Luiss
School of European Political Economy

Heterogeneity, Bubbles and Monetary Policy

Jacopo Bonchi and Salvatore Nisticò

Working Paper 2/2022

February 8, 2022

Heterogeneity, Bubbles and Monetary Policy*

Jacopo Bonchi
LUISS Guido Carli

Salvatore Nisticò
Sapienza University of Rome

February 8, 2022

Abstract

Using a tractable New Keynesian model with heterogeneous agents, we analyze the interplay between heterogeneity and rational bubbles, and their implications for monetary policy. Households are infinitely-lived and heterogeneous because of two sources of idiosyncratic uncertainty, which makes them stochastically cycle in and out of segmented asset markets, and in and out of employment. We show that bubbles can emerge in equilibrium despite the fact that households are infinitely lived, because of the structural heterogeneity that affects their activity in asset and labor markets. The elasticity of an endogenous labor supply, the heterogeneity in asset-market participation and the level of long-run monopolistic distortions are shown to affect the size of equilibrium bubbles and their cyclical implications. We also show that a central bank concerned with social welfare faces an additional tradeoff implied by bubbly fluctuations which makes, in general, strict inflation targeting a suboptimal monetary-policy regime.

JEL codes: E21, E32, E44, E58
Keywords: Wealth inequality, Rational bubbles, Optimal monetary policy

*We are grateful to Pierpaolo Benigno, Oreste Tristani, Alejandro Van der Ghote for insightful discussions and comments and to seminar and conference participants at the ECB, the 11th "RCEA Money, Macro and Finance" Conference, the 27th "Computing in Economics and Finance" International Conference. S. Nisticò gratefully acknowledges the financial support from the PRIN grant 2020LKAZAH.

J. Bonchi: LUISS Guido Carli, Department of Economics and Finance and School of European Political Economy, Viale di Villa Emiliani, 14, 00197 Rome, email: jbonchi'at'luiss.it. S. Nisticò: Sapienza Università di Roma, Department of Social and Economic Sciences, piazzale Aldo Moro 5, 00185 Rome, email: salvatore.nistico'at'uniroma1.it

1 Introduction

The past 40 years have been characterized by a secular downward trend in long-term interest rates that started much before the recent financial and pandemic crises. On top of this, advanced economies have spent most of the past 15 years at exceptionally low levels of interest rates at all maturities – even compared to that trend – ever since the burst of the housing bubble in the U.S. and the ensuing global financial crisis forced most central banks to slash their policy rates down to zero. The strong surge in asset prices that followed such massive monetary expansion is fueling the idea that a new bubble might be in the making, with potentially large risks for an already shaken global economy. At the same time, the financial and pandemic crises have brought in the spotlight of the public and academic debate the increasing level of inequalities in wealth and income globally, and central banks worldwide seem ever more concerned with the distributional consequences of their policy actions.[1]

In this paper, we study how rational bubbles in asset prices can emerge in a low interest rate environment populated by heterogeneous agents, their aggregate and distributional effects on the economy and their normative implications for monetary policy.

Despite their relevance in the public debate, the analysis of bubble-driven fluctuations in modern monetary models is somewhat limited, mainly because in the workhorse New Keynesian model (widely used for monetary policy analysis) the assumption of a representative and infinitely-lived agent implies that the transversality condition ensuring solvency at the individual level necessarily holds for the whole economy as well, thus preventing the existence of bubbles in equilibrium. For this reason, rational bubbles have been studied mostly in OLG models, where the assumption of finite lives prevents the transversality condition from holding at the economy level, and bubbles can emerge if a declining path of labor income implies $r < g$ and excess savings to be absorbed.[2] In a recent paper, Galí (2021) modifies the New Keynesian framework to include the basic mechanism of this class of models with finite lives and studies the positive implications of rational bubbles for monetary policy. A second class of theoretical frameworks study rational bubbles in (real) infinite-horizon models with financial constraints.[3] In this case, as shown by Miao and Wang (2018), bubbles carry a "collateral yield", making their growth rate lower than the real interest rate. Thus, bubbles can exist even if $r > g$ and the transversality condition holds.[4]

The framework used in this paper is a tractable New Keynesian model with heterogeneous agents, and thus belongs to the first class of models above, with which we share some key structural assumptions allowing for the existence of bubbles in equilibrium.[5] At the same time, we depart from this literature by focusing on the role of heterogeneity (rather than finite lives) for the rise of rational

[1] Among other public displays of such concerns, see Yves Mersch of the Executive Board of the ECB, at the Corporate Credit Conference in 2014: "*...we need to be aware that there are distributional consequences of our actions – and these may well be particularly significant at times of exceptionally low interest rates and non-standard measures.*"

[2] See, e.g., Samuelson (1958) and Tirole (1985).

[3] See, e.g., Kocherlakota (1992), Miao and Wang (2012, 2014, 2018) and Hirano and Yanagawa (2017).

[4] See Santos and Woodford (1997) for an analysis of the general conditions for the existence of rational bubbles. Instead, a comparison between the two approaches to the study of rational bubbles can be found in Miao (2014).

[5] In particular, the paper most related to ours is Galí (2021).

bubbles, and on the normative implications of such heterogeneity for monetary policy, derived in a linear-quadratic perspective within a fully microfounded, general equilibrium model.[6] We analyze the role of households' heterogeneity in the rise of bubble-driven fluctuations and how these fluctuations affect in turn the cross-sectional consumption dispersion and social welfare. Households in the economy are heterogeneous because of two sources of idiosyncratic uncertainty, which makes them stochastically cycle in and out of segmented asset markets, and in and out of employment. We build the analysis on the stochastic asset-market participation model developed in Nisticò (2016), extended to account also for the transition in and out of the labor market and for the rise of bubbles in equilibrium.

Our main results and contributions to the literature can be summarized as follows.

First, we show that, despite agents in our economy being infinitely-lived, their structural heterogeneity implies a finite planning horizon in asset markets that is isomorphic to an OLG structure. The transversality condition holds for individual asset-market participants, but not necessarily for the aggregate economy. Accordingly, the type of heterogeneity implied by the stochastic participation in asset and labor markets characterizing our economy satisfies the same conditions for the existence of bubbles in equilibrium derived in related OLG models.[7]

Second, in our general model with endogenous labor supply, we emphasize the role of complementarity effects of labor on consumption, labor supply elasticity, heterogeneity in asset-market participation and long-run monopolistic distortions in characterizing the equilibrium size of bubbles in the balanced-growth path, as well as the shape of bubbly fluctuations over the business cycle.

Third, we exploit the tractability of our New Keynesian model with heterogeneous agents to derive a simple second-order approximation of social welfare around the efficient balanced-growth path to characterize the welfare-maximizing monetary policy in the face of bubble-driven fluctuations. Such an approach allows evaluating, from a welfare perspective, the cyclical implications of fluctuations in the rational bubble, taking into explicit account the distributional consequences of the latter among the heterogeneous agents that populate our economy, and that are relevant for welfare. We show that, despite "divine coincidence" holds from the supply-side of the economy, so that output gap and inflation can be simultaneously stabilized, an endogenous policy tradeoff emerges that makes strict inflation targeting in the presence of bubble-driven fluctuations suboptimal. This additional tradeoff is more stringent for the central bank when the economy fluctuates around a bubbleless balanced-growth path, when monetary policy cannot affect bubbles directly through its policy rate. Moreover, such a bubbleless balanced-growth path is necessarily globally

[6] Related New Keynesian models with heterogeneous agents are analyzed by Bilbiie (2018) and Debortoli and Galí (2018), among others, which extend the Two-Agent New Keynesian (TANK) or Limited Asset-Market Participation (LAMP) settings (see, e.g., Galí et al. 2007 and Bilbiie, 2008) but do not analyze the implications for bubbly fluctuations.

[7] Both assumptions of finite planning horizon and finite lives imply infinitely many traders who can trade in the bubbly asset, a feature discussed in Tirole (1982). Indeed, the bubble can exist only if the buyer can sell it in the future to someone else, who is the future new entrant in the asset markets in the model with stochastic asset-market participation and the future generation in the OLG model. An alternative way of introducing bubbles in infinite-horizon models is Weil (1989), where individual planning horizons and individual lifetimes are both infinite, but where the population size is growing over time.

stable, with the further implication that bubble fluctuations can arise from self-fulfilling revisions in expectations about the value of pre-existing bubbly assets. In response to this type of bubbly fluctuations, we show in particular that the optimal monetary policy requires deviating from strict inflation targeting, with the specific type of policy response depending also on the nature of the bubble shock.

The paper is structured as follows. Section 2 presents the model; in Section 3 we discuss the implications for equilibrium bubbles along the balanced-growth paths and in a linear version of our model. Section 4 analyzes the monetary policy tradeoffs implied by bubbly fluctuations, and their normative implications. Section 5 concludes.

2 The Model Economy

The economy is populated by infinitely-lived households consuming a bundle of differentiated goods and supplying labor for their production. A continuum of firms produces the differentiated goods using labor services and technology, and faces a positive default probability. The public sector consists of a fiscal authority which imposes taxes and provides transfers within a balanced budget, and a central bank in charge of monetary policy.

2.1 Households

A continuum of infinitely-lived households spans the interval [0,1]. Households face two types of idiosyncratic uncertainty, related to their participation in asset and labor markets. Agents are therefore heterogeneous along three respects: *i*) their participation status in asset markets, where they can smooth consumption over time, *ii*) their longevity in asset markets, which implies a non-uniform cross-sectional distribution of financial wealth, *iii*) their employment status.

With respect to the first layer of heterogeneity, we build on the stochastic asset-market participation framework developed in Nisticò (2016): a share ϑ of the population has access to the financial market and smooths consumption over time while $1-\vartheta$ does not and consumes its net labor income period by period. We refer to the former as "market participants", or "financially active" agents, and denote them with the superscript p, while we refer to the latter as "rule-of-thumber", or "financially inactive" agents, and denote them with the superscript r. Each agent's status in the financial market evolves over time as an independent two-state Markov chain: each period, each agent learns whether or not she will be active in asset markets, where the relevant probability is only dependent on the agent's current state. Each market participant remains financially active with probability $\gamma \in (0,1]$, while with probability $1-\gamma$ she becomes a rule-of-thumber. Participants turning rule-of-thumbers have an incentive to enter into an insurance contract à la Blanchard (1985), in order to smooth the effects of the transition out of the asset market over the time span in which they are active, in the form of an extra return on their financial portfolio. Rule-of-thumbers remain financially inactive with probability $\varrho \in [0,1]$, and turn active with probability $1-\varrho$. Therefore, the outflow from financial markets each period has mass $\vartheta(1-\gamma)$ while the inflow has mass

$(1-\vartheta)(1-\varrho)$: assuming $\vartheta(1-\gamma) = (1-\vartheta)(1-\varrho)$ ensures that the shares of participants and rule-of-thumbers remain constant over time. Defining a "cohort" as the set of agents experiencing a transition in the same period, the time$-t$ size of the cohort that became financially active at time $s \leq t$ is $m^{p}_{t|s} \equiv \vartheta\,(1-\gamma)\,\gamma^{t-s}$, and the size of the cohort that became inactive at time $s \leq t$ is $m^{r}_{t|s} \equiv (1-\vartheta)\,(1-\varrho)\,\varrho^{t-s}$.

The second layer of heterogeneity is related to the employment status. To keep things simple and reduce the state space while still allowing the model to display the relevant features that support bubbly equilibria, we model the transition into and out of employment as follows: the transition out of employment occurs only for financially active agents, while the transition into employment occurs only for financially inactive ones. In particular, each employed market participant keeps her job every period with probability ν, and loses it with probability $1-\nu$. Instead, rule-of-thumbers keep their employment status until they are hit by the idiosyncratic shock that makes them financially active, in which case they also become employed with conditional probability 1.[8] Transition into market participation, therefore, also implies transition into employment (for the unemployed rule-of-thumbers). Denoting with the superscripts e and u respectively employed and unemployed agents, the time-t mass of employed market participants belonging to cohort s is $m^{pe}_{t|s} \equiv \vartheta\,(1-\gamma)\,(\gamma\nu)^{t-s}$, while the time-$t$ mass of unemployed ones is $m^{pu}_{t|s} \equiv \vartheta\,(1-\gamma)\,\gamma^{t-s}\left(1-\nu^{t-s}\right)$. Accordingly, the share of market participants (as well as of the overall population) that is employed in each period is

$$\alpha \equiv \sum_{s=-\infty}^{t} \frac{m^{pe}_{t|s}}{\vartheta} = \frac{1-\gamma}{1-\gamma\nu} \in [0,\,1].$$

Likewise, the time-t mass of employed and unemployed rule-of-thumbers in the cohort s is, respectively, $m^{re}_{t|s} \equiv (1-\vartheta)\,(1-\varrho)\,\alpha\varrho^{t-s}$ and $m^{ru}_{t|s} \equiv (1-\vartheta)\,(1-\varrho)\,(1-\alpha)\,\varrho^{t-s}$.

Finally, note that the stochastic transition into and out of the financial market implies heterogeneity not only *between* market participants and rule-of-thumbers, but also *within* the set of market participants, related to the cross-sectional distribution of financial wealth associated with the different longevities in the asset market. On the contrary, rule-of-thumbers hold zero wealth and are thus identical within their employment status, independently of their longevity out of the financial market.

As discussed in Nisticò (2016), this framework nests as special cases most popular models used for the analysis of the business cycle. In particular, the limiting case where $\vartheta = 1$ here nests the perpetual-youth economy considered in Galí (2021) – where γ is the probability of dying – extended to account for endogenous labor-supply decisions and aggregate wage schedule.

Let $j \in \mathcal{T} \equiv \{pe,\,pu,\,re,\,ru\}$ index the individual type with respect to the first two layers of heterogeneity, and $s \in (-\infty,\,t]$ index the cohort, thus capturing the third one. The economy-wide

[8] As will soon become clear, this assumption has no consequence for our results.

aggregate of a generic variable X is therefore a mass-weighted average across types and cohorts:

$$X_t \equiv \sum_{j\in\mathcal{T}} \sum_{s=-\infty}^{t} m_{t|s}^{j} X_{t|s}^{j} \tag{1}$$

$$= \vartheta X_t^p + (1-\vartheta)\, X_t^r \tag{2}$$

$$= \vartheta \left[\alpha X_t^{pe} + (1-\alpha)\, X_t^{pu}\right] + (1-\vartheta)\left[\alpha X_t^{re} + (1-\alpha)\, X_t^{ru}\right], \tag{3}$$

where $X_t^p = \sum_{j\in\{pe,\,pu\}} \sum_{s=-\infty}^{t} \frac{m_{t|s}^{j}}{\vartheta} X_{t|s}^{j}$ is the average per-capita level across participants and, analogously, $X_t^{pe} = \sum_{s=-\infty}^{t} \frac{m_{t|s}^{pe}}{\vartheta\alpha} X_{t|s}^{pe}$ and $X_t^{pu} = \sum_{s=-\infty}^{t} \frac{m_{t|s}^{pu}}{\vartheta(1-\alpha)} X_{t|s}^{pu}$ the average per-capita level across employed and unemployed participants, respectively. Finally, since rule-of-thumbers have zero financial wealth, we have $X_{t|s}^{re} = X_t^{re}$ and $X_{t|s}^{ru} = X_t^{ru}$ for all $s \in (-\infty,\, t]$.

2.1.1 Preferences

Households have preferences in the class introduced by Greenwood, Hercowitz and Huffman (1998) modified to ensure consistency with a balanced-growth path, in the spirit of Jaimovic and Rebelo (2009):

$$U_{t|s}^{j} = \log\left(C_{t|s}^{j} - V(N_{t|s}^{j})\right) = \log \widetilde{C}_{t|s}^{j},$$

where $\widetilde{C}_{t|s}^{j} \equiv C_{t|s}^{j} - V(N_{t|s}^{j})$ denotes *adjusted* consumption and $V(N_{t|s}^{j})$ the disutility of labor. These preferences imply complementarity effects of labor on consumption, so that we can think of $V(N_{t|s}^{j})$ also as the "subsistence" level of consumption, at or below which utility would be undefined. We specify the disutility of labor as

$$V(N_{t|s}^{j}) \equiv \frac{\delta\Gamma^t}{1+\varphi}\left(N_{t|s}^{j}\right)^{1+\varphi},$$

where $\delta \geq 0$, Γ^t is an index of labor productivity, growing at the rate $\Gamma \equiv (1+g) \geq 1$, and φ is the inverse Frisch elasticity of labor supply, which here also captures (inversely) the complementarity effects of labor efforts on consumption, which will play an important role in the analysis.

Agents consume a composite bundle of a mass α of differentiated brands

$$C_{t|s}^{j} \equiv \left[\left(\frac{1}{\alpha}\right)^{\frac{1}{\epsilon}} \int_{i\in\mathcal{F}} \left(C_{t|s}^{j}(i)\right)^{\frac{\epsilon-1}{\epsilon}} di\right]^{\frac{\epsilon}{\epsilon-1}},$$

where $\mathcal{F}$ denotes the set of firms producing these brands, and $\epsilon > 1$ the elasticity of substitution between any two of such brands. Each brand sells at price $P(i)$, implying that the consumption-based aggregate price index is

$$P_t \equiv \left[\frac{1}{\alpha}\int_{i\in\mathcal{F}} P_t(i)^{1-\epsilon} di\right]^{\frac{1}{1-\epsilon}}.$$

The optimal allocation of spending across differentiated goods implies the equilibrium demand for brand i for individual of type j in cohort s

$$C^j_{t|s}(i) = \frac{1}{\alpha}\left(\frac{P_t(i)}{P_t}\right)^{-\epsilon} C^j_{t|s}$$

for all $i \in \mathcal{F}$. This allows to write the aggregate individual spending for consumption as

$$\int_{i\in\mathcal{F}} P_t(i) C^j_{t|s}(i) di = P_t C^j_{t|s},$$

and the aggregate demand faced by firm producing brand i as

$$C_t(i) = \frac{1}{\alpha}\left(\frac{P_t(i)}{P_t}\right)^{-\epsilon} C_t, \tag{4}$$

where we used aggregator (3).

2.1.2 Rule-of-thumbers

A mass α of rule-of-thumbers is employed, and maximizes her utility each period facing the budget constraint[9]

$$C^{re}_t = W_t N^{re}_t - T^{re}_t,$$

where W_t is the real wage and T^{re}_t denotes lump-sum taxes net of transfers. The equilibrium labor supply is

$$N^{re}_t = \left(\frac{w_t}{\delta}\right)^{\frac{1}{\varphi}}, \tag{5}$$

where $w_t \equiv \frac{W_t}{\Gamma^t}$ denotes the real wage relative to productivity.

