(y^*) - \varphi(y^*)] + x - b\} - \beta k\theta.$$

The frictions in the goods market, as captured by α^s, have an impact on the labor market, i.e., as α^s decreases, θ goes down. However, consumer search creates no competitive pressure due to the Diamond paradox.

Next, consider the interior case where $\mu \in (0, 1)$. From (32), θ is a decreasing function of Z. As the value of consumers' outside options increase, the profits of the firm decrease. From (33), we can solve for Z in closed form and obtain

$$Z = \frac{\alpha^b(\theta)(1 - \mu)}{\rho + \lambda + \gamma + \alpha^b(\theta)(1 - \mu)} [\upsilon(y^*) - \varphi(y^*)]. \tag{35}$$

The value of consumers' outside options increases with θ and decreases with μ. As frictions vanish, $\alpha^b \to +\infty$, and $Z \to \upsilon(y^*) - \varphi(y^*)$. We can substitute the expression for Z into (32) to reduce an equilibrium to a single equation in θ,

$$(\rho + \delta) \frac{k\theta}{f(\theta)} + \beta k\theta = (1 - \beta) \left\{ \frac{\alpha^s(\theta)\mu (\rho + \lambda + \gamma)}{\rho + \lambda + \gamma + \alpha^b(\theta)(1 - \mu)} [\upsilon(y^*) - \varphi(y^*)] + x - b \right\}. \tag{36}$$

It is easy to check that an equilibrium of the cashless economy exists and is unique. The following proposition shows how the determinants of market power in the goods market, λ and γ, affect labor market outcomes.

Proposition 2 ***(Unemployment and consumer search in a cashless economy)*** *Suppose $\chi^d = 1$ and $\mu \in (0, 1)$. As λ or γ increases, the value of consumers' outside options (Z) decreases, labor market tightness (θ) increases, wages (w) increase, and unemployment (u) decreases.*

If λ increases, i.e., consumers do not stay idle long, or γ increases, i.e., the desire to consume vanishes quickly, then the opportunity cost of accepting a trade decreases, which makes the search for an alternative producer less profitable and raises producers' market power. As a result, market tightness in the labor market increases, $\partial\theta/\partial\lambda > 0$ and $\partial\theta/\partial\gamma >$

0, wages increase, $\partial w/\partial\lambda > 0$ and $\partial w/\partial\gamma > 0$, and unemployment decreases, $\partial u/\partial\lambda < 0$ and $\partial u/\partial\gamma < 0$.

3.5 Pure currency economies without consumer search

The economy in BMW is a pure currency economy, $\chi^m = 1$, in which there is no opportunity cost for the consumer to complete a trade, $\lambda = +\infty$. From (15), all buyers are active at all points in time, $\omega_1 = \omega$. From (27), market tightness in the goods market is

$$q = \frac{f(\theta)}{\omega\left[\delta + f(\theta)\right]}. \tag{37}$$

From (30), $Z = 0$ and an equilibrium can be reduced to a pair (θ, y) that is solution to

$$(\rho+\delta)\frac{k\theta}{f(\theta)} = (1-\beta)\left\{\alpha^s(\theta)\mu\left[\upsilon(y)-\varphi(y)\right]+x-b\right\}-\beta k\theta \quad \text{and} \tag{38}$$

$$\frac{\upsilon'(y)-\varphi'(y)}{\mu\upsilon'(y)+(1-\mu)\varphi'(y)} = \frac{i}{(1-\mu)\alpha^b(\theta)}. \tag{39}$$

Given (θ, y), the wage is given by

$$w = \beta\left\{\alpha^s(\theta)\mu\left[\upsilon(y)-\varphi(y)\right]+x\right\}+(1-\beta)b+\beta k\theta. \tag{40}$$

Suppose $x < b$. Equation (38) gives a positive relationship between θ and y, with $\theta = 0$ when $y = 0$ and $\theta = \bar{\theta} > 0$ when $y = y^*$. Equation (39) gives a positive relationship between y and θ, with $y = 0$ if $\theta < \underline{\theta}$, where $\underline{\theta}$ solves $\alpha^b(\underline{\theta}) = [i\mu/(1-\mu)]$ and $y \to y^*$ as $\theta \to +\infty$. There always exists a non-active equilibrium with $y = \theta = 0$ and, generically, if an active equilibrium exists, then the number of active equilibria is even.[12] In the following, we focus on the equilibrium with the highest θ. In our model, inflation, π, affects allocations via $i \equiv \rho + \pi$, making it equivalent to study changes in π or changes in i.

Proposition 3 ***(Long-run Phillips curve in the absence of consumer search)*** *Assume $x < b$ and $\lambda = +\infty$ and focus on the equilibrium with the highest market tightness. An increase in i leads to a decrease in market tightness (θ), a decrease in wages (w), and an increase in unemployment (u).*

We provide a graphical proof (see Figure 5) as the result is similar to the one in BMW. At the high equilibrium, the MD curve representing (39) in the (θ, y) space intersects the

[12] These results mirror those of Rocheteau and Wright (2005) who describe a pure currency economy with free entry of sellers.

JC curve representing (38) from above. As i increases, the MD curve (39) shifts downward. Hence, y and θ decrease. So the model predicts a positive relationship between unemployment and inflation in the long run, i.e., the long-run Phillips curve is upward sloping.

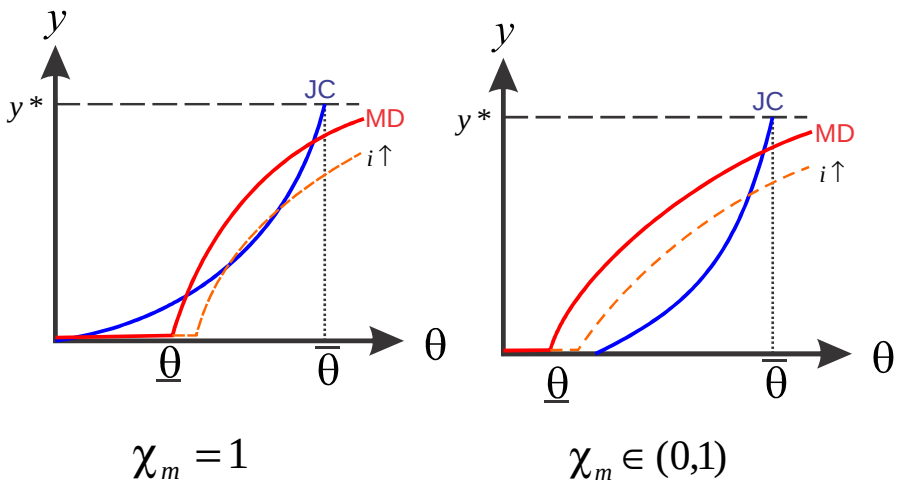


Figure 5: Left: Pure monetary economy. Right: Economy with money and credit.

The logic is based on the following *real-balance effect* of inflation. As i increases, consumers reduce their real money holdings, which lowers the gains from trade that can be extracted in the goods market. Firms, whose profits fall, post fewer vacancies and the unemployment rate increases. The result in Proposition 3 is robust if we reintroduce some credit trades, $\chi^d \in (0,1)$. What matters is that consumers have no outside option when $\lambda = +\infty$ so that inflation affects θ only through the real-balance effect.

4 Positive and normative implications

We now turn to the effects of inflation in an economy with consumer search, $\lambda < +\infty$. We will show that introducing consumers' endogenous outside options generates new insights for the slope of the Phillips curve, the convergence to perfect competition outcomes as frictions vanish, and the optimal monetary policy.

4.1 The non-monotone long-run Phillips curve

We study increases in the inflation rate of different magnitudes: (1) a small increase in the neighborhood of the Friedman rule, $i = 0$, or (2) a large increase from $i = 0$ to $i = +\infty$ that reduces the value of money to zero.

Proposition 4 ***(The non-monotone long-run Phillips curve)*** *Suppose* $\chi^m \in (0,1)$.

1. *A small increase in i, starting from $i = 0^+$, leads to: an increase in labor market tightness (θ); a decrease in the unemployment rate (u); an increase in wages (w); and an increase in stock prices (Π) and market capitalization (K).*

2. *A large increase in i from $i = 0^+$ to $i = +\infty$ leads to: a decrease in labor market tightness (θ); an increase in the unemployment rate (u); a decrease in wages (w); and a decrease in stock prices (Π) and market capitalization (K).*

An increase in inflation has two effects on the labor market. First, it reduces consumers' real balances, which tightens liquidity constraints and reduces the amount of goods firms can sell to consumers. As shown in Proposition 3, this effect reduces market tightness, raises unemployment, and reduces wages. There is a second effect according to which an increase in i raises the cost for consumers to search for an alternative producer since they have to carry real money balances until they find a new opportunity to trade. When i is close to 0, y is close to y^*, and the first effect – *the real-balance effect* – is negligible.[13] Only the second effect on consumers' outside options – *the market-power effect* – is of first order magnitude, i.e., inflation raises firms' market power by making it more costly for consumers to exercise their outside option.[14]

We illustrate the two effects graphically in Figure 6 that plots the outcome of bargaining in pairwise meetings in the goods market. As inflation increases, a decreases, and the Pareto frontier of the bargaining problem shifts downward. It reduces the profits of the firm, U^s. But as inflation increases, Z decreases, thereby shifting the origin of the rent sharing line to the left, which increases U^s. At the Friedman rule, $i = 0$, the Pareto frontier is linear when it intersects the rent sharing line, and this linear part does not shift as a decreases. Hence, only the shift of Z matters for the allocations in pairwise meetings.

When i is sufficiently large, the two effects described above are first order. We can show that when i is so high that consumers do not hold real balances, i.e. all trades are conducted with credit, then firms are worse-off compared to the equilibrium at the Friedman rule, i.e. θ is lower when $i = +\infty$ relative to $i = 0$. This guarantees that the relation between θ and i is non-monotone; it is first rising and eventually decreases for sufficiently large values of i.

[13] For this result, it is important that the bargaining solution is monotone and implements y^* when $i = 0$. See the discussion in Aruoba et al. (2007), and Hu and Rocheteau (2020) for why the monotonicity assumption is natural. Under Nash bargaining, $i = 0$ does not implement y^* and hence a deviation from the FR has a first-order effect on y. The two effects of inflation would still be present but it would not be guaranteed that the market power effect dominates at low inflation rates. However, the gradual Nash solution of Rocheteau et al. (2021) is monotone and would generate the same results.

[14] In Online Appendix D, we consider a version with horizontally differentiated products and show that as inflation increases, consumers become less picky and purchase varieties of the good that they value less – the standard 'hot potato' effect of inflation arises.

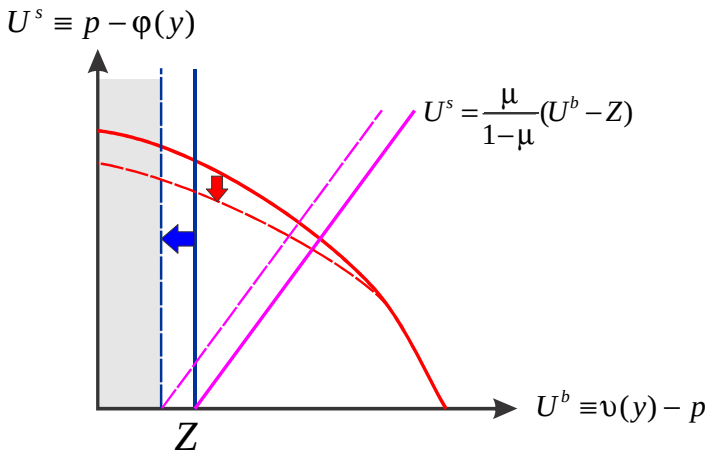


Figure 6: Effects of inflation on the bargaining outcome

Since real wages increase with firms' profits (we show in the proof of Proposition 4 that w and θ comove as i changes), the relationship between w and i is also non-monotone. A small, anticipated inflation rate pushes *real* wages up while a large inflation rate depresses *real* wages. Hence, the unemployment minimizing level of i maximizes real wages.

Since the market power effect of inflation exists when $\lambda < +\infty$ and vanishes as $\lambda \to +\infty$, one might conjecture that the passthrough of monetary policy to the labor market, from i to θ, decreases monotonically in λ. In Online Appendix G we show it is not necessarily the case by studying how $\partial\theta/\partial i|_{i=0}$ varies with λ. We consider the special case where agents are infinitely patient, $\rho = 0$, and the job separation rate vanishes, $\delta \to 0$. An increase in λ has two opposing effects on $\partial\theta/\partial i$ when evaluated at $i = 0$. First, as λ rises, the outside option, Z, is closer to 0, and thus Z is less sensitive to changes in inflation, thereby reducing $\partial\theta/\partial i$. Second, as Z falls, consumers carry more real money balances. This effect through a increases $\partial\theta/\partial i$ because buyers' cost of searching for producers becomes more sensitive to inflation. In general, $\partial\theta/\partial i$ is non-monotone in λ due to these two opposing forces.

4.2 Frictionless limits

We now characterize equilibrium outcomes when search frictions i) in the goods market vanish, ii) in labor market vanish, and iii) vanish in both markets simultaneously. We interpret vanishing frictions in terms of matching technologies that become infinitely efficient at pairing agents. We show that the results of Gale (1986a,b, 1987) regarding frictionless limits apply to our economy.

Before we state our results for the generic case, $\lambda \in (0, +\infty)$, it is useful to consider as a reference point the BMW economy, in which consumers are always active ($\lambda = +\infty$). We

write the matching function in the goods market as $\alpha(q) = A\bar{\alpha}(q)$, where $A > 0$, and we take the limit as $A \to +\infty$.

Proposition 5 ***(Frictionless limits: The BMW model.)*** *Suppose $\lambda = +\infty$, $\chi^d > 0$, $\mu \in (0,1)$, and $\alpha(q) = A\bar{\alpha}(q)$. As $A \to +\infty$, $y \to y^*$, $q \to 1/\omega$, $\theta \to +\infty$, $\alpha^b \to +\infty$, and $\alpha^s \to +\infty$. Rents in pairwise meetings approach their maximum value, $\upsilon(y^*) - \varphi(y^*) > 0$. Markups remain bounded away from zero, $MKUP \to \mu\left[\upsilon(y^*)/\varphi(y^*) - 1\right] > 0$.*

In Proposition 5, we show output in pairwise meetings is undistorted at the frictionless limit because the cost of holding money is irrelevant when consumers can find a seller almost instantly.[15] Since $Z = 0$, surpluses, $\upsilon(y^*) - \varphi(y^*)$, are maximum, i.e., there are rents to be bargained over even when frictions vanish and those rents are as large as they can possibly be. Labor market tightness becomes unbounded because consumers are always active and can find sellers from whom to buy at infinite speed. The consumer base of each firm, defined as the expected number of consumers served by a firm per unit of time, α^s, grows unbounded as A explodes.

We represent the outcome of bargaining in the left panel of Figure 7. The Pareto frontier is the farthest away from the origin, i.e., rents are maximum, and the rent sharing line intersects the Pareto frontier in its linear part, i.e., $y = y^*$. Following Makowski and Ostroy (2001), we define perfect competition as a situation where rents are zero. By that criterion, the equilibrium outcome of BMW does not converge to a perfect competition outcome as frictions vanish. Another way to establish the lack of convergence to perfect competition is by computing the markup at the limit, which is positive at the frictionless limit.

We now contrast these results to the ones in an economy where consumers have meaningful outside options in both the goods and labor markets.

Proposition 6 ***(Frictionless limits: The general case)*** *Suppose $\lambda \in (0, +\infty)$, $\alpha(q) = A\bar{\alpha}(q)$ and $f(\theta) \equiv B\overline{f}(\theta)$.*

1. ***Limit as the goods market becomes frictionless****. Consider the limit as $A \to +\infty$. Then, $Z \to \upsilon(y^*) - \varphi(y^*)$ and $MKUP \to 0$. If $x > b$, then $q \to +\infty$ and $\theta \to \theta^g_\infty$, where $\theta^g_\infty > 0$ is the unique solution to*

$$(\rho + \delta)\frac{k\theta^g_\infty}{f(\theta^g_\infty)} + \beta k\theta^g_\infty = (1-\beta)(x-b). \tag{41}$$

   If $x \leq b$, then $\theta \to 0$.

[15] The proof is available in an online appendix.

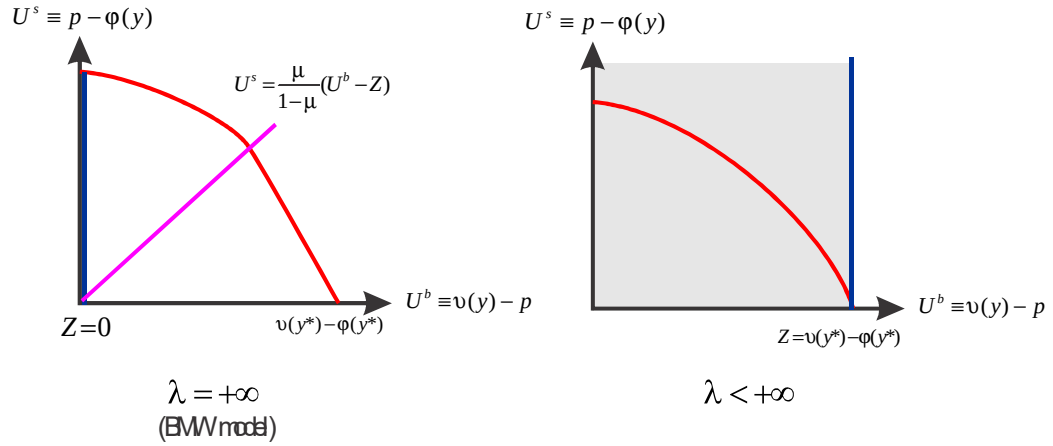


Figure 7: Outcome of bargaining when the goods market becomes frictionless ($A \to +\infty$). Left panel: BMW economy ($\lambda = +\infty$). Right panel: Economy with consumers' outside options ($\lambda < +\infty$).

