Resolving the Missing Deflation Puzzle*

Jesper Lindé†  Mathias Trabandt‡

September 23, 2019

Abstract

We propose a resolution of the missing deflation puzzle. Our resolution stresses the importance of nonlinearities in price- and wage-setting when the economy is exposed to large shocks. We show that a nonlinear macroeconomic model with real rigidities resolves the missing deflation puzzle, while a linearized version of the same underlying nonlinear model fails to do so. In addition, our nonlinear model reproduces the skewness of inflation and other macroeconomic variables observed in post-war U.S. data. All told, our results caution against the common practice of using linearized models to study inflation and output dynamics.

JEL Classification: E30, E31, E32, E37, E44, E52

Keywords: Great Recession, financial crisis, inflation dynamics, monetary policy, liquidity trap, zero lower bound, linearized model solution, nonlinear model solution, strategic complementarities, real rigidities, skewness of inflation.

*We are grateful for comments and suggestions by our discussants Borağan Aruoba, Edouard Challe, Nils Gornemann, Ben Johannsen and Luigi Paciello and to Gregor Boehl as well as to participants at the 2018 Federal Reserve Board of Governors – Federal Reserve Bank of New York conference on Developments in Empirical Macroeconomics, the 2018 Konstanz Seminar on Monetary Theory and Policy, 2018 Macroeconomic Model Comparison Network conference at Stanford University, the 2018 CEBRA annual research conference in Frankfurt am Main, the 7th Banca d’Italia-CEPR conference in Rome, the 15th NBER Workshop on Methods and Applications for DSGE Models at the Federal Reserve Bank of Chicago, the 2018 Northwestern – Ensai Conference in Rennes, the 2019 EACBN-NBP conference in Warsaw, and seminar participants at the Bank of Canada, Sveriges Riksbank, International Monetary Fund, Banque de France, University of Maastricht, Bank of Israel, Bank of Estonia, Verein für Socialpolitik Committee for Monetary Theory and Policy, University of Bern, Goethe University Frankfurt am Main, European Central Bank and Toulouse School of Economics. The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of Sveriges Riksbank.

†Sveriges Riksbank, Research Division, SE-103 37 Stockholm, Sweden and CEPR, E-mail: jesper.linde@riksbank.se.

‡Freie Universität Berlin, School of Business and Economics, Chair of Macroeconomics, Boltzmannstrasse 20, 14195 Berlin, Germany, German Institute for Economic Research (DIW) and Halle Institute for Economic Research (IWH), E-mail: mathias.trabandt@gmail.com.
1. Introduction

The Great Recession has left macroeconomists with many puzzles. One such puzzle is the alleged breakdown of the relationship between inflation and the output gap – also known as the Phillips curve. The Great Recession generated an extraordinary decline in U.S. GDP of about 10 percent relative to its pre-crisis trend while inflation dropped only by about 1.5 percent (see. e.g. Christiano, Eichenbaum and Trabandt, 2015). The modest decline in inflation was surprising to many macroeconomists. For instance, New York Fed President John C. Williams (2010, p. 8) wrote: “The surprise [about inflation] is that it’s fallen so little, given the depth and duration of the recent downturn. Based on the experience of past severe recessions, I would have expected inflation to fall by twice as much as it has”.

The small drop in inflation has been referred to as the “missing deflation puzzle” and has attracted considerable interest among academics and policymakers. Hall (2011) argues that since inflation fell so little in the face of the large contraction in GDP, one might view inflation as being essentially exogenous to the economy. Specifically, Hall (2011) argues that popular DSGE models based on the simple New Keynesian Phillips curve “cannot explain the stabilization of inflation at positive rates in the presence of long-lasting slack”. Similarly, Ball and Mazumder (2011) argue that Phillips curves estimated for the post-war pre-crisis period in the United States cannot explain the modest fall in inflation during the years of the financial crisis from 2008 through 2010. They argue that the fit of the standard Phillips curve deteriorates sharply during the crisis. A further challenge to the New Keynesian Phillips curve is raised by King and Watson (2012), who find a large discrepancy between actual inflation and inflation predicted by the workhorse Smets and Wouters (2007) model. Interestingly, all above contributions use linearized variants of the New Keynesian Phillips curve in their analysis.