The unemployed rule-of-thumbers, of relative mass $1-\alpha$, consume each period the unemployment benefit T^{ru}_t received by the fiscal authority, which is set in such a way to equalize the marginal utility of consumption across all financially inactive agents regardless of the employment status:[10]

$$\widetilde{C}^{ru}_t = C^{ru}_t = T^{ru}_t = \widetilde{C}^{re}_t.$$

The unemployment benefit is financed partly with the lump-sum tax on employed rule-of-thumbers and partly with a tax on market participants T^r_t:

$$(1-\alpha)\, T^{ru}_t = \alpha T^{re}_t + T^r_t. \tag{6}$$

It follows that, at equilibrium, the average per-capita level of consumption for financially inactive

[9]Since financially inactive agents are homogeneous across cohorts, henceforth we drop the index s.

[10]This subsidy effectively acts as an insurance mechanism against unemployment risk, analogous to the one provided by complete markets for asset-market participants, as shown in the next subsection.

agents is[11]

$$C_t^r = \alpha \delta^{-\frac{1}{\varphi}} w_t^{\frac{1+\varphi}{\varphi}} \Gamma^t + T_t^r. \tag{7}$$

2.1.3 Market Participants

Market participants can borrow and/or save in the financial market to smooth consumption over time. Since agents stochastically cycle in and out of asset markets, those financially active (though infinitely-lived) take savings decisions using a finite planning horizon, and therefore discount utility flows both for impatience (β) and the probability of remaining in the financial market next period (γ). At time t, an employed agent who has been financially active since $s \leq t$ maximizes

$$E_t \sum_{t=0}^{\infty} (\beta\gamma)^t \, U_{t|s}^{pe}$$

subject to a sequence of budget constraints, expressed in real terms, of the form

$$C_{t|s}^{pe} + E_t \left\{ \Lambda_{t,t+1} Z_{t+1|s}^e \right\} + \int_{i \in \mathcal{F}} \left[Q_t^F(i) - \left(1 - \tau^D\right) D_t(i) \right] Z_{t+1|s}^{Fe}(i) \, di + Q_t^B Z_{t+1|s}^{Be} \\ = A_{t|s}^e + W_t N_{t|s}^{pe} - T_t^{pe}, \tag{8}$$

where $\mathcal{F}$ is the set of monopolistic firms producing a mass α of differentiated brands and issuing equity shares that are traded in a stock market; Z^e is a portfolio of state-contingent assets, for which markets are complete, with $\Lambda_{t,t+1}$ the (unique) relevant stochastic discount factor for one-period-ahead real payoffs; $Z^{Fe}(i)$ is the equity share in firm of brand i, paying off real dividends $D(i)$ taxed at rate τ^D and selling at (real) price $Q^F(i)$; Z^{Be} is the share in bubbles available in the current period, selling at (real) price Q^B; N^{pe} denotes hours worked, remunerated at the real wage W; T^{pe} are lump-sum taxes net of transfers, in real terms, that are independent of the specific longevity in the type ($T_{t|s}^{pe} = T_t^{pe}$ for all $s \leq t$), and A is the real market value of the overall financial portfolio at the beginning of the period.

The latter, for incumbent agents who have been financially active since period $s < t$, is defined as:

$$A_{t|s}^e \equiv \frac{1}{\gamma} \left[Z_{t|s}^e + \int_{i \in \mathcal{F}^*} Q_t^F(i) \, Z_{t|s}^{Fe}(i) \, di + B_t Z_{t|s}^{Be} \right], \tag{9}$$

which pays the extra-return $\gamma/(1-\gamma)$ granted by the insurance contract *à la* Blanchard (1985), and where B is the real market value of bubbles available in the previous period.

Following Galí (2021), we assume that each firm defaults with probability $1 - \nu\gamma$ and exits the economy before a new period starts: accordingly, $\mathcal{F}^*$ is the set of firms that were active in the previous period and have not defaulted, and has mass $\alpha\gamma\nu$. At the beginning of each period, a mass $1 - \alpha\gamma\nu$ of new firms is set up, which replace defaulted ones, and the corresponding shares are

[11] In particular, $T_t^r \equiv \tau^D \frac{\Gamma^t}{1-\vartheta}(d_t - d)$ denotes a transfer through which the fiscal authority redistributes to rule-of-thumbers part of the revenues from a dividend-tax on market participants, where $d_t \equiv D_t/\Gamma^t$ denotes the productivity-adjusted level of aggregate real dividends in the stochastic equilibrium and d its level along the balanced-growth path.

distributed to newcomers in asset markets.[12]

For newcomers, turning financially active in the current period ($s = t$), the portfolio at the beginning of the period includes all the shares in the newborn firms at time t, whose total real market capitalization is $\alpha(1-\gamma\nu)Q^F_{t|t}$, and the newly created bubbly assets, whose total value is U_t, both distributed uniformly among the $\vartheta(1-\gamma)$ newcomers:

$$A^e_{t|t} \equiv \frac{Q^F_{t|t}}{\vartheta} + \frac{U_t}{\vartheta(1-\gamma)}, \tag{10}$$

where we used $\alpha(1-\gamma\nu) = (1-\gamma)$ from the definition of α.

The problem of unemployed market participants, where the relevant variables are denoted with an apex u instead of e, is identical to that of employed ones, except for $N^{pu}_{t|s} = T^{pu}_{t|s} = 0$ and for the fact that the set of unemployed newcomers has zero mass.

The optimality conditions for employed and unemployed market participants imply the equilibrium one-period-ahead stochastic discount factor

$$\Lambda_{t,t+1} = \beta \frac{\widetilde{C}^{pe}_{t|s}}{\widetilde{C}^{pe}_{t+1|s}} = \beta \frac{C^{pu}_{t|s}}{C^{pu}_{t+1|s}}, \tag{11}$$

which is unique because of complete markets, and thus equals the intertemporal marginal rate of substitution in individual consumption across all cohorts; notice that the assumption of complete markets additionally implies equal marginal utility of consumption between employed and unemployed agents within the same cohort, i.e. $\widetilde{C}^{pu}_{t|s} = C^{pu}_{t|s} = \widetilde{C}^{pe}_{t|s}$; the equilibrium fundamental value of equity shares for each brand $i \in [0, \alpha]$

$$Q^F_t(i) = \left(1-\tau^D\right) D_t(i) + \gamma\nu E_t\left\{\Lambda_{t,t+1} Q^F_{t+1}(i)\right\}, \tag{12}$$

related to current dividends (net of taxes) and its own future expected discounted value conditional upon survival of the firm (with probability $\gamma\nu$); the equilibrium market value for the rational bubble

$$Q^B_t = E_t\left\{\Lambda_{t,t+1} B_{t+1}\right\}, \tag{13}$$

related only to its own expected discounted future value, as bubbles are intrinsically worthless; the equilibrium labor supply schedule for employed agents

$$N^{pe}_t = N^{pe}_{t|s} = \left(\frac{w_t}{\delta}\right)^{\frac{1}{\varphi}}, \tag{14}$$

which simply relates hours worked to the productivity-adjusted real wage, and it is therefore common across all market participants, regardless of the longevity in the type, and also equal to the one arising from the financially inactive agents, as shown by equation (5).[13] Finally, a set of no-Ponzi

[12] Alternatively and equivalently, each agent gaining access to asset markets in period t also sets up a new firm.

[13] A labor supply schedule of this kind, with no wealth effects relating individual hours worked to individual con-

game conditions also holds in equilibrium

$$\lim_{k\to\infty} E_t \left\{\Lambda_{t,t+k}\gamma^k A^e_{t+k|s}\right\} = \lim_{k\to\infty} E_t \left\{\Lambda_{t,t+k}\gamma^k A^u_{t+k|s}\right\} = 0,$$

for all $s \in (-\infty,\, t]$.

Using the equilibrium conditions above, we can relate individual current consumption to the stock of financial (both fundamental and bubbly) and human wealth for employed and unemployed agents, respectively:

$$C^{pe}_{t|s} = (1-\beta\gamma)\left(A^e_{t|s} + H_t\right) + V\left(N^{pe}_t\right) \tag{15}$$

$$C^{pu}_{t|s} = (1-\beta\gamma)\, A^u_{t|s}, \tag{16}$$

where the stock of human wealth H includes the expected discounted stream of disposable labor income net of the disutility from working

$$H_t \equiv E_t \left\{\sum_{k=0}^{\infty} (\gamma\nu)^k\, \Lambda_{t,t+k}\left[W_{t+k}N^{pe}_{t+k} - T^{pe}_{t+k} - V(N^{pe}_{t+k})\right]\right\}$$

and is common across all employed market participants. Finally, for all $s \in (-\infty,\, t]$, complete markets imply $A^e_{t|s} + H_t = A^u_{t|s}$.

We can compute the equilibrium per-capita consumption of market participants by taking the mass-weighted average across cohorts and employment statuses, using the aggregators introduced in the previous subsection. Accordingly, the equilibrium per-capita consumption of market participants can be cast in the form

$$C^p_t = (1-\beta\gamma)\,(A_t + \alpha H_t) + \alpha V\left(N^{pe}_t\right) \tag{17}$$

$$= (1-\beta\gamma)\left(\frac{Q^F_t + Q^B_t}{\vartheta} + \alpha H_t\right) + \alpha V\left(N^{pe}_t\right), \tag{18}$$

where in the second equality we use the asset-market clearing condition implying

$$\vartheta A_t = \int_{i\in\mathcal{F}^*} Q^F_t(i)di + (1-\gamma)Q^F_{t|t} + B_t + U_t = Q^F_t + Q^B_t,$$

with

$$Q^B_t = B_t + U_t \tag{19}$$

capturing the current aggregate value of bubbly assets (including both newly created and pre-existing ones) and $Q^F_t \equiv \int_{i\in\mathcal{F}^*} Q^F_t(i)di + \alpha(1-\gamma\nu)Q^F_{t|t}$ the aggregate stock-market value, following

$$Q^F_t = (1-\tau^D)E_t\left\{\sum_{k=0}^{\infty}(\gamma\nu)^k\Lambda_{t,t+k}D_{t+k}\right\}.$$

sumption, is a direct implication of the class of preferences in the GHH class.

For future reference, note that the partition of market participants in newcomers (with mass $1-\gamma$) and incumbents (with mass γ) imply $C_t^p = \gamma C_{t|in}^p + (1-\gamma) C_{t|nc}^p$, where the aggregate consumption of incumbents is

$$\begin{aligned}
\gamma C_{t|in}^p &= (1-\beta\gamma)\left[\frac{1}{\vartheta}\left(\int_{i\in\mathcal{F}^*} Q_t^F(i)di + B_t\right) + \alpha\gamma\nu H_t\right] + \alpha\gamma\nu V\left(N_t^{pe}\right) \\
&= (1-\beta\gamma)\left(\frac{\gamma\nu Q_t^F + B_t}{\vartheta} + \alpha\gamma\nu H_t\right) + \alpha\gamma\nu V\left(N_t^{pe}\right), \qquad (20)
\end{aligned}$$

and that of newcomers is

$$\begin{aligned}
(1-\gamma) C_{t|nc}^p &= (1-\beta\gamma)\left[\frac{\alpha(1-\gamma\nu)Q_{t|t}^F + U_t}{\vartheta} + \alpha(1-\gamma\nu)H_t\right] + \alpha(1-\gamma\nu)V\left(N_t^{pe}\right) \\
&= (1-\beta\gamma)\left[\frac{(1-\gamma\nu)Q_t^F + U_t}{\vartheta} + \alpha(1-\gamma\nu)H_t\right] + \alpha(1-\gamma\nu)V\left(N_t^{pe}\right), \qquad (21)
\end{aligned}$$

where the second equalities in equations above reflect the assumption that newly created firms are homogeneous with the ones they replace.[14]

2.2 Firms

The economy is also populated by a continuum of monopolistic firms: they are in mass α and have access to the linear technology $Y_t(i) = \Gamma^t N_t(i)$ for each brand $i \in [0, \alpha]$; each period a share $\gamma\nu$ of firms remains active, while a share $1-\gamma\nu$ defaults, after producing, and it is replaced by an equal mass of new entrants; newly created firms set their price equal to the past-period's aggregate level while incumbent firms face a probability θ of having to keep their price unchanged, following Calvo (1983); they maximize the expected discounted stream of their profits subject to the demand for their brand (4). In addition, we also assume that employment is subsidized by the government at rate τ^F, to ensure that the aggregate level of output along the balanced-growth path is efficient.

Given these assumptions, the equilibrium price level P^* set by optimizing firms at time t satisfies

$$E_t\left\{\sum_{k=0}^{\infty}(\theta\gamma\nu)^k\left[\Lambda_{t,t+k}C_{t+k}\frac{1}{\alpha}\left(\frac{P_t^*}{P_{t+k}}\right)^{-\epsilon}\left(\frac{P_t^*}{P_{t+k}} - (1+\mu)MC_{t+k}\right)\right]\right\} = 0,$$

where $\mu \equiv (\epsilon-1)^{-1}$ and real marginal costs at time t are equal to the productivity-adjusted real wage, net of the employment subsidy:

$$MC_t = (1-\tau^F)w_t.$$

[14] In particular, homogeneity implies that the stock-market value of a new entrant firm is equal to the average stock-market value across all firms: $Q_{t|t}^F = \alpha^{-1}\int_{i\in\mathcal{F}} Q_t^F(i)di$.

2.3 The Government and the Aggregate Equilibrium

The fiscal authority collects lumps-sum taxes from the employed market participants (T^{pe}) and employed rule-of-thumbers (T^{re}), as well as a tax on dividends, and uses those resources to finance the employment subsidy to firms, a transfer to all rule-of-thumbers, and the unemployment benefit for unemployed rule-of-thumbers:

$$\alpha\vartheta T_t^{pe} + \alpha(1-\vartheta)T_t^{re} + \tau^D D_t = \tau^F W_t N_t + (1-\alpha)(1-\vartheta)T_t^{ru}.$$

Using (6), we can then write

$$\alpha\vartheta T_t^{pe} + \tau^D D_t = \tau^F W_t N_t + (1-\vartheta)T_t^{r}. \tag{22}$$

Moreover, let $y_t \equiv Y_t/\Gamma^t$ denote the productivity-adjusted level of real output, and analogously for all variables inheriting a deterministic trend let a lower-case letter denote the productivity-adjusted level of the corresponding upper-case one.[15] The aggregate stationary equilibrium then also features the resource constraint

$$y_t = \vartheta c_t^p + (1-\vartheta)c_t^r, \tag{23}$$

the aggregate production function

$$y_t \Delta_t^p = N_t, \tag{24}$$

where $\Delta_t^p \equiv \alpha^{-1}\int_{i\in\mathcal{F}}(P_t(i)/P_t)^{-\epsilon}di$ is an index of cross-sectional price dispersion across firms and $N_t \equiv \int_{i\in\mathcal{F}} N_t(i)di$ is the aggregate level of hours worked, the aggregate level of dividends

$$d_t = y_t - (1-\tau^F)w_t N_t \tag{25}$$

and the aggregate stock-market valuation equation

$$q_t^F = (1-\tau^D)d_t + \gamma\nu\Gamma E_t\left\{\Lambda_{t,t+1}q_{t+1}^F\right\}. \tag{26}$$

Finally, note that equations (5) and (14) imply that equilibrium hours worked are identical for rule-of-thumbers and market participants, $N_t^{pe} = N_t^{re} = N_t/\alpha$. The supply-side of the labor market is therefore described by the same wage schedule relating aggregate hours worked to the real productivity-adjusted wage only as in Galí (2021), although here it arises endogenously as an equilibrium condition:

$$w_t = \delta\left(\frac{N_t}{\alpha}\right)^{\varphi}. \tag{27}$$

[15] Therefore, $c_t^j \equiv \frac{C_t^j}{\Gamma^t}$ for $j \in \mathcal{T}$, and note, in particular, $v(N_t^j) \equiv \frac{V(N_t^j)}{\Gamma^t}$.

3 Equilibrium Bubbles in the BGP and the Linear Model

The set of equilibrium conditions useful to characterize the implications for bubbles in the balanced-growth path (BGP henceforth) can be cast in the following form:

$$\tilde{c}_t^p = (1-\beta\gamma)\left(\frac{q_t^B}{\vartheta} + x_t\right) \tag{28}$$

$$x_t = \tilde{c}_t^p + \gamma\nu\Gamma E_t\left\{\Lambda_{t,t+1}x_{t+1}\right\} \tag{29}$$

$$q_t^B = \Gamma E_t\left\{\Lambda_{t,t+1}q_{t+1}^B\right\} - \Gamma E_t\left\{\Lambda_{t,t+1}u_{t+1}\right\}, \tag{30}$$

in which $\tilde{c}_t^p = c_t^p - \alpha v(N_t/\alpha)$ is the consumption of market participants net of the "subsistence" level, and x denotes the productivity-adjusted stock of fundamental wealth

$$\begin{aligned} x_t &\equiv \frac{q_t^F}{\vartheta} + \alpha h_t \\ &= E_t\left\{\sum_{k=0}^{\infty}(\gamma\nu\Gamma)^k\Lambda_{t,t+k}\left[\frac{1-\tau^D}{\vartheta}d_{t+k} + \alpha\left(w_{t+k}N_{t+k}^{pe} - t_{t+k}^{pe} - v(N_{t+k}^{pe})\right)\right]\right\} \\ &= E_t\left\{\sum_{k=0}^{\infty}(\gamma\nu\Gamma)^k\Lambda_{t,t+k}\left[c_{t+k}^p - \alpha v(N_{t+k}^{pe})\right]\right\}, \end{aligned} \tag{31}$$

where the last equality follows from aggregation of the budget constraints (8) across all market participants, implying $c_t^p = \frac{1-\tau^D}{\vartheta}d_t + \alpha\left(w_t N_t^{pe} - t_t^{pe}\right)$.