2. ***Limit as the labor market becomes frictionless***. *Consider the limit as* $B \to +\infty$. *Then,* $(\theta, a, Z) \to \left(\theta_\infty^\ell, a_\infty^\ell, Z_\infty^\ell\right)$, *where* $\left(\theta_\infty^\ell, a_\infty^\ell, Z_\infty^\ell\right)$ *is the solution to*

$$\begin{aligned} \beta k \theta_\infty^\ell &= (1-\beta)\left\{\alpha^s(\theta_\infty^\ell)\mu\left[\chi^m S^m(a_\infty^\ell, Z_\infty^\ell) + \chi^d S^d(Z_\infty^\ell)\right] + x - b\right\}, && (42) \\ a_\infty^\ell &\in \arg\max_{a \geq 0}\left\{-ia + \alpha^b(\theta_\infty^\ell)(1-\mu)\chi^m S^m(a, Z_\infty^\ell)\right\}, \text{ and} && (43) \\ (\rho + \lambda + \gamma) Z_\infty^\ell &= \max_{a \geq 0}\left\{-ia + \alpha^b(\theta_\infty^\ell)(1-\mu)\left[\chi^m S^m(a, Z_\infty^\ell) + \chi^d S^d(Z_\infty^\ell)\right]\right\}. && (44) \end{aligned}$$

Moreover, $w \to \hat{x}_\infty^\ell$, *where*

$$\hat{x}_\infty^\ell = \alpha^s(\theta_\infty^\ell)\mu\left[\chi^m S^m(a_\infty^\ell, Z_\infty^\ell) + \chi^d S^d(Z_\infty^\ell)\right] + x, \tag{45}$$

and $MKDOWN \to 0$.

3. ***Limit as both markets become frictionless***. *Consider the limit as* $A \to +\infty$ *and* $B \to +\infty$. *Then,* $Z \to \upsilon(y^*) - \varphi(y^*)$ *and* $w \to x$. *If* $x > b$, *then* $\theta \to \theta_\infty$, *where*

$$\theta_\infty = \frac{(1-\beta)(x-b)}{\beta k}. \tag{46}$$

If $x \leq b$, *then* $\theta \to 0$.

As the frictions in the goods market vanish (Part 1 of Proposition 6), the value of consumers' outside options exhaust the gains from trade, $Z \to \upsilon(y^*) - \varphi(y^*)$. As a result, rents in pairwise meetings, $S^m(a, Z)$ and $S^d(Z)$, go to zero. This finding is illustrated in the right panel of Figure 7 that shows that the outcome of bargaining is pinned down by consumers'

outside options. It is a definition of perfect competition in the goods market (e.g., Makowski and Ostroy, 2001). The quantities are efficient, $y \to y^*$, and the payment just covers the production cost, $p \to \varphi(y^*)$, i.e., the average markup goes to zero. Provided $x > b$, market tightness is positive and bounded at the frictionless limit, but it is independent of monetary policy, i.e., the long-run Phillips curve becomes vertical.

The results above, which hold for all $\lambda < +\infty$, are in sharp contrast with the ones of the BMW economy described earlier.[16] While in the BMW economy, rents and markups remain positive at the frictionless limit, they all disappear when consumers incur an opportunity cost from trading. Moreover, the consumer base remains bounded when $\lambda < +\infty$ because as consumers trade faster, the measure of active consumers shrinks.

Our results can be related to Lauermann (2013) that provides a necessary and sufficient condition for decentralized market equilibria under quasi-linear preferences to converge to a Walrasian outcome as the rate of time preference approaches zero. He shows that convergence occurs if and only if the economy features *competitive pressure*, in the sense that buyers can secure a positive trade surplus in at least some meetings, and the surplus increases as frictions vanish. In our model, competitive pressure only exists if $Z > 0$, which requires both that consumers obtain a positive surplus in meetings, $\mu < 1$, and they have outside options, $\lambda < +\infty$. As A increases, consumers can realize their outside options faster, which exacerbates the competitive pressure. In contrast, if $Z = 0$ then there is no competitive pressure, irrespective of A.

As the frictions in the labor market vanish (Part 2 of Proposition 6), market tightness approaches a finite and positive limit, defined in (42). The job finding rate, $B\overline{f}(\theta^{\ell}_{\infty})$, and the vacancy filling rate, $B\overline{f}(\theta^{\ell}_{\infty})/\theta^{\ell}_{\infty}$, go to infinity. Hence, unemployment vanishes asymptotically and the Phillips curve becomes vertical at $u = 0$. The real wage approaches the average productivity of a worker so that the markdown tends to 0.

Finally, in Part 3 of Proposition 6, we describe the limit as both frictions in the labor and goods market vanish. A competitive outcome is obtained in both markets. In the goods market, the value of consumers' outside options exhausts all rents, $Z \to \upsilon(y^*) - \varphi(y^*)$, and markups tend to zero. In the labor market, the wage approaches workers' productivity, $w \to x$, i.e., markdowns go to zero. From (46), market tightness is bounded at the limit and it decreases with workers' bargaining power.

[16] One implication of these results is that the order according to which we take the limits, $\lambda \to +\infty$ and $A \to +\infty$, matters. If we take the limit $\lambda \to +\infty$ first, we obtain the BMW economy. As $A \to +\infty$, the economy grows unbounded but remains imperfectly competitive. In contrast, if we take the limit $A \to +\infty$ first then the equilibrium outcome converges to a perfect-competition outcome that is independent of λ. So, as λ goes to $+\infty$, allocations remain competitive and labor market tightness stays bounded.

4.3 Welfare

In the presence of search externalities in decentralized goods and labor markets, the monetary policy that minimizes the unemployment rate is not necessarily the one that maximizes social welfare. As shown in Proposition 4, an increase in π above the Friedman rule reduces unemployment by raising firms' market power. If firms' market power is at a level that generates efficient entry, then a deviation from the Friedman rule might be welfare-reducing. We explore this conjecture by first characterizing the constrained-efficient allocations and by comparing them to the equilibrium allocations.

4.3.1 Constrained-efficient allocations

The planner's problem is

$$\max \int_0^{+\infty} e^{-\rho t} \varpi_t dt, \tag{47}$$

subject to

$$\dot{\omega}_{1,t} = \lambda(\omega - \omega_{1,t}) - \left[\alpha\left(\frac{n_t}{\omega_{1,t}}\right) + \gamma\right]\omega_{1,t}, \quad \text{and} \tag{48}$$

$$\dot{n}_t = f(\theta_t)(1 - n_t) - \delta n_t, \tag{49}$$

where n_0 and $\omega_{1,0}$ are given and the instantaneous social welfare, ϖ_t, is given by

$$\begin{aligned} \varpi_t = {} & \omega_{1,t}\alpha\left(\frac{n_t}{\omega_{1,t}}\right)\left\{\chi^m\left[\upsilon(y_{m,t}) - \varphi(y_{m,t})\right] + \chi^d\left[\upsilon(y_{d,t}) - \varphi(y_{d,t})\right]\right\} \\ & + n_t(x - b) + b - (1 - n_t)\theta_t k. \end{aligned} \tag{50}$$

The state variables are $\omega_{1,t}$ and n_t. The control variables are $y_{m,t}$, $y_{d,t}$, and θ_t. According to (50), welfare is the sum of the trade surpluses in pairwise meetings plus the production of the numeraire good by employed workers, x, and unemployed workers, b, net of the vacancy posting costs. (Here, b is not treated as a transfer).

We denote ξ_t as the current-value co-state variable associated with n_t, and ζ_t the current-value co-state variable associated with $\omega_{1,t}$. So ξ_t is the shadow value of a match in the labor market while ζ_t is the shadow value of an active buyer. The social surplus from a trade in a pairwise meeting is $\upsilon(y_{s,t}) - \varphi(y_{s,t}) - \zeta_t$ where $s \in \{m, d\}$. The first-order conditions with respect to $y_{m,t}$ and $y_{d,t}$ give $y_{m,t} = y_{d,t} = y^*$ for all t. The first-order condition with respect to θ_t gives

$$k = \xi_t f'(\theta_t). \tag{51}$$

The planner equalizes the cost of posting a vacancy with its marginal benefit as measured by the product of the shadow value of an employed worker, ξ_t, and the marginal increase in the job finding rate, $f'(\theta_t)$. As shown in the proof of Proposition 7 below, at a stationary solution to the planner's problem, ζ and θ solve

$$(\rho+\lambda+\gamma)\,\zeta = \alpha(q)\left[1-\epsilon_\alpha(q)\right]\left[\upsilon(y^*)-\varphi(y^*)-\zeta\right] \quad \text{and} \tag{52}$$

$$(\rho+\delta)\,\frac{k\theta}{f(\theta)} = \epsilon_f(\theta)\left\{\frac{\alpha(q)}{q}\epsilon_\alpha(q)\left[\upsilon(y^*)-\varphi(y^*)-\zeta\right]+x-b\right\}-\left[1-\epsilon_f(\theta)\right]k\theta, \tag{53}$$

where $\epsilon_\alpha \equiv \alpha'(q)\,q/\alpha(q)$ and $\epsilon_f \equiv f'(\theta)\theta/f(\theta)$ denote the elasticities of the matching rates in goods and labor markets, respectively. Equation (52) is the analog to (33) that determines Z in equilibrium. Equation (53) is the analog to (32) that determines θ in equilibrium. In the proof of Proposition 7, we show that there is a unique $(\theta^s, q^s, n^s, \omega_1^s)$ that is a solution to (26), (27), (52), and (53). It is the solution to the planner's optimal control problem when $n_0 = n^s$ and $\omega_{1,0} = \omega_1^s$.

Proposition 7 ***(Constrained efficiency)***. *A decentralized equilibrium implements the constrained-efficient allocation if* $i_t = 0$*,*

$$\mu = \epsilon_\alpha(q^s), \tag{54}$$

$$1-\beta = \epsilon_f(\theta^s), \tag{55}$$

and the initial values n_0 *and* $\omega_{1,0}$ *equal the steady state values* n^s *and* ω_1^s*.*

A decentralized equilibrium maximizes social welfare when monetary policy implements the Friedman rule, $i_t = 0$, and the Hosios conditions in both the labor and goods markets hold.[17] It should be noted that imposing the Hosios condition in each market is not necessary to implement the constrained-efficient allocation. Indeed, from (36), labor market tightness at the Friedman rule is the unique solution to

$$(\rho+\delta)\,\frac{k\theta}{f(\theta)} = (1-\beta)\left\{\alpha^s(\theta)\frac{\mu(\rho+\lambda+\gamma)}{\rho+\lambda+\gamma+\alpha^b(\theta)(1-\mu)}\left[\upsilon(y^*)-\varphi(y^*)\right]+x-b\right\}-\beta k\theta. \tag{56}$$

[17] The proof for Proposition 7 is available in an online appendix. This result is a generalization of Berentsen et al. (2007) to an economy where both goods and labor markets are frictional. The logic of this result requires that the Friedman rule implements y^* which is the case if the bargaining solution is monotone but would not be true for the Nash solution. See Aruoba et al. (2007) for a discussion. Relatedly, Mangin and Julien (2021) derive a "generalized Hosios condition" for a static version of BMW. Petrosky-Nadeau and Wasmer (2017) obtain a similar condition in a nonmonetary economy where credit, labor, and goods markets are frictional. In Online Appendix F, we show that if there is competitive search in both the goods and labor markets, then the Hosios conditions are satisfied and the Friedman rule is optimal.

From Proposition 7, the constrained-efficient allocation is achieved for $\beta = 1 - \epsilon_f(\theta^{sp})$ and $\mu = \epsilon_\alpha[q(\theta^{sp})]$, where θ^{sp} is the constrained-efficient labor market tightness. There are other combinations of (β, μ) that achieve the efficient level of market tightness. For instance, if β is slightly above $1 - \epsilon_f(\theta^{sp})$, then a μ slightly below $\epsilon_\alpha(q^{sp})$ will implement the same θ^{sp}. This multiplicity arises because the planner can control two bargaining shares to target a single variable, θ.

4.3.2 Optimal monetary policy

We now study the optimality of the Friedman rule. In order to simplify the analysis, we follow Hosios (1990) and Pissarides (2000) and assume that agents are infinitely patient, i.e. $\rho \to 0$, which allows us to focus on steady-state welfare. The planner maximizes ϖ in (50) by choosing the inflation rate (or the nominal interest rate) subject to the law of motion of workers and consumers and the free entry of firms.

Proposition 8 ***(Optimal monetary policy)***

1. *If $\lambda = +\infty$, then the Friedman rule, $i = 0$, is optimal.*
2. *If $\lambda < +\infty$, then there exists a $\bar{\beta}(\mu) \in [0, 1)$ such that $i > 0$ is locally optimal if and only if $\beta > \bar{\beta}(\mu)$. The threshold, $\bar{\beta}(\mu)$, rises in μ.*

When $\lambda = +\infty$, the Friedman rule is optimal regardless of whether labor market tightness is inefficiently high or inefficiently low (relative to the planner's solution). If θ is too high, because firms have too much bargaining power, a deviation from the Friedman rule can lower firm entry, thereby improving welfare. However, there is a negative effect on quantities traded, which reduces welfare. The second effect dominates so that the Friedman rule is always optimal.[18]

When $\lambda < +\infty$, a small deviation from the Friedman rule, i.e. $i > 0$ (or, equivalently, $\pi > -\rho$), has a first-order impact on firm entry by reducing the value of consumers' outside option, Z. This deviation improves welfare if the market tightness, θ, is inefficiently low at the steady-state equilibrium, which happens when firms have too little bargaining power.

In the left panel of Figure 8, we plot the welfare-maximizing inflation rate as a function of β and μ. The choice of the parameter values, which are calibrated to the US economy, and functional forms will be discussed in greater detail in the next section. For now, it suffices to know here that the matching functions, $f(\theta)$ and $\alpha(q)$, are Cobb-Douglas with

[18] A similar result is derived by Rocheteau and Wright (2005) in the context of a search-and-bargaining monetary model with free entry of sellers.

elasticities equal to 0.6 in the labor market and 0.5 in the goods market. The line $\bar{\beta}(\mu)$ in Proposition 8 is illustrated as an upward sloping black line, which starts at a point where β is a little less than 0.2. It contains the combinations of (β, μ) that implement the constrained-efficient market tightness defined in (56). One point on this line corresponds to the Hosios condition in both markets, namely $\beta = 0.4$ and $\mu = 0.5$. Therefore, along the black line, the Friedman rule implements the constrained-efficient allocation. The black line is upward sloping because if θ^* is implemented for a pair (β, μ) and if the worker's bargaining power in the labor market increases, then the same optimal market tightness can be maintained by raising the bargaining power of firms in the goods market.

To the left of this black line, the average bargaining power of firms across goods and labor markets is too high. In that case, it is optimal to set the money growth rate at the Friedman rule since an increase in the inflation rate would exacerbate the inefficiently high entry of firms. To the right of the black line, the average bargaining power of firms is too low. In that case, the optimal policy is $i > 0$ and we illustrate the welfare-maximizing inflation rate with contour lines. In this region, firm entry and labor market tightness are too low compared to the planner's solution in (52)-(53). Increasing inflation above the Friedman rule reduces consumers' outside options, thereby increasing firms' market power and incentivizing entry. It also raises the cost of holding money, but, in the neighborhood of $y = y^*$, the effect on social welfare is second order.

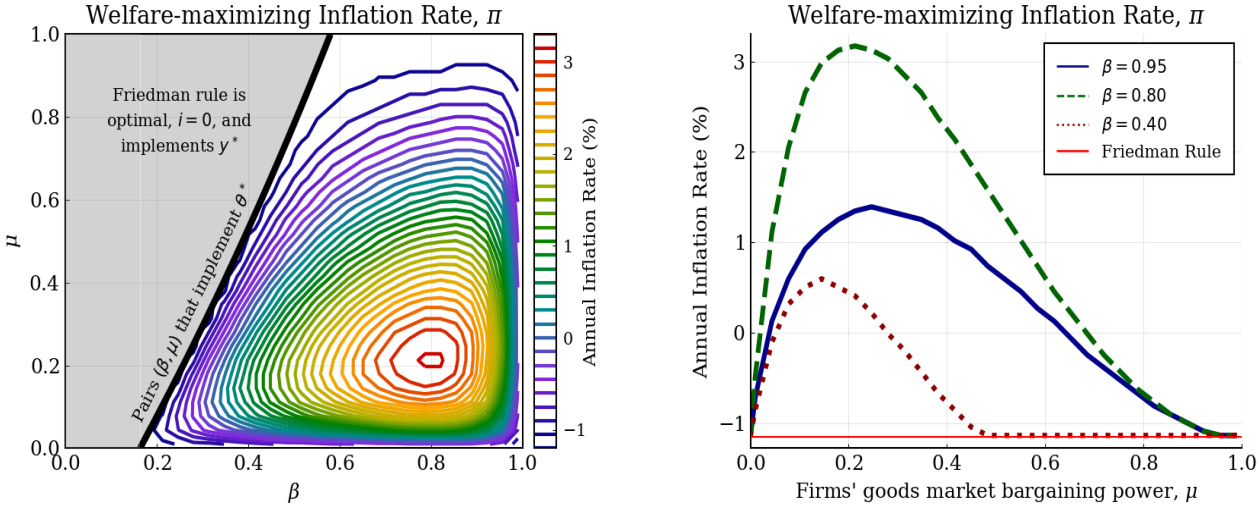


Figure 8: Welfare-maximizing inflation rate, given μ and β.

In the right panel of Figure 8, the Friedman rule is represented by the horizontal red line, which is $\pi_{FR} \approx -0.012$ at an annual frequency in this parameterization. For a given β, the welfare-maximizing inflation rate is non-monotone in firms' bargaining power in the goods market, μ. When $\mu = 0$, firms receive no surpluses from trade. Hence an increase in inflation

reduces consumers' outside options but does not increase firm entry, i.e. the market power effect of inflation is not in operation. In this case, the welfare-maximizing inflation rate is $\pi = \pi_{FR} \equiv -\rho$. When $\mu \in (0,1)$, the market power effect of inflation exists and it is optimal to raise π above $-\rho$. When $\mu \to 1$, the optimal π returns to π_{FR} because consumers' outside options $Z \to 0$ and thus the market power effect again vanishes. This example illustrates that the market power effect is non-monotone in firms' bargaining power in general.

5 Quantitative assessment

We now carry out a quantitative analysis by calibrating our model to the US between 1955-2005. We generalize the model by introducing heterogeneity across consumers in terms of λ's and μ's in order to capture realistic dispersion in consumption frequency and markups across components of household consumption.

