ECONOMIC RESEARCH

FEDERAL RESERVE BANK OF ST. LOUIS

WORKING PAPER SERIES

QE: When and How Should the Fed Exit?

Federal Reserve Bank of St. Louis, Research Division, P.O. Box 442, St. Louis, MO 63166

QE: When and How Should the Fed Exit?*

Yi Wen

Federal Reserve Bank of St. Louis

&

Tsinghua University

July 21, 2014

Abstract

The essence of Quantitative Easing (QE) is to reduce the costs of private borrowing through large-scale purchases of privately issue debts, instead of public debts (Ben Bernanke, 2009). Notwithstanding the effectiveness of this highly unconventional monetary policy in reviving private investment and the economy, it is time to think about the likely impacts of the unwinding of QE (or the reversed private-asset purchases) on the economy. In a standard economic model, if monetary injections can increase aggregate output and employment, then the reversed action will likely undo such effects. Would this imply that the U.S. economy will dive into a recession once the Fed starts its large-scale asset sales (under the assumption that QE has successfully pulled the economy out of the Great Recession)? This paper shows that three aspects of the Federal Reserve's exit strategy matter for achieving (or maintaining) maximum gains in aggregate output and employment under QE (if any): (i) the timing of exit, (ii) the pace of exit, and (iii) the private sector's expectations of when and how the Fed will exit.

Keywords: Large-Scale Asset Purchases, Unconventional Monetary Policies, Quantitative Easing, Qualitative Easing, Optimal Exit Strategies.

JEL codes: E50, E52.

1 Introduction

Since the onset of the financial crisis in late 2007, the Federal Reserve has injected an astronomical amount of money into the economy through its large-scale asset purchase (LSAP) programs. According to former Fed Chairman Ben Bernanke (2009, p5), the essence of LSAP is "credit easing" (CE)—that is, reducing the cost of private borrowing by direct purchases of privately issued debt instead of government debt. However, given that the government has no intention to hold private debt on its balance sheet forever, at some point the Fed must sell it.

The goal of this article is to answer the following questions about the likely effects of the Fed's exit strategies:

- Will the Fed's exit from quantitative easing (QE) undo the gains of LSAPs (if any)?

- Do the timing and pace of the exit matter?

- Should the exit be state dependent, and how would the economy respond to a fully anticipated exit compared with an unanticipated exit?

In this article, I use a calibrated general equilibrium model to shed light on these questions. The model's framework features explicit asset purchases by the government that mimic the real-world scenario. In this article, QE, CE, and LSAPs are considered synonymous. To highlight the general equilibrium effects of QE and facilitate the study of exit strategies, I use a real model in which long-run inflation is fully anchored, consistent with the fact that the U.S. inflation rate has remained stable and below the 2 percent target since the implementation of QE in 2008. In the real model, all transactions and payments are conducted by the exchange/transfer of goods, so there is no need to distinguish monetary authority from fiscal authority. More specifically, I assume there is a consolidated government that can purchase private assets using revenues raised by lump-sum taxes or sales of public debt.[1]

My basic findings can be summarized as follows. Assuming that inflation is fully anchored, the longer and more massively the Fed can hold private debt on its balance sheet before the adverse financial shocks dissipate, the more likely QE will be able to stimulate aggregate investment and employment. In other words, QE is unlikely to have any significant effect on the economy if its scope is too small and too transitory. However, when the aggregate shocks are not permanent, QE may lower the steady-state output if the Fed never exits QE. Consequently,

---

[1]To simplify the analysis, public debt is not modeled in this paper. Interested readers are referred to Wen (2013) for the case when LSAPs involving both public and private debt.

not only does an optimal timing of the Fed's exit exist (depending on the persistence of the financial shocks), but the pace of the exit and whether the exit is anticipated or unanticipated also matter. In particular, it is optimal for the exit to be completely unanticipated by the public to preserve the maximum gains on aggregate output and employment achieved under QE. However, once the exit starts, it is better for it to be quick rather than gradual. Accordingly, it would be a mistake for the Fed to pre-announce or discuss in advance the timing of an exit from QE soon after it was implemented; this sort of announcement would shorten the effective duration of QE anticipated by the public. On the other hand, if the exit is too gradual, the long-run adverse effects of QE would arise and thus offset the benefits from QE gained in the earlier periods.

My model is a fairly standard off-the-shelf model based on the recent macro-finance literature.[2] A key feature of this class of models is that an endogenously determined distribution of heterogeneous creditors/debtors (instead of households' time preference per se) pins down the real interest rate and asset prices in the asset market through the demand and supply of privately issued debts. QE influences the real economy by affecting the allocation (distributions) of credit/debt in the asset markets. Two key assumptions in the model dictate my results:

- Debtors are relatively more productive than creditors—in other words, more productive agents opt to issue debts and less productive agents opt to lend.

- Financial markets are incomplete—that is, agents face uninsurable idiosyncratic shocks and are borrowing-constrained.

Under these fairly standard assumptions, the following properties emerge naturally from the model. First, the demand for liquid assets increases despite their low returns relative to capital investment. Second, when the cost of borrowing is reduced, marginal creditors self-select to become debtors, which can raise the quantity of aggregate debt but also unambiguously lower the average quality (efficiency) of loans.

Given these core properties, it is clear that QE's main effect on aggregate output is changing the distribution of credit/debt in the financial market—that is, QE pushes more creditors to become debtors, which in turn increases the total quantity of loans but at the same time decreases the average efficiency of loans. When economic activities depend not only on the extent and scope of credit but also on the quality of loans, such a trade-off between quantity and quality implies two things: (i) Aggregate output and employment are insensitive to small-scale temporary asset purchases even with relatively large changes in the real interest rate and asset prices (Wen, 2013). And (ii) QE's positive quantitative effect on aggregate investment

---

[2]See, for example, Wang and Wen (2009, 2012, 2013) and Wen (2013).

may dominate its adverse qualitative effect in the short run to mitigate negative financial shocks if asset purchases are sufficiently large and persistent relative to the magnitudes of financial shocks. This property renders QE much less effective if its exit is fully anticipated. On the other hand, since permanent QE may reduce the steady-state output when financial shocks are not permanent, it is desirable to exit not only at a certain point in time, but also as quickly as possible once the exit starts.

2 The Model

The key actors in the model are firms, which make production and investment decisions in a uncertain world. There is a credit market where firms can lend/borrow from each other by issuing/purchasing private debt. Firms face idiosyncratic uncertainty in the rate of returns to their investment projects, modeled specifically as an idiosyncratic shock to the marginal efficiency of firm-level investment (as specified below). In any period, some firms opt to lend and some opt to borrow, depending on their draws of the idiosyncratic shock to the return on investment projects. The real interest rate will then be determined endogenously by the supply and demand of private debt.

2.1 Government

The consolidated government uses lump-sum taxes on household income to finance purchases of private debt. Total private debt purchased by the government in period t is denoted by B_{t+1} and the market price of private debt by $\frac{1}{1+r_t}$, where r_t is the real interest rate on private debt. Total money supply in end of period t is denoted by M_{t+1}, the aggregate price level by P_t, and the inflation rate by $1 + \pi_t = \frac{P_t}{P_{t-1}}$. The government budget constraint in each period is given by

$$ G_t + \frac{1}{1+r_t} B_{t+1} = B_t + \frac{(M_{t+1} - M_t)}{P_t} + T_t, \tag{1} $$

where the left-hand side (LHS) is total government expenditures and the right-hand side (RHS) is total government revenues. Government outlays include government spending G_t and new purchases of private debt B_{t+1} at price $\frac{1}{1+r_t}$. Total government revenues include debt repayment B_t from the private sector, real seigniorage income $\frac{(M_{t+1}-M_t)}{P_t}$, and lump-sum taxes T_t.