We can use equations (28)–(29) to derive an IS-type relation for aggregate (adjusted) consumption of market participants

$$\tilde{c}_t^p = \frac{\nu\Gamma}{\beta}E_t\left\{\Lambda_{t,t+1}\tilde{c}_{t+1}^p\right\} + \frac{1-\beta\gamma}{\beta\gamma\vartheta}\left[q_t^B - \gamma\nu\Gamma E_t\left\{\Lambda_{t,t+1}q_{t+1}^B\right\}\right]. \tag{32}$$

Equation (32) shows that in this economy the wealth effect relevant for the dynamics of aggregate consumption is related to *bubbly* wealth only. This is a notable difference with respect to related frameworks, such as Nisticò (2016), where also *fundamental* financial wealth affects the dynamics of aggregate consumption. This difference is a direct implication of the assumption that the default probability for equity shares $(1-\nu\gamma)$ is equal to the probability that a household loses either her job or access to the asset market. Indeed, this effectively equates the rates at which people discount future dividends and future disposable labor income – as shown by the second line of equation (31) – and allows to express the overall fundamental wealth in the simple recursive formulation (29).

Notice that, on the one hand, equations (28) and (31) show that microfounding the wage equation (27), and the entailed complementarity between labor and consumption, implies that the definition of fundamental wealth that is relevant for consumption decisions also accounts for the discounted disutility of labor over the planning horizon, which captures the complementarity effects on future consumption. Therefore, a permanently higher disutility of labor tends to *increase* the desire to save, through a negative wealth effect on current consumption. On the other hand, equation (28) shows that a permanently higher disutility of labor also tends to *decrease* the desire to save, for a given stock of total wealth, through a positive complementarity effect on current consumption.

Using the definition above in equations (20)–(21) finally allows us to decompose the per-capita *adjusted* consumption in that of incumbents

$$\tilde{c}^{p}_{t|in} = (1-\beta\gamma)\left(\frac{b_t}{\gamma\vartheta} + \nu x_t\right), \tag{33}$$

and that of newcomers

$$\tilde{c}^{p}_{t|nc} = (1-\beta\gamma)\left[\frac{u_t}{\vartheta(1-\gamma)} + \frac{1-\gamma\nu}{1-\gamma}x_t\right]. \tag{34}$$

and to derive the "consumption gap" between the two groups

$$\begin{aligned}\tilde{c}^{p}_{t|in} - \tilde{c}^{p}_{t|nc} &= \frac{(1-\beta\gamma)}{\gamma}\left[\frac{b_t}{\vartheta} - \left(\frac{\gamma}{1-\gamma}\right)\frac{u_t}{\vartheta} - \left(\frac{1-\alpha}{\alpha}\right)x_t\right] &(35)\\ &= \frac{(1-\beta\gamma)}{\gamma}\left[\frac{q^B_t}{\vartheta} - \frac{u_t}{\vartheta\,(1-\gamma)} - \left(\frac{1-\alpha}{\alpha}\right)x_t\right]. &(36)\end{aligned}$$

The last equation plays a crucial role in the analysis below. It is the measure of consumption inequality that is relevant for social welfare, and it reflects the underlying wealth inequality between incumbents, who are relatively richer in terms of *bubbly* wealth given their higher longevity in asset markets, and newcomers, who are relatively richer in terms of *fundamental* wealth, since they are endowed with the shares of newly-set up firms and are all employed. Equation (36) shows that the effect of bubbly fluctuations on the "consumption gap" depends on the nature of the bubble and reflects the underlying heterogeneity among asset-market participants. Changes in pre-existing bubbles have opposite effects on the consumption gap compared to those in newly created bubbles, as the former only affect the consumption of incumbents while the latter only that of newcomers. On the other hand, changes in fundamental wealth affect both incumbents and newcomers, but the latter relatively more than the former.

3.1 The Balanced-Growth Paths

In a perfect-foresight BGP, productivity-adjusted variables (and hours worked) are constant. In particular, marginal costs are $(1+\mu)^{-1}$ and the productivity-adjusted real wage equals

$$w = \frac{1}{(1+\mu)(1-\tau^F)} = 1-\varpi,$$

where $\varpi \in [0,1)$ in the second equality defines the overall amount of monopolistic distortions along the BGP.[16] Along the BGP, equilibrium output – normalizing by productivity – is given by $y = N = \alpha\left(\frac{\delta}{1-\varpi}\right)^{-\frac{1}{\varphi}}$. It follows that an optimally set employment subsidy τ^F implies $\varpi = 0$ and

[16] Specifically, for non-negative employment subsidies, $\varpi \in [0, 1/\epsilon]$, with $\varpi = 0$ when $\tau^F = 1/\epsilon$ and $\varpi = 1/\epsilon = \frac{\mu}{1+\mu}$ when $\tau^F = 0$, the latter being the case in Galí (2021).

implements an efficient level of output along the BGP, given by $y = N = \alpha\delta^{-\frac{1}{\varphi}}$. Since the fiscal redistribution from market participants to rule-of-thumbers is zero along the BGP, from equation (7) it follows $c^r = (1-\varpi)y$ and from equation (23) $c^p = \left(1 + \varpi\frac{1-\vartheta}{\vartheta}\right) y$, further implying that an optimal employment subsidy also offsets the distributional consequences of monopolistic distortions along the BGP and implements a uniform cross-sectional distribution of average consumption *between* market participants and rule-of-thumbers: $c^r = c^p = y$.[17] Finally, the BGP-level of the disutility of labor is $v(N/\alpha) = (1-\varpi)\frac{y/\alpha}{1+\varphi}$, implying the following level for market participants' *adjusted* consumption:

$$\tilde{c}^p = c^p - \alpha v(N/\alpha) = \eta y \tag{37}$$

and for the fundamental wealth:

$$x = \frac{\eta}{1-\gamma\nu\Gamma\Lambda} y, \tag{38}$$

where $\eta \equiv \left[\frac{\varpi}{\vartheta} + (1-\varpi)\frac{\varphi}{1+\varphi}\right]$.

Using the above in the system (28)–(30) and some algebra finally yield

$$q^B = \eta\vartheta\frac{\gamma(\beta R - \nu)}{(1-\beta\gamma)(R-\gamma\nu)} \tag{39}$$

and

$$u = (1-R)\, q^B, \tag{40}$$

where we denote with q^B and u the BGP-level of the aggregate bubble-output and newly-created bubble-output ratios, respectively, and with $R \equiv (\Gamma\Lambda)^{-1} = \frac{1+r}{1+g}$ the ratio between the gross real interest rate and the gross growth rate of the economy.

Compared to the discussion in Galí (2021), we derive three additional results that underline the role of the additional margins arising in our economy and affecting the nature of bubbly BGPs, through the term $\eta\vartheta$ in equation (39).

Result 1 *The theoretical underpinnings of the wage schedule are not inconsequential for the size of rational bubbles along the BGP. Endogenous labor supply shrinks the size of equilibrium bubbles, the more so the higher the Frisch elasticity.*

Consider the polar case where $\vartheta = 1$, which nests Galí's economy extended with a microfounded labor supply. Then, equation (39) shows that the BGP-level of the bubble-output ratio depends also on inverse-Frisch elasticity of labor supply $\varphi > 0$:

$$q^B = \frac{\varpi + \varphi}{1+\varphi}\frac{\gamma(\beta R - \nu)}{(1-\beta\gamma)(R-\gamma\nu)}. \tag{41}$$

Therefore, finite values for φ imply that the size of the bubble-output ratio is smaller when the

[17] The cross-sectional distribution of consumption *within* the set of market participants is, instead, not uniform, since, otherwise, no room for bubbles would arise.

labor supply is more elastic (i.e. the smaller φ is), for any $\varpi \in [0, 1/\epsilon]$.[18]

To better highlight the intuition behind the role of φ, notice that the equilibrium average demand for consumption by market participants, in the absence of bubbles, can in general be written as[19]

$$c^p = \left(\eta \frac{1 - \beta\gamma}{1 - \gamma\nu\Gamma\Lambda} + \frac{1 - \varpi}{1 + \varphi} \right) y, \tag{42}$$

where the first term captures the fundamental-wealth effect and the second one the complementarity effect of labor, previously discussed, while their average income is

$$y^p = \left(1 + \varpi \frac{1 - \vartheta}{\vartheta} \right) y. \tag{43}$$

Using the definition of η then yields the economy-wide excess savings in the absence of bubbles:[20]

$$\vartheta(y^p - c^p) = \vartheta\eta \left(1 - \frac{1 - \beta\gamma}{1 - \gamma\nu\Gamma\Lambda} \right) y. \tag{44}$$

Therefore, the condition $\nu\Gamma\Lambda = \beta$ supports a bubbleless BGP equilibrium in our economy because per-capita consumption by market participants c^p equals their income y^p and all desired savings are thus absorbed by fundamental wealth. Instead, if $\nu\Gamma\Lambda < \beta$, then consumption falls short of income ($c^p < y^p$), as the discounted value of fundamental wealth goes down, and a role for bubbles to absorb the excess savings arises.

These excess savings, however, are smaller if φ is lower. On the one hand, a lower φ implies stronger complementarity effects on *future* consumption which *reduce* the desire to consume today because of a negative fundamental-wealth effect (through a lower η in (42)). On the other hand, it implies a stronger complementarity effect on *current* consumption as well, which instead *increases* the current desire to consume (second term in equation (42)). Since the relative weight on the negative fundamental-wealth effect is $\frac{1-\beta\gamma}{1-\gamma\nu\Gamma\Lambda}$ in (44), then $\nu\Gamma\Lambda < \beta$ implies that the positive complementarity effect on current consumption is always relatively stronger than the negative wealth effect, which explains the net fall in excess savings and the smaller room for bubbles.[21]

Result 2 *Heterogeneity in asset-market participation affects the size of rational bubbles along the BGP, and in general implies smaller equilibrium bubbles.*

To show how the stochastic participation in the asset market reduces the bubble size compared to Galí (2021), we consider now his original assumption of an *ad hoc* wage schedule, which here is equivalent to the limiting case $\varphi \to \infty$.[22] Then, equation (39) shows that the BGP-level of the

[18] The magnitude of the bubble-output ratio would be the same as in Galí (2021) – where the wage schedule is appended ex-post – only in the limiting case of an inelastic labor supply, i.e. $\varphi \to \infty$.

[19] This equation can be obtained from (28), taken at the BGP, using (38) and the aggregate disutility of labor.

[20] Recall that financially inactive agents have zero savings.

[21] Equation (43) shows that per-capita income is instead invariant with φ.

[22] We should clarify that the equivalence between the two models in the limiting case $\varphi \to \infty$ is limited to the BGP.

bubble-output ratio is scaled down depending on the share of market participants, $\vartheta \in [0,1]$:

$$q^B = [\varpi + (1-\varpi)\vartheta] \frac{\gamma(\beta R - \nu)}{(1-\beta\gamma)(R-\gamma\nu)}. \tag{45}$$

In general, equation (39) suggests (perhaps not surprisingly) that economies with larger shares of financially constrained agents are exposed to smaller bubble-output ratios along the BGP, regardless of the amount of distortions $\varpi \in [0, 1/\epsilon]$. Indeed, the size of the aggregate excess savings in the absence of bubbles is necessarily restricted to the share of the population ϑ that has access to financial markets, as also shown by equation (44), where $\vartheta\eta$ is an increasing function of ϑ.

Result 3 *As long as the economy features either i) endogenous and elastic labor supply (i.e. with finite φ) or ii) stochastic asset-market participation (i.e. $\vartheta < 1$), the extent to which the policy maker seeks to offset monopolistic distortions along the BGP also affects the size of BGP rational bubbles, and in general implies smaller equilibrium bubbles.*

Indeed, equation (39) directly implies that if $\vartheta = 1$ and $\varphi \to \infty$ the bubble-output ratio is independent of the BGP distortions ϖ, and equal to the case in Galí (2021). In all other – and more general – cases, however, q^B is an increasing function of ϖ because larger distortions reduce the complementarity effects of labor on consumption and stimulate savings. Hence, q^B is a decreasing function of the employment subsidy τ^F, with the minimum value associated with the optimal τ^F implying $\varpi = 0$.

We are interested in drawing normative implications of bubbly fluctuations in a linear-quadratic framework, using a second-order approximation of expected social welfare around the BGP. For the BGP to be consistent with an equilibrium allocation around which a quadratic Taylor expansion of expected social welfare is a valid second-order approximation of expected welfare when evaluated using only first-order-approximated equilibrium conditions, we assume the existence of an optimal employment subsidy in our baseline economy.

In our economy with stochastic asset-market participation, endogenous and elastic labor supply and optimal employment subsidy, therefore, the equilibrium bubble-output ratio along the BGP is:

$$q^B = \frac{\vartheta\varphi}{1+\varphi} \frac{\gamma(\beta R - \nu)}{(1-\beta\gamma)(R-\gamma\nu)}, \tag{46}$$

where all three additional margins are at work, compared to Galí (2021), thus implying a lower q^B.

The additional conditions for the existence of BGP equilibria with non-negative q^B and u hinge on the relative magnitude of ν and β, and on the real interest rate relative to the growth rate of the economy (with the aforementioned condition $r < g$ holding). As equations (40) and (46) clearly show, the conditions discussed in Galí (2021) – i.e. $R \in [\nu/\beta, 1]$ for $\nu \leq \beta$ – ensure $q^B \geq 0$ and $u \geq 0$ in our economy as well, since the assumption of finite lives is isomorphic to finite planning horizon for infinitely-lived agents due to stochastic asset-market participation.[23]

[23] Equations (40) and (46) would actually imply a positive value for both q^B and u also for levels of $R < \gamma\nu < \nu/\beta$.

More specifically, as discussed in the Appendix, the continuum of BGPs associated to the interval $R \in [R_0,\, 1]$, with $R_0 \equiv \nu/\beta$, can be partitioned into a subset of *stable* BGPs, for $\in [R_0,\, R^*]$ and *unstable* ones, for $R \in (R^*,\, 1]$, with

$$R^* = \gamma\nu + \sqrt{(\gamma\nu)^2 + \frac{\nu}{\beta}\left[1 - \gamma(\beta + \nu)\right]}. \tag{47}$$

3.2 The Linear Model

Consider a BGP of our economy with stochastic asset-market participation, endogenous labor supply and optimal employment subsidies, where the relative interest rate R lies in the range consistent with non-negative aggregate and new bubbles, i.e. $R \in [R_0,\, 1]$.

Taking a first-order approximation of the relevant equilibrium conditions around such a BGP, we can describe the private sector of our economy with the following system of five log-linear equations, where for a generic variable Z, we use the notation $\widehat{z}_t \equiv \log\left(\frac{Z_t}{Z_t^*}\right) = \log\left(\frac{Z_t}{z\Gamma^t}\right) = \log\left(\frac{z_t}{z}\right)$:[24]

$$\widehat{x}_t = \Phi E_t \widehat{x}_{t+1} - \frac{\varphi}{1+\varphi}\frac{\Phi}{1-\beta\gamma\Phi}\widehat{r}_t + \frac{1-\beta\gamma}{\vartheta\beta\gamma}\widehat{q}_t^B \tag{48}$$

$$\widehat{y}_t = \Theta\left(\frac{\widehat{q}_t^B}{\vartheta} + \widehat{x}_t\right) \tag{49}$$

$$\widehat{q}_t^B = \frac{\beta}{\nu}\Phi E_t \widehat{b}_{t+1} - q^B \widehat{r}_t \tag{50}$$

$$\widehat{q}_t^B = \widehat{b}_t + \widehat{u}_t \tag{51}$$

$$\widehat{\pi}_t = \beta\gamma\Phi E_t \widehat{\pi}_{t+1} + \kappa\widehat{y}_t, \tag{52}$$

in which $\Phi \equiv \frac{\nu\Gamma\Lambda}{\beta} = \frac{\nu}{R\beta}$, $\tau \equiv \frac{\tau^D(\varphi-\mu)}{(1-\vartheta)(1+\mu)}$, $\Theta \equiv \frac{\vartheta(1-\beta\gamma)}{(1-\vartheta)(\tau-\varphi)}$ and $\kappa \equiv \varphi\frac{(1-\theta)(1-\gamma\nu\Gamma\Lambda\theta)}{\theta}$ are composite parameters and $\widehat{r}_t \equiv \widehat{\imath}_t - E_t\widehat{\pi}_{t+1}$ defines the real interest rate, with $\widehat{\imath}_t \equiv \log\left(\frac{1+i_t}{1+r}\right)$ the nominal one.

Equation (48) describes the dynamics of fundamental wealth as a function of the real interest rate and the aggregate bubble; equation (49) determines the equilibrium level of the output gap, given the stock of total (fundamental and bubbly) wealth; equations (50) and (51) determine the law of motion of the aggregate bubble and its decomposition in pre-existing and new components, and equation (52) is the familiar New-Keynesian Phillips Curve describing the price-setting behavior of firms, in which the relative weights on expected inflation and marginal costs reflect the several layers of heterogeneity, compared to the standard New Keynesian model.

Result 4 *The theoretical underpinnings of the labor supply affect the extent to which monetary policy can control fluctuations in the bubble using its policy rate. An endogenous and elastic labor supply in general implies a lower interest-rate elasticity of equilibrium rational bubbles.*

These values of R are however ruled out because they are not consistent with a non-negative stock-market value q^F, as implied by the BGP-version of equation (26).

[24] The exceptions to this rule are: $\widehat{q}_t^B \equiv \frac{q_t^B}{y} - q^B$, $\widehat{b}_t \equiv \frac{b_t}{y} - b$, $\widehat{u}_t \equiv \frac{u_t}{y} - u$, $\widehat{x}_t \equiv \frac{x_t - x}{y}$, $\widehat{d}_t \equiv \frac{d_t - d}{y}$. Please refer to the Appendix for a full list of the non-linear and log-linear equilibrium conditions describing our economy.

This result is straightforward from equation (50), which shows that monetary policy can affect rational bubbles (in a first-order approximation) through the valuation effects that a change in the nominal interest rate implies. Since, however, these valuations effects are proportional to the aggregate bubble-output ratio along the BGP, q^B, any structural factor affecting the size of this ratio also affects the ability of the central bank to control bubbly fluctuations through its policy rate.

To highlight the role of the stochastic asset-market participation, and the heterogeneity that it implies, we establish the following two results.

