Bubbly Firm Dynamics and Aggregate Fluctuations

Haozhou Tang* Donghai Zhang†
Bank of Mexico *University of Bonn*

October 11, 2021

Abstract

This paper studies the transmission channel of asset bubbles in a heterogeneous firm model with endogenous entry and exit. We highlight the effects of asset bubbles along the extensive margin: the aggregate bubble can boost real economic activities by affecting firms' entry and exit decisions. Moreover, the model predicts the selection effect of bubbles: bubbly firms—firms with asset bubbles—are less productive than bubble-less firms. Finally, we provide empirical evidence that supports bubbles' effects along the extensive margin.

Keywords: Rational Bubbles, Extensive Margin, Firm Dynamics, Heterogeneous Firms
JEL Classification: E32, E44, E47

*This study supersedes an earlier paper by Haozhou Tang "Asset Price Bubbles and the Distribution of Firms". We thank Adrien Auclert, Luigi Bocola, Pavel Brendler, Andrea Caggese, Davide Debortoli, Jordi Galí, Manuel García-Santana, Calixto López, Alberto Martin, Nuno Palma, Francisco Queirós, Josep Pijoan-Mas, Michael Reiter, Pontus Rendahl, Edouard Schaal, Aleh Tsyvinski, Jaume Ventura, Pengfei Wang, and the participants at the Bank of Mexico, the XXV Meeting of the Central Bank Researchers Network, 2021 Annual Conference of the International Association for Applied Econometrics, and the IV Workshop of the Spanish Macroeconomics Network for the insightful discussions. We thank Luis Calderon for proofreading. The views expressed herein are the sole responsibility of the authors and do not necessarily reflect the views of the Bank of Mexico.

*Bank of Mexico, htangp@banxico.org.mx.

†Institute for Macroeconomics and Econometrics—University of Bonn, donghai.zhang@uni-bonn.de.

1 Introduction

Researchers have focused on the effects of asset bubbles on the aggregate economy because of the recent boom and bust of asset prices and their association with business cycles, e.g., Farhi and Tirole (2012), Martin and Ventura (2012), Miao and Wang (2018), Ikeda and Phan (2019), and Asriyan *et al.* (2020).[1] These previous studies emphasize asset bubble's effects along the *intensive margin*: bubbles stimulate firms' production (e.g., by relaxing financial constraint) without affecting the number of firms. However, asset bubbles are linked to the values of firms, and therefore, entry and exit—the two most important decisions for firms—are linked to the acquisition/loss of bubbles. The current paper focuses on the effects of asset bubbles along the *extensive margin*: asset bubbles affect the aggregate economy by influencing firms' entry and exit decisions. Is the extensive margin a relevant transmission channel of asset bubbles? What are the implications of asset bubbles for the distribution of firms?

This study addresses these hitherto unaddressed questions through a theoretical model. Moreover, we provide empirical evidence supporting the effects of asset bubbles along the extensive margin.

Theoretically, this study builds a heterogeneous firm model with endogenous entry and exit, allowing for rational bubbles.[2] The framework helps in analyzing the effects of rational bubbles on real economic activities through firms' entry and exit decisions. The model nests the standard bubble-less firm dynamics model as a special case.[3] Firms make entry (exit) decisions by comparing their value of operating in the market, which is composed of a fundamental and a bubble component, with the entry cost (outside option of exiting). Bubbles can influence decisions through their direct effects on firms' values. Specifically, for bubbly incumbents (i.e., existing firms whose value carries a bubble component), exiting the market leads to the loss of bubbles. In other words, bubbles act as an *exiting tax*, making it unattractive for a firm to exit. For potential entrants, bubbles act as an *entry subsidy*: ceteris paribus, the value of entering the market is higher because of an asset bubble. Hence, bubbles stimulate new firms to enter the market.

The calibrated model confirms the relevance of bubbles' effects along the extensive margin. Particularly, the total number of firms and aggregate output increase after a

[1]See also Barlevy (2012) and Martin and Ventura (2018) for reviews of the literature.

[2]There are various definitions of bubbles. We define a bubble as a pyramid scheme that pays no dividend, but is traded at a positive price: see e.g., Blanchard and Watson (1982), Tirole (1985), Martin and Ventura (2012), and Galí (2014).

[3]For example: Hopenhayn (1992) and Clementi and Palazzo (2016).

positive bubble shock (**Result 1**). Moreover, due to the extensive margin, the effects of an asset bubble on real economic activities are *slow-moving* and *hump shaped*, whereas bubble shocks affect asset prices *on impact* (**Result 2**).

Furthermore, bubbles have micro-level implications. Firms are heterogeneous in productivity, capital, and bubble component. We compare the characteristics of bubbly firms to their bubble-less counterparts along the balanced-growth path (BGP) in the model. On average, a bubbly firm is characterized by lower productivity and is shown to accumulate less capital than the corresponding variables of a bubble-less firm over the entire life cycle. The underlying intuition is that bubbles act as an exit tax and thus prevent low-productivity firms from exiting. Moreover, by acting as an entry subsidy, bubbles induce low-productivity firms to enter the market; therefore, bubbly firms are less productive on average, namely, the *selection effect* of bubbles. Consequently, bubbly firms accumulate less capital.

Finally, we provide empirical evidence that is consistent with our model by estimating a structural vector autoregressive (SVAR) model. Controlling for fundamental shocks, such as productivity, monetary policy, and risk premium shocks, we construct "bubble shocks" using the medium-run restriction, as pioneered by Uhlig (2003, 2004).[4] The medium-run restriction identifies "bubble shocks" as the shocks that, among unobserved fundamental shocks that are not controlled for in the VAR, maximize the forecast error variance decomposition of our proxy of the non-fundamental component at the frequency of business cycles.[5]

Using this empirical strategy, we uncover that an expansionary "bubble shock" increases the total number of firms by increasing the number of new firms and reducing firms' exit rates. Moreover, while the "bubble shock" has peak effects on asset prices and its non-fundamental component on impact (within a year), its effects on real economic activities are slow-moving and hump-shaped. These empirical findings are consistent with the effects of bubbles along the extensive margin that the model predicts, namely **Result 1** and **Result 2**. In addition, we provide evidence that the identified "bubble shock" is not confounded with fundamental shocks such as risk premium shocks, shocks to the short-term interest rate, and credit supply shocks.

[4]The medium-run restriction is widely used to identify news shocks (Barsky and Sims 2011) and other structural shocks; see Zeev and Pappa (2017) and Ben Zeev *et al.* (2017). More recently, Levchenko and Pandalai-Nayar (2020) use this strategy to identify sentiment shocks in an open economy.

[5]This paper uses the price-fundamental differential component and the non-fundamental component interchangeably to denote the deviations of an asset's price from its fundamental value. The fundamental component is defined as the present value of future real dividends.

Literature Review The study contributes to a growing body of literature on asset bubbles. Tirole (1985) views bubbles as pyramid schemes or assets without intrinsic value. He argues that bubbles crowd out investment and lower output. Caballero and Krishnamurthy (2006), Farhi and Tirole (2012), Martin and Ventura (2012), Ventura (2012), Hirano and Yanagawa (2017), Miao and Wang (2018), Bengui and Phan (2018), Ikeda and Phan (2019), Biswas *et al.* (2020), Guerrón-Quintana *et al.* (2020), and Asriyan *et al.* (2020) argue that bubbles relax financial constraints.[6] Hellwig and Lorenzoni (2009), Martin and Ventura (2016), and Domeij and Ellingsen (2018) study bubbles in debt markets. Olivier (2000),[7] and Queiros (2019) claim that imperfect competition in the goods market also allows bubbles to boost real economic activities.[8] The existing works rely on market frictions for bubbles to boost real economic activities and focus on the effects of bubbles along the intensive margin. We emphasize an alternative transmission channel of asset bubbles through firms' endogenous entry and exit decisions. This hitherto unexplored mechanism gives rise to the co-movement of asset prices and real economic activities even in competitive markets. Moreover, we find the selection effect of bubbles through which bubbles affect the distribution of firm-level variables, such as productivity and capital.

This study is related to previous studies that emphasize the macroeconomic implications of firm-level heterogeneity. Caballero and Engel (1999), Thomas (2002), Khan and Thomas (2008), Bachmann *et al.* (2013), and Winberry (2020) study the implications of capital adjustment costs at the micro level for the behavior of aggregate investment. Khan and Thomas (2013), Senga *et al.* (2017), Arellano *et al.* (2019), and Ottonello and Winberry (2018) investigate the implications of financial frictions in models with heterogeneous firms. Bachmann and Bayer (2014) and Bloom *et al.* (2018) consider the impacts of uncertainty shocks in models with heterogeneous firms. Our model is closely related to studies on heterogeneous firms that emphasize the endogenous entry and exit, pioneered by Hopenhayn (1992). Jaimovich and Floetotto (2008), Bilbiie *et al.* (2012), Lee and Mukoyama (2015), Clementi and Palazzo (2016), Sedláček and Sterk (2017) and Bilbiie and Melitz (2020) find that firm entry and/or exit justify important features of business

[6] See also Miao *et al.* (2015) and Larin (2019) for the quantitative importance of bubble shocks as the source of business cycles.

[7] Olivier (2000) models R&D, which can somehow be interpreted as firm entry; however, the lack of firm heterogeneity removes the selection effect that changes the distribution of entrants. Importantly, studies have not considered endogenous firm exit, which plays a key role in our model.

[8] Galí (2014, 2020) and Domeij and Ellingsen (2021) analyze the interaction between monetary policy and rational bubbles, in the presence of goods market frictions. See also Vuillemey and Wasmer (2020), who study the effects of bubbles in a frictional labor market.

cycles. In addition, Jermann and Quadrini (2007) relate asset price booms to aggregate productivity gains in a model with decreasing returns to scale and financial frictions. Our framework contributes to this literature by generalizing a firm dynamics model, allowing for rational bubbles. In our model, the bubble size is an additional dimension of firm heterogeneity. Through the selection effect, we show that bubbles can affect the distribution of firm-level productivity and capital. Moreover, the present study shows that empirical evidence uncovered in this paper can be reconciled by allowing for the existence of bubbles in an otherwise standard firm dynamics model.

The remainder of the paper is structured as follows. Section 2 provides a simple example of the existence of rational bubbles. Section 3 introduces the full-fledged model, and Section 4 presents the predictions of the model. Section 5 provides empirical evidence. Section 6 discusses the assumptions made in the current paper. Finally, Section 7 concludes.

2 An Example of a Rational Bubble

This section provides a simple example of the emergence of rational bubbles at the firm level and the condition for the aggregate bubble to be sustainable.

2.1 Rational Bubbles

Consider a friction-less, small open economy with risk-neutral households. Households plant trees that produce $d > 0$ unit of fruit in every period. Without the loss of generality, we assume that planting trees incurs no cost, and we consider fruits as the numeraire in the economy. There is no uncertainty. Households plant new trees and the total number of trees increases at the rate g. We assume that the world's output grows at the same rate g so that the small open economy remains small compared with the world's economy. The following equation gives the value of a tree at time t:

$$V_t = d + \frac{1}{R} V_{t+1}, \tag{1}$$

where $R > 1$ denotes the gross risk-free interest rate.

Solving equation (1) gives

$$V_t = \frac{R}{R-1}d + \lim_{\tau \to \infty} \left(\frac{1}{R}\right)^{\tau} V_{t+\tau}. \tag{2}$$

The following equation is a popular assumption made in the literature to exclude bubbles:

$$\lim_{\tau \to \infty} \left(\frac{1}{R}\right)^{\tau} V_{t+\tau} = 0, \tag{3}$$

which we call the no-bubble condition. This condition implies that there exists a unique solution for the value of trees:

$$V = F \equiv \frac{R}{R-1}d, \tag{4}$$

where F denotes the value of a tree under the assumption of a no-bubble condition. In the absence of a bubble, the value of a tree is determined to be equal to its net present value of fruits. However, there exist other solutions to equation (1) insofar as we relax the no-bubble condition (3). Assume the following solution:

$$V_t = F + B_t, \tag{5}$$

where F is defined in (4) and B_t is a process that satisfies the following property:

$$B_t = \left(\frac{1}{R}\right) B_{t+1}. \tag{6}$$

It is straightforward to verify that equations (5) and (6) indeed solve equation (1). In equation (5), the value of a tree is decomposed into two components. The first component F represents the net present value of fruits, as defined in equation (4). We label F as the fundamental value of trees/assets. The second component B_t is a pyramid scheme. Hereafter, we label it as a bubble. Free-disposal implies that $B_t \geq 0$. Note that equation (5) nests equation (4) as a special case when $B_t = 0, \forall t$.

There exists a unique fundamental component F. However, infinitely many processes exist for B_t that are consistent with (6). Any solution satisfying equations (5) and (6) also satisfies the conditions that household and firms have rational expectations and solve their problems optimally. Hence, bubbles (B_t) are labeled as rational bubbles in the literature.

2.2 Aggregate Bubble at the Balanced-growth Path

It follows from (6) that for every tree, a rational bubble grows over time. In aggregate, if bubbles persistently outgrow output, the world resource will ultimately become insufficient to purchase the bubbles. Under this condition, as argued by Tirole (1985), rational expectations exclude bubbles because investors expect bubbles to be unsustainable. Therefore, the existence of bubbles requires the aggregate bubble to remain stationary relative to the aggregate output at the BGP.

The number of trees M_t grows at a constant rate g:

$$M_{t+1} = M_t + g \cdot M_t \tag{7}$$

Conjecture that there exists an equilibrium where every new tree starts with an initial bubble $0 < \bar{B} < \infty$. Owing to the entry of new trees, the aggregate bubble B_t^a evolves as follows:

$$B_{t+1}^a = R \cdot B_t^a + g \cdot M_t \cdot \bar{B}. \tag{8}$$

The right hand side of equation (8) comprises two parts: $R \cdot B_t^a$ represents the existing bubbles from the existing old trees, and $g \cdot M_t \cdot \bar{B}$ represents the new bubbles brought by new trees. The law of motion (8) can be rewritten as

$$\frac{B_{t+1}^a}{M_{t+1}} = \frac{R}{1+g} \cdot \frac{B_t^a}{M_t} + \frac{g}{1+g} \cdot \bar{B}. \tag{9}$$

We assume that $R < 1 + g$ such that $\frac{B_t^a}{M_t}$ is stationary. Otherwise, bubbles are unsustainable. In fact, on the BGP: $\frac{B_t^a}{M_t} = \frac{g\bar{B}}{1+g-R}$. It is straightforward to show that the market capitalization of trees divided by the aggregate output is also a constant. Specifically, bubbles are sustainable (i.e., the market for trees clears on the BGP) and consistent with rational investment behavior. Thus the conjecture is verified, and the equilibrium exists.

Because $\bar{B}$ can take different values, multiple equilibria exist in which the value of trees follows (5) and (6), and every new tree is attached with an initial bubble $0 < \bar{B} < \infty$. The realization of $\bar{B}$ depends on market sentiment, or market psychology, as per Asriyan *et al.* (2020).

Briefly, the condition $R < 1 + g$ ensures the existence of BGP along which rational bubbles arise.

These results are shared by the complete model in Section 3. However, two features of the complete model are absent in the simple example. First, a standard model with an infinitely living representative household features an equilibrium real interest rate that is higher than the growth rate ($R > 1 + g$), which violates the existence of rational bubbles. To overcome this issue, the model in Section 3 features overlapping generations of households with each generation living for a finite horizon.[9] Second, in this stylized example, bubbles do not affect F. In the complete model, bubbles affect firms' entry and exit decisions and therefore, affect the fundamental components of firms' values.

3 Model

Our model builds on standard firm dynamics models with competitive markets.[10] We extend the existing models by introducing a firm-level bubble component that raises the equity price of a firm above its net present value of dividends.[11] In the model, bubbles influence the firms' selection mechanism. As a firm's exit incurs the loss of bubbles, bubbles discourage incumbents from exiting. Moreover, prospective entrants are encouraged to enter the market to obtain bubbles.

3.1 Preferences

Time is discrete and the horizon is infinite. There is no aggregate uncertainty, and the economy is populated with overlapping generations. In each period, a new generation of individuals who live for T periods joins the economy. A representative household i in the new generation maximizes lifetime utility:

$$U_{it} = E_{it} \sum_{\tau=0}^{T-1} \beta^{\tau} C_{it+\tau}, \tag{10}$$

where $\beta \in (0,1)$ is the discount factor and $C_{it} \geq 0$ denotes the consumption. The new cohort's size N_t follows:

$$N_{t+1} = (1+g)\, N_t. \tag{11}$$

[9] See Domeij and Ellingsen (2018) and Galí (2020) for models of infinite-horizon agents with sustainable pyramid schemes.

[10] See Hopenhayn (1992) and Clementi and Palazzo (2016).

[11] We model a bubble as a pyramid scheme that pays no dividend but is traded at a positive price: see, e.g., Blanchard and Watson (1982), Tirole (1985), Martin and Ventura (2012), Galí (2014), Domeij and Ellingsen (2018), and Ikeda and Phan (2019).

Consequently, the size of the entire population also grows at the rate g.

3.2 Firms

Firms produce a homogeneous good in a perfectly competitive market. The production technology features decreasing returns to scale:

$$y_{jt} = A_t \varphi_{jt} k_{jt}^{\alpha} \tag{12}$$

with $\alpha \in (0,1)$. A_t denotes the common productivity component that is identical across firms. φ_{jt} is an idiosyncratic productivity shock. k_{jt} denotes the capital stock, which is predetermined. The idiosyncratic productivity shock φ_{jt} follows a Markov process:

$$\log \varphi_{jt+1} = \rho \log \varphi_{jt} + \varepsilon_{jt+1}, \tag{13}$$

where $\rho \in (0,1)$ and $\varepsilon_{jt} \sim N\left(0,\sigma^2\right)$ $\forall t$, $\forall j$.

Firms are owned by households. A firm's equity is issued and traded in a frictionless market. For each firm, we normalize the number of shares to one. Firms operate indefinitely and are subject to endogenous exit. The start-of-period value of a firm is equal to

$$V(\lambda,\mu,k) = A\varphi k^{\alpha} - c^f + \max\left\{V^c(\lambda,\mu,k), V^x(k)\right\}, \tag{14}$$

where λ includes aggregate state variables, μ contains the idiosyncratic state variables other than capital stock k, and c^f denotes a fixed operation cost. We suppress the subscripts to reduce the notational burdens. A variable X' denotes the variable X in the next period. After production, the firms' owners decide whether or not to exit from the market by comparing their value of continuation V^c with the value of exit V^x.

If firms exit, they divest all the capital stock, and the owners will then receive the scrap value. Therefore, the value of exit is

$$V^x(k) = (1-\delta)k - g(k,0), \tag{15}$$

where δ denotes the depreciation rate during production and $g(k,k')$ denotes the cost of adjusting capital stock from k to k'. The adjustment cost follows a standard functional form:

$$g\left(k,k'\right) = c_0 \mathbb{1}\left\{k' \neq (1-\delta)k\right\} k + c_1 \left(\frac{k' - (1-\delta)k}{k}\right)^2 k, \tag{16}$$

where $c_0, c_1 \in (0,1)$, and $\mathbb{1}\{k' \neq (1-\delta)k\}$ is the indicator function that equals 1 if firms invest and equals 0 otherwise.

Firms continue to operate and invest if and only if $V^c(\lambda,\mu,k) \geq V^x(k)$. If a firm remains in the market, the owners can trade the shares in a friction-less market and decide how much to invest for future production. The continuation value of a firm is equal to

$$V^c(\lambda,\mu,k) = \max_{k'} \left\{ (1-\delta)k - k' - g(k,k') + \frac{1}{R} \int V(\lambda',\mu',k')\, dJ(\lambda',\mu'|\lambda,\mu) \right\}, \tag{17}$$

where $J(\lambda',\mu'|\lambda,\mu)$ denotes the transition probability of λ and μ,[12] and R denotes the return on firm equity.[13] The factors that the vector μ includes depend on whether the no-bubble condition is imposed.

The Fundamental Component The continuation value $V^c(\lambda,\mu,k)$ can be expressed as

$$V^c(\lambda,\mu,k) = F^c(\lambda,\mu,k) + V^c_\infty(\lambda,\mu,k). \tag{18}$$

Equation (18) is analogous to equation (2): $F^c(\lambda,\mu,k)$ is the net present value of cash inflows on continuation, while $V^c_\infty(\lambda,\mu,k)$ is the discounted expected continuation value in the limit. We label $F^c(\lambda,\mu,k)$ as the fundamental component of the continuation value $V^c(\lambda,\mu,k)$. The formal definition of $F^c(\lambda,\mu,k)$ and $V^c_\infty(\lambda,\mu,k)$, as well as the derivation of (18), are presented in Appendix B.1.