2.2 Firms' Problem

There is a continuum of firms indexed by $i \in [0, 1]$. A firm i's objective is to maximize the present value of its discounted future dividends,

$$V_t(i) = \max E_t \sum_{\tau=0}^{\infty} \beta^\tau \frac{\Lambda_{t+\tau}}{\Lambda_t} d_{t+\tau}(i), \tag{2}$$

where $d_t(i)$ is firm i's dividend in period t and Λ_t is the representative household's marginal utility, which firms take as given. The production technology is given by the constant returns to scale function

$$y_t(i) = A_t k_t(i)^\alpha n_t(i)^{1-\alpha}, \tag{3}$$

where A_t represents aggregate technology level, and $n_t(i)$ and $k_t(i)$ are firm-level employment and capital, respectively. Firms accumulate their own capital stock through the law of motion,

$$k_{t+1}(i) = (1 - \delta)k_t(i) + \varepsilon_t(i) x_t(i), \tag{4}$$

where investment $x_t(i)$ is irreversible:

$$x_t(i) \geq 0, \tag{5}$$

and an i.i.d. idiosyncratic shock to the marginal efficiency of investment is denoted by $\varepsilon_t(i)$. In each period t, a firm needs to pay wages $W_t n_t(i)$, decide whether to invest in fixed capital, and distribute dividends $d_t(i)$ to households. Firms' investment is financed by internal cash flow and external funds. Firms raise external funds by issuing one-period debt (bonds), $b_{t+1}(i)$, which pays the competitive market interest rate r_t. Note that $b_{t+1}(i) < 0$ when a firm holds bonds issued by other firms (i.e., $b_{t+1}(i)$ can be either positive or negative).

A firm's dividend in period t is then given by

$$d_t(i) = y_t(i) + \frac{b_{t+1}(i)}{1 + r_t} - i_t(i) - w_t n_t(i) - (1 - \mathbf{P}_t) b_t(i), \tag{6}$$

where the probability of default or the aggregate default risk of private debt is denoted by $\mathbf{P}_t$. For ease of exposition, we temporarily set $\mathbf{P}_t = 0$ and defer the discussions to Section (4.2). Firms cannot pay negative dividends,

$$d_t(i) \geq 0; \tag{7}$$

which is the same as saying that fixed investment is financed entirely by internal cash flow $(y_t(i) - W_t n_t(i))$ and external funds net of loan repayment $(\frac{b_{t+1}(i)}{1+r_t} - b_t(i))$. The idiosyncratic shock to investment efficiency has the cumulative distribution function $F(\varepsilon)$.

Loans are subject to collateral constraints. That is, firm i is allowed to pledge a fraction $\theta \in (0, 1]$ of its fixed capital stock $k_t(i)$ at the beginning of period t as collateral. In general, the parameter θ represents the extent of financial market imperfections or the tightness of the financial market. At the end of period t, the pledged collateral is priced by the market value of newly installed capital, so the market value of collateral is simply Tobin's q, denoted by q_t, which is equivalent to the expected value of a firm that owns collateralizable capital stock $\theta k_t(i)$. The borrowing constraint is thus given by

$$b_{t+1}(i) \leq \theta q_t k_t(i), \tag{8}$$

which specifies that any new debt issued cannot exceed the collateral value (q_t) of a firm with the pledged capital stock $\theta k_t(i)$. When $\theta = 0$ for all t, the model is identical to one that prohibits external financing.

Firms affect both the supply and demand of the private debt, and they may also affect the demand for money when the real rate of return on money dominates that on private debt. To simplify the analysis, I start with the equilibrium condition that $\frac{1}{1+\pi} < 1 + r$, so that firms hold only private debts and no money. When this condition is violated (i.e., when $\frac{1}{1+\pi} = 1 + r$), the two assets become perfect substitutes. In this case, firms are indifferent between holding money and private debt, and their portfolios are determined in equilibrium by the aggregate supply of each asset. This situation is called a "liquidity trap".[3]

2.3 The Household's Problem

There is a representative household, and it is assumed that the household is subject to the cash-in-advance (CIA) constraint for consumption purchases, $C_t \leq \frac{M_t}{P_t}$. Since firms may also hold money as an alternative store of value, the CIA constraint implies that the household is always the residual claimant of the aggregate money stock whenever the CIA constraint binds. Because there is a liquidity premium on the privately issued bonds and households do not face idiosyncratic risk and incomplete financial markets, the rate of return to private bonds is

---

[3]However, the liquidity trap is not addressed in this paper; interested readers are referred to Wen (2013). Ignoring the liquidity trap has no effect on the conclusions here because exiting QE tends to increase the interest rate and, thus, relax the constraint of the zero lower bound on the economy.

dominated by that of equity. Hence, the household does not hold private bonds in equilibrium.[4] The representative household chooses nominal money demand M_{t+1}, consumption plan C_t, labor supply N_t, and share holdings $s_{t+1}(i)$ of each firm i to solve

$$\max \sum_{t=0}^{\infty} \beta^t \left\{ \log C_t - \frac{N_t^{1+\gamma}}{1+\gamma} \right\} \tag{9}$$

subject to the constraints,

$$C_t \leq \frac{M_t}{P_t} \tag{10}$$

$$C_t + \frac{M_{t+1}}{P_t} + \int_{i=0}^{1} s_{t+1}(i) \left[V_t(i) - d_t(i) \right] di \leq \frac{M_t}{P_t} + W_t N_t + \int_{i=0}^{1} s_t(i) V_t(i) di - T_t, \tag{11}$$

where T_t denotes lump-sum income taxes, $s_t(i) \in [0,1]$ denotes firm i's equity shares, and $V_t(i)$ denotes the value of the firm (stock price). Let Λ_t be the Lagrangian multiplier of budget constraint (11), the first-order condition for $s_{t+1}(i)$ is given by

$$V_t(i) = d_t(i) + E_t \beta \frac{\Lambda_{t+1}}{\Lambda_t} V_{t+1}(i). \tag{12}$$

Equation (12) implies that the stock price $V_t(i)$ of firm i is determined by the present value of the firm's discounted future dividends, as in equation (2).

3 Competitive Equilibrium

Given the initial money balance of the household M_0, the initial level of government holding of private debt B_0, the initial distributions of private debt $b_0(i)$ and capital stocks $k_0(i)$ across firms, a competitive equilibrium consists of the sequences and distributions of quantities $\{C_t, N_t, M_t\}_{t=0}^{\infty}$, $\{x_t(i), n_t(i), y_t(i), k_{t+1}(i), b_{t+1}(i)\}_{t\geq 0}$ for $i \in [0,1]$ and the sequences of prices $\{P_t, W_t, V_t(i), r_t\}_{t=0}^{\infty}$ such that

(i) Given prices $\{W_t, r_t\}_{t\geq 0}$, the sequences $\{x_t(i), n_t(i), y_t(i), k_{t+1}(i), b_{t+1}(i)\}_{t\geq 0}$ solve problem (2) for all firms subject to constraints (3) through (8).