Result 5 *In an economy with full asset-market participation (as in Galí, 2021) and endogenous labor supply, where the fiscal authority implements an efficient BGP through an optimal employment subsidy, the equilibrium path of the output gap is indeterminate.*

To see the intuition behind Result 5, which follows from Result 3, let us focus on equation (49) and consider the economy analyzed in Galí (2021), where $\vartheta = 1$, $\varpi > 0$ and $\tau^D = 0$. Microfounding the wage schedule (27) in that economy introduces an active role of the complementarity effects of labor on consumption in shaping the demand equation (49). Moreover, these complementarity effects depend on the amount of monopolistic distortions along the BGP ϖ, as already discussed. As a consequence, equation (49) in such an economy reads

$$\widehat{y}_t = \frac{1-\beta\gamma}{\varpi}\left(\widehat{q}_t^B + \widehat{x}_t\right).$$

Using the above in the dynamic equation for fundamental wealth (48) yields the following dynamic IS-type equation for the equilibrium output gap:

$$\widehat{y}_t = \Phi E_t \widehat{y}_{t+1} - \frac{\Phi}{\varpi}\left(\frac{\varpi+\varphi}{1+\varphi}\right)\left(\frac{1-\beta\gamma}{1-\beta\gamma\Phi}\right)\widehat{r}_t + \frac{1-\beta\gamma}{\varpi\beta\gamma}\left(\widehat{q}_t^B - \beta\gamma\Phi E_t \widehat{q}_{t+1}^B\right). \tag{53}$$

The above equation clarifies that both the interest-rate and bubble elasticities of the output gap in such an economy are inversely related to the amount of distortions in the BGP, and can therefore be quite (and counterfactually) large if the government is successful in reducing these distortions. In the limit, if the fiscal authority implements efficiency along the BGP through an optimal employment subsidy that implies $\varpi = 0$, these elasticities tend to infinity, like the multipliers on any demand shock possibly affecting equation (53).[25] The result would be that while the real interest rate and aggregate bubble evolve as in the flexible-price equilibrium discussed in Galí (2021), the equilibrium dynamics of the output gap is indeterminate.

Result 6 *In an economy with stochastic asset-market participation and endogenous labor supply, the equilibrium path of the output gap can be determinate even if the fiscal authority implements an efficient BGP through an optimal employment subsidy.*

[25] This implication derives directly from the choice of preferences that we need to make in order to microfound a labor supply with no wealth effects and therefore the wage equation (27). See also Auclert et al. (2021) for a related discussion of fiscal multipliers with GHH preferences.

In an economy with stochastic asset-market participation, indeed, the marginal propensity to consume out of total wealth, Θ, reflects the aggregation of the demand for consumption of both financially active and inactive agents, thus breaking the tight link between aggregate consumption and the complementarity effect of labor on the consumption of market participants. As a consequence, in our baseline economy the interest-rate and bubble elasticities of the output gap are finite, also in the case of an efficient BGP, and the implied IS-type equation is[26]

$$\widehat{y}_t = \Phi E_t \widehat{y}_{t+1} - \frac{\varphi}{1+\varphi}\frac{\Theta\Phi}{1-\beta\gamma\Phi}\widehat{r}_t + \frac{\Theta}{\vartheta\beta\gamma}\left(\widehat{q}_t^B - \beta\gamma\Phi E_t \widehat{q}_{t+1}^B\right). \tag{54}$$

4 Rational Bubbles and Monetary Policy Tradeoffs

In this Section, we discuss the normative implications of rational bubbles for monetary policy in our baseline economy with stochastic asset-market participation and endogenous labor supply.

In particular, we are interested in the Ramsey policy that maximizes the expected social welfare

$$\mathcal{W}_{t_0} \equiv E_{t_0}\left\{\sum_{t=t_0}^{\infty}\beta^{t-t_0}U_t\right\}, \tag{55}$$

where the period-utility U_t is a weighted average of the individual utilities in the economy at time t

$$U_t \equiv \sum_{j\in\mathcal{T}}\sum_{s=-\infty}^{t}\chi_s^j U_{t|s}^j \tag{56}$$

and $\{\chi_s^j\}$ is a system of Pareto-weights, with $j \in \mathcal{T} = \{pe,\, pu,\, re,\, ru\}$ indexing the agent type, and $s = -\infty, ..., t-1, t$ the generic time of transition in or out of financial markets, and such that

$$\sum_{j\in\mathcal{T}}\sum_{s=-\infty}^{t}\chi_s^j = 1.$$

To evaluate the implied tradeoffs and derive the optimal monetary policy, we can use a purely quadratic loss function deriving from a second-order approximation of (55) given (56) around an efficient BGP.[27] The BGP, in turn, is efficient if it is consistent with the solution of the Ramsey problem that maximizes (55) given (56) under the resource and technological constraint

$$\Gamma^t\left[\sum_{s=-\infty}^{t} m_{t|s}^{pe}N_{t|s}^{*pe} + \sum_{s=-\infty}^{t} m_{t|s}^{re}N_{t|s}^{*re}\right] = Y_t^* = C_t^* = \sum_{j\in\mathcal{T}}\sum_{s=-\infty}^{t} m_{t|s}^{j}C_{t|s}^{*j}, \tag{57}$$

[26] The non-separability of preferences also exacerbates the tendency of the model to display the *"inverted aggregate demand logic"* discussed in Bilbiie (2008) due to the limited asset-market participation, which here would also imply a negative bubble-elasticity of output. The redistribution of the dividend-tax revenues to rule-of-thumbers allows us to focus on the (arguably more realistic) case of positive bubble-elasticity of output and *"standard aggregate demand logic"*.

[27] To be more accurate, we focus on a limited-efficient BGP, insofar as we impose that a subset of the agents in the economy is unemployed, as shown by contraint (57).

where X_t^* denotes the BGP-level of generic variable X, and $m_{t|s}^j$ is the relative mass of agents of type j and cohort $s \leq t$, with

$$\sum_{j \in \mathcal{T}} \sum_{s=-\infty}^{t} m_{t|s}^j = 1.$$

As shown in the Appendix, the efficiency of the BGP requires an appropriate system of Pareto-weights that supports a given initial cross-sectional distribution of wealth and consumption across different agent-types, and an appropriate employment subsidy that offsets monopolistic distortions and thus implies $\varpi = 0$. Under these two restrictions, a quadratic Taylor expansion of (55) is a valid second-order approximation of expected social welfare that can be evaluated using only first-order-approximated equilibrium conditions.

In the Appendix, we show that such second-order Taylor expansion of expected social welfare leads to the following quadratic loss function:

$$\mathcal{L}_{t_0} \equiv -\mathcal{W}_{t_0} = \frac{1}{2} \frac{\varepsilon \varphi}{\kappa} E_{t_0} \left\{ \sum_{t=t_0}^{\infty} \beta^{t-t_0} \left(\widehat{\pi}_t^2 + \alpha_y \widehat{y}_t^2 + \alpha_\omega \widehat{\omega}_t^2 \right) \right\}, \tag{58}$$

where $\widehat{\omega}_t$ captures the welfare losses coming from variations in consumption dispersion among market participants relative to the BGP,[28] and the relative welfare weights are defined as

$$\alpha_y \equiv \frac{\kappa}{\varphi \varepsilon} \left[\varphi + \left(\frac{1+\varphi}{\varphi} \right) \left(\frac{1-\vartheta}{\vartheta} \right) (\tau - \varphi)^2 \right] \tag{59}$$

$$\alpha_\omega \equiv \frac{\kappa \vartheta}{\varepsilon \varphi} \left(\frac{1+\varphi}{\varphi} \right) \frac{(1-\gamma)(1-\beta\gamma)}{\gamma}. \tag{60}$$

Cross-sectional consumption dispersion originates from two of the three layers of heterogeneity characterizing our economy: the dispersion *between* financially active and inactive agents, and the dispersion *within* the set of market participants, related to the individual longevity in the type. The former is proportional to squared output gap, and is reflected in the second addendum in its welfare weight, (59). As shown in the Appendix, instead, the cross-sectional consumption dispersion *within* the set of market participants $\widehat{\Delta}_{c,t}^p$ evolves according to the following law of motion:

$$\widehat{\Delta}_{c,t}^p = \gamma \widehat{\Delta}_{c,t-1}^p + \frac{1-\gamma}{\gamma} \left[\frac{(1+\varphi)(1-\beta\gamma)}{\varphi} \right]^2 \widehat{\omega}_t^2 \tag{61}$$

[28] Since losses come symmetrically from lower and higher consumption inequality in the monetary-policy loss function (58), the central bank does not aim to reduce structural consumption/wealth inequality, consistently with the view of Bernanke (2015). Rather, it aims to dampen temporary fluctuations in wealth components distorting its overall distribution (and that of consumption) relative to the long-run counterpart.

and thus ultimately depends on the "consumption gap" defined in equation (36)

$$\widehat{\omega}_t \equiv \frac{\widehat{b}_t}{\vartheta} - \frac{\gamma}{\vartheta(1-\gamma)}\widehat{u}_t - \frac{1-\alpha}{\alpha}\widehat{x}_t \tag{62}$$

$$= \frac{1}{\alpha\vartheta}\widehat{q}_t^B - \frac{\widehat{u}_t}{\vartheta(1-\gamma)} - \frac{1-\alpha}{\alpha\Theta}\widehat{y}_t \tag{63}$$

$$= \frac{\gamma}{1-\beta\gamma}\left(\widehat{\widetilde{c}}^p_{t|in} - \widehat{\widetilde{c}}^p_{t|nc}\right). \tag{64}$$

The additional term $\widehat{\omega}_t$ in the welfare criterion (58), therefore, responds to bubbly fluctuations along two dimensions: *i*) the relative size of fluctuations in *pre-existing* versus *new* bubbles, and *ii*) the relative size of fluctuations in *bubbly* versus *fundamental* wealth. Indeed, changes in existing bubbles affect only the consumption of incumbents, while changes in new bubbles only affect the consumption of newcomers. On the other hand, changes in fundamental wealth affect the consumption of both, but more than proportionately the one of newcomers, which are entitled to a larger per-capita share of human wealth (being entirely employed) and of fundamental financial wealth (holding the whole lot of new shares).

A straightforward implication of loss (58) is the following result.

Result 7 *Strict inflation targeting is generally not an optimal monetary-policy regime. Despite the "divine coincidence" implied by equation (52), bubbly fluctuations imply an endogenous tradeoff between inflation/output-gap stability on the one hand, and consumption dispersion on the other.*

Indeed, given the specification of the firm problem and the ensuing Phillips Curve, the "divine coincidence" applies in our economy, so that output gap and inflation can be stabilized simultaneously. Nevertheless, a straightforward implication of the loss function (58) is that pursuing the flexible-price allocation by stabilizing inflation, and thus the output gap, is generally not an optimal policy from a welfare perspective, and an endogenous tradeoff arises.[29]

Given the definition of $\widehat{\omega}_t$, indeed, the flexible-price allocation maximizes social welfare only in one of two cases. The first is when there are no bubble fluctuations whatsoever ($\widehat{q}_t^B = \widehat{u}_t = 0$ for all t). In this case, stabilizing the output gap not only achieves stabilization of inflation, but also of fundamental wealth $\widehat{x}_t$, as implied by (49), and ultimately implies zero welfare losses, i.e. $\widehat{y}_t = \widehat{\pi}_t = \widehat{\omega}_t = 0$.

The second case is the fortuitous one in which $\widehat{q}_t^B = \frac{\alpha\widehat{u}_t}{1-\gamma}$ for all t, whereby again $\widehat{y}_t = \widehat{\pi}_t = \widehat{\omega}_t = 0$. In this case, fluctuations in the old bubble are such that they perfectly offset those in the new bubble leaving the consumption gap unaffected along dimension *i*). Pursuing stability of the output gap finally ensures that consumption dispersion is also unaffected along dimension *ii*).

In all other and more general cases, instead, a welfare-maximizing central bank has an incentive to allow fluctuations in output (and inflation) in order to reduce the effects of bubbly fluctuations on cross-sectional consumption dispersion. Whether this incentive translates into actual deviations

[29] See also Nisticò (2016), for an analogous tradeoff arising in a related environment.

from strict inflation targeting under optimal policy, depends also on the nature of the BGP around which the economy fluctuates.

Result 8 *In an economy where bubble fluctuations can only arise from pre-existing bubbles (i.e. $\widehat{u}_t = 0$ for all t), the optimality of strict inflation targeting from a welfare perspective depends on the global stability properties of the BGP around which the economy fluctuates:*

a) if the BGP is globally unstable, price stability is an optimal policy regime, and the associated rational expectations equilibrium rules out bubble fluctuations altogether

b) if the BGP is globally stable, price stability is not an optimal policy regime, as the associated rational expectations equilibrium cannot rule out sunspot fluctuations in existing bubbles.

In case a) of Result 8, indeed, the only stationary equilibrium with rational expectations also implies full stabilization of pre-existing bubbles, i.e. $\widehat{q}_t^B = \widehat{b}_t = 0$, and thus zero welfare losses. To see this, notice that imposing $\widehat{u}_t = \widehat{y}_t = \widehat{\pi}_t = 0$ for all t in system (48)–(52) implies the following equilibrium condition for the old bubble

$$\widehat{b}_t = \Psi E_t \widehat{b}_{t+1}, \tag{65}$$

with $\Psi \equiv \Phi \left[1 + (\Lambda\Gamma - 1)\frac{1-\beta\gamma}{1-\beta\gamma\Phi}\right]$, which equals the equilibrium condition arising in Galí (2021) under flexible prices for the aggregate bubble, when new bubbles are unpredictable. Equation (65) admits $\widehat{b}_t = 0$ as the *unique* stationary solution only when $\Psi < 1$. To see how this restriction coincides with the BGP being *unstable*, note that we can use the definitions $R \equiv (\Lambda\Gamma)^{-1}$ and $\Phi \equiv \frac{\nu\Lambda\Gamma}{\beta} = \nu/\beta R$ to rewrite Ψ as a decreasing function of R:

$$\Psi = \frac{\nu}{\beta R}\left[1 + (1-R)\frac{1-\beta\gamma}{R-\nu\gamma}\right].$$

Now, note that, when the relative interest rate is at its highest level consistent with non-negative bubbles along the BGP, $R = 1$, then $\Psi = \nu/\beta$, which is less than one provided that $\nu < \beta$, as we are assuming throughout. On the other hand, when the relative interest rate is at the lower end of its admissible levels, $R = \nu/\beta$, then $\Psi = \beta/\nu$, which is in turn higher than one.

For all the levels in between, it can be easily shown that the threshold value for the relative interest rate implying $\Psi = 1$ solves

$$\beta R^2 - 2\beta\gamma\nu R - \nu(1 - \beta\gamma - \nu\gamma) = 0,$$

which is the same polynomial admitting R^* as a root, as we show in the Appendix. Therefore, $\Psi < 1$ requires $R \in (R^*, 1]$, and thus the BGP to be *unstable*.

On the other hand, were the BGP *stable* – i.e. $R \in [\nu/\beta, R^*]$, and thus $\Psi > 1$, as in case b) of Result 8 – a multiplicity of stationary *sunspot* solutions would arise, triggered by any unanticipated change in pre-existing bubbles – arguably the most realistic case when it comes to bubbly fluctuations – which would then make the strict-inflation targeting regime no longer optimal. This implies that a low interest rate environment is not only delicate because it makes the rise of bubbles possible in equilibrium, but also because of its monetary policy implications, potentially requiring deviations from inflation targeting.

In order to evaluate this case and other more general ones (for example, new bubbles fluctuations), we now turn to the analysis of optimal monetary policy.

4.1 Optimal Monetary Policy

The optimal monetary policy problem can be characterized as the minimization of loss (58) under the system of constraints (48)-(52). Under discretion, the optimal policy chooses output and inflation in order to minimize the period-loss function

$$\frac{1}{2}\left(\widehat{\pi}_t^2 + \alpha_y \widehat{y}_t^2 + \alpha_\omega \widehat{\omega}_t^2\right),$$

subject to the constraints

$$\widehat{y}_t = \frac{\Theta}{\vartheta}\frac{1+\chi}{\chi}\widehat{q}_t^B + K_{y,t}\widehat{\pi}_t \qquad\qquad = \kappa\widehat{y}_t + K_{\pi,t},$$

given definition (63), and where the first constraint can be derived by combining (48)–(51).[30] Moreover, $K_{x,t}$ and $K_{\pi,t}$ collect expectational terms that are unaffected in the discretionary equilibrium, and $\chi \equiv (1-\Phi)\,\frac{\beta\gamma}{1-\beta\gamma}$. The solution to this problem implies the optimal targeting rule

$$\Theta\alpha_y\widehat{y}_t + \Theta\kappa\widehat{\pi}_t = \alpha_\omega \frac{1-\alpha(1+\chi)}{\alpha(1+\chi)}\widehat{\omega}_t, \tag{66}$$

which disciplines how to optimally trade off output and inflation stability for less consumption dispersion across market participants.

The targeting rule (66), given the definition of χ, Φ and α, directly implies the following

Result 9 *The stringency of the additional tradeoff implied by bubbly fluctuations is increasing in the transition probability out of asset markets $1-\gamma$, and decreasing in the BGP-level of the relative interest rate R, through the effect of these parameters on the BGP-level of the bubble-output ratio q^B. In particular, such tradeoff is most stringent when the BGP is bubbleless, which requires $R = \nu/\beta < 1$, $\Phi = 1$ and therefore $q^B = \chi = 0$.*

Indeed, for $\nu < \beta$, $\Phi \in [\nu/\beta, 1]$ is inversely related to the bubble-output ratio along the BGP, (46), and thus it reaches its highest value (and χ its lowest value in (66)) in a bubbleless BPG. We view the result that the implied tradeoff is most stringent when the BGP is bubbleless as particularly insightful for two reasons. First, we can think of a bubbleless BGP as a reasonably realistic

[30] We report here the analysis of the discretionary equilibrium. The equilibrium under either constrained or unconstrained commitment does not add much to the insights we are able to derive, analytically, under discretion. Details are available upon request.

description of an economy where boom-and-bust cycles in asset prices imply bubbly fluctuations that eventually revert back to a bubbleless long-run equilibrium. The second reason is that a bubbleless BGP associated with a low interest rate environment is necessarily globally *stable* and thus allows for sunspot fluctuations in the aggregate bubble. Therefore, the additional tradeoff implied by bubbly fluctuations is most stringent in a case that is not only the arguably most realistic one, but also the one where it is most relevant, given that strict inflation targeting is in general not optimal around a globally *stable* BGP, as suggested by Result 8.

The intuition behind Result 9 is straightforward and is related to the role of valuation effects for the ability of monetary policy to directly affect fluctuations in the aggregate bubble. Indeed, as implied by equation (50), a bubbleless BGP implies that these valuation effects are nil in a first-order approximation of the model, and monetary policy is therefore unable to affect the bubble directly by changing its policy rate. As a consequence, when the BGP is bubbleless, monetary policy can only lever on the output gap and fundamental wealth to dampen the effects of bubbly fluctuations on consumption dispersion, clearly implying potentially larger deviations from price stability. On the contrary, if the BGP is bubbly, monetary policy can affect bubble fluctuations directly in order to optimally steer consumption dispersion, thereby requiring a smaller deviation from output-gap and price stability.