Our proposed resolution of the missing deflation puzzle rests on two key elements. First, we argue that it is key to introduce real rigidities in price- and wage-setting. To do this, we follow Dotsey and King (2005) and Smets and Wouters (2007) and use the Kimball (1995) aggregator instead of the standard Dixit-Stiglitz (1977) aggregator. The Kimball aggregator introduces additional strategic complementarities in the price- and wage-setting, which lowers the sensitivity of prices and wages to the relevant wedges for a given degree of price- and wage-stickiness. As such, the Kimball aggregator is commonly used in New Keynesian models, see e.g. Smets and Wouters (2007), as it allows to simultaneously account for the macroeconomic evidence of a low Phillips curve slope and the microeconomic evidence of frequent price changes.
Second, we argue that the standard procedure of linearizing all equilibrium equations around the steady state, except for the zero lower bound (ZLB) constraint on the nominal interest rate, introduces large approximation errors when large shocks hit the economy as was the case during the Great Recession. Implicit in the linearization procedure is the assumption that the linearized solution is accurate even far away from the steady state. Our analysis shows that the linearized solution is very inaccurate far away from the steady state. We show that one ought to use the nonlinear model solution instead of the linearized model solution to understand the output-inflation dynamics during the Great Recession. We show that the nonlinearity implied by the Kimball aggregator is a key model feature that accounts for the differences between the linearized and nonlinear model solutions: it is crucial to resolve the missing deflation puzzle. The Kimball aggregator implies that the demand elasticity for intermediate goods is state-dependent, i.e. firms’ demand elasticity is an increasing function of its relative price and the demand curve is quasi-kinked.\(^1\) While the fully nonlinear model takes the state-dependence of the quasi-kinked demand curve explicitly into account, a linear approximation replaces this key nonlinearity by a linear function. When the economy is exposed to large shocks the state-dependence of the quasi-kinked demand curve becomes quantitatively important and the linear approximation ceases to provide accurate results.

A key contribution of our work is that we provide a structural general equilibrium model which can account for the alleged breakdown of the Phillips curve during the Great Recession.\(^2\) That is, a nonlinear formulation of our model generates a very modest fall of inflation in the wake of a large and persistent contraction of output. By contrast, the linearized version of the same underlying nonlinear model predicts a much larger fall of inflation and thereby fails to account for the output-inflation dynamics in the data. Therefore, our model provides a resolution to the missing deflation puzzle: the puzzle arises only if one uses the linearized solution to predict inflation in the Great Recession. Section 2.1.2 provides the intuition and Figure 4 illustrates this key result graphically. In a nutshell, our model implies nonlinear price and wage Phillips curves, i.e. nonlinear relationships between price and wage inflation and the output gap. The slopes of our price and wage Phillips curves are notably flatter in recessions than in booms, i.e. our Phillips curves have a banana- or boomerang-shape as in the seminal paper by Phillips (1958). Consequently and consistent with the data, our nonlinear model implies that inflation falls relatively little in deep recessions. By contrast, the linearized model erroneously predicts a much larger fall in inflation by extrapolating

\(^1\) See Dupraz (2018) for a microfoundation of a kinked demand curve theory.
\(^2\) Thus, we provide a structural interpretation of the emerging body of empirical literature that provides macroevidence in favor of nonlinearities in the Phillips curve, see for instance Doser et al. (2018), Gagnon and Collins (2019) and the references therein.
local decision rules around the steady state to deep recessions far away from the steady state.\(^3\)

We establish our main results in a variant of the Erceg, Henderson and Levin (2000) (EHL henceforth) benchmark model. The EHL model, however, does not allow for endogenous capital accumulation and other real rigidities like habit formation in consumer preferences and investment adjustment costs. Therefore, we examine the robustness of our results in an estimated medium-sized New Keynesian model based on Christiano, Eichenbaum and Evans (2005) and Smets and Wouters (2007) which includes the features listed above. We estimate this extended model using Bayesian maximum likelihood techniques. We show that the insights from the EHL model carry over to the estimated medium-sized model. In addition, by performing stochastic simulations with the estimated nonlinear model, we also establish two further important insights.