(ii) Given prices $\{P_t, W_t, V_t(i)\}_{t\geq 0}$, the sequences $\{C_t, N_t, M_t, s_{t+1}(i)\}_{t\geq 0}$ maximize the house-

---

[4]Introducing idiosyncratic risk into the household sector does not affect the main findings qualitatively but may substantially complicate the analysis by making the model analytically intractable.

hold's lifetime utility (9) subject to its budget constraint (11) and the CIA constraint (10).

(iv) All markets clear:

$$\int b_{t+1}(i)\, di = B_{t+1} \tag{13}$$

$$s_{t+1}(i) = 1 \qquad \text{for all } i \in [0,1] \tag{14}$$

$$N_t = \int n_t(i)\, di \tag{15}$$

$$C_t + \int x_t(i)\, di + G_t = \int y_t(i)\, di \tag{16}$$

$$M_t = \bar{M}, \tag{17}$$

where equation (13) states that the net supply of private bonds issued by all firms equals the total purchase of private bonds by the government. Note that if $B_{t+1} = 0$, then the government holds zero private debt and all bonds issued by firms are circulated only among themselves with zero net supply/demand.

3.1   Firms' Decision Rules

Under constant returns to scale technology a firm's labor demand is proportional to its capital stock. Hence, a firm's net cash flow (revenue minus wage costs) is also a linear function of its capital stock,

$$y_t(i) - W_t n_t(i) = \alpha A_t \left(\frac{(1-\alpha) A_t}{w_t} \right)^{\frac{1-\alpha}{\alpha}} k_t(i) \equiv R(W_t, A_t) k_t(i), \tag{18}$$

where R_t depends only on the aggregate state. With this notation of R_t, we have the following two propositions:

Proposition 1 *The decision rule for investment is characterized by an optimal cutoff ε_t^* such that a firm undertakes capital investment if and only if $\varepsilon_t(i) \geq \varepsilon_t^*$:*

$$x_t(i) = \begin{cases} \left[R_t + \frac{\theta_t q_t}{1+r_t} \right] k_t(i) - b_t(i). & \text{if } \varepsilon_t(i) \geq \varepsilon_t^* \\ \\ 0 & \text{if } \varepsilon_t(i) < \varepsilon_t^* \end{cases}, \tag{19}$$

8

where the cutoff ε_t^* is a function of the aggregate state space only, is independent of the individual firm's history, and is a sufficient statistic for characterizing the distribution of the firm's actions.

Proof. See Appendix I. ∎

Proposition 2 *The equilibrium interest rate of private debt satisfies the following relation:*

$$\frac{1}{1+r_t} = \beta E_t \frac{\Lambda_{t+1}}{\Lambda_t} Q(\varepsilon_{t+1}^*), \tag{20}$$

where $Q(\varepsilon_t^*) \equiv \int_{\varepsilon(i) < \varepsilon_t^*} d\mathrm{F}(\varepsilon) + \int_{\varepsilon(i) \geq \varepsilon_t^*} \frac{\varepsilon_t(i)}{\varepsilon_t^*} dF(\varepsilon).$

Proof. See Appendix II. ∎

3.2 Aggregation

Proposition 3 *Define aggregate capital stock as $K_t = \int k_t(i) di$, aggregate employment as $N_t = \int n_t(i) di$, aggregate output as $Y_t = \int y_t(i) di$, and aggregate investment expenditure as $I_t = \int x_t(i) di$. Since the cutoff ε_t^* is a sufficient statistic for characterizing the distribution of firms, the model's equilibrium can be fully characterized as the sequences of aggregate variables $\{C_t, K_{t+1}, I_t, Y_t, N_t, R_t, \varepsilon_t^*, r_t, W_t, P_t\}_{t=0}^{\infty}$, which can be solved by the following system of nonlinear equations (given the path of any aggregate shocks, money supply $\{\bar{M}_t\}_{t=0}^{\infty}$, the distribution function $F(\varepsilon)$, and the initial distributions of private debts $b_0(i)$ and capital stocks $k_0(i)$):*

$$C_t = \frac{M_t}{P_t} \tag{21}$$

$$\Lambda_t W_t = N_t^\gamma \tag{22}$$

$$\frac{1}{C_t} = \Lambda_t + \Theta_t \tag{23}$$

$$\Lambda_t = \frac{\beta}{1+\pi} E_t (\Lambda_{t+1} + \Theta_{t+1}) \tag{24}$$

$$C_t + I_t + G_t = Y_t \tag{25}$$

$$\frac{1}{\varepsilon_t^*} = \beta E_t \frac{\Lambda_{t+1}}{\Lambda_t} \left\{ R_{t+1} Q(\varepsilon_{t+1}^*) + \frac{\theta_{t+1}}{\varepsilon_{t+1}^*} \frac{[Q(\varepsilon_{t+1}^*) - 1]}{1 + r_{t+1}^c} + \frac{(1-\delta)}{\varepsilon_{t+1}^*} \right\} \tag{26}$$

$$\frac{1}{1 + r_t} = \beta E_t \frac{\Lambda_{t+1}}{\Lambda_t} Q(\varepsilon_{t+1}^*) \tag{27}$$

$$I_t = \left\{ \left(R_t + \frac{\theta_t}{(1 + r_t)\, \varepsilon_t^*} \right) K_t - B_t \right\} [1 - F(\varepsilon_t^*)] \tag{28}$$

$$K_{t+1} = (1 - \delta) K_t + Z(\varepsilon_t^*) I_t \tag{29}$$

$$R_t = \alpha \frac{Y_t}{K_t} \tag{30}$$

$$W_t = (1 - \alpha) \frac{Y_t}{N_t} \tag{31}$$

$$Y_t = A_t K_t^\alpha N_t^{1-\alpha}, \tag{32}$$

where $Q(\varepsilon_t^*) \equiv \int \max \left\{ \frac{\varepsilon}{\varepsilon_t^*}, 1 \right\} dF(\varepsilon)$ *and* $Z(\varepsilon_t^*) \equiv \left[\int_{\varepsilon \geq \varepsilon_t^*} \varepsilon f(\varepsilon) d\varepsilon \right] [1 - F(\varepsilon_t^*)]^{-1}$.

Proof. See Appendix III. ■

Equations (21) through (24) are the household's first-order conditions; equation (25) is the aggregate resource identity derived from the household's budget constraint; equations (26) through (28) are derived from firm's decision rules based on the law of large numbers; equation (29) is the law of motion for the aggregate capital stock, where $P(\varepsilon^*)$ denotes aggregate investment efficiency; equations (30) and (31) relate firm's marginal products to factor prices; and equation (32) is the aggregate production function.

4   Macroeconomic Effects of QE

In the model, QE affects aggregate output indirectly through its direct impact on the distribution of credit/debt. In particular, QE affects aggregate output by affecting the aggregate capital stock (in the absence of a productivity change, employment is determined entirely by the capital stock, which determines the marginal product of labor). However, equation (29) shows that the level of aggregate capital stock depends on two margins: the volume of aggregate investment I_t and the average efficiency of firm-level investment $Z_t \equiv \left[\int_{\varepsilon \geq \varepsilon_t^*} \varepsilon f(\varepsilon) d\varepsilon \right] / [1 - F(\varepsilon_t^*)]$. The product $Z_t I_t$ can be called the effective aggregate investment.