5.1 The generalized model

Suppose there are $J \in \mathbb{N}$ categories of goods and J types of consumers. A consumer of type $j \in \{1, ..., J\}$ is specialized in the consumption of good j. The measure of consumers of type j is ω_j with $\sum_{j=1}^{J} \omega_j = \omega$. What distinguishes goods is the frequency at which they are consumed, λ_j, and the bargaining power of consumers when purchasing them, $1 - \mu_j$. For tractability, firms are ex-ante identical, can produce all categories of goods, and are randomly matched with consumers of different types.[19]

From (30), the value of the outside option of consumer j is determined by

$$(\rho + \lambda_j + \gamma) Z_j = \max_{a \geq 0} \left\{ -ia + \alpha(1 - \mu_j) \left[\chi^m S^m(a, Z_j) + \chi^d S^d(Z_j) \right] \right\}. \tag{57}$$

Consumers of different types carry different amounts of real balances because they have different market powers. The measure of active consumers of type j is

$$\omega_{1,j} \equiv \frac{\omega_j \lambda_j}{\lambda_j + \gamma + \alpha(q)}, \tag{58}$$

and the measure of all active consumers is $\omega_1 \equiv \sum_{j=1}^{J} \omega_{1,j}$. Market tightness in the goods market is $q = n/\omega_1$. Using that $n = f(\theta)/[\delta + f(\theta)] = q \sum_{j=1}^{J} \omega_{1,j}$, we can express tightness

[19] One can re-interpret the environment as one populated with large households composed of workers and consumers/shoppers. Each consumer is specialized in terms of the goods they purchase. The utility of the household is the sum of the utilities of its members.

in the labor market as a function of tightness in the goods market as

$$\frac{f(\theta)}{\delta + f(\theta)} = \sum_{j=1}^{J} \frac{\lambda_j \omega_j q}{\lambda_j + \gamma + \alpha(q)}. \tag{59}$$

This gives an implicit solution $q = Q(\theta)$. From the free entry condition,

$$(\rho + \delta) \frac{k\theta}{f(\theta)} = (1 - \beta) \left\{ \frac{\alpha\left[Q(\theta)\right]}{Q(\theta)} \sum_{j=1}^{J} \frac{\omega_{1,j}}{\omega_1} \mu_j \left[\chi^m S^m(a_j, Z_j) + \chi^d S^d(Z_j) \right] + x - b \right\} - \beta k\theta. \tag{60}$$

An equilibrium can be reduced to a list, $\{(a_j, Z_j, \omega_{1,j})\}_{j=1}^{J}$, and two positive real numbers, q and θ, that are solutions to (57), (58), (59), and (60).

5.2 Calibration strategy

A unit of time corresponds to one month. We assume the matching functions in both labor and goods markets are Cobb-Douglas. Specifically, the job finding rate is given by $f(\theta) = \Xi\theta^{\eta}$ and the matching rate for consumers in the goods market is given by $\alpha(q) = \Psi q^{\psi}$. The cost of production in bilateral meetings is $\varphi(y) = Gy^g$ and utility is $\upsilon(y) = Hy^h$, for $h \in (0, 1)$.

We set $\gamma = 0$ so that when the desire to consume arrives, it never disappears. To calibrate the set $\{\lambda_j, \omega_j, \mu_j\}_{j \in J}$ we target cross-sectional moments related to households' frequency of purchases, expenditure shares, and markups across major household expenditure categories. We estimate the average frequency of purchases and expenditure shares using the Consumer Expenditure Survey (CEX). We set $J = 15$, which represents expenditure categories ranging from "shelter" and "food-at-home" to "personal care products and services" and "reading". Table 2 lists each category. The CEX reports the fraction of households who spent a positive amount in more narrowly-defined expenditure categories (e.g. prepared flour mixes as a sub-component of food-at-home) in a given time period.[20] We interpret these fractions as the probability a representative household purchases a product in that category in a given time period. We then aggregate to the j-category level (e.g. food-at-home) by computing the average, expenditure-weighted probability for purchasing any product in that category. We use expenditure shares directly from the CEX. To derive a measure of markups for each expenditure category we map each category in the CEX to a corresponding North American Industry Classification System (NAICS) sector. For those expenditure categories that fall within the retail trade sector (NAICS 44), we estimate the markup using gross profit margins

[20] In the CEX, the time period varies between weekly and quarterly across categories, depending on if the dairy or interview survey was used.

reported in the Census' Retail Trade Survey. For those expenditure categories that fall outside retail trade, we use estimated markups in De Loecker et al. (2020) for two-digit NAICS industries (e.g. for CEX category "food away from home" we use the estimated markup for NAICS 72 "accommodation and food services").[21]

In terms of moments in the model, the probability of a representative consumer making at least one purchase in category j over a unit of time (under $\gamma = 0$) is given by

$$\frac{\omega_j \lambda_j}{\lambda_j + \alpha}(1 - e^{-\alpha}) + \frac{\omega_j \alpha}{\lambda_j + \alpha}(1 - e^{-\lambda_j})(1 - e^{-\alpha}).$$

The first term corresponds to the probability a consumer is of type j and active, $\omega_j \lambda_j/(\lambda_j+\alpha)$, times the probability that this consumer leaves the active state in a unit of time (calibrated to one month), $1 - e^{-\alpha}$. Likewise, the second term is the probability that a consumer is of type j and inactive, $\omega_j \alpha/(\lambda_j + \alpha)$, times the joint probability of making two transitions from idle to active, then from active to idle, over a unit of time. Expenditure shares in the model are given by $\omega_j a_j / \sum_{k=1}^{J} \omega_k a_k$. The expression of the markup is given by (7). We report the set of estimated parameters in Table 2.

	Data Moments			**Calibrated Parameters**		
Category	Pr(Purchase) monthly	Expenditure Share (%)	Markup	λ_j	ω_j	μ_j
Shelter	0.74	24.4	1.10	1.29	0.18	0.871
Vehicle purchases and expenses	0.12	16.2	1.24	0.12	0.14	0.996
Healthcare	0.30	10.0	1.44	0.34	0.11	0.998
Food at home	0.99	9.66	1.37	4.24	0.08	0.996
Utilities, fuels, and public services	0.71	7.76	1.24	1.16	0.06	0.920
Entertainment	0.12	6.56	1.30	0.12	0.10	0.999
Food away from home	0.84	5.57	1.12	1.75	0.06	0.998
Household furnishings and equipment	0.07	4.96	1.36	0.07	0.06	0.998
Gasoline and motor oil	0.53	3.95	1.29	0.72	0.03	0.994
Apparel and services	0.16	3.22	1.80	0.17	0.07	0.999
Household operations	0.19	3.01	1.24	0.20	0.04	0.998
Alcoholic beverages	0.24	1.64	1.28	0.26	0.13	0.995
Housekeeping supplies	0.92	1.48	1.91	2.44	0.03	0.999
Personal care products	0.31	1.42	1.44	0.36	0.02	0.998
Reading	0.09	0.21	1.67	0.09	0.01	0.999
Mean (expenditure weighted)	-	-	1.42	0.94	-	0.970

Table 2: Cross sectional moments and parameters

On average, the frequency of consumption opportunities is $\mathbb{E}\lambda_j = 0.94$, which implies the average good is purchased around once per month. Some categories are purchased

[21] In Online Appendix B, we discuss in more details the CEX data and how we map expenditure categories in the CEX to NAICS industries.

frequently, like food at home ($\lambda_j = 4.24$) or housekeeping supplies ($\lambda_j = 2.44$). However, many categories are purchased infrequently, like household furnishings and equipment ($\lambda_j = 0.07$) or reading ($\lambda_j = 0.09$), which includes newspapers, magazines, or books. The largest expenditure categories are shelter – purchased relatively frequently $\lambda_j = 1.29$ – and vehicle purchases and other vehicle expenses – purchased relatively infrequently $\lambda_j = 0.12$. In terms of market power, the heterogeneity in λ_j captures some degree of the cross-sectional variation in markups since a larger λ_j tends to reduce consumers' outside option Z and increase a firm's markup. However, the residual is accounted for by heterogeneity in bargaining power μ_j.[22]

The remaining parameters are fixed using an approach close to that in BMW. We set the discount rate $\rho = 0.001$ so that the real interest rate in the model matches the difference between the rate on Aaa bonds and realized inflation, on average. We use the vacancy posting cost k and job destruction rate δ to match the average unemployment rate and unemployment-to-employment (UE) transition rate. The level parameter of the labor market matching function, Ξ, is normalized so that the vacancy rate is 1. The elasticity of the labor market matching function, η, targets the regression coefficient of labor market tightness θ on the UE rate of 0.6. Firms' bargaining power in the labor market, $1 - \beta$, is set to match the average wage markdown, $1 - \mathbb{E}[w_j/\hat{x}]$, where $\hat{x} = \alpha^s \sum_j \frac{\omega_{1,j}}{\omega_1} \mu_j \left(\chi^m S^m + \chi^d S^d\right) + x$ is average output per worker. Following evidence in Yeh et al. (2022), we target an average wage markdown of 0.35.

Unemployment benefits b represent both unemployment income and the value of non-work. We follow Hall and Milgrom (2008), and set $b = 0.71\hat{x}$.[23] We set the level of the goods market matching function to $\Psi = 1$, and assume the matching rate of firms, α^s, and that of consumers, α^b, have the same elasticity, i.e. $\psi = 0.5$.

To calibrate $v(y)$ and $\varphi(y)$, we first normalize the level of utility to $H = 1$ and set the curvature $h = 0.90$ to target an elasticity of substitution across goods of 10, within the range of estimates provided in Redding and Weinstein (2020).[24] We then calibrate the level and curvature of the cost of production, (G, g), to match the relationship between money demand M/pY and i in the data, using the adjusted M1 series in Lucas and Nicolini (2015)

[22] Notice in Table 2 there does not appear to be much variation in bargaining powers across expenditure categories, which might imply that heterogeneity in λ_j accounts for most of the variation in markups. However, the theoretical mapping from bargaining power μ_j to markups is highly non-linear, so it does not take much variation in μ_j to generate variation in markups.

[23] The effects of inflation on unemployment are channeled through labor productivity. Hence, we aim to capture the extent to which movements in labor productivity affect unemployment, as studied in the large literature following Shimer (2005). Our target for b represents a conservative estimate between the calibrations of Shimer (2005), Hall and Milgrom (2008), and Hagedorn and Manovskii (2008).

[24] Since the utility of a large household is the weighted sum of CRRA functions $v(y_j)$, the elasticity of the ratio of any two consumption goods to the household-level marginal rate of substitution between the same goods is given by $1/(1-h)$.

Parameter	Description	Targets	Value
ρ	Rate of time preference	Average real interest rate	0.001
k	Vacancy cost	Average unemployment rate	0.49
δ	Job destruction rate	Unemployment-to-employment rate	0.03
Ξ	Level of labor market matching	Average vacancies (normalization)	0.08
η	Elasticity of labor market matching	Elasticity of UE rate	0.60
Ψ	Level of goods market matching	—	1.00
ψ	Elasticity of goods market matching	Equal contribution to matching	0.50
β	Bargaining power of worker	Wage markdown	0.03
b	Unemployment benefits	$b = 0.71\hat{x}$	0.40
γ	Desire to consume disappears	—	0
χ^m	Fraction of monetary meetings	Fraction of credit card transactions	0.80
G	Level of production cost $\varphi(y)$	Level of money demand	0.58
g	Elasticity of production cost $\varphi(y)$	Elasticity of money demand	1.28
H	Level of $\upsilon(y)$	Normalization	1.0
h	Curvature of $\upsilon(y)$	Elasticity of substitution across goods	0.90

Table 3: Calibrated parameters

as our measure of money. Finally, we set $\chi^m = 0.8$ because in the Atlanta Fed data discussed by Foster et al. (2013), credit cards account for 23% of purchases in volume. In the Bank of Canada data discussed by Arango and Welte (2012), this number is 19%.

The targets and parameter values discussed above are summarized in Table 3.

5.3 Quantitative results

In the top row of Figure 9, we illustrate how long-run inflation affects the aggregate unemployment rate (solid blue line in top-left panel), real money balances (top-middle panel) and consumers' outside options (top-right panel) across expenditure categories, under the baseline calibration. The Phillips curve is non-monotone in π, following the prediction in Proposition 4 for homogenous consumers. Consumers' outside options, Z_j, decrease with inflation since money holdings become more costly – *the market power effect of inflation*. Even though lower money holdings constrain the payment that consumers can make to firms – *the real balance effect of inflation* – their lower outside options improve firms' market power across goods markets. Markups increase, as illustrated in the bottom-left panel of Figure 9, as well as firms' expected revenue from a filled vacancy.[25]

In the top-left panel of Figure 9, we quantify the real balance and market power effects

[25] The model produces a quantitatively-similar relationship between inflation and firms' markups to that documented during the 2021 inflation surge. For instance, Glover et al. (2023) document that firm-level markups increased by 3.4 percentage points while Personal Consumption Expenditure inflation increased by 2.9 percentage points. For the same increase in inflation, our model predicts that firm markups should increase by 4.6 percentage points.

of inflation by computing counter-factual Phillips curves as responses of unemployment to inflation due to changes only in i) real money balances (dash-dotted red line) or ii) consumers' outside options (dashed green line), respectively. Specifically, we recompute u using (26) and (60) keeping either $\{Z_j\}_j$ or $\{a_j\}_j$ constant. The outside option effect is quantitatively dominant for annual inflation rates between the Friedman rule and -0.2%, the latter representing the unemployment-minimizing inflation rate. As inflation rises beyond -0.2%, the effect on liquidity constraints dominates and the Phillips curve turns upward-sloping.

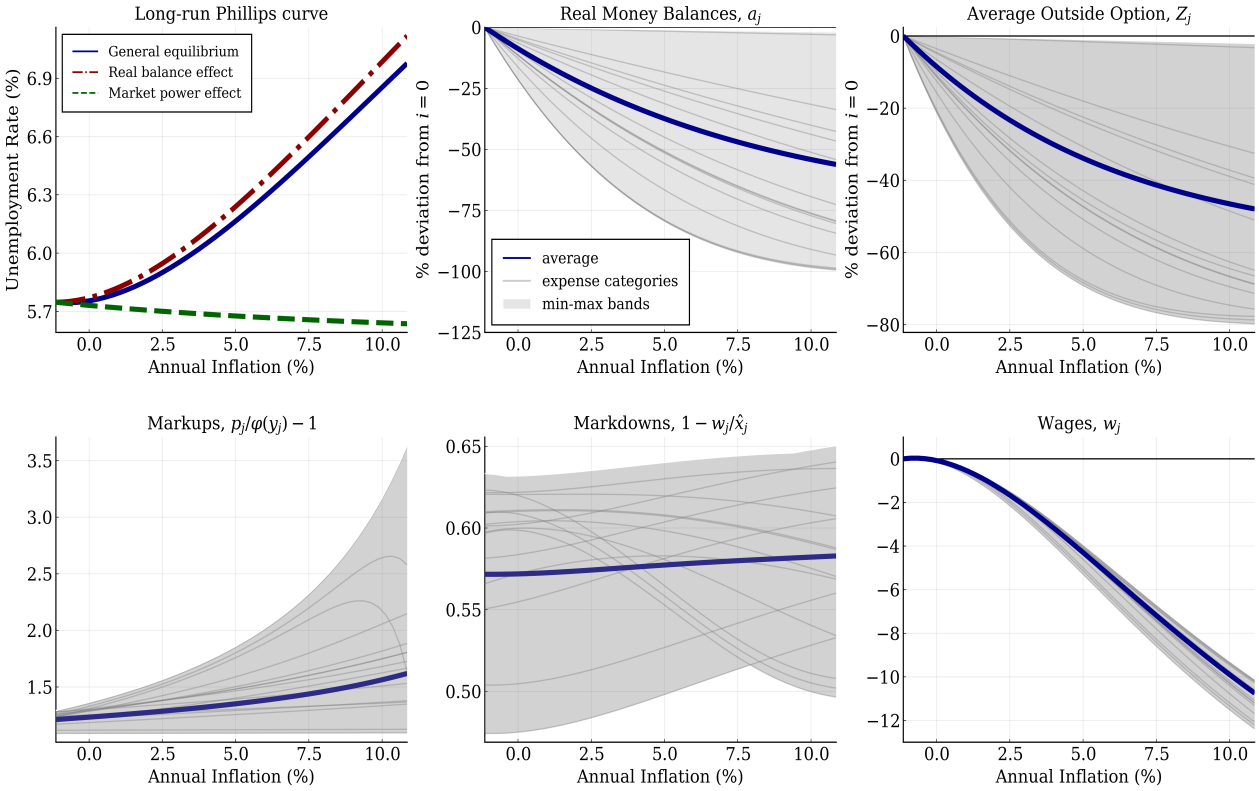


Figure 9: The effects of inflation on unemployment (top-left), real money balances (top-middle), consumers' outside options (top-right), markups (bottom-left), markdowns (bottom-middle), and wages (bottom-right)

There is substantial heterogeneity in the response of outside options and markups to inflation across sectors. The range of responses are illustrated with a gray band in Figure 9 while individual sector responses are shown as gray lines. For instance, in high markup sectors such as "housekeeping supplies" and "apparel and services", increasing inflation from zero to 5% can decrease consumers' outside options by more than 75%, which in turn increases markups from around 25% to 80%. However in low markup sectors, such as "real estate" and "utilities and public services", inflation has negligible effects on outside options or markups.

The bottom-middle and bottom-right panels of Figure 9 illustrate the effects of inflation on wage markdowns and wages, respectively. On average, inflation slightly raises firms' markdowns. However, inflation can have non-monotone effects on markdowns across sectors. There are several factors at work. Higher inflation tends to decrease wages (except for very low inflation rates) since it reduces consumers' payment capacity and, in turn, firms' profits. However, in some sectors, firms' expected revenue $\hat{x}$ can increase with inflation while wages fall. This implies their markdown rises. In other sectors, expected revenue falls faster than wages leading to a reduction in measured wage markdowns.

The strength of the market power effect depends on the rate at which the desire to consume arrives, λ_j. If it occurs relatively infrequently, then consumers' opportunity cost of consumption is relatively large since the value of returning to being idle is relatively low. If, however, consumers spend little time idle, the opportunity cost of consumption is low. In Figure 10, we illustrate the effects of changing λ_j across sectors on the shape of the long-run Phillips curve (left panel) and the unemployment-minimizing level of inflation (right panel). To do so, we shift the distribution of λ_j by a constant and report the effects over $\mathbb{E}\lambda_j$.