First, a large body of recent work – see for instance IMF (2016, 2017) – has been aimed towards understanding the absence of upward pressure of price and wage inflation during the recovery from the Great Recession, also called the missing inflation puzzle. The nonlinear formulation of our model offers an explanation for this phenomenon. Given the nonlinear Phillips curves in our model, price and wage inflation remain subdued until the level of economic activity has recovered sufficiently relative to its potential. Put differently, even though economic growth may resume after a deep recession, price and wage inflation will only increase modestly until economic slack has subsided sufficiently. Therefore, our model does not only resolve the missing deflation puzzle, it also offers an explanation for the missing inflation puzzle. According to our model, subdued price and wage inflationary pressures as observed during the recovery from the Great Recession are indicative of substantial economic slack.\(^4\) While the linearized model suggests that price and wage inflation should rise sharply during a recovery from a deep recession, the nonlinear model suggests otherwise: price and wage inflation will not rise until economic slack has narrowed sufficiently from its elevated recession level.

Second, our nonlinear model can be used to understand the positive skewness in post-war U.S. price inflation, i.e. that price inflation scares are much more common than deflationary episodes. We establish this result by comparing the skewness for inflation in the data and our model. Consistent with the data, our nonlinear model implies positive skewness in inflation. In addition, our nonlinear model also generates positive skewness for wage inflation as well as for the nominal policy interest rate as observed in the data. By contrast, the linearized version of the model fails to account for

\(^3\) Recent work by Boneva, Braun, and Waki (2016) and Lindé and Trabandt (2018) suggests that linearization can also produce misleading results about the effects of fiscal policy.

\(^4\) See e.g. Erceg and Levin (2014) who argue that economic slack was bigger than commonly thought following the Great Recession.
the skewness of these time series. All told, our nonlinear model allows us to simultaneously explain the quantitative inflation performance prior to the Great Recession, the missing deflation during the Great Recession, as well as the missing inflation in recent years.

The remainder of the paper is organized as follows. Section 2 presents the stylized New Keynesian model with stickiness and real rigidities in price- and wage-setting. Section 3 discusses the results based on the stylized model. Section 4 examines the robustness of our results in an estimated medium-sized New Keynesian model. Section 5 discusses related literature on micro- and macroeconomic evidence. Section 6 provides concluding remarks.

2. A Stylized New Keynesian Model

The simple model we use is very similar to the EHL model with gradual price and wage adjustment. We deviate from EHL in two ways. First, by allowing for Kimball (1995) aggregators in price- and wage-setting (with the standard Dixit and Stiglitz (1977) specification used by EHL as a special case). Second, by including a discount factor shock (or more generally a savings shock) into the model.

2.1. Model

2.1.1. Firms and Price Setting

Final Goods Production. Competitive firms use intermediate goods \( Y_{f,t} \) to produce final goods \( Y_t \) using the production technology \( \int_0^1 G_Y \left( \frac{Y_{f,t}}{Y_t} \right) \, df = 1 \). Following Dotsey and King (2005) and Levin, Lopez-Salido and Yun (2007), we assume that \( G_Y (\cdot) \) is given by the following concave and increasing function:

\[
G_Y \left( \frac{Y_{f,t}}{Y_t} \right) = \frac{\omega_p}{1 + \psi_p} \left( \frac{1}{1 + \psi_p} \left( 1 + \psi_p \right) \frac{Y_{f,t}}{Y_t} - \psi_p \right)^{-\frac{1}{\psi_p}} + 1 - \frac{\omega_p}{1 + \psi_p}, \tag{1}
\]

where \( \omega_p = \frac{\phi_p (1 + \psi_p)}{1 + \phi_p \psi_p} \). \( \phi_p > 1 \) denotes the gross markup of intermediate goods firms. The parameter \( \psi_p \leq 0 \) governs the degree of curvature of the intermediate firm’s demand curve. If \( \psi_p = 0 \), \( G_Y (\cdot) \) is the standard Dixit and Stiglitz (1977) aggregator. If \( \psi_p < 0 \), \( G_Y (\cdot) \) is the Kimball (1995) aggregator as used in e.g. Smets and Wouters (2007). Note that \( G_Y (1) = 1 \), implying constant returns to scale when all intermediate firms produce the same amount.