A popular assumption to exclude bubbles in the literature is to impose $V^c_\infty(\lambda,\mu,k) = 0$, which is also called the no-bubble condition. Under this assumption, a firm's value solely consists of its fundamental component. In this scenario, the idiosyncratic state μ consists of the idiosyncratic productivity shock φ. We relax this assumption in the study. Consequently, the bubble component of the $V^c(\lambda,\mu,k)$ arises, and the firm-level bubble component emerges as a new idiosyncratic state variable.

The Bubble Component By relaxing the no-bubble condition, the continuation value of a firm can be written as

$$V^c(\lambda,\mu,k) = F^c(\lambda,\mu,k) + B, \tag{19}$$

[12]The notation used in this section can also be applied to an environment with aggregate uncertainty.

[13]Households are risk-neutral, and therefore the equity market clears only if all firms have the same expected return.

where B represents the deviation of the continuation value from its fundamental component, which is labeled as the bubble component or a bubble. Following the studies on the rational bubbles, we model B as a a pyramid scheme:

$$B = \frac{1}{R} \int_{\Phi'} B' dJ \left(\left(\lambda', \mu' | \lambda, \mu \right) \right), \tag{20}$$

where $\Phi' \equiv \{ (\lambda', \mu') \,| V^c (\lambda', \mu', k') \geq V^x (k') \}$. In Appendix B.2, we verify that (19) and (20) are consistent with (18). Only the bubbles of surviving firms are tradable. If a firm is expected to exit in the subsequent period, the present value of its bubble component is equal to zero. Bubbles can be considered as assets without intrinsic value: the assets that do not generate cash inflows. The size of a bubble depends solely on the expectation of its future size. We assume free disposal of bubbles such that $B \geq 0$. The solution under the no bubble condition $V^c (\lambda, \mu, k) = F^c (\lambda, \mu, k)$ is a special case when $B = 0$ for all firms in each period.

We assume that bubbles are stochastic and subject to investor sentiment shocks. Bubble-less firms ($B = 0$) continue to be bubble-less in subsequent periods. Regarding a bubbly firm ($B > 0$) that does not exit the market, its bubble component can collapse (hit by a negative sentiment shock) with a probability $1 - p^b$. With probability p^b, the bubble component rolls over. Firms are subject to the negative sentiment shock with a common probability. However, the realization of the bubble crashes is independent across firms. Among those firms whose bubble components carry over to the next period, we assume that their bubble components evolve according to $B' = \left(\frac{1}{R} \cdot p^b \cdot p^o (\lambda, \mu, k') \right)^{-1} B$; thus, the process that governs the dynamics of the bubble component is

$$B' = \begin{cases} 0, & \text{with probability } 1 - p^b \\ \left(\frac{1}{R} \cdot p^b \cdot p^o (\lambda, \mu, k') \right)^{-1} B, & \text{with probability } p^b \end{cases} \tag{21}$$

where $p^o (\lambda, \mu, k') \equiv \int_{\Phi'} dJ^o (\lambda', \mu' | \lambda, \mu)$. $J^o (\lambda', \mu' | \lambda, \mu)$ denotes the transition probability of λ and μ conditional on the roll-over of the bubble component. $p^o (\lambda, \mu, k')$ is the survival probability if the firm does faces no bubble crash in the next period.

The discounted value of the next period's bubble is equal to

$$\begin{aligned}\frac{1}{R}\int_{\Phi'} B' dJ\left(\left(\lambda',\mu'|\lambda,\mu\right)\right) &= p^b\cdot\frac{1}{R}\int\limits_{\Phi'}\left(\frac{1}{R}\cdot p^b\cdot p^o\left(\lambda,\mu,k'\right)\right)^{-1} BdJ^o\left(\lambda',\mu'|\lambda,\mu\right)\\ &= \left(p^o\left(\lambda,\mu,k'\right)\right)^{-1} B\int\limits_{\Phi'} dJ^o\left(\lambda',\mu'|\lambda,\mu\right)\\ &= B.\end{aligned}$$

Hence, equation (21) guarantees equation (20) to hold. The process (21) ensures that the bubbles are consistent with investor optimality and rational expectations while maintaining the model's tractability. As in Galí (2014) and Martin and Ventura (2016), bubbles follow a "backward-looking" process so that the size of future bubbles (if bubble crashes do not occur) can be specified, given the current bubble size. The "backward" feature of the process facilitates our analysis because we can record the dynamics of bubbles on a given route.[14]

A key feature of (21) is the idiosyncratic bubble crashes. Hence, our model allows for the coexistence of bubbly and bubble-less firms. Equation (21) also implies that the growth rate of bubbles is increasing in the probability of a bubble crash and firm exit. Intuitively, bubbles must grow fast enough to compensate for the possible loss if firms are likely to exit or if bubbles are likely to crash. Alternative processes of bubbles consistent with investor rationality exist. However, these alternative processes do not affect the existence of bubbles' effects along the extensive margin and the selection effects of bubbles emphasized in this study. Our approach is a simple method for incorporating idiosyncratic bubble crashes.[15]

Discussions We now discuss three features of the model. First, by relaxing the no-bubble condition, idiosyncratic states μ of a firm include the size of its bubble B, in addition to idiosyncratic productivity φ. To confirm this feature, recall that firms exit if

$$V^c\left(\lambda,\mu,k\right) < V^x\left(k\right).$$

[14] Galí (2014) and Martin and Ventura (2016) state that there are multiple possible backward-looking solutions of bubbles that are consistent with investor rationality. See also Galí (2020) and Asriyan *et al.* (2020). In contrast, Farhi and Tirole (2012), Miao and Wang (2018) characterize the size of bubbles through forward-looking solutions. Miao *et al.* (2019) discuss the differences between the two approaches.

[15] As stated by Galí (2014) and Martin and Ventura (2018), how to choose the process is still debated.

Bubbles enhance the value of continuation and reduce the probability of firms exiting.[16] The model could be solved by imposing/assuming that the policy function and value functions do not depend on B. Such a solution is valid. However, this is equivalent to imposing the no-bubble condition.[17] This study relaxes the no-bubble condition and focuses on the case with $B > 0$.

Second, studies on rational bubbles focus on discrete aggregate bubble shocks: shocks that shift the entire economy between a bubbly state and a bubble-less state.[18] In our model, bubble shocks can be captured by continuous changes in p^b. A decrease in p^b implies a lower likelihood to roll over the bubble component. If p^b drops to 0, the entire economy shifts to a bubble-less state. In Section 4, we study the transition dynamics (perfect-foresight equilibrium paths) following unexpected changes in p^b.

Finally, sustainable bubbles are essential for a bubbly equilibrium; that is, the aggregate bubble-output ratio has to be stationary. The economy would eventually deplete its resources if bubbles persistently outgrow its output, and thereafter such equilibrium does not exist. Section 3.4 discusses the condition under which a bubbly BGP is sustainable.

3.3 New Cohorts and Firm Entry

Only newborn households have the opportunities to create new firms. We assume that each individual of the new generation has access to a unit continuum of projects. Projects can become firms that eventually produce. These projects are called *potential entrants*. The problems of potential entrants are similar to that of incumbents: they choose whether to produce in the following period, that is, to enter the market or not. If they decide to produce, they choose the amount of investment. Each potential entrant has its own realization of productivity and bubble. A potential entrant's productivity is drawn from the distribution

$$\varphi_t \sim \log N\left(-m\frac{\sigma}{\sqrt{1-\rho^2}}, \frac{\sigma^2}{1-\rho^2}\right),$$

[16]One might be concerned that V^c is possibly independent from B. Conversely, assume that V^c is independent from B. Then, it can be inferred from equations (14)-(17) that the fundamental component F^c is independent from B, too. Hence, $V^c - F^c$ is independent from B, which contradicts $V^c = F^c + B$ for $B > 0$.

[17]For example, recall that in Section 2, if V is expressed as a function of d, there exists a unique solution to equation (1): $V_t = V(d) = \frac{R}{R-1}d$. However, if V is expressed as a function of d and B_t, then $V_t = V(d, B_t) = \frac{R}{R-1}d + B_t$ with $B_t = \frac{1}{R}B_{t+1}$.

[18]See, e.g., Martin and Ventura (2012) and Miao and Wang (2018).

where m captures the possible difference between the initial productivity draw and the idiosyncratic productivity process (13).[19] Without the loss of generality, we assume that the investors are optimistic about a potential entrant with probability p^b—the same parameter that governs the survival of existing bubbles. If investors are optimistic, the potential entrants receive an initial bubble B_0 before their entry.

Analogous to incumbents, potential entrants can only issue tradable equity if they decide to produce in the subsequent period. Potential entrants enter the market if and only if the value of entry is non-negative:

$$V^e(\lambda, \mu) \geq 0. \tag{22}$$

Potential entrants start with zero capital stock. Hence, the value of entry is:

$$V^e(\lambda, \mu) = V^c(\lambda, \mu, 0). \tag{23}$$

Following (19), the value of entry can be decomposed into the sum of a fundamental and bubble component. The latter is equal to B_0 or 0 depending on the realization of market sentiment. We define *entrants* as the potential entrants who decide to enter.

3.4 Bubbly Balanced-Growth Path

A dynamic equilibrium consists of value functions, decision rules, prices, and the measure of firms, which satisfy the optimality conditions of firms and households, and clear all markets. Besides, in equilibrium, the measure (number) of firms' law of motion is consistent with the decision rules of investment, entry, and exit. We introduce a formal definition of equilibrium in Appendix B.3.

We characterize the equilibrium path for a given pair of p^b and B_0, which are exogenously determined. Appendix B.4 describes the solution method to obtain the equilibrium path.[20] Here, we discuss the condition under which a bubbly BGP is sustainable. Along the BGP, the relative composition of firms remains stable, while the total amount

[19] Although potential entrants draw productivity, they do not produce prior to entry. One may label the initial draw as the advance information about future productivity upon entering the market.

[20] For prefixed B_0 and process (21), we take as given that there exists a unique equilibrium path. We follow the typical procedure adopted in the literature by verifying that our iterative numerical solution leads to the same results, given different initial guesses of value functions. As in Galí (2014) and Asriyan *et al.* (2020), our model features multiple equilibria that are characterized by various combinations of p^b and B_0: when feeding the model different p^b and B_0, we find a different equilibrium paths that satisfies the definition in Appendix B.3.

of firms increases at rate g. In Section 4, we calibrate the BGP and study the transition dynamics. The discussions in the remaining of the current section are similar to those provided in Section 2.2.

At the firm-level, no upper bound exists for the size of bubbles, as indicated by equation (21). However, the aggregate bubble must be bounded relative to the aggregate output for the equilibrium to be sustainable. In the presence of firm entry and exit, the firm-level explosiveness of bubbles does not lead to the explosiveness of an aggregate bubble because in every period, firms with large bubbles exit the market and new firms with smaller bubbles replace them. Next, we show that, if $R < 1 + g$, a bubbly BGP is sustainable.

In the equity market, the total amount of tradable bubbles B_A is equal to

$$B_A = B_I + B_N,$$

where B_I and B_N denote the amount of bubbles attached to surviving firms and entrants, respectively. Since there is no aggregate uncertainty, according to (21), we have

$$\begin{aligned} B_I' &= \int \left[\left(\frac{1}{R} \cdot p^b \cdot p^o \left(\lambda, \mu, k' \right) \right)^{-1} B \right] \cdot \left[p^b \cdot p^o \left(\lambda, \mu, k' \right) \right] d\eta \left(\varphi, B, k' \right) \\ &= R \int B \cdot d\eta \left(\varphi, B, k' \right) = R B_A, \end{aligned} \tag{24}$$

where $\eta\left(\varphi, B, k'\right)$ represents the distribution of the incumbents and entrants. Therefore, the law of motion for the aggregate bubble is given by

$$B_A' = R B_A + B_0 M'. \tag{25}$$

Here, $B_N = B_0 M$, where M denotes the number of bubbly entrants. The bubble-output ratio follows

$$\frac{B_A'}{Y'} = R \frac{B_A}{Y} \frac{Y}{Y'} + \frac{B_0 M'}{Y'}. \tag{26}$$

On the BGP, both the measure of bubbly entrants M and the aggregate output Y grow at the rate g. Thus, (26) can be rewritten as

$$b' = \frac{R}{1+g} b + b_0,$$

where $b \equiv \frac{B_A}{Y}$ denotes the bubble-output ratio and $b_0 \equiv \frac{B_0 M_0}{Y_0}$. The bubble-output ratio follows a stationary process as long as $R < 1 + g$. Indeed, the ratio stays at a constant $\left(1 - \frac{R}{1+g}\right)^{-1} b_0$ on the BGP.

The bubble–output ratio stays at a constant on the BGP. At the firm level, bubbles grow faster than the rate of return on investment because of the likelihood of a bubble crash and firm exit. In the aggregate, however, owing to firm exits and idiosyncratic bubble crashes, the aggregate bubble from continuing incumbents grows at the same rate as the investment return. A BGP is sustainable insofar as the growth rate of output exceeds the rate of return from investments.

4 Results

Calibration We calibrate our model on a BGP. Table 1 presents the parameter values. The model is in annual frequency. Panel A lists the pre-assigned parameters. We set the discount factor $\beta = 0.98$, the decreasing returns to scale $\alpha = 0.65$, and the capital depreciation rate $\delta = 0.1$. The processes of the idiosyncratic productivity shocks are calibrated to the estimates provided by İmrohoroğlu and Tüzel (2014): $\rho = 0.7$ and $\sigma = 0.38$. We set the growth rate g to 2.81%, which corresponds to the average growth rate of real GDP in the United States between 1976 and 2019.

In our analysis of BGP and transition dynamics, we focus on the scenario where $R = \beta^{-1}$.[21] T does not affect the target moments or the impulse responses studied. This study assumes a sufficiently large T such that the computed equilibrium paths are consistent with household rationality.[22] The BGP is sustainable because $R < 1 + g$.

Panel B reports the remaining parameters estimated to minimize the distance between the selected targets and their model-implied counterparts. First, using data from the Business Dynamics Statistics, we target the share of new establishments, and those of age 4. These two moments are related to business dynamism and firm life cycles.[23] Second, we target PE ratios and consider two groups of firms in Compustat: firms above the 75th percentile of assets and those below. We treat the groups as two portfolios and calculate their

[21] In equilibrium R cannot be smaller than β^{-1}; otherwise the equity market cannot clear. We do not consider the case $R > \beta^{-1}$ in which finding R in equilibrium is computationally expensive.

[22] Household rationality requires households to consume all their wealth before they die and thus the aggregate consumption to be at least as great as the wealth of the oldest cohort. Appendix B.5 demonstrates that the aggregate consumption is at least as great as the wealth of the oldest cohort provided that T is sufficiently large.

[23] We calculate the shares of tradable firms, i.e., the firms that exist at the end of a period.

Panel A: Fixed Parameters

Parameter	Description	Value
α	Decreasing returns to scale	0.65
ρ	Idiosy. shock persistence	0.70
σ	Idiosy. shock volatility	0.38
δ	Depreciation rate	0.10
β	Discount factor	0.98
g	Growth rate	2.81%

Panel B: Estimated Parameters

Parameter	Description	Value
m	Mean shift of entrants' prod.	1.63
b_0	Initial bubble component	443.57
c_f	Fixed cost of production	15.57
c_0	Fixed adjustment cost	$8.60 \cdot 10^{-4}$
c_1	Variable adjustment cost	0.039
p_b	Surviving probability of a bubble	0.92

Table 1: Parameters

average PE ratio between 1976 and 2016.[24] In the model, we match the two moments with the corresponding PE ratios for the firms that have survived at least 20 years. Finally, we target investment-related moments, including the average investment rate, inaction rate, and dispersion of investment rate, as reported by Zwick and Mahon (2017).[25] Table 2 reports the targeted moments in the data and the counterparts generated in the model. Overall, our model fits the selected moments well, except that the calibrated model underpredicts the PE ratio of big firms. However, the model matches the qualitative pattern in the data that the PE ratio decreases in size.

4.1 Firms' Life Cycles

Figure 1 plots the evolution of the exit rate, productivity, and capital of an average firm by age. We plot the life cycles of an average bubbly and bubble-less firms. The bubbly

[24] Firms' earnings can be negative in the data, which distorts the ranking of the PE ratio as a proxy for the degree of over-value of a firm. In the calculation of the two PE ratios, we exclude observations with negative earnings. In the calibration of the model, we apply the same filter.

[25] These moments can be found in Zwick and Mahon (2017)'s Appendix, Table B.1, Balanced Sample. To mimic their balanced sample, model moments are computed from firms that are alive for at least 12 years (the time dimension of their balance panel data) in the simulated panel data.

Moment	Data	Model
Share of new establishments	0.087	0.077
Share of four-year-old establishments	0.045	0.056
PE ratio of firms above 75th percentile of assets	16.36	8.50
PE ratio of firms below 75th percentile of assets	18.93	20.25
Investment inaction rate	0.237	0.220
Average investment rate	0.104	0.141
Standard deviation of investment rate	0.160	0.138

Table 2: Calibration Targets and Model Fit

labels those firms whose asset prices exceed their fundamental values. For both groups of firms, their capital and productivity increase in age. Therefore, older firms are less likely to exit.

Within an age cohort, an average bubbly firm features a lower exit rate, less capital, and a lower productivity across the entire life cycle than the respective variables of an average bubble-less firm. These results reflect that bubbles encourage the entry of the firms that otherwise would not enter the market without bubbles. Furthermore, with an asset bubble, firms have less incentive to exit. Consequently, bubbly firms are, on average, less productive and they accumulate less capital over the life cycle.

4.2 Impulse Responses Functions

We now analyze the impacts of a bubble shock. To this end, we feed in an unexpected innovation to p^b at $t = 0$, and the innovation decays at rates $\rho_p = 0.5$.[26] The economy is assumed to return to its steady state (BGP) at $t = 100$. The IRFs are then computed under the perfect foresight assumption.

Figure 2 plots the IRFs of selected variables. The bubble shock is normalized to have a unit impact on the aggregate real asset price, which is defined as the average price of firms having survived for at least 20 years. The choice of aggregate asset price is to mimic the use of S&P 500 as the aggregate asset price in the data, as well as in the calibration of the model. Two results are obtained. First, a bubble shock affects real economic activities. Second, while as a bubble shock has peak effects on asset prices and its non-fundamental component in the short run, its effects on real economic activities are slow-moving and hump-shaped.

[26] Results are robust for using alternative values of $\rho_p \in [0,1)$.

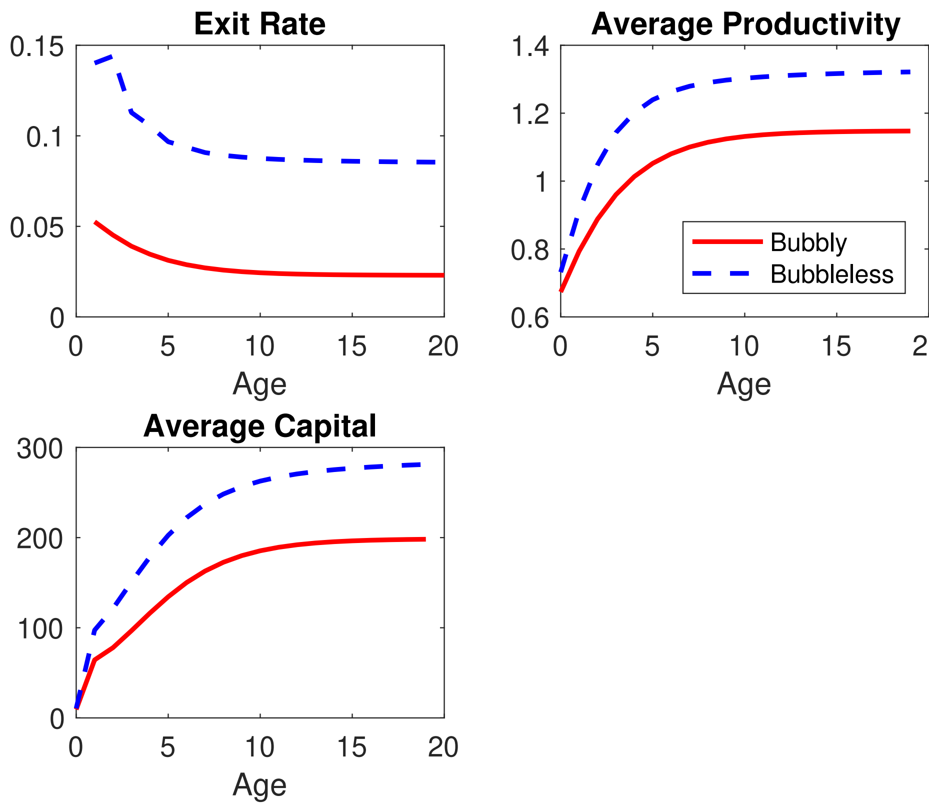

Figure 1: This figure plots the evolution of the exit rate, productivity and capital of an average firm by age.