QE affects both the volume and the efficiency level of investment in the economy by affecting the distribution of credit/debit in the private asset market. On the one hand, QE can lower the interest rate r and raise the market value of the firm q, thus boosting firm-level investment

along the intensive margin (see equation (19)). In addition, because of a lower borrowing cost and a lower rate of return to saving (due to a lower interest rate), more creditors self-select to become debtors, thus boosting aggregate investment along the extensive margin (i.e., the cutoff, ε_t^*, decreases). On the other hand, as more creditors self-select to become debtors, the average rate of return to investment (as measured by the aggregate efficiency Z_t) declines because these new investors (debtors) are less productive, offsetting the positive impact of investment volume on the formation of aggregate capital (see equation (29)). Thus, the effective level of aggregate capital stock can be lowered by QE if the average efficiency Z_t is reduced more than the increase in the investment volume I_t.

In other words, QE works by pushing more creditors to become debtors, which in turn increases the total demand for loans but decreases their average efficiency. In addition, since the number of voluntary lenders is reduced, the increased aggregate demand for loans can only be met by the government's supply of credit through QE (or a lump-sum tax on the consumers). When economic activities depend not only on the extent and scope of credit/debt but also on the quality of loans, such a trade-off between quantity and quality implies the following: (i) Aggregate output and employment may be *insensitive* to small-scale asset purchases even with relatively large changes in the real interest rate and asset prices. And (ii) the positive *quantitative* effect of QE on aggregate investment volume may or may not dominate their adverse *qualitative* effect on aggregate investment efficiency in the steady state. Thus, aggregate output may either increase or decrease under QE in the steady state, depending on parameter values of the model.

4.1 Calibration

Let the time period be one quarter, the time discount rate $\beta = 0.99$, the rate of capital depreciation $\delta = 0.025$, the capital income share $\alpha = 0.36$, and the inverse labor supply elasticity $\gamma = 0.5$. In the U.S., the total private debt-to-GDP ratio of nonfinancial firms doubled from 23 percent to 48 percent over the past half century. The model-implied private debt-to-output ratio is about 25 percent when $\theta = 0.1$ and about 50 percent when $\theta = 0.5$. Assume that the idiosyncratic shock ε follows the power distribution $F(\varepsilon) = \left(\frac{\varepsilon}{\varepsilon_{\max}}\right)^\eta$ with $\varepsilon \in [0, \varepsilon_{\max}]$ and $\eta > 0$. The shape parameter is set to $\eta = \frac{\bar{\varepsilon}}{\varepsilon_{\max} - \bar{\varepsilon}}$ to easily control the mean ($\bar{\varepsilon}$) and conduct mean-preserving experiments on the variance of idiosyncratic shocks by changing the upper bound $\varepsilon_{\max}$. The distribution becomes uniform when the mean $\bar{\varepsilon} = \frac{1}{2}\varepsilon_{\max}$. I choose the steady-state ratio of private asset purchase to GDP $\bar{b} = 0.4$ as the benchmark value. This

large value is chosen to make the effects of QE large enough for the qualitative analysis. With this parameterization, the positive mitigating effect of QE on output is significant in the short run. However, although QE can mitigate the negative impact of financial shocks on GDP in the short run, it can also permanently lower the level of GDP in the steady state if QE never ends. In this case, it simplifies recognition of the differential effects of anticipated exits compared with unanticipated exits on output and the optimal timing of exit. Table 1 summarizes the calibrated parameter values.

Table 1. Parameter Values

Parameter	β	δ	γ	α	$\bar{\theta}$	π	$\varepsilon_{\max}$	$\bar{\bar{\varepsilon}}$	$\bar{b}$	ρ_θ	ρ_A	$\rho_{\mathbf{p}}$
Calibration	0.99	0.025	0.5	0.36	0.5	0.03	2.0	1.0	0.4	0.9	0.95	0.9

4.2 State-Contingent QE

Three aggregate shocks are introduced into the benchmark model to evaluate the effects of the central bank's unconventional monetary policy for combating a simulated financial crisis. To this purpose, we assume (i) the debt limit θ is a stochastic process with the law of motion,

$$\log \theta_t = (1 - \rho_\theta) \log \bar{\theta} + \rho_\theta \log \theta_{t-1} + \varepsilon_{\theta t}; \tag{33}$$

(ii) total factor productivity (TFP) is a stochastic process with the law of motion,

$$\log A_t = (1 - \rho_A) \log \bar{A} + \rho_A \log A_{t-1} + \varepsilon_{At}; \tag{34}$$

and (iii) the default risk $\mathbf{P}$ is a stochastic process with the law of motion,

$$\log \mathbf{P}_t = (1 - \rho_{\mathbf{p}}) \log \bar{\mathbf{P}} + \rho_{\mathbf{p}} \log \mathbf{P}_{t-1} + \varepsilon_{\mathbf{p}t}. \tag{35}$$

We introduce the default risk shock in the following way. A firm i solves

$$V_t(i) = \max E_t \sum_{\tau=0}^{\infty} \beta^\tau \frac{\Lambda_{t+\tau}}{\Lambda_t} d_{t+\tau}(i) \tag{36}$$

subject to

$$d_t(i) \equiv R_t k_t(i) - x_t(i) + \frac{b_{t+1}(i)}{1 + r_t} - (1 - \mathbf{P}_t) b_t(i) \geq 0, \tag{37}$$

$$k_{t+1}(i) = (1 - \delta)k_t(i) + \varepsilon_t(i) i_t(i), \tag{38}$$

$$b_{t+1}(i) \leq \theta_t (1 - \mathbf{P}_t) q_t k_t(i) \tag{39}$$

$$x_t(i) \geq 0, \tag{40}$$

where $\mathbf{P}_t$ denotes systemic default risk of private debt. When $\mathbf{P}_t$ increases, each firm's existing debt level is reduced from $b_t(i)$ to $(1 - \mathbf{P}_t) b_t(i)$, which also reduces the firm's ability to pledge collateral by the factor of $(1 - \mathbf{P}_t)$. Thus, firms' ability to issue debt is severely hindered when the aggregate default risk rises. In the extreme case of a 100 percent default probability, firms are no longer able to issue debt, so the asset market shuts down and the real interest rate shoots up to infinity.

All three shocks—a negative shock to θ_t and A_t and a positive shock to $\mathbf{P}_t$—can generate financial-crisis-like effects on output, consumption, investment, and employment: They all decline sharply. The real interest rate, however, decreases under a negative shock to either the credit limit θ_t or TFP (as in the United States), but increases under a positive shock to default risk (as in Europe during the recent debt crisis). The asset price (q_t) increases under θ_t and $\mathbf{P}_t$ shocks but decreases under a TFP shock.

A state-contingent QE policy is specified as

$$\hat{B}_{t+1} = \rho_B \hat{B}_t + \sigma_x \hat{\mathbf{X}}_t, \tag{41}$$

where $\rho_B \in [0, 1]$ measures the persistence of QE, $\sigma_x = \begin{bmatrix} \sigma_\theta & \sigma_A & \sigma_P \end{bmatrix}$ is a 1×3 row vector, and $\hat{\mathbf{X}} = \begin{bmatrix} \hat{\theta}_t & \hat{A}_t & \hat{\mathbf{P}}_t \end{bmatrix}'$ is a 3×1 column vector. Notice that $\rho_B = 1$ implies that QE never ends.

4.3 Exit Strategies

Consider the following types of exit strategies:

- A one-time exit:
$$\hat{B}_{t+1} = \hat{B}_t + \sigma_x \hat{\mathbf{X}}_t \quad \text{for } t = 0, 1, ...T - 1; \tag{42}$$

$$\hat{B}_{t+1} = 0 \quad \text{for } t \geq T; \tag{43}$$

where T is the number of periods of QE.