Hence, in a bubbleless BGP where $q^B = \chi = 0$ and $\Phi = 1$, the aggregate bubble evolves autonomously and independently of monetary policy, given the absence of valuation effects

$$\widehat{q}_t^B = \frac{\beta}{\nu} E_t \widehat{q}_{t+1}^B - \frac{\beta}{\nu} E_t \widehat{u}_{t+1} \tag{67}$$

and the optimal targeting rule requires

$$\Theta \alpha_y \widehat{y}_t + \Theta \kappa \widehat{\pi}_t = \alpha_\omega \left(\frac{1-\alpha}{\alpha} \right) \widehat{\omega}_t. \tag{68}$$

Moreover, if we focus on the case $\nu < \beta$ (which is required for bubble fluctuations to arise as an equilibrium outcome in the first place) and considering that the bubbleless BGP is globally *stable*, equation (67) admits stationary solutions of the form

$$\widehat{q}_t^B = R_0 \widehat{q}_{t-1}^B + e_t, \tag{69}$$

where $R_0 \equiv \nu/\beta < 1$ and $e_t \equiv \widehat{b}_t - E_{t-1}\{\widehat{b}_t\} + \widehat{u}_t$ is a martingale difference process.[31]

Therefore, self-fulfilling revisions in expectations about the future size of currently existing bubbles are able, through e_t, to exert identical *positive* implications for the dynamics of the aggregate bubble, regardless of whether these revisions apply to bubbles that have just arisen in the current period, $\widehat{u}_t$, or ones that are surviving from the past, $\widehat{b}_t$. On the contrary, this difference can be relevant when it comes to the *normative* implications of these sunspot shocks.

[31] We are restricting attention to the case where future new bubbles are always unpredictable, i.e. $E_t\{\widehat{u}_{t+k}\} = 0$ for $k = 1, 2, \ldots$ and all t.

Result 10 *The optimal policy response to bubble shocks critically depends on the nature of the innovation, and on the way in which bubble fluctuations affect the cross-sectional consumption distribution across heterogeneous agents. Compared to the equilibrium outcome under an inflation-targeting regime, given the structure of our economy:*

a) a shock to pre-existing bubbles requires a persistent accommodation of the inflationary and expansionary effects of the bubble shock

b) a shock to newly created bubbles requires leaning against the inflationary and expansionary effects on impact, and accommodating them in the transition.

Consider first an unexpected transitory shock to pre-existing bubbles, i.e. $e_t = e_t^b = \widehat{b}_t > 0$ and $E_{t-1}\{\widehat{b}_t\} = \widehat{u}_t = 0$. Using the targeting rule (68) in the system of constraints, along with the definition of $\widehat{\omega}_t$, we obtain the optimal state-contingent path for the welfare-relevant variables:

$$\widehat{\pi}_t = \frac{\psi_\pi^q R_0}{\vartheta}\widehat{q}_{t-1}^B + \frac{\psi_\pi^b}{\vartheta}e_t^b \tag{70}$$

$$\widehat{y}_t = \frac{\psi_y^q R_0}{\vartheta}\widehat{q}_{t-1}^B + \frac{\psi_y^b}{\vartheta}e_t^b \tag{71}$$

$$\widehat{\omega}_t = \frac{\psi_\omega^q R_0}{\vartheta}\widehat{q}_{t-1}^B + \frac{\psi_\omega^b}{\vartheta}e_t^b \tag{72}$$

where

$$\psi_\pi^q \equiv \frac{\kappa\Xi_q}{1-\nu\gamma+\kappa\Xi_\pi}, \qquad \psi_y^q \equiv \Xi_q - \Xi_\pi\psi_\pi^q, \qquad \psi_\omega^q \equiv \frac{1}{\alpha} - \frac{\psi_y^q(1-\alpha)}{\alpha\Theta}$$

given

$$\Xi_q \equiv \frac{\alpha_\omega\frac{1-\alpha}{\alpha^2\Theta}}{\alpha_y + \alpha_\omega\left(\frac{1-\alpha}{\alpha\Theta}\right)^2} > 0 \qquad \Xi_\pi \equiv \frac{\kappa}{\alpha_y + \alpha_\omega\left(\frac{1-\alpha}{\alpha\Theta}\right)^2} > 0$$

and $\psi_\pi^b = \psi_\pi^q$, $\psi_y^b = \psi_y^q$, $\psi_\omega^b = \psi_\omega^q$. Note that, under inflation targeting (IT) – which here would arise if the central banker chose $\alpha_\omega = 0$ – the state-contingent path of the welfare-relevant variables follow the system (70)–(72) with coefficients

$$\psi_\pi^{q,IT} = \psi_\pi^{b,IT} = 0 \qquad \psi_y^{q,IT} = \psi_y^{b,IT} = 0 \qquad \psi_\omega^{q,IT} = \psi_\omega^{b,IT} = \frac{1}{\alpha}.$$

Moreover, simple algebra shows that ψ_π^q, $\psi_y^q > 0$ and $\psi_\omega^q < 1/\alpha$.

In this class of models, therefore, in response to an upward revision in the expected value of bubbles that were already traded in asset markets, a welfare-maximizing central bank allows for the inflationary and expansionary effect of the bubble to partially pass through, in order to dampen the effect on consumption dispersion. This response is markedly different from that of an inflation-targeting central bank, which would instead increase the policy rate more aggressively in order to fully stabilize inflation and the output gap, at the cost of more volatile consumption dispersion.[32] Moreover, note that, despite the transitory nature of the bubble innovation e_t^b, the strong persistence in the aggregate bubble implied by (69) is reflected in the optimal deviation

[32]The implied response of the real interest rate can be easily derived using equation (54).

from inflation targeting: inflation and the output gap are persistently higher, while consumption dispersion and the real interest rate are persistently lower, both on impact and during the transition.

The optimal policy in response to a sunspot shock to pre-existing bubbles is therefore less contractionary than under inflation targeting. The intuition behind this response can be understood by focusing on the implied response of fundamental wealth. Under inflation targeting, the need to stabilize the output gap requires cutting fundamental wealth as much as needed to completely offset the increase in the aggregate bubble, as implied by equation (49), regardless of the nature of the underlying sunspot shock. Under optimal policy, instead, such nature is key. If the revision in expectations is related to pre-existing bubbles, indeed, this has an expansionary effect on the consumption of incumbents only, thereby raising the "consumption gap", as shown by (62). On the other hand, cutting fundamental wealth so as to stabilize output gap would reduce the consumption of both incumbents and newcomers, but the latter relatively more – as implied by equations (33)–(34) – thus further increasing consumption dispersion. Under optimal policy, therefore, fundamental wealth falls less than under inflation targeting in order to dampen the response of consumption dispersion, and can even increase, depending on the relative size of ψ_y^b and Θ:

$$\widehat{x}_t = -\frac{1}{\vartheta}\left(1 - \frac{\psi_y^q}{\Theta}\right)\left(R_0 \widehat{q}_{t-1}^B + e_t^b\right).$$

To see how bubbles of different nature may have different normative implications, consider now an unexpected transitory shock to newly created bubbles, i.e. $e_t = e_t^u = \widehat{u}_t > 0$ and $E_{t-1}\{\widehat{b}_t\} = \widehat{b}_t = 0$. Using this with the targeting rule (68) and the definition of $\widehat{\omega}_t$ in the system of constraints, we can show that the optimal state-contingent path for the welfare-relevant variables now reads:

$$\widehat{\pi}_t = \frac{\psi_\pi^q R_0}{\vartheta} \widehat{q}_{t-1}^B - \frac{\psi_\pi^u}{\vartheta} e_t^u \tag{73}$$

$$\widehat{y}_t = \frac{\psi_y^q R_0}{\vartheta} \widehat{q}_{t-1}^B - \frac{\psi_y^u}{\vartheta} e_t^u \tag{74}$$

$$\widehat{\omega}_t = \frac{\psi_\omega^q R_0}{\vartheta} \widehat{q}_{t-1}^B - \frac{\psi_\omega^u}{\vartheta} e_t^u \tag{75}$$

where

$$\psi_\pi^u \equiv \frac{\alpha+\gamma-1}{1-\gamma}\psi_\pi^q \frac{\kappa \Xi_\pi}{1+\kappa\Xi_\pi}, \qquad \psi_y^u \equiv \frac{\alpha+\gamma-1}{1-\gamma}\Xi_q - \Xi_\pi \psi_\pi^u, \qquad \psi_\omega^u \equiv \frac{\alpha+\gamma-1}{\alpha(1-\gamma)} - \frac{\psi_y^u(1-\alpha)}{\alpha\Theta},$$

and ψ_π^q, ψ_y^q, ψ_ω^q, Ξ_q and Ξ_π have been previously defined. Moreover, it can easily be shown that $\psi_\pi^u > \psi_\pi^{u,IT}$, $\psi_y^u > \psi_y^{u,IT}$ and $\psi_\omega^u < \psi_\omega^{u,IT}$, where $\psi_\pi^{u,IT} = 0$, $\psi_y^{u,IT} = 0$ and $\psi_\omega^{u,IT} = \frac{\alpha+\gamma-1}{\alpha(1-\gamma)}$ are the corresponding response coefficients under inflation targeting.

Two implications of system (73)–(75) are particularly worth noticing, compared to the case of innovation in the old bubble. First, in order to dampen the effect on consumption dispersion, the optimal response on impact to an upward revision in the expected value of bubbles that are newly created leans against its inflationary and expansionary effects. This is shown by the second term in each of equations (73)–(74). The intuition behind this result is straightforward, and again it is instructive to focus on the response of fundamental wealth. An increase in the value of new bubbles raises the consumption of newcomers only, reducing consumption dispersion below the efficient BGP-level, *ceteris paribus*. As a consequence, a welfare-maximizing central bank finds it optimal to induce a larger fall in fundamental wealth (compared to the inflation-targeting regime), since such a fall lowers the consumption of newcomers relatively more than that of incumbents, thus dampening the effect on cross-sectional consumption dispersion:

$$\widehat{x}_t = -\frac{R_0}{\vartheta}\left(1 - \frac{\psi_y^q}{\Theta}\right)\widehat{q}_{t-1}^B - \frac{1}{\vartheta}\left(1 + \frac{\psi_y^u}{\Theta}\right)e_t^u.$$

Second, the optimal response in the transition has the opposite sign with respect to that on impact. Indeed, while in period t the bubble shock impacts the consumption of newcomers only, from period $t+1$ onward, the persistency of the aggregate bubble dynamics affects the consumption of incumbents, making the analysis of the previous case again relevant for the transition. This is clearly shown by the first term in each of equations (73)–(75), which are indeed identical to those in equations (70)–(72).

5 Conclusions

We analyze the interplay between agents' heterogeneity and rational bubbles in asset prices, and study the normative implications for monetary policy in a tractable New Keynesian model with heterogeneous agents, where the heterogeneity stems from stochastic participation in asset and labor markets.

Although agents in our economy are infinitely-lived, their structural heterogeneity satisfies the same conditions derived in related OLG models for the rise of rational bubbles in equilibrium. Additionally, we emphasize the role of endogenous labor supply decisions, heterogeneity in asset-market participation and monopolistic distortions in determining the equilibrium size of bubbles in the balanced-growth path and the shape of bubbly fluctuations over the business cycle.

We also derive a monetary-policy loss function microfounded on a second-order approximation of social welfare, which emphasizes the relevance of bubbly fluctuations as an additional target of policy, through their effect on cross-sectional consumption dispersion. We show that, in this environment, strict inflation targeting is generally not an optimal monetary-policy regime, despite the "divine coincidence" that characterizes our economy, and that bubbly fluctuations imply an endogenous tradeoff between inflation/output-gap stability on the one hand, and cross-sectional inequality on the other.

This additional tradeoff is more stringent for monetary policy when business cycle fluctuations occur around a balanced-growth path with small or no bubbles, as in this case the central bank has limited or no ability to affect bubbles directly in the pursuit of its welfare objectives. When interest rates are very low in the balanced-growth path, moreover, our analysis suggests that bubble fluctuations can arise from self-fulfilling revisions in expectations about the value of pre-existing bubbly assets. This type of bubbly fluctuations in particular requires a welfare-maximizing central bank to deviate from strict inflation targeting, with the type of policy response depending also on the specific nature of the bubble innovations.

Several extensions seem promising avenues for future research. Our work points to cross-sectional wealth and consumption distribution as the main channel through which bubble fluctuations may require a monetary policy response from a welfare perspective. Moreover, in our economy different kinds of bubbles may affect inequality in different ways, requiring different policy responses. This implies that the specific cross-sectional distribution of bubbly assets in the economy is key to understanding what the optimal monetary response to bubble shocks should be. In this respect, our results are sensitive to the specific assumptions made about the distribution of newly-created bubbly assets. A more comprehensive – or empirically accurate – account of this margin would be interesting to explore.[33]

Analogously, a more general version of the model in which the default probability of firms is not equal to the probability that a household loses either her job or access to the asset market would imply an additional margin relevant from a distributional perspective. In that case, indeed, the wealth effect on aggregate consumption dynamics includes also *fundamental* financial wealth, which also has an additional relevance for welfare, as in Nisticò (2016). This would likely imply an additional channel through which monetary policy can affect consumption dispersion, and might change the optimal policy response to bubbly fluctuations.

In this paper, we implicitly focus on a single monetary policy tool – the nominal interest rate – in a cashless economy. Introducing monetary aggregates in our framework would be an interesting extension for at least two reasons. The first is that it would introduce an alternative policy tool, such as asset purchase programs, that can have non-trivial distributional effects on heterogeneous agents. The second is that cash is a *public* bubble that could be used to replace the *private* ones, with the additional benefit of being more easily controllable by the monetary authority.[34]

Finally, our paper focuses on a stylized description of the economy that allows to derive analytical results and focus on the theoretical mechanisms that relate bubbly fluctuations to agents' heterogeneity and the implied margins that are relevant for monetary policy from a welfare perspective. A fully-fledged quantitative version of our model would help in shedding light on the quantitative relevance of these margins, and the extent to which simpler monetary policy regimes such as inflation targeting are still a good approximation of optimal policy also in the face of bubbly fluctuations affecting wealth inequality.

[33]Data on the U.S., for example, show that housing wealth is concentrated at the middle of the wealth distribution, unlike stock holdings that are concentrated at the top (Kuhn et al., 2020). Through the lens of our model, this suggests that bubbly episodes related to these two segments of the asset market might require potentially different policy responses from a welfare perspective.

[34]See Asriyan et al. (2021).

References

[1] **Asriyan, Vladimir, Fornaro, Luca, Martin, Alberto and Jaume Ventura**. 2021. "Monetary Policy for a Bubbly World", *Review of Economic Studies*, 88 (3): 1418-1456.

[2] **Auclert, Adrien, Bardóczy, Bence and Matthew Rognlie**. 2021. "MPCs, MPEs, and Multipliers: A Trilemma for New Keynesian Models", *Review of Economics and Statistics*, forthcoming.

[3] **Bernanke, Ben.** 2015. "Monetary Policy and Inequality", Brookings Blog, June 1.

[4] **Bilbiie, Florin.** 2008. "Limited Asset Markets Participation, Monetary Policy and (Inverted) Aggregate Demand Logic", *Journal of Economic Theory*, 140 (1): 162-196.

[5] **Bilbiie, Florin.** 2018. "Monetary Policy and Heterogeneity: An Analytical Framework", CEPR Discussion Papers 12601, C.E.P.R. Discussion Papers.

[6] **Blanchard, Olivier.** 1985. "Debt, Deficits, and Finite Horizons", *Journal of Political Economy*, 93 (2): 223-247.

[7] **Calvo, Guillermo**. 1983. "Staggered Prices in a Utility-Maximizing Framework", *Journal of Monetary Economics*, 12 (3): 983-998.

[8] **Debortoli, Davide and Jordi Galí.** 2018. "Monetary Policy with Heterogeneous Agents: Insights from TANK models", Mimeo.

[9] **Galí, Jordi**. 2014. "Monetary Policy and Rational Asset Price Bubbles", *American Economic Review*, 104 (3): 721-752.

[10] **Galí, Jordi**. 2021. "Monetary Policy and Bubbles in a New Keynesian Model with Overlapping Generations", *American Economic Journal: Macroeconomics*, 13 (2): 121-167.

[11] **Galí, Jordi, López-Salido, David and Javier Vallés**. 2007. "Understanding the Effects of Government Spending on Consumption", *Journal of the European Economic Association*, 5 (1): 227-270.

[12] **Greenwood, Jeremy, Hercowitz, Zvi and Gregory Huffman**. 1988. "Investment, Capacity Utilization, and the Real Business Cycle", *American Economic Review*, 78 (3): 402-417.

[13] **Hirano, Tomohiro and Noriyuki Yanagawa.** 2017. "Asset Bubbles, Endogenous Growth and Financial Frictions", *Review of Economic Studies*, 84 (1): 406-443.

[14] **Jaimovich, Nir and Sergio Rebelo.** 2009. "Can News about the Future Drive the Business Cycle?", *American Economic Review*, 99 (4): 1097-1118.

[15] **Kocherlakota, Narayana.** 1992. "Bubbles and Constraints on Debt Accumulation", *Journal of Economic Theory*, 57: 245-256.

[16] **Kohn, Donald.** 2009. "Monetary Policy and Asset Prices Revisited", *Cato Journal*, 29 (1): 31-44.

[17] **Kuhn, Moritz, Schularick, Moritz and Ulrike Steins.** 2020. "Income and Wealth Inequality in America, 1949â€“2016" *Journal of Political Economy*, 128 (9): 3469-3519.

[18] **Miao, Jianjun.** 2014. "Introduction to Economic Theory of Bubbles," *Journal of Mathematical Economics*, 53: 130-136.

[19] **Miao, Jianjun, and Pengfei Wang**. 2012. "Bubbles and Total Factor Productivity", *American Economic Review*, 102: 82-87.

[20] **Miao, Jianjun, and Pengfei Wang**. 2014. "Sectoral Bubbles, Misallocation and Endogenous Growth", *Journal of Mathematical Economics*, 53: 153-163.

[21] **Miao, Jianjun, and Pengfei Wang**. 2018. "Asset Bubbles and Credit Constraints,", *American Economic Review*, 108 (9), 2590-2628.

