Inequality, Liquidity, and Optimal Monetary Policy

Florin O. Bilbiie\textsuperscript{II} \quad Xavier Ragot\textsuperscript{III}

\textit{PSE and CEPR} \quad \textit{Sciences Po and OFCE}

March 2017 (First draft February 2016)

Abstract

When inequality is of the essence, optimal monetary policy requires deviations from price stability over the cycle. We build a sticky-price model where aggregate demand depends on liquidity, as heterogeneous consumers hold money in face of uninsurable risk and participate infrequently in financial markets. The model is tractable and can be solved in closed form. A novel trade-off for Ramsey-optimal monetary policy emerges between inequality and standard stabilization objectives. Inequality appears in the "loss function" (a second-order approximation to aggregate welfare). Price stability has significant welfare costs: inflation volatility is optimal because it insures constrained agents by hindering their consumption volatility.

\textbf{JEL Codes:} D14, D31, E21, E3, E4, E5

\textbf{Keywords:} inequality; optimal (Ramsey) monetary policy; heterogenous agents; incomplete markets; liquidity constraints; limited participation; money.

\textsuperscript{I}Previously circulated as part of "Monetary Policy and Inequality when Aggregate Demand Depends on Liquidity" and presented since March 2016. We are especially grateful to our discussant Benjamin Moll and to Mirko Wiederholt for extensive comments, to Adrien Auclert, Paul Beaudry, Jess Benhabib, Wei Cui, Ester Faia, Nobu Kiyotaki, Jordi Gali, Alejandro Justiniano, Dirk Krueger, Eric Leeper, Francesco Lippi, Ralph Luetticke, Roland Meeks, Cyril Monnet, Morten Ravn, Pontus Rendahl, Gilles Saint-Paul, Vincent Sterk, Paolo Surico, Gianluca Violante, Pierre-Olivier Weill, and Michael Woodford for comments and suggestions, and to other participants in several conferences and seminars. Bilbiie thanks (without implicating) Banque de France for financial support through the eponymous Chair at the Paris School of Economics, Institut Universitaire de France, and New York University for hospitality during part of writing this paper.

\textsuperscript{II}Paris School of Economics, 48 Boulevard Jourdan 75014 Paris. Email: florin.bilbiie@parisschoolofeconomics.eu. URL: http://florin.bilbiie.googlepages.com.

\textsuperscript{III}Sciences Po, CNRS, and OFCE URL:http://www.parisschoolofeconomics.com/ragot-xavier/
1 Introduction

Should central banks care about inequality? Central bankers certainly do care—so much so that the last two Chairs of the Federal Reserve explicitly called for more research on its monetary policy implications: Bernanke (2007, 2015) and Yellen (2014). Echoes come from the other side of the Atlantic: the President of the European Central Bank recently dedicated a whole speech to the issue (Draghi, 2016, October), as did previously other board members (Coëuré, 2012; Mersch, 2014).

Yet the literature on optimal monetary policy does not provide a simple welfare-based reason for including inequality as a policy objective for the monetary authority. This is what this paper does: in a model where inequality matters for aggregate demand, we identify a new trade-off: between reducing inequality, and standard stabilization of inflation and real activity. Quantitatively, this novel trade-off implies significant optimal deviations from price stability over the cycle.\(^1\)

On the empirical front, several papers studied the impact of monetary policy on inequality (e.g. through the redistribution effects of inflation), using a variety of methods.\(^2\) A general and robust conclusion seems to be that looser monetary policy is associated with less inequality. And monetary policy has certainly been loose in the aftermath of the 2008 financial crisis and the ensuing Great Recession, most notably through unprecedented liquidity expansion: to take one example, the year-on-year growth rate of M1 quadrupled (from 2.5 to 11 percent on average) in the post-crisis period as compared to the 2000-2008 interval.\(^3\)

Our goal is to contribute to the understanding of optimal monetary policy in a model that belongs to a new synthesis that is under way at the time of our writing. A very recent quantitative literature that we review below analyzes monetary policy transmission in incomplete-market, heterogeneous-agent New Keynesian models—abbreviated "HANK" by Kaplan, Moll, and Violante (2014). These contributions can speak to the aforementioned empirical findings, and are also consistent with recent microeconometric evidence on the heterogeneity of marginal propensities to consume MPCs, its relation to liquidity constraints, and income and wealth distributions.\(^4\)

We thus revisit standard "New Keynesian" optimal monetary policy—in a tractable general equilibrium model where households are heterogeneous and subject to liquidity constraints, and liquidity is used to self-insure against uninsurable risk: financial markets are incomplete as in Bewley,\(^5\)

\(^1\) We review below other recent (dated 2017) studies that deal with the topic.
\(^2\) Starting from Doepke and Schneider (2006), these include i.a. Adam and Zhu (2014), Coibion, Gorodnichenko, Kueng, and Silvia (2013), Adam and Tzamourani (2016), Deutsche Bundesbank (2016), and Furceri, Loungani, and Zdzienicka (2016).
\(^3\) Since nominal GDP has actually fallen during the crisis and growth thereafter does not nearly match money growth, it follows that velocity sank during the crisis and kept falling.
\(^4\) Recent empirical evidence using micro data from various sources supports the hypothesis that high MPCs correspond to households who are liquidity constrained (rather than, say, income-poor); see Kaplan and Violante (2014), Cloyne et al (2016), Jappelli and Pistaferri (2014), Misra and Surico (2014) and Surico and Trezzi (2016).
and participation is limited (infrequent) in the Baumol-Tobin tradition.

In equilibrium, aggregate demand depends on liquidity, and inequality matters. A clarification regarding our use of these two key words—liquidity and inequality—is in order. We define liquidity as the nominal asset which is used by agents to self-insure, and call this asset "money"—but it can be any asset whose return is affected by monetary policy. We use inequality as in the Bewely-Huggett-Aiyagari (heterogeneous-agent) literature, i.e. as the endogenous outcome of uninsurable shocks combined with agents’ ability to self-insure. Admittedly, by focusing on short-run business cycles and stabilization policy, our framework does not capture other important aspects of inequality, such as human capital accumulation, inequality along the age-dimension, and others—some of which operate in the richer models reviewed below. But here we focus on the notion of inequality that has a long tradition, going back at least to Friedman’s (1969) analysis of the redistributive effect of monetary policy and to Bewley’s (1983) formalization of that analysis.

Like many others, we consider that monetary policy is the relevant tool at business cycle (quarterly) frequency to improve the distorted market outcome. Thus, we analyze the residual trade-offs for monetary policy after the (imperfect) use of any fiscal tools (without considering time-varying fiscal tools as a policy instrument). Liquidity is an equilibrium phenomenon in our economy precisely because inequality subsists, as fiscal policy does not insure agents perfectly. The degree of liquidity demanded is thus an indirect metric of the insurance job left undone by fiscal policy. We in fact calibrate the degree of inequality in the model—which as we shall see is the main determinant of optimal inflation—to match the fall in consumption at unemployment observed in the data (which takes into account any fiscal transfers).

We study Ramsey-optimal monetary policy in this framework, and unveil—what is to the best of our knowledge—a novel trade-off between reducing inequality and stabilizing inflation and aggregate demand. Indeed, in our model inequality emerges as one of the objectives of optimal policy. We illustrate this point analytically by providing a second-order approximation to the aggregate welfare function à la Woodford (2003, Ch. 6). There is scope for a planner to decrease inequality and provide consumption insurance—an objective that is costly to achieve through inflation when prices are sticky (and absent a full set of fiscal instruments). This trade-off operates in the long-run, as in any monetary model, making deflation optimal. But more importantly, and unlike other monetary

---

5Our framework can hence accommodate nominal bonds, if they are used to self-insure and thus have a liquidity premium. See also the discussion of alternative choices to model liquidity in Section 2.1.

6See Krueger, Mitman and Perri (2017) for a recent contribution and review of what is now a vast literature.

7Note that this is the standard in monetary policy analysis: even in the baseline, textbook NK model, if the fiscal authority had enough lump-sum instruments and the ability to use them at quarterly frequency, any cost-push type shocks could be accommodated through variations in, e.g., labor-tax rates or sale subsidies—thus redistributing from firms to consumers. Similarly, the zero lower bound could be avoided by appropriate saving taxes (consumption subsidies). Standard analysis assumes that such perfect redistribution is unfeasible, which is what we also assume.
sticky-price frameworks, it also operates in the short run: insofar as there is long-run inequality, optimal policy requires volatile inflation. Moreover, this inflation volatility matters for welfare: a central bank that stabilizes inflation, albeit around an optimal long-run target, incurs a large welfare cost—consumers would pay (around 0.1 to 0.5 percent of consumption) to live in the economy with volatile inflation. Such deviations and welfare effects are larger than those encountered in existing monetary models with nominal rigidities.

Inflation volatility is beneficial in our economy because it dampens the consumption volatility of constrained agents without much affecting unconstrained agents who can self-insure. The optimal policy consists of providing liquidity, which insures constrained agents, and inflating away some of its value, in order to give the right intertemporal incentives to unconstrained agents to hold this liquidity for precautionary purposes.

1.1 Related Literature

Our paper is related to three literatures that we now review. We start this review in inverse chronological order, by placing our contribution in the context of several papers whose first draft appeared (and of which we became aware of) after we circulated and presented ours. All these independent contributions constitute the first sub-literature: optimal monetary policy with heterogeneous agents and sticky prices.

Perhaps most related in this realm, at least with regards to the main policy conclusion, is a recent paper by Bhandari, Evans, Golosov and Sargent (2017) that studies optimal monetary policy in a fully-fledged HANK model. The model is solved numerically using techniques that allow to bring the state space to finite dimensionality. What distinguishes our framework is the introduction of limited participation in financial markets as a microfoundation for money/liquidity; we thus consider two assets, at the cost of a simplification of the cross-sectional distribution. We therefore focus on and isolate a novel trade-off between inequality and standard stabilization, and abstract from several other mechanisms that are at work in Bhandari et al in order to focus on this channel. We find closed-form solutions for the case of exogenous policy and when solving for optimal policy, we summarize the trade-offs through a "loss function"—which allows a transparent illustration of the mechanism at work. The Ramsey problem that we solve is much simpler but illustrates that some form of long-run inequality, in the form of imperfect consumption insurance, is enough—and hence essential—to motivate large optimal deviations from price stability.

Two other recent papers analyze optimal policy in different incomplete-market models. Nuno and Thomas (2017) solve for optimal policy in a model with incomplete markets but only idiosyncratic (as opposed to aggregate) risk. Challe (2017) studies optimal monetary policy in a model where the

---

8 Limited participation being a pervasive fact in the US data, the heterogeneity of returns that we consider is likely important for the link between monetary policy and inequality.
uninsurable unemployment risk is endogenous—a channel that we abstract from—but there is no equilibrium trade and no endogenous liquidity.

Several earlier studies analyzed optimal monetary policy in different heterogeneous-agents models, focusing on other channels. In the realm of two-agent models, Bilbiie (2008) derives optimal policy in a model with hand-to-mouth agents, and Curdia and Woodford (2009) and Nistico (2015) in models with infrequent participation and borrowers and savers. The setup of these last two papers shares similarities to ours, in particular concerning the "infrequent participation" structure that draws on an earlier monetary theory literature; but in the domain of optimal policy, these studies focus on the case where there is perfect insurance in steady state, thus abstracting from the "inequality" channel that gives rise to the novel trade-off we emphasize.

We owe much debt to the literature that, building on the seminal paper of Lucas and Stokey (1983), shaped our understanding of optimal policy in NK models. Some of the key contributions include Khan, King, and Wollman (2003), Adao, Correia, and Teles (2003), Woodford (2003, Ch. 6), Benigno and Woodford (2005, 2012), and Schmitt-Grohe and Uribe (2004, 2007).

In our framework, significant deviations from price stability are optimal, and not only in the long run—the cited papers also imply, when relying on (other, different) money demand theories, some convex combination between the Friedman rule and a zero inflation long-run prescription. Our framework also gives rise, more surprisingly, to significant optimal deviations from price stability over the cycle—in response to shocks that in those frameworks do not generate such deviations. A welfare-maximizing central bank relies on inflation volatility optimally, as this inflation volatility provides insurance and contributes to reducing inequality. Renouncing this volatility (by adopting a policy of constant deflation at the optimal asymptotic rate) thus has a large welfare cost in our model, whereas it is innocuous in the NK models with money demand reviewed above. The key to this difference is inequality.

The model that we use for studying optimal policy integrates two streams of monetary economics that evolved divergently over the past two decades: New Keynesian (NK) models with nominal rigidities, and microfounded models of money demand with flexible prices. Within these frameworks, we connect their two subsets that focus on heterogeneity, market incompleteness and limited participation.

---

9Braun and Nakajima (2012) prove an aggregation result, showing the conditions under which the optimal policy in an incomplete-market economy is the same as that of a representative-agent model.

10An important difference between our model and Curida and Woodford’s is also that our non-participant agents are liquidity constrained, and consume a liquidity injection that relaxes this constraint. Whereas in their setup, non-participant borrowers can borrow but subject to a spread.

11Money demand in the NK model is generically residual when money is introduced in the utility function, through a cash-in-advance constraint, or through shopping-time distortions. This has nevertheless important consequences for optimal policy, which we review in due course.
The second sub-literature that we build upon consists of monetary theory models with limited participation and incomplete markets in the Bewley and Baumol-Tobin tradition. In our model, money is used to self-insure against idiosyncratic shocks as in Bewley models, but only for non-participating agents as in the Baumol-Tobin literature. Some of the key contributions, all with flexible prices, include Bewley (1983); Scheinkman and Weiss (1986); Lucas, (1990); Kehoe, Levine, and Woodford (1992); Algan, Challe, and Ragot (2010); Alvarez and Lippi (2014); Khan and Thomas (2015); Cao et al (2016); Lippi, F., S. Ragni and N. Trachter (2015); Gottlieb (2015); Rocheteau, Weill and Wong (2015, 2016); and Ragot (2016).

The third sub-literature studies heterogeneous agents in NK models; an earlier literature introduced "hand-to-mouth" consumers (or limited participation in asset markets) to study aggregate demand and monetary policy—one could call this "first-generation HANK". Gali, Lopez-Salido and Valles (2004, 2007) and Bilbiie (2004, 2008) are two early examples of such models, where a subset of agents are employed hand-to-mouth and have unit MPC. Compared to these models, we allow for temporarily-binding credit constraints and allow constrained agents to self-insure. So does the more recent literature referred to as HANK above: quantitative models with household heterogeneity and incomplete markets that are consistent with microeconomic heterogeneity and data on household finances, and replicate plausible distributions of wealth and marginal propensities to consume.