A positive bubble shock affects the aggregate output through its effect on the total number of firms. The output per firm decreases because, through the selection effect of bubbles, the average productivity decreases at the firm-level. However, the aggregate TFP, measured as the Solow residual (in natural logarithm: $tfp = y_t - \alpha k_t$), increased after a positive bubble shock because of the decreasing return to scale and the increase in the number of firms. Firms that otherwise would exit the market or not enter the market in the absence a bubble shock contributes positively to the aggregate TFP because of the decreasing return to scales. Expansionary bubble shock increases household' wealth, which, in turn, improves social welfare (bottom panel of Figure 2). Social welfare is defined as the discounted sum of the lifetime utility of all living households $U_t \equiv \int U_{it} di$, where $U_{it} = E_{it} \sum_{\tau=0}^{T-1} \beta^{\tau} C_{it+\tau}$.[27]

Bubble shocks affect real economic activities along the extensive margin. An increase in the total number of firms is a result of the decline in the rate of exit and more entrants

[27] Appendix B.6 shows that the social welfare at time t is equal to the aggregate wealth held by living individuals at time t.

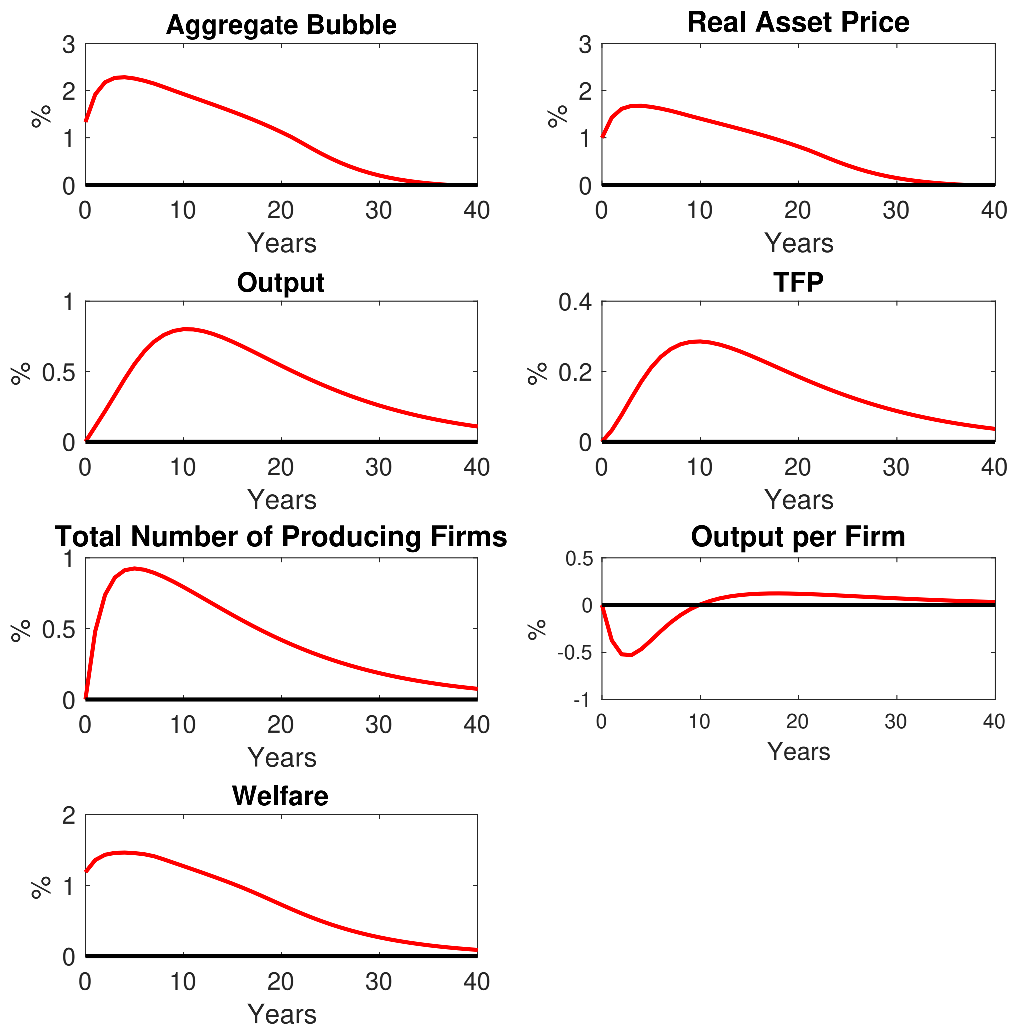

Figure 2: This figure plots the IRFs to a positive bubble shocks in the model. The shock is normalized to increase the real asset price on impact by 1%.

(Figure 3). In the presence of bubbles, firms have more incentive to enter the market and are less likely to exit. Interestingly, the entry rate shows an overshooting pattern because, in response to an expansionary bubble shock, there is a faster increase in the number of entrants than in the total number of firms in the short run but slower in the medium-run. The total number of exiting firms decreases in the short run, followed by a persistent increase in the medium-run. This result is driven by the increase in the total number of firms: holding the exit rate constant, an increase in the total number of firms leads to a higher number of firms that exit.

Overall, a heterogeneous firms model is developed with endogenous entry and exit that allows for asset bubbles. This study shows how aggregate bubble shocks affect the

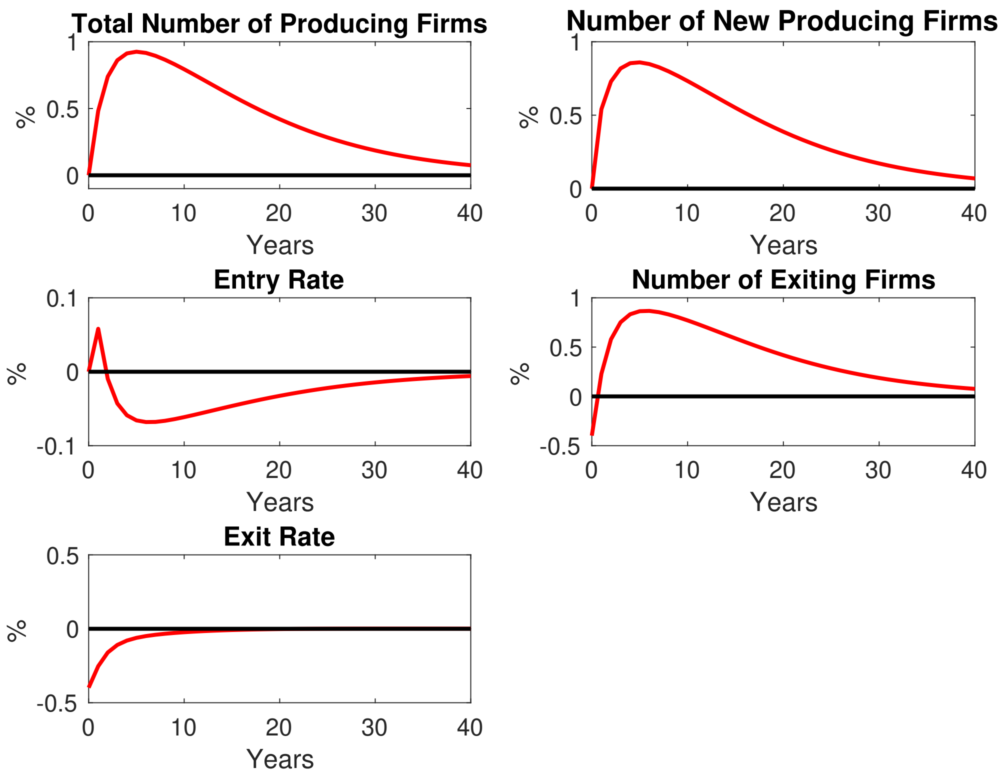

Figure 3: This figure plots the IRFs to a positive bubble shocks in the model. The shock is normalized to increase the real asset price on impact by 1%

economy in this framework. A standard model without bubbles is a special case in this framework. Theoretically, it is possible that the actual economy does not feature bubbles. In the remaining of the paper, empirical exercises are conducted to support the existence of bubbles in the data and their transmission mechanism.

5 Empirical Evidence

Section 5.1 constructs a proxy for the fundamental component (and the corresponding non-fundamental component) of the asset prices. Section 5.2 discusses the identification strategy. Section 5.3 presents empirical findings. Section 5.4 discusses the confounding issue that might arise in the baseline specification, and provides robustness checks of the main empirical findings by controlling for risk premium shocks, monetary policy shocks, and credit supply shocks.

5.1 A Proxy for the Fundamental Component

We study the portfolio of the top 500 firms (S&P 500) listed in the United States. The exact composition of this portfolio changes over time without affecting the representativeness of this asset. Let P_t denote the value of the portfolio that yields a stream of dividend (D_t). As discussed previously, the value (price) of such a portfolio can be decomposed into a fundamental component (F_t) and a bubble component (B_t)

$$P_t = F_t + B_t, \tag{27}$$

The fundamental component is the net present value of future dividends:

$$F_t \equiv E_t \left\{ \sum_{h=1}^{\infty} \left(\prod_{j=0}^{h-1} \Lambda_{t+j} \right) D_{t+j} \right\}, \tag{28}$$

where Λ_{t+j} is the discount factor. Log-linearizing Equation (28) leads to

$$f_t = c + \sum_{h=0}^{\infty} \left(\frac{\Gamma}{R} \right)^h \left[\left(1 - \frac{\Gamma}{R} \right) E_t\{d_{t+h+1}\} - E_t\{r_{t+h}\} \right], \tag{29}$$

where a variable in lower case denotes its log value and r_t denotes $-log(\Lambda_t)$. Γ and R denote the growth rate of dividend and $1/\Lambda_t$ in the steady state, respectively. This paper uses the price-fundamental differential component and the non-fundamental component interchangeably to describe the term: $p_t - f_t$.

Our identification of "bubble shocks" relies crucially on the measure of the price fundamental differential component; however, the fundamental component f_t is not directly observed in the data. Following Campbell and Shiller (1988) and Galí and Gambetti (2015), we construct a proxy of this unobserved component using VAR forecasts, and thus obtain a proxy for the corresponding price-fundamental differential component. We consider a VAR consisting of the following variables: TFP, real GDP (y_t), real dividend (d_t), real stock price (p_t), real interest rate (r_t) and the total number of firms (N_t). The choice of variables contained in our VAR is motivated by our theoretical model.

Let $Y_t \equiv [TFP_t, y_t, d_t, p_t, r_t, N_t]'$, the reduced form representation of our VAR model is

$$Y_t = B(L)Y_t + U_t, \tag{30}$$

where L is the lag operator and $B(L)$ is the matrix of lag order polynomials. Using the VAR's forecasts as the empirical proxies for $E_t\{d_{t+h}\}$ and $E_t\{r_{t+h}\}$, the construction of a proxy for the fundamental component according to (29) is straightforward. Appendix D provides the details regarding the source and sample coverage of the data set used in this study. In Appendix E, we demonstrate that the linear VAR adopted in this paper is a good approximation of our non-linear model introduced in Section 3. Moreover, using the data simulated from the theoretical model discussed in Section 4, Appendix E provides a simulation exercise illustrating that the *proxy* for the fundamental component constructed based on the VAR mimics the *true fundamental* in the model.

Having mentioned these, the constructed measure of $p_t - f_t$ does not necessarily coincide with the true price-fundamental differential for two reasons. First, the VAR's forecasts of d_{t+h} and r_{t+h} at time t might be different from the rational expectations of the corresponding variables. Second, in the empirical application, we use the real rate constructed from Moody's Baa Corporate Bond Yield to approximate the discount factor.[28] As will be discussed later, the medium-run restriction searches for the shock that maximizes the forecast error variance decomposition of our proxy of the non-fundamental component, controlling for fundamental shocks. Despite the proxy of the price-fundamental differential being merely a proxy, the identification strategy remains valid as long as the uncontrolled fundamental shocks are not as important as bubble shocks in explaining our proxy of the non-fundamental component. Moreover, in Section 5.4, we conduct robustness checks controlling for risk premium shocks, monetary policy shocks, and credit supply shocks; the results are similar to our baseline findings. Figure 6 plots the decomposition of the S&P 500 based on the estimation of our VAR.

5.2 The Identification of "Bubble Shocks"

This section constructs the "bubble shocks" using the medium-run restriction, as pioneered by Uhlig (2003, 2004). In the literature, the medium-run restriction is widely used to identify news shocks (Barsky and Sims 2011) and other structural shocks (Zeev and Pappa 2017 and Ben Zeev *et al.* 2017). More recently, Levchenko and Pandalai-Nayar (2020) use this strategy to identify sentiment shocks in an open economy.

The "bubble shocks" are identified as the shocks that maximize the forecast error variance decomposition (FEVD) of the constructed price–fundamental differential ($p_t - f_t$) in

[28] Galí and Gambetti (2015) use risk free rate as the discount factor. Our empirical findings are robust to this alternative choice.

the subsequent H periods, controlling for productivity shocks (unanticipated and anticipated shocks in the near future) in the baseline; selected structural shocks such as risk premium shocks, and credit supply shocks, and shocks to the short-term interest rate in robustness checks.[29] The identification strategy assumes that: fluctuations in our proxy of the non-fundamental component are mainly driven by sentiment/bubble shocks; fundamental shocks, particularly the ones that are *not controlled* for in the VAR, have limited impacts on the proxy of the non-fundamental component. Appendix C provides in-depth discussions of the identification issue and the construction of the "bubble shocks".

There are two points to be clarified. First, our identification strategy targets the *proxy* of the non-fundamental component instead of the true non-fundamental component. The noise introduced in the construction of the price-fundamental component does not detriment the identification of bubble shocks if the latter explain most of the fluctuations in the constructed non-fundamental component. Section 5.4 conducts robustness checks by including additional variables—measures of risk premium, credit supply, and monetary policy—into the baseline VAR. Estimation results confirm that the baseline findings are robust to alternative proxies of the non-fundamental component arising from different VAR specifications. .

Second, in the baseline specification, we do allow other shocks, such as risk premium shocks, to have significant impacts on *asset prices*. The validity of the identification requires the response of the proxy of the *fundamental component* to be similar to the response of *asset prices* so that the overall response of the constructed price-fundamental differential is small. Nevertheless, there might exist fundamental shocks that do not conform to this requirement. To address this issue, we control for productivity shocks (unanticipated and anticipated) in the baseline, and shocks to interest rate, risk premium, and credit supply in robustness checks. Overall, the empirical findings are similar both qualitatively and quantitatively across specifications. More importantly, the empirical effects uncovered in this section are consistent with the impulse responses displayed in Section 4.2.

We control for anticipated productivity shocks (news shocks) up to 3 years ahead, and we set $H = 40$. The number of lags is chosen to be 4. $\frac{\Gamma}{R}$ is calibrated to 0.98 in the empirical analysis. Results are robust to alternative calibrations of $\frac{\Gamma}{R}$.

[29] In contrast to Jordà *et al.* (2015) and Brunnermeier *et al.* (2020), who construct dates for the boom-bust of asset prices, our approach provides a time-series of "bubble shocks".

5.3 Results

Figure 7 plots the identified "bubble shocks". The shaded areas indicate the five recessions occurring in the sample. Each recession is associated with a big negative "bubble shock".

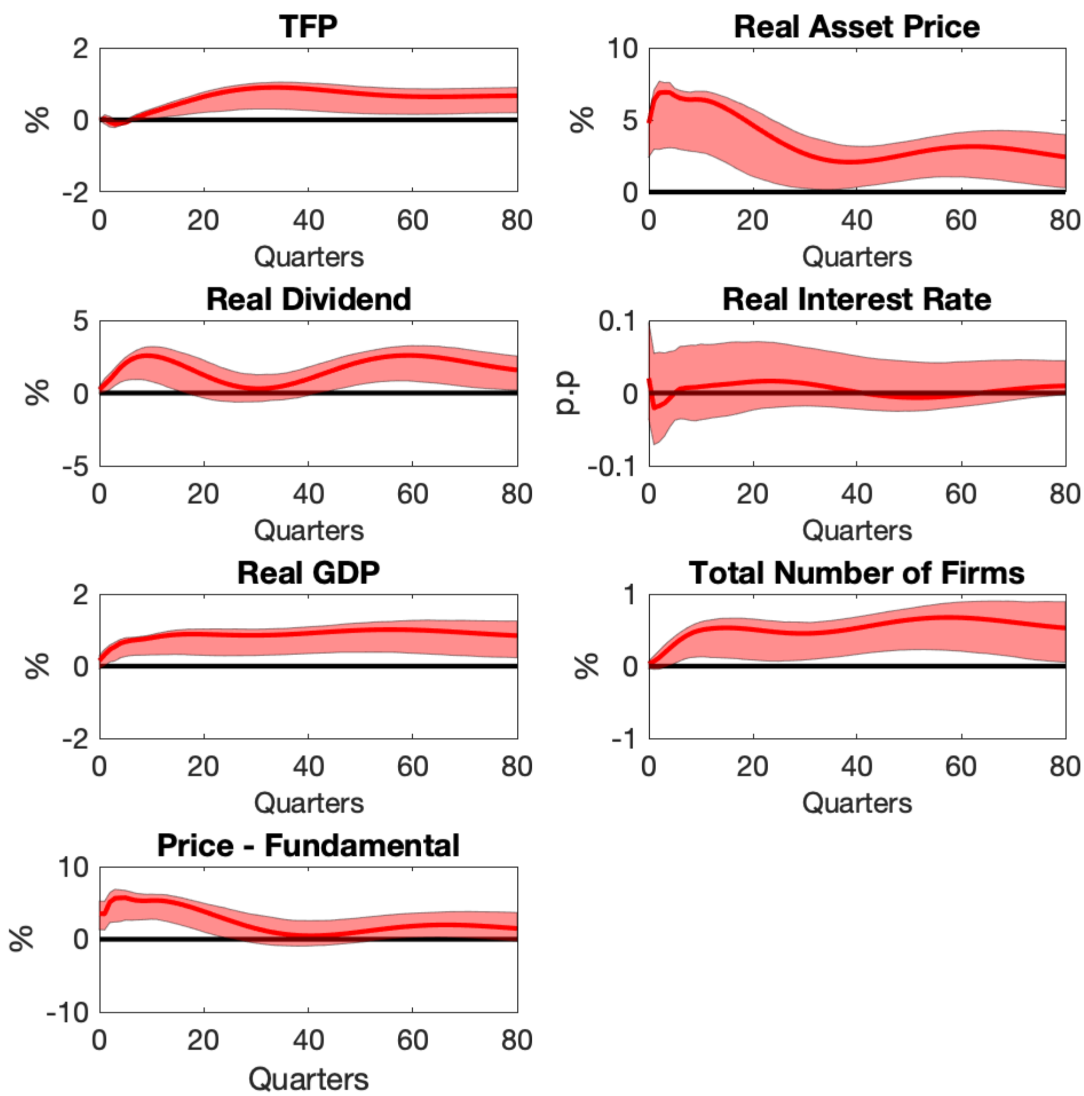

Figure 4: This figure plots the impulse response functions to a one standard deviation "bubble shock" estimated based on the baseline SVAR. The shaded areas indicate the 95% confidence bands.

Figure 4 reports the impacts of positive "bubble shocks" on the aggregate economy in our baseline. Following Kilian (1998), we construct standard errors from 2000 biascorrected bootstraps. The shaded areas indicate the 95% confidence bands. Two features should be highlighted from these IRFs. First, a positive "bubble shock" has expansionary effects: it has positive impacts on real asset price, real GDP, and the aggregate TFP. Second, while the "bubble shock" has peak effects on asset prices and its non-fundamental component on impact (within a year), its effects on real economic activities are slow-moving and hump-shaped. These results are consistent with the effects of bubble shocks in the model.

More importantly, the transmission of "bubble shocks" in the data along the extensive margin is relevant: a one standard deviation "bubble shock" increases the total number of firms by roughly 0.5%, persistently.