---

5 The complete specification of the nonlinear and linearized formulation of the model is provided in the Technical Appendix available at: https://sites.google.com/site/mathiastrabandt/home/downloads/LindeTrabandt_Inflation_TechApp.pdf

6 The parameter used in Smets and Wouters (2007) to characterize the curvature of the Kimball aggregator can be mapped to our model using the following formula: \( \epsilon_p = -\frac{\phi_p}{\phi_p + 1} \psi_p \).
Final goods producers minimize the cost of producing a given quantity of output $Y_t$ taking as given the price $P_{f,t}$ of each intermediate good $Y_{f,t}$. Final goods producers sell units of final output at a price $P_t$ which they also take as given.

Formally, producers solve: $\max_{Y_t,Y_{f,t}} P_t Y_t - \int_0^1 P_{f,t} Y_{f,t} df$ subject to $\int_0^1 G_Y \left( \frac{Y_{f,t}}{Y_t} \right) df = 1$ where $G_Y(\cdot)$ is given by equation (1). The first order conditions can be written as:

$$\frac{Y_{f,t}}{Y_t} = \frac{1}{1+\psi_p} \left( \frac{P_{f,t}}{P_t \vartheta_t^p} \right)^{\varphi_p} + \frac{\psi_p}{1+\psi_p} \quad (2)$$

$$1 = \int_0^1 \left( \frac{P_{f,t}}{P_t \vartheta_t^p} \right)^{\varphi_p} df \quad (3)$$

$$\vartheta_t^p = 1 + \psi_p - \psi_p \int_0^1 \frac{P_{f,t}}{P_t \vartheta_t^p} df \quad (4)$$

where $\varepsilon_p = \frac{\phi_p(1+\psi_p)}{1-\phi_p}$ and $\vartheta_t^p$ denotes the Lagrange multiplier on the aggregator constraint (1). Equation (2) denotes the demand for goods, equation (3) is the aggregate price index and equation (4) is the zero profit condition for final goods firms. Note that for $\psi_p = 0$ we obtain the standard Dixit-Stiglitz demand equation and aggregate price index.

**Intermediate Goods Production.** A continuum of intermediate goods $Y_{f,t}$ for $f \in [0,1]$ is produced by monopolistically competitive firms, each of which produces a single differentiated good. Each intermediate good producer faces the demand curve derived above.

Production of intermediate good firms is given by $Y_{f,t} = K_f^{\alpha} N_t^{1-\alpha}$. Intermediate good firms rent capital and labor services from economy-wide factor markets. Aggregate capital, $K$, is assumed to be fixed. Even though the aggregate capital stock is fixed, shares of it can be freely allocated across intermediate good firms, i.e. $K \equiv \int K_f df$. Given this, real marginal cost, $MC_{f,t}/P_t$ is identical across firms and equal to $MC_{f,t}/P_t \equiv \frac{W_t/P_t}{MPL_t} = \frac{W_t/P_t}{(1-\alpha)(K/N_t)}$. We discuss the determination of the aggregate labor-index $N_t$ in Section 2.1.3.

Prices of intermediate goods are determined by Calvo-Yun (1996) style staggered nominal contracts. In each period, each firm $f$ faces a probability $1 - \xi_p$ of being able to choose its profit maximizing price. Firms solve:

$$\max_{P_{f,t}} \mathbb{E}_t \sum_{s=0}^{\infty} \left( \beta \xi_p \right)^s \varsigma_{t+s} \Lambda_{t+s} \left( \Pi^s P_{f,t}^{\text{opt}} - MC_{t+s} \right) Y_{f,t+s} \quad (5)$$

where $\Lambda_{t+s}$ is household marginal utility, $\varsigma_{t+s}$ is a household discount factor shock discussed in section 2.1.3 and demand $Y_{f,t+s}$ is given by equation (2).