Kaplan, Violante, and Moll (2014) revisit the transmission mechanism of monetary policy in such a model with liquid and illiquid assets. In contrast to representative-agent NK models where monetary policy works mainly through intertemporal substitution, in their model monetary policy works mainly through what they label an "indirect effect" (the endogenous, general-equilibrium response of output). Ravn and Sterk (2013) analyze an incomplete-markets model where search and matching frictions generate unemployment risk; job uncertainty in their model can generate deep and lasting recessions through aggregate demand amplification. Gornemann, Kuester, and Nakajima (2012) also study monetary transmission when markets are incomplete and unemployment risk endogenous, focusing on the distributional welfare effects on households with different wealth levels. Challe, Matheron, Ragot, and Rubio-Ramirez (2015) estimate a model with endogenous unemployment risk (search and matching) using Bayesian methods, and assess the quantitative importance of the link

---

12Recent empirical work argues that such frictions are needed to explain money demand, including the distribution of money holdings across agents (i.a. Alvarez and Lippi 2009; Cao et al 2012; Ragot 2014).

between precautionary saving and aggregate demand. McKay, Nakamura, and Steinsson (2015) use a similar model to those mentioned above, but with exogenous unemployment risk, to show that forward guidance is less powerful than in the standard model—mostly because an incomplete-markets model implies a form of "discounting" of aggregate demand. Auclert (2015) analyzes the role of redistribution for the transmission mechanism and decomposes it into three channels that are related to households’ asset positions.

Very few models in this realm consider money demand at all. Bayer, Luetticke, Pham-Dao, and Tjaden (2015) use a HANK model where money is used as precautionary liquidity—their underlying model of money demand thus being conceptually similar to ours; they show how higher idiosyncratic uncertainty increases precautionary money demand, thus leading to a recession. They analyze the effects of different money-growth rules, including the differential welfare effects on different groups of agents, depending on their wealth. Den Haan, Rendhal and Riegler (2016) study a model with uninsurable unemployment risk and sticky wages and show that these two features taken together provide amplification and can cause a deflationary spiral; a key element is (precautionary) demand for money, where money enters the utility function.\(^{14}\)

Our framework captures some key features and mechanisms of these recent quantitative HANK models. In building our simplified framework, we trade off some (relevant and important, but thoroughly analyzed elsewhere) complexity for analytical tractability; this allows us to analyze monetary policy transmission and optimal monetary policy, which are an integral part of the state-of-the-art NK framework. To fix ideas, one could argue that while the existing literature in this realm puts more emphasis on the "heterogeneous-agent" part of HANK, our framework does the opposite—it simplifies heterogeneity to put more emphasis on the latter part of HANK. In our opinion, the two approaches are complementary.

2 A Monetary NK Model with Heterogeneous Agents

We build a simple, tractable, heterogeneous-agent, New Keynesian model with money: heterogeneous households hold money to self-insure against unemployment risk, markets are incomplete, participation is limited (infrequent), and price adjustment is costly.

**Households.** There is a mass 1 of households, indexed by \( j \in [0, 1] \), who discount the future at rate \( \beta \) and derive utility from consumption \( c^j_t \) and disutility from labor supply \( l^j_t \). The period utility function is:

\[
    u(c^j_t) = \chi \left( \frac{(l^j_t)^{1+\varphi}}{1+\varphi} \right),
\]

\(^{14}\)They also argue that there is a significant role for unemployment insurance in their model; see also McKay and Reis (2015).
with \( u(c) = (c^{1-\gamma} - 1) / (1 - \gamma) \). Households have access to three assets: money (with zero nominal return), public debt (with nominal return \( \delta_t > 0 \)), and shares in monopolistically competitive firms. Money is held despite being a dominated asset because financial frictions give it a consumption-smoothing, insurance role. These frictions are: uninsurable idiosyncratic risks and infrequent participation in financial markets. Such frictions customarily generate a large amount of heterogeneity: the economy is characterized by a continuous distribution of wealth, which is very hard to study with aggregate shocks and sticky prices.

To simplify the problem (and thus enable us to perform the analysis previewed in the Introduction), we use tools developed in the incomplete-markets literature to reduce the amount of heterogeneity. These simplifications keep the essence of intertemporal trade-offs and of redistributive effects of monetary policy in general equilibrium, and can be viewed as a simple generalization of the Lucas (1990) multiple-member household metaphor. As we shall see, in our economy the key intertemporal trade-offs are captured by agents’ Euler equations for money and other assets; at the same time, a relevant but limited amount of heterogeneity captures the redistributive effects of inflation and money creation. The gain of this modeling strategy is that standard tools used in New Keynesian economics can be used; in particular, we can compute Ramsey-optimal policy with aggregate shocks—to the best of our knowledge, for the first time in an incomplete-markets economy with uninsurable risk and limited participation.

Households participate infrequently in financial markets. When they do, they can freely adjust their portfolio and receive dividends from firms. When they do not, they can use only money to smooth consumption. Denote by \( \alpha \) the probability to keep participating in period \( t + 1 \), conditional upon participating at \( t \) (hence, the probability to switch to not participating is \( 1 - \alpha \)). Likewise, call \( \rho \) the probability to keep non-participating in period \( t + 1 \), conditional upon not participating at \( t \) (hence, the probability to become a participant is \( 1 - \rho \)). The fraction of participating households is \( n = (1 - \rho) / (2 - \alpha - \rho) \), and the fraction \( 1 - n = (1 - \alpha) / (2 - \alpha - \rho) \) does not participate.

Furthermore, households belong to a family whose head maximizes the intertemporal welfare of family members using a utilitarian welfare criterion (all households are equally weighted), but faces some limits to the amount of risk sharing that it can do. Households can be thought of as being in two states or "islands". All households who are participating in financial markets are on the

---

15 See also Curdia and Woodford (2009) for another application of the "infrequent participation" structure in a different context (with savers and borrowers) with sticky prices.

16 Bilbiie (2008) analyzes optimal monetary policy in a sticky-price model when some agents are (employed) hand-to-mouth, and there is no self-insurance other than labor. Curdia and Woodford (2009) also look at optimal policy in their model, around a steady-state with perfect insurance.

17 The relevance of the family head and island metaphors is studied further, in a different context, in Le Grand and Ragot (2016)—where the latter occurs as the endogenous outcome of an economy where the set of insurance contracts is incomplete. See also Khan and Thomas (2011) for a decentralization of the family head assumption with limited participation.
same island, called $P$. All households who are not participating in financial markets are on the same island, called $N$. The family head can transfer all resources across households within the island, but cannot transfer some resources between islands.

Households in the participating island work at real wage $w_t$, whereas households in the non-participating island work to get a fixed home-production amount $\delta$ (which is also their fixed labor supply). To simplify the exposition, we assume that financial risks (participating or not) and labor market risks (employed or not) are perfectly correlated. This isolates the channel that we want to emphasize: self-insurance through money in face of unemployment risk.

The timing is the following. At the beginning of the period, the family head pools resources within the island. The aggregate shocks are revealed and the family head determines the consumption/saving choice for each household in each island. Then households learn their next-period participation status and have to move to the corresponding island accordingly, taking only money with them. The key assumption is that the family head cannot make transfers to households after the idiosyncratic shock is revealed, and will take this as a constraint for the consumption/saving choice.

The flows across islands are as follows. The total measure of households leaving the $N$ island each period is the number of households who participate next period: $(1 - n) (1 - \rho)$. The measure of households staying on the island is thus $(1 - n) \rho$. In addition, a measure $(1 - \alpha) n$ leaves the $P$ island for the $N$ island at the end of each period.

Total welfare maximization implies that the family head pools resources at the beginning of the period in a given island and implements symmetric consumption/saving choices for all households in that island. Denote as $b_{t+1}^P$ and $M_{t+1}^P$ the per-capita period $t$ bonds and money balances respectively, in the $P$ island, after the consumption-saving choice. The real money balances are $m_{t+1}^P = M_{t+1}^P / \mathcal{P}_t$, where $\mathcal{P}_t$ is the price level. The end-of-period per capita real values (after the consumption/saving choice but before agents move across islands) are $\tilde{b}_{t+1}^P$ and $\tilde{m}_{t+1}^P$. Denote as $m^N_t$ the per capita beginning-of-period capital money in the $N$ island (where the only asset is money). The end-of-period values (before agents move across islands) are $\tilde{m}_{t+1}^N$. We have the following relations, after simplification (as bonds do not leave the $P$ island, we have $b_{t+1}^P = \tilde{b}_{t+1}^P$):

$$
m_{t+1}^P = \alpha \tilde{m}_{t+1}^P + (1 - \alpha) \tilde{m}_{t+1}^N
$$

$$
m_{t+1}^N = (1 - \rho) \tilde{m}_{t+1}^P + \rho \tilde{m}_{t+1}^N.
$$

The program of the family head is (with $\pi_t = (\mathcal{P}_t - \mathcal{P}_{t-1}) / \mathcal{P}_{t-1}$ denoting the net inflation rate):

$$
W \left( b_t^P, m_t^P, m_t^N, X_t \right) = \max_{\left( c_t^P, b_{t+1}^P, m_t^P, m_t^N, \tilde{m}_{t+1}^P, \tilde{m}_{t+1}^N, c_{t+1}^P \right)} \left[ u \left( c_t^P \right) - \chi \frac{\left( b_t^P \right)^{1+\varphi}}{1 + \varphi} \right] + (1 - n) \left[ u \left( c_t^N \right) - \chi \frac{\delta^{1+\varphi}}{1 + \varphi} \right] + \beta EW \left( b_{t+1}^P, m_{t+1}^P, m_{t+1}^N, X_{t+1} \right)
$$
subject to:

\[ c^P_t + b^P_{t+1} + \tilde{m}^P_{t+1} = w^P_t - \tau^P_t \]

\[ + \frac{1 + \rho_t}{1 + \pi_t} b^P_t + \frac{m^P_t}{1 + \pi_t} + \frac{1}{n} d_t, \]

\[ \tilde{m}^N_{t+1} + c^N_t = \delta - \tau^N_t + \frac{m^N_t}{1 + \pi_t} \]

\[ \tilde{m}^P_{t+1}, \tilde{m}^N_{t+1} \geq 0 \]

and the laws of motion for money flows outlined above, relating \( m^j_{t+1} \) to \( \tilde{m}^j_{t+1} \). Equation (1) is the per capita budget constraint in the \( P \) island: \( P \)-households (who own all the firms) receive dividends \( d_t/n_t \), and the real return on money and bond holdings. With these resources they consume and save in money and bonds, and pay taxes/receive transfers \( \tau^P_t \) (lump-sum taxes include any new money created or destroyed). Equation (2) is the budget constraint in the \( N \) island. Finally (3) are positive constraints on money holdings and are akin to credit constraints in the heterogeneous-agent literature.

The variable \( X_t \) in the value function refers to all relevant period \( t \) information necessary to form rational expectations. Using the first-order and envelope conditions, we have:

\[ u'(c^P_t) \geq \beta E \frac{1 + \rho_t}{1 + \pi_{t+1}} u'(c^P_{t+1}) \text{ or } b^P_{t+1} = 0 \]

\[ u'(c^P_t) \geq \beta E \left[ \alpha u'(c^P_{t+1}) + (1 - \alpha) u'(c^N_{t+1}) \right] \frac{1}{1 + \pi_{t+1}} \text{ or } \tilde{m}^P_{t+1} = 0 \]

\[ u'(c^N_t) \geq \beta E \left[ (1 - \rho) u'(c^P_{t+1}) + \rho u'(c^N_{t+1}) \right] \frac{1}{1 + \pi_{t+1}} \text{ or } \tilde{m}^N_{t+1} = 0 \]

\[ w_t u'(c^P_t) = \chi \left( \frac{1}{t^P_t} \right)^{\varepsilon} \]

The first Euler equation corresponds to the choice of bonds: there is no self-insurance motive, for they cannot be carried to the \( N \) island: the equation is the same as with a representative agent. The money choice of \( P \)-island agents is governed by (5), which takes into account that money can be used when moving to the \( N \) island. The third equation (6) determines the money choice of agents in the \( N \) island, and the last equation labor supply.

The important implication of this market structure is that the Euler equations (5) and (6) have the same form as in a fully-fledged incomplete-markets model of the Bewely-Huggett-Aiyagari type. In particular, the probability \( 1 - \alpha \) measures the uninsurable risk to switch to "low income" (unemployment) next period, risk for which money is the only means to self-insure. This is why money is held in equilibrium for self-insurance purposes, despite being a dominated asset.

**Production and Price Setting.** The final good is produced by a firm using intermediate goods as inputs. The final sector production function is \( Y_t = \left( \int_0^{y_t(z)} (y_t(z))^{1-\varepsilon} dz \right)^{\frac{1}{1-\varepsilon}} \), where \( y_t \) is the amount of intermediate good \( z \) used in production. Denote as \( P_t(z) \) the price of intermediate goods \( z \). Demand for an individual product is \( Y_t(z) = (P_t(z)/P_t)^{-\varepsilon} Y_t \) with the welfare-based price
index $\mathcal{P}_t = \left( \int_0^1 \mathcal{P}_t(z)^{1-\varepsilon} \, dz \right)^{1/\varepsilon}$. Each individual good is produced by a monopolistic competitive firm, indexed by $z$, using a technology given by: $Y_t(z) = A_t I_t(z)$. Cost minimization, taking the wage as given, implies that the real marginal cost is $W_t / (A_t \mathcal{P}_t)$. The problem of producer $z$ is to maximize the present value of future profits, discounted using the stochastic discount factor of their shareholders, the participants.