To better understand the effects of "bubble shocks" along the extensive margin, we estimate the responses of other measures of firm dynamics to "bubble shocks". Particularly, we include the number of new firms, the entry rate, the number exiting firms, and the exit rate into the empirical analysis. To avoid the curse of dimensionality, we include one variable of firm dynamics at a time. The red lines and the shaded area in Figure 5 show how the "bubble shocks" affect other measures of firm dynamics. Each row reports the results estimated from a separate SVAR model. The first row in Figure 5 reports the IRFs of the selected variables estimated from the baseline VAR considered earlier, and these results have been discussed before. They are repeated in Figure 5 to facilitate comparisons. The last four rows report the results estimated from separate VAR models by replacing the total number of firms by the number of new firms, the entry rate, the number of exiting firms, and the exit rate, respectively. The "bubble shocks" are normalized to have unit positive effects on real GDP at the horizon 10. The real effects of "bubble shocks" on real GDP revealed in the baseline are robust to alternative specifications (left column, Figure 5).

Figure 5 demonstrates that in a boom driven by our identified "bubble shocks", both the total number of firms and the number of new firms increase. While the number of new firms increases in the short run, the increase in the total number of firms is more persistent in the medium-run. Therefore, the entry rate increases in the short run and declines persistently in the medium-run. The exit rate decreases. The total number of exiting firms decreases in the short run, followed by a persistent increase in the medium-run. The last result is driven by the rise in the total number of firms: Holding the exit rate constant, an increase in the total number of firms leads to a higher number of exiting firms.

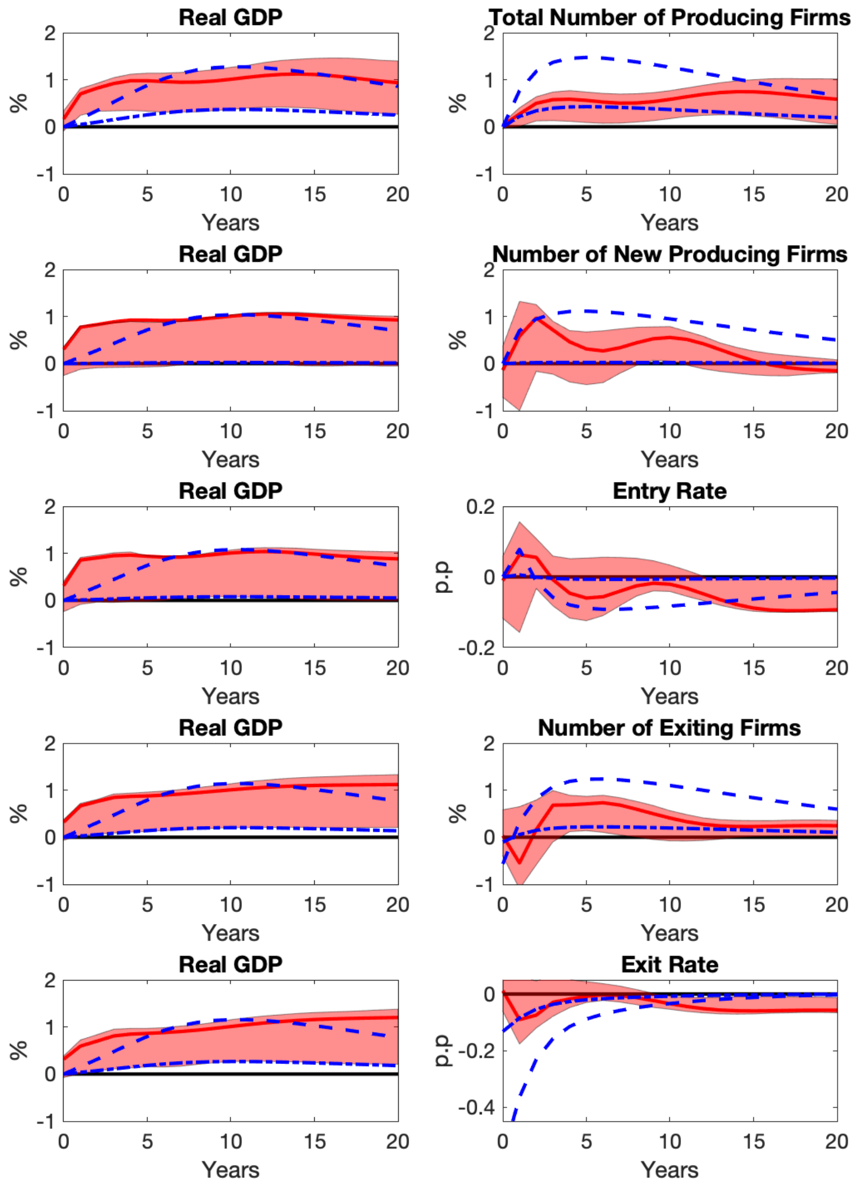

Figure 5: This figure plots the IRFs to positive bubble shocks in the model (blue lines) together with the estimated IRFs from the data (red lines). The shocks are normalized to have unit positive effects on real GDP at the horizon 10 for point estimates. The first row reports the IRFs estimated from the baseline VAR considered in Section 5.1. The last four rows report results estimated from separate SVARs, replacing the total number of firms by the number of new firms, the entry rate, the number of exiting firms, and the exit rate, respectively. The shaded areas indicate the 95% confidence bands.

Figure 5 plots the predictions of the model (blue lines) together with the empirical findings (red lines). We introduce bubble shocks in the model such that the response of the real GDP in the model matches those estimated in the data (left column, Figure 5). Then, we compare the IRFs of firm dynamics variables to a bubble shock in the model to the counterpart estimated in the data in response to the "bubble shock" (right column, Figure 5). The empirical IRFs are constructed with estimation uncertainty: confidence bounds. For comparison, we compute two sets of IRFs from the model. The blue dashed lines report the responses to a, relatively bigger, bubble shock that matches the upper-bound effects on real GDP revealed in the data. Similarly, the blue dash-dotted lines are constructed by shocking the model economy with a, relatively smaller, bubble shock that matches the lower-bound effects on real GDP revealed in the data.

The blue lines (model's predictions) in Figure 5 overlap with the red shaded area (data). Specifically, the IRFs of firm dynamics variables to a bubble shock in the model qualitatively and quantitatively match the IRFs to the identified shock in the data.

Simulation exercises conducted in Appendix E demonstrate that applying the medium-run restriction to the data simulated using the model uncovers the true IRFs in Section 4.2. Those IRFs are comparable to their empirical counterparts in Figure 4; however, conducting the same exercise using the data simulated from the model without bubbles fails to produce the empirical findings presented in Figure 4.

5.4 Robustness Checks

This section demonstrates that the baseline empirical findings are robust to alternative specifications that control for risk premium shocks, monetary policy shocks, and credit supply shocks, separately.

Risk Premium Shocks The main concern about the identification strategy in the baseline specification is that there might exist other fundamental shocks that contribute significantly to the fluctuations in the constructed price-fundamental differential component. One important candidate of the confounding shocks is the risk premium shock.

One might question the validity of the identification strategy basing on the argument that the risk premium shock is one of the main drivers of *asset price* fluctuations. This criticism is not entirely accurate. For the risk premium shock to be troublesome, it needs to have significant contribution to the variations in the *proxy of the price-fundamental differential component*. In the baseline specification, we do allow for the possibility that risk premium shocks *can* have significant impacts on asset prices. The implicit assumption made in the baseline analysis is that the response of the proxy of the fundamental component to risk premium shocks is similar to the asset prices' response so that the overall response of the constructed price-fundamental differential is small. Nevertheless, this might still be a strong assumption.

To address this concern, we conduct a robustness check controlling for risk premium shocks. Specifically, we control for shocks that have impact effects on the Volatility Index (VIX). Carr and Wu (2009) and Bekaert *et al.* (2013) document that the VIX contains a risk premium component that reflects risk aversion in financial markets. Therefore, controlling for the VIX mitigates the confounding problem with risk premium shocks described above. If omitting risk premium were indeed a problem, we would expect to see big changes in the estimated IRFs; however, Figure 8 shows that this is not the case. By controlling for risk premium shocks, the estimated IRFs, particularly the responses of the key variables that this article emphasizes—real GDP and the total number of firms—are similar to those obtained in the baseline.

Monetary Policy Shocks Omitting monetary policy shocks could also lead to the confounding problem. To address this, we construct "bubble shocks" controlling for shocks to the short-term interest rate. Figure 9 presents the estimation results: the IRFs of the key variables are similar to those obtained in the baseline model.

Other Financial Shocks: the Credit Supply Shocks Shocks to the credit market—the excess bond premium measure (credit supply shocks) proposed by Gilchrist and Zakrajsek 2012)—can have large effects on asset prices, and possibly, contribute to the fluctuations in the constructed price-fundamental differential component significantly. We construct the "bubble shocks" by controlling for shocks that have impact effects on the excess bond premium. Figure 10 shows that the main message of the paper remains unchanged. Finally, Figures 11 and 12 show that the empirical findings are robust to alternative calibrations of $\frac{\Gamma}{R}$.

6 Discussions and Future Research

Public Listing In the model, we assume that the value of any firm, including new firms, can have a bubble component. This assumption seemingly contradicts the common narrative that only publicly listed firms might be traded with bubbles in the real world. We offer two arguments that defend the assumption.

First, when a non-publicly-listed firm is to be sold, its value can contain a rational bubble component as long as the buyer believes that the bubble component would roll over and future buyers would pay for this bubble component.

Second, conditional on the existence of rational bubbles, the common narrative that *only* publicly-listed firms have the bubble component is inconsistent with rational expectations. Consider a slightly more complicated version of the model: firms can become publicly listed once their capital stocks reach a cutoff level, and the process of IPO incurs zero cost. It would be a mistake to assume that bubble components can only be obtained when firms become listed. The reason is the following: Even if a firm does not reach the cutoff level of capital to be listed, as long as it has a positive probability to be listed in the future, rational expectations ensure that a positive bubble component is attached to the value of this firm.[30] Therefore, rational expectations imply that both listed firms and non-listed firms contain bubble components.

Friction-less Markets One important distinction between our paper and other recent works of rational bubbles is our assumption of plain-vanilla competitive markets.[31] The vast literature of macro-finance highlights the importance of financial frictions in the propagation of financial shocks. Previous studies of rational bubbles typically feature financial frictions as well.[32] The literature argues that bubbles act as collateral or liquidity, and thus could relax financial constraints. One example of how this mechanism works is the following: When firms are subject to collateral constraint, bubbles can increase firms' collateral value and raise their borrowing capacity. Consequently, an increase in the aggregate bubble raises the aggregate output along the intensive margin.[33]

[30]Note that the reasoning above does not require zero cost of IPO. As long as the bubble component to be obtained upon IPO is higher than the cost of IPO, firms always choose to be listed once they reach the threshold. Therefore, positive bubble components always exist on the firms that will possibly become listed in the future.

[31]Although our model is abstract from market frictions, it does feature a real friction: capital adjustment cost. We incorporate capital adjustment cost to better account for firm life cycles. It is worth noting that bubbles affect firm entry and exit regardless of the existence of capital adjustment cost.

[32]For example, see Martin and Ventura (2012), Miao and Wang (2018), Ikeda and Phan (2019). See, also, Olivier (2000) and Queiros (2019) who model asset bubbles in frictional goods markets.

[33]In a model with financial frictions, it is indeed questionable whether bubble shocks can be disentangled from other financial shocks, like the shocks studied by Jermann and Quadrini (2012), and Gilchrist and Zakrajsek (2012), since a shock to bubbles is also a shock to financial constraints. One may thus interpret bubble shocks as a way to model financial shocks in a frictional environment.

The framework in the current study shows that, even in the absence of market frictions, investor optimism can give rise to an asset price boom and an aggregate output expansion. This result rests on the mechanism that bubble components directly affect the entry and exit decisions of firms. The mechanism is a natural outcome after relaxing the no-bubble condition in an otherwise standard firm dynamics model. In fact, bubbles in our model can affect the real economic activities only through their effects on firm entry and exit. Bubble components would have no real effects if one substitutes our assumptions of endogenous firm entry and exit by assuming inelastic entry and exit, i.e., firms' entry and exit are independent to the continuation value.

Results presented in Section 4.2 can be viewed as a friction-less benchmark for the true impulse responses of an aggregate bubble shock in the real world. To incorporate frictional financial intermediaries is beyond the scope of this paper.[34] As it is already uncovered in previous studies, in the presence of financial frictions, bubbles relax firms' financial constraints and thus, boost the real economic activities along the intensive margin. Therefore, financial frictions might further amplify the effects of bubbles on real economic activities.

Alternative Interpretation of Empirical Findings The empirical findings support the transmission channel of rational bubbles discussed in this study; however, our model may not be the only explanation of the data. Section 5.4 presents evidence that three well-known financial shocks—risk premium shocks, credit supply shocks, and shocks to the short-term interest rate—that we control for in the VAR do not drive the empirical findings. Nevertheless, we can not fully exclude the existence of other financial shocks that may also explain the findings. Further studies that address this potential issue of observational equivalence can enhance our understanding of the source and consequences of fluctuations in the asset prices.

Having said this, asset booms caused by rational bubbles differ from booms originating from other financial shocks in terms of policy implications. For example, Galí (2014) demonstrates that bubbles and fundamental components react differently to monetary policy shocks. Particularly, a positive monetary policy shock increases bubbles, but depresses fundamental components. Therefore, extending our framework to conduct monetary policy analysis might offer unique insights otherwise absent in models featur-

[34] In a companion paper, Tang and Zhang (2021), we develop a model of heterogeneous firms that have access to a frictional debt market, but no access to an equity market, and study the bubbles in the debt market.

ing asset booms of a different origin. The joint analysis of firm dynamics and monetary policy in a bubbly versus a bubble-less economy needs further investigation.

7 Conclusion

The present study examines the effects of bubbles in a heterogeneous firm model with endogenous entry and exit. The framework introduces a new transmission channel through which asset bubbles affect real economic activities. Bubbles have both micro- and macro-level implications. At the firm-level, this study shows the selection effect of bubbles: low-productivity firms enter/stay in the market due to the presence of bubbles. At the aggregate level, we emphasize the effects of bubbles along the extensive margin: an increase in the aggregate bubble enhances real economic activities by increasing the number of new firms and decreasing the exit rate.

Basing on a structural vector autoregressive model, we identify "bubble shocks" as the shocks that maximize the forecast error variance decomposition of the constructed non-fundamental component of the S&P 500, controlling for fundamental shocks such as the TFP, risk premium, monetary policy, and credit supply shocks. In the data, "bubble shocks" affect the real economic activities, and their effects along the extensive margin are comparable to the effects of bubble shocks in the model.

This study contributes to understanding the source of fluctuations in the asset prices and the consequences of changes in the bubble component. The novel transmission channel of asset bubbles proposed in this study has empirical support in the data. Policymakers should consider these effects while designing policies addressing asset bubbles.

References

ARELLANO, C., BAI, Y. and KEHOE, P. J. (2019). Financial Frictions and Fluctuations in Volatility. *Journal of Political Economy*, **127** (5), 2049–2103.

ASRIYAN, V., FORNARO, L., MARTÍN, A. and VENTURA, J. (2020). Monetary Policy for a Bubbly World. *Review of Economic Studies forthcoming*.

BACHMANN, R. and BAYER, C. (2014). Investment Dispersion and the Business Cycle. *American Economic Review*, **104** (4), 1392–1416.

—, CABALLERO, R. J. and ENGEL, E. M. R. A. (2013). Aggregate implications of lumpy investment: New evidence and a dsge model. *American Economic Journal: Macroeconomics*, **5** (4), 29–67.

BARLEVY, G. (2012). Rethinking theoretical models of bubbles. *New Perspectives on Asset Price Bubbles*, p. 41.

BARSKY, R. B. and SIMS, E. R. (2011). News Shocks and Business Cycles. *Journal of Monetary Economics*, **58** (3), 273–289.

BEKAERT, G., HOEROVA, M. and DUCA, M. L. (2013). Risk, Uncertainty and Monetary Policy. *Journal of Monetary Economics*, **60** (7), 771–788.

BEN ZEEV, N., PAPPA, E. and VICONDOA, A. (2017). Emerging economies business cycles: The role of commodity terms of trade news. *Journal of International Economics*, **108** (C), 368–376.

BENGUI, J. and PHAN, T. (2018). Asset pledgeability and endogenously leveraged bubbles. *Journal of Economic Theory*, **177** (C), 280–314.

BILBIIE, F. O., GHIRONI, F. and MELITZ, M. J. (2012). Endogenous Entry, Product Variety, and Business Cycles. *Journal of Political Economy*, **120** (2), 304–345.

— and MELITZ, M. J. (2020). *Aggregate-Demand Amplification of Supply Disruptions: The Entry-Exit Multiplier*. Tech. rep., National Bureau of Economic Research.

BISWAS, S., HANSON, A. and PHAN, T. (2020). Bubbly Recessions. *Forthcoming American Economic Journal: Macroeconomics*.

BLANCHARD, O. J. and QUAH, D. (1989). The Dynamic Effects of Aggregate Demand and Supply Disturbances. *The American Economic Review*, **79** (4), 655–673.

— and WATSON, M. W. (1982). Bubbles, Rational Expectations and Financial Markets .

BLOOM, N., FLOETOTTO, M., JAIMOVICH, N., EKSTEN, I. S. and TERRY, S. J. (2018). Really Uncertain Business Cycles. *Econometrica*, **86** (3), 1031–1065.

BOPPART, T., MITMAN, K. and KRUSELL, P. (2018). Exploiting mit shocks in heterogeneous-agent economies: The impulse response as a numerical derivative. *Journal of Economic Dynamics and Control*, **89**, 68–92.

BRUNNERMEIER, M., ROTHER, S. and SCHNABEL, I. (2020). Asset Price Bubbles and Systemic Risk. *The Review of Financial Studies*, **33** (9), 4272–4317.

CABALLERO, R. J. and ENGEL, E. M. (1999). Explaining investment dynamics in us manufacturing: a generalized (s, s) approach. *Econometrica*, **67** (4), 783–826.

— and KRISHNAMURTHY, A. (2006). Bubbles and Capital Flow Volatility: Causes and Risk Management. *Journal of Monetary Economics*, **53** (1), 35–53.

CAMPBELL, J. Y. and SHILLER, R. J. (1988). Stock Prices, Earnings, and Expected Dividends. *The Journal of Finance*, **43** (3), 661–676.

CARR, P. and WU, L. (2009). Variance Risk Premiums. *The Review of Financial Studies*, **22** (3), 1311–1341.

CLEMENTI, G. L. and PALAZZO, B. (2016). Entry, Exit, Firm Dynamics, and Aggregate Fluctuations. *American Economic Journal: Macroeconomics*, **8** (3), 1–41.

DOMEIJ, D. and ELLINGSEN, T. (2018). Rational bubbles and public debt policy: A quantitative analysis. *Journal of Monetary Economics*, **96**, 109–123.

— and — (2021). Bubbles and Kinks: A Monetary Model of Recessions.

FARHI, E. and TIROLE, J. (2012). Bubbly Liquidity. *Review of Economic Studies*, **79** (2), 678–706.

FAVARA, G., GILCHRIST, S., LEWIS, K. F. and ZAKRAJSEK, E. (2016). *Updating the Recession Risk and the Excess Bond Premium*. FEDS Notes 2016-10-06, Board of Governors of the Federal Reserve System (U.S.).

FERNALD, J. (2014). A quarterly, Utilization-adjusted Series on Total Factor Productivity. Federal Reserve Bank of San Francisco.

GALÍ, J. (1999). Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations? *American economic review*, **89** (1), 249–271.

— (2014). Monetary Policy and Rational Asset Price Bubbles. *The American Economic Review*, **104** (3), 721–752.

— (2020). Monetary Policy and Bubbles in a New Keynesian Model with Overlapping Generations. *Forthcoming American Economic Journal: Macroeconomics*.

— and GAMBETTI, L. (2015). The Effects of Monetary Policy on Stock Market Bubbles: Some Evidence. *American Economic Journal: Macroeconomics*, **7** (1), 233–57.

GARÍN, J., LESTER, R. and SIMS, E. (2019). Are Supply Shocks Contractionary at the ZLB? Evidence from Utilization-adjusted TFP Data. *Review of Economics and Statistics*, **101** (1), 160–175.

GILCHRIST, S. and ZAKRAJSEK, E. (2012). Credit Spreads and Business Cycle Fluctuations. *American Economic Review*, **102** (4), 1692–1720.

GUERRÓN-QUINTANA, P., HIRANO, T. and JINNAI, R. (2020). Recurrent bubbles and economic growth. *Available at SSRN 3350097*.