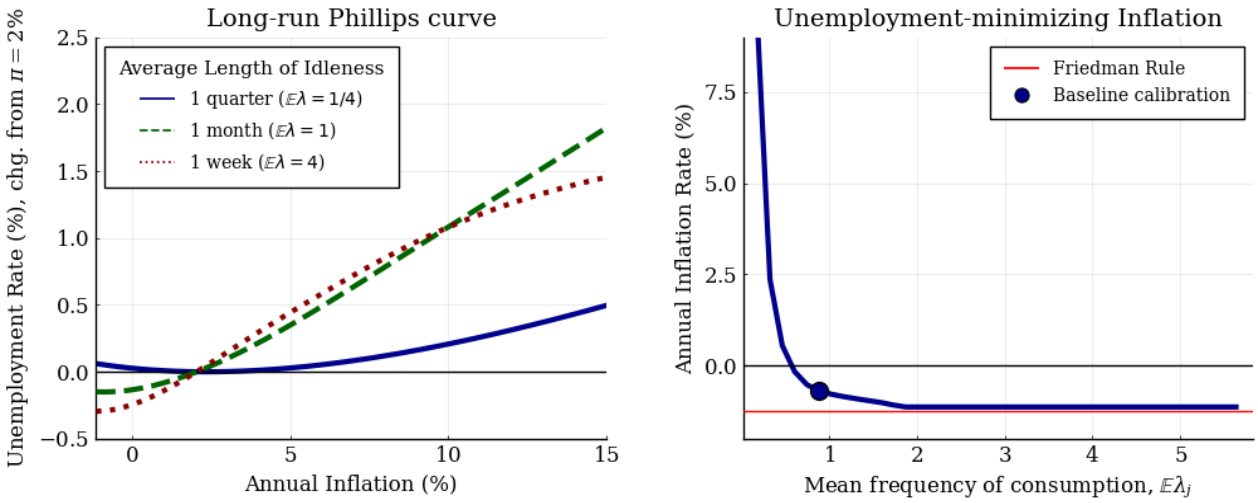


Figure 10: The Long-run Phillips curve and λ (left); Unemployment-minimizing inflation rate and λ (right)

In the left panel of Figure 10, in the y-axis we plot the changes in the unemployment rate relative to the baseline calibration, which corresponds to $\pi = 2\%$. This panel shows that for relatively low values of $\mathbb{E}\lambda_j$ the market power effect is quantitatively strong. For instance, if the desire to consume occurs on average once per quarter, illustrated in the solid-blue curve, the long-run Phillips curve is downward sloping for annual inflation rates up to 4%. The right panel illustrates how the unemployment-minimizing inflation rate changes as $\mathbb{E}\lambda_j$ varies from close to zero to 6. As $\mathbb{E}\lambda_j$ approaches zero, the unemployment-minimizing inflation rate is just

under 9%. As the speed of the desire to consume increases, the unemployment-minimizing inflation rate decreases towards the Friedman Rule, without hitting it.

6 Conclusion

In this paper we made a simple, but robust, observation regarding the long-run trade-off between unemployment and inflation. In the class of models pioneered by BMW, where goods and labor markets are frictional, and money plays an essential role to facilitate the exchange of goods and services, the relation between unemployment and inflation is non-monotone. As a result, the inflation rate that minimizes unemployment is above the one prescribed by the Friedman rule. This result is robust in that it does not require parametric conditions to hold once one introduces consumer search in order to endogenize consumer outside options and firms' market power. We also show that consumers' outside options are crucial for the optimality of the Friedman rule and the convergence of equilibria to a competitive outcome, where rents, markups, and markdowns vanish at the frictionless limit.

References

Alvarez, F., A. Atkeson, and P. J. Kehoe (2002). Money, interest rates, and exchange rates with endogenously segmented markets. *Journal of Political Economy 110*(1), 73–112.

Alvarez, F., R. E. Lucas Jr, and W. E. Weber (2001). Interest rates and inflation. *American Economic Review 91*(2), 219–225.

Anderson, S. P. and R. Renault (1999). Pricing, product diversity, and search costs: A Bertrand-Chamberlin-Diamond model. *RAND Journal of Economics*, 719–735.

Arango, C. and A. Welte (2012). The Bank of Canada's 2009 methods-of-payment survey: Methodology and key results. *Bank of Canada Working Paper*.

Aruoba, S., G. Rocheteau, and C. Waller (2007). Bargaining and the value of money. *Journal of Monetary Economics 54*(8), 2636–2655.

Ascari, G., P. Bonomolo, and Q. Haque (2022). The long-run Phillips curve is... a curve. *Working paper*.

Bajaj, A. and S. Mangin (2023). Consumer choice and the cost of inflation. *Working paper*.

Banerjee, A. and B. Russell (2001, May). The relationship between the markup and inflation in the G7 economies and australia. *Review of Economics and Statistics 83*(2), 377–384.

Banerjee, A. and B. Russell (2005, June). Inflation and measures of the markup. *Journal of Macroeconomics 27*(2), 289–306.

Benabou, R. (1988). Search, price setting and inflation. *Review of Economic Studies 55*(3), 353–376.

Benati, L. (2015). The long-run phillips curve: A structural var investigation. *Journal of Monetary Economics 76*, 15–28.

Berentsen, A., G. Menzio, and R. Wright (2011). Inflation and unemployment in the long run. *American Economic Review 101*(1), 371–98.

Berentsen, A., G. Rocheteau, and S. Shi (2007). Friedman meets Hosios: Efficiency in search models of money. *Economic Journal 117*(516), 174–195.

Bethune, Z., M. Choi, and R. Wright (2020). Frictional goods markets: Theory and applications. *Review of Economic Studies 87*(2), 691–720.

Bethune, Z., G. Rocheteau, and P. Rupert (2015). Aggregate unemployment and household unsecured debt. *Review of Economic Dynamics 18*(1), 77–100.

Branch, W. A., N. Petrosky-Nadeau, and G. Rocheteau (2016). Financial frictions, the housing market, and unemployment. *Journal of Economic Theory 164*, 101–135.

Bullard, J. and J. W. Keating (1995). The long-run relationship between inflation and output in postwar economies. *Journal of Monetary Economics 36*(3), 477–496.

Burdett, K. and K. L. Judd (1983). Equilibrium price dispersion. *Econometrica*, 955–969.

Chirinko, R. and S. Fazzari (2000, August). Market power and inflation. *Review of Economics and Statistics 82*(3), 509–513.

Choi, M. and G. Rocheteau (2021). New Monetarism in continuous time: Methods and applications. *Economic Journal 131*(634), 658–696.

Cooley, T. F. and V. Quadrini (2004). Optimal monetary policy in a Phillips-curve world. *Journal of Economic Theory 118*(2), 174–208.

De Grauwe, P. and C. Costa Storti (2004). The effects of monetary policy: a meta-analysis. *SSRN Working Paper 567102*.

De Loecker, J., J. Eeckhout, and G. Unger (2020). The rise of market power and the macroeconomic implications. *Quarterly Journal of Economics 135*(2), 561–644.

DePillis, L. (2022). Is 'greedflation' rewriting economics, or do old rules still apply? *New York Times*.

Devereux, M. B. and J. Yetman (2002). Menu costs and the long-run output–inflation tradeoff. *Economics Letters 76*(1), 95–100.

Diamond, P. A. (1971). A model of price adjustment. *Journal of Economic Theory 3*(2), 156–168.

Diamond, P. A. (1993). Search, sticky prices, and inflation. *Review of Economic Studies 60*(1), 53–68.

Duffie, D., N. Gârleanu, and L. H. Pedersen (2005). Over-the-counter markets. *Econometrica 73*(6), 1815–1847.

Foster, K., S. Schuh, and H. Zhang (2013). The 2010 survey of consumer payment choice. *Federal Reserve Bank of Boston Research Data Report*.

Friedman, M. (1977). Nobel lecture: Inflation and unemployment. *Journal of Political Economy 85*(3), 451–472.

Gale, D. (1986a). Bargaining and competition part I: Characterization. *Econometrica*, 785–806.

Gale, D. (1986b). Bargaining and competition part II: Existence. *Econometrica*, 807–818.

Gale, D. (1987). Limit theorems for markets with sequential bargaining. *Journal of Economic Theory 43*(1), 20–54.

Galenianos, M. and P. Kircher (2009). Directed search with multiple job applications. *Journal of Economic Theory 144*(2), 445–471.

Glover, A., J. M. del Rio, and A. von Ende-Becker (2023). How much have record corporate profits contributed to recent inflation. *Federal Reserve Bank of Kansas City Economic Review First Quarter*, 23–35.

Gu, C., J. H. Jiang, and L. Wang (2023). Credit condition, inflation and unemployment. *Working paper*.

Hagedorn, M. and I. Manovskii (2008). The cyclical behavior of equilibrium unemployment and vacancies revisited. *American Economic Review 98*(4), 1692–1706.

Hall, R. E. and P. R. Milgrom (2008). The limited influence of unemployment on the wage bargain. *American Economic Review 98*(4), 1653–74.

Hansen, N.-J., F. Toscani, and J. Zhou (2023). Euro area inflation after the pandemic and energy shock: Import prices, profits, and wages. IMF Working Paper 23/131.

Head, A. and A. Kumar (2005). Price dispersion, inflation, and welfare. *International Economic Review 46*(2), 533–572.

Hosios, A. (1990). On the efficiency of matching and related models of search and unemployment. *Review of Economic Studies 57*(2), 279–298.

Hu, T.-W. and G. Rocheteau (2020). Bargaining under liquidity constraints: Unified strategic foundations of the Nash and Kalai solutions. *Journal of Economic Theory 189*.

Julien, B., J. Kennes, and I. King (2008). Bidding for money. *Journal of Economic Theory 142*(1), 196–217.

Kalai, E. (1977). Proportional solutions to bargaining situations: Interpersonal utility comparisons. *Econometrica 45*(7), 1623–1630.

King, R. and A. L. Wolman (1996). Inflation targeting in a St. louis model of the 21st century. Federal Reserve Bank of St. Louis Review, 95.

King, R. G. and M. W. Watson (1994). The post-war US Phillips curve: a revisionist econometric history. In *Carnegie-Rochester Conference Series on Public Policy*, Volume 41, pp. 157–219. Elsevier.

King, R. G. and M. W. Watson (1997). Testing long-run neutrality. *FRB Richmond Economic Quarterly 83*(3), 69–101.

Lagos, R. and R. Wright (2005). A unified framework for monetary theory and policy analysis. *Journal of Political Economy 113*(3), 463–484.

Lauermann, S. (2013). Dynamic matching and bargaining games: A general approach. *American Economic Review 103*(2), 663–689.

Lehmann, E. (2012). A search model of unemployment and inflation. *Scandinavian Journal of Economics 114*(1), 245–266.

Lehmann, E. and B. Van der Linden (2010). Search frictions on product and labor markets: Money in the matching function. *Macroeconomic Dynamics 14*(1), 56–92.

Lucas, R. E. and J. P. Nicolini (2015). On the stability of money demand. *Journal of Monetary Economics 73*, 48–65.

Lucas, R. E. and E. C. Prescott (1974). Equilibrium search and unemployment. *Journal of Economic Theory 7*(2), 188–209.

Makowski, L. and J. M. Ostroy (2001). Perfect competition and the creativity of the market. *Journal of Economic Literature 39*(2), 479–535.

Mangin, S. and B. Julien (2021). Efficiency in search and matching models: A generalized hosios condition. *Journal of Economic Theory 193*, 105208.

McCall, J. J. (1970). Economics of information and job search. *Quarterly Journal of Economics*, 113–126.

Michaillat, P. and E. Saez (2015). Aggregate demand, idle time, and unemployment. *Quarterly Journal of Economics 130*(2), 507–569.

Mortensen, D. T. and C. A. Pissarides (1994). Job creation and job destruction in the theory of unemployment. *Review of Economic Studies 61*(3), 397–415.

Neiss, K. (2001). The markup and inflation: Evidence in OECD countries. *Canadian Journal of Economics 34*(2), 570–587.

Petrosky-Nadeau, N. and E. Wasmer (2015). Macroeconomic dynamics in a model of goods, labor, and credit market frictions. *Journal of Monetary Economics 72*, 97–113.

Petrosky-Nadeau, N. and E. Wasmer (2017). *Labor, Credit, and Goods Markets: The macroeconomics of search and unemployment*. MIT Press.

Pissarides, C. A. (2000). *Equilibrium unemployment theory*. MIT press.

Redding, S. and D. Weinstein (2020). Measuring aggregate price indicies with taste shocks: Theory and evidence for CES preferences. *Quarterly Journal of Economics 135*(1), 503–560.

Rocheteau, G., T.-W. Hu, L. Lebeau, and Y. In (2021). Gradual bargaining in decentralized asset markets. *Review of Economic Dynamics 42*, 72–109.

Rocheteau, G. and A. Rodriguez-Lopez (2014). Liquidity provision, interest rates, and unemployment. *Journal of Monetary Economics 65*, 80–101.

Rocheteau, G., P. Rupert, and R. Wright (2007). Inflation and unemployment in general equilibrium. *Scandinavian Journal of Economics 109*(4), 837–855.

Rocheteau, G. and R. Wright (2005). Money in search equilibrium, in competitive equilibrium, and in competitive search equilibrium. *Econometrica 73*(1), 175–202.

Rocheteau, G., R. Wright, and S. X. Xiao (2018). Open market operations. *Journal of Monetary Economics 98*, 114–128.

Shi, S. (1998). Search for a monetary propagation mechanism. *Journal of Economic Theory 81*(2), 314–352.

Shimer, R. (2005). The cyclical behavior of equilibrium unemployment and vacancies. *American Economic Review 95*(1), 25–49.

Wang, L. (2016). Endogenous search, price dispersion, and welfare. *Journal of Economic Dynamics and Control 73*, 94–117.

Wang, L., R. Wright, and L. Q. Liu (2020). Sticky prices and costly credit. *International Economic Review 61*(1), 37–70.

Williamson, S. D. (2006). Search, limited participation, and monetary policy. *International Economic Review 47*(1), 107–128.

Williamson, S. D. (2012). Liquidity, monetary policy, and the financial crisis: A New Monetarist approach. *American Economic Review 102*(6), 2570–2605.

Williamson, S. D. (2015). Keynesian inefficiency and optimal policy: A New Monetarist approach. *Journal of Money, Credit and Banking 47*(S2), 197–222.

Wolinsky, A. (1986). True monopolistic competition as a result of imperfect information. *Quarterly Journal of Economics 101*(3), 493–511.

Yeh, C., C. Macaluso, and B. Hershbein (2022). Monopsony in the US labor market. *American Economic Review 112*(7), 2099–2138.

A Proofs of propositions and lemmas

Proof of Proposition 2. Equation (36) determines θ. The left side is increasing in θ from 0 to $+\infty$ while the right side is decreasing in θ from $+\infty$ to $(1-\beta)(x-b)$. Hence, θ is unique. The right side of (36) is increasing in $\lambda+\gamma$. Hence, θ increases with λ and γ. It follows that the unemployment rate, $u=\delta/[\delta+f(\theta)]$, decreases with λ and γ. From (32), θ is a decreasing function of Z, so Z decreases with λ and γ. Finally, by (34), w increases in λ and in γ. ■

Proof of Proposition 4. Part 1: Labor market tightness solve $\Gamma(\theta;i)=0$ where

$$\Gamma(\theta;i)\equiv(1-\beta)\left\{\alpha^s(\theta)\mu\left\{\chi^m S^m\left[a^*(\theta;i),Z(\theta;i)\right]+\chi^d S^d\left[Z(\theta;i)\right]\right\}+x-b\right\}-\beta k\theta-(\rho+\delta)\frac{k\theta}{f(\theta)}.$$

When $i=0$,

$$\Gamma(\theta;0)\equiv(1-\beta)\left\{\frac{\alpha^s(\theta)\mu(\rho+\lambda+\gamma)}{\rho+\lambda+\gamma+\alpha^b(\theta)(1-\mu)}\left[\upsilon(y^*)-\varphi(y^*)\right]+x-b\right\}-\beta k\theta-(\rho+\delta)\frac{k\theta}{f(\theta)}. \tag{61}$$

It is monotone decreasing in θ, so the equilibrium at the Friedman rule ($i=0$) is unique.

We now show that $\Gamma(\theta;i)$ is increasing in i in the neighborhood of the Friedman rule. From (8), $\frac{\partial S^m(a,Z)}{\partial a}\equiv\left[\upsilon'(y)-\varphi'(y)\right]\frac{\partial y}{\partial a}$. In the neighborhood of $i=0^+$, $y=y^*$ and $\upsilon'(y)-\varphi'(y)=0$. Hence, $\partial S^m(a,Z)/\partial a=0$. The effect of a change in a^* on the match surplus, induced by an increase in i, is second order when i is close to 0 because the match surplus is maximum. However, from (8)-(9), when y is in the neighborhood of y^*, $\frac{\partial S^m(a,Z)}{\partial Z}=\frac{\partial S^d(Z)}{\partial Z}=-1$. From (30), $\frac{\partial Z}{\partial i}=\frac{-a^*}{\rho+\lambda+\gamma+\alpha^b(\theta)(1-\mu)}$, where, at the Friedman rule, by (12) and (35)

$$a^*=\varphi(y^*)+\frac{(\rho+\lambda+\gamma)\mu}{\rho+\lambda+\gamma+\alpha^b(\theta)(1-\mu)}\left[y^*-\varphi(y^*)\right]>0.$$

Combining these results, $\chi^m S^m\left[a^*(\theta;i),Z(\theta;i)\right]+\chi^d S^d\left[Z(\theta;i)\right]$ is increasing in i, and hence $\Gamma(\theta;i)$ is also increasing in i. It follows that θ, such that $\Gamma(\theta;i)=0$, is increasing in i. The unemployment rate, $u=\delta/[\delta+f(\theta)]$, is decreasing in θ and hence decreasing in i. By the definition of $\Gamma(\theta;i)$ and (31), we can reexpress $\Gamma(\theta;i)$ as

$$\Gamma(\theta;i)=\frac{(1-\beta)}{\beta}(w-b)-k\theta-(\rho+\delta)\frac{k\theta}{f(\theta)}.$$

Since $\Gamma(\theta;i)=0$ in equilibrium, w and θ must comove as i changes. Therefore, w rises in i.