When price adjustment is frictionless, prices of all firms are equal to a constant markup over the nominal marginal cost—the real marginal cost is constant $W_t / (A_t \mathcal{P}_t) = (\varepsilon - 1) / \varepsilon$. We assume that firms are subject to nominal rigidities as in Rotemberg (1982): to change their prices, firms incur a quadratic adjustment cost that is homogenous across firms. Profits of each firm are thus given by $d_t = \left( 1 - \frac{w_t}{A_t} - \frac{\nu}{2} \pi_t^2 \right) Y_t$, anticipating that the equilibrium is symmetric. Maximization of their present discounted value gives rise to the nonlinear forward-looking "New Keynesian Phillips curve", whose derivation is described in detail in the Appendix—where we replaced the labor supply schedule $w_t = \chi \left( l_t^P \right) \phi \left( c_t^P \right) \gamma$:

$$\pi_t (1 + \pi_t) = \beta E_t \left[ \left( \frac{c_t^P}{P_t} \right)^\gamma \frac{Y_{t+1}}{Y_t} \pi_{t+1} (1 + \pi_{t+1}) \right] + \frac{\varepsilon}{\nu} \left( \frac{\chi \left( l_t^P \right) \phi \left( c_t^P \right) \gamma}{A_t} + \Phi - 1 \right),$$

where $\Phi \equiv 1 - (\varepsilon - 1) (1 + \sigma) / \varepsilon$ captures the steady-state distortion and $\sigma$ is a corrective sales subsidy. In particular, when the subsidy is equal to the desired net markup $\sigma = (\varepsilon - 1)^{-1}$, there is no steady-state distortion associated with monopolistic competition and elastic labor, $\Phi = 0$. These considerations will be useful when studying the Ramsey policy below.

**Money Creation and the Government Budget.** New money is created through "helicopter drops", and we consider uniform taxation $\tau_t^N = \tau_t^N = \tau_t$. Denote by $m_{t+1}^{CB}$ the (real value of) new money created in period $t$, and by $M_{t+1}^{\text{tot}}$ the total nominal quantity of money in circulation at the end of each period. In nominal terms, $M_{t+1}^{\text{tot}} = M_{t+1}^{\text{tot}} + \mathcal{P}_t m_{t+1}^{CB}$, and in real terms:

$$m_{t+1}^{\text{tot}} = \frac{m_{t}^{\text{tot}}}{1 + \pi_t} + m_{t+1}^{CB}$$

Hence, the total period $t$ net taxes/transfers are $\tau_t = -m_{t+1}^{CB}$.

**Market clearing and equilibrium.** Since there is no public debt, the period $t$ market for bonds is $n b_{t+1}^P = 0$. The money market clears $m_{t+1}^{\text{tot}} = (1 - n) m_{t+1}^N + n \tilde{m}_{t+1}^P$ and so does the labor

---

18Different ways of money creation, i.e. open market operations have some starkly different positive and normative implications that are the subject of a companion paper (Bibbiie and Ragot, 2016). In that paper, we also leave open the possibility of exogenous redistribution by choosing type-specific transfers: $\tau_t^P = \frac{\omega}{n} \pi_t; \tau_t^N = \frac{1 - \omega}{1 - n} \pi_t$ where $\omega$ is the share of total taxes paid by all type-N agents. When $\omega = n$ we have the case of entirely lump-sum transfers. At the other extreme, it can be easily shown that there exist processes $\tau_t^I$ that redistribute money so as to replicate Woodford’s cashless limit. Finally, there exists a value of $\omega$ that restores neutrality and Wallace’s 1981 logic—i.e. "keeping fiscal policy constant" in the sense of finding an (exogenous) redistribution that un-does the endogenous redistribution triggered by a monetary policy shock in our framework.
market $l_t = nl_t^P$. Denoting by $c_t$ total consumption and $Y_t = A^i l_t$ market-produced output (or earned total income), we have that the goods market will also clear, by Walras’ Law:

$$c_t \equiv nc_t^P + (1 - n) c_t^N = \left(1 - \frac{\nu}{2} \pi_t^2\right) Y_t + (1 - n) \delta.$$  (10)

Note for further use that there is a resource cost of changing prices (inflation), which is isomorphic to the welfare cost of relative price dispersion in a Calvo-type model, see e.g. Woodford (2003). In Appendix A we provide the summary of model equations and the equilibrium definition.

**Steady state.** The analysis of the model’s steady state (defined as an allocation where real variables are constant and nominal variables grow at a constant rate $\pi$) provides a series of first insights into its monetary structure. The Euler equation for bonds implies that their real return is always equal to the inverse of the discount factor:

$$\frac{1 + i}{1 + \pi} = \beta^{-1}.$$  

Defining the new variable *consumption inequality* as $q_t \equiv c_t^P / c_t^N$, the self-insurance Euler equation delivers:

$$q \equiv \frac{c_t^P}{c_t^N} = \left(\frac{1 + \pi}{\beta} - \alpha\right)^{\frac{1}{\alpha}} > 1.$$  

Letting the steady-state share of unemployment benefits (home production) in average consumption be $\delta_c \equiv \delta/c$, and the share of $N$ households’ consumption in total be $h$:

$$h \equiv \frac{c_t^N}{c} = \frac{1}{1 + n (q - 1)},$$  

we find (as long as it is positive) the steady-state money demand share, or inverse consumption velocity of money:

$$\mu \equiv \frac{m_{\text{tot}}}{c} = \frac{h - \delta_c}{2 - \rho - \alpha},$$  

Subject to a caveat of existence of a monetary equilibrium, discussed in detail in Appendix B, steady-state money demand is equal to the share of (non-home-produced) consumption of $N$, divided by a parameter capturing the degree of overall churning, the sum of the transition probabilities from one state to another. Under the restriction $\alpha + \rho > 1$ (which we return to below), this parameter is between 0 and 1. For a given level of home production, this expression implicitly defines upper bounds on the degree of market incompleteness (as described by $\alpha$ and $\rho$) so that steady-state money demand is positive.\(^\dagger\) Conversely, for given $\alpha$ and $\rho$ there exists a threshold $\delta$ beyond which $P$ agents choose not to hold money: the outside option is too good and there is no need to self-insure. Thus, $\delta_c$ captures the degree of insurance provided by (unmodelled) fiscal transfers: were it high enough,

\(^{19}\)The formal restriction is, for the case of zero steady-state inflation and treating $n$ as a parameter: $\alpha < 1 - n \frac{\beta^{-1} - 1}{\delta_c - 1} < \beta^{-1}$. In terms of the original parameter $\rho$ we have $\frac{1 - \alpha}{1 - \rho} > \sqrt{\frac{1}{2} + \frac{(1 - \delta c) \beta}{(1 - \rho) \beta (1 - \delta)} - \frac{1}{2}}$. 12
no liquidity would be traded $\mu = 0$ and there would be no role for monetary policy in this model beyond its standard role in cashless models. We will focus on the case with equilibrium liquidity and inequality, $\delta_c < h < 1$.

### 2.1 Simple Monetary NK Model with Heterogeneous Agents

It is instructive to pause and compare the household side of our model with that of the seminal HANK papers reviewed in the introduction. This helps understand how ours is a simplified version of that framework—what mechanisms it still captures, and what it misses in order to gain tractability. Take first our concept of *liquidity*, which differs from that of Kaplan, Moll, and Violante (2014)—where bonds are liquid and equity and housing illiquid. In assuming that bonds and equity are illiquid while money is liquid, we follow the definition of the monetary theory that we reviewed: a similar notion of liquidity is also used by the quantitative HANK model of Bayer, Luetticke, Pham-Dao, and Tjaden (2015). Second, our constrained unit-MPC households are *wealthy hand-to-mouth*, similarly to Kaplan, Moll, and Violante’s—their wealth is located just on the $P$ island, where they have a positive probability of going (back). Third, unlike in Kaplan, Moll, and Violante, our constrained households are *unemployed*. An earlier literature already clarified the amplifying, Keynesian effect on monetary transmission of hand-to-mouth households who are employed and thus have endogenous income (see Bilbiie, 2004; 2008 and the discussion in the Introduction). We abstract from that well-understood general equilibrium channel to isolate and better understand another, which we emphasize below—endogenous movements in liquidity; this is also consistent with the uninsurable risk being related to unemployment.

Lastly, the assumptions we used to reduce *heterogeneity* and *history-dependence* have a close counterpart in the sticky-price literature that is probably clear to readers well-seasoned in NK models: our participation/insurance scheme is conceptually similar to the Calvo model of price stickiness. Whereas Kaplan, Moll, and Violante's portfolio decision based on a quadratic transaction cost for illiquid assets is conceptually close to the Rotemberg model of price stickiness; because state variables enter this decision complex distributional dynamics occur that our simplification abstracts from.

Thus, our model *cannot* fit the detailed distribution of asset holdings and wealth, nor reproduce movements in portfolio shares or realistic idiosyncratic income processes—it does not capture the rich household wealth dynamics of fully-fledged Bewley-Aiyagari-Huggett models; in particular, it does

---

20 Tongue in cheek, one may label this a MONK model, as in "monetarist New Keynesian". See Weill (2007), Rocheteau and Weill (2011), Kiyotaki and Moore (2012), and Cui and Sadde (2016) for recent sophisticated refinements (as well as reviews) of the concept of liquidity in recent monetary theory, including a different view of liquidity based on asset resalability (while our is on limited participation).

not capture tails of the distribution—households who have a long stream of good (or bad) luck. But it does captures market incompleteness by one parameter through which, as we shall see, "history matters"—even though for just one period. The simplifications "buy" us the ability to compute optimal policy. In that sense, we see our framework as complementary to fully-specified HANK models, and as an extra step in the direction of a synthesis—since in a fully-fledged HANK-type model solving the Ramsey policy would require keeping track of a state variable (wealth) that has infinite dimension.\textsuperscript{22}

3 Inspecting HANK Transmission: Inequality, Aggregate Demand, and Liquidity

In this section, we use our simple and tractable model to shed light on some of its properties that are key for understanding optimal monetary policy. First, we assume that liquidity provision is exogenous—the central bank follows a money supply (growth) rule—and ask what is the effect and transmission of a liquidity injection when prices are sticky. Then, we study whether (endogenous) liquidity provision can be used to neutralize the effect of aggregate shocks on inequality, that is to provide insurance. If so, what are the inflationary consequences of such a policy? Before exploring these questions we provide a local approximation of the model around a steady state with zero inflation $\pi = 0$ (a summary of all loglinearized equilibrium conditions around an arbitrary inflation rate is in Appendix B). Denote log-deviations of any variable by a hat, unless specified otherwise.

The Euler equation of participants and the self-insurance equation are given by, respectively

$$\hat{c}_t^P = E_t \hat{c}_{t+1}^P - \gamma^{-1} (i_t - E_t \hat{\pi}_{t+1})$$ \hspace{1cm} (11)

$$\hat{c}_t^P = \alpha \beta E_t \hat{c}_{t+1}^P + (1 - \alpha \beta) E_t \hat{c}_t^N + \gamma^{-1} E_t \hat{\pi}_{t+1}$$ \hspace{1cm} (12)

Let $\hat{x}_t \equiv (m_{t+1}^{CB} - m^{CB}) / m^{tot}$ be the deviation of new money issued by the central bank today, as a fraction of steady-state total money. The equation governing money growth is hence:

$$\hat{x}_t = \hat{m}_{t+1}^{tot} - \hat{m}_t^{tot} + \hat{\pi}_t$$ \hspace{1cm} (13)

The linearized budget constraint of non-participants is:

$$\hat{c}_t^N = \frac{1 - \alpha \mu}{1 - n \hat{h}} (\hat{m}_t^{tot} - \hat{\pi}_t) + \frac{\mu}{\hat{h}} \hat{x}_t.$$ \hspace{1cm} (14)

\textsuperscript{22}In a separate paper, we concentrate on the positive implications: we loglinearize this model, solve it in closed-form, and analyze the monetary transmission mechanism. We show that the Taylor principle fails dramatically in this economy: the central bank cannot stick to a Taylor rule that is otherwise reasonable in the representative-agent model. Augmenting the rule with inequality or a measure of liquidity restores determinacy, while a money growth rule is even better.
New money $\dot{x}_t$ reaches $N$ agents within the period (because money is issued through helicopter drops), and the Pigou effect reduces the value of their outstanding real balances. Finally, the price-setting equation is the loglinearized version of (8):\(^23\)

$$\hat{\pi}_t = \beta E_t \hat{\pi}_{t+1} + \psi \left[ (\varphi np + \gamma) \hat{c}^P_t + \hat{\phi} (1 - n) h \hat{c}^N_t - (1 + \varphi) \hat{a}_t \right]. \quad (15)$$

with $\psi \equiv \varepsilon^{-1}_r$ ranging from 0 (fixed prices) to $\infty$ (flexible prices) and $\hat{\phi} = \varphi / (1 - (1 - n) \delta_c).$\(^24\)

A local rational expectations equilibrium consists of a vector of processes $\hat{c}_t^N, \hat{c}_t^P, \hat{\pi}_t, \hat{x}_t, \hat{\pi}_t, \hat{x}_t, \hat{m}_{t+1}^{\text{tot}}$ that satisfy the equations (11) to (15). The reduced-form model is thus reminiscent of both "old" monetarist and Keynesian "dynamic ISLM" models—such as Sargent and Wallace (1975), or Sargent (1987)—yet it relies on a specification with incomplete markets and limited participation, and captures some key features of more recent HANK models. To close the model, we need to specify how monetary policy is conducted. In this section, we sketch the effects of monetary injections and analyze their transmission under the assumption that the central bank chooses money growth (under which, as in Sargent and Wallace, the equilibrium is determinate). Subsequently, we solve for the optimal policy of the central bank: liquidity will then be determined endogenously.\(^25\)

### 3.1 Inequality, Aggregate Demand, and Inflation

Combining (12) and (11), we obtain a core equation of our model, which captures the link between interest rates (the price of liquidity) and inequality in consumption defined as $\check{q}_t \equiv \hat{c}^P_t - \hat{c}^N_t$:

$$E_t \check{q}_{t+1} = E_t \hat{c}^P_{t+1} - E_t \hat{c}^N_{t+1} = \frac{\gamma^{-1}}{1 - \alpha \beta} \hat{a}_t$$ \quad (16)

This illustrates the link between monetary policy and inequality in our model: more liquidity (lower interest rates) leads to less inequality: as the opportunity cost of holding liquidity falls, P agents hold more of it, leading to higher consumption for N (and lower for P) agents tomorrow: hence,

---

\(^23\)We used that aggregate labor supply is proportional to participants’ labor supply $\dot{l}_t = \dot{l}^P_t$; the linearized labor supply equation $\varphi \dot{l}_t = \dot{w}_t - \gamma \dot{c}^P_t$, the economy resource constraint $\dot{c}_t = (1 - (1 - n) \delta_c) \left( \dot{l}_t + \dot{a}_t \right)$ with aggregate consumption (denoting $p \equiv e_t / c = q_h$): $\dot{c}_t = np \dot{c}^P_t + (1 - n) h \dot{c}^N_t$. Finally, $\hat{\phi} = \varphi / (1 - (1 - n) \delta_c).$

\(^24\)Both agents consume the output, so movements along the labor supply curve concern them both: thus, the real wage depends on total consumption with elasticity $\varphi$. However, since only participants work, they are the only ones subject to the income effect: thus, the real wage depends only on consumption of participants with elasticity equal to the income effect $\gamma$. This generates an asymmetry in the inflationary effects of consumption of the two agents.