If a firm does not optimize its price in a given period, it updates its price according to $\tilde{P}_t = \Pi P_{t-1}$, where $\Pi$ is the steady-state gross inflation rate, i.e. $\Pi = 1 + \bar{\pi}$ where $\bar{\pi}$ the is net steady state price inflation rate.
2.1.2. Intuition

To provide intuition about the implications of the Kimball vs. Dixit-Stiglitz aggregator, we consider the flexible price version of the price setting problem of firms.

The upper left panel in Figure 1 shows how relative demand in equation (2) is affected by the relative price under alternative assumptions about $\psi_p$, given a gross markup of $\phi_p = 1.1$ and assuming that $\delta^p_t = 1$. When $\psi_p = 0$, the demand curve exhibits the standard Dixit-Stiglitz constant demand elasticity, implying a log-linear relationship between relative demand and relative price, i.e. $\log y_f = \varepsilon_p \log p_f$ where $y_f \equiv Y_f/Y$ and $p_f \equiv P_f/P$. With the Kimball specification $\psi_p < 0$ – as in e.g. Smets and Wouters (2007) – a firm instead faces a quasi-kinked demand curve, implying that a drop in its relative price triggers only a small increase in demand. On the other hand, a rise in its relative price generates a larger fall in demand compared to the $\psi_p = 0$ case. Relative to the standard Dixit-Stiglitz aggregator, this introduces more strategic complementarities (or real rigidities) in price setting.

The upper right panel of Figure 1 depicts the period-by-period profit function $\Phi_f(p_f) = (p_f - mc) y_f(p_f)$ of intermediate good firm $f$ taking into account relative demand (equation 2). More strategic complementarities in the Kimball specification imply that the profit function $\Phi_f(p_f)$ becomes notably more curved when $\psi_p < 0$, especially when the firm’s relative price exceeds unity.

The two middle panels in Figure 1 show how 5 percent positive and negative changes in marginal cost affect firm profits $\Phi_f(p_f)$ and the optimal profit maximizing relative price $p^*_f$. The left panel shows the effects for the Kimball specification and the right panel shows the effects for the Dixit-Stiglitz specification. The profit function shifts up when marginal cost decline and shifts down when marginal cost rise.

In the Dixit-Stiglitz case ($\psi_p = 0$), the optimal relative price is changed by exactly the same percent change in marginal costs. This result reflects the optimal markup pricing rule implied by maximizing profits subject to the demand curve:

\[
\text{Dixit-Stiglitz (nonlinear): } p^*_f = \phi_p mc.
\]

Log-linearizing around steady state yields:

\[
\text{Dixit-Stiglitz (log-linear): } \frac{p^*_f - \bar{p}^*_f}{\bar{p}^*_f} = \frac{mc - \bar{mc}}{\bar{mc}}
\]

where ‘bar’ variables denote steady state variables. Note that the nonlinear and log-linearized optimal pricing rules under Dixit-Stiglitz have the same implications: the optimal relative price changes one to one – in percent terms – with marginal cost.
With the Kimball specification ($\psi_p < 0$), the results can differ sharply from the Dixit-Stiglitz case. Maximizing profits subject to the demand curve yields:

$$\text{Kimball (nonlinear): } p_f^{\text{opt}} \left(1 - \frac{\theta_p \psi_p}{1 + \psi_p + \theta_p \psi_p} \left(1 + \frac{\theta_p}{\psi_p} \right) \right) = \frac{(1 + \psi_p)(1 + \theta_p)}{1 + \psi_p + \theta_p \psi_p} \cdot mc. \quad (6)$$

where $\theta_p$ denotes the net price markup, i.e. $\phi_p = 1 + \theta_p$. The left middle panel in Figure 1 shows that a negative shock to $mc$ implies a smaller change in $p_f^{\text{opt}}$ in absolute value (decline from 1 to 0.997) than a positive shock to $mc$ (increase from 1 to 1.0045) in the nonlinear framework. The asymmetric response of $p_f^{\text{opt}}$ to negative and positive shocks to $mc$ in the nonlinear model is apparent from equation (6). The equation implies that the asymmetry will be stronger the smaller (i.e. negative) $\psi_p$ and the larger the change in $mc$.