Part 2: We now consider the limiting case $i=+\infty$. Since agents do not carry money, the outcome is similar to a pure credit economy but with $\chi^d<1$. From (29), $a^*=0$. By

the steps leading to (35), $Z = \frac{\alpha^b(\theta)\chi^d(1-\mu)}{\rho+\lambda+\gamma+\alpha^b(\theta)\chi^d(1-\mu)}\left[\upsilon(y^*)-\varphi(y^*)\right]$ and hence

$$\Gamma(\theta;+\infty) \equiv (1-\beta)\left\{\frac{\alpha^s(\theta)\mu\chi_d\left(\rho+\lambda+\gamma\right)}{\rho+\lambda+\gamma+\alpha^b(\theta)\chi^d(1-\mu)}\left[\upsilon(y^*)-\varphi(y^*)\right]+x-b\right\}-\beta k\theta-(\rho+\delta)\frac{k\theta}{f(\theta)}. \tag{62}$$

From (61) and (62), $\Gamma(\theta;+\infty) < \Gamma(\theta;0)$ for all $\theta > 0$. So θ_0 solution to $\Gamma(\theta;0) = 0$ is larger than $\theta_{+\infty}$ solution to $\Gamma(\theta;+\infty) = 0$. Hence, the unemployment rate when $i = 0$ is lower than the unemployment rate when $i = +\infty$. Since w and θ comove, w is lower when $i = +\infty$ than when $i = 0$. Finally, note that there is a finite upper bound for i above which a monetary equilibrium does not exist. This upper bound is $i = \chi^m(1-\mu)\alpha^b(\theta)/\mu$ where θ is bounded above by the market tightness of a pure credit economy with $Z = 0$. This upper bound of θ is finite and it solves $(1-\beta)\left\{\alpha^s(\theta)\mu\left[\upsilon(y^*)-\varphi(y^*)\right]+x-b\right\}-\beta k\theta-(\rho+\delta)\frac{k\theta}{f(\theta)}=0$. ■

Proof of Proposition 6. From (27), $\frac{q}{\bar{\alpha}(q)} \to \frac{Af(\theta)}{\lambda\omega[\delta+f(\theta)]} = +\infty$ for all $\theta > 0$. So, if $\theta > 0$, $q \to +\infty$ and the firm matching rate with a consumer tends to $A\bar{\alpha}(q)/q \to \lambda\omega\left[\delta+f(\theta)\right]/f(\theta)$. Since $A\bar{\alpha}(q) \to +\infty$, from (36), $Z \to \upsilon(y^*)-\varphi(y^*)$. Using that $S^m(a^*,Z) \to 0$ and $S^d(Z) \to 0$, from (28), θ solves $(\rho+\delta)\frac{k\theta}{f(\theta)} = (1-\beta)(x-b)-\beta k\theta$. It is consistent with $\theta > 0$ iff $x > b$. Hence, if $x \leq b$ then $\theta \to 0$ as $A \to +\infty$. Suppose $q \to q_\infty > 0$. Then, $\alpha^b = A\bar{\alpha}(q) \to +\infty$ and, from (30), $Z \to \upsilon(y^*)-\varphi(y^*)$. If $q_\infty = 0$, then $A\bar{\alpha}(q)/q \to +\infty$ and, from (28), $\theta = 0$ implies $Z \to \upsilon(y^*)-\varphi(y^*)$. ■

Proof of Proposition 8. <u>*Part 1:*</u> When $\lambda = +\infty$, by (28)-(31), $\omega_{1,t} = \omega$, $Z = 0$, and a steady state equilibrium can be reduced to a 3-tuple, (θ, a, w), that is a solution to

$$\begin{aligned}
(\rho+\delta)\frac{k\theta}{f(\theta)} &= (1-\beta)\left\{\alpha^s(\theta)\mu\left[\chi^m S^m(a,0)+\chi^d S^d(0)\right]+x-b\right\}-\beta k\theta, \qquad (63)\\
a &\in \arg\max_{\hat{a}\geq 0}\left\{-ia+\alpha^b(\theta)(1-\mu)\chi^m S^m(a,0)\right\}, \text{ and}\\
w &= \beta\left\{\alpha^s(\theta)\mu\left[\chi^m S^m(a,0)+\chi^d S^d(0)\right]+x\right\}+(1-\beta)b+\beta k\theta.
\end{aligned}$$

By (50), the instantaneous surplus becomes

$$\begin{aligned}
\varpi &= \omega\alpha\left(\frac{n}{\omega}\right)\left\{\chi^m\left[\upsilon(y_m)-\varphi(y_m)\right]+\chi^d\left[\upsilon(y_d)-\varphi(y_d)\right]\right\} \qquad (64)\\
&\quad +n(x-b)+b-(1-n)\theta k.
\end{aligned}$$

The planner maximizes ϖ by choosing the policy rate i, subject to the law of motion of workers (26) and the firms' free entry condition (63).

We first eliminate y^m in ϖ with the free-entry condition. By (63),

$$\alpha^s(\theta)\left[\chi^m S^m(a,0)+\chi^d S^d(0)\right] = \frac{1}{\mu(1-\beta)}\left[(\rho+\delta)\frac{k\theta}{f(\theta)}+\beta k\theta\right]-\frac{x-b}{\mu}. \tag{65}$$

Since $S^m(a,0) = \upsilon(y_m)-\varphi(y_m)$ and $S^d(0) = y_d-\varphi(y_d)$, we can eliminate the trade surplus

in the planner's objective function in (64) by (65). Also, we can eliminate n by using $n = f(\theta)/[\delta + f(\theta)]$. Therefore, we can rewrite the planner's problem as $\max_\theta \varpi(\theta)$, where

$$\varpi(\theta) \equiv \frac{f(\theta)}{\delta + f(\theta)} \left[\frac{k\theta}{\mu(1-\beta)} \left(\frac{\rho + \delta[1-\mu(1-\beta)]}{f(\theta)} + \beta \right) - \left(\frac{1}{\mu} - 1 \right) (x-b) \right] + b.$$

The expression in the square bracket rises in θ. When the square bracket is positive, the right side strictly rises in θ. Therefore, the maximizer of $\varpi(\theta)$, that we call $\bar{\theta}$, is the largest element in the set of feasible market tightness, provided that $\varpi(\bar{\theta}) > b$. Since the free-entry condition (63) defines a positive relationship between θ and y_m, the maximum tightness, $\bar{\theta}$, is achieved if $y_m = y_d = y^*$, which corresponds to the Friedman rule.

Now we argue $\varpi(\bar{\theta}) > b$. Let θ_0 be the value of θ such that the right side of (65) is 0. At $\theta = \theta_0$, by (64),

$$\varpi(\theta_0) = \frac{\theta k}{\delta + f(\theta)} \left(\frac{\rho + \beta[\delta + f(\theta)]}{1-\beta} \right) + b.$$

Since $\varpi(\theta_0) > b$, $\varpi(\theta)$ rises in θ for all $\theta \geq \theta_0$. Since $\bar{\theta} > \theta_0$, $\varpi(\bar{\theta}) > b$.

Part 2: When $\lambda < +\infty$, we can rewrite (50) by replacing $\omega_1 = \lambda/[\alpha(\theta) + \gamma + \lambda]$ and $n = f(\theta)/[\delta + f(\theta)]$:

$$\begin{aligned} \varpi(\theta, y_m, y_d) &= \frac{\lambda\alpha(\theta)}{\alpha(\theta) + \gamma + \lambda} \left\{ \chi^m \left[y_m - \varphi(y_m) \right] + \chi^d \left[y_d - \varphi(y_d) \right] \right\} \\ &\quad + \frac{f(\theta)}{f(\theta) + \delta}(x-b) + b - \frac{\delta\theta}{\delta + f(\theta)} k. \end{aligned}$$

At the Friedman rule, an increase in i reduces y_m and Z, but the change in y^m only has a second-order effect on firms' profits and market tightness. The only first-order effect is due to the drop in Z. As Z falls, firms get more profits from trade and thus θ rises. Hence, a local derivation from the Friedman rule exists when the derivative of $\varpi(\theta, y_m, y_d)$ with respect to θ is positive at the Friedman rule, fixing $y_m = y_d = y^*$, i.e.

$$\frac{\partial \varpi(\theta, y^*, y^*)}{\partial \theta} \propto \epsilon_f(\theta) \left[\frac{\alpha^s(\theta) S(y^*, 0)(\gamma + \lambda)}{\gamma + \lambda + \alpha(\theta)(1 - \epsilon_\alpha)} \epsilon_\alpha(\theta) + (x-b) \right] - \left(\frac{1}{n} - \epsilon_f(\theta) \right) k\theta, \qquad (66)$$

is strictly positive, and $\propto$ means the left and right side have the same sign. The right side of (66) is equivalent to that of (72) when $\rho = 0$. Hence the right side of (66) falls in θ. Since we have assumed $f'(0) = +\infty$, the right side explodes as $\theta \to 0$. Thus the Friedman rule is suboptimal if and only if θ is sufficiently small.

The labor tightness, θ, at the Friedman rule solves (56). There is a unique solution of θ because the right side of (56) falls and the left side rises in θ. It is easy to check that the solution of θ rises in μ and falls in β. Given μ, if $\beta \to 1$, then $\theta \to 0$. Hence, there exists a $\bar{\beta}(\mu) < 1$ such that the Friedman rule is locally suboptimal if and only if $\beta > \bar{\beta}(\mu)$. Moreover, $\bar{\beta}(\mu)$ rises in μ because the market tightness θ at the FR rises in μ by (56). ■

Online Appendix to "Inflation and Unemployment in the Long Run Revisited"

A Omitted proofs

Proof of Proposition 1.

Define $Z(\theta)$, the solution to (30), as a function of θ. To see that $Z(\theta)$ exists and is unique, note that the left hand side of (30) is linear, increasing in Z, while the right hand side (RHS) is decreasing for all $Z \in (0, \upsilon(y^*) - \varphi(y^*))$. If $Z = 0$, then $RHS \geq \alpha^b(\theta)(1-\mu)\chi^d [\upsilon(y^*) - \varphi(y^*)] > 0$ if $\theta > 0$. If $Z = \upsilon(y^*) - \varphi(y^*)$, then $RHS = 0$. Hence, for all $\theta > 0$, there is a unique $Z \in (0, \upsilon(y^*) - \varphi(y^*))$ solving (30). Since $\alpha^b(\theta)$ is increasing in θ, it follows that RHS is increasing in θ, and hence $Z'(\theta) > 0$. Moreover, $\alpha^b(0) = 0$ implies $Z(0) = 0$.

We have seen that the set of solutions to (11) is either the corner solution 0, the interior solution given by (12), which we denote by $a_I^*(\theta)$, or both. Note that $a_I^*(\theta)$ is increasing continuously in θ, meaning the best response transitions from 0 to $a_I^*(\theta)$ as θ rises, and not the other way around. Hence, $S^m [a^*(\theta), Z(\theta)]$ can only jump upward as θ rises. We now define the following function:

$$\Gamma(\theta) \equiv \left\{(1-\beta)\left\{\alpha^s(\theta)\mu\left\{\chi^m S^m [a^*(\theta), Z(\theta)] + \chi^d S^d [Z(\theta)]\right\} + x - b\right\}\right. \\ \left. -\beta k\theta - (\rho+\delta)\frac{k\theta}{f(\theta)}\right\}. \tag{67}$$

When both 0 and $a_I^*(\theta)$ are optimal for consumers, we assume $a^*(\theta) = a_I^*(\theta)$, and hence $\Gamma(\theta)$ is right continuous. An equilibrium can be reduced to a θ solution to $\Gamma(\theta) = 0$. Since $S^m [a^*(\theta), Z(\theta)]$ only jumps upward, so does $\Gamma(\theta)$. As $\theta \to 0$, $a^*(\theta) \to 0$, so $\Gamma(\theta)$ converges to the value

$$(1-\beta)\left\{\alpha^s(0)\mu\chi^d [\upsilon(y^*) - \varphi(y^*)] + x - b\right\}.$$

Since $\alpha^s(0) = +\infty$, $\Gamma(0) = +\infty$. As $\theta \to +\infty$, $\alpha^s(\theta) \to 0$, $\theta/f(\theta) \to 1/f'(+\infty) = +\infty$ and hence $\Gamma(\theta) \to -\infty$. Since $\Gamma(\theta)$ can only jump upward, it must cut the x-axis and so a steady-state equilibrium exists. ■

Proof of Proposition 5 . From the assumption $\lambda = +\infty$, $Z = 0$. Then from (28) and $\chi^d, \mu > 0$, $\theta \to +\infty$ as $A \to +\infty$. When $\lambda = +\infty$, buyers' money holding decision is given by (39) even when $\chi^d > 0$. Then from (39), for all $\theta > 0$, as $A \to +\infty$, $y^m \to y^*$. Hence

surpluses in all pairwise meetings are equal to $\upsilon(y^*) - \varphi(y^*)$. We now turn to labor market tightness and matching rates. As $A \to +\infty$. From (37),

$$\alpha^s(\theta) = A\frac{\alpha\left[q(\theta)\right]}{q(\theta)}, \text{ where } q(\theta) = \frac{f(\theta)}{\omega\left[\delta + f(\theta)\right]}.$$

Since $f(\theta) \to +\infty$, tightness in the goods market, q, tends to $1/\omega$. Since $\alpha\left[q(\theta)\right]/q(\theta) \to \omega\alpha(\omega^{-1})$, from the expression above, $\alpha^s(\theta) \to +\infty$. The limit for the markup follows directly from (7). ■

Proof of Proposition 7. The current-value Hamiltonian is

$$\begin{aligned}\mathcal{H} \equiv\ & \omega_1\alpha\left(\frac{n}{\omega_1}\right)\left\{\chi^m\left[\upsilon(y_m) - \varphi(y_m)\right] + \chi^d\left[\upsilon(y_d) - \varphi(y_d)\right]\right\} + n(x-b) + b - (1-n)\theta k \\ & +\xi\left\{f(\theta)(1-n) - \delta n\right\} \\ & +\zeta\left\{\lambda(\omega - \omega_1) - \left[\alpha\left(\frac{n}{\omega_1}\right) + \gamma\right]\omega_1\right\}.\end{aligned}$$

From Pontryagin's Maximum Principle,

$$(\theta_t, y_{m,t}, y_{d,t}) \in \arg\max \mathcal{H}(\theta_t, y_{m,t}, y_{d,t}, n_t, \omega_{1,t}) \text{ for all } t.$$

Hence, θ_t solves (51) and $y_{m,t} = y_{d,t} = y^*$ for all t. The law of motion for the co-state variable, ξ_t, is given by $\rho\xi_t = \partial\mathcal{H}_t/\partial n_t + \dot{\xi}_t$, i.e.,

$$\rho\xi_t = \alpha'\left(q_t\right)\left[\upsilon(y^*) - \varphi(y^*) - \zeta_t\right] + x - b + \theta_t k - \left[f(\theta_t) + \delta\right]\xi_t + \dot{\xi}_t. \tag{68}$$

Similarly, the law of motion for ζ_t is given by $\rho\zeta_t = \partial\mathcal{H}_t/\partial\omega_{1,t} + \dot{\zeta}_t$, i.e.,

$$\left(\rho + \lambda + \gamma\right)\zeta_t = \left[\alpha\left(q_t\right) - q_t\alpha'\left(q_t\right)\right]\left[\upsilon(y^*) - \varphi(y^*) - \zeta_t\right] + \dot{\zeta}_t. \tag{69}$$

Combining (51) and (69), the optimality condition for labor market tightness can be rewritten as

$$\begin{aligned}\left(\rho + \delta\right)\frac{k\theta_t}{f(\theta_t)} =\ & \epsilon_f(\theta_t)\left\{\frac{\alpha\left(q_t\right)}{q_t}\epsilon_\alpha(q_t)\left[\upsilon(y^*) - \varphi(y^*) - \zeta_t\right] + x - b\right\} \\ & - \left[1 - \epsilon_f(\theta_t)\right]k\theta_t - \frac{f''(\theta_t)}{f(\theta_t)f'(\theta_t)}k\theta_t\dot{\theta}_t.\end{aligned} \tag{70}$$

Using the definition of ϵ_α, and the law of motion for ζ_t, (69) can be rewritten as

$$\left(\rho + \lambda + \gamma\right)\zeta_t = \alpha\left(q_t\right)\left[1 - \epsilon_\alpha\left(q_t\right)\right]\left[\upsilon(y^*) - \varphi(y^*) - \zeta_t\right] + \dot{\zeta}_t. \tag{71}$$

From (70) and (71), a constant solution to the planner's problem is a pair (θ, ζ) that is a solution to (52)-(53), provided that $n_0 = n_s \equiv f(\theta)/[f(\theta) + \delta]$ and $\omega_{1,0} = \omega_1^s \equiv \lambda/[\alpha(\theta) + \gamma + \lambda]$.

To show that the stationary solution is unique, we eliminate ζ from (53) by using (52). Then, θ solves

$$(\rho + \delta)\frac{k\theta}{f(\theta)} = \epsilon_f(\theta)\left\{\frac{\alpha(q)}{q}\epsilon_\alpha(q)\frac{\rho + \lambda + \gamma}{\rho + \lambda + \gamma + \alpha(q)(1 - \epsilon_\alpha)}\left[\upsilon(y^*) - \varphi(y^*)\right] + x - b\right\} - \left[1 - \epsilon_f(\theta)\right]k\theta. \tag{72}$$

By (27), q falls in θ. Since $\alpha'', f'' < 0$, the elasticities ϵ_α and ϵ_f decrease in θ. Therefore, the right hand side of (72) falls and the left hand side rises in θ. Hence, the stationary solution is unique.

We now show that the constant solution is a solution to the planner's problem by invoking the Mangasarian sufficiency condition. First, the current-value Hamiltonian is jointly concave in (ι, n, ω_1), where $\iota \equiv (1-n)\theta$. To see this, note that $\omega_1\alpha\left(n/\omega_1\right)$ is jointly concave in (ω_1, n). Also, $f(\theta)(1-n) = f(\iota/u)u$ is jointly concave in (ι, u). Finally, the constant solution satisfies the following transversality conditions,

$$\begin{aligned}\lim_{t\to+\infty} e^{-\rho t}\xi_t n_t &= 0 \text{ and} \\ \lim_{t\to+\infty} e^{-\rho t}\zeta_t \omega_{1,t} &= 0.\end{aligned}$$

We now compare the optimality conditions with the equilibrium conditions. Since the planner requires $y_{m,t} = y^*$, monetary policy must implement the Friedman rule, i.e., $i_t = 0$. From (14), Z at the Friedman rule solves

$$(\rho + \lambda + \gamma)Z = \alpha(1 - \mu)\left[\upsilon(y^*) - \varphi(y^*) - Z\right]. \tag{73}$$

From (22), θ at the Friedman rule solves

$$(\rho + \delta)\frac{k\theta}{f(\theta)} = (1 - \beta)\left\{\frac{\alpha(q)}{q}\mu\left[\upsilon(y^*) - \varphi(y^*) - Z\right] + x - b\right\} - \beta k\theta. \tag{74}$$

The comparison of the planner's optimality conditions, (52)-(53), and the equilibrium conditions (73)-(74), show that they coincide if the Hosios conditions in the goods and labor markets, (54)-(55), hold. ■

B Further details on calibration

Estimation Procedure for the Frequency of Purchases in the Consumer Expenditure Survey (CEX) Consumer Expenditure Survey (CEX) Table R-1 reports the fraction of households who spent a positive amount in detailed expenditure categories either in a given week (in the diary survey) or in a given quarter (in the interview survey). We use this measure of the extensive margin of spending to estimate the probability of a purchase by a representative household over a month for a good or service in one of $J = 15$ major expenditure categories (reported in Table 2). To form the 15 categories, we predominately use CEX level-2 categories. We combine "Alcoholic beverages" (UCC code $ALCBEVG$) and "Tobacco products and smoking supplies" (UCC code $TOBACCO$) to form our category of "Alcoholic beverages and tobacco products". We also combine "Vehicle purchases (UCC code $VEHPURCH$), "Other vehicle expenses" (UCC code $VEHOTHXP$), and "Public transportation" (UCC code $PUBTRANS$). Finally, we drop some expenditure categories that our model does not capture well, including "Education" (UCC code $EDUCATN$), "Cash contributions" (UCC code $CASHCONT$), and "Personal insurance and pensions" (UCC code $INSPENSN$).