\(^25\)In a separate paper, we analyze the other case: monetary policy via Taylor-type interest rate rules, and ask: how does a central bank ensure equilibrium determinacy in an incomplete-markets economy where liquidity (money creation) $\dot{x}_t$ is endogenous. It turns out that endogenous fluctuations in precautionary liquidity seriously challenge the central bank’s control of aggregate demand and question the appropriateness of Taylor rules. For moderate market incompleteness, the Taylor coefficients required for determinacy are in the double digits—but responding to inequality or liquidity can restore conventional wisdom in the form of the "Taylor principle"
lower future inequality. This effect is stronger, the more intertemporal substitution there is (higher $\gamma^{-1}$) and the higher is $\alpha$. This equilibrium outcome of our model is consistent with the empirical findings documenting a positive correlation between expansionary, inflationary monetary policy and redistribution (less inequality); see for example Doepke and Schneider (2006), Adam and Zhu (2014), and Coibion et al (2013).

**Monetary-NK IS curve.** Aggregate demand in our economy is made of the demand of the two types, participants and nonparticipants. The demand of participants is determined by an Euler equation, but in contrast to the standard RA model (and to models with hand-to-mouth agents) this Euler equation includes an insurance/precautionary saving motive (12). That equation thus links the two components of aggregate demand: participants’ (employed agents) and non-participants (unemployed). The demand of non-participants is determined by the previous accumulation of money balances, and by the money transfer received, as in (14). Inflation has an impact on both agents’ demand: realized inflation reduces the real value of money balances (and hence, the income and consumption) of N agents, while expected future inflation influences the insurance decision of P agents.

Combining (14), and (13), we obtain an equation linking the aggregate demand of N agents to money transfers and inflation:

$$
(1 - n)\ h\ c_t^N = \mu (\alpha - n)\ x_t + \mu (1 - \alpha)\ h\ c_{t+1}^N \\
= \mu (1 - n)\ h\ c_{t+1}^N - \mu (\alpha - n)\ (h\ c_t^N - \tilde{\pi}_t) 
$$  \hspace{1cm} (17)

The key measure of market incompleteness in our model is:

$$\alpha - n = (1 - n)\ (\alpha + \rho - 1) > 0,$$

which captures the direct effect on non-participants’ demand of an increase in liquidity $x_t$. Indeed, $\alpha - n$ captures the idea that the *conditional* probability of remaining $P$ is higher than the *unconditional* probability of becoming $P$, i.e. the share of $P$ in the total population. The parameter thus measures the incumbents’ advantage, the "memory" of the process, or the trials’ not being independent: $\alpha > 1 - \rho$ implies that it is more likely for an employed agent to keep their job than for an unemployed agent to become employed, which is a natural restriction. In equilibrium, $\alpha - n$ is hence the elasticity (integrated across all N agents) of N agents’ consumption to a monetary transfer (for given future real money balances). The same parameter captures also the elasticity of N’s aggregate demand to inflation, for given real money balances—that is, the Pigou effect discussed previously.

The aggregate IS curve of our economy puts this (17) together with the Euler equation of participants (11) and the expression of aggregate consumption $\tilde{c}_t = np\ h\ c_t^P + (1 - n)\ h\ c_t^N$. 

16
\[ \hat{c}_t = E_t \hat{c}_{t+1} - np \gamma^{-1} (\hat{i}_t - E_t \hat{\pi}_{t+1}) + \mu (1 - \alpha) E_t \hat{\pi}_{t+1} \]
\[ + \mu (\alpha - n) \hat{x}_t - \mu (1 - n) E_t \hat{x}_{t+1} \]  
\hspace{1cm} (18)

Through what we could call the monetary-New Keynesian IS curve (18), aggregate demand depends on money (liquidity), interest, and prices (inflation)—hat tip to Patinkin (1956). There are three main differences with respect to the aggregate IS curve of a standard representative-agent economy, corresponding to these three components.

First, money (liquidity) creation affects aggregate demand directly, through its impact on aggregate demand of \( N \) agents discussed in detail above. This effect is proportional to \( \alpha - n > 0 \), which captures market incompleteness in our model as explained above.\(^{26}\)

Second, the interest-elasticity of aggregate demand is lower than in a representative-agent economy: \( np \gamma^{-1} < \gamma^{-1} \) and decreasing with the share of constrained agents. This is the opposite with respect to a model in which nonparticipants are employed, and hence have endogenous labor income.\(^{27}\) In that model, the interest elasticity of aggregate demand is increasing with the share of hand-to-mouth nonparticipants: in response to a cut in interest rates, demand expands, labor demand shifts, and the wage increases; the income of the constrained increases, leading to a further amplification on demand. We abstract from these (by now well understood) considerations and focus on the role of money for self-insurance against unemployment risk, thus assuming that all nonparticipants are unemployed.\(^{28}\)

Lastly, expected inflation matters for aggregate demand over and above its effect through the ex-ante real interest rate. Higher expected inflation creates more demand today (through \( \mu (1 - \alpha) E_t \hat{\pi}_{t+1} \)) by intertemporal substitution, because it diminishes the real value of liquidity tomorrow. This expected inflation channel is "as if" \( N \) agents were at the zero lower bound permanently. Combining the last two effects (at given nominal interest), it is clear that expected inflation has asymmetric effects on aggregate demand.

\(^{26}\)A particular exogenous redistribution through transfers \( n \tau^P = \alpha \tau \) is one instance of a Wallace (1981)-type "constant fiscal policy" that will undo the effect of monetary injections; we study this issue in more detail in our companion paper.

\(^{27}\)See Bilbiie (2004, 2008) for a full analysis of a cashless model with employed nonparticipants (including the case where at high values of the share of non-participants, the interest elasticity of aggregate demand changes sign). See Gali, Lopez-Salido, and Valles (2004, 2007) for related models with hand-to-mouth agents focusing on different issues. Eggertsson and Krugman (2012) use a similar aggregate demand structure to analyze deleveraging and liquidity traps.

\(^{28}\)This is more in line with McKay, Nakamura and Steinsson (2015). See also Bilbiie (2017) for further discussion of the difference between the two aggregate demand models and their different implications for forward guidance.
3.2 Closed-form solution with horizontal aggregate supply (fixed prices)

Rewriting the self-insurance equation using the definition of inequality and replacing the budget constraint of N agents, we obtain and equation that links inequality to present and future liquidity, and expected inflation:

\[ q_t = \alpha \beta E_t q_{t+1} + \frac{\alpha - n \mu}{1 - n \frac{h}{h}} \tilde{x}_t + \frac{\mu}{h} E_t \tilde{x}_{t+1} + \left( \gamma^{-1} - \frac{1 - \alpha \mu}{1 - n \frac{h}{h}} \right) E_t \tilde{\pi}_{t+1}, \]  

(19)

Expected inflation has two opposing effects on present inequality, keeping future inequality (and hence the nominal interest rate) fixed. On the one hand, it tells P agents to consume more today, for money will have a lower payoff tomorrow—this is driven by intertemporal substitution. On the other hand, the Pigou effect on N agents tomorrow tells P agents (who might become N tomorrow) to save more for precautionary reasons, i.e. hold more liquidity and consume less—an income effect that reduces inequality today. With log utility the elasticity to inflation is positive and less than unity, namely \(0 < \frac{\delta}{\theta} < 1\) (as required by positive steady-state money demand).

Since under a money growth rule liquidity is exogenous, we can solve for the entire path of inequality and inflation using equations (19) and (15), appropriately rewritten as:

\[ \tilde{\pi}_t = \beta E_t \tilde{\pi}_{t+1} + \psi \left( (\tilde{\varphi} np + \gamma) q_t + (\tilde{\varphi} + \gamma) \left( \frac{1 - \alpha \mu}{1 - n \frac{h}{h}} (\tilde{m}_t^{ot} - \tilde{\pi}_t) + \frac{\mu}{h} \tilde{x}_t \right) \right) - (1 + \varphi) \tilde{a}_t, \]

where \(\tilde{a}_t\) is the log-deviation of the technology level \(A_t\). The money equation then determines the path of real money balances, while the nominal interest rate is proportional to expected future inequality as explained above in (16). Note that there is a liquidity effect if expected inequality falls when issuing money.

Solving the model in closed form is not possible in the general case under a money growth rule; this is a property shared with even the simplest textbook NK model with money demand under a money growth rule, e.g. Gali (2008). To obtain closed-form solutions that help our understanding of the model, we take the extreme case of fixed prices \(\psi = 0\).\(^{29}\) Under this assumption, inflation does not move \(\tilde{\pi}_t^{f} = 0\), so one equation (19) is enough to determine equilibrium, which is locally unique because \(\alpha \beta < 1\); solving it forward under the assumption that money growth is AR(1), \(E_t \tilde{x}_{t+1} = \theta \tilde{x}_t\), we obtain:

\[ q_t^{f} = -\left( \frac{\alpha - n}{1 - \alpha \beta \frac{h}{h}} \right) \tilde{x}_t \]

(20)

The response depends on our key parameter, \(\alpha - n\): the larger it is, the larger the effect of liquidity on aggregate demand (through demand of the constrained), and the larger the ensuing fall in inequality. If the shock is "too" persistent, inequality can increase though: agents correctly anticipate the future windfall and self-insure less. The path of nominal interest rates is immediately determined through

\(^{29}\)A closed-form solution can also be obtained in the other polar case of flexible prices, \(\psi \to \infty\), but without obtaining much additional intuition for our purposes in the case of exogenous liquidity.
in particular, there is a liquidity effect (interest rates fall) if and only if 1. the increase in liquidity is persistent but 2. not too persistent, so that inequality goes down: $0 < \theta < \frac{\alpha - n}{1 - n}$.

3.3 Perfect insurance and inflation

In order to help our intuition for the full optimal policy considered next, it is instructive to consider the mechanics of endogenous liquidity, given an allocation. In particular, we consider the endogenous path of liquidity $\hat{x}_t$ necessary to implement two specific allocations when the economy is hit by aggregate shocks $a_t$. The first allocation we consider is the perfect-insurance benchmark—the first-best limit of our economy. We compare this with a policy of perfectly stabilizing inflation. Subsequently, we conduct a rigorous Ramsey-optimal policy exercise—but the purpose here is to elucidate the mechanism at work using simple closed-form expressions.

Consider first the policy implementing perfect insurance ($q_t = 0$) under flexible prices and starting from steady state with $q = 1$ ($h = p = 1$). Assuming further log utility, the solution for inequality in the simplest iid case is, defining $\Omega \equiv \frac{1+\varphi}{1+\varphi} (1 - \delta_c) > 0$:

$$q_t = -\frac{\alpha - n}{1 - \alpha} \Omega \hat{x}_t + \delta_c a_t.$$ 

Denoting with a double star the economy with no inequality variations $q_t^{**} = 0$, the level of endogenous liquidity injection that achieves this is:

$$\hat{x}_t^{**} = \frac{1 - \alpha}{\alpha - n \Omega} \delta_c a_t,$$

which is always positive because $1 > \delta_c > 0$. It is decreasing with $\alpha - n$ because, as shown above, $\alpha - n$ captures the elasticity of inequality and aggregate demand to liquidity; the higher this elasticity, the lower the necessary liquidity injection.

One key implication of this policy is that inflation varies, namely:

$$\frac{d\pi_t^{**}}{da_t} = \frac{1}{\Omega} \left( \frac{1-n}{\alpha-n} \delta_c - 1 \right)$$

$$\frac{dE_t \pi_{t+1}^{**}}{da_t} = \frac{1 + \varphi}{1 + \varphi}$$

In particular, expected inflation gives agents the right intertemporal incentives to hold the extra liquidity for self-insurance purposes. Thus, the equilibrium is one with inflation volatility. Finally,

---

30The result that sticky-price models deliver a liquidity effect is emphasized by Christiano, Eichenbaum, and Evans (2005). The same authors compared sticky-price and limited-participation models’ ability to deliver a liquidity effect in previous work—see Christiano (1991), Christiano and Eichenbaum (1992, 1995); see also Fuerst (1992).

31Notice that the welfare objective is not merely perfect insurance in deviations; as our objective function derived below shows clearly, there is a rationale for increasing $c^N$ in levels.
the consumption of $N$ agents under this policy is increased by the liquidity injection but decreased by the Pigou effect if there is current inflation, for a total effect of

$$\frac{dc_t^{N**}}{da_t} = \frac{1 + \varphi}{1 + \bar{\varphi}}.$$ 

Consider now the other extreme, of **strict inflation targeting**: a policy (denoted by superscript $0$) of stabilizing inflation around the perfect-insurance steady-state (note that the degree of price stickiness plays no role as inflation is constant). Imposing $\frac{d\tilde{\pi}_t^{0}}{da_t} = 0$ at all times, we obtain that inequality is inversely directly related to liquidity, $q_t^0 = -\frac{\mu}{\mu} \tilde{x}_t^0$, and liquidity and the consumption of $N$ agents in this equilibrium are:

$$\frac{dc_t^{N0}}{da_t} = \mu \frac{d\tilde{x}_t^{0}}{da_t} = \frac{1 + \varphi}{\bar{\varphi} (1 - \mu)}.$$ 

It then follows immediately by direct inspection that $dc_t^{N0} > dc_t^{N**}$: consumption of $N$ agents responds more, and is thus more volatile under the zero-inflation policy. In other words, inflation is a means to insure constrained agents against aggregate risk. We will see that this general insight holds more broadly when we analyze the policy trade-offs rigorously by means of a Ramsey-optimal policy analysis.