[22] **Miao, Jianjun, Wang, Pengfei, and Zhouxiang Shen**. 2019. "Monetary Policy and Rational Asset Bubbles: Comments,", *American Economic Review*, 109, 1969-1990.

[23] **Nisticó, Salvatore**. 2016. "Optimal Monetary Policy and Financial Stability in a Non-Ricardian Economy", *Journal of the European Economic Association*, 14 (5): 1225-1252.

[24] **Samuelson, Paul.** 1958. "An Exact Consumption-Loan Model of Interest with or without the Social Contrivance of Money", *Journal of Political Economy*, 66 (6): 467-482.

[25] **Santos, Manuel and Michael Woodford**. 1997. "Rational Asset Pricing Bubbles", *Econometrica*, 65 (1): 19- 58.

[26] **Tirole, Jean**. 1982. "On the Possibility of Speculation Under Rational Expectations," *Econometrica*, 50: 1163-1182.

[27] **Tirole, Jean**. 1985. "Asset Bubbles and Overlapping Generations", *Econometrica*, 53 (6): 1499- 1528.

[28] **Weil, Philippe**. 1989. "Overlapping Families of Infinitely-Lived Agents" *Journal of Public Economics*, 38: 183-198.

Appendix

A The Complete Model

The set of equilibrium conditions – in terms of productivity-adjusted variables – describing our baseline economy with stochastic asset-market participation, microfounded labor supply and optimal employment subsidies, whose BGP satisfies $c^r = c^p = c = y = N = \alpha\delta^{-\frac{1}{\varphi}}$, are as follows.

$$y_t = c_t \tag{76}$$

$$w_t = \delta\left(\frac{N_t}{\alpha}\right)^{\varphi} \tag{77}$$

$$y_t\Delta_t^p = N_t = y w_t^{1/\varphi} \tag{78}$$

$$d_t = y_t - \frac{y}{1+\mu} w_t^{\frac{1+\varphi}{\varphi}} \tag{79}$$

$$\alpha v_t \equiv \alpha v\left(\frac{N_t}{\alpha}\right) = \frac{y}{1+\varphi} w_t^{\frac{1+\varphi}{\varphi}} \tag{80}$$

$$c_t^r = y w_t^{\frac{1+\varphi}{\varphi}} + \frac{\tau^D}{1-\vartheta}(d_t - d) \tag{81}$$

$$c_t^p = (1-\beta\gamma)\left(\frac{q_t^B}{\vartheta} + x_t\right) + \frac{y}{1+\varphi} w_t^{\frac{1+\varphi}{\varphi}} \tag{82}$$

$$\tilde{c}_t^r = c_t^r - \alpha v\left(\frac{N_t}{\alpha}\right) \tag{83}$$

$$\tilde{c}_t^p = c_t^p - \alpha v\left(\frac{N_t}{\alpha}\right) \tag{84}$$

$$\tilde{c}_t = y_t - \alpha v\left(\frac{N_t}{\alpha}\right) \tag{85}$$

$$x_t = \gamma\nu\Gamma E_t\left\{\Lambda_{t,t+1}x_{t+1}\right\} + \tilde{c}_t^p \tag{86}$$

$$= \frac{\Phi}{\Lambda} E_t\left\{\Lambda_{t,t+1}x_{t+1}\right\} + \frac{1-\beta\gamma}{\vartheta\beta\gamma} q_t^B \tag{87}$$

$$y_t = (1-\beta\gamma)\left(q_t^B + \vartheta x_t\right) + \left[\frac{1+(1-\vartheta)\varphi}{1+\varphi}\right] y w_t^{\frac{1+\varphi}{\varphi}} + \tau^D(d_t - d) \tag{88}$$

$$q_t^B = \Gamma E_t\left\{\Lambda_{t,t+1}b_{t+1}\right\} \tag{89}$$

$$q_t^B = b_t + u_t \tag{90}$$

$$0 = E_t\left\{\sum_{k=0}^{\infty}(\theta\gamma\nu\Gamma)^k\left[\Lambda_{t,t+k}\frac{y_{t+k}}{\alpha}\left(\frac{P_t^*}{P_{t+k}}\right)^{-\epsilon}\left(\frac{P_t^*}{P_{t+k}} - (1+\mu)MC_{t+k}\right)\right]\right\}, \tag{91}$$

where $\Phi \equiv \frac{\nu\Gamma\Lambda}{\beta} \in (0,1]$, given the conditions for the existence of non-negative bubbles.

A.1 A First-Order Approximation Around the Efficient BGP

Take a first-order Taylor expansion of the above equilibrium conditions around a BGP in which the optimal employment subsidy completely offsets the monopolistic distortions, and denote with a "hat" the corresponding log-deviation, such that, for generic a variable Z, $\widehat{z}_t \equiv \log\left(\frac{Z_t}{Z_t^*}\right) = \log\left(\frac{Z_t}{z\Gamma^t}\right) = \log\left(\frac{z_t}{z}\right)$.[35] The approximated equilibrium conditions describing the model economy then read:

$$\widehat{y}_t = \widehat{c}_t \tag{92}$$

$$\widehat{w}_t = \varphi \widehat{N}_t \tag{93}$$

$$\widehat{y}_t = \widehat{N}_t = \frac{1}{\varphi}\widehat{w}_t \tag{94}$$

$$\widehat{d}_t = \frac{\mu-\varphi}{1+\mu}\widehat{y}_t \tag{95}$$

$$\alpha\widehat{v}_t = \frac{1}{\varphi}\widehat{w}_t = \widehat{y}_t \tag{96}$$

$$\widehat{c}_t^r = (1+\varphi-\tau)\,\widehat{y}_t \tag{97}$$

$$\widehat{c}_t^p = (1-\beta\gamma)\left(\frac{\widehat{q}_t^B}{\vartheta} + \widehat{x}_t\right) + \widehat{y}_t = \left[1 + \frac{(1-\vartheta)}{\vartheta}(\tau-\varphi)\right]\widehat{y}_t \tag{98}$$

$$\widehat{\widetilde{c}}_t^r = \frac{1+\varphi}{\varphi}\left(\widehat{c}_t^r - \alpha\widehat{v}_t\right) = \frac{1+\varphi}{\varphi}(\varphi-\tau)\,\widehat{y}_t \tag{99}$$

$$\widehat{\widetilde{c}}_t^p = \frac{1+\varphi}{\varphi}\left(\widehat{c}_t^p - \alpha\widehat{v}_t\right) = \frac{1+\varphi}{\varphi}\left(\frac{1-\vartheta}{\vartheta}\right)(\tau-\varphi)\,\widehat{y}_t. \tag{100}$$

$$\widehat{\widetilde{c}}_t = \frac{1+\varphi}{\varphi}\left(\widehat{c}_t - \alpha\widehat{v}_t\right) = 0 \tag{101}$$

$$\widehat{x}_t = \Phi E_t\left\{\widehat{x}_{t+1}\right\} - \frac{\varphi}{1+\varphi}\frac{\Phi}{1-\beta\gamma\Phi}\widehat{r}_t + \frac{1-\beta\gamma}{\vartheta\beta\gamma}\widehat{q}_t^B \tag{102}$$

$$\widehat{y}_t = \Theta\left(\frac{\widehat{q}_t^B}{\vartheta} + \widehat{x}_t\right) \tag{103}$$

$$\widehat{q}_t^B = \frac{\beta}{\nu}\Phi E_t\widehat{b}_{t+1} - q^B\widehat{r}_t \tag{104}$$

$$\widehat{q}_t^B = \widehat{b}_t + \widehat{u}_t \tag{105}$$

$$\widehat{\pi}_t = \beta\gamma\Phi E_t\widehat{\pi}_{t+1} + \kappa\widehat{y}_t, \tag{106}$$

where $\tau \equiv \left(\frac{\tau^D}{1-\vartheta}\right)\left(\frac{\varphi-\mu}{1+\mu}\right)$, $\Theta \equiv \frac{\vartheta(1-\beta\gamma)}{(1-\vartheta)(\tau-\varphi)}$ and $\kappa \equiv \varphi\frac{(1-\theta)(1-\gamma\nu\Gamma\Lambda\theta)}{\theta}$.

B Stability of the BGPs

To analyze the continuum of BGPs and characterize their stability properties, consider a perfect-foresight version of system (28)–(30), and define the ratios $\tilde{x}_t \equiv \eta\frac{x_t}{\tilde{c}_t^p}$, $\tilde{q}_t^B \equiv \eta\frac{q_t^B}{\tilde{c}_t^p}$, $\tilde{u}_t \equiv \eta\frac{u_t}{\tilde{c}_t^p}$ and

[35] The exceptions to this rule are: $\widehat{q}_t^B \equiv \frac{q_t^B}{y} - q^B$, $\widehat{b}_t \equiv \frac{b_t}{y} - b$, $\widehat{u}_t \equiv \frac{u_t}{y} - u$, $\widehat{x}_t \equiv \frac{x_t - x}{y}$, $\widehat{d}_t \equiv \frac{d_t - d}{y}$.

$\tilde{\Lambda}_{t,t+1} \equiv \frac{\Lambda_{t,t+1}\tilde{c}^p_{t+1}}{\tilde{c}^p_t}$, where $\eta \equiv \left[\frac{\varpi}{\vartheta} + (1-\varpi)\frac{\varphi}{1+\varphi}\right]$.[36] Accordingly, system (28)–(30) implies

$$\begin{aligned} \tilde{q}^B_t &= \eta\frac{\vartheta\beta\gamma}{1-\beta\gamma} + \gamma\nu\Gamma\tilde{\Lambda}_{t,t+1}\tilde{q}^B_{t+1} - \eta\frac{\vartheta\gamma\nu\Gamma}{1-\beta\gamma}\tilde{\Lambda}_{t,t+1} \\ &= \eta\frac{\vartheta\beta\gamma}{1-\beta\gamma} + \gamma\nu\frac{\tilde{q}^B_t\tilde{q}^B_{t+1}}{\tilde{q}^B_{t+1}-\tilde{u}_{t+1}} - \eta\frac{\vartheta\gamma\nu}{1-\beta\gamma}\frac{\tilde{q}^B_t}{\tilde{q}^B_{t+1}-\tilde{u}_{t+1}}, \end{aligned} \tag{107}$$

where the second line uses $\Gamma\tilde{\Lambda}_{t,t+1} = \frac{\tilde{q}^B_t}{\tilde{q}^B_{t+1}-\tilde{u}_{t+1}}$ as implied by equation (30).

Consistently with the analysis in Galí (2014) and Miao et al. (2019), consider a constant value $\tilde{u}$ for the ratio between new bubbles and adjusted consumption and use equation (107) to define the following mapping $f(\cdot)$ from current levels of the bubble-to-adjusted-consumption ratio $\tilde{q}^B_t$ to the next-period one $\tilde{q}^B_{t+1}$:

$$\tilde{q}^B_{t+1} = \frac{\tilde{q}^B_t\left[\eta\vartheta\gamma\nu - (1-\beta\gamma)\tilde{u}\right] + \eta\vartheta\beta\gamma\tilde{u}}{\eta\vartheta\beta\gamma - (1-\beta\gamma)(1-\gamma\nu)\tilde{q}^B_t} = f\left(\tilde{q}^B_t,\, \tilde{u},\, \eta\right). \tag{108}$$

The implied fixed point is therefore:

$$q^B = \frac{\eta\vartheta\gamma(\beta-\nu) + (1-\beta\gamma)\tilde{u} \pm \sqrt{\left[\eta\vartheta\gamma(\beta-\nu) + (1-\beta\gamma)\tilde{u}\right]^2 - 4\eta\vartheta\beta\gamma(1-\beta\gamma)(1-\gamma\nu)\tilde{u}}}{2(1-\beta\gamma)(1-\gamma\nu)}. \tag{109}$$

To highlight the implications of equation (108), notice that $f(\cdot)$ is twice continuously differentiable in $\tilde{q}^B_t$ for $0 \le \tilde{q}^B_t < \underline{q}^B \equiv \frac{\eta\vartheta\beta\gamma}{(1-\beta\gamma)(1-\gamma\nu)}$, and it also has the following properties:

$$f\left(0,\, 0,\, \eta\right) = 0 \tag{110}$$

$$f\left(0,\, \tilde{u},\, \eta\right) = \tilde{u} \tag{111}$$

$$f_1\left(\tilde{q}^B_t,\, \tilde{u},\, \eta\right) \equiv \frac{\partial f\left(\tilde{q}^B_t,\, \tilde{u},\, \eta\right)}{\partial \tilde{q}^B_t} = \frac{\eta\vartheta\gamma^2\nu\beta\left[\eta\vartheta - (1-\beta\gamma)\tilde{u}\right]}{\left[\eta\vartheta\beta\gamma - (1-\beta\gamma)(1-\gamma\nu)\tilde{q}^B_t\right]^2} > 0 \tag{112}$$

$$f_{11}\left(\tilde{q}^B_t,\, \tilde{u},\, \eta\right) \equiv \frac{\partial^2 f\left(\tilde{q}^B_t,\, \tilde{u},\, \eta\right)}{\partial \tilde{q}^B_t\partial \tilde{q}^B_t} = 2\frac{\eta\vartheta\gamma^2\nu\beta\left[\eta\vartheta - (1-\beta\gamma)\tilde{u}\right]}{\left[\eta\vartheta\beta\gamma - (1-\beta\gamma)(1-\gamma\nu)\tilde{q}^B_t\right]^3}(1-\beta\gamma)(1-\gamma\nu) > 0 \tag{113}$$

$$f_{12}\left(\tilde{q}^B_t,\, \tilde{u},\, \eta\right) \equiv \frac{\partial^2 f\left(\tilde{q}^B_t,\, \tilde{u},\, \eta\right)}{\partial \tilde{q}^B_t\partial \tilde{u}} = -\frac{\eta\vartheta\gamma^2\nu\beta(1-\beta\gamma)}{\left[\eta\vartheta\beta\gamma - (1-\beta\gamma)(1-\gamma\nu)\tilde{q}^B_t\right]^2} < 0 \tag{114}$$

in which $f_1\left(\tilde{q}^B_t,\, \tilde{u},\, \eta\right)$ and $f_{11}\left(\tilde{q}^B_t,\, \tilde{u},\, \eta\right)$ are positive under the restriction $\tilde{u} < \frac{\eta\vartheta}{1-\beta\gamma}$,[37] for $0 \le \tilde{q}^B_t < \underline{q}^B$ and $\lim_{\tilde{q}^B_t \to \underline{q}^B} f\left(\tilde{q}^B_t, \tilde{u}, \eta\right) = +\infty$.

The above properties imply that the mapping $f(\cdot)$, capturing the equilibrium dynamics of the aggregate bubble for given (constant) new bubbles, is strictly increasing and strictly convex. Figure 1 displays such mapping with the 45-degree line, for alternative values of $\tilde{u}$. The fixed points in $f(\cdot)$

[36] We normalize by $\frac{\tilde{c}^p_t}{\eta}$ rather than just $\tilde{c}^p_t$ because, along a BGP, $\tilde{c}^p = \eta y$, as implied by equation (37), and thus this normalization conveniently implies $\tilde{x} = x$, $\tilde{q}^B = q^B$, and $\tilde{u} = u$.

[37] Such restriction always holds in BGPs associated with non-negative aggregate bubbles.

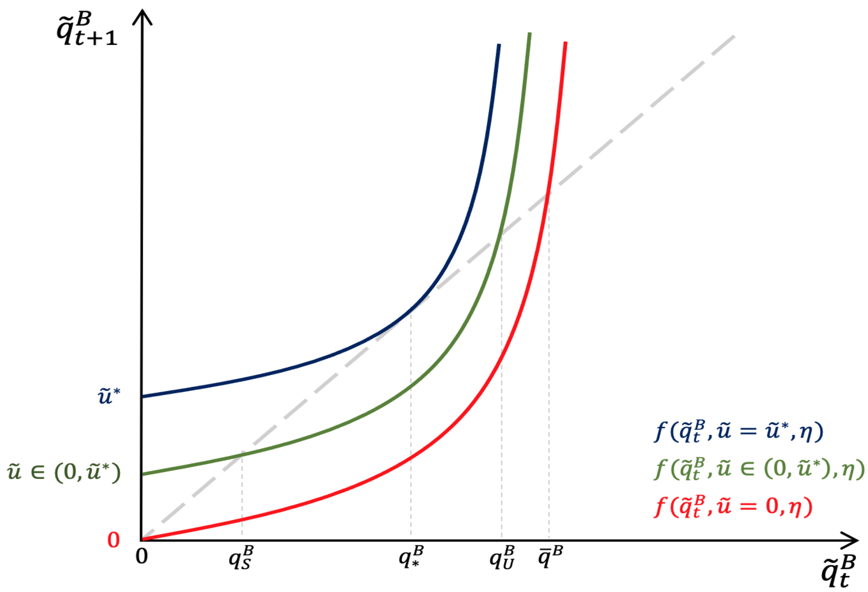

Figure 1: Equilibrium dynamics for the aggregate bubble under perfect foresight, for different values of the (constant) ratio of new bubbles to adjusted consumption of market participants, $\tilde{u}$. The dashed grey line is the 45-degree line.

then identify the BGPs associated with a non-negative aggregate bubble.

As the figure shows, there exists an upper bound on the aggregate bubble that is the larger of the solutions to equation (109) when $\tilde{u} = 0$:

$$\bar{q}^B = \frac{\eta\vartheta\gamma(\beta - \nu)}{(1 - \beta\gamma)(1 - \gamma\nu)} = \left[\varpi + (1 - \varpi)\frac{\vartheta\varphi}{1 + \varphi}\right] \frac{\gamma(\beta - \nu)}{(1 - \beta\gamma)(1 - \gamma\nu)}, \tag{115}$$

where the second equality uses the definition of η. Moreover, any BGP characterized by a bubble-to-output ratio $q^B_S \in [0,\, q^B_*)$ is globally stable because $f_1\left(q^B,\, \tilde{u},\, \eta\right) < 1$ (where $q^B = \tilde{q}^B$ along a BGP), while those characterized by a bubble-to-output ratio $q^B_U \in [q^B_*,\, \bar{q}^B]$ are globally unstable because $f_1\left(q^B,\, \tilde{u},\, \eta\right) > 1$. The threshold between stable and unstable BGPs, in turn, corresponds to the value of $\tilde{u}$, $\tilde{u}^*$, for which the two solutions to equation (109) coincide:

$$q^B_* = \frac{\eta\vartheta\gamma(\beta - \nu) + (1 - \beta\gamma)\tilde{u}^*}{2(1 - \beta\gamma)(1 - \gamma\nu)} = \frac{\eta\vartheta\gamma(\beta - \nu)}{(1 - \beta\gamma)(1 + R - 2\gamma\nu)}, \tag{116}$$

where the second equality uses equation (40) and the fact that, along a BGP, $\tilde{u} = u$.