- A gradual exit:

$$\hat{B}_{t+1} = \hat{B}_t + \sigma_x \hat{\mathbf{X}}_t, \qquad \text{for } t = 0, 1, ... T - 1; \tag{44}$$

$$\hat{B}_{t+1+j} = \frac{N-j}{N} \hat{B}_t, \qquad \text{for } t = T \text{ and } j = 1, ..., N; \tag{45}$$

$$\hat{B}_{T+N+1+h} = 0 \qquad \text{for } h \geq 0; \tag{46}$$

where T is the number of QE periods and $N \geq 1$ is the number of periods over which the exit will occur.

In both exit experiments, I consider anticipated and unanticipated exits. Figure 1 shows the impulse response functions of output, total asset prices, the real interest rate, and total asset purchases by the government to a persistent credit crunch shock (θ_t). Three different scenarios are compared for each variable: (i) no government intervention (no QE< shown by dark-blue dashed lines on the figure), (ii) no exit (red dotted lines), and (iii) a fully anticipated one-time exit after 20 periods of QE intervention (light blue-dash-dotted lines). Therefore, in each panel there are three impulse responses.
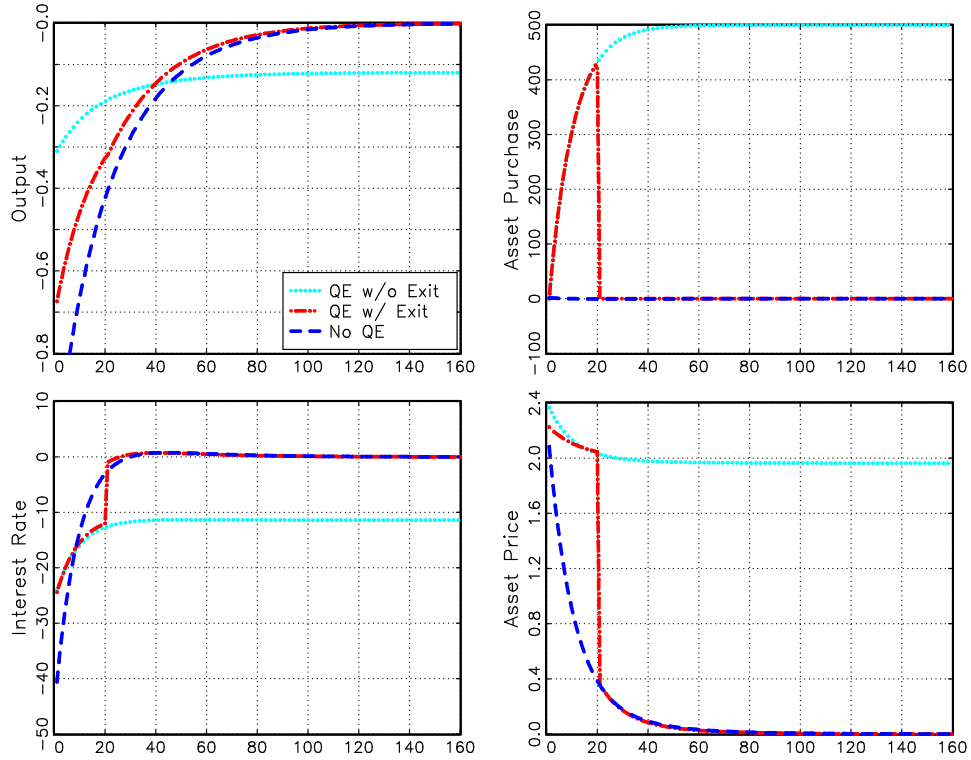

Figure 1. Impulse Response to θ_t with/without Exit Strategy.

The top-left panel shows clearly that the sharp drop in GDP under a credit crunch can be significantly mitigated by QE with or without an exit. However, the short-run mitigation

14

effect is much stronger if there is no exit. On the other hand, there would be permanent GDP losses in the long run if there is never an exit from QE. This sharp contrast between the short- and long-run effects of QE on output is the consequence of the trade-off between the positive quantity effect of QE on aggregate investment volume and the negative quality effect of QE on the average investment efficiency. Because the capital stock is relatively fixed in the short run, the quantity effect dominates in the short run because a higher total investment volume has a strong demand-side effect on the economy. In the long run, however, despite a larger volume of total investment, the average efficiency of investment is lowered by QE, leading to a lower (instead of a higher) effective capital stock. A lower capital stock in turn leads to lower labor demand and, hence, lower aggregate output. This interesting trade-off between the quantity and the quality of investment generates a dynamic trade-off between the short-run and long-run mitigating effects of QE on the economy, leading to interesting and surprising implications for the optimal timing and manner of the exit strategies under different shocks with state-contingent QE policies.

Figure 2 shows the impulse responses of the same variables as in Figure 1 (output, total asset prices, the real interest rate, and total asset purchases by the government) to a persistent TFP shock. The pattern of the output responses to the TFP shock are similar to that under a credit crunch with respect to the mitigating effects of QE policies. Notice that, as before, exiting QE in the 20th period has no visible impact on output, although such an impact is obvious in the other three variables. This is also due to the trade-off between the quantity effect and the quality effect of QE on aggregate investment.
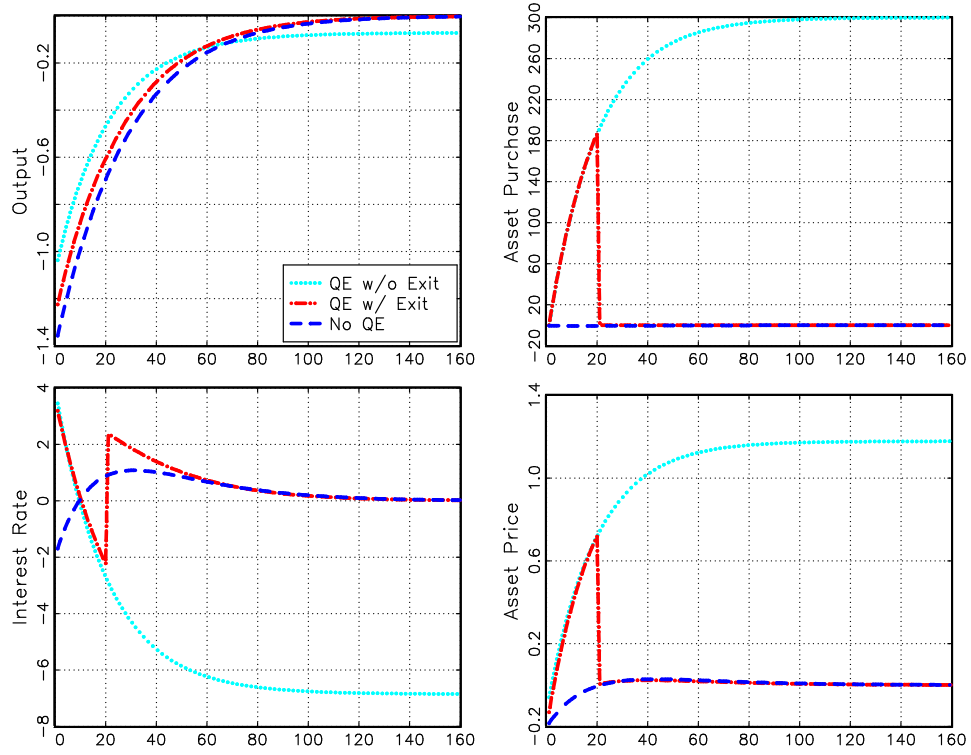
Figure 2. Impulse Response to A_t with/without Exit Strategy.