### 4 Optimal Monetary Policy: Inequality and Inflation

Optimal monetary policy needs to find the balance between two main distortions in this economy. The first is **inequality**: a scope for providing insurance, specific to an incomplete-markets, limited-participation setup like ours. The second is **costly price adjustment**: the standard distortion that operates in a representative-agent NK model. This section analyzes how this trade-off is resolved in our model. We solve the full Ramsey-optimal policy, provide a second-order approximation à la Woodford (2003) that is useful to understand the policy trade-offs, and analyze optimal policy quantitatively. The general theme is that inequality implies large optimal deviations from price stability.

#### 4.1 Ramsey-Optimal Policy

Following a long tradition started by Lucas and Stokey (1983), we assume that the central bank acts as a Ramsey planner who maximizes aggregate welfare. In our economy, this entails calculating the welfare of the two agents and weighting them by their shares in the population. The constraints of the planner are the rearranged private equilibrium conditions: self-insurance (5), the Phillips curve (8), and the economy resource constraint (10).\(^{32}\) We denote the system of these three constraints by

\(^{32}\)Since the nominal interest rate enters only the Euler equation for bonds, the problem can be regarded as one where the planner chooses the allocation directly; once the consumption of participants and inflation are known, the optimal
\( \Gamma_t(c^P_t, c^N_t, l^P_t, \pi_t) \). As it is by now well understood, the optimal policy problem of the central bank can be written as choosing the allocation \( \{c^P_t, c^N_t, l^P_t, \pi_t\} \) to maximize the following Lagrangian:

\[
\max_{\{c^P_t, c^N_t, l^P_t, \pi_t\}} E_0 \sum_{t=0}^{\infty} \beta^t \left\{ n \left[ u(c^P_t) - \chi \left( \frac{l^P_t}{1 + \varphi} \right)^{1+\varphi} \right] + (1 - n) \left[ u(c^N_t) - \chi \left( \frac{\delta}{1 + \varphi} \right) + \omega_t \Gamma_t \right] \right\} \tag{21}
\]

where \( \omega_t \) is the vector of three costate, Lagrange multipliers, one for each constraint in \( \Gamma_t \). The first-order conditions of this problem are outlined in Appendix B.4, as is the proof of the following Proposition.

**Proposition 1** *The optimal long-run inflation rate is such that \( \beta - 1 \leq \pi^* \leq 0 \).*

As in other NK models incorporating different theories of money demand (e.g. Khan et al, 2003; Schmitt-Grohe and Uribe, 2004, 2007), the long-run inflation rate ranges from the Friedman rule under flexible prices and optimal subsidy, to zero inflation under sticky prices and inelastic labor. A low inflation rate allows agents to self-insure, but generates price adjustment costs. An inflation rate close to 0 minimizes price adjustment costs, but decreases the ability of agents to self-insure, as the return on money decreases.

But unlike in other NK models, including those incorporating money demand, in our economy the central bank also uses inflation optimally over the cycle—to provide insurance and decrease inequality. We first illustrate formally the trade-off faced by a central bank by deriving a second-order approximation to the aggregate utility function, which contains an "inequality" motive, and we then explore the quantitative significance of this novel trade-off.

### 4.2 A second-order approximation to welfare

To understand the relevant policy trade-offs, we derive a second-order approximation à la Woodford (2003, Ch. 6) to the aggregate welfare function, around a steady-state with imperfect insurance \((\bar{p} > 1 > \bar{h})\), an optimal subsidy inducing marginal-cost pricing in steady state \((\Phi = 0 \text{ in (8)})\), and arbitrary steady-state inflation.\(^{33}\) In Appendix B.5 we prove the following:

**Proposition 2** *Solving the welfare maximization problem is equivalent to solving:*

\[
\min_{\{c^P_t, c^N_t, q_t, \pi_t\}} \frac{1}{2} E_0 \sum_{t=0}^{\infty} \beta^t \left\{ \Lambda_x \tilde{\pi}_t^2 + \Lambda_c \tilde{c}_t^2 + \Lambda_q \tilde{q}_t^2 - \Lambda \left[ \tilde{c}_t^N + \frac{1 - \gamma}{2} (\tilde{c}_t^N)^2 \right] \right\}
\]

The interest rate is determined by the Euler equation. By similar reasoning, once the consumption of non-participants is determined, along with inflation, the quantity of real money balances is fully determined too. These simplifications apply only when money is issued via helicopter drops; our companion paper compares this with open market operations implementation; see Appendix B.4.

\(^{33}\)Bilbiie (2008) also derives a quadratic loss function in a cashless economy with hand-to-mouth agents. Curdia and Woodford (2009) and Nistico (2015) also do this for cashless models with infrequent access to credit markets; unlike us, they focus on an efficient equilibrium with insurance when calculating optimal policy.
where the optimal relative weights are:

\[ \Lambda_\pi = \frac{\nu}{1 - \frac{\nu}{2} \pi^2}; \quad \Lambda_c = \gamma - 1 + \frac{1 + \varphi}{1 - (1 - n) \delta_c} \]

\[ \Lambda = 2 (1 - n) h (q^\gamma - 1); \quad \Lambda_q = (\gamma - 1) np (1 - n) h. \]

The Proposition transparently illustrates the novel trade-off implied by our framework between inequality and stabilization (of inflation and aggregate demand).\(^{34}\) The last two terms pertain to inequality. The last one represents the first-order welfare benefit of increasing N agents’ welfare by increasing their consumption level—intuitively, its weight \( \Lambda \) is proportional to the steady-state inequality distortion captured by \( (q^\gamma - 1) \), since the first-order benefit exists only insofar as the steady state is distorted to start with. This is analogous to the linear benefit of increasing output above the natural rate when the steady-state is first-order distorted in the standard New Keynesian model (see the next footnote). The distortion vanishes when the steady-state is egalitarian \( (p = h, \text{perhaps through a steady-state insurance scheme, if enough fiscal lump-sum instruments are available to undertake such policy}) \) or, trivially, when \( n = 1 \) (the standard cashless representative-agent NK model). Replacing the equilibrium \( q \) reveals that \( \Lambda \) is proportional to \( \beta^{-1} - 1 + \pi \): the distortion also becomes arbitrarily small when the steady state tends toward the Friedman rule, \( \pi \to 1 - \beta^{-1}. \)

The other new term is \( \Lambda_q q^2_t \), which captures the welfare cost of the volatility of inequality (naturally, this drops out if all agents are identical, \( n = 0 \) or \( n = 1 \)).

Because of the linear term in the loss function, second-order terms in the private constraints matter for welfare\(^{35}\): as long as there is steady-state inequality, inflation and aggregate demand volatility will matter for welfare beyond their direct effects through \( \Lambda_\pi \) and \( \Lambda_c \). The reason is by now intuitively clear: when the steady state has \( q > 1 \), increasing the consumption of non-participants provides a first-order welfare benefit; the only way to achieve this benefit, absent fiscal instruments, is monetary. As we show next, a quantitative analysis of optimal policy in a calibrated version of our model suggests that inflation volatility is desirable in this framework. Pursuing price stability instead, even around an optimally chosen inflation target, has large welfare costs.

\(^{34}\)Note that \( \tilde{\pi}_t = \pi_t + \pi \) is the inflation level, so the function is written so that target inflation is zero absent aggregate shocks, \( \tilde{\pi}_t = 0 \). The optimal target is the optimal long-run inflation found in the Ramsey problem above, the equivalent of which is here the steady state of the solution of the relevant linear-quadratic problem.

\(^{35}\)This is analogous to the linear benefit of increasing output above the natural rate when the steady-state is first-order distorted in the standard New Keynesian model. See Woodford (2003; Ch. 6), Benigno and Woodford (2005, 2012) and Schmitt-Grohe and Uribe (2007) for an analysis of this when the distortion pertains to monopolistic distortion, i.e. \( \Phi > 0 \), including explanations of the second-order corrections that are necessary to correctly evaluate welfare. The quadratic term \( \frac{1 - \gamma}{2} (\tilde{\epsilon}_t^N)^2 \) represents merely a second-order correction.
4.3 Optimal deviations from price stability and inequality: a quantitative evaluation

To analyze optimal policy in our model we calibrate the model at quarterly frequency. We follow, for common parameters pertaining to preferences and the supply side, the classic papers in optimal policy in NK models, Khan, King, and Wollman (2003) and Schmitt-Grohé and Uribe (2007): the inverse elasticity of labor supply is $\varphi = 0.25$, and $\gamma = 1$. The elasticity of substitution between goods is $\varepsilon = 6$, and we introduce the steady-state subsidy $\sigma = 1/(\varepsilon - 1)$ to avoid steady-state distortions due to monopolistic competition—thus isolating our novel channel as a motivation for deviations from price stability. Both cited papers use different models of staggered pricing and assume that prices stay unchanged on average for 5 periods; this implies a Phillips curve slope (our $\psi$) of around 0.05. Given our $\varepsilon$, the price adjustment cost parameter that delivers the same $\psi$ is $\nu = 100$. The discount factor is $\beta = 0.98$, as in other studies with heterogeneous agents (Eggertsson and Krugman, 2012; Curdia and Woodford, 2009); we consider larger values for robustness below. We use the same labor productivity process as Khan et al, with autocorrelation 0.95 and standard deviation 1%.

Three parameters pertain to market incompleteness and money demand: the probabilities to keep participating ($\alpha$) and non-participating ($\rho$), and home production when non-participating (or unemployment benefits) $\delta$. Since we perfectly correlated financial market and labor market participation to obtain our tractable model, two calibrations are possible: one that targets financial market participation and money demand, and the other labor market variables. We use the former as a benchmark and report the latter for robustness.

We target three data features in our benchmark calibration. First, the number of participants $n$: in the US economy roughly half of the population participates in financial markets, either directly or indirectly (Bricker et al, 2014), and this is stable over time. We thus take $n = 0.5$, which implies the restriction $\alpha = \rho$. Second, the velocity of money (roughly speaking, $\mu^{-1}$ in our notation): considering a broad money aggregate, the quarterly velocity ($GDP/M2$) is around 2 over the period 1982—2007 (chosen to avoid the zero lower bound period). Third, consumption inequality $q$ between participating and non-participating agents captures the lack of insurance due to market incompleteness. Since agents participate infrequently in financial markets (Vising-Jorgensen, 2002) and one cannot keep track of their participation status, it is hard to find an exact empirical counterpart to $q$. We take as a proxy the fall of nondurable consumption when becoming unemployed, which is estimated between 10% and 20% (see e.g. Chodorow-Reich and Karabarbounis, 2014) and target the conservative value of 10% for this object ($q^{-1} - 1$ in our model). These three targets jointly imply $\alpha = \rho = 0.9$, and $\delta = 0.783$. Table 1 presents our parameters and the implied Ramsey steady-state values for our target variables—which are determined by the exact Ramsey equilibrium conditions outlined in the Appendix.
Optimal long-run deviations from price stability

The optimal asymptotic (steady-state) inflation rate is $\pi = -0.79\%$. As expected, this is higher than the inflation implied by the Friedman Rule (which is $-2\%$), because prices are sticky, just as in standard monetary models with sticky prices, e.g. Khan et al (2003). More equilibrium deflation occurs if prices are more flexible, labor is more elastic, and $\alpha$ is higher. The first two elements are standard (the former was first noticed by Chari Christiano Kehoe, 1997; see also Schmitt-Grohe and Uribe, 2004). The last part has a standard interpretation too: at given $n$, higher $\alpha$ implies more elastic money demand. As we will show below, less elastic money demand (lower $\alpha - n$) implies less optimal deflation—as in Khan, King and Wollman, although for a different theory of money demand.

<table>
<thead>
<tr>
<th>Preferences</th>
<th>Production and price setting</th>
<th>Heterogeneity</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>$\gamma$</td>
<td>$\varphi$</td>
</tr>
<tr>
<td>0.98</td>
<td>1</td>
<td>0.25</td>
</tr>
<tr>
<td>$\varepsilon$</td>
<td>$\nu$</td>
<td>$\rho^a$</td>
</tr>
<tr>
<td>6</td>
<td>100</td>
<td>0.95</td>
</tr>
<tr>
<td>$\sigma_a$</td>
<td>$\alpha$</td>
<td>$\rho$</td>
</tr>
<tr>
<td>0.01</td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td>$\delta$</td>
<td></td>
<td>0.78</td>
</tr>
</tbody>
</table>

Model outcome

<table>
<thead>
<tr>
<th>GDP/M2</th>
<th>$n$</th>
<th>$(c^N - c^P)/c^P$</th>
<th>$\pi$</th>
<th>$c^p$</th>
<th>$c^N$</th>
<th>$l$</th>
<th>$n_t^{tot}$</th>
<th>$q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>50%</td>
<td>-11%</td>
<td>-0.79%</td>
<td>0.98</td>
<td>0.87</td>
<td>1.07</td>
<td>0.46</td>
<td>1.12</td>
</tr>
</tbody>
</table>

Table 1: Baseline calibration

What is the welfare cost of inflation? This is a classic question in monetary economics, going back at least to Bailey’s 1956 calculation. The welfare cost of (steady-state) inflation in our framework can be easily calculated, in the Lucas (1987) tradition—we review this briefly in the Appendix. We find that moving from a steady-state annualized inflation rate of 2% (0.5% quarterly) to the optimal rate of $-3.2\%$ ($-0.79\%$ quarterly) is equivalent to a permanent increase in consumption of 0.61%—this is in line with numbers found in the literature, e.g. Lucas (2000) and Imorohoglu (1992), although slightly larger.

Our model’s implications for optimal policy in the long run are thus rather standard. But in the short run, things are different and this is intimately related to inequality in our model: with long-run inequality ($q = 1.12$), the steady state is distorted and this has implications for short-run optimal policy.

36Brunnemeier and Sannikov (2016) provide an example of a flexible-price monetary model where the Friedman rule is not optimal, because there is a distorted portfolio decision between money and physical capital.

37A large literature analyzed this question using a variety of frameworks. To cite just some prominent examples, Lucas (2000) found that reducing inflation from 10 to 0 percent annually results in a 1 percent increase in consumption. Analyzing a monetary framework closer in spirit to the one our model embeds (based on the Bewley model), Imorohoglu (1992) showed that the welfare effects of inflation are larger in incomplete-markets economies. See Doepke and Schneider (2006), Erosa and Ventura (2002) and Ragot (2014) for reviews.
Optimal short-run deviations from price stability

The inequality channel requires the central bank to accommodate some inflation volatility, as doing otherwise leads to large welfare losses. This is true in our economy even when the source of business cycles is a shock that, in the standard NK model with money demand but no inequality, generates no such trade-off: a plain vanilla labor productivity shock.