To compute the probability of a purchase in category $j \in \{1, \ldots, J\}$, we start with the reported fraction of households who spent a positive amount in the most narrowly-defined CEX expenditure categories (e.g. "prepared flour mixes" as a subcomponent of "food at home").[26] Let $\pi^f_{k,j} \in [0, 1]$ stand for this fraction, where $f \in \{w, q\}$ stands for weekly or quarterly frequency and $k \in K_j$, where K_j stands for the set of all narrowly-defined expenditure categories within one of the major categories $j \in \{1, \ldots, J\}$. We map weekly probabilities into monthly using $\pi^m_{k,j} = 1-(1-\pi^w_{k,j})^4$ and quarterly probabilities into monthly using $\pi^m_{k,j} = 1-(1-\pi^q_{k,j})^{(1/3)}$. We then compute the probability of a purchase in a major expenditure category $j \in \{1, \ldots, J\}$ as an expenditure-weighted average probability, $\pi^m_j = \sum_{k \in K_j} \omega_{k,j} \pi^m_{k,j}$, where $\omega_{k,j}$ is the expenditure share of sub-category $k \in K_j$ and $\sum_{k \in K_j} \omega_{k,j} = 1$. The second column of Table 2 reports π^m_j.

Markup imputation We use various sources to obtain estimates of the markup for different spending categories. For most categories, we use estimates of the retail gross margin as a percentage of sales from the 1993-2021 Retail Trade Survey reported by the U.S. Census. The Retail Trade Survey reports gross margins for industries comprising NAICS code 44. Since the CEX does not use the NAICS classification system, we map each NAICS sub-industry in the Retail Trade Survey to one of our 15 major CEX expenditure categories. For instance, "vehicle purchases (net outlay) and other vehicle expenses" corresponds to NAICS

[26] The CEX does not directly report the frequency of purchases for level-2 expenditure categories.

code 441 “Motor vehicle and parts dealers". The CEX category “household furnishings and equipment" corresponds to NAICS codes 442 “Furniture and home furnishings stores", 443 “Electronics and appliance stores", and 444 “Building material, garden equipment, and supplies dealers". There are five categories that do not align well within retail trade. For these, we match them with corresponding NAICS industries and estimated markups in De Loecker et al. (2020). We use the estimated markup for NAICS 53 “Real Estate and Rental and Leasing" for the CEX category “Shelter", the markup for NAICS 22 “Utilities" for the CEX category “Utilities fuels and public services", the markup for NAICS 71 “Arts, Entertainment, and Recreation" for the CEX category “Entertainment", the markup for NAICS 72 “Accommodation and Food Services" for the CEX category “Food away from home", and the markup for NAICS 81 “Other Services" for the CEX category “Household operations".

Calibration with $\lambda_j = \lambda$ and $\mu_j = \mu$ For some of our quantitative results, we use the baseline economy, without heterogeneity in λ and μ. We calibrate this economy using the same procedure outlined in Section 5.2 where we calibrate (λ, μ) to target the average probability of purchasing a good/service across categories of 0.94 and the average markup of 1.42.

C Nominal interest rates and unemployment

So far, monetary policy took the form of a constant money growth rate, π, that could be mapped one-to-one into the nominal interest rate of an illiquid bond, $i = \rho + \pi$. As a result, the relationship between u and π was isomorphic to the one between u and i. In its current practice, however, the central bank sets the nominal interest rate on highly liquid, short-term bonds. Hence, in this section, we revisit the relationship between unemployment and nominal interest rates by distinguishing between illiquid and liquid bonds. We determine how changes in the short-term nominal interest rate affect firms' market power and unemployment, and we characterize the combination of two policy instruments, the money growth rate and the short-term nominal interest rate, that minimize unemployment.

C.1 Liquid bonds and market segmentation

Liquid government bonds are of the pure discount (or zero coupon) variety. A bond pays one unit of numéraire when it matures at Poisson rate $\sigma > 0$. So, $1/\sigma$ is the expected maturity of the bond. The real interest rate of the bond is denoted r_g and its price in terms of the numéraire is ϕ_g. Hence, by standard asset pricing,

$$r_g \phi_g = \sigma(1 - \phi_g) \Rightarrow \phi_g = \frac{\sigma}{r_g + \sigma}.$$

We consider the limit as the maturity of the bond is close to 0, i.e., $\sigma \to +\infty$, which implies $\phi_g = 1$. The nominal interest rate on liquid bonds is denoted $i_g = r_g + \pi$. The total supply of bonds is denoted A_g. Since $\phi_g = 1$, A_g is also the real value of the bond supply.

According to the rate-of-return dominance puzzle, if bonds are perfectly liquid, then $r_g = -\pi$ and $i_g = 0$, i.e., bonds do not bear interest. We address this puzzle by assuming limited participation in the bonds market. We divide consumers into two types, $\kappa \in \{1, 2\}$, where types differ in terms of the assets – money or bonds – they can hold in their portfolios. A fraction ψ_1 of consumers, those of type $\kappa = 1$, can only use money to finance the consumption of good y. The remaining fraction, $\psi_2 = 1 - \psi_1$, corresponds to type-2 consumers who can use both bonds and money as means of payment. One can think of type-1 consumers as unsophisticated investors who do not participate in the bonds market while type-2 consumers are sophisticated investors who can hold financial assets.[27]

[27] One could endogenize participation in the bond market, and hence the measure ψ_κ, by introducing a distribution of participation costs across buyers. See, e.g., Rocheteau et al. (2018).

C.2 Definition of equilibrium

Let a_g be a consumer's holding of bonds measured in numéraire and a_m her real balances. The total liquid wealth of an agent is denoted $a \equiv a_m + a_g$. As before, in a fraction χ^d of meetings, buyers have access to credit. In the remaining fraction, $\chi^\ell \equiv 1 - \chi^d$, buyers can use their liquid assets (money for type-1 buyers, and money and bonds for type-2 buyers). The value function of a buyer of type $\kappa \in \{1, 2\}$ is denoted $V^b_\kappa(a) = a + V^b_\kappa$.

The decision problem of type-1 consumers is the same as that in the baseline model. Their real balances are denoted a^1_m, and their real bond holdings are $a^1_g = 0$. The HJB equation for a type-2 consumer in a steady state is

$$\rho V^b_2 = \max_{a_m, a_g \geq 0} \left\{ -s_m a_m - s_g a_g + \tau + \alpha(1-\mu) \left[\chi^\ell S^\ell(a, Z_2) + \chi^d S^d(Z_2) \right] - \gamma(V^b_2 - W^b_2) \right\}, \tag{75}$$

where $s_g \equiv \rho - r_g$ is the spread between the real interest rate on an illiquid bond and the real interest rate on a liquid government bond, i.e., it is the opportunity cost of holding liquid bonds. Similarly, the spread between the real interest rate on an illiquid bond and the real interest rate on money (r_m) is denoted $s_m \equiv \rho - r_m$, and it is equal to $i = \rho + \pi$. By the logic leading to (14), the outside option of the consumer, Z_2, solves

$$(\rho + \lambda + \gamma) Z_2 = \max_{a_m, a_g \geq 0} \left\{ -s_m a_m - s_g a_g + \alpha(1-\mu) \left[\chi^\ell S^\ell(a, Z_2) + \chi^d S^d(Z_2) \right] \right\}, \tag{76}$$

which defines a negative relationship between s_g and Z_2. The first-order conditions with respect to a_j, $j \in \{m, g\}$, are

$$s_j \geq \frac{\alpha \chi^\ell (1-\mu) \left[\upsilon'(y_2) - \varphi'(y_2) \right]}{\mu \upsilon'(y_2) + (1-\mu)\varphi'(y_2)} \qquad \text{“} = \text{” if } a_j > 0, \text{ for } j \in \{m, g\},$$

where $y_2 = y(a, Z_2)$ is defined by (6). For type-2 consumers, money and bonds are perfect substitutes as means of payment. Hence, if $s_g < s_m$, they go cashless. They hold both money and bonds only if $s_g = s_m$, i.e., bonds do not bear interest.

By market clearing, bonds must be held, and therefore $s_g \leq s_m$. Hence, the relevant holding cost of liquid assets for type-2 consumers is s_g. It follows from (76) that we can rewrite the outside options of type-1 and type-2 consumers, Z_1 and Z_2, as solutions to

$$(\rho + \lambda + \gamma) Z_\kappa = \max_{a \geq 0} \left\{ -s_j a + \alpha(1-\mu) \left[\chi^\ell S^\ell(a, Z_\kappa) + \chi^d S^d(Z_\kappa) \right] \right\}, \ \ \kappa \in \{1, 2\}. \tag{77}$$

The only difference between the outside options of type-1 and type-2 consumers is the holding cost of liquid assets — type-1 faces s_m while type-2 faces s_g. Sophisticated consumers have

better outside options than unsophisticated ones because they can hold liquidity at a lower cost, thereby reducing the cost of searching for an alternative seller.

In a steady-state equilibrium, the market-clearing condition of the bonds market implies

$$A_g = \psi_2 \omega_1 a_g^2,$$

where $\psi_2 \omega_1$ is the measure of active type-2 consumers and a_g^2 represents their bond holdings. By fixing the supply of bonds, A_g, the government can control r_g, and thus s_g. The real balances held by an active type-2 consumer are denoted a_m^2. Hence, aggregate real balances are

$$\phi_m M = \omega_1(\psi_1 a_m^1 + \psi_2 a_m^2).$$

By the same logic as before, market tightness in the labor market solves

$$(\rho + \delta)\frac{k\theta}{f(\theta)} = (1-\beta)\left\{\alpha^s \mu \sum_{\kappa \in \{1,2\}} \psi_\kappa \left[\chi^\ell S^\ell(a^\kappa, Z_\kappa) + \chi^d S^d(Z_\kappa)\right] + x - b\right\} - \beta k\theta, \quad (78)$$

where $a^\kappa = a_m^\kappa + a_g^\kappa$ is a type-κ consumer's choice of liquidity. The equilibrium wage solves

$$w = \beta\left\{\alpha^s \mu \sum_{\kappa \in \{1,2\}} \psi_\kappa \left[\chi^\ell S^\ell(a^\kappa, Z_\kappa) + \chi^d S^d(Z_\kappa)\right] + x\right\} + (1-\beta)b + \beta k\theta. \quad (79)$$

An equilibrium is a tuple, $(\theta, a_m^1, a_m^2, a_g^2, Z_1, Z_2, w)$, where θ solves (78), $\left(a_m^1, a_m^2, a_g^2\right)$ and Z_κ solve (77) for $\kappa \in \{1,2\}$, and w solves (79). By the logic used in the proof of Proposition 1, there exists an active steady-state equilibrium provided that $\chi^d > 0$.

C.3 Typology of equilibria

We distinguish two types of equilibria. There is a *liquidity trap* equilibrium where $s_g = s_m = i$, i.e., the nominal interest on liquid bonds is $i_g = 0$.[28] In such equilibria, type-2 consumers hold a portfolio of money and bonds as they are indifferent between the two assets. Both types of consumers have the same lifetime utility and the same outside options. Liquidity-trap equilibria are isomorphic to the monetary equilibria studied in Section 3.

A decrease in A_g reduces the bond holdings of type-2 consumers but it has no impact on the value of their portfolio, a^2. Buyers compensate the decrease in a_g^2 by raising a_m^2 so as to keep their liquid wealth constant. Therefore, changes in A_g do not affect production and

[28] Williamson (2012) provides another example of a New Monetarist model that delivers a liquidity trap equilibrium.

unemployment. A liquidity trap equilibrium occurs when

$$A_g \leq a(i)\psi_2\omega_1,$$

where $a(i)$ is the choice of liquid wealth of a consumer when the user cost of liquidity is i. So a liquidity trap equilibrium exists when the supply of liquid bonds is lower than a threshold.

If $A_g > a(i)\psi_2\omega_1$, then the economy is outside of the liquidity trap and $s_g < s_m$, i.e. $i_g > 0$. In that case, type-2 consumers only hold interest-bearing liquid bonds. By market clearing and bargaining,

$$a_g^2 = \frac{A_g}{\psi_2\omega_1} = (1-\mu)\varphi(y_2) + \mu\upsilon(y_2) - \mu Z_2, \quad \text{where } \frac{\alpha\chi^\ell(1-\mu)\left[\upsilon'(y_2) - \varphi'(y_2)\right]}{\mu\upsilon'(y_2) + (1-\mu)\varphi'(y_2)} = s_g. \quad (80)$$

Note that y_2 is a function of s_g by the first-order condition, and Z_2 is also a function of s_g. So, the supply of bonds determines the liquidity spread of government bonds, i.e., changes in A_g affect s_g and allocations.

C.4 The relation between unemployment and the nominal interest rate

In the following proposition, we describe a *modified* Phillips curve that gives the relationship between u and the short-term nominal interest rate, i_g. Moreover, we characterize the choice of two policy instruments, π and i_g, to minimize the unemployment rate.

Proposition 9 ***(Unemployment and the short-term nominal interest rate.)***

1. ***(The modified Phillips curve)*** *Given $i > 0$, a small decrease in i_g starting from $i_g = i^-$ leads to a decrease in the unemployment rate (u). Moreover, there exists $\bar{\imath} > 0$ such that if $i > \bar{\imath}$, then a large decrease in i_g from $i_g = i^-$ to $i_g = 0$ leads to an increase in u.*

2. ***(Unemployment miniminzing policy)*** *Unemployment is minimized when $s_m = s_g > 0$, i.e., $i > 0$ and $i_g = 0$.*

The first claim of part 1 of Proposition 9 states that in the neighborhood of $i_g = i^-$, u falls as i_g decreases. Indeed, when $i_g = i$, the opportunity cost of holding liquid bonds is zero, and hence type-2 buyers hold enough bonds to purchase the first-best level of output, i.e., $y_2 = y^*$. Therefore, if i_g falls slightly below i, y_2 falls as well, but it only has a second-order effect on the output net of the production cost, $\upsilon(y) - \varphi(y)$. However, the increase in s_g

above 0 has a first-order effect on consumers' outside options, Z_2, which raises firms' market power. Consequently, it induces more firm entry and a lower unemployment rate.

The second claim of part 1 states that the relationship between u and i_g is non-monotone when inflation, or i, is sufficiently large. By the same logic as in the pure monetary economy, we can establish that the unemployment rate is lower when $s_g = 0$ compared to $s_g = +\infty$. However, s_g is bounded above by i, i.e., i_g is bounded below by 0. So, in order for u to be non-monotone as i_g decreases from i to 0, it must be that i (or π) is sufficiently large. We illustrate these findings in Figure 11 by plotting the modified Phillips curve for low and high inflation rates.

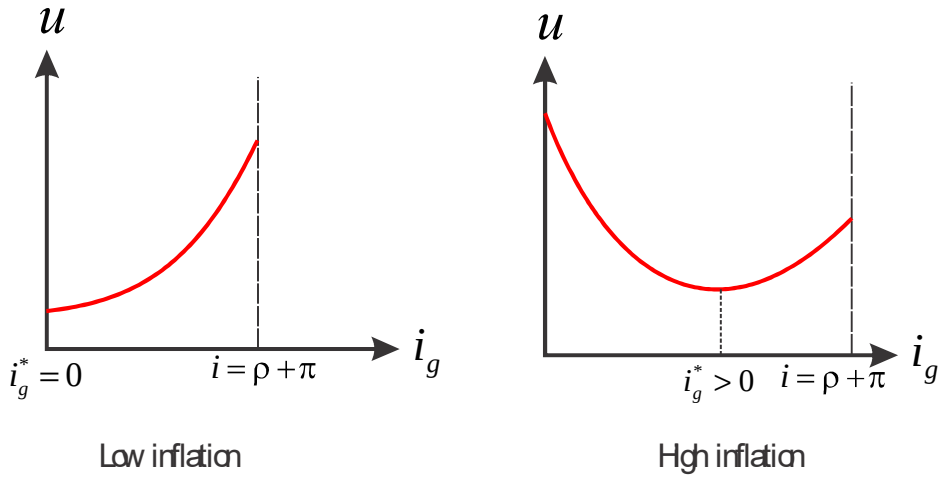


Figure 11: The modified Phillips curve: The relationship between u and i_g at low and high inflation rates

Part 2 of Proposition 9 determines the monetary policy mix, in terms of inflation and short-term nominal interest rate, that minimizes unemployment. From (78), the policymaker chooses the spreads s_m and s_g as follows:

$$s_j \in \arg\max \left\{ \chi^{\ell} S^{\ell} \left[a^{\kappa}(s_j), Z_{\kappa}(s_j) \right] + \chi^{d} S^{d} \left[Z_{\kappa}(s_j) \right] \right\}, \text{ for } \ (j, \kappa) \in \{m, g\} \times \{1, 2\}.$$

The problems that determine s_m and s_g are independent since the profits arising from the two types of buyers are additively separable and identical. Hence, if s_m maximizes the profits from type-1 consumers, then $s_g = s_m$ maximizes the profits from type-2 consumers. An implication of this result is that the unemployment-minimizing nominal interest rate, i_g, is zero, i.e., when the unemployment rate is minimum, the economy is in a liquidity trap. Note that the reverse is not always true. An economy can be in a liquidity trap but the inflation rate might not be at the level that minimizes unemployment. In addition, as in Proposition 4, the unemployment rate is minimum when interest rate spreads are positive,

i.e., the inflation rate is above the Friedman rule.

We illustrate these effects in Figure 12 under the calibration in Section 5 with $\lambda_j = \lambda = 0.94$ and $\mu_j = \mu = 0.97$. We set the fraction of type-1 consumers to 0.4 to match the fraction of households in the Survey of Consumer Finances with no financial assets (i.e. they only hold currency, transaction accounts, or durable assets). The left panel illustrates the relationship between unemployment and the liquid bond rate, for a given inflation rate. For inflation rates below -0.2%, unemployment is increasing in the liquid bond rate and the unemployment-minimizing i_g is achieved at $i_g = 0$. For higher inflation rates, unemployment is U-shaped in i_g and the unemployment-minimizing liquid bond rate increases in inflation (middle panel). The middle panel illustrates the unemployment-minimizing i_g as a function of the inflation rate while the right panel illustrates the resulting minimum level of unemployment. The minimum-unemployment joint policy (π, i_g) is achieved at $\pi = -0.2\%$ and $i_g = 0$ (the dots in middle and right panel).