HELLWIG, C. and LORENZONI, G. (2009). Bubbles and Self-enforcing Debt. *Econometrica*, **77** (4), 1137–1164.

HIRANO, T. and YANAGAWA, N. (2017). Asset Bubbles, Endogenous Growth, and Financial Frictions. *The Review of Economic Studies*, **84** (1), 406–443.

HOPENHAYN, H. A. (1992). Entry, Exit, and Firm Dynamics in Long Run Equilibrium. *Econometrica*, **60** (5), 1127–1150.

IKEDA, D. and PHAN, T. (2019). Asset Bubbles and Global Imbalances. *American Economic Journal: Macroeconomics*, **11** (3), 209–251.

JAIMOVICH, N. and FLOETOTTO, M. (2008). Firm dynamics, markup variations, and the business cycle. *Journal of Monetary Economics*, **55** (7), 1238–1252.

JERMANN, U. J. and QUADRINI, V. (2007). Stock Market Boom and the Productivity Gains of the 1990s. *Journal of Monetary Economics*, **54** (2), 413–432.

— and — (2012). Macroeconomic Effects of Financial Shocks. *American Economic Review*, **102** (1), 238–271.

JORDÀ, Ò., SCHULARICK, M. and TAYLOR, A. M. (2015). Leveraged Bubbles. *Journal of Monetary Economics*, **76**, S1–S20.

KHAN, A. and THOMAS, J. K. (2008). Idiosyncratic Shocks and the Role of Nonconvexities in Plant and Aggregate Investment Dynamics. *Econometrica*, **76** (2), 395–436.

— and — (2013). Credit Shocks and Aggregate Fluctuations in an Economy with Production Heterogeneity. *Journal of Political Economy*, **121** (6), 1055–1107.

KILIAN, L. (1998). Small-Sample Confidence Intervals For Impulse Response Functions. *The Review of Economics and Statistics*, **80** (2), 218–230.

KOPECKY, K. A. and SUEN, R. M. H. (2010). Finite State Markov-Chain Approximations to Highly Persistent Processes. *Review of Economic Dynamics*, **13**, 701–714.

LARIN, B. (2019). *Bubble-Driven Business Cycles*. Tech. rep.

LEE, Y. and MUKOYAMA, T. (2015). Entry and exit of manufacturing plants over the business cycle. *European Economic Review*, **77** (C), 20–27.

LEVCHENKO, A. A. and PANDALAI-NAYAR, N. (2020). TFP, News, and "Sentiments": The International Transmission of Business Cycles. *Journal of the European Economic Association*, **18** (1), 302–341.

LORIA, F., MATTHES, C. and ZHANG, D. (2019). Assessing Macroeconomic Tail Risk.

MARTIN, A. and VENTURA, J. (2012). Economic Growth with Bubbles. *American Economic Review*, **102** (6), 3033–58.

— and — (2016). Managing Credit Bubbles. *Journal of the European Economic Association*, **14** (3), 753–789.

— and — (2018). The Macroeconomics of Rational Bubbles: A User's Guide. *Annual Review of Economics*, **10** (1), 505–539.

MIAO, J., SHEN, Z. and WANG, P. (2019). Monetary policy and rational asset bubbles: Comments. *American Economic Review*, **109**, 1969–1990.

— and WANG, P. (2018). Asset Bubbles and Credit Constraints. *American Economic Review*, **108** (9), 2590–2628.

—, — and XU, Z. (2015). A Bayesian Dynamic Stochastic General Equilibrium Model of Stock Market Bubbles and Business Cycles. *Quantitative Economics*, **6** (3), 599–635.

OLIVIER, J. (2000). Growth-Enhancing Bubbles. *International Economic Review*, **41** (1), 133–151.

OTTONELLO, P. and WINBERRY, T. (2018). *Financial Heterogeneity and the Investment Channel of Monetary Policy*. NBER Working Papers 24221, National Bureau of Economic Research, Inc.

QUEIROS, F. (2019). *Asset Bubbles and Product Market Competition*. 2018 Meeting Papers 462, Society for Economic Dynamics.

RAMEY, V. A. (2016). Macroeconomic shocks and their Propagation. In *Handbook of macroeconomics*, vol. 2, Elsevier, pp. 71–162.

SEDLÁČEK, P. and STERK, V. (2017). The Growth Potential of Startups over the Business Cycle. *American Economic Review*, **107** (10), 3182–3210.

SENGA, T., THOMAS, J. and KHAN, A. (2017). *Default Risk and Aggregate Fluctuations in an Economy with Production Heterogeneity*. 2017 Meeting Papers 889, Society for Economic Dynamics.

SHILLER, R. J. (2015). *Irrational Exuberance: Revised and Expanded Third Rdition*. Princeton university press.

TANG, H. and ZHANG, D. (2021). Credit Bubble Cycles .

THOMAS, J. (2002). Is lumpy investment relevant for the business cycle? *Journal of Political Economy*, **110** (3), 508–534.

TIROLE, J. (1985). Asset Bubbles and Overlapping Generations. *Econometrica*, **53** (6), 1499–1528.

UHLIG, H. (2003). What Drives GNP? *Unpublished manuscript, Euro Area Business Cycle Network*.

— (2004). Do technology Shocks Lead to a Fall in Total Hours Worked? *Journal of the European Economic Association*, **2** (2-3), 361–371.

VENTURA, J. (2012). Bubbles and Capital Flows. *Journal of Economic Theory*, **147** (2), 738–758.

VUILLEMEY, G. and WASMER, E. (2020). Frictional Unemployment with Stochastic Bubbles. *European Economic Review*, **122**, 103352.

WINBERRY, T. (2020). *Lumpy Investment, Business Cycles, and Stimulus Policy*. Tech. rep.

ZEEV, N. B. and PAPPA, E. (2017). Chronicle of a War Foretold: The Macroeconomic Effects of Anticipated Defence Spending Shocks. *Economic Journal*, **127** (603), 1568–1597.

ZWICK, E. and MAHON, J. (2017). Tax policy and heterogeneous investment behavior. *American Economic Review*, **107** (1), 217–48.

İMROHOROĞLU, A. and TÜZEL, S. (2014). Firm-Level Productivity, Risk, and Return. *Management Science*, **60** (8), 2073–2090.

A Figures

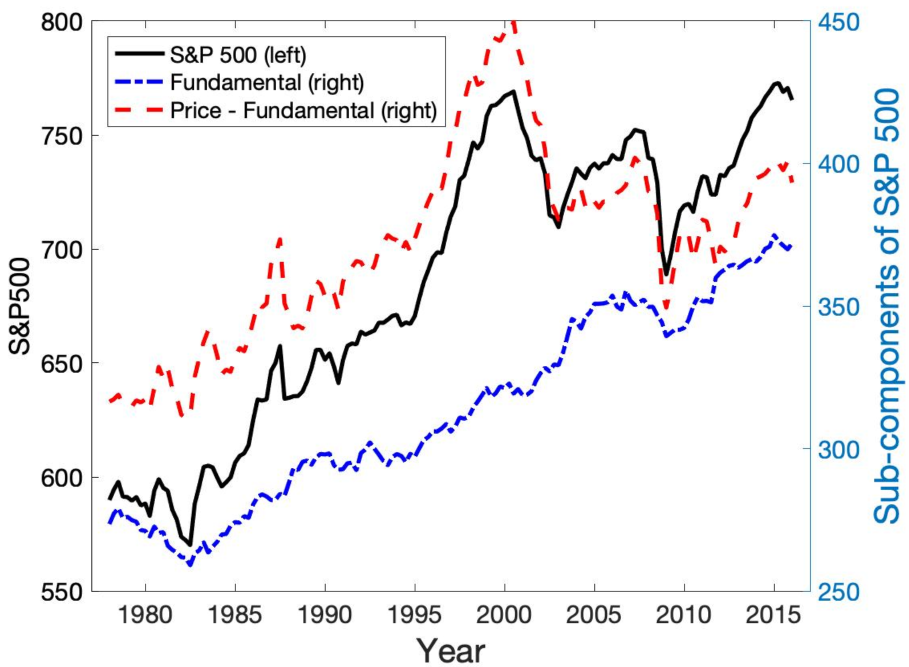

Figure 6: This figure plots the evolution of the real value (in log) of Standard and Poors (S&P) 500 Index (solid black line with the y-axis on the left-hand side) together with its decomposition (with y-axis on the right-hand side): the fundamental component (dash-dot blue line) and the price-fundamental differential (dash red line). The fundamental component is constructed as the present value of future real dividends, where expected future dividends are constructed in real-time using forecasts based on a vector autoregressive model. Authors' own calculation.

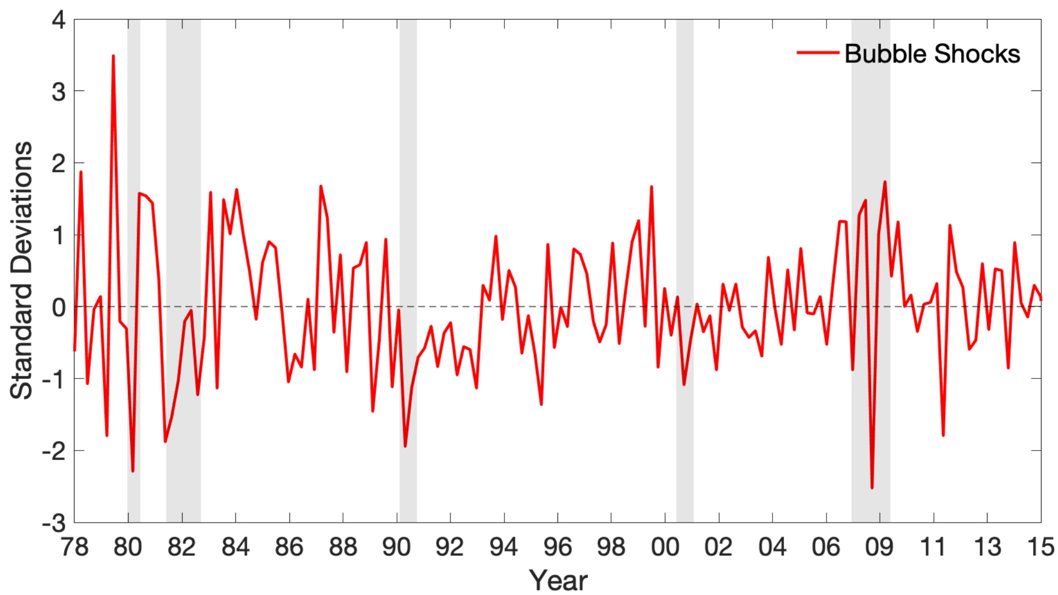

Figure 7: This figure plots the identified "bubble shocks". The shaded area indicates the five recessions occurred in the sample.

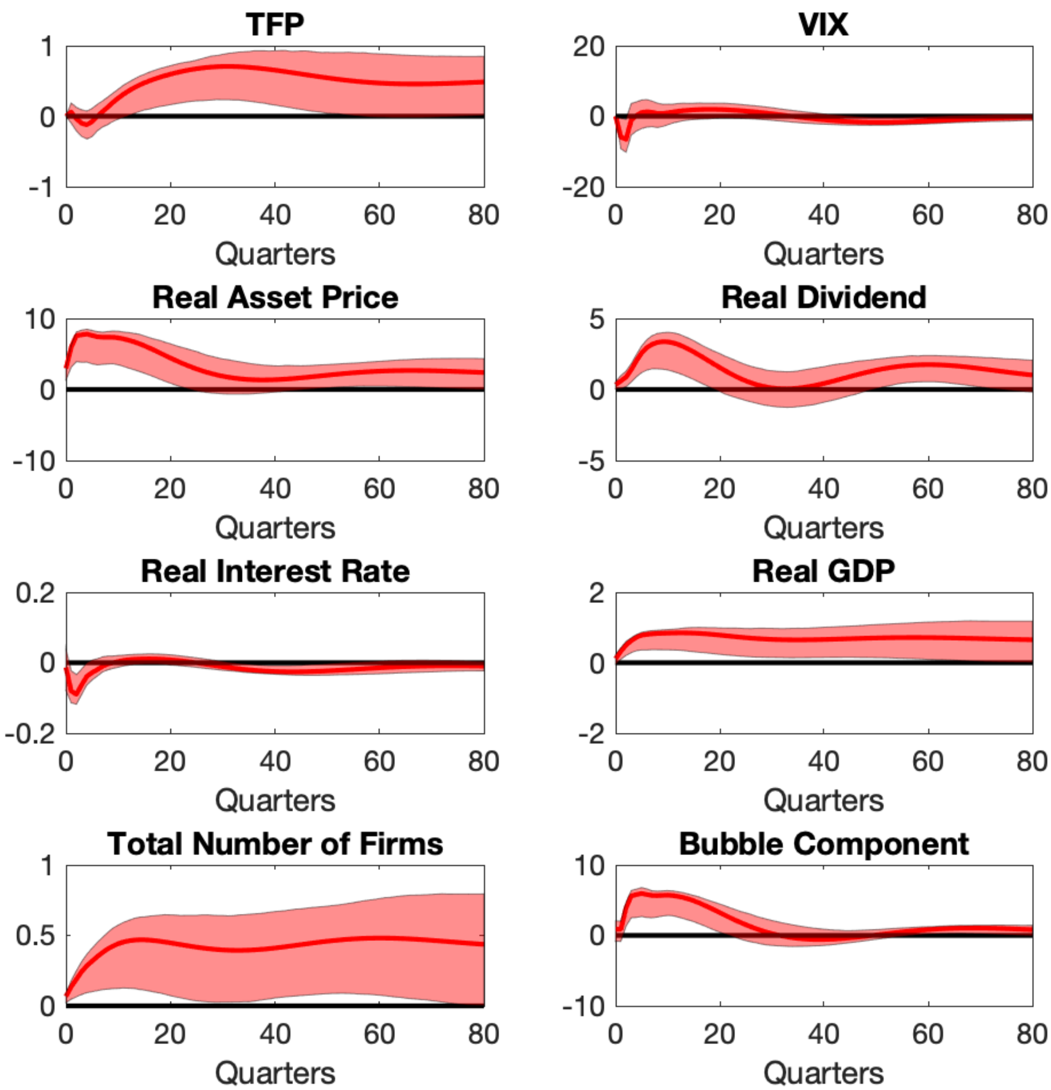

Figure 8: Robustness check I: controlling for risk premium shocks. This figure plots the impulse response functions to a one standard deviation "bubble shock" estimated based on the a SVAR that controls for TFP shocks and risk premium shocks. The shaded areas indicate the 95% confidence bands.

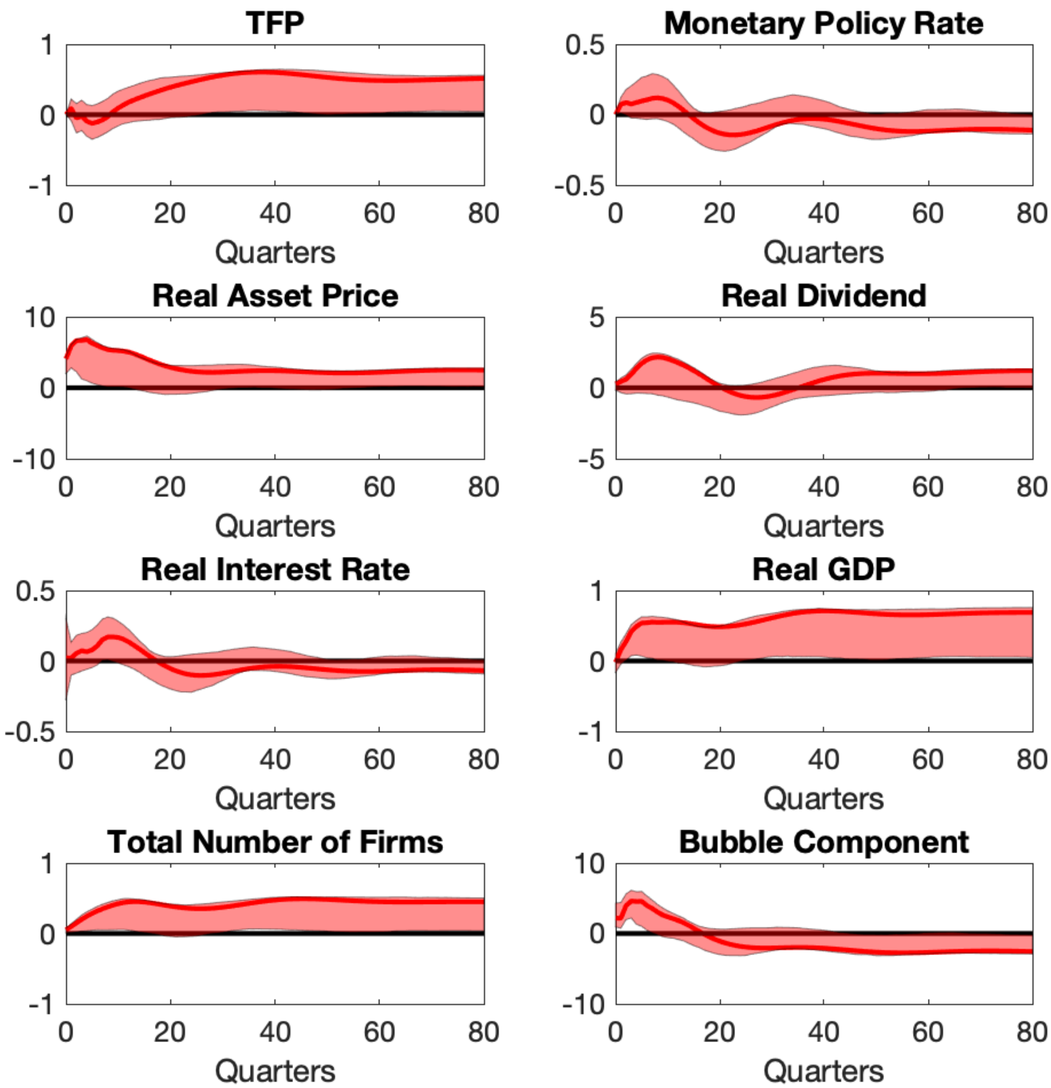

Figure 9: Robustness check II: controlling for shocks to the short-term interest rate. This figure plots the impulse response functions to a one standard deviation "bubble shock" estimated based on the a SVAR that controls for TFP shocks and shocks to the short-term interest rate. The shaded areas indicate the 95% confidence bands.

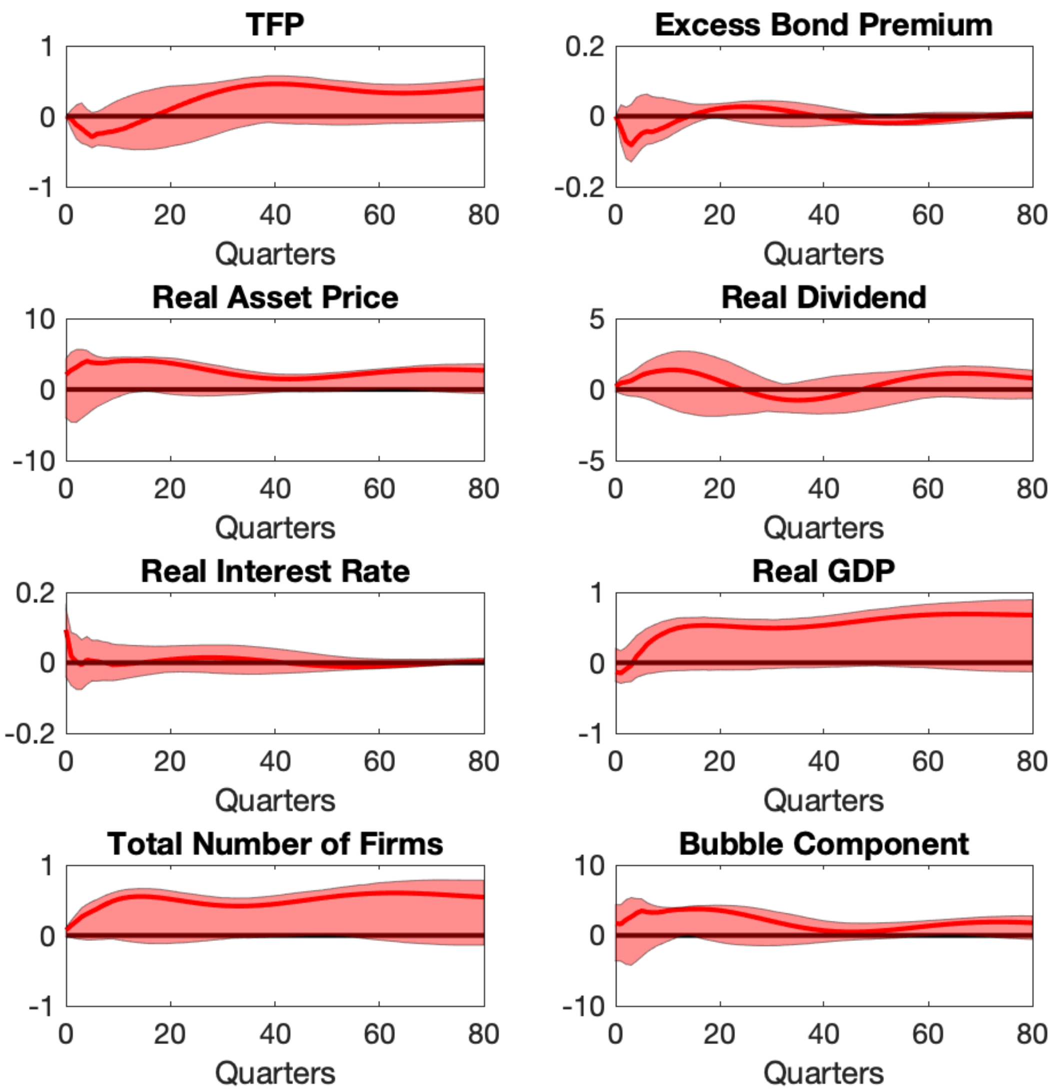

Figure 10: Robustness check III: controlling for credit supply shocks. This figure plots the impulse response functions to a one standard deviation "bubble shock" estimated based on the a SVAR that controls for TFP shocks and credit supply shocks. The shaded areas indicate the 95% confidence bands.