Figure 3 shows the impulse response functions of output, real interest rate, asset prices, and total asset purchases by the government to a persistent default-risk shock. As before, I compare three different scenarios. Again, the broad pattern of the output responses to default risk shock are similar to those under a credit crunch with respect to the mitigating effects of QE policies.
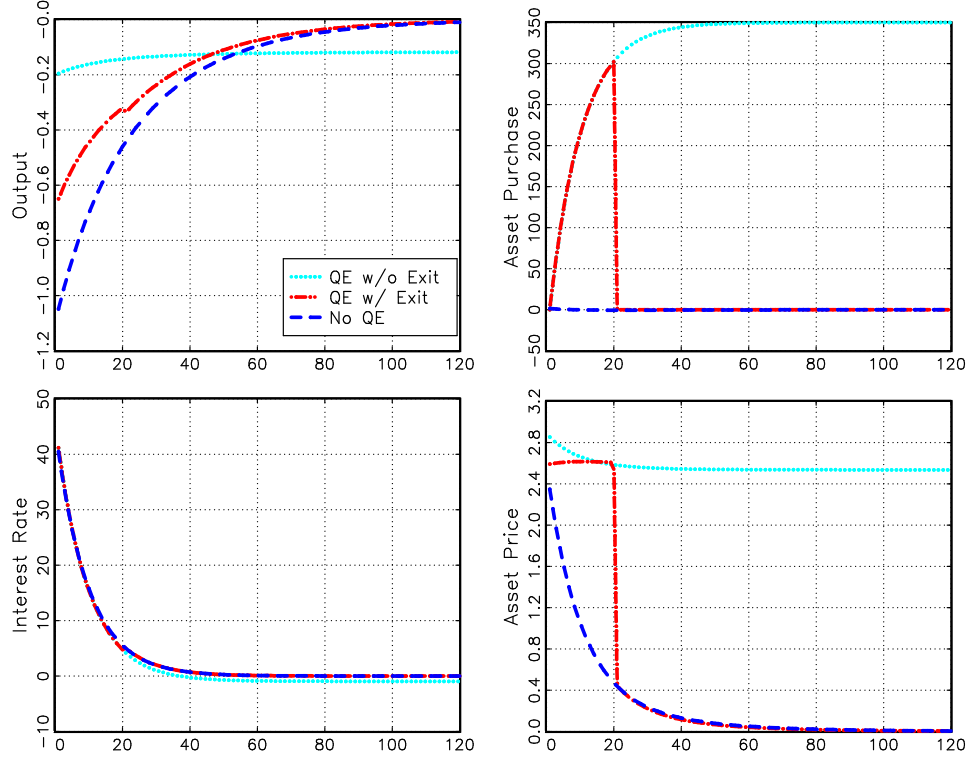
Figure 3. Impulse Response to $\mathbf{P}_t$ with/without Exit Strategy.

Figure 4 shows the differential effects of an unanticipated one-time exit and an anticipated gradual exit under a credit crunch, along with the other cases considered earlier. The solid light-blue lines in each panel shown in Figure 4 are impulse responses of different variables to θ_t shock when there is no QE. This case serves as our benchmark. Note that in this case, total asset purchases remain in the steady state (top-right panel), and GDP decreases sharply by 1 percent on impact and gradually returns to the steady state.

The red dotted lines in each panel of Figure 4 show the scenario for permanent QE no exit. In this case, total asset purchases increase permanently from 0 percent to 500 percent above the steady-state level under QE (top right panel). The GDP level drops by 0.3 percent on impact (top-left panel), showing a significant mitigating effect of QE. However, in the long run, GDP remains more than 0.1 percent below the steady state, suggesting that QE has an adverse long-run effect on output.
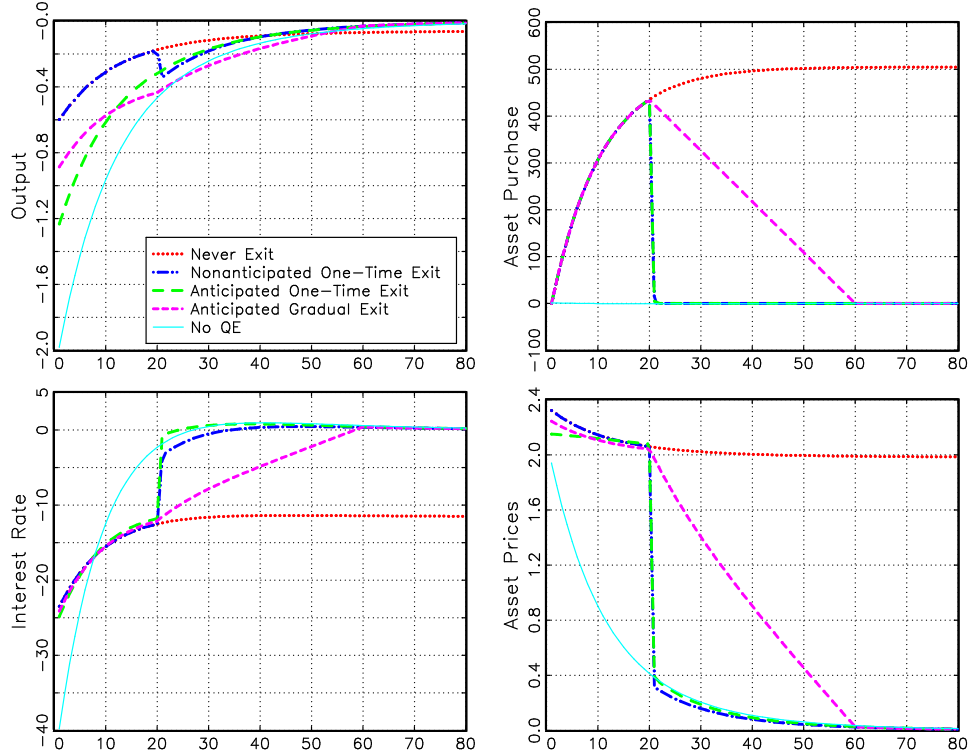
Figure 4. Effects of Different Exit Strategies.

Two of the exit scenarios consider a one-time complete exit of QE in period 20 after QE is implemented. In one case, the exit is completely unanticipated (dash-dotted dark-blue lines in Figure 4). In the other case, the exit is fully anticipated in period 0 (dashed green lines) as in Figure 1. In both cases, total asset purchases suddenly drop back to the steady state level after 20 periods (top-right panel). However, GDP (top-left panel) behaves quite differently under the two scenarios: With an unanticipated exit, the negative impact of a credit crunch on GDP is significantly mitigated in the first 20 periods (as in the case with permanent QE with no exit), because in this case agents treat QE as a permanent policy with no exit before the unanticipated exit in period 20. With an anticipated exit, the mitigating effect of QE is much weaker in the first 20 periods (albeit still stronger than the case with no QE). Because of the sudden unanticipated one-time exit of QE, output drops sharply in period 20, unlike the case with an anticipated one-time exit.