What is the intuition for the asymmetric response of $p_f^{\text{opt}}$ when $\psi_p < 0$? The nonlinearity embedded in the Kimball aggregator and the resulting nonlinear demand curve are key to understand the asymmetric price response. Consider a fall in marginal cost, i.e. a recession. Recall the profit function $\Phi_f(p_f) = (p_f - mc) y_f(p_f)$. Keep in mind that the demand elasticity falls when the relative price of a firm falls. Thus, at the margin, a firm’s ability to increase its demand $y_f(p_f)$ by cutting its price is limited, especially the deeper the recession. Large price cuts result in lower profits because at the margin, demand $y_f(p_f)$ rises only by little while revenues $p_f y_f(p_f)$ fall substantially. Therefore, firms have little incentive to cut prices a lot. The incentive to cut prices becomes weaker the larger the fall in marginal cost since the lower relative price of a firm reduces the demand elasticity implying less demand can be crowded-in by the firm when cutting its price. All told, firms cut their price by less at the margin the deeper the recession, i.e. an asymmetric price response. By continuity, firms change their prices by more at the margin the larger the boom in the economy.

Log-linearizing equation (6) around steady state yields:

$$\text{Kimball (log-linear): } \frac{p_f^{\text{opt}} - \bar{p}_f^{\text{opt}}}{\bar{p}_f^{\text{opt}}} = \frac{1}{1 - (1 + \theta_p)} \frac{mc - \bar{mc}}{mc}$$

Note that in contrast to Dixit-Stiglitz, the optimal relative price does not move one to one in percent terms with marginal costs under the Kimball specification. With Kimball, firm’s adjust their prices less than with Dixit-Stiglitz since $\frac{1}{1 - (1 + \theta_p) \psi_p} < 1$ when $\psi_p < 0$.

How different are the nonlinear and log-linearized pricing rules implied by Kimball vs. Dixit-Stiglitz quantitatively? The bottom left panel of Figure 1 plots the optimal relative price $p_f^{\text{opt}}$ as a function of marginal cost in the nonlinear vs. log-linear Kimball specification. The bottom right
panel shows the Dixit-Stiglitz case. Several important results emerge. First, the nonlinear Kimball specification generates substantial differences between the nonlinear and log-linearized pricing rules. The nonlinear pricing rule implies a banana- or boomerang type pricing schedule: small price adjustments following deep recessions (declines in $mc$) and large price adjustments following big booms (increases in $mc$). Second, the nonlinearity becomes stronger the deeper the recession or the bigger the boom. Hence, linearization can produce large approximation errors. Third, the slope of the linearized optimal pricing schedule under Kimball is noticeably flatter compared to the Dixit-Stiglitz case. Fourth, the nonlinear and log-linear pricing schedules under Dixit-Stiglitz are identical. Fifth and finally, when prices are flexible, Kimball implies a countercyclical markup while Dixit-Stiglitz implies a constant markup.

2.1.3. Households and Wage Setting

Labor Contractors. Competitive labor contractors aggregate specialized labor inputs $N_{j,t}$ supplied by households into homogenous labor $N_t$ which is hired by intermediate good producers. Labor contractors maximize profits

$$\max_{N_{j,t},N_t} W_t N_t - \int_0^1 W_{j,t} N_{j,t} dj$$

where $W_{j,t}$ is the wage paid by the labor contractor to households for supplying type $j$ labor. $W_t$ denotes the wage paid to the labor contractor for homogenous labor. Profit maximization is subject to

$$R \geq N_{j,t} N_t = 1$$

where

$$G_N \left( \frac{N_{j,t}}{N_t} \right) = \frac{\omega_w}{1 + \psi_w} \left( (1 + \psi_w) \frac{N_{j,t}}{N_t} - \psi_w \right) + 1 - \frac{\omega_w}{1 + \psi_w}$$

is the Kimball aggregator specification as used in Dotsey and King (2005) or Levin, Lopez-Salido and Yun (2007) adapted for the labor market. Note that $\omega_w = \frac{(1 + \psi_w) \phi_w}{1 + \phi_w \psi_w}$ where $\phi_w \geq 1$ denotes the gross wage markup and $\psi_w \leq 0$ is the Kimball parameter that controls the degree of complementarities in wage-setting.