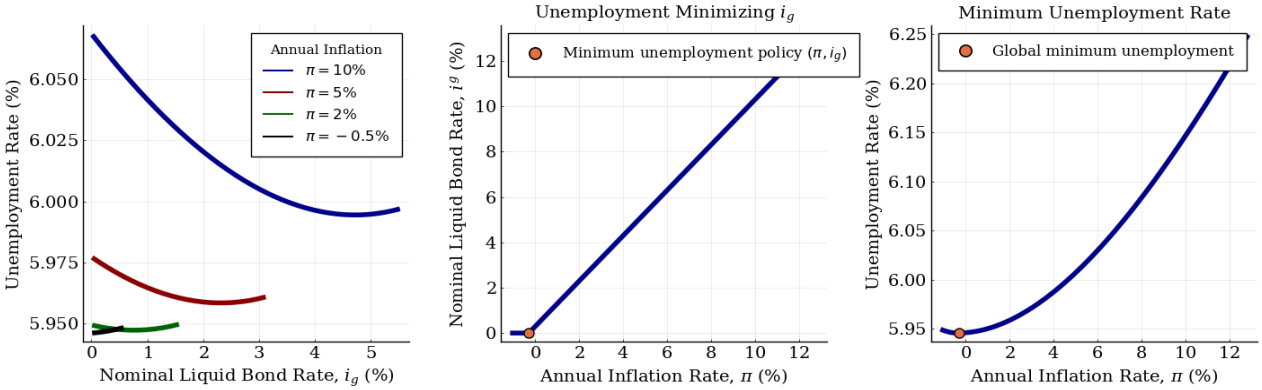


Figure 12: The effect of the nominal rate i^g on unemployment (left), the unemployment-minimizing nominal rate i^g given π (middle), and the minimum unemployment achieved given π (right).

So far, we have assumed that consumers of type 1 and 2 were identical in terms of their preferences and opportunities to consume. Suppose now that type-1 and type-2 consumers receive preference shocks at different rates, λ_1 and λ_2. This heterogeneity can lead to an optimal policy such that $s_m \neq s_g$ and hence $i_g > 0$. To see this, note that for $(j, \kappa) \in \{(m, 1), (g, 2)\}$,

$$\left.\frac{\partial Z_\kappa}{\partial s_j}\right|_{s_j = 0^+} = \frac{-p(y^*)}{\rho + \lambda_\kappa + \gamma + \alpha(1-\mu)}.$$

Intuitively, in absolute value, the impact of s_j on Z_κ is smaller when λ_κ is larger. Hence, when λ_κ is larger, the outside option is less responsive to changes in s_j, and so an increase in s_j is less effective in reducing unemployment. Thus, as λ_κ increases, the spread that minimizes unemployment should decrease. Therefore, we conjecture that if $\lambda_2 > \lambda_1$, then

$s_g < s_m$ and $i_g > 0$. The next proposition shows that this conjecture is true when λ_1 is sufficiently large, and we numerically illustrate that the conjecture holds in our calibrated economy.

Proposition 10 ***(Asymmetric consumers.)*** *If $\lambda_2 > \lambda_1$ and λ_1 is sufficiently large, then the unemployment minimizing policy features $s_g < s_m$ and thus $i_g > 0$.*

In Figure 13, we illustrate the unemployment-minimizing joint policy, either as a combination of nominal interest rates (i, i_g) or as a combination of spreads (s_m, s_g). We fix λ_2 and vary λ_1 along the x-axis. For $\lambda_1 < \lambda_2$, unemployment is minimized for positive inflation rates $i > 0$ and positive nominal liquid bond rates $i_g > 0$. As λ_1 increases, both i and i_g fall until λ_1 rises above λ_2, in which case $i_g = 0$ minimizes unemployment. These results illustrate that when $\lambda_1 < \lambda_2$, a policy maker who wants to minimize unemployment might prefer $i, i_g > 0$.

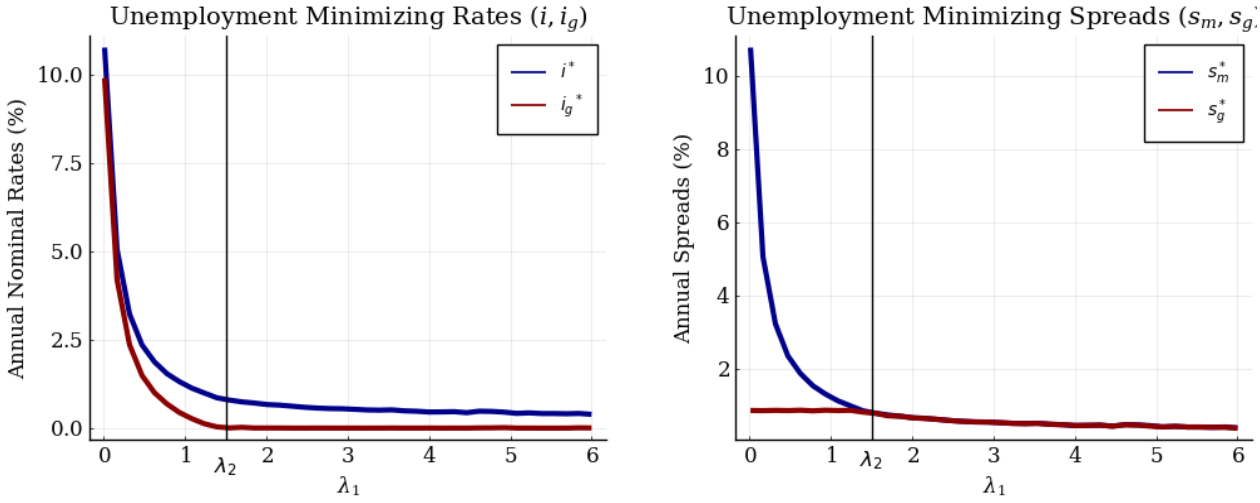


Figure 13: Unemployment-minimizing policy under asymmetric preferences (λ_1, λ_2)

C.5 Proofs

Proof of Proposition 9. *Part 1*. From (78), for given s_m, market tightness increases with the expected surplus in type-2 matches,

$$\chi^{\ell} S^{\ell} \left[a^2(s_g), Z_2\right] + \chi^{d} S^{d}(Z_2).$$

If $s_g = 0$, i.e., $i_g = i$, from (80), a small increase in s_g reduces y_2 below y^*, which has a second-order effect on S^{ℓ}. From (76), it reduces Z_2, which has a first-order and positive effect on S^{ℓ} and S^{d}. As a result, θ increases and u decreases. By the same reasoning as in

the proof of Part 2 of Proposition 4, market tightness when $s_g = 0$ is larger than market tightness when $s_g = +\infty$,

$$\theta|_{s_g=0} > \theta|_{s_g=+\infty}.$$

Since $s_g \leq i$, there exists a threshold for i, denoted $\bar{\imath}$, such that if $i > \bar{\imath}$, then $\theta|_{s_g=0} > \theta|_{s_g=i}$ and $u|_{s_g=0} < u|_{s_g=i}$.

Part 2. The minimum level of unemployment is implemented when θ solution to (78) is maximum. It follows that the optimal spreads are such that

$$s_j^* \in \arg\max_{s_j \geq 0} \left\{ \chi^\ell S^\ell \left[a^\kappa(s_j), Z_\kappa(s_j) \right] + \chi^d S^d [Z_\kappa(s_j)] \right\}, \quad (j, \kappa) \in \{m, g\} \times \{1, 2\}.$$

By the same reasoning as before, a solution exists and it is such that $s_j^* > 0$. ■

Proof of Proposition 10. Let the expected trade surplus generated by a type-κ consumer be

$$\mathcal{S}_\kappa(s_j) \equiv \chi^\ell S^\ell \left[a^\kappa(s_j), Z_\kappa(s_j) \right] + \chi^d S^d \left[Z_\kappa(s_j) \right], \quad (j, \kappa) \in \{m, g\} \times \{1, 2\}.$$

We assume s_j^* is the maximizer of $\mathcal{S}_\kappa(s_j)$. To show $s_g^* < s_m^*$, differentiate $\mathcal{S}_\kappa(s_j)$ with respect to s_j, taking the market tightness θ as given:

$$\begin{aligned} \frac{\partial \mathcal{S}_\kappa(s_j)}{\partial s_j} &= -\frac{\partial Z_\kappa}{\partial s_j} + \chi^\ell \frac{\partial S^\ell \left[a^\kappa(s_j), Z_\kappa \right]}{\partial a^\kappa(s_j)} \frac{\partial a^\kappa(s_j)}{\partial s_j} \\ &= -\frac{\partial Z_\kappa}{\partial s_j} + \frac{s_j}{\alpha(1-\mu)} \frac{\partial a^\kappa(s_j)}{\partial s_j}. \end{aligned}$$

Next consider the cross derivative by differentiating the expression above with respect to λ_κ:

$$\begin{aligned} \frac{\partial^2 \mathcal{S}_\kappa(s_j)}{\partial s_j \partial \lambda_\kappa} &= -\frac{\partial^2 Z_\kappa}{\partial s_j \partial \lambda_\kappa} + \frac{s_j}{\alpha(1-\mu)} \frac{\partial^2 a^\kappa(s_j)}{\partial s_j \partial \lambda_\kappa} \\ &= \frac{1}{\rho + \lambda_\kappa + \gamma + \alpha(1-\mu)} \frac{\partial Z_\kappa}{\partial s_j} \left[1 - \frac{\mu Z_\kappa}{a^\kappa(s_j)} + \frac{s_j}{\alpha(1-\mu)} \mu \right]. \end{aligned}$$

By the logic of monotone comparative statics, s_j^* falls in λ_κ if the cross derivative is negative for all relevant choice of s_j. The derivative $\partial Z_\kappa / \partial s_j$ is negative. As $\lambda_\kappa \to \infty$ (i.e. when λ_1 is sufficiently large), $Z_\kappa \to 0$ and the square bracketed term becomes positive. Hence, the optimal choice of s_j^* falls in λ_κ when λ_1 is large. ■

D Equilibrium consumer search

In the baseline model of Section 3, consumers who are matched with a firm never exert their option to search on the equilibrium path. Searching for alternative producers is a threat that reduces firms' market power and that affects the surplus of the match and the terms of trade. In this section, we extend our model with horizontally differentiated products to introduce search on the equilibrium path.

The preference of a consumer for good y is now $\varepsilon v(y)$ where ε is a random variable capturing the idiosyncratic taste of the consumer. When a consumer meets a firm, $\varepsilon \in [0, \bar{\varepsilon}]$ is drawn from a cumulative distribution $F(\varepsilon)$ and is common-knowledge in the match.[29] The consumer can then decide to bargain with the firm or to keep searching for an alternative producer.

Consider a match between a consumer with a real balances and a firm when the realization of the preference shock is ε. There are gains from trade if

$$\max_{y \geq 0} \{\varepsilon v(y) - \varphi(y) : \varphi(y) \leq a\} > Z, \tag{81}$$

which has a similar interpretation as (4). The next lemma characterizes the threshold for ε, above which condition (81) holds.

Lemma 1 ***(Optimal search)*** *The threshold for ε above which gains from trade are positive is:*

$$\begin{aligned} \varepsilon_R(a, Z) &= \hat{\varepsilon}(Z) \quad \textit{if } \hat{\varepsilon}(Z) \leq \tilde{\varepsilon}(a), \\ &= \frac{a + Z}{v[\varphi^{-1}(a)]} \quad \textit{otherwise,} \end{aligned} \tag{82}$$

where $\hat{\varepsilon}(Z)$ is the solution to $\varepsilon v(y_\varepsilon^*) - \varphi(y_\varepsilon^*) = Z$ and $\tilde{\varepsilon}(a) \equiv \varphi'[\varphi^{-1}(a)]/v'[\varphi^{-1}(a)]$.

The threshold ε_R can take two values. If a is large, it is equal to $\hat{\varepsilon}(Z)$ which only depends on a consumer's outside options. It is the value of ε such that the match surplus is zero when the quantity traded, y, is efficient. If a is small, the liquidity constraint binds and ε_R depends on both a and Z. It increases with Z, i.e., if consumers' outside options improve, consumers become pickier. It decreases with a, i.e., if consumers hold more real balances, then they are willing to buy varieties that they value less because they can compensate the lowest match quality by purchasing larger quantities.

[29] One can give several interpretations for ε. For instance, firms produce different varieties of good y and consumers value these different varieties differently. Alternatively, the intensity for the desire to consume could be varying over time.

For all $\varepsilon \geq \varepsilon_R$, the determination of the terms of trade is given by

$$\max_{p,y} \{p - \varphi(y)\} \quad \text{s.t. } p - \varphi(y) = \mu \left[\varepsilon \upsilon(y) - \varphi(y) - Z\right] \quad \text{and} \quad p \leq a. \tag{83}$$

We denote $y_\varepsilon(a, Z)$ as the solution to (83). The surpluses in monetary and credit matches are

$$S_\varepsilon^m(a, Z) \equiv \varepsilon \upsilon\left[y_\varepsilon(a, Z)\right] - \varphi\left[y_\varepsilon(a, Z)\right] - Z \quad \text{if } \varepsilon \geq \varepsilon_R(a, Z), \text{ and} \tag{84}$$

$$S_\varepsilon^d(Z) \equiv \varepsilon \upsilon(y_\varepsilon^*) - \varphi(y_\varepsilon^*) - Z \quad \text{if } \varepsilon \geq \hat{\varepsilon}(Z). \tag{85}$$

If there are no gains from trade, then $S_\varepsilon^m = S_\varepsilon^d = 0$. Following the same reasoning as above, the outside option of the consumer, Z, solves

$$(\rho + \lambda + \gamma) Z = \max_{a \geq 0} \left\{ -ia + \alpha(1 - \mu) \int \left[\chi^m S_\varepsilon^m(a, Z) + \chi^d S_\varepsilon^d(Z)\right] dF(\varepsilon) \right\}. \tag{86}$$

This equation has a similar interpretation as (14). It admits a unique solution, $Z \in [0, \bar{\varepsilon} y_{\bar{\varepsilon}}^* - \varphi(y_{\bar{\varepsilon}}^*)]$, which is an increasing function of α and a decreasing function of μ. If the optimal choice of real balances is interior, it solves

$$\int_{\varepsilon_R(a,Z)}^{\bar{\varepsilon}} \frac{\alpha \chi^m (1 - \mu) \left[\varepsilon \upsilon'(y_\varepsilon) - \varphi'(y_\varepsilon)\right]}{\mu \varepsilon \upsilon'(y_\varepsilon) + (1 - \mu) \varphi'(y_\varepsilon)} dF(\varepsilon) = i, \tag{87}$$

where $y_\varepsilon = y_\varepsilon(a, Z)$. The left side, which represents the marginal benefits from holding real balances, is decreasing in a. So, if a solution to (87) exists, it is unique.

The free-entry condition (22) in the labor market is generalized to give

$$(\rho + \delta) \frac{k\theta}{f(\theta)} + \beta k \theta = (1 - \beta) \times \left\{ \alpha^s(\theta) \mu \left[\chi^m \int_{\varepsilon_{R(a,Z)}}^{\bar{\varepsilon}} S_\varepsilon^m(a, Z) dF(\varepsilon) + \chi^d \int_{\hat{\varepsilon}(Z)}^{\bar{\varepsilon}} S_\varepsilon^d(Z) dF(\varepsilon) \right] + x - \ell(\theta) \right\} \tag{88}$$

The two integrals on the right side represent firm surpluses in all monetary and credit matches where the gains from trade are positive.

The steady-state measure of active consumers solves

$$\left\{\gamma + \alpha \chi^m \left[1 - F(\varepsilon_R)\right] + \alpha \chi^d \left[1 - F(\hat{\varepsilon})\right]\right\} \omega_1 = \lambda(\omega - \omega_1). \tag{89}$$

The left side is the flow of consumers who become inactive, either because their desire for consumption vanishes at rate γ or because they meet a firm that produces a variety of the

good that they want to consume. The right side represents the flow of inactive consumers that become active at rate λ. Using that $q\omega_1 = n$, (89) can be rewritten to obtain the following relationship between q and θ:

$$\frac{\lambda\omega q}{\gamma + \alpha(q)\chi^m \left[1 - F(\varepsilon_R)\right] + \alpha(q)\chi^d \left[1 - F(\hat{\varepsilon})\right] + \lambda} = \frac{f(\theta)}{\delta + f(\theta)}. \tag{90}$$

The tightness of the goods market increases with θ but decreases with ε_R and $\hat{\varepsilon}$.

An equilibrium is defined as a list, $(\varepsilon_R, Z, a, \theta, q)$, that is a solution to (82), (86), (87), (88), and (90). In the following, we consider equilibria in the neighborhood of the Friedman rule, i.e. $i = 0^+$.

Proposition 11 ***(Long-run Phillips curve and consumer search)*** *Assume $x > b$, $\rho + \lambda + \gamma$ is small, and $i = 0^+$.*

1. *Conditional on a trade taking place, quantities traded are efficient, $y_\varepsilon = y_\varepsilon^*$ for all $\varepsilon \geq \varepsilon_R(a, Z) = \hat{\varepsilon}(Z)$, where a and Z are the equilibrium values when $i = 0^+$.*

2. *An increase in λ or γ reduces the value of consumer search, Z, and makes consumers less picky, i.e., ε_R decreases. It raises labor market tightness, θ, and reduces unemployment, u.*

3. *A small increase in i from $i = 0^+$ generates an increase in θ, and a decrease in u, Z, and ε_R.*

The proof utilizes some substitutions to reduce an equilibrium to a pair, (θ, Z), that solves (86) and (88). The assumption that the effective discount rate, $\rho + \lambda + \gamma$, is small allows us to focus on equilibria where the curve representing (88) cuts the curve representing (86) from above in the (θ, Z) space, as shown in Figure 14. We also focus on monetary policy in the neighborhood of the Friedman rule so that, conditional on a trade, quantities are efficient, as shown by the first part of the proposition.

The second part of Proposition 11 establishes the links between market power, consumer search, and unemployment. As λ or γ rises, the value of searching falls and hence firms' market power increases. Consumers spend their real balances on goods that they value less, i.e., ε_R decreases. Firms' expected revenue rises, which induces them to open more vacancies that leads to lower unemployment.

The third part of Proposition 11 shows that the result according to which the long-run Phillips curve is downward sloping at low inflation rates is robust when one introduces differentiated goods and ex-post match heterogeneity. As the inflation rate increases, consumer search becomes more costly. As a result, consumers become less choosy, ε_R decreases, and

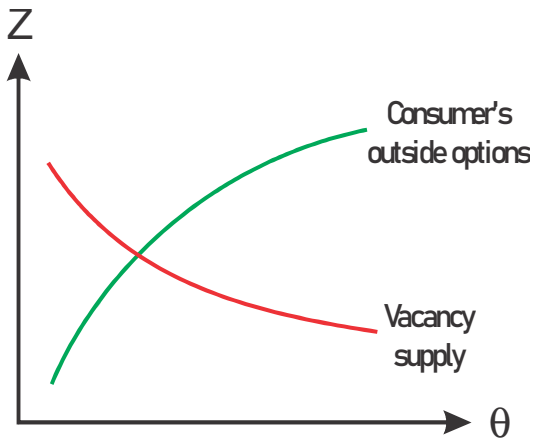


Figure 14: Equilibrium with consumer search when $i = 0^+$

the value of their outside options decrease. As firms' market power increases, they open more vacancies, θ increases, and the unemployment rate decreases.