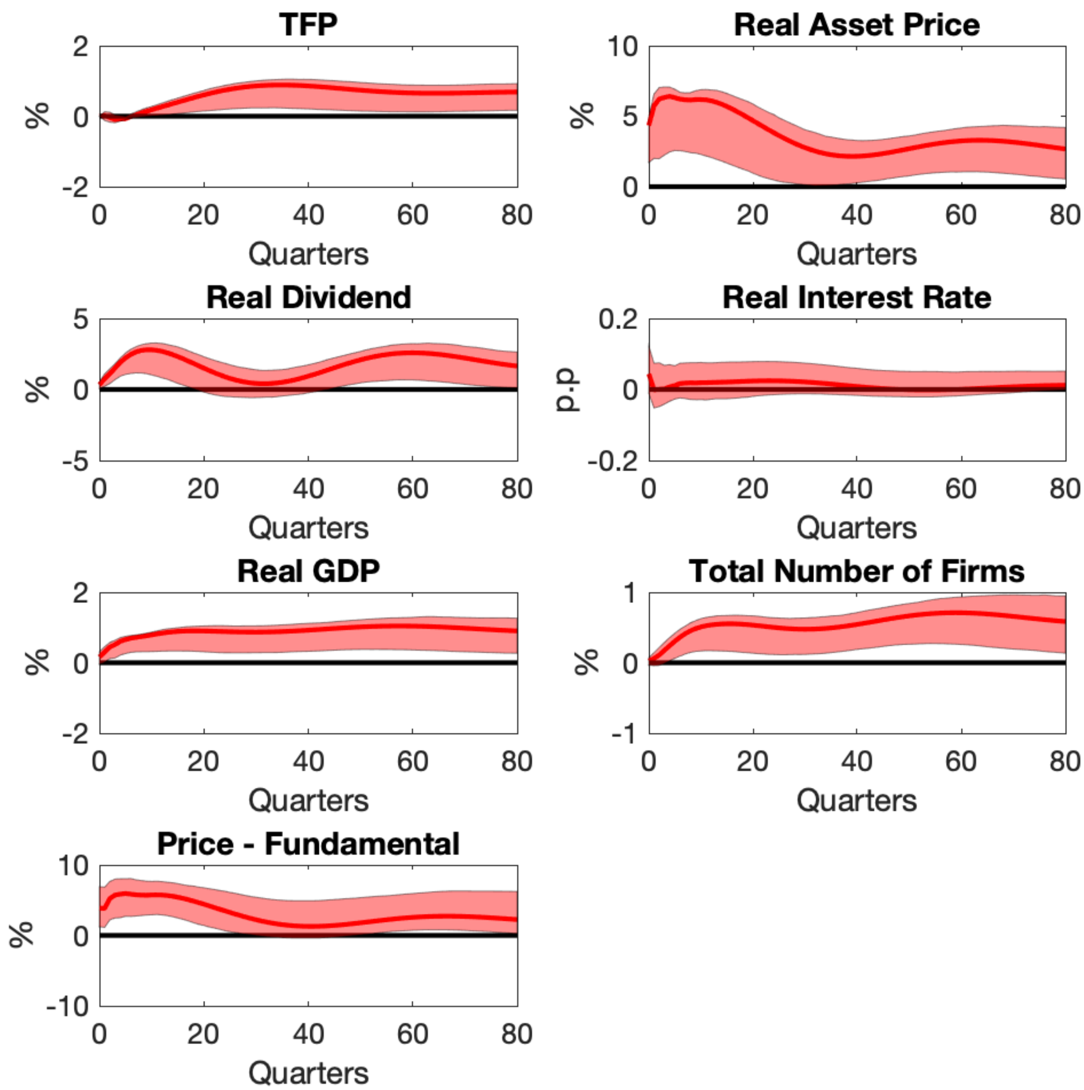

Figure 11: Robustness Check IV: using alternative calibration of $\frac{\Gamma}{R} = 0.99$. This figure plots the impulse response functions to a one standard deviation "bubble shock" estimated based on the SVAR. The shaded areas indicate the 95% confidence bands.

Figure 12: Robustness Check V: using alternative calibration of $\frac{\Gamma}{R} = 0.75$. This figure plots the impulse response functions to a one standard deviation "bubble shock" estimated based on the SVAR. The shaded areas indicate the 95% confidence bands.

B Appendix for the Model

B.1 The Decomposition of Firms' Values

Equations (14) and (17) can be further rewritten into

$$V^c(\lambda,\mu,k) = (1-\delta)k - k^{(1)} - g\left(k,k^{(1)}\right) + \frac{1}{R}\int V\left(\lambda^{(1)},\mu^{(1)},k^{(1)}\right)dJ\left(\lambda^{(1)},\mu^{(1)}|\lambda,\mu\right), \tag{31}$$

$$V(\lambda,\mu,k) = A\varphi k^{\alpha} - c^f + 1\{V^c(\lambda,\mu,k) < V^x(k)\}V^x(k) + 1\{V^c(\lambda,\mu,k) \geq V^x(k)\}V^c(\lambda,\mu,k), \tag{32}$$

where a variable $X^{(i)}$ denotes the variable X after i periods, specifically, $k^{(1)} = k^*(\lambda,\mu,k)$, $k^{(i)} = k^*\left(\lambda^{(i-1)}\mu^{(i-1)},k^{(i-1)}\right)$ for all $i > 1$, k^* denotes the policy function for k.

Moreover, we can expand equation (31) into

$$\begin{aligned} V^c(\lambda,\mu,k) = {} & (1-\delta)k - k^{(1)} - g\left(k,k^{(1)}\right) + \frac{1}{R}\int\left[A^{(1)}\varphi^{(1)}(k^{(1)})^{\alpha} - c^f\right]dJ\left(\lambda^{(1)},\mu^{(1)}|\lambda,\mu\right) \\ & + \frac{1}{R}\int_{\Psi^{(1)}} V^x\left(k^{(1)}\right)dJ\left(\lambda^{(1)},\mu^{(1)}|\lambda,\mu\right) + \frac{1}{R}\int_{\Phi^{(1)}} V^c\left(\lambda^{(1)},\mu^{(1)},k^{(1)}\right)dJ\left(\lambda^{(1)},\mu^{(1)}|\lambda,\mu\right). \end{aligned} \tag{33}$$

where $\Psi^{(i)} \equiv \left\{\left(\lambda^{(i)},\mu^{(i)}\right)|V^c\left(\lambda^{(i)},\mu^{(i)},k^{(i)}\right) < V^x\left(k^{(i)}\right)\right\}$,
and $\Phi^{(i)} \equiv \left\{\left(\lambda^{(i)},\mu^{(i)}\right)|V^c\left(\lambda^{(i)},\mu^{(i)},k^{(i)}\right) \geq V^x\left(k^{(i)}\right)\right\}$.

Define cash inflow $\pi(\lambda,\mu,k)$ as:

$$\begin{aligned} \pi(\lambda,\mu,k) \equiv {} & (1-\delta)k - k^{(1)} - g\left(k,k^{(1)}\right) + \frac{1}{R}\int\left[A^{(1)}\varphi^{(1)}(k^{(1)})^{\alpha} - c^f\right]dJ\left(\lambda^{(1)},\mu^{(1)}|\lambda,\mu\right) \\ & + \frac{1}{R}\int_{\Psi^{(1)}} V^x\left(k^{(1)}\right)dJ\left(\lambda^{(1)},\mu^{(1)}|\lambda,\mu\right). \end{aligned}$$

By repeatedly expanding the last term on the right-hand side of (33), the continuation value $V^c(\lambda,\mu,k)$ can be expressed as

$$V^c(\lambda,\mu,k) = F^c(\lambda,\mu,k) + V^c_{\infty}(\lambda,\mu,k), \tag{34}$$

where

$$
\begin{aligned}
F^{c}(\lambda,\mu,k) \equiv\ & \pi(\lambda,\mu,k) \\
& + \frac{1}{R} \int\limits_{\Phi^{(1)}} \pi\left(\lambda^{(1)},\mu^{(1)},k^{(1)}\right) dJ\left(\lambda^{(1)},\mu^{(1)}|\lambda,\mu\right) \\
& + \frac{1}{R}\frac{1}{R^{(1)}} \int\limits_{\Phi^{(1)}} \int\limits_{\Phi^{(2)}} \pi\left(\lambda^{(2)},\mu^{(2)},k^{(2)}\right) dJ\left(\lambda^{(2)},\mu^{(2)}|\lambda^{(1)},\mu^{(1)}\right) dJ\left(\lambda^{(1)},\mu^{(1)}|\lambda,\mu\right) \\
& \vdots \\
& + \lim_{i\to\infty}\left(\prod_{j=0}^{i-1}\frac{1}{R^{(j)}}\right) \int\limits_{\Phi^{(1)}} \cdots \int\limits_{\Phi^{(i)}} \pi\left(\lambda^{(i)},\mu^{(i)},k^{(i)}\right) dJ\left(\lambda^{(i)},\mu^{(i)}|\lambda^{(i-1)},\mu^{(i-1)}\right) \\
& \cdots dJ\left(\lambda^{(1)},\mu^{(1)}|\lambda,\mu\right),
\end{aligned}
\tag{35}
$$

$$
\begin{aligned}
V^{c}_{\infty}(\lambda,\mu,k) \equiv\ & \lim_{i\to\infty}\left(\prod_{j=0}^{i-1}\frac{1}{R^{(j)}}\right) \int\limits_{\Phi^{(1)}} \cdots \int\limits_{\Phi^{(i)}} V^{c}\left(\lambda^{(i)},\mu^{(i)},k^{(i)}\right) dJ\left(\lambda^{(i)},\mu^{(i)}|\lambda^{(i-1)},\mu^{(i-1)}\right) \\
& \cdots dJ\left(\lambda^{(1)},\mu^{(1)}|\lambda,\mu\right),
\end{aligned}
\tag{36}
$$

in which $R^{(0)} = R$.

B.2 The Bubble Component

We need to show:

$$
\begin{aligned}
& F^{c}(\lambda,\mu,k) + B = F^{c}(\lambda,\mu,k) \\
& + \lim_{i\to\infty}\left(\prod_{j=0}^{i-1}\frac{1}{R^{(j)}}\right) \int\limits_{\Phi^{(1)}} \cdots \int\limits_{\Phi^{(i)}} [F^{c}\left(\lambda^{(i)},\mu^{(i)},k^{(i)}\right) + B^{(i)}] dJ\left(\lambda^{(i)},\mu^{(i)}|\lambda^{(i-1)},\mu^{(i-1)}\right) \\
& \cdots dJ\left(\lambda^{(1)},\mu^{(1)}|\lambda,\mu\right).
\end{aligned}
$$

Note that (20) can be expanded into:

$$
\begin{aligned}
B &= \frac{1}{R}\int_{\Phi^{(1)}} B^{(1)} dJ\left(\left(\lambda^{(1)},\mu^{(1)}|\lambda,\mu\right)\right) \\
&= \frac{1}{R}\frac{1}{R^{(1)}}\int\limits_{\Phi^{(1)}}\int\limits_{\Phi^{(2)}} B^{(2)} dJ\left(\lambda^{(2)},\mu^{(2)}|\lambda^{(1)},\mu^{(1)}\right) dJ\left(\lambda^{(1)},\mu^{(1)}|\lambda,\mu\right) = \cdots \\
&= \lim_{i\to\infty}\left(\prod_{j=0}^{i-1}\frac{1}{R^{(j)}}\right)\int\limits_{\Phi^{(1)}}\cdots\int\limits_{\Phi^{(i)}} B^{(i)} dJ\left(\lambda^{(i)},\mu^{(i)}|\lambda^{(i-1)},\mu^{(i-1)}\right)\cdots dJ\left(\lambda^{(1)},\mu^{(1)}|\lambda,\mu\right).
\end{aligned}
\tag{37}
$$

What remains to be proven is:

$$
\lim_{i\to\infty}\left(\prod_{j=0}^{i-1}\frac{1}{R^{(j)}}\right)\int\limits_{\Phi^{(1)}}\cdots\int\limits_{\Phi^{(i)}} F^{c}\left(\lambda^{(i)},\mu^{(i)},k^{(i)}\right) dJ\left(\lambda^{(i)},\mu^{(i)}|\lambda^{(i-1)},\mu^{(i-1)}\right)\cdots dJ\left(\lambda^{(1)},\mu^{(1)}|\lambda,\mu\right) = 0.
\tag{38}
$$

According to (35), fundamental component can be written into:

$$
\begin{aligned}
&F^{c}(\lambda,\mu,k) = \pi(\lambda,\mu,k) + \frac{1}{R}\int\limits_{\Phi^{(1)}} \pi\left(\lambda^{(1)},\mu^{(1)},k^{(1)}\right) dJ\left(\lambda^{(1)},\mu^{(1)}|\lambda,\mu\right) \\
&+ \frac{1}{R}\frac{1}{R^{(1)}}\int\limits_{\Phi^{(1)}}\int\limits_{\Phi^{(2)}} \pi\left(\lambda^{(2)},\mu^{(2)},k^{(2)}\right) dJ\left(\lambda^{(2)},\mu^{(2)}|\lambda^{(1)},\mu^{(1)}\right) dJ\left(\lambda^{(1)},\mu^{(1)}|\lambda,\mu\right) + \cdots \\
&+ \lim_{i\to\infty}\left(\prod_{j=0}^{i-1}\frac{1}{R^{(j)}}\right)\int\limits_{\Phi^{(1)}}\cdots\int\limits_{\Phi^{(i)}} \pi\left(\lambda^{(i)},\mu^{(i)},k^{(i)}\right) dJ\left(\lambda^{(i)},\mu^{(i)}|\lambda^{(i-1)},\mu^{(i-1)}\right)\cdots dJ\left(\lambda^{(1)},\mu^{(1)}|\lambda,\mu\right) \\
&+ \lim_{i\to\infty}\left(\prod_{j=0}^{i-1}\frac{1}{R^{(j)}}\right)\int\limits_{\Phi^{(1)}}\cdots\int\limits_{\Phi^{(i)}} F^{c}\left(\lambda^{(i)},\mu^{(i)},k^{(i)}\right) dJ\left(\lambda^{(i)},\mu^{(i)}|\lambda^{(i-1)},\mu^{(i-1)}\right)\cdots dJ\left(\lambda^{(1)},\mu^{(1)}|\lambda,\mu\right).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
F^{c}(\lambda,\mu,k) = F^{c}(\lambda,\mu,k) \\
&+ \lim_{i\to\infty}\left(\prod_{j=0}^{i-1}\frac{1}{R^{(j)}}\right)\int\limits_{\Phi^{(1)}}\cdots\int\limits_{\Phi^{(i)}} F^{c}\left(\lambda^{(i)},\mu^{(i)},k^{(i)}\right) dJ\left(\lambda^{(i)},\mu^{(i)}|\lambda^{(i-1)},\mu^{(i-1)}\right) \\
&\cdots dJ\left(\lambda^{(1)},\mu^{(1)}|\lambda,\mu\right).
\end{aligned}
$$

It is now straightforward to see that (38) holds. Q.E.D.

B.3 Definition of the Equilibrium

A recursive competitive bubbly equilibrium consists of:[35] i) value functions $V(\lambda_t, \varphi_t, B_t, k_t)$, $V^c(\lambda_t, \varphi_t, B_t, k_t)$, $V^e(\lambda_t, \varphi_t, B_t)$, ii) policy function $k^*(\lambda_t, \varphi_t, B_t, k_t)$, iii) the measure of entrants $\chi_t(\varphi_t, B_t)$, the measure of potential entrants $\omega(\varphi_t, B_t)$, the measure of surviving firms $\eta_t^s(\varphi_t, B_t, k_t)$, the measure of producing firms $\eta_t^p(\varphi_t, B_t, k_t)$, iv) the process of B_t, and v) the investment return R_t, such that for all t,

1. $V(\lambda_t, \varphi_t, k_t, B_t)$ and $V^c(\lambda_t, \varphi_t, k_t, B_t)$ solve the optimization problem of incumbent firms, which is described by

$$V(\lambda_t, \varphi_t, B_t, k_t) = A_t \varphi_t k_t^\alpha - c^f + \max\{V^c(\lambda_t, \varphi_t, B_t, k_t), V^x(k_t)\},$$

$$V^c(\lambda_t, \varphi_t, B_t, k_t) =$$
$$\max_{k_{t+1}} \left\{ (1-\delta) k_t - k_{t+1} - g(k_t, k_{t+1}) + \frac{1}{R_t} \int V(\lambda_{t+1}, \varphi_{t+1}, B_{t+1}, k_{t+1})\, dJ(\lambda_{t+1}, \varphi_{t+1}, B_{t+1} | \lambda_t, \varphi_t, B_t) \right\},$$

where

$$g(k_t, k_{t+1}) = c_0 1\{k_{t+1} \neq (1-\delta) k_t\} k + c_1 \left(\frac{k_{t+1} - (1-\delta) k_t}{k_t} \right)^2 k_t,$$
$$V^x(k_t) = (1-\delta) k_t - g(k_t, 0),$$

$k^*(\lambda_t, \varphi_t, B_t, k_t)$ is the associated policy function for optimal k_{t+1}.

2. The value of firm entry is given by

$$V^e(\lambda_t, \varphi_t, B_t) = V^c(\lambda_t, \varphi_t, B_t, 0).$$

3. The process of bubble satisfies

$$B_t = \frac{1}{R_t} \int_{\Phi_{t+1}} B_{t+1} dJ((\lambda_{t+1}, \varphi_{t+1}, B_{t+1} | \lambda_t, \varphi_t, B_t)),$$

where $\Phi_{t+1} \equiv \{(\lambda_{t+1}, \varphi_{t+1}, B_{t+1}) | V^c(\lambda_{t+1}, \varphi_{t+1}, B_{t+1} k_{t+1}) \geq V^x(k_{t+1})\}$, $k_{t+1} = k^*(\lambda_t, \varphi_t, B_t, k_t)$.

4. The measure of entrants is

$$\chi_t(\varphi_t, B_t) = N_t \int \mathbb{1}\{V^e(\lambda_t, \varphi_t, B_t) \geq 0\}\, d\omega(\varphi_t, B_t).$$

[35] In this Appendix we modify the state space of functions by replacing μ with φ and B. For some functions we keep only the relevant states. We also add time subscripts.