Finally, consider the scenario with an anticipated but gradual exit of QE starting in period 20, with the total exit time equal to 40 periods. The diagonal dashed line in the top-right panel shows that total asset purchases start to decline in period 20 and gradually reach zero in period 60. Because the exit of QE is gradual, QE generates a larger mitigating effect on GDP than an anticipated one-time exit in the initial 10 periods (top-left panel). However, GDP performs worse than in the case of a one-time anticipated exit afterward because the long-run adverse

18

effect of QE begins even before the exit of QE takes place. Clearly, since the exit is nonetheless finished in finite time periods, GDP does not suffer from permanent losses as in the case of permanent QE.

To summarize, these scenarios suggest that (i) QE can mitigate the negative impact of financial shocks on GDP in the short run if it is aggressive enough; (ii) there is an optimal timing to exit with respect to maximizing the mitigating effect of QE, which is around 35 to 40 periods under the current parameter configuration; (iii) an unanticipated exit works better than an anticipated exit, everything else equal; and (iv) a one-time exit is likely to work better than a gradual exit provided the timing of the exit is not too early compared with the optimal timing.[5]

5 Conclusion

Despite the popularity of QE among central banks since the financial crisis, few studies exist to explicitly model and study the macroeconomic effects of QE and its various exit strategies. This article fills this void by constructing a general equilibrium model featuring explicitly large-scale private asset purchases. I show that both the timing and the manner in which central banks unwind and reverse their asset purchase programs matter greatly for the economy. An anticipated exit that is too early can render QE ineffective in mitigating the financial crisis. On the other hand, an exit that is too late may also damage the economy because highly persistent (or permanent) QE promotes risk-taking behavior that is too intense (i.e., it encourages too many less-productive firms to undertake investment) and generates long-run inefficiency.

---

[5]An obvious future project is to mathematically design an optimal state-contingent exit strategy that can maximize the mitigating effects of QE, which is beyond the scope of this paper.

References

[1] Bernanke, Ben, 2009, "Credit Easing versus Quantitative Easing." Federalreserve.gov (13 January 2009).

[2] Wang, Pengfei and Yi Wen. "Financial Development and Economic Volatility: A Unified Explanation." Federal Reserve Bank of St. Louis Working Paper, 2009-022C.

[3] Wang, Pengfei and Yi Wen. "Speculative Bubbles and Financial Crisis." American Economic Journal: Macroeconomics, July 2012, 4(3): pp. 184-221.

[4] Wang, Pengfei and Yi Wen, "Financial Development and Long-Run Volatility Trends." Federal Reserve Bank of St. Louis Working Paper 2013-003A.

[5] Wen, Yi, "Evaluating Unconventional Monetary Policies: Why Aren't They more Effective?" Federal Reserve Bank of St. Louis Working Paper, 2013.

[6] Wen, Yi, "QE: When and How Should the Fed Exit?" Federal Reserve Bank of St. Louis Working Paper, 2014.

Appendices
Appendix I. Proof of Proposition 1

Applying the definition in equation (18), the firm's problem can be rewritten as

$$\max_{\{i_t(i), b_{t+1}(i), k_{t+1}(i)\}} E_0 \sum_{t=0}^{\infty} \beta^t \frac{\Lambda_t}{\Lambda_0} \left(R_t k_t(i) - i_t(i) + \frac{b_{t+1}(i)}{1 + r_t^c} - b_t(i) \right) \tag{47}$$

subject to

$$k_{t+1}(i) = (1 - \delta) k_t(i) + \varepsilon_t(i) i_t(i) \tag{48}$$

$$i_t(i) \geq 0 \tag{49}$$

$$i_t(i) \leq R_t k_t(i) + \frac{b_{t+1}(i)}{1 + r_t} - b_t(i) \tag{50}$$

$$b_{t+1}(i) \leq \theta_t q_t k_t(i). \tag{51}$$

Notice that if $r_t > \frac{1}{1+\pi_t} - 1$, firms do not hold money. On the other hand, if $r_t = \frac{1}{1+\pi_t} - 1$, firms are indifferent between holding private debt and money. Which case prevails depends on the steady-state supply of debt and the inflation rate (Wen, 2013). I proceed by assuming $\frac{1}{1+\pi} < 1 + r < \frac{1}{\beta}$ in equilibrium and refer readers to Wen (2013) for the other cases.

Denoting $\{\lambda_t(i), \pi_t(i), \mu_t(i), \phi_t(i)\}$ as the Lagrangian multipliers of constraints (48) through (51), respectively, the firm's first-order conditions for $\{i_t(i), k_{t+1}(i), b_{t+1}(i)\}$ are given, respectively, by

$$1 + \mu_t(i) = \varepsilon_t(i) \lambda_t(i) + \pi_t(i), \tag{52}$$

$$\lambda_t(i) = \beta E_t \frac{\Lambda_{t+1}}{\Lambda_t} \left\{ [1 + \mu_{t+1}(i)] R_{t+1} + (1 - \delta) \lambda_{t+1}(i) + \theta_{t+1} q_{t+1} \phi_{t+1}(i) \right\}, \tag{53}$$

$$\frac{1 + \mu_t(i)}{1 + r_t^c} = \beta E_t \left\{ \frac{\Lambda_{t+1}}{\Lambda_t} [1 + \mu_{t+1}(i)] \right\} + \phi_t(i). \tag{54}$$

The complementarity slackness conditions are $\pi_t(i) i_t(i) = 0$, $[R_t k_t(i) - i_t(i) + b_{t+1}(i)/(1 + r_t^c) - b_t(i)] \mu_t(i) = 0$, and $\phi_t(i)[\theta q_t k_t(i) - b_{t+1}(i)] = 0$.

Proof. Consider two possible cases for the efficiency shock $\varepsilon_t(i)$.

Case A: $\varepsilon_t(i) \geq \varepsilon_t^*$. In this case, firm i receives a favorable shock. Suppose this shock induces the firm to invest, resulting in $i_t(i) > 0$ and $\pi_t(i) = 0$. By the law of iterated expectations,

equations (52) and (53) then become

$$\frac{1 + \mu_t(i)}{\varepsilon_t(i)} = \beta E_t \frac{\Lambda_{t+1}}{\Lambda_t} \left\{ [1 + \bar{\mu}_{t+1}] R_{t+1} + (1 - \delta) \bar{\lambda}_{t+1} + \theta_{t+1} q_{t+1} \bar{\phi}_{t+1} \right\}. \tag{55}$$

Since the multiplier $\mu_t(i) \geq 0$, this equation implies

$$\varepsilon_t(i) \geq \left[\beta E_t \frac{\Lambda_{t+1}}{\Lambda_t} \left\{ [1 + \bar{\mu}_{t+1}] R_{t+1} + (1 - \delta) \bar{\lambda}_{t+1} + \theta_{t+1} q_{t+1} \bar{\phi}_{t+1} \right\} \right]^{-1} \equiv \varepsilon_t^*. \tag{56}$$