D.1 Proofs

Proof of Lemma 1. The inequality (81) holds iff $\varepsilon > \varepsilon_R$, where the reservation value for the preference shock, $\varepsilon_R(a^*, Z)$, solves

$$\max_{y \geq 0} \{\varepsilon_R \upsilon(y) - \varphi(y)\} = Z \quad \text{s.t.} \quad \varphi(y) \leq a^*. \tag{91}$$

We distinguish two cases depending on whether or not the constraint, $\varphi(y) \leq a^*$, binds.

Case #1. If the constraint $\varphi(y) \leq a^*$ is not binding, then $\varepsilon_R = \hat{\varepsilon}(Z)$ is the solution to $\varepsilon \upsilon(y_\varepsilon^*) - \varphi(y_\varepsilon^*) = Z$ where $y_\varepsilon^* = \arg\max \{\varepsilon \upsilon(y) - \varphi(y)\}$. We now check the condition under which $\varphi(y_{\hat{\varepsilon}}^*) \leq a^*$ is slack. Using that φ is an increasing bijection, we have

$$\varphi(y_{\hat{\varepsilon}}^*) \leq a^* \iff y_{\hat{\varepsilon}}^* \leq \varphi^{-1}(a^*).$$

Apply the increasing function φ'/υ' on both sides to obtain

$$\varphi(y_{\hat{\varepsilon}}^*) \leq a^* \iff \frac{\varphi'(y_{\hat{\varepsilon}}^*)}{\upsilon'(y_{\hat{\varepsilon}}^*)} \leq \frac{\varphi' \circ \varphi^{-1}(a^*)}{\upsilon' \circ \varphi^{-1}(a^*)}.$$

Use the definition of $y_{\hat{\varepsilon}}^*$, i.e., $\varphi'(y_{\hat{\varepsilon}}^*) = \hat{\varepsilon} \upsilon'(y_{\hat{\varepsilon}}^*)$, to rewrite the inequality above as

$$\varphi(y_{\hat{\varepsilon}}^*) \leq a^* \iff \hat{\varepsilon} \leq \tilde{\varepsilon}(a^*) \equiv \frac{\varphi' \circ \varphi^{-1}(a^*)}{\upsilon' \circ \varphi^{-1}(a^*)}.$$

Case #2. If the constraint, $\varphi(y) \leq a^*$, is binding then $y = \varphi^{-1}(a^*)$ so that ε_R solves

$\varepsilon_R \upsilon[\varphi^{-1}(a^*)] - a^* = Z$. Solving for ε_R, we obtain $\varepsilon_R = (a^* + Z)/\upsilon[\varphi^{-1}(a^*)]$. ■

Proof of Proposition 11. *Part 1*. From (87), $y_\varepsilon = y_\varepsilon^*$ for all $\varepsilon \geq \varepsilon_R$, i.e., agents trade the efficient quantities in all matches where there are gains from trade. This requires $a^* = \varphi(y_{\bar{\varepsilon}}^*) + \mu\left[\bar{\varepsilon}\upsilon(y_{\bar{\varepsilon}}^*) - \varphi(y_{\bar{\varepsilon}}^*) - Z\right]$.

We now prove that $\varepsilon_R(a^*, Z) = \hat{\varepsilon}(Z)$. From (86), $Z < \bar{\varepsilon}\upsilon(y_{\bar{\varepsilon}}^*) - \varphi(y_{\bar{\varepsilon}}^*)$ since otherwise $S_\varepsilon^m(a^*, Z) = S_\varepsilon^d(Z) = 0$ and

$$(\rho + \lambda + \gamma) Z = \alpha(1-\mu) \int \left[\chi^m S_\varepsilon^m(a^*, Z) + \chi^d S_\varepsilon^d(Z)\right] dF(\varepsilon) = 0,$$

which is a contradiction. Using that $Z < \bar{\varepsilon}\upsilon(y_{\bar{\varepsilon}}^*) - \varphi(y_{\bar{\varepsilon}}^*)$, $\varphi^{-1}(a^*) > y_{\bar{\varepsilon}}^*$ and $\tilde{\varepsilon}(a^*) \equiv \varphi'[\varphi^{-1}(a^*)]/\upsilon'[\varphi^{-1}(a^*)] > \varphi'(y_{\bar{\varepsilon}}^*)/\upsilon'(y_{\bar{\varepsilon}}^*) = \bar{\varepsilon}$. Moreover, $\hat{\varepsilon}(Z) < \bar{\varepsilon}$. Hence, by Lemma 1, $\varepsilon_R(a^*, Z) = \hat{\varepsilon}(Z)$.

We now show how to reduce an equilibrium to a pair (θ, Z) that is a solution to two equations. From (90),

$$\frac{\lambda\omega q}{\gamma + \alpha(q)\left[1 - F(\hat{\varepsilon})\right] + \lambda} = \frac{f(\theta)}{\delta + f(\theta)}. \tag{92}$$

From (92), $q = Q(\theta, \hat{\varepsilon})$, where Q is increasing in θ and decreasing in $\hat{\varepsilon}$. From (86), Z solves

$$(\rho + \lambda + \gamma) Z = \alpha\left[Q(\theta, \hat{\varepsilon}(Z))\right](1-\mu) \int_{\hat{\varepsilon}(Z)}^{\bar{\varepsilon}} S_\varepsilon^d(Z) dF(\varepsilon). \tag{93}$$

From (93), Z is an increasing function of θ. From (88), θ solves

$$(\rho + \delta)\frac{k\theta}{f(\theta)} + \beta k\theta = (1-\beta)\left[\alpha^s\left[Q(\theta, \hat{\varepsilon}(Z))\right]\mu \int_{\hat{\varepsilon}(Z)}^{\bar{\varepsilon}} S_\varepsilon^d(Z) dF(\varepsilon) + x - b\right]. \tag{94}$$

When Z is close to $\bar{\varepsilon}\upsilon(y_{\bar{\varepsilon}}^*) - \varphi(y_{\bar{\varepsilon}}^*)$, the effect of a change in Q on the right side is negligeable. In that case, θ is a decreasing function of Z. From (93), as $\rho + \lambda + \gamma$ approaches 0, for all $\theta > 0$, Z approaches $\bar{\varepsilon}\upsilon(y_{\bar{\varepsilon}}^*) - \varphi(y_{\bar{\varepsilon}}^*)$. From (94), θ approaches the positive solution to

$$(\rho + \delta)\frac{k\theta}{f(\theta)} + \beta k\theta = (1-\beta)(x-b).$$

Part 2. An increase in λ or γ shifts the curve representing (93) downward in the space (θ, Z). See Figure 14. As a result Z decreases while θ increases. It follows that $\varepsilon_R = \hat{\varepsilon}(Z)$ decreases.

Part 3. Consider now a small increase in i from $i = 0^+$. We still have $\varepsilon_R(a^*, Z) = \hat{\varepsilon}(Z)$ where $q = Q(\theta, \hat{\varepsilon})$ is implicitly defined by (92). Any change in a^* only has a second-order

effect on $S_\varepsilon^m(a, Z)$, hence (86) is approximated by

$$(\rho + \lambda + \gamma) Z = -ia^* + \alpha [Q(\theta, \hat{\varepsilon})] (1 - \mu) \int S_\varepsilon^d(Z) dF(\varepsilon).$$

As i increases, the curve representing (86) shifts downward in the (θ, Z) space. The equilibrium condition (88) is still approximated by (94). Since we start from an equilibrium where the curve representing (94) intersects the curve representing (93) by above, θ increases, Z decreases, and $\varepsilon_R = \hat{\varepsilon}$ decreases. ■

E Alternative interpretation of production cost

Previously, we have assumed the production cost $\varphi(y)$ was a disutility paid by the entrepreneur, or manager. An alternative interpretation would be that $\varphi(y)$ is a production cost paid by the worker, who is compensated by the wage w. The interpretation of $\varphi(y)$ does not matter for allocations, but it affects the definition of the wage w in (18), and thus the expression of the wage markdown. In this appendix, we determine the new equations for the wage, and the wage markdown, under this alternative interpretation.

Given the new interpretation of w and $\varphi(y)$, the worker now pays the variable cost of production, and equation (18) becomes

$$\rho E = w - \alpha^s \left[\chi^m \varphi(y^m) + \chi^d \varphi(y^*)\right] - \delta\beta J. \tag{95}$$

By the logic leading to (23), the wage w can be reexpressed as

$$w = \beta \left\{\alpha^s \mu \left[\chi^m S^m(a^*, Z) + \chi^d S^d(Z)\right] + x\right\} + \alpha^s \left[\chi^m \varphi(y^m) + \chi^d \varphi(y^*)\right] + (1-\beta)b + \beta k\theta. \tag{96}$$

When compared with (23), the key novelty is the presence of the second term on the right side, which represents the compensation to the worker for the variable cost of production.

We can compute the wage markdown as in (25). But since the variable cost of production is paid by the worker and compensated by wages, we do not need to subtract the variable cost when calculating the net expected revenue of a firm. Hence, $\hat{x} = \mathbb{E}[p] + x$, or equivalently

$$\hat{x} = \alpha^s \left\{\chi^m \varphi(y^m) + \chi^d \varphi(y^*) + \mu \left[\chi^m S^m(a^*, Z) + \chi^d S^d(Z)\right]\right\} + x. \tag{97}$$

The markdown is

$$\begin{aligned} MKDOWN &\equiv \frac{\hat{x} - w}{\hat{x}} \\ &= \frac{(1-\beta)\left\{\alpha^s \mu \left[\chi^m S^m(a^*, Z) + \chi^d S^d(Z)\right] + x - b\right\} - \beta k\theta}{\hat{x}}. \end{aligned} \tag{98}$$

The numerator is the same as that in (25), which represents the net profit of the firm-worker match. But now the net expected revenue of the firm in the denominator includes the variable cost of production, as captured by the difference between (24) and (97).

F Competitive search

In the main text we argue that one can implement constrained efficiency allocations if the Hosios conditions are satisfied in the goods and labor markets, and the monetary policy follows the Friedman rule. Now we illustrate this possibility by considering a version of our model where the terms of trade in the goods and labor markets are determined by competitive search, instead of random search. The goal of the exercise is to show that the Hosios conditions (54) and (55) are satisfied under competitive search. In this exercise we assume that the nominal interest rate is at the Friedman rule level, $i_t = 0$.

F.1 Goods market

The goods market is similar to that in Rocheteau and Wright (2005). The dual problem of the firm's profit maximization problem is such that it posts a list of quantities, payments, and market tightness to maximize consumers' surplus, subject to the profit being higher than the equilibrium level of profit Π^*.

$$\begin{aligned} &\max_{p_m,p_d,y_m,y_d,n} \{\alpha(n)[\chi^m(\upsilon(y_m) - p_m - Z) + \chi^d(\upsilon(y_d) - p_d - Z)]\} \\ s.t. \quad &\frac{\alpha(n)}{n}\{\chi^m[p_m - \varphi(y_m)] + \chi^d[p_d - \varphi(y_d)]\} = \Pi^*. \end{aligned}$$

By using the constraint, we can eliminate $\chi^m p_m + \chi^d p_d$ from the objective function. Then, the optimal quantities are $y_m = y_d = y^*$, and the first-order condition with respect to n is given by

$$\alpha'(n)[\upsilon(y^*) - \varphi(y^*) - Z] = \Pi^*.$$

By substituting this expression back into the constraint of the optimization problem, the share of surplus to the firm equals the elasticity of the matching function, i.e.

$$\frac{\chi^m p_m + \chi^d p_d - \varphi(y^*)}{\upsilon(y^*) - \varphi(y^*) - Z} = \frac{\alpha'(n)n}{\alpha(n)} \equiv \epsilon_\alpha.$$

Hence the Hosios condition is satisfied in the goods market.

F.2 Labor market

Now we derive the firm's surplus in the labor market. The continuation value of an employed worker is

$$\rho E = w + \delta(U - E).$$

As mentioned in the previous subsection, the surplus in the goods market is given by $S^d(Z) = \upsilon(y^*) - \varphi(y^*) - Z$ and the firm's share of the surplus is ϵ_α. Therefore, the value of the firm-worker pair, J, is given by

$$(\rho + \delta)J = \alpha^s \epsilon_\alpha S^d(Z) + x - \rho U. \tag{99}$$

Similar to the goods market, under competitive search, firms maximize workers' expected payoff subject to the free-entry condition, namely

$$\begin{aligned} &\max_{\theta,w}\{f(\theta)\,(E-U)\}, \\ s.t. \quad &\frac{f(\theta)}{\theta}\left\{J - E + U\right\} = k. \end{aligned}$$

Using the constraint to eliminate $E - U$ in the objective function, the firm's problem can be rewritten as an unconstrained problem

$$\max_{\theta}\left\{f(\theta)J - k\theta\right\}.$$

By (99), the matched value J is independent of the firm's choice of θ and w. Therefore, the first-order condition with respect to θ implies the firm's share of the surplus is

$$\frac{J - E + U}{J} = \frac{\theta f'(\theta)}{f(\theta)} \equiv \epsilon_f.$$

Hence the Hosios condition is satisfied in the labor market. Since Proposition 7 shows that $\mu = \epsilon_\alpha$, and $\beta = 1 - \epsilon_f$ lead to a constrained-efficient outcome under the Friedman rule, we conclude that the combination of the Friedman rule and competitive search can implement the constrained-efficient outcome.

G Magnitude of the market power effect

We now study how the magnitude of the market power effect of inflation depends on λ. Since the market power effect vanishes as λ explodes, one might expect it to decrease monotonically in λ. We argue not. We do so by considering how λ affects the magnitude of the derivative $\partial\theta/\partial i|_{i=0}$. For tractability, we consider the special case when agents are infinitely patient, $\rho = 0$, and the job separation rate vanishes, $\delta \to 0$. As δ vanishes, $n \to 1$ and $q \to 1/\omega$. The unemployment rate is $u = 0$ but θ is still endogenous and determined by the free-entry condition. This limit is not necessarily realistic but is instructive because of its tractability. At this limit, by (28) and (30), θ and Z are given by

$$\beta k\theta = (1-\beta)\left\{\alpha^s(1/\omega)\mu\left[\chi^m S^m(a,Z) + \chi^d S^d(Z)\right] + x - b\right\} \text{ and} \tag{100}$$

$$(\lambda+\gamma)\,Z = -ia + \alpha^b(1/\omega)(1-\mu)\left[\chi^m S^m(a,Z) + \chi^d S^d(Z)\right]. \tag{101}$$

The market tightness, θ, is linear in the surplus of a firm-work pair. Hence, the derivative of θ with respect to i, at $i = 0$, is proportional to the change in Z, namely

$$\left.\frac{\partial\theta}{\partial i}\right|_{i=0} = \frac{(1-\beta)\alpha^s(1/\omega)\mu}{\beta k}\left(-\left.\frac{\partial Z}{\partial i}\right|_{i=0}\right). \tag{102}$$

Therefore, the change in θ is determined solely by the market power effect and is proportional to the change in the outside option. The impact of inflation on the outside option, Z, can be derived using (30), i.e.

$$\left.\frac{\partial Z}{\partial i}\right|_{i=0} = \frac{-a}{\lambda+\gamma+\alpha^b(1/\omega)(1-\mu)}. \tag{103}$$

Money holdings, a, and the outside option, Z, are given by

$$a = (1-\mu)\varphi(y^*) + \mu\upsilon(y^*) - \mu Z \text{ and } Z = \frac{\alpha^b(1/\omega)(1-\mu)[\varphi(y^*) - \upsilon(y^*)]}{\lambda+\gamma+\alpha^b(1/\omega)(1-\mu)}.$$

According to (103), $\partial Z/\partial i$ is proportional to $-a$. It is because an infinitesimal increase in i reduces Z via raising the cost of money holding, by an amount that is equal to a. As λ rises, there are two opposing effects on $\theta/\partial i$. First, a rises because the outside option, Z, falls and thus buyers must carry more money for payment. This effect raises the magnitude of $\partial Z/\partial i$ and $\partial\theta/\partial i$. On the other hand, as λ rises, the denominator in (103) rises because Z is smaller and becomes less sensitive to changes in i. This effect reduces the magnitude of $\partial Z/\partial i$ and $\partial\theta/\partial i$. Altogether, because of these opposing forces, $\partial\theta/\partial i$ and $\partial Z/\partial i$ can be non-monotone in λ, as shown in the following claim.

Claim 1 *Assume $i = \rho = 0$, and consider the limit $\delta \to 0$. The derivative $\partial\theta/\partial i$ is non-negative and falls in β. As $\lambda + \gamma$ rises from 0 to $+\infty$, $\partial\theta/\partial i$ falls monotonically if $\mu \leq \varphi(y^*)/[\upsilon(y^*) - \varphi(y^*)]$ and otherwise it is hump-shaped. It vanishes as λ explodes.*

Proof. We only show the claim regarding increases in $\lambda + \gamma$ because other claims are straightforward implications of (102). Substituting the expressions for a and Z into (102),

$$\frac{\partial\theta}{\partial i} \propto \frac{1}{\lambda + \gamma + \alpha(q)(1-\mu)} \left[(1-\mu)\varphi(y^*) + \mu\upsilon(y^*) - \mu \frac{\alpha(q)(1-\mu)[\varphi(y^*) - \upsilon(y^*)]}{\lambda + \gamma + \alpha(q)(1-\mu)} \right],$$

where the sign $\propto$ means the left and right side have the same sign. The derivative of the right side with respect to λ is positive if and only if

$$\frac{2\alpha(q)(1-\mu)\mu[\upsilon(y^*) - \varphi(y^*)]}{(1-\mu)\varphi(y^*) + \mu\upsilon(y^*)} > \lambda + \gamma + \alpha(q)(1-\mu),$$

which fails when $\lambda + \gamma$ is sufficiently large. Hence $\partial\theta/\partial i$ is either decreasing or hump-shaped in λ. When $\lambda + \gamma = 0$, the inequality holds if and only if $\mu \leq \varphi(y^*)/[\upsilon(y^*) - \varphi(y^*)]$. ■

According to Claim 1, the size of the market power effect, $\partial\theta/\partial i$, is hump-shaped in λ provided that the bargaining power of firms in the goods market, μ, is sufficiently large.