5. The measure of surviving firms follows

$$\begin{aligned}\eta_{t+1}^{s}\left(\varphi_{t+1},B_{t+1},k_{t+1}\right) = &\int \mathbb{1}\left\{V^{c}\left(\lambda_{t+1},\varphi_{t+1},B_{t+1},k_{t+1}\right) \geq V^{x}\left(k_{t+1}\right)\right\} \times \mathbb{1}\left\{k_{t+1} = k^{*}\left(\lambda_{t},\varphi_{t},B_{t},k_{t}\right)\right\} \\ &\times j\left(\lambda_{t+1},\varphi_{t+1},B_{t+1}|\lambda_{t},\varphi_{t},B_{t}\right) d\eta_{t}^{s}\left(\varphi_{t},B_{t},k_{t}\right) \\ &+ \int \mathbb{1}\left\{V^{c}\left(\lambda_{t+1},\varphi_{t+1},B_{t+1},k_{t+1}\right) \geq V^{x}\left(k_{t+1}\right)\right\} \times \mathbb{1}\left\{k_{t+1} = k^{*}\left(\lambda_{t},\varphi_{t},B_{t},0\right)\right\} \\ &\times j\left(\lambda_{t+1},\varphi_{t+1},B_{t+1}|\lambda_{t},\varphi_{t},B_{t}\right) d\chi_{t}\left(\varphi_{t},B_{t}\right),\end{aligned}$$

where $j\left(\lambda_{t+1},\varphi_{t+1},B_{t+1}|\lambda_{t},\varphi_{t},B_{t}\right)$ denotes the density of transition probability $J\left(\lambda_{t+1},\varphi_{t+1},B_{t+1}|\lambda_{t},\varphi_{t},B_{t}\right)$.

6. The measure of producing firms follows

$$\begin{aligned}\eta_{t+1}^{p}\left(\varphi_{t+1},B_{t+1},k_{t+1}\right) = &\int \mathbb{1}\left\{k_{t+1} = k^{*}\left(\lambda_{t},\varphi_{t},B_{t},k_{t}\right)\right\} \times j\left(\lambda_{t+1},\varphi_{t+1},B_{t+1}|\lambda_{t},\varphi_{t},B_{t}\right) d\eta_{t}^{s}\left(\varphi_{t},B_{t},k_{t}\right) \\ &+ \int \mathbb{1}\left\{k_{t+1} = k^{*}\left(\lambda_{t},\varphi_{t},B_{t},0\right)\right\} \times j\left(\lambda_{t+1},\varphi_{t+1},B_{t+1}|\lambda_{t},\varphi_{t},B_{t}\right) d\chi_{t}\left(\varphi_{t},B_{t}\right).\end{aligned}$$

7. The investment return

$$R_t \geq \beta^{-1}.$$

8. The aggregate output is defined by

$$Y_t \equiv \int A_t \varphi_t k_t^{\alpha} d\eta_t^{p}\left(\varphi_t,B_t,k_t\right).$$

The aggregate equity value is defined by

$$V_t^{a} \equiv \int V^{c}\left(\lambda_t,\varphi_t,B_t,k_t\right) d\eta_t^{s}\left(\varphi_t,B_t,k_t\right),$$

and the aggregate consumption is given by

$$\begin{aligned}C_t = &Y_t - c^{f} \int d\eta_t^{p}\left(\lambda_t,\varphi_t,B_t,k_t\right) - \\ &\int \left[k^{*}\left(\lambda_t,\varphi_t,B_t,k_t\right) - \left(1-\delta\right)k_t - g\left(k_t,k^{*}\left(\lambda_t,\varphi_t,B_t,k_t\right)\right)\right] d\eta_t^{p}\left(\varphi_t,B_t,k_t\right),\end{aligned}$$

which are subject to

$$0 < \frac{V_t^{a}}{Y_t} < \infty,$$

$$0 < \frac{C_t}{Y_t} < \infty.$$

B.4 Solution Method

The key to solve the model is to find out fundamental component $F^c(\lambda, \mu, k)$, which follows

$$F^c(\lambda,\mu,k) = (1-\delta)k - k^{(1)} - g\left(k,k^{(1)}\right) + \frac{1}{R}\int F\left(\lambda^{(1)},\mu^{(1)},k^{(1)}\right) dJ\left(\lambda^{(1)},\mu^{(1)}|\lambda,\mu\right), \tag{39}$$

$$F(\lambda,\mu,k) = A\varphi k^{\alpha} - c^f + 1\{V^c(\lambda,\mu,k) < V^x(k)\} V^x(k) + 1\{V^c(\lambda,\mu,k) \geq V^x(k)\} F^c(\lambda,\mu,k). \tag{40}$$

Note that the size of B is independent to the choice of future capital. Therefore, (39) can be rewritten into

$$F^c(\lambda,\mu,k) = \max_{k'}\left\{(1-\delta)k - k' - g(k,k') + \frac{1}{R}\int F(\lambda',\mu',k')\, dJ(\lambda',\mu'|\lambda,\mu)\right\}, \tag{41}$$

Equation (41) can also be derived from (17) if we remove bubbles from both sides of the equation. We can implement value function iteration to solve for $F^c(\lambda,\mu,k)$.

To facilitate our discussion, throughout this appendix we remove the aggregate state λ and replace the idiosyncratic states μ with φ and B. It is worth noting that in our analysis, we take as given that there exist unique $F^c(\lambda,\mu,k)$, $k^*(\lambda,\mu,k)$, and $p^s(\lambda,\mu,k')$ for a given path of B. To check for the uniqueness of the solution, we follow the typical procedure adopted in the literature by verifying that the solution method leads to the same solution given different initial guesses.

Balanced-growth-path Hereafter, we focus on a BGP. Given a set of parameters, the BGP is solved by the following steps.

1. We start by defining grids over idiosyncratic state variables φ, B, and k. As for φ grid, we use 9 points that are exponentially spaced, so that the log productivity points are equally spaced. The process of φ is approximated by Rouwenhorst method.[36] We use 20 equally spaced points for B grid. We construct a denser grid for k since it is chosen by firms: we use 90 points that are exponentially spaced.

[36] See Kopecky and Suen (2010).

2. Make guesses for two functions: 1) fundamental component, $F^c(\varphi, B, k)$; 2) the survival probability in the subsequent period when a firm faces no bubble crash, $p^o(\varphi, B, k')$. $p^o(\varphi, B, k')$ is used to pin down the future bubble, if it does not crash. We interpolate fundamental component $F^c(\varphi', B', k')$ at $B' = \left(\frac{1}{R} \cdot p^b \cdot p^o(\lambda, \mu, k')\right)^{-1} B$, and combine it with the transition probability of φ, to get

$$
\begin{aligned}
F^e(\varphi, B, k') &\equiv \frac{1}{R} \int F(\varphi', B', k')\, dJ(\varphi', B' | \varphi, B) \\
&= p^b \frac{1}{R} \int F\left(\varphi', \left(\beta \cdot p^b \cdot p^o(\lambda, \mu, k')\right)^{-1} B, k'\right) \xi(\varphi' | \varphi) \\
&+ (1 - p^b) \frac{1}{R} \int F(\varphi', 0, k')\, \xi(\varphi' | \varphi),
\end{aligned}
$$

where $\xi(\varphi' | \varphi)$ denotes the transition probability of idiosyncratic productivity. $F^e(\varphi, B, k')$ is the expected value $\frac{1}{R} \int F(\varphi', B', k')\, dJ(\lambda', \mu' | \lambda, \mu)$ in (41), given φ, B, and k'.

3. Interpolate $F^e(\varphi, B, k')$ at $k' = (1 - \delta) k$, and calculate the continuation value if a firm does not adjust its capital stock.

4. Calculate the continuation value if a firm adjusts its capital stock, and find out the associated optimal k'.

5. Compare values from 3 and those from 4, update the guess of continuation value, and characterize the policy function for k'.

6. Provided the new guess of $F^c(\varphi, B, k)$ and the transition probability of φ, we can compute probability

$$
p^s(\varphi, B', k') \equiv \int_{\Phi^s} d\xi(\varphi' | \varphi),
$$

where $\Phi^s \equiv \{\varphi' | V^c(\varphi', B', k') \geq V^x(k')\}$. Interpolate $p^s(\varphi, B', k')$ at $B' = \left(\frac{1}{R} \cdot p^b \cdot p^o(\varphi, B, k')\right)^{-1} B$ to update the guess of $p^o(\varphi, B, k')$.

7. Repeat Step 3-6 until the guesses converge.

8. Substitute k' in $p^o(\varphi, B, k')$ with the policy function, and denote the new function with $\widetilde{p}^o(\varphi, B, k)$. Now given the transition matrix of φ, $\widetilde{p}^o(\varphi, B, k)$, and the policy function for k', we can construct the transition matrix, Γ, of idiosyncratic states φ, B, and k. Combining the transition matrix with the detrended vector of entrants, $\widetilde{\zeta}$, we can characterize the stationary distribution vector of φ, B, and k. Iterate the following equation

$$
\tilde{\eta}_{t+1} = \frac{1}{1+g} \Gamma \times \tilde{\eta}_t + \tilde{\zeta}
$$

until $\tilde{\eta}_t$ converges. The fixed point of $\tilde{\eta}_t$ is the detrended vector of the stationary distribution on the BGP.

Transition Dynamics Transition dynamics are perfect foresight equilibrium paths following unexpected shocks. Prior to $t=0$, the economy is at the BGP. Shocks are modeled as innovations to p^b, β, or A at $t=0$, and the innovations gradually decay to zero. We assume that the economy returns to the BGP at $t=100$.

Since p^b and A follow deterministic processes after the shocks are realized, and $F_t^c(\varphi_t, B_t, k_t)$ at $t=100$ is equal to the fundamental component $F^c(\varphi, B, k)$ along the BGP, we can use backward induction to pin down fundamental component $F_t^c(\varphi_t, B_t, k_t)$, probability $p_t^o(\varphi_t, B_t, k_{t+1})$, and policy function $k_{t+1}^*(\varphi_t, B_t, k_t)$ back from $t=99$ to $t=0$. Starting from $t=99$

1. We firstly make a guess regarding $p_t^o(\varphi_t, B_t, k_{t+1})$, which is used to predict the associated bubble size at $t+1$. Provided $F_{t+1}^c(\varphi_{t+1}, B_{t+1}, k_{t+1})$ and the transition probability of φ, we can compute probability

$$p_t^s(\varphi_t, B_{t+1}, k_{t+1}) \equiv \int_{\Phi_{t+1}^s} d\xi(\varphi_{t+1}|\varphi_t),$$

where $\Phi_{t+1}^s \equiv \left\{\varphi_{t+1}|V_{t+1}^c(\varphi_{t+1}, B_{t+1}, k_{t+1}) \geq V^x(k_{t+1})\right\}$. Interpolate $p_t^s(\varphi_t, B_{t+1}, k_{t+1})$ at $B_{t+1} = \left(\frac{1}{R_t} \cdot p^b \cdot p_t^o(\varphi_t, B_t, k_{t+1})\right)^{-1} B_t$ to update the guess of $p_t^o(\varphi_t, B_t, k_{t+1})$. Repeat this process until the guess converges.

2. Use backward induction to compute $F_t^c(\varphi_t, B_t, k_t)$ and $k_{t+1}^*(\varphi_t, B_t, k_t)$.

3. Iterate (1)-(2) back from $t=99$ to $t=0$.

Finally, using fundamental component $F_t^c(\varphi_t, B_t, k_t)$, probability $p_t^o(\varphi_t, B_t, k_{t+1})$, and policy function $k_{t+1}^*(\varphi_t, B_t, k_t)$, we compute the transition matrix Γ_t and the vector of entrants $\widetilde{\zeta}_t$ from $t=0$ to $t=99$. We can characterize the distribution of idiosyncratic states along the transition dynamics following

$$\tilde{\eta}_{t+1} = \frac{1}{1+g}\Gamma_t \times \tilde{\eta}_t + \tilde{\zeta}_{t+1},$$

where $\tilde{\eta}_0$ is the stationary distribution on the BGP.

B.5 Verification of $R = \beta^{-1}$ in the Equilibrium

We set $R = \beta^{-1}$ in our computation of BGP and we can calculate the implied aggregate output. Suppose that, on the BGP, households of age $< T$ save everything to consume when they are old. Relative to the aggregate output, the wealth of an age cohort evolves at the rate $\frac{R}{(1+g)}$. The ratio of the oldest cohort's wealth to the aggregate output is decreasing at T: if T is large enough, the ratio is close to zero. As the saving of households of age $< T$ cannot exceed their wealth, the ratio of the wealth of the oldest cohort to the aggregate output is equal to zero if $T \to \infty$.

We examine that along the calibrated BGP, the aggregate consumption (relative to the aggregate output) is positive. As R is fixed in our computation, the size of T does not directly affect the calibration results. T can indeed be arbitrarily large. In our calibrated model, if T is large enough, the wealth of oldest individuals is smaller than the aggregate consumption.

This reasoning also applies to transition dynamics. We fix $R = \beta^{-1}$ in our analysis of transition dynamics. We can examine that along the transition path, the aggregate consumption remains positive relative to the aggregate output. Using the reasoning above, we know that if T is large enough, at any time on the transition path, the aggregate consumption is higher than the consumption of the oldest cohort.

B.6 Social Welfare

Define social welfare as

$$U_t \equiv \int U_{it} di.$$

The social welfare is the sum of the lifetime utility of all existing households. Incumbent individuals have the following budget constraint

$$C_{it} + \int_{\Phi_t} \left\{ V^c \left(\lambda_t, \mu_t, k_t\right) + \left[k_{t+1} - (1-\delta) k_t + g\left(k_t, k_{t+1}\right)\right] \right\} d\theta_{it+1} \left(\mu_t, k_t\right) = \int V\left(\lambda_t, \mu_t, k_t\right) d\theta_{it} \left(\mu_t, k_t\right),$$

where $\theta_{it} \left(\mu_t, k_t\right)$ denotes the distribution (CDF) of shares of incumbent firms. Note that for the oldest individuals, $\theta_{it+1} \left(\mu_t, k_t\right) = 0$, as they consume all their wealth. Newly born individuals have the following budget constraint

$$C_{it} + \int_{\Phi_t} \left\{ V^c \left(\lambda_t, \mu_t, k_t\right) + \left[k_{t+1} - (1-\delta) k_t + g\left(k_t, k_{t+1}\right)\right] \right\} d\theta_{it+1} \left(\mu_t, k_t\right) = \int_{\Phi_t^e} V^e \left(\lambda_t, \mu_t\right) d\theta_{it}^e \left(\mu_t\right),$$

where $\Phi_t^e \equiv \{(\lambda_t, \mu_t) | V^e(\lambda_t, \mu_t) \geq 0\}$, $\theta_{it}^e(\lambda_t, \mu_t)$ denotes the distribution (CDF) of potential entrants.

The budget constraint for an oldest individual can be written as

$$C_{it} = \int V(\lambda_t, \mu_t, k_t)\, d\theta_{it}(\mu_t, k_t).$$

As for an individual of age $T-1$, since there is no aggregate uncertainty, and the return on the investment portfolio is equal to $\frac{1}{\beta}$, the budget constraint can be written as

$$C_{it} + I_{it} = \int V(\lambda_t, \mu_t, k_t)\, d\theta_{it}(\mu_t, k_t),$$

and

$$C_{it+1} = \frac{1}{\beta} I_{it},$$

where

$$I_{it} \equiv \int_{\Phi_t} \{V^c(\lambda_t, \mu_t, k_t) + [k_{t+1} - (1-\delta)k_t + g(k_t, k_{t+1})]\}\, d\theta_{it+1}(\mu_t, k_t).$$

Hence their lifetime utility is equal to

$$U_{it} = C_{it} + \beta C_{it+1} = \int V(\lambda_t, \mu_t, k_t)\, d\theta_{it}(\mu_t, k_t).$$

Keep iterating, and we can obtain that for every incumbent individual i,

$$U_{it} = \int V(\lambda_t, \mu_t, k_t)\, d\theta_{it}(\mu_t, k_t).$$

As for newly born individuals,

$$U_{it} = \int_{\Phi_t^e} V^e(\lambda_t, \mu_t)\, d\theta_{it}^e(\mu_t).$$

In aggregate

$$U_t = \int V(\lambda_t, \mu_t, k_t)\, d\eta_t^p(\mu_t, k_t) + \int_{\Phi_t^e} V^e(\lambda_t, \mu_t)\, d\chi_t(\mu_t).$$

In our model, the social welfare is equal to the aggregate wealth held by all existing individuals.

In Figure 2 we plot the impulse response of welfare following a positive bubble shock. Bubbles have a wealth effect. A positive bubble shock can raise aggregate wealth and therefore lifetime consumption. It might seem obvious that bubbles are distortionary as they act as "subsidies" to firm entry and "taxes" to firm exit. However, this is mistaken. In a bubbly equilibrium, the social benefit of continuation is the value of continuation instead of its fundamental component. Firms make non-distorted decisions when they compares the exit value (the cost of continuation) and continuation value, no matter whether the continuation value is different from its fundamental component.

C SVAR: Medium run restriction

C.1 The identification Issue

We now move to the discussion about the construction of the main driver of the non-fundamental component, which we refer to as the bubble shock. Note that U_t in (30) includes the vector of reduced form residuals, which are linear combinations of structural shocks:

$$U_t = B_0^{-1}\zeta_t, \tag{42}$$

where ζ_t denotes a vector of structural shocks normalized to have unit variances, $(B_0^{-1})'B_0^{-1} = \Sigma_u$, where Σ_u is the covariance matrix of the reduced form residuals. It is trivial to get an unbiased estimate $\hat{\Sigma}_u$. However, the identification issue arises because there are more parameters to be estimated in B_0^{-1} than the number of knowns contained in $\hat{\Sigma}_u$. Therefore, structural assumptions are required to overcome this identification problem. One way to understand how to impose structural assumptions is to rewrite B_0^{-1} as:

$$B_0^{-1} = AQ, \tag{43}$$

where A is a lower triangular matrix with $AA' = \Sigma_u$, e.g., the Cholesky decomposition of $\hat{\Sigma}_u$ is a natural candidate for $\hat{A}$, and Q is a orthonormal matrix such that $QQ' = I$. Solving the identification problem boils down to find the orthonormal matrix Q such that identification assumptions are satisfied.

We are interested in one shock, labeled as a bubble shock, hence it is sufficient to identify one column of B_0^{-1} associated with the bubble shock. To do so, we rely on the medium-run restriction. As explained above, the medium-run restriction is problematic if we fail to control for other shocks that might be, albeit unlikely, the main driver of the price-fundamental differential. In the baseline, we control for both unanticipated and anticipated productivity shocks that are, arguably, the main drivers of the business cycle.

To control for unexpected TFP shocks, following Levchenko and Pandalai-Nayar (2020), we include Fernald (2014)'s measure of utility adjusted TFP in our VAR. The first difference of this variable is widely used in the empirical literature as a measure of unexpected productivity shock, see e.g., Garín, Lester and Sims (2019) and Loria, Matthes and Zhang (2019) for recent applications, and Ramey (2016) for a survey.[37] We assume that unanticipated shocks are the only shocks that affect TFP contemporaneously. With TFP ordered the first, a simple Cholesky decomposition of $\hat{\Sigma}_u$ gives us a lower triangular matrix $\hat{A}$ whose first column is associated with the TFP shocks:

$$B_0^{-1}\zeta_t = \underbrace{\begin{bmatrix} * & 0 & 0 & 0 & 0 & 0 \\ * & * & 0 & 0 & 0 & 0 \\ * & * & * & 0 & 0 & 0 \\ * & * & * & * & 0 & 0 \\ * & * & * & * & * & 0 \\ * & * & * & * & * & * \end{bmatrix}}_{\hat{A}} \begin{bmatrix} \epsilon_t^{TFP} \\ u_{2,t} \\ u_{3,t} \\ u_{4,t} \\ u_{5,t} \\ u_{6,t} \end{bmatrix} \tag{44}$$

In addition, we control for anticipated productivity shocks, i.e., news shocks. Those shocks are identified as a linear combination (Q_1) of remaining reduced form residuals, $[u_{2,t}, ..., u_{6,t}]'$, such that it contributes the most to the cumulative sum of the square of the forecast error of TFP, i.e., $\sum_{h=0}^{H_1}(FE_{t+h|t}^{tfp})^2$. With this Q_1, the second shock is identified as news shocks:

$$B_0^{-1}\zeta_t = \hat{A}Q_1 \begin{bmatrix} \epsilon_t^{TFP} & \epsilon_t^{news} & u_{3,t} & u_{4,t} & u_{5,t} & u_{6,t} \end{bmatrix}'. \tag{45}$$

In the next section, we explain the procedure to derive Q_1.