Thus, equation (53) implies $\lambda_t(i) = \frac{1}{\varepsilon_t^*}$. Since $\pi(i) = 0$, equation (52) then becomes

$$\frac{1 + \mu_t(i)}{\varepsilon_t(i)} = \frac{1}{\varepsilon_t^*}. \tag{57}$$

Hence, $\mu_t(i) > 0$ if and only if $\varepsilon_t(i) > \varepsilon_t^*$. It follows that under Case A firm i opts to invest at full capacity,

$$i_t(i) = R_t k_t(i) + \frac{b_{t+1}(i)}{1 + r_t^c} - b_t(i), \tag{58}$$

and pays no dividend. Also, since $\mu_t(i) \geq 0$, equation (54) implies

$$\phi_t(i) \geq \frac{1}{1 + r_t^c} - \beta E_t \frac{\Lambda_{t+1}}{\Lambda_t} \left\{ [1 + \bar{\mu}_{t+1}] \right\} \equiv \phi_t^*, \tag{59}$$

where the right-hand side defines the cutoff ϕ_t^*, which is independent of i. Note that $\phi_t^* \geq 0$ because it is the value of the Lagrangian multiplier when $\mu_t(i) = 0$. Hence, equation (54) can also be written as

$$\phi_t(i) = \frac{\varepsilon_t(i) - \varepsilon_t^*}{\varepsilon_t^*} \frac{1}{1 + r_t^c} + \phi_t^*. \tag{60}$$

Because $\phi_t^* \geq 0$, $\phi_t(i) > 0$ when $\varepsilon_t(i) > \varepsilon_t^*$, which means that under Case A firms are willing to borrow up to the borrowing limit $b_{t+1}(i) = \theta q_t k_t(i)$ to finance investment. Therefore, the optimal investment equation (58) can be rewritten as

$$i_t(i) = \left[R_t + \frac{\theta q_t}{1 + r_t^c} \right] k_t(i) - b_t(i). \tag{61}$$

22

Case B: $\varepsilon_t(i) < \varepsilon_t^*$. In this case, firm i receives an unfavorable shock, so the firm opts to underinvest, $i_t(i) < R_t k_t(i) + \frac{b_{t+1}}{1+r_t^c} - b_t$, then the multiplier $\mu_t(i) = 0$. Equation (52) implies that $\pi_t(i) = \frac{1}{\varepsilon_t(i)} - \frac{1}{\varepsilon_t^*} > 0$. Thus, the firm opts not to invest at all, $i_t(i) = 0$. Since $\int_0^1 b_{t+1}(i)di = 0$, and $b_{t+1}(i) = \theta q_t k_t(i) > 0$ when $\varepsilon_t(i) > \varepsilon_t^*(i)$, there must exist firms indexed by j such that $b_{t+1}(j) < 0$ if $\varepsilon_t(j) < \varepsilon_t^*$. It then follows that $\phi_t(j) = \phi_t^* = 0$ under Case B. That is, firms that receive unfavorable shocks will not invest in fixed capital but instead will opt to invest in financial assets in the bond market by lending a portion of their cash flows to other (more productive) firms.

A firm's optimal investment policy is thus given by the decision rules in Proposition 1, and the Lagrangian multipliers must satisfy:

$$
\pi_t(i) = \begin{cases} 0 & \text{if } \varepsilon_t(i) \geq \varepsilon_t^* \\[2mm] \frac{1}{\varepsilon(i)} - \frac{1}{\varepsilon^*} & \text{if } \varepsilon_t(i) < \varepsilon_t^* \end{cases}, \tag{62}
$$

$$
\mu_t(i) = \begin{cases} \frac{\varepsilon_t(i) - \varepsilon_t^*}{\varepsilon_t^*} & \text{if } \varepsilon_t(i) \geq \varepsilon_t^* \\[2mm] 0 & \text{if } \varepsilon_t(i) < \varepsilon_t^* \end{cases}, \tag{63}
$$

$$
\phi_t(i) = \begin{cases} \frac{\varepsilon_t(i) - \varepsilon_t^*}{\varepsilon_t^*} \frac{1}{1+r_t^c} & \text{if } \varepsilon_t(i) \geq \varepsilon_t^* \\[2mm] 0 & \text{if } \varepsilon_t(i) < \varepsilon_t^* \end{cases} = \frac{\mu_t(i)}{1+r_t^c}. \tag{64}
$$

Using equations (62) to (64) and equations $\lambda_t(i) = \frac{1}{\varepsilon_t^*}$ and (53), the cutoff ε_t^* can be expressed as a recursive equation:

$$
\frac{1}{\varepsilon_t^*} = \beta E_t \frac{\Lambda_{t+1}}{\Lambda_t} \left\{ R_{t+1}^L Q(\varepsilon_{t+1}^*) + \frac{\theta_{t+1}}{\varepsilon_{t+1}^*} \frac{[Q(\varepsilon_{t+1}^*) - 1]}{1+r_{t+1}^c} + \frac{(1-\delta)}{\varepsilon_{t+1}^*} \right\}, \tag{65}
$$

which determines the cutoff as a function of aggregate states only. Finally, equations (62) to (64) also imply that all the Lagrangian multipliers $\{\lambda_t(i), \pi_t(i), \mu_t(i), \phi_t(i)\}$ depend only on aggregate states and the current idiosyncratic shock $\varepsilon_t(i)$. Hence, their expected values $\{\bar{\lambda}_t, \bar{\pi}_t, \bar{\mu}_t, \bar{\phi}_t\}$ are independent of individual history and i. ∎

Appendix II. Proof of Proposition 2

Proof. Using equation (63), equation (54) can be rewritten as

$$\frac{[1 + \mu_t(i)]}{r_t^c} = \beta E_t \frac{\Lambda_{t+1}}{\Lambda_t} Q(\varepsilon_{t+1}^*) + \phi_t(i). \tag{66}$$

Evaluating this equation for firms with $\varepsilon_t(i) < \varepsilon_t^*$ yields equation (20). ∎

Appendix III. Proof of Proposition 3

Proof. By definition, the aggregate investment is $I_t \equiv \int i_t(j)dj$. Integrating equation (19) gives equation (28). The aggregate capital stock evolves according to

$$K_{t+1} = (1 - \delta)K_t + \int_{\varepsilon_t(j) \geq \varepsilon_t^*} i_t(j)\varepsilon_t(j)dj, \tag{67}$$

which by the firm's investment decision rule implies

$$K_{t+1} = (1 - \delta)K_t + I_t[1 - \mathbf{F}(\varepsilon_t^*)]^{-1} \int_{\varepsilon_t(j) \geq \varepsilon_t^*} \varepsilon_t(j)dj. \tag{68}$$

Defining $Z(\varepsilon_t^*) \equiv \left[\int_{\varepsilon \geq \varepsilon_t^*} \varepsilon d\mathbf{F}(\varepsilon) \right] [1 - \mathbf{F}(\varepsilon_t^*)]^{-1}$ as the measure of aggregate (or average) investment efficiency gives (29). Equation (18) implies $(1 - \alpha) \left[\frac{Y_t}{N_t} \right]^{1-\sigma} A_t^\sigma = w_t$. Since the capital-to-labor ratio is identical across firms, it must be true that $\frac{k(i)}{n(i)} = \frac{K}{N}$. It follows that the aggregate production function is given by $Y_t = A_t K_t^\alpha N_t^{1-\alpha}$. By the property of constant returns to scale, the defined function $R(w_t, A_t)$ in equation (18) is then the capital share, $R_t = \alpha \left(\frac{Y_t}{K_t} \right)^{1-\sigma}$, which equals the marginal product of aggregate capital. Because $\int_0^1 b_t(i)di = B_t$ and the equity share $s_{t+1}(i) = 1$ in equilibrium, the aggregate dividend and profit income are given by $D_t + \Pi_t = [Y_t - I_t - w_t N_t] + \left[\frac{B_{t+1}}{1+r_t^c} - B_t \right]$. Hence, given the government budget constraint, the household resource constraint becomes $C_t + I_t + G_t = Y_t$, as in equation (25). ∎