Recall what happens in the baseline NK model in response to this shock: not much. A welfare-maximizing central bank keeps prices unchanged and inflation at zero, as this shock creates no trade-off: the central bank can close the output gap costlessly, a well-known result labeled "divine coincidence" by Blanchard and Gali (2007). This result changes but only slightly when the steady state is distorted ($\Phi > 0$ in our notation), as analyzed in detail by Benigno and Woodford (2005): productivity shocks then have a "cost-push" dimension, creating a trade-off. Quantitatively, however, this is moot—subject to one caveat mentioned in the next footnote. The same is true in models incorporating a variety of other frictions—in particular, in models with monetary frictions such as Khan, King, and Wollman (2003) and Schmitt-Grohe and Uribe (2004, 2007): price stability is a robust policy prescription. Even though these models do imply inflation volatility under the optimal Ramsey policy, the welfare cost of eliminating such volatility is generically negligible.\(^{38}\)

This is no longer the case in our model: optimal Ramsey policy requires volatile inflation, and this volatility matters for welfare. To see the first part of this argument, consider the impulse responses to a productivity shock presented in Figure 1, for three economies. With a black solid line, we have our economy under optimal policy—obtained by solving (21). With a blue dashed line, we have our monetary economy under what we label "Strict inflation targeting" (SIT): the central bank perfectly stabilizes inflation around the Ramsey-optimal steady state inflation (this is implemented by a Taylor rule with large $\phi_\pi$ and the optimal $\pi^*$ target). Finally, we show with a red circle line optimal policy in a standard cashless equilibrium,\(^ {39}\) a comparison with which illustrates the extent of risk-sharing provided by money in our model. All variables are in percentage deviation from steady state, except the inflation and interest rates, which are in deviation from steady state.

\(^{38}\)See for instance Table 2 in Schmitt-Grohé and Uribe (2007); see also Bilbiie, Fujiwara, and Ghironi (2014) for a result on optimal short-run price stability in a model with entry and variety, and a review of the literature using other distortions. As Benigno and Woodford (2005) show analytically, this result changes—price stability ceases to be optimal—if, on top of $\Phi > 0$, the share of government spending in steady-state output is also non-zero.

\(^{39}\)Since in the non-monetary equilibrium the steady-state inflation rate is 0, we recalibrated it to have the same steady state allocation. In particular, we reduce output by $\frac{\Phi}{2} \pi^2$ and introduce a transfer between $N$ and $P$ households, such that the steady-state consumption and labor supply are the same in the monetary and non-monetary equilibrium, and only the steady-state inflation is different.
Figure 1: Responses to a labor productivity shock under optimal Ramsey policy in our model (solid black), strict inflation targeting in our model (blue dash), and optimal policy in cashless model (red circles).

The responses of the cashless model are standard: inflation does not move, and output is equal to its natural rate. Since labor productivity affects only P agents, their consumption increases (and so does inequality), and the nominal interest rate goes down.

In our monetary economy, the planner reduces inequality by providing insurance: compared to the red circle line, the black solid line shows that the consumption of N agents increases, and inequality decreases. The planner issues money and interest rates fall; the result is inflation (due to the demand effect on firms), which erodes N agents’ purchasing power (money balances) via the Pigou effect.

Consider now the allocation when this inflation is absent (blue dashed line): more money is issued, and the real value of balances is much higher: thus, the consumption of N agents responds more, and is more volatile. Since the consumption of P agents is largely unchanged, the same is true for inequality. We will now show that this extra volatility is costly in terms of aggregate welfare.\footnote{\textsuperscript{40} It is by now well known, starting with the influential paper of King and Wolman (1999), that welfare calculations depend crucially upon the initial values of the Lagrange multipliers—which can be set to 0, or to their Ramsey steady-state values. Under the former choice, policy is not timeless-optimal: initial period $t_0$ inflation has no consequence for prior expectations, thus the policy chosen in any later period is not a continuation of $t_0$ policy. In the second case, policy is timeless-optimal in the sense of King and Wolman (1999) and Khan et al.(2003) (Woodford 2003 uses...}
Table 2 reports the standard deviations of the main variables for the (Ramsey-)optimal and SIT allocation. The volatility of inflation is comparable to that obtained by Khan et al (2003). Because of limited risk-sharing, $N$ agents’s consumption volatility is higher than $P$ agents’. More importantly, $N$ agents’s consumption volatility is higher under strict inflation targeting than under optimal policy—as a result, the volatility of our inequality measure is twice as large. This difference in volatilities translates into a large welfare cost of price stability (around the optimal asymptotic inflation rate): agents need to be compensated by 0.08% of consumption every period in order to live in an economy with stable prices, rather than in one with optimal policy and inflation volatility.41

This welfare cost is high, much higher than those usually found in the literature reviewed above, even though our optimal inflation volatility is of comparable size; it is of the same magnitude as the total welfare cost of uncertainty in incomplete-markets models (see Krusell and Smith, 1998, and Lucas, 2003). The reason is by now clear: our long-run equilibrium is one with inequality (imperfect insurance), which for a planner is a distortion. In standard NK models with distorted steady state, this is not enough to generate significant costs of price stability. Here, it is, because volatility has a first-order effect through the level of $N$ agents’ consumption. In terms of our second-order approximation, this effect makes it "as if" the weight on inflation volatility in true Ramsey loss function were smaller than $\Lambda_\pi$.

<table>
<thead>
<tr>
<th>Economies</th>
<th>$\hat{c}^P$</th>
<th>$\hat{c}^N$</th>
<th>$\hat{q}$</th>
<th>$\pi$</th>
<th>$\Delta W$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>2.6</td>
<td>3.3</td>
<td>0.6</td>
<td>0.05</td>
<td>–</td>
</tr>
<tr>
<td>Ramsey</td>
<td>2.6</td>
<td>3.7</td>
<td>1.2</td>
<td>0</td>
<td>0.08</td>
</tr>
<tr>
<td>SIT</td>
<td>2.5</td>
<td>3.2</td>
<td>0.6</td>
<td>0.05</td>
<td>–</td>
</tr>
<tr>
<td>SIT</td>
<td>2.5</td>
<td>3.8</td>
<td>1.3</td>
<td>0</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Table 2: Standard deviations and welfare losses (percent)

That small inflation volatility translates into high welfare gains in our model is due not so much to volatility itself as to inequality. To illustrate this, consider an economy where the steady-state

different definition). The numbers we report are for the former, $t_0$-optimal case; in the timeless-optimal case, the welfare losses are very close to zero in all cases; see also Bilbiie, Fujiwara, and Ghironi (2014) for further discussion in a different context.

41 The welfare losses are similar for the two types of agents: thus, our simplified heterogeneity misses some of the distributional effects emphasized by Krusell and Smith (1998)—in their framework and the subsequent literature (see Lucas 2003 for a review) the welfare benefits of eliminating uncertainty are asymmetric among the poor, the rich, and the middle class. One would expect that such an asymmetry occurs in a richer model of the wealth distribution also for the welfare costs that we calculate.
inequality distortion vanishes, \( q \to 1 \), which amounts to taking \( \beta \to 1 \) and re-calibrating \( \delta \) to get the same steady-state money velocity; evidently, the optimal long-run inflation rate converges to 0. As the bottom panel of Table 2 illustrates, the volatility of inflation under Ramsey policy in this economy is unchanged. Nevertheless, this volatility no longer means welfare: without an inequality trade-off, the central bank can safely and costlessly pursue price stability (just as in Khan et al., 2003, and Schmitt-Grohe and Uribe, 2007).

How do simple interest rate rules perform in this framework? To answer this, we optimized over simple rules of the form studied in the previous section, following the method outlined by Schmitt-Grohe and Uribe (2007). An inequality-augmented rule with \( \phi_q = 0.1 \) reproduces the same allocation as strict inflation targeting, but with a much smaller inflation response (\( \phi_\pi = 3 \)), which is desirable for determinacy.\(^{42}\)

**Robustness**

As a first robustness check we report the same outcomes for economies with more flexible prices \((\nu = 50)\) and less elastic labor \((\varphi = 1)\).\(^{43}\) The upper panel of Table 3 contains the results.

<table>
<thead>
<tr>
<th>Economies</th>
<th>( \pi^{SS} ) (%)</th>
<th>( \hat{c}^P )</th>
<th>( \hat{c}^N )</th>
<th>( \hat{q} )</th>
<th>( \pi )</th>
<th>( \Delta W )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \nu = 50 )</td>
<td>-1.54</td>
<td>2.7</td>
<td>3.2</td>
<td>0.5</td>
<td>0.08</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.6</td>
<td>3.8</td>
<td>1.2</td>
<td>0</td>
<td>0.13</td>
</tr>
<tr>
<td>( \varphi = 1 )</td>
<td>-0.6</td>
<td>2.1</td>
<td>2.7</td>
<td>0.6</td>
<td>0.04</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.0</td>
<td>3.0</td>
<td>0.9</td>
<td>0</td>
<td>0.03</td>
</tr>
<tr>
<td>Labor market calibration</td>
<td>-0.36</td>
<td>3.0</td>
<td>5.1</td>
<td>2.1</td>
<td>0.02</td>
<td>-</td>
</tr>
<tr>
<td>((\alpha = 0.95; \rho = 0.5; \delta = 0.5))</td>
<td></td>
<td>3.1</td>
<td>5.5</td>
<td>2.4</td>
<td>0</td>
<td>0.06</td>
</tr>
</tbody>
</table>

| Table 3: Robustness analysis |

Both the inflation volatility and its welfare benefit increase as prices become more flexible and labor supply more elastic. The reason is that with more flexible prices (lower \( \nu \)), the cost of using inflation is lower: in the limit, as prices become flexible, inflation essentially becomes a lump-sum tax—an insight originally due to Chari, Christiano, and Kehoe (1997) and also discussed by Schmitt-Grohe and Uribe (2004).

The second alternative calibration we consider is based on labor market risk. Instead of matching financial market variables \((n, \mu, \text{and } q)\) as in our previous calibration, we draw on the labor market

\(^{42}\)We also optimized over liquidity rules and found that they generate much lower welfare—the intuition being that they preclude liquidity provision, which is what is needed to implement Ramsey policy.

\(^{43}\)For each economy, in order to perform meaningful welfare comparisons we calibrate the discount factor \( \beta \) and home production \( \delta \), to start from the same steady state: this gives 0.973 and 0.79 for the first and 0.982 and 0.765 for the second calibration (results are similar when we keep these parameters unchanged).
literature, in particular Shimer (2005) to find parameter values for $\alpha, \rho,$ and $\delta$. At quarterly frequency, the job loss probability is $5\%$ and the average job finding probability $50\%$ for the post-war period—these two numbers imply $\alpha = 0.95$, $\rho = 0.5$ and thus $n = 0.94$; the gross replacement ratio is set to $\delta/w = 50\%$ (see also Challe and Ragot, 2014). The lower panel of Table 3 contains the results, assuming that all other parameters are as in the baseline; apart from the reported numbers, it is worth mentioning that the quarterly velocity of money is somewhat higher (2.33), and the fall in consumption when becoming unemployed is now $24\%$—in the upper range of the empirical estimates discussed above.

The optimal steady-state inflation rate is $-0.36\%$: there is less deflation than in the baseline calibration, for there is less money in circulation. This is similar to the optimal deflation rate obtained by Khan et al for their calibration with low money demand elasticity (obtained by estimating money demand over a shorter sample); indeed, since $\alpha$ is very close to $n$, our calibration also implies low money demand elasticity. The similarities go further: as in that model, optimal policy also implies lower inflation volatility under this calibration; but the parallel stops here, for this smaller volatility is still associated with a large welfare cost in our model. Agents are willing to sacrifice $0.06\%$ of consumption every period in order to live in an economy with optimally volatile inflation, rather than in an economy with stable prices. Our result thus survives even in this economy with very low unemployment risk.

All our previous calibrations assumed that there is an optimal subsidy that undoes the steady-state monopolistic distortion, $\Phi = 0$; this allows isolating the novel channel that operates in our framework. We now report one last set of robustness checks, assuming that there is no such subsidy $\sigma = 0$.

![Table 4](image)

As Table 4 shows, the welfare losses are now much larger. The two long-run distortions complement each other and generate significant losses from price stability. The notable exception is the case when there is no steady-state inequality: the welfare loss is, again, zero; as our second-order approximation showed, the linear term in the loss function disappears in this case. This result is related to Benigno and Woodford, who showed that a distorted steady state implies significant deviations from price stability only when the steady-state government spending share is non-zero. Our
framework thus identifies another channel—which, when shut off, makes price stability again optimal even when the monopolistic distortion is large. But when our channel is at work, i.e. when long-run inequality matters \((q > 1)\), the optimal deviations from price stability can be very large indeed if supply-side distortions are also an issue \((\Phi > 0)\).

5 Conclusions

In monetary policy analysis, a new synthesis looms: the integration of sticky-price, New Keynesian models and models of heterogeneous agents, incomplete markets and limited participation. This very active research area (which we reviewed in the Introduction) started in the early 2000s and is going at full speed. We hope to have contributed to these convergence efforts a fully-fledged, NK-style optimal monetary policy analysis, in a tractable framework that captures key mechanisms of heterogeneous-agent incomplete-markets models and includes a deep reason for money as a self-insurance device. Liquidity and inequality are intimately related in our model: today’s liquidity (lower interest rates) implies tomorrow’s lower inequality. Aggregate demand in this model depends on: money, or liquidity, which relaxes the constraint of non-participating agents; interest, because of intertemporal substitution by participating agents; and prices, or inflation, because a Pigou effect operates for non-participating agents, and (expected) inflation is the relevant return for holding liquidity.

Inequality is the keystone for optimal monetary policy: in this model, a novel trade-off arises between reducing inequality and stabilization of inflation and real activity. We first illustrate this trade-off analytically by means of a second-order approximation to the aggregate welfare function, in the New Keynesian tradition pioneered by Woodford (2003): there is a first-order benefit to reducing inequality, insofar as the long-run equilibrium is characterized by it.