Figure 2 displays the role of the three additional factors affecting the nature of the bubbly BGPs, discussed in Section 3.1. Notice from equations (108)–(114) that these three additional margins – the stochastic asset-market participation, the endogenous labor supply, and the employment subsidy offsetting monopolistic distortion – are captured by either η or ϑ, and affect all the relevant properties only jointly, through the term $\eta\vartheta = \left[\varpi + (1 - \varpi)\frac{\vartheta\varphi}{1+\varphi}\right]$, which is increasing in all ϑ, φ, and ϖ. Contrasting the solid and dashed lines in Figure 2 then shows how a lower share of market participants, a lower concavity of the utility of leisure or a lower amount of monopolistic distortions

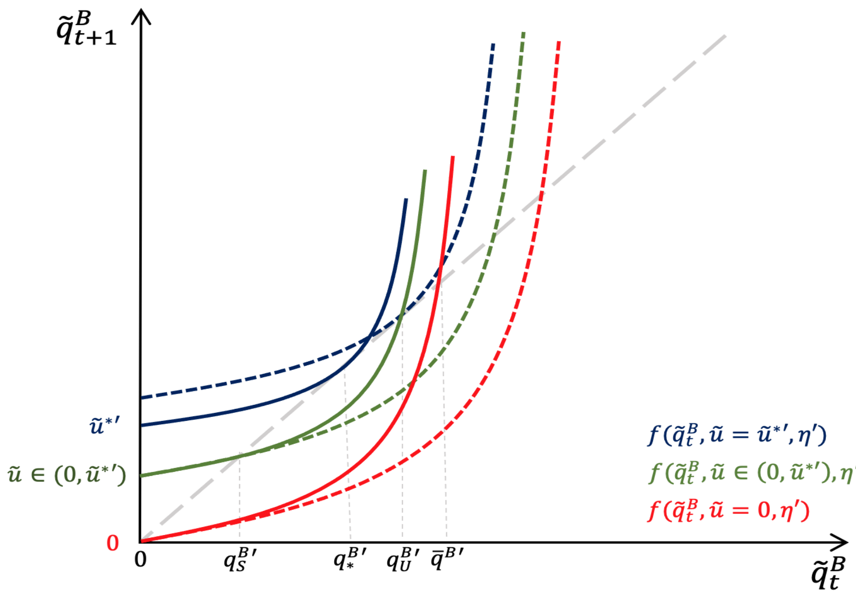

Figure 2: Equilibrium dynamics for the aggregate bubble under perfect foresight: the role of stochastic asset-market participation, the endogenous labor supply and the employment subsidy. The dashed grey line is the 45-degree line.

are all associated with a smaller aggregate bubble.

Notice, however, that while the size of the equilibrium aggregate bubble is affected by these three additional margins, the relevant interval and thresholds for the relative real interest rate are not. Indeed, the stable BGPs are associated with an equilibrium real interest rate (relative to the growth rate of the economy) $R \in [\nu/\beta,\, R^*]$, while the unstable ones are associated with $R \in (R^*,\, 1]$.

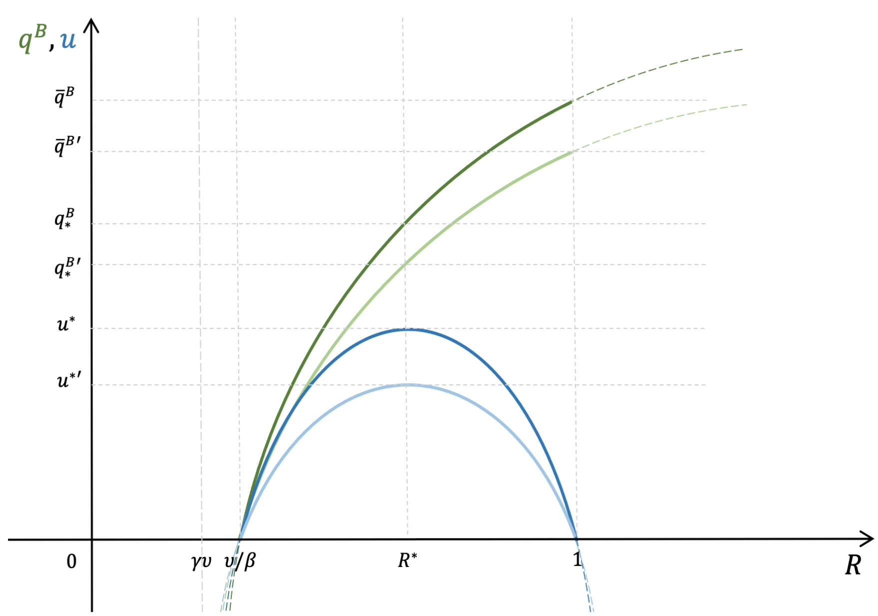

Figure 3: Equilibrium size of aggregate (q^B) and new (u) bubbles along the BGP, as a function of the relative real interest rate R. Solid lines are the relevant part of the mapping, corresponding to non-negative q^B and u. Lighter lines correspond to lower values of η.

Moreover, equations (39) and (116) jointly imply that the threshold level of the relative interest rate R^* that separates stable and unstable BGPs solves

$$\beta R^2 - 2\beta\gamma\nu R - \nu(1 - \beta\gamma - \nu\gamma) = 0 \tag{117}$$

and is therefore independent of both η and ϑ, as also shown by equation (47). This implication is further shown in Figure 3, which displays equations (40) and (46) as functions of R.

C The Welfare-Based Monetary-Policy Loss Function

We evaluate alternative policies using a second-order approximation of social welfare around the efficient BGP where, for a generic variable X, we use the notation $X_t^* \equiv x\Gamma^t$. To derive the latter, consider the system of Pareto-weights $\{\chi_s^j\}$, with $j \in \mathcal{T} = \{pe,\, pu,\, re,\, ru\}$ indexing the agent type, and $s = -\infty$, ..., $t-1$, t the generic time of transition in that type and such that

$$\sum_{j\in\mathcal{T}} \sum_{s=-\infty}^{t} \chi_s^j = 1,$$

and the following Ramsey problem:

$$\max \qquad E_{t_0}\left\{\sum_{t=t_0}^{\infty} \beta^{t-t_0} U_t\right\}$$

with

$$U_t \equiv \sum_{j\in\mathcal{T}} \sum_{s=-\infty}^{t} \chi_s^j U_{t|s}^j \tag{118}$$

subject to the aggregate production function and the resource constraint:

$$\Gamma^t \left[\sum_{s=-\infty}^{t} m_{t|s}^{pe} N_{t|s}^{pe} + \sum_{s=-\infty}^{t} m_{t|s}^{re} N_{t|s}^{re}\right] = Y_t = C_t = \sum_{j\in\mathcal{T}} \sum_{s=-\infty}^{t} m_{t|s}^{j} C_{t|s}^{j}, \tag{119}$$

where $m_{t|s}^j$ denotes the relative mass of agents transited into type j at time $s \leq t$, with

$$\sum_{j\in\mathcal{T}} \sum_{s=-\infty}^{t} m_{t|s}^j = 1.$$

An efficient BGP satisfies the following first-order conditions for the Ramsey allocation:

$$\chi_s^{pe} U_{C^*,t|s}^{pe} = \lambda_t^* m_{t|s}^{pe} = \lambda_t^* \vartheta(1-\gamma)(\gamma\nu)^{t-s} \tag{120}$$
$$\chi_s^{re} U_{C^*,t|s}^{re} = \lambda_t^* m_{t|s}^{re} = \lambda_t^* (1-\vartheta)(1-\varrho)\alpha\varrho^{t-s} \tag{121}$$
$$\chi_s^{pu} U_{C^*,t|s}^{pu} = \lambda_t^* m_{t|s}^{pu} = \lambda_t^* \vartheta(1-\gamma)\gamma^{t-s}(1-\nu^{t-s}) \tag{122}$$
$$\chi_s^{ru} U_{C^*,t|s}^{ru} = \lambda_t^* m_{t|s}^{ru} = \lambda_t^* (1-\vartheta)(1-\varrho)(1-\alpha)\varrho^{t-s} \tag{123}$$
$$\chi_s^{pe} U_{N^*,t|s}^{pe} = -\lambda_t^* \Gamma^t m_{t|s}^{pe} = -\lambda_t^* \Gamma^t \vartheta(1-\gamma)(\gamma\nu)^{t-s} \tag{124}$$
$$\chi_s^{re} U_{N^*,t|s}^{re} = -\lambda_t^* \Gamma^t m_{t|s}^{re} = -\lambda_t^* \Gamma^t (1-\vartheta)(1-\varrho)\alpha\varrho^{t-s} \tag{125}$$

for each $s = -\infty$, ..., $t-1$, t, where λ_t^* is the BGP-level of the Lagrange multiplier associated to the constraint (119). Dividing (124) by (120) and (125) by (121) verifies that the intratemporal efficiency condition holds:

$$MRS_{t|s} = -\frac{U_{N^*,t|s}^{pe}}{U_{C^*,t|s}^{pe}} = -\frac{U_{N^*,t|s}^{re}}{U_{C^*,t|s}^{re}} = \Gamma^t = MPN_t. \tag{126}$$

Moreover, note that since hours worked are constant along a BGP, the equations (124) and (125) imply

$$\lambda_t^* = \frac{\overline{\lambda}}{\Gamma^t},$$

for some $\overline{\lambda} > 0$. Therefore, (120)–(125), in an efficient BGP, jointly imply

$$\chi_s^j U_{C^*,t|s}^j \Gamma^t = \overline{\lambda} m_{t|s}^j \tag{127}$$

for any s, j and all t. Recall that preferences are of the type

$$U_{t|s}^j = \log\left(C_{t|s}^j - V(N_{t|s}^j)\right) = \log \widetilde{C}_{t|s}^j$$

with $\widetilde{C}_{t|s}^j \equiv C_{t|s}^j - V(N_{t|s}^j)$ denoting *adjusted* consumption and $V(N_{t|s}^j) \equiv \frac{\delta\Gamma^t}{1+\varphi}(N_{t|s}^j)^{1+\varphi}$ the disutility of labor. As a consequence, the marginal utilities of both consumption and *adjusted* consumption are the same (both on and off the BGP):

$$U_{C^*,t|s}^j = U_{\widetilde{C}^*,t|s}^j = \frac{1}{\widetilde{C}_{t|s}^{j,*}}. \tag{128}$$

Now, consider a second-order Taylor expansion of the period utility (118) around the efficient BGP, disregarding terms of higher order or independent of policy:

$$U_t = U_t^* + \sum_{j\in\mathcal{T}} \sum_{s=-\infty}^{t} \chi_s^j \left[U_{\widetilde{C}^*,t|s}^j \left(\widetilde{C}_{t|s}^j - \widetilde{C}_{t|s}^{j,*}\right) + \frac{1}{2} U_{\widetilde{C}^*\widetilde{C}^*,t|s}^j \left(\widetilde{C}_{t|s}^j - \widetilde{C}_{t|s}^{j,*}\right)^2 \right]. \tag{129}$$

The economy therefore converges to an efficient BGP if we impose two restrictions: i) a set of Pareto weights satisfying condition (127) and ii) an appropriate employment subsidy $\tau^F = \frac{\mu}{1+\mu}$, implementing condition (126). In particular, using the appropriate Pareto weights in (129) implies:

$$
\begin{aligned}
U_t - U_t^* &= \frac{\overline{\lambda}}{\Gamma^t} \sum_{j \in \mathcal{T}} \sum_{s=-\infty}^{t} m_{t|s}^j \left[\left(\widetilde{C}_{t|s}^j - \widetilde{C}_{t|s}^{j,*} \right) + \frac{1}{2} \frac{U^j_{\widetilde{C}^* \widetilde{C}^*, t|s}}{U^j_{\widetilde{C}^*, t|s}} \left(\widetilde{C}_{t|s}^j - \widetilde{C}_{t|s}^{j,*} \right)^2 \right] \\
&= \overline{\lambda} \sum_{j \in \mathcal{T}} \sum_{s=-\infty}^{t} m_{t|s}^j \left[\left(\widetilde{c}_{t|s}^j - \widetilde{c}_s^j \right) - \frac{1}{2} \frac{1}{\widetilde{c}_s^j} \left(\widetilde{c}_{t|s}^j - \widetilde{c}_s^j \right)^2 \right] \\
&= \overline{\lambda} E_{sj} \left[\left(\widetilde{c}_{t|s}^j - \widetilde{c}_s^j \right) - \frac{1}{2} \frac{1}{\widetilde{c}_s^j} \left(\widetilde{c}_{t|s}^j - \widetilde{c}_s^j \right)^2 \right],
\end{aligned} \tag{130}
$$

where the second line uses (128) and the normalizations for aggregate productivity, $\widetilde{c}_{t|s}^j \equiv \frac{\widetilde{C}_{t|s}^j}{\Gamma^t}$ and $\widetilde{c}_s^j \equiv \frac{\widetilde{C}_{t|s}^{j,*}}{\Gamma^t}$, while the last line uses the definition of mass-weighted cross-sectional mean across all agents in the economy, regardless of the type and the longevity in the type:

$$
E_{sj} x_{t|s}^j \equiv \sum_{j \in \mathcal{T}} \sum_{s=-\infty}^{t} m_{t|s}^j x_{t|s}^j
$$

for any generic variable x.

Focusing on the first-order term in (130), and considering $N_{t|s}^{pe} = N_{t|s}^{re} = N_t/\alpha$ and $N_{t|s}^{pu} = N_{t|s}^{ru} = 0$ for all s, we can write

$$
\begin{aligned}
E_{sj} \left(\widetilde{c}_{t|s}^j - \widetilde{c}_s^j \right) &= c_t - \frac{\alpha\delta}{1+\varphi} \left(\frac{N_t}{\alpha} \right)^{1+\varphi} - \left[c - \frac{\alpha\delta}{1+\varphi} \left(\frac{N}{\alpha} \right)^{1+\varphi} \right] \\
&= y_t - \frac{\alpha\delta}{1+\varphi} \left(\frac{y_t \Delta_{p,t}}{\alpha} \right)^{1+\varphi} - \left[y - \frac{\alpha\delta}{1+\varphi} \left(\frac{y}{\alpha} \right)^{1+\varphi} \right],
\end{aligned} \tag{131}
$$

where $c_t \equiv C_t/\Gamma^t$ and the second line uses the aggregate resource constraint and aggregate production function, $N_t = y_t \Delta_{p,t}$, with $y_t \equiv Y_t/\Gamma^t$.

Now, let $\widehat{y}_t \equiv \log \left(\frac{Y_t}{y\Gamma^t} \right) = \log \left(\frac{y_t}{y} \right)$ and $\widehat{\Delta}_{p,t} \equiv \log \Delta_{p,t}$ and consider that, in a second-order approximation

$$
\begin{aligned}
y_t &= y \left(1 + \widehat{y}_t + \frac{1}{2} \widehat{y}_t^2 \right) \\
N_t &= y_t \Delta_t^p = y \left(1 + \widehat{y}_t + \frac{1}{2} \widehat{y}_t^2 + \widehat{\Delta}_{p,t} \right).
\end{aligned}
$$

We can use the above equations, together with the expression for equilibrium output in the efficient BGP, $y = \alpha\delta^{-1/\varphi}$, to evaluate (131) as:

$$
E_{sj} \left(\widetilde{c}_{t|s}^j - \widetilde{c}_s^j \right) = -\frac{y}{2} \left(\varphi \widehat{y}_t^2 + 2\widehat{\Delta}_{p,t} \right). \tag{132}
$$

Now, focus on the second-order term in (130)

$$E_{sj}\left[\frac{1}{\widetilde{c}_s^j}\left(\widetilde{c}_{t|s}^j-\widetilde{c}_s^j\right)^2\right]\equiv\sum_{j\in\mathcal{T}}\sum_{s=-\infty}^{t}m_{t|s}^j\left[\frac{1}{\widetilde{c}_s^j}\left(\widetilde{c}_{t|s}^j-\widetilde{c}_s^j\right)^2\right]$$
$$=\sum_{j\in\mathcal{T}}\sum_{s=-\infty}^{t}m_{t|s}^j\widetilde{c}_s^j\left(\widehat{\widetilde{c}}_{t|s}^j\right)^2, \tag{133}$$

where the second line uses the first-order approximation

$$\widehat{\widetilde{c}}_{t|s}^j\equiv\log\left(\frac{\widetilde{C}_{t|s}^j}{\widetilde{c}_s^j\Gamma^t}\right)=\log\left(\frac{\widetilde{c}_{t|s}^j}{\widetilde{c}_s^j}\right)=\frac{\widetilde{c}_{t|s}^j-\widetilde{c}_s^j}{\widetilde{c}_s^j}.$$

Note that, along the efficient BGP, the cross-sectional mean of the *adjusted* consumption is proportional to aggregate output

$$E_{sj}\widetilde{c}_s^j=\sum_{j\in\mathcal{T}}\sum_{s=-\infty}^{t}m_{t|s}^j\widetilde{c}_s^j=\frac{\varphi}{1+\varphi}y,$$

which implies that we can define the following cross-sectional mean operator, for a given variable x:

$$\widetilde{E}_{sj}x_s^j\equiv\frac{1+\varphi}{\varphi}\sum_{j\in\mathcal{T}}\sum_{s=-\infty}^{t}m_{t|s}^j\frac{\widetilde{c}_s^j}{y}x_s^j.$$

Using the last two expressions in (133), we can write

$$E_{sj}\left[\frac{1}{\widetilde{c}_s^j}\left(\widetilde{c}_{t|s}^j-\widetilde{c}_s^j\right)^2\right]=\frac{\varphi y}{1+\varphi}\widetilde{E}_{sj}\left[\left(\widehat{\widetilde{c}}_{t|s}^j\right)^2\right]$$
$$=\frac{\varphi y}{1+\varphi}\left[\left(\widetilde{E}_{sj}\widehat{\widetilde{c}}_{t|s}^j\right)^2+\widetilde{var}_{sj}\widehat{\widetilde{c}}_{t|s}^j\right]$$
$$=\frac{\varphi y}{1+\varphi}\widetilde{var}_{sj}\widehat{\widetilde{c}}_{t|s}^j, \tag{134}$$

where the second line uses $E(x^2)=[E(x)]^2+var(x)$ and the third line uses

$$\widetilde{E}_{sj}\widehat{\widetilde{c}}_{t|s}^j=\frac{1+\varphi}{\varphi}\sum_{j\in\mathcal{T}}\sum_{s=-\infty}^{t}m_{t|s}^j\frac{\widetilde{c}_s^j}{y}\widehat{\widetilde{c}}_{t|s}^j$$
$$=\frac{1+\varphi}{\varphi}\left[\sum_{j\in\mathcal{T}}\sum_{s=-\infty}^{t}m_{t|s}^j\frac{c_s^j}{y}\widehat{c}_{t|s}^j-\sum_{s=-\infty}^{t}\left(m_{t|s}^{pe}+m_{t|s}^{re}\right)\frac{1}{\alpha}\widehat{N}_t\right]$$
$$=\frac{1+\varphi}{\varphi}\left(\widehat{c}_t-\widehat{N}_t\right)=\frac{1+\varphi}{\varphi}\left(\widehat{y}_t-\widehat{y}_t\right)=0,$$

where the second line uses the first-order approximation of *adjusted* consumption for employed agents ($\widetilde{c}_s^j\widehat{\widetilde{c}}_{t|s}^j=c_s^j\widehat{c}_{t|s}^j-\frac{y}{\alpha}\widehat{N}_t$, for $j=pe$, re) and for unemployed ones ($\widetilde{c}_s^j\widehat{\widetilde{c}}_{t|s}^j=c_s^j\widehat{c}_{t|s}^j$, for $j=pu$, ru), and the last line uses a first-order approximation of the resource constraint ($\widehat{c}_t=\widehat{y}_t$) and the aggregate production function ($\widehat{N}_t = \widehat{y}_t$).