The last identification assumption assumes that, once controlled for unanticipated and anticipated news shocks, bubble shocks are the ones that maximize the forecast error variance decomposition of price-fundamental differential. That is, bubble shocks are identi-

[37] An alternative popular approach to identify productivity shock replies on long-run restriction (Blanchard and Quah 1989 and Galí 1999), this is not suitable as we show both empirically and theoretically that bubble shocks have very persistent effect on productivity.

fied as a linear combination (Q_2) of remaining reduced form residuals, $[u_{3,t}, u_{4,t}, u_{5,t}, u_{6,t}]'$, such that it contributes the most to the cumulative sum of the square of the forecast error of $p_t - f_t$, i.e., $\sum_{h=0}^{H_2}(FE_{t+h|t}^{p/f})^2$, where f_t is defined as in (29). With such an orthonormal matrix Q_2 in place, the third shock is identified as the bubble shock:

$$B_0^{-1}\zeta_t = \hat{A}Q_1Q_2\left[\epsilon_t^{TFP} \quad \epsilon_t^{news} \quad \epsilon_t^{b} \quad u_{4,t} \quad u_{5,t} \quad u_{6,t}\right]'. \tag{46}$$

In what follows, we illustrate the formal procedure to derive Q_2 and we demonstrate how to control for additional shocks.

C.2 Technical Details

The staring point is to obtain impulse responses to reduced form residuals. Those can be obtained by considering our VAR(p) in companion form:

$$\mathbb{Y}_t = \mathbb{A}\mathbb{Y}_{t-1} + \mathbb{U}_t,$$

where

$$\mathbb{Y}_t \equiv \begin{bmatrix} Y_t \\ \vdots \\ Y_{t-p+1} \end{bmatrix}, \mathbb{A} \equiv \begin{bmatrix} B_1 & B_2 & \dots & B_{p-1} & B_p \\ I_K & 0 & & 0 & 0 \\ 0 & I_K & & 0 & 0 \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & I_K & 0 \end{bmatrix}, \mathbb{U}_t \equiv \begin{bmatrix} U_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

where K is the number of variables. Solve this equation forward:

$$\mathbb{Y}_{t+h} = \mathbb{A}^{h+1}\mathbb{Y}_{t-1} + \sum_{j=0}^{h}\mathbb{A}^j\mathbb{U}_{t+h-j},$$

Multiply this equation by $\mathbb{J} \equiv [I_K, 0_{K\times K(p-1)}]$ yields:

$$\begin{aligned} Y_{t+h} &= \mathbb{J}\mathbb{A}^{h+1}\mathbb{Y}_{t-1} + \sum_{j=0}^{h} \mathbb{J}\mathbb{A}^{j}\mathbb{U}_{t+h-j} \\ &= \mathbb{J}\mathbb{A}^{h+1}\mathbb{Y}_{t-1} + \sum_{j=0}^{h} \mathbb{J}\mathbb{A}^{j}\mathbb{J}'\mathbb{J}\mathbb{U}_{t+h-j} \\ &= \mathbb{J}\mathbb{A}^{h+1}\mathbb{Y}_{t-1} + \sum_{j=0}^{h} \mathbb{J}\mathbb{A}^{j}\mathbb{J}' U_{t+h-j}. \end{aligned}$$

Therefore, the response of the variable $j = 1, \ldots, K$ to a reduced form residual u_{kt} that occurred h periods ago, is given by:

$$\Phi_h \equiv [\phi_{jk,h}] \equiv \mathbb{J}\mathbb{A}^{h}\mathbb{J}'.$$

The h-step ahead forecast error is:

$$Y_{t+h} - Y_{t+h|t} = \sum_{i=0}^{h-1} \Phi_i U_{t+h-i}.$$

Hence the MSFE at horizon h is:

$$MSFE_h = \sum_{i=0}^{h-1} \Phi_i \Sigma_u \Phi_i'.$$

Identifying News Shocks By imposing structural assumptions through an orthonormal matrix Q, the structural impulse response is given by $\Phi_h AQ$. And recall that A is the lower triangular matrix resulting from the Cholesky decomposition of Σ_u. With TFP ordered the first, the first shock is identified as unanticipated productivity shock under the assumptions that no other shocks can affect TFP contemporaneously. News shocks are identified as the linear combination (Q_1) of remaining reduced form residuals that contribute the most to the MSFE of TFP at horizons up to H_1.

$$Q_1 = \operatorname{argmax} \frac{\boldsymbol{e_i}'(\sum_{h=0}^{H_1} \Phi_h \hat{A} Q_1 \boldsymbol{e_j} \boldsymbol{e_j}' Q_1' \hat{A}' \Phi_h') \boldsymbol{e_i}}{\boldsymbol{e_i'}(\sum_{h=0}^{H_1} MSFE_h) \boldsymbol{e_i}}, \tag{47}$$

s.t.

$$Q_1 \equiv \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & Q_1(2,2) & & Q_1(2,p) \\ \vdots & & \ddots & \\ 0 & Q_1(K,2) & & Q_1(K,p) \end{bmatrix}$$

$$Q_1 Q_1' = 1$$

$\boldsymbol{e_i}$ denotes selection vectors with one in the ith place and zeros elsewhere. In our empirical application, we set $i = 1$ and $j = 2$ to label the second shock as the one that maximize forecast error variance of the first variable. Note that the first column and the first row of Q_1 are specified in this way to select non-productivity shocks at the same time guarantee its orthogonality.

Identifying Bubble Shocks Once we have identified the first two shocks, we identify the bubble shocks as the linear combination (Q_2) of the remaining residuals that contribute the most to the MSFE of $p_t - f_t$ at horizons up to H_2.

The response of $\sum_{i=0}^{\infty} \Lambda^i E_t(Y_{t+i})$ to a shock that occurred h periods ago, upon a successful selection of the matrix Q_2, is given by:

$$\widetilde{\Phi}_{h,0} = \mathbb{J}\mathbb{A}^{h-1}(1 - \Lambda\mathbb{A})^{-1}\mathbb{J}' A Q_1 Q_2$$

The response of $\sum_{i=0}^{\infty} \Lambda^i (1 - \Lambda) E_t(Y_{t+i+1})$ to a shock that occurred h periods ago is given by:

$$\widetilde{\Phi}_{h,1} = (1 - \Lambda)\Lambda\mathbb{J}\mathbb{A}^{h}(1 - \Lambda\mathbb{A})^{-1}\mathbb{J}' A Q_1 Q_2$$

Hence, the response of the variable f_t to a shock that occurred h periods ago, is given by:

$$\widetilde{\Phi}_h^f = \widetilde{\Phi}_{h,1}(d) - \widetilde{\Phi}_{h,0}(r),$$

where $\widetilde{\Phi}_{h,1}(d)$ and $\widetilde{\Phi}_{h,0}(r)$ selects, respectively, the vector of response associated with d_t and r_t. The reponse of $p_t - f_t$ to a shock that occurred h periods ago, is given by:

$$\widetilde{\Phi}_h^{p/f} = \Phi_h(p) - \widetilde{\Phi}_h^f.$$

Therefore, the MSFE of $p_t - f_t$ at horizon h is given by:

$$MSFE_h^{p/f} = \sum_{i=0}^{h-1} \widetilde{\Phi}_i^{p/f} \Sigma_u (\widetilde{\Phi}_i^{p/f})'.$$

We identify the bubble shocks as the linear combination (Q_2) of the remaining residuals that contribute the most to the MSFE of $p_t - f_t$ at horizons up to H_2. Formally:

$$Q_2 = \text{argmax} \frac{\sum_{h=0}^{H_2} \widetilde{\Phi}_h^{p/f} \boldsymbol{e_j e_j}' (\widetilde{\Phi}_h^{p/f})'}{\sum_{h=0}^{H_1} MSFE_h^{p/f}}, \tag{48}$$

s.t.

$$Q_2 \equiv \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & Q_2(3,3) & & Q_2(3,p) \\ \vdots & & & \ddots & \\ 0 & 0 & Q_2(K,3) & & Q_2(K,p) \end{bmatrix}$$

$$Q_2 Q_2' = 1.$$

Controlling for other Shocks Our baseline identification strategy can be easily extended to control for more shocks. In our empirical application, as robustness checks, we control for credit supply, uncertainty and risk premium shocks, and shocks to the short term nominal interest rate. To do so, we include (i) uncertainty and risk premium — VIX (Bekaert *et al.* 2013), (ii) credit supply — Gilchrist and Zakrajsek (2012)'s excess bond premium, and (iii) the short term nominal rate into our VAR, separately to avoid the curse of dimensionality. These variables are ordered the second in the VAR system. Shocks to these control variables are identified using the classical recursive restrictions: on impact, apart from the structural shock in question, no other structural shocks affects the variable from which the structural shocks is identified. Our bubble shocks are then identified analogically to the procedure described above: as the linear combination of reduced form residuals, excluding unanticipated and anticipated TFP shocks and the additional shock that we control for, that maximize the FEVD of the price fundamental differential.

D Data

This section provides a brief overview of the data we use throughout this paper. The stock price, dividend and earning of the SP500 are taken from Shiller (2015), which are updated and made available on the author's website. We take the Gilchrist and Zakrajsek (2012) excess bond premium (EBP henceforth) updated by Favara *et al.* (2016). We include Fernald (2014)'s utility adjusted TFP as a measure of TFP. The remaining variables are taken from FRED: the real gross domestic product per capita (real GDP), the gross private domestic investment, and the civilian unemployment rate that are seasonally adjusted, Moody's Baa Corporate Bond Yield (to construct real interest rate r_t), CBOE Volatility Index (VIX), 3-month treasury bill rate, and GDP deflator. All nominal variables are rescaled by GDP deflator to obtain their real values. The Business Dynamics Statistics (BDS) provide information about firms' entry and exit rates at annual frequency from 1977 to 2016. We interpolate them into quarterly frequency, using a simple linear interpolation.

E Linking the VAR with the Model

In this section, we apply the empirical strategies adopted in the previous section using data simulated from our structural model for validating the empirical strategies and providing plausible structural interpretations to the resulting reduced-form estimates.

We simulate the time series following Boppart *et al.* (2018). We include three types of shocks: (1) shocks to p^b, which correspond to the bubbles shocks in the baseline VAR, (2) shocks to A_t, which correspond to the technology shocks in the baseline VAR, and (3) shocks to β (i.e., shocks to R), which corresponds to the interest rate shocks. All shocks follow AR(1) processes. Without loss of generality, we set the auto-correlations of all types of shocks to 0.5. The relative standard deviations of technology, bubble, and interest rate shocks are chosen such that they account for 25%, 60%, and 10% of the FEVD of real output at horizon 40, respectively. This choice corresponds to the main driver of the non-fundamental component justifying more than 60% of the FEVD of the real GDP at 40 quarters ahead. Our results are robust to different parameterizations of shocks processes. We add small measurement errors (with standard errors equal to 4% of the standard deviation of interest rate shocks) to each simulated data to avoid dynamic singularity.

The Decomposition of Asset Prices using VAR We treat the simulated data similarly as we treat the data in Section 5.1 with the same sample size (160 periods) to obtain the estimated price–fundamental differential components. We repeat this exercise 1000 times using 1000 simulated samples. Figure 13 plots the distribution of correlations between the true price–fundamental differential component of the aggregate asset price in the model and the estimated one. A correlation that equals 1 indicates a perfect match. Our simulation shows that our decomposition strategy based on a linear VAR successfully captures the truth in a world described by our nonlinear structural model.

What Is the Main Driver of the Non-Fundamental Component? The following simulation exercise shows that in the model, the medium-run restriction identifies bubble shocks. We estimate the SVAR with the medium-run restriction using 1000 sample simulated from the model.

The blue dashed lines in Figure 14 plot the effects of a positive bubble shock in the model, called the true IRFs. The red lines in Figure 14 plot the estimated IRFs by applying our SVAR model to the simulated data. The solid red line indicates the median of the estimated IRFs from 1000 simulations. The dotted red lines represent the 90% intervals of the estimated IRFs from the 1000 simulations. The median of the estimated IRFs (out of 1000 simulations) matches the true IRFs underlying the model.

We conduct the same exercise using the data simulated from the model without bubbles. Figure 15 reports the results. Without bubbles, the model fails to produce empirical findings presented in Figure 4.

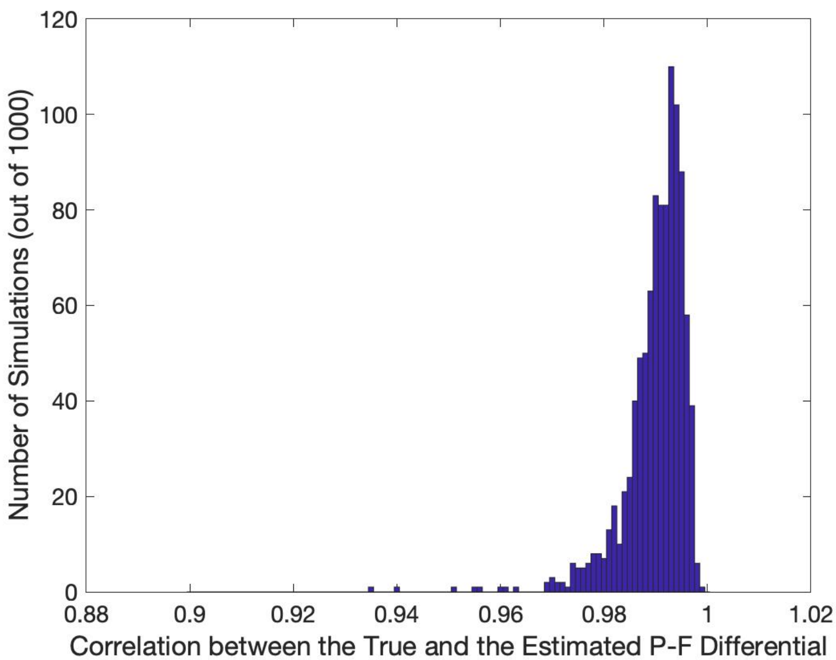

Figure 13: This figure plots the distribution of the correlation between the true and the estimated price-fundamental differential component based on 1000 simulations of the model.

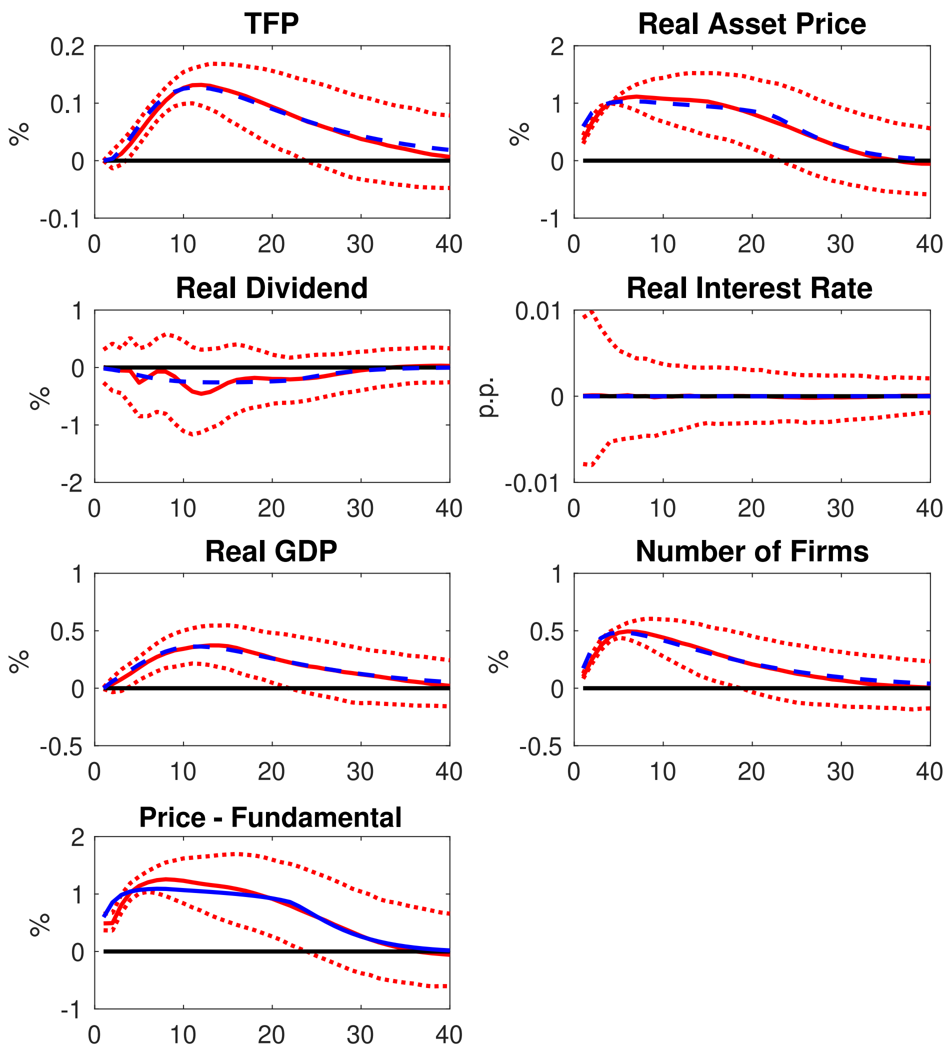

Figure 14: This figure plots the true IRFs to a positive bubble shocks in the model (blue dashed line) together with the estimated IRFs by applying our SVAR model to the simulated data (red lines). The solid red line indicates the median of the estimated IRFs from 1000 simulations. The dotted red lines represent the 90% intervals out of the 1000 simulations. Each simulated data set consists of 160 periods—the same sample length as in the data.

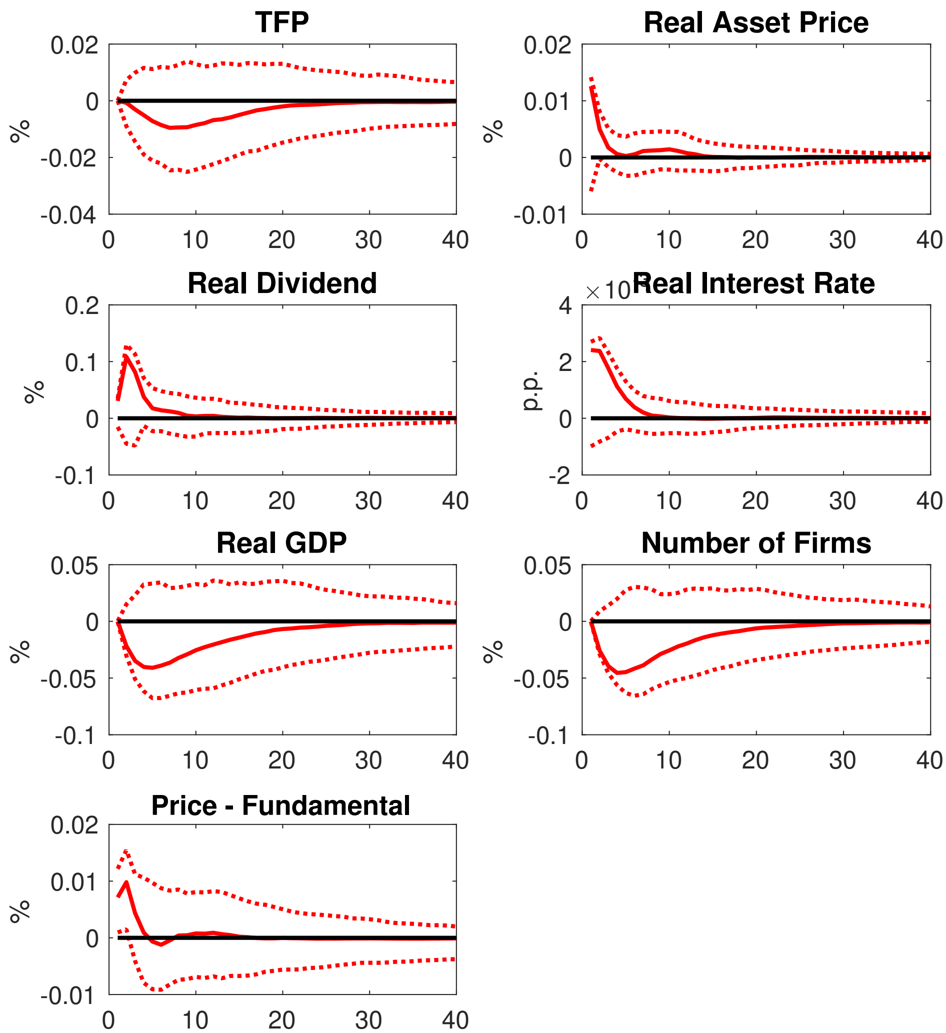

Figure 15: This figure plots the estimated IRFs to the shocks by applying our SVAR model to data simulated from a model without bubbles. The solid red line indicates the median of the estimated IRFs from 1000 simulations. The dotted red lines represent the 90% intervals out of the 1000 simulations. Each simulated data set consists of 160 periods—the same sample length as in the data.