We then move to a quantitative assessment of this trade-off: in an economy with a long-run equilibrium characterized by consumption inequality, deviations from price stability are optimal. This holds, first, in the long run: the optimal inflation target should be between zero and the Friedman rule; this is no surprise—it is true in most monetary models. But in our framework, unlike in others, it is also true in the short run. Optimal policy implies inflation volatility in response to (productivity) shocks that otherwise create no trade-off.

What is more, this volatility matters for welfare. A policy of stabilizing prices (albeit around the optimal inflation target) incurs a large welfare loss. This happens because, when there is long-run inequality, short-run volatility has a first-order effect on constrained agents: optimal policy requires giving less weight to inflation stabilization—which de facto implies giving more weight to constrained agents.

While we view our study as a step in the direction of the new synthesis that we mention at the
outset of these concluding comments, we think such efforts should continue, for much remains to be done. Our tractable framework allows the calculation of optimal policy, but it inherently misses several other, surely important links between inequality and monetary policy, as we acknowledged in Section 2.1 above. Incorporating such channels as those present in the "HANK" models reviewed above (more realistic wealth distributions; endogenous portfolio shares; nominal debt; endogenous unemployment risk; etc.) is paramount in order to attain a thorough understanding of how monetary policy works and how it should be conducted in a world where inequality matters.

6 References


Adam, K. and J. Zhu, 2014 "Price Level Changes and the Redistribution of Nominal Wealth Across the Euro Area" *Journal of the European Economic Association*


Auclert, A. 2016, "Monetary Policy and the Redistribution Channel" Mimeo Stanford


Bernanke, B. S. 2007, "The Level and Distribution of Economic Well-Being," speech delivered at the Greater Omaha Chamber of Commerce, Omaha, Nebraska

Bernanke, B. S. 2015, "Monetary Policy and Inequality", Brookings.edu


Cao, S., C. A. Meh, J.V. Rios-Rull and Y. Terajima. 2012. "Inflation, Demand for Liquidity and


Cloyne, J., C. Ferreira, and P. Surico 2015 "Monetary Policy when Households have Debt: New Evidence on the Transmission Mechanism" CEPR Discussion Paper


Cui, W. and S. Radde, 2016, "Search-Based Endogenous Asset Liquidity and the Macroeconomy", Mimeo UCL and ECB


Den Haan, W., Rendahl, P. and M. Riegler 2016, "Unemployment (Fears) and Deflationary Spirals", working paper.


Draghi, M., 2016 “Stability, Equity and Monetary Policy” 2nd DIW Europe Lecture, DIW Berlin


Friedman, M. 1969, The Optimum Quantity of Money, Macmillan


Furceri, D., P. Loungani,and A. Zdzienicka, 2016 "The Effects of Monetary Policy Shocks on Inequality" International Monetary Fund Working Paper


Iacoviello, M., 2005 House Prices, Borrowing Constraints, and Monetary Policy in the Business Cycle, American Economic Review


Kaplan, G., B. Moll and G. Violante 2015, "Monetary Policy According to HANK", mimeo Princeton and NYU


Mersch, Y., 2014, "Monetary policy and economic inequality", Keynote Speech at Corporate Credit Conference, Zurich


Nuno, G. and C. Thomas, 2017, "Optimal Monetary Policy with Heterogeneous Agents" Mimeo Bank of Spain


Rocheteau, G., P. O. Weill and T. Wong, 2015, "A Tractable Model of Monetary Exchange with Ex-Post Heterogeneity" Mimeo UC Irvine, UCLA and Bank of Canada


*Journal of Monetary Economics* 54, 6

*Brookings Papers on Economic Activity*, 301-373


Sterk, V. and S. Tenreyro, 2016, "The Transmission of Monetary Policy through Redistributions and Durable Purchases" Mimeo UCL and LSE


A Model Summary

The equations describing our model in the general case are:

\[ u' (c_t^P) = \beta E \frac{1 + \pi_{t+1}}{1 + \pi_{t+1}} u' (c_{t+1}^P) \]

\[ u' (c_t^N) \geq \beta E \left[ (1 - \rho) u' (c_t^P) + \rho u' (c_t^N) \right] \frac{1}{1 + \pi_{t+1}} \]

\[ c_t^P + \tilde{m}_{t+1}^P = w_t u_t^P - \tau_t^P + \frac{1}{n} d_t + \frac{\alpha}{1 + \pi_t} \tilde{m}_t^P + \frac{1 - \alpha}{1 + \pi_t} \tilde{m}_t^N \]

\[ \tilde{m}_{t+1}^N + c_t^N = \delta - \tau_t^N + \frac{1 - \rho}{1 + \pi_t} \tilde{m}_t^P + \frac{\rho}{1 + \pi_t} \tilde{m}_t^N \]

\[ \pi_t (1 + \pi_t) = \beta E_t \left[ \left( \frac{c_t^P}{c_{t+1}^P} \right)^\gamma \frac{Y_{t+1}}{Y_t} \right] \frac{1}{1 + \pi_{t+1}} + \frac{\varepsilon - 1}{\nu} \left[ \frac{\varepsilon}{\varepsilon - 1} A_t - (1 + \sigma) \right] \]

\[ Y_t = n A_t l_t^P \]

\[ w_t = \chi \left( \Pi_t^{P} \right)^{\nu} (c_t^P)^{\gamma} \]

\[ d_t = \left( 1 - \frac{w_t}{A_t} - \frac{\nu}{2} \pi_t^2 \right) Y_t \]

\[ n \tilde{m}_{t+1}^P + (1 - n) \tilde{m}_{t+1}^N = \frac{n \tilde{m}_t^P + (1 - n) \tilde{m}_t^N}{1 + \pi_t} + m_{t+1}^{CB} \]

\[ \tau_t = -m_{t+1}^{CB} \]

\[ \tau_t^P = \frac{\omega \tau_t^{N}}{n}; \tau_t^N = \frac{1 - \omega}{1 - n} \tau_t \]

\[ \tilde{m}_{t+1}^P, \tilde{m}_{t+1}^N \geq 0 \]

where \( m_{t+1}^{CB} \) is a period \( t \) monetary shock (new money created) and \( A_t \) exogenous labor productivity.

The economy resource constraint follows by Walras’ law:

\[ n c_t^P + (1 - n) c_t^N = \left( 1 - \frac{\nu}{2} \pi_t^2 \right) Y_t + (1 - n) \delta \]

and can replace (for instance) the P agents’ budget constraint in the system above.

An equilibrium of the economy is a sequence \( \{ c_t^P, c_t^N, c_t \pi_t, m_{t+1}^{CB}, \tilde{m}_{t+1}^P, \tilde{m}_{t+1}^N, w_t, l_t^P, \tau_t^P, \tau_t^N, \tau_t, d_t, Y_t \} \) satisfying the previous conditions. Assuming that nominal bonds are in zero net supply, we guess-and-verify the structure of the equilibrium with \( \tilde{m}_t^N = 0 \), i.e. non-participating households never hold money at the end of the period. The conditions for households in the N island not to hold money, which we check holds in the equilibrium we consider, is:

\[ u' (c_t^N) > \beta E \left[ (1 - \rho) u' (c_{t+1}^P) + \rho u' (c_{t+1}^N) \right] \frac{1}{1 + \pi_{t+1}}. \]
Consider the conditions for a monetary equilibrium to exist. In a monetary steady-state, $N$ agents do not hold money when $1 + \pi > \beta$, while $P$ agents save in money comparing the gain to self-insure and the opportunity cost (deflation). Thus, we have $c^P > c^N$ and $q > 1$. Since it is costly for $P$ agents to save (the return on money is lower than the discount factor), they rationally choose not to perfectly self-insure. Using this inequality in the condition (6), we have $u' (c^N) > [(1 - \rho) u' (c^P) + \rho u' (c^N)] \frac{\beta}{1 + \pi}$: $N$ agents do not hold money at the end of each period, and $\bar{m}^N = 0$. In steady state, positive money demand requires the restriction that the outside option not be too good:

$$\delta_c < \bar{\delta} = h = \left[1 + n \left(\frac{1 + \pi - \alpha \beta}{\beta (1 - \alpha)}\right)^\frac{\gamma}{\beta} - 1\right]^{-1} \quad (22)$$

Under a Taylor rule, the steady-state inflation rate $\pi$ is determined by the central bank’s target and the above condition is parametric. Under Ramsey policy, $\pi$ is endogenous (it depends, among other things, on $\delta$) and the above condition defines a threshold implicitly.

B Derivations and Proofs

B.1 New Keynesian Phillips curve

The intermediate goods producers solve:

$$\max_{P_t(z)} \sum_{t=0}^{\infty} E_0 Q_{0,t}^s \left[ (1 + \sigma) P_t(z) Y_t(z) - W_t l_t(z) - \frac{\nu}{2} \left( \frac{P_t(z)}{P_{t-1}(z)} - 1 \right)^2 P_t Y_t \right],$$

where $Q_{0,t}^s \equiv \beta_s^t (P_0 c_0^P / P_t c_t^P)^\gamma$ is the marginal rate of intertemporal substitution of participants between times 0 and $t$, and $\sigma$ is a sales subsidy. Firms face demand for their products from two sources: consumers and firms themselves (in order to pay for the adjustment cost); the demand function for the output of firms $z$ is $Y_t(z) = (P_t(z)/P_t)^{-\varepsilon} Y_t$. Substituting this into the profit function,

44Monetary variables are generally not uniquely determined. There always exists an equilibrium of our model where money has no value. If agents anticipate that money is not traded in the future, they do not accept money today and the price of money is 0. The reason for the existence of a non-monetary equilibrium is the same as in the monetary overlapping-generations model of Samuelson (1958). In such a cashless equilibrium, the consumption of $N$ agents is $c^N = \delta_t$ in each period. The consumption of $P$ agents is easily determined; this is akin to the standard cashless New Keynesian model studied in Woodford (2003).
the first-order condition is, after simplifying:

\[ 0 = Q_{0,t} \left( \frac{P_t(z)}{P_t} \right)^{-\varepsilon} Y_t \left[ (1 + \sigma) (1 - \varepsilon) + \varepsilon W_t \left( \frac{P_t(z)}{P_t} \right)^{-1} \right] - Q_{0,t} \nu P_t Y_t \left( \frac{P_t(z)}{P_{t-1}(z)} - 1 \right) \frac{1}{P_{t-1}(z)} + \]

\[ + E_t \left\{ Q_{0,t+1} \left[ \nu P_{t+1} Y_{t+1} \left( \frac{P_{t+1}(z)}{P_t(z)} - 1 \right) \frac{P_{t+1}(z)}{P_t(z)^2} \right] \right\} \]

In a symmetric equilibrium all producers make identical choices (including \( P_t(z) = P_t \)); defining net inflation \( \pi_t \equiv (P_t/P_{t-1}) - 1 \), and noticing that \( Q_{0,t+1} = Q_{0,t}\beta \left( c_t^P/c_{t+1}^P \right)^{\gamma} (1 + \pi_{t+1})^{-1} \), equation (23) becomes:

\[ \pi_t (1 + \pi_t) = \beta E_t \left[ \left( \frac{c_t^P}{c_{t+1}^P} \right)^{\gamma} \frac{Y_{t+1}}{Y_t} \pi_{t+1} (1 + \pi_{t+1}) \right] + \]

\[ + \frac{\varepsilon - \frac{1}{\nu}}{\varepsilon - 1} \left( \frac{\varepsilon}{A_t} - (1 + \sigma) \right) \]

### B.2 Loglinearized equilibrium conditions

Table 1 outlines the equilibrium conditions loglinearized around an arbitrary steady state, denoting \( \hat{\pi}_t = \ln \frac{1+\pi_t}{1+\pi} \) and \( \hat{i}_t = \ln \frac{1+i_t}{1+i} \).

<table>
<thead>
<tr>
<th>Table 1: Summary of loglinearized equilibrium conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Euler P \quad \hat{c}<em>t^P = E_t c</em>{t+1}^P - \gamma^{-1} (\hat{i}<em>t - E_t \hat{\pi}</em>{t+1})</td>
</tr>
<tr>
<td>Self-insurance/money demand \quad \hat{c}<em>t^P = \frac{\alpha \beta}{1+\pi} E_t c</em>{t+1}^P + \left( 1 - \frac{\alpha \beta}{1+\pi} \right) E_t \hat{c}<em>{t+1}^N + \gamma^{-1} E_t \hat{\pi}</em>{t+1}</td>
</tr>
<tr>
<td>Budget constraint N \quad \hat{c}<em>{t+1}^N = \frac{1-\rho}{\pi+1} \left( \hat{m}</em>{t+1}^n - \hat{\pi}_t \right) - \frac{\rho}{\pi+1} \hat{\tau}_t + \frac{\delta}{\pi} \hat{\tau}_t</td>
</tr>
<tr>
<td>Ec res cons (replace P’s BC) \quad \hat{c}_t = (1 - (1 - n) \delta_c) \left( \hat{a}_t + \hat{i}_t - \frac{\nu (1+\pi) \hat{\pi}_t}{1+\pi} \right) + (1 - n) \delta_c \hat{\pi}_t</td>
</tr>
<tr>
<td>Aggregate labor \quad \hat{l}_t = \hat{\rho}_t</td>
</tr>
<tr>
<td>Labor supply \quad \varphi \hat{l}_t = \hat{w}_t - \gamma \hat{c}_t^P</td>
</tr>
<tr>
<td>Aggregate C \quad \hat{c}_t^P = \frac{1}{np} \hat{\pi}<em>t - \frac{(1-n)h}{np} \hat{c}</em>{t+1}^N</td>
</tr>
<tr>
<td>Transfer \quad -\hat{\tau}_t = \hat{\tau}_t</td>
</tr>
<tr>
<td>Money growth \quad \hat{m}<em>t = \hat{m}</em>{t+1}^n - \hat{\pi}_t</td>
</tr>
<tr>
<td>Phillips curve \quad \hat{\pi}<em>t = \beta E_t \hat{\pi}</em>{t+1} + \frac{1+\pi}{1+2\pi} \frac{\varepsilon - \frac{1}{\nu}}{\varepsilon - 1} \left( \hat{w}_t - \hat{a}_t \right)</td>
</tr>
</tbody>
</table>

### B.3 Friedman Rule with flexible prices

First we show that, as in other monetary economies, the price level is indeterminate at the Friedman rule. For \( i = 0 \), the steady state implies \( 1 + \pi = \beta \); \( c^P = c^N = c \) and \( m^{CB} = (1 - \beta^{-1}) \left( n \hat{m}^P + (1 - n) \hat{m}^N \right) \).