Substituting (134) and (132) in (130) yields

$$-\frac{U_t - U_t^*}{\overline{\lambda} y} = \widehat{\Delta}_{p,t} + \frac{\varphi}{2}\widehat{y}_t^2 + \frac{1}{2}\frac{\varphi}{1+\varphi}\widetilde{var}_{sj}\widehat{\widetilde{c}}_{t|s}^{j}, \tag{135}$$

which emphasizes that the social welfare loss does not only depend on relative-price dispersion and output-gap volatility, as in the benchmark New Keynesian model, but it is also increasing in the cross-sectional consumption dispersion, reflecting the several layers of households' heterogeneity characterizing the economy.

So let us focus on this latter term. We can first decompose it into *between* and *within* groups, using the law of total variance, to get

$$\widetilde{var}_{sj}\widehat{\widetilde{c}}_{t|s}^{j} = \widetilde{E}_j\left(\widetilde{var}_s\widehat{\widetilde{c}}_{t|s}^{j}\right) + \widetilde{var}_j\left(\widetilde{E}_s\widehat{\widetilde{c}}_{t|s}^{j}\right), \tag{136}$$

where j indexes the groups of agents, and s the longevity in each group. Moreover, note that the assumption of complete markets for financially active agents and the redistribution scheme among financially inactive ones imply that, *within the two agent types*, the adjusted consumption of any two agents *with the same longevity in the type* is the same, regardless of their employment status: $\widehat{\widetilde{c}}_{t|s}^{pe} = \widehat{\widetilde{c}}_{t|s}^{pu} = \widehat{\widetilde{c}}_{t|s}^{p}$ and $\widehat{\widetilde{c}}_{t|s}^{re} = \widehat{\widetilde{c}}_{t|s}^{ru} = \widehat{\widetilde{c}}_{t|s}^{r}$. Therefore, the first relevant partition to consider to evaluate the law of total variance, is the one between market participants and rule-of-thumbers, with relative mass equal to ϑ and $1-\vartheta$, respectively. Accordingly, we can write the first term in (136) as

$$\widetilde{E}_j\left(\widetilde{var}_s\widehat{\widetilde{c}}_{t|s}^{j}\right) = \vartheta\widetilde{var}_s\widehat{\widetilde{c}}_{t|s}^{p} + (1-\vartheta)\widetilde{var}_s\widehat{\widetilde{c}}_{t|s}^{r} = \vartheta\widetilde{var}_s\widehat{\widetilde{c}}_{t|s}^{p}, \tag{137}$$

where the second equality reflects the homogeneity within the set of rule-of-thumbers, implying $\widehat{\widetilde{c}}_{t|s}^{r} = \widehat{\widetilde{c}}_{t}^{r}$ for all s, and therefore $\widetilde{var}_s\widehat{\widetilde{c}}_{t|s}^{r} = 0$.

As to the second term in (136), we can use $var(x) = E(x^2) - [E(x)]^2$ to write it as

$$\begin{aligned}\widetilde{var}_j\left(\widetilde{E}_s\widehat{\widetilde{c}}_{t|s}^{j}\right) &= \vartheta\left(\widetilde{E}_s\widehat{\widetilde{c}}_{t|s}^{p}\right)^2 + (1-\vartheta)\left(\widetilde{E}_s\widehat{\widetilde{c}}_{t|s}^{r}\right)^2 - \left(\widetilde{E}_{sj}\widehat{\widetilde{c}}_{t|s}^{j}\right)^2 \\ &= \vartheta\left(\widehat{\widetilde{c}}_{t}^{p}\right)^2 + (1-\vartheta)\left(\widehat{\widetilde{c}}_{t}^{r}\right)^2 \qquad (138) \\ &= \left(\frac{1+\varphi}{\varphi}\right)^2\left(\frac{1-\vartheta}{\vartheta}\right)(\varphi-\tau)^2\,\widehat{y}_t^2, \qquad (139)\end{aligned}$$

where the second line uses $\widetilde{E}_{sj}\widehat{\widetilde{c}}_{t|s}^{j} = 0$, derived above, and the definition of the *within*-group cross-sectional means

$$\widetilde{E}_s \widehat{\widetilde{c}}^p_{t|s} \equiv \frac{1+\varphi}{\varphi} \sum_{s=-\infty}^{t} \frac{m^p_{t|s}}{\vartheta} \frac{\widetilde{c}^p_s}{y} \widehat{\widetilde{c}}^p_{t|s} = \widehat{\widetilde{c}}^p_t \tag{140}$$

$$\widetilde{E}_s \widehat{\widetilde{c}}^r_{t|s} \equiv \frac{1+\varphi}{\varphi} \sum_{s=-\infty}^{t} \frac{m^r_{t|s}}{1-\vartheta} \frac{\widetilde{c}^r_s}{y} \widehat{\widetilde{c}}^r_{t|s} = \widehat{\widetilde{c}}^r_t \,, \tag{141}$$

where $m^p_{t|s} = m^{pe}_{t|s} + m^{pu}_{t|s}$ and $m^r_{t|s} = m^{re}_{t|s} + m^{ru}_{t|s}$ for all s, $\widetilde{c}^p_s = \widetilde{c}^{pe}_s = \widetilde{c}^{pu}_s$, and $\widetilde{c}^r_s = \widetilde{c}^{re}_s = \widetilde{c}^{ru}_s$, while the third line uses (99)–(100).

Using (136), (137) and (139), we can further simplify (135) into

$$-\frac{U_t - U^*_t}{\bar{\lambda} y} = \widehat{\Delta}_{p,t} + \frac{\varphi}{2}\left[1 + (1+\varphi)\left(\frac{1-\vartheta}{\vartheta}\right)\left(1 - \frac{\tau}{\varphi}\right)^2\right]\widehat{y}^2_t + \frac{1}{2}\frac{\varphi}{1+\varphi}\vartheta \widehat{\Delta}^p_{c,t}, \tag{142}$$

which emphasizes that the heterogeneity *between* agent types is proportional to the squared output gap, while the heterogeneity *within* agent types – and in particular within market participants, captured by the cross-sectional consumption dispersion $\widehat{\Delta}^p_{c,t} \equiv \widetilde{var}_s \widehat{\widetilde{c}}^p_{t|s}$ – is instead a source of additional and independent welfare loss.

To dig deeper into the meaning of this last term, consider the partition of the set of market participants between new-coming agents in the type – of mass $(1-\gamma)$ – and incumbent agents – of mass γ – to decompose (140) into the cross-sectional average *between* these two subsets:

$$\widetilde{E}_s \widehat{\widetilde{c}}^p_{t|s} = \widetilde{E}_s \left(\widehat{\widetilde{c}}^p_{t|s} \middle| s \leq t \right) = (1-\gamma) \widetilde{E}_{s=t} \widehat{\widetilde{c}}^p_{t|s} + \gamma \widetilde{E}_{s<t} \widehat{\widetilde{c}}^p_{t|s}, \tag{143}$$

and *within* each of them:

$$\widetilde{E}_{s=t} \widehat{\widetilde{c}}^p_{t|s} \equiv \widetilde{E}_s \left(\widehat{\widetilde{c}}^p_{t|s} \middle| s = t \right) = \frac{1+\varphi}{\varphi} \widehat{\widetilde{c}}^p_{t|nc} \tag{144}$$

$$\widetilde{E}_{s<t} \widehat{\widetilde{c}}^p_{t|s} \equiv \widetilde{E}_s \left(\widehat{\widetilde{c}}^p_{t|s} \middle| s \leq t-1 \right) = \frac{1+\varphi}{\varphi} \widehat{\widetilde{c}}^p_{t|in} \tag{145}$$

where $\widehat{\widetilde{c}}^p_{t|nc}$ denotes the average adjusted consumption of *newcomers* (nc) in deviation from the BGP as a ratio to aggregate output

$$\widehat{\widetilde{c}}^p_{t|nc} \equiv \frac{\widetilde{c}^p_t}{y} \widehat{\widetilde{c}}^p_{t|t} = \frac{\widetilde{c}^p_{t|s=t} - \widetilde{c}^p_{s=t}}{y}$$

and $\widehat{\widetilde{c}}^p_{t|in}$ denotes the average adjusted consumption of *incumbent* agents (in) in deviation from the BGP as a ratio to aggregate output

$$\widehat{\widetilde{c}}^p_{t|in} \equiv \sum_{s=-\infty}^{t-1} \frac{m^p_{t|s}}{\vartheta\gamma} \frac{\widetilde{c}^p_s}{y} \widehat{\widetilde{c}}^p_{t|s} = \sum_{s=-\infty}^{t-1} \frac{m^p_{t|s}}{\vartheta\gamma} \left(\frac{\widetilde{c}^p_{t|s} - \widetilde{c}^p_s}{y} \right) = \frac{\widetilde{c}^p_{t|in} - \widetilde{c}^p_{in}}{y}.$$

The definitions above can be used to decompose $\widehat{\Delta}^p_{c,t}$ by means of the law of total variance:

$$\begin{aligned}\widehat{\Delta}^p_{c,t} &\equiv \widetilde{var}_s \widehat{\widetilde{c}}^p_{t|s} = \widetilde{var}_s \left(\widehat{\widetilde{c}}^p_{t|s}\middle| s \leq t\right) \\ &= \gamma \widetilde{var}_{s<t}\widehat{\widetilde{c}}^p_{t|s} + (1-\gamma)\left(\widetilde{E}_{s=t}\widehat{\widetilde{c}}^p_{t|s}\right)^2 + \gamma\left(\widetilde{E}_{s<t}\widehat{\widetilde{c}}^p_{t|s}\right)^2 - \left(\widetilde{E}_s\widehat{\widetilde{c}}^p_{t|s}\right)^2 \\ &= \gamma\widehat{\Delta}^p_{c,t-1} + \left[(1-\gamma)\left(\widetilde{E}_{s=t}\widehat{\widetilde{c}}^p_{t|s}\right)^2 + \gamma\left(\widetilde{E}_{s<t}\widehat{\widetilde{c}}^p_{t|s}\right)^2 - \left(\widetilde{E}_s\widehat{\widetilde{c}}^p_{t|s}\right)^2\right],\end{aligned} \tag{146}$$

where in the second line we use the homogeneity of newcomers within their subset, implying $\widetilde{var}_{s=t}\widehat{\widetilde{c}}^p_{t|s} = 0$, and in the third line a first-order approximation of the Euler equation of market participants, implying

$$\begin{aligned}\widetilde{var}_{s<t}\widehat{\widetilde{c}}^p_{t|s} &= \widetilde{var}_s \left(\widehat{\widetilde{c}}^p_{t|s}\middle| s \leq t-1\right) \\ &= \widetilde{var}_s \left(\widehat{\widetilde{c}}^p_{t-1|s} - \widehat{\Lambda}_{t-1,t}\middle| s \leq t-1\right) \\ &= \widetilde{var}_s \left(\widehat{\widetilde{c}}^p_{t-1|s}\middle| s \leq t-1\right) = \widehat{\Delta}^p_{c,t-1}.\end{aligned}$$

To evaluate the term in squared brackets in (146), note that

$$\begin{aligned}\widetilde{E}_{s=t}\widehat{\widetilde{c}}^p_{t|s} &= \frac{1+\varphi}{\varphi}\widehat{\widetilde{c}}^p_{t|nc} = \frac{(1+\varphi)(1-\beta\gamma)}{\varphi}\left[\frac{\widehat{u}_t}{\vartheta(1-\gamma)} + \frac{\widehat{x}_t}{\alpha}\right] \\ \widetilde{E}_{s<t}\widehat{\widetilde{c}}^p_{t|s} &= \frac{1+\varphi}{\varphi}\widehat{\widetilde{c}}^p_{t|in} = \frac{(1+\varphi)(1-\beta\gamma)}{\varphi}\left(\frac{\widehat{b}_t}{\vartheta\gamma} + \nu\widehat{x}_t\right) \\ \widetilde{E}_s\widehat{\widetilde{c}}^p_{t|s} &= \frac{(1+\varphi)(1-\beta\gamma)}{\varphi}\left(\frac{\widehat{b}_t + \widehat{u}_t}{\vartheta} + \widehat{x}_t\right),\end{aligned}$$

where $\widehat{b}_t \equiv \frac{b_t}{y} - b$, $\widehat{u}_t \equiv \frac{u_t}{y} - u$ and $\widehat{x}_t \equiv \frac{x_t - x}{y}$. , Substituting the last three equations into (146), after some algebra, yields

$$\widehat{\Delta}^p_{c,t} = \gamma\widehat{\Delta}^p_{c,t-1} + \frac{1-\gamma}{\gamma}\left[\frac{(1+\varphi)(1-\beta\gamma)}{\varphi}\right]^2 \widehat{\omega}_t^2, \tag{147}$$

which is the law of motion of the cross-sectional consumption dispersion among market participants, and where we defined

$$\begin{aligned}\widehat{\omega}_t &\equiv \frac{1}{\vartheta}\widehat{q}^B_t - \frac{\widehat{u}_t}{\vartheta(1-\gamma)} - \frac{1-\alpha}{\alpha}\widehat{x}_t \\ &= \gamma\left[\frac{\widehat{b}_t}{\vartheta\gamma} - \frac{\widehat{u}_t}{\vartheta(1-\gamma)} - \frac{1-\nu}{1-\gamma}\widehat{x}_t\right] \\ &= \frac{\gamma}{1-\beta\gamma}\left(\widehat{\widetilde{c}}^p_{t|in} - \widehat{\widetilde{c}}^p_{t|nc}\right).\end{aligned}$$

Moving from an arbitrary initial level $\widehat{\Delta}^c_{t_0-1}$, which is independent of policies implemented from

$t = t_0$ onward, we can write the consumption dispersion among participants at time t as

$$\widehat{\Delta}_{c,t}^{p} = \gamma^{t-t_0+1}\widehat{\Delta}_{t_0-1}^{c} + \frac{1-\gamma}{\gamma}\left[\frac{(1+\varphi)(1-\beta\gamma)}{\varphi}\right]^2 \sum_{T=t_0}^{t} \gamma^{t-T}\widehat{\omega}_T^2 \tag{148}$$

and the discounted value over all periods $t > t_0$ (ignoring terms independent of policy) as

$$\sum_{t=t_0}^{\infty} \beta^{t-t_0}\widehat{\Delta}_{c,t}^{p} = \frac{(1-\gamma)(1-\beta\gamma)}{\gamma}\left(\frac{1+\varphi}{\varphi}\right)^2 \sum_{t=t_0}^{\infty} \beta^{t-t_0}\widehat{\omega}_t^2. \tag{149}$$

Finally, taking the time$-t_0$ conditional expectation of the discounted stream of future period social losses yields the welfare-based loss function $\mathcal{L}_{t_0}$, expressed as a share of steady-state aggregate output. Ignoring the terms independent of policy and those of third or higher order, we can write it as

$$\mathcal{L}_{t_0} \equiv -E_{t_0}\left\{\sum_{t=t_0}^{\infty} \beta^{t-t_0}\left(\frac{U_t - U_t^*}{\bar{\lambda} y}\right)\right\} = \frac{1}{2}\frac{\varepsilon\varphi}{\kappa}E_{t_0}\left\{\sum_{t=t_0}^{\infty} \beta^{t-t_0}\left(\widehat{\pi}_t^2 + \alpha_y \widehat{y}_t^2 + \alpha_\omega \widehat{\omega}_t^2\right)\right\}, \tag{150}$$

where we use (142), (149), and

$$\widehat{\Delta}_{p,t} \approx \frac{\varepsilon}{2} var_i p_t(i)$$

$$E_{t_0}\sum_{t=t_0}^{\infty} \beta^{t-t_0} var_i p_t(i) = \frac{\theta}{(1-\theta)\,(1-\gamma\nu\Gamma\Lambda\theta)} \sum_{t=t_0}^{\infty} \beta^{t-t_0}\widehat{\pi}_t^2,$$

and the relative welfare weights are defined as

$$\alpha_y \equiv \frac{\kappa}{\varphi\varepsilon}\left[\varphi + \left(\frac{1+\varphi}{\varphi}\right)\left(\frac{1-\vartheta}{\vartheta}\right)(\tau-\varphi)^2\right]$$

$$\alpha_\omega \equiv \frac{\kappa\vartheta}{\varepsilon\varphi}\left(\frac{1+\varphi}{\varphi}\right)\frac{(1-\gamma)(1-\beta\gamma)}{\gamma}$$

with

$$\kappa \equiv \varphi\frac{(1-\theta)(1-\gamma\nu\Gamma\Lambda\theta)}{\theta}$$

$$\tau \equiv \left(\frac{\tau^D}{1-\vartheta}\right)\left(\frac{\varphi-\mu}{1+\mu}\right).$$