The real allocation is determined by \( 1 = \chi \left( l^P \right)^{\phi} \left( c^P \right)^{\gamma} \); \( c = nl^P + (1 - n) \delta \) and

\[ c = \delta + \frac{1}{\beta} \left( 1 - \rho - (1 - (1 - \beta) \phi) \left( \frac{1}{\beta} - 1 \right) n \right) \hat{m}^P + \frac{1}{\beta} \left( \rho - \beta - (1 - (1 - \beta) \phi) \left( \frac{1}{\beta} - 1 \right) (1 - n) \right) \hat{m}^N \]
There is indeterminacy, even though the real variables \(c\) and \(l\) are uniquely determined as the steady-state first-best values: the monetary variables \(\tilde{m}^P\) and \(\tilde{m}^N\) must satisfy only one equation, so the real quantity of money is indeterminate.

Second, we show (in the nonlinear model) convergence to the first-best allocation (when \(\nu = 0\) if \(2 - \alpha - \rho > \beta^{-1} - 1\); the steady-state allocation converges to the first best when \(i \to 0^+\). In this case, \(1 + \pi \to \beta^+\). For \(0 < k < 1\), define \(\hat{l}_t(k)\) as the unique solution to the equation:

\[
(n + (1 - n) k) \left( \frac{A_t}{\chi} \right)^{\frac{1}{\gamma}} \left( l_t^P(k) \right)^{-\frac{\nu}{\gamma}} = A_t \hat{l}_t^P(k) + (1 - n) \delta_t
\]

As the left hand side is decreasing and the right hand side increasing in \(l_t^P\), there always exists a positive solution to the previous equation, whatever \(A_t, \delta_t > 0\). Define \(\hat{c}_t^P(k)\) as:

\[
\hat{c}_t^P(k) = \frac{nA_t l_t^P + (1 - n) \delta_t}{n + (1 - n) k}
\]

For any \(k < 1\), we show that this can reach allocations where \(c_t^P = \hat{c}_t^P(k)\), \(c_t^N = k \hat{c}_t^P(k)\) and \(l_t = \hat{l}_t^P(k)\). When \(k\) equals 1, the allocation is exactly the first-best allocation. When \(k\) approaches 1, the allocation can be made arbitrarily close to the first-best allocation and the nominal interest rate \(i_t\) tends toward \(0^+\). Take now the model equations from Appendix A, for the case of flexible prices \(\nu = 0\) and using the money market equilibrium to substitute for \(\tilde{m}^P_{t+1}\). We proceed by guess and verify. At any period, the variables \(m_t^{CB}\) and \(\tilde{m}_t^P\) are predetermined. As a consequence, assume that the period \(t\) money creation \(m_t^{CB}\) is determined by the following law:

\[
m_t^{CB} = k \hat{c}_t^P - \delta - \frac{1}{\beta} \frac{u'(\hat{c}_{t-1}^P)}{u'(\hat{c}_t^P)} \frac{1 - \rho}{\alpha + (1 - \alpha) k^{-\gamma}} \tilde{m}_t^P
\]

It is easy to show that the allocation \(c_t^P = \hat{c}_t^P; c_t^N = k \hat{c}_t^P; i_t = \alpha + (1 - \alpha) k^{-\gamma}; \tau_t = -m_t^{CB}\) and

\[
1 + \pi_t = \beta \left[ \alpha + (1 - \alpha) k^{-\gamma} \right] \frac{u'(\hat{c}_t^P)}{u'(\hat{c}_{t-1}^P)}
\]

is an equilibrium of the model, because it satisfies all equations. The equilibrium is locally unique, which we show by standard perturbation methods in a more general case in our companion paper.
B.4 Optimal Ramsey Policy

The constraints of the Ramsey planner are (these are the model equations, with relevant substitutions and using the economy resource constraint instead of the P budget constraint):

\[ u' \left( c_t^P \right) = \beta E_t \left[ \frac{\alpha u' \left( c_{t+1}^P \right) + (1 - \alpha) u' \left( c_{t+1}^N \right)}{1 + \pi_{t+1}} \right] \]

\[ nc_t^P + (1 - n) c_t^N = \left( 1 - \frac{\nu}{2} n \pi_t^2 \right) n A_t l_t^P + (1 - n) \delta \]

\[ \pi_t (1 + \pi_t) = \beta E_t \left[ \left( \frac{c_t^P}{c_{t+1}^P} \right)^{\gamma} \frac{A_{t+1} l_{t+1}^P}{A_t l_t^P} \pi_{t+1} (1 + \pi_{t+1}) + \frac{1}{1 + \pi_t} \right] \frac{1 - \omega}{1 - n} \left( m_{t+1}^{\text{tot}} - \frac{m_t^{\text{tot}}}{1 + \pi_t} \right) + \frac{1}{1 + \pi_t} \frac{1 - \alpha}{1 - n} m_{t+1}^{\text{tot}} \]

\[ u' \left( c_t^P \right) = \beta E_t \frac{1 + i_t}{1 + \pi_{t+1}} u' \left( c_{t+1}^P \right) \]

As long as money is created through helicopter-drop, within-period transfers, only the first three equations above are constraints for the Ramsey planner.\(^{45}\) Indeed, once \( c^P \) and \( \pi \) are known, \( i \) follows from the Euler equation for bonds (which hence will not bind as a constraint). Similarly, once the allocation of the consumption of \( N \) and inflation have been chosen, the quantity of money delivering it can be recovered through the following equation:

\[ c_t^N = \delta_t + m_{t+1}^{\text{tot}} - \frac{\alpha - n}{1 - n} \frac{m_t^{\text{tot}}}{1 + \pi_t}, \]

where, implicitly, we concentrate only on equilibria where money is used.

The central bank chooses \( c^P, c^N, l^P, \pi \) to maximize the objective defined in the text, subject to the above system of 3 constraints which we denoted in text by \( \Gamma_t \) and write here explicitly for reference:

\[ \max_{\{c_t^P, c_t^N, l_t^P, \pi_t\}} E_0 \sum_{t=0}^{\infty} \beta^t \left\{ n \left[ u \left( c_t^P \right) - \frac{\chi (l_t^P)^{1+\varphi}}{1 + \varphi} \right] + (1 - n) \left[ u \left( c_t^N \right) - \frac{\chi \delta_t^{1+\varphi}}{1 + \varphi} \right] \right\} \]

\[ + \omega_1 t \left[ (1 + \pi_{t+1}) \left( c_t^P \right)^{-\gamma} - \beta \alpha \left( c_{t+1}^P \right)^{-\gamma} - \beta (1 - \alpha) \left( c_{t+1}^N \right)^{-\gamma} \right] \]

\[ + \omega_2 t \left[ nc_t^P + (1 - n) c_t^N - \left( 1 - \frac{\nu}{2} n \pi_t^2 \right) n A_t l_t^P - (1 - n) \delta \right] \]

\[ + \omega_3 t \left[ \pi_t (1 + \pi_t) - \beta E_t \left[ \left( \frac{c_t^P}{c_{t+1}^P} \right)^{\gamma} \frac{Y_{t+1}}{Y_t} \pi_{t+1} (1 + \pi_{t+1}) \right] - \frac{\varepsilon}{\nu} \left( \frac{\chi (l_t^P)^{\varphi} (c_t^P)^{\gamma}}{A_t} + \Phi - 1 \right) \right] \}

The solution is a system of 4 first-order conditions and 3 constraints, for 4 variables and 3 co-states (the Lagrange multipliers on the constraints). The first-order conditions of the Ramsey problem are for each variable respectively:

\(^{45}\)Whereas with open-market operations, all of the above equations are constraints; we analyze this case and provide a welfare comparison in the companion paper.
plus the three constraints with complementary slackness.

A steady-state of the Ramsey problem is defined by:

\[ \omega_1 = 0 \quad \text{or} \quad (c^P)^{-\gamma} = \frac{1 - \alpha}{1 + \pi - \alpha} \left( c^N \right)^{-\gamma} \]
\[ \omega_2 = 0 \quad \text{or} \quad n c^P + (1 - n) c^N = \left( 1 - \frac{\nu}{2} \pi^2 \right) n P + (1 - n) \delta \]
\[ \omega_3 = 0 \quad \text{or} \quad \pi (1 + \pi) = \frac{\varepsilon}{\nu (1 - \beta)} \left[ \chi \left( l^P \right)^{\varphi} (c^P)^{\gamma} - (1 - \Phi) \right] \]

The Proof of the proposition pertaining to optimal long-run inflation is now immediate. With flexible prices \( \nu = 0 \) and optimal subsidy, the only solution to the above system of equations is perfect insurance through the Friedman Rule:

\[ \frac{1 + \pi}{\beta} = 1 \rightarrow c^P = c^N = c \]

With sticky prices and inelastic labor \( \varphi \rightarrow \infty \), the intratemporal optimality condition disappears from the set of constraints, labor is fixed, and it can be easily shown that inflation tends to zero \( (\pi = 0 \text{ solves the above system}) \).

**Computing the welfare cost** To calculate the welfare cost of inflation, we proceed in the standard way pioneered by Lucas (1987). Denote with an upper-script SS the allocation for the inflation rate \( \pi^{SS} \) and no shock. We denote the welfare of an economy where inflation is, say \( \pi^* \), as \( V^* \). We then
compute the proportional decrease in consumption for all households in the economy with inflation rate \( \pi^* \) to equalize the two welfare measures. Formally we compute \( \Delta W \) to have:

\[
E_0 V^* = E_0 \sum_{t=0}^{\infty} \beta^t \left\{ n \left[ u \left( (1 - \Delta^W) c_t^{PSS} \right) - \chi \left( \frac{P_t^{SSt}}{1 + \varphi} \right)^{1+\varphi} \right] + (1 - n) \left[ u \left( (1 - \Delta^W) c_t^{NSS} \right) - \chi \left( \frac{P_t^{SSt}}{1 + \varphi} \right)^{1+\varphi} \right] \right\}
\]

### B.5 Welfare function and second-order approximation

The second-order approximation technique used is described in detail in Woodford (2003, Chapter 6), Benigno and Woodford (2005, 2012), and Bilbiie (2008) for the case of two agents. A second-order approximation of \( P \) agents' utility delivers:

\[
U_t^P - U^P = \left( c^P \right)^{1-\gamma} \left( \hat{c}_t^P + \frac{1-\gamma}{2} \left( \hat{c}_t^P \right)^2 \right) + \left( U_L^P \right)^{1-\gamma} \left( \hat{i}_t^P + \frac{1+\varphi}{2} \left( \hat{i}_t^P \right)^2 \right)
\]

where the second equality used SS under subsidy \( w = 1 \), \( U_L^P = -U_C^P \).

For \( N \) agents:

\[
U_t^N - U^N = \left( c^N \right)^{1-\gamma} \left( \hat{c}_t^N + \frac{1-\gamma}{2} \left( \hat{c}_t^N \right)^2 \right)
\]

Aggregating:

\[
U_t - U = \left( c^P \right)^{-\gamma} \left( nc^P \hat{c}_t^P - nl^P \hat{i}_t^P + \left( \frac{c^N}{c^P} \right)^{-\gamma} (1 - n) c^N \hat{c}_t^N \right)
\]

\[
+ n \left( c^P \right)^{1-\gamma} \left( \frac{1-\gamma}{2} \left( \hat{c}_t^P \right)^2 - \frac{l^P}{c^P} \left( \frac{1+\varphi}{2} \left( \hat{i}_t^P \right)^2 \right) \right) + (1 - n) \left( c^N \right)^{1-\gamma} \frac{1-\gamma}{2} \left( \hat{c}_t^N \right)^2
\]

Take the linear term first:

\[
\left( c^P \right)^{-\gamma} \left( c \hat{c}_t - l \hat{i}_t + \left( \frac{c^N}{c^P} \right)^{-\gamma} - 1 \right) \left( 1 - n \right) c^N \hat{c}_t^N
\]

The economy resource constraint to second order is (denote \( \Delta_t = 1 - \frac{\nu}{2 \pi_t^2} \)):

\[
\hat{c}_t = (1 - (1 - n) \delta_c) \left( a_t + \hat{l}_t + \hat{\Delta}_t \right) \text{ where } \hat{\Delta}_t = - \frac{\nu \pi}{1 - \frac{\nu}{2 \pi_t^2}} - \frac{1}{2} \frac{\nu}{1 - \frac{\nu}{2 \pi_t^2}} \pi_t^2
\]

Note that under zero inflation the linear term disappears. The squared term captures the welfare cost of inflation.

The linear term becomes hence:

\[
\left( c^P \right)^{-\gamma} c \left( \hat{\Delta}_t + a_t + (q^\gamma - 1) (1 - n) h \hat{c}_t^N \right)
\]
where we recall $q^\gamma = 1 + \frac{(1+\pi)^{\beta-1}-1}{1-\alpha}$; at the Friedman rule this is unity, and the linear term drops out. Otherwise, it is larger than 1 and the linear term has a positive coefficient – increasing the consumption of $N$ closes the inequality gap, providing a first-order benefit.

The quadratic term is (ignoring price dispersion because in quadratic terms it becomes third or fourth order):

$$
\left(\frac{c^P}{2}\right)^{-\gamma} c \left( (1 - \gamma) \left( np \left( \hat{c}_t^P \right)^2 + (1 - n) hq^\gamma \left( \hat{c}_t^N \right)^2 \right) - \frac{1 + \varphi}{1 - (1 - n) \delta_c} \hat{c}_t^2 \right)
$$

Thus the loss function becomes, rearranging and ignoring terms independent of policy:

$$
\mathcal{L} = \left(\frac{c^P}{2}\right)^{-\gamma} c \left( -\hat{\Delta}_t - (q^\gamma - 1) (1 - n) h\hat{c}_t^N + \frac{\gamma - 1}{2} \left( np \left( \hat{c}_t^P \right)^2 + (1 - n) hq^\gamma \left( \hat{c}_t^N \right)^2 \right) + \frac{1 + \varphi}{1 - (1 - n) \delta_c} \hat{c}_t^2 \right)
$$

Adding and subtracting the steady-state inflation constant and ignoring all terms independent of policy, we obtain the loss function in the text.