Let $\partial_{\psi}^w$ denote the multiplier on the labor contractor’s constraint. Optimization results in the following optimality conditions:

$$\frac{N_{j,t}}{N_t} = \frac{1}{1 + \psi_w} \left( \frac{W_{j,t}}{W_t \partial_{\psi}^w} \right) \varepsilon_w + \frac{\psi_w}{1 + \psi_w}$$

$$1 = \int_0^1 \left( \frac{W_{j,t}}{W_t \partial_{\psi}^w} \right) \varepsilon_w \, dj$$

$$\partial_{\psi}^w = 1 + \psi_w - \psi_w \int_0^1 \frac{W_{j,t}}{W_t} \, dj$$

where $\varepsilon_w = \frac{\phi_w (1 + \psi_w)}{1 - \phi_w}$. Equation (7) denotes the demand for labor, equation (8) is the aggregate wage index and equation (9) is the zero profit condition for labor contractors. Note that for $\psi_w = 0$ we get the standard Dixit-Stiglitz demand curve and aggregate wage index.
Households. There is a continuum of households $j \in [0,1]$ in the economy. Each household supplies a specialized type of labor service $j$ to the labor market, for which it is a monopolistic supplier. The $j^{th}$ household maximizes

$$E_0 \sum_{t=0}^{\infty} \beta^t \zeta_t \left( \ln C_{j,t} - \omega \frac{N_{j,t}^{1+\chi}}{1+\chi} \right)$$

subject to

$$P_t C_{j,t} + B_{j,t} = W_{j,t} N_{j,t} + R^k_t K_j + (1 + i_{t-1}) B_{j,t-1} - T_{j,t} + \Gamma_{j,t} + A_{j,t}$$

where the choice variables of the $j^{th}$ household are consumption $C_{j,t}$ and risk-free government debt $B_t$. The $j^{th}$ household also chooses the wage $W_{j,t}$ subject to Calvo sticky prices as in EHL. The household understands that when choosing $W_{j,t}$ it must supply the amount of labor $N_{j,t}$ demanded by a labor contractor according to equation (7).

The variable $\zeta_t$ is an exogenous shock to the discount factor $0 < \beta < 1$. We assume that $\delta_t = \frac{\delta_t}{\delta_{t-1}}$ is exogenous with $\delta = 1$ in steady state. $P_t$ denotes the aggregate price level. $i_t$ denotes the net nominal interest rate on bonds purchased in period $t - 1$ which pay off in period $t$. $R^k_t$ is the rental rate of the fixed aggregate capital stock that the households rents to goods producing firms. $T_{j,t}$ are lump-sum taxes net of transfers and $\Gamma_{j,t}$ denotes the share of profits that the household receives. $A_{j,t}$ denotes payments and receipts associated with insurance in the presence of wage stickiness. $\omega > 0$ and $\chi \geq 0$ are parameters.

Utility maximization for consumption and government bond holdings yields the standard aggregate consumption Euler equation:

$$1 = \beta E_t \left( \delta_{t+1} \frac{1 + i_t}{1 + \pi_{t+1}} \frac{C_t}{C_{t+1}} \right)$$

where $1 + \pi_{t+1} = P_{t+1}/P_t$.

Wage Setting. Households face a standard monopoly (labor union) problem of selecting $W_{j,t}$ to maximize utility (10) subject to the demand for labor (7). Following EHL, we assume that households experience Calvo-style frictions in its choice of $W_{j,t}$. In particular, with probability $1 - \xi_w$ the $j^{th}$ household has the opportunity to choose its optimal wage $W_{j,t}^{\text{opt}}$. With the complementary probability, the household must set its wage rate according to $\tilde{W}_{j,t} = \Pi_w W_{j,t-1}$ where $\Pi_w$ denotes

\footnote{In principle, the presence of wage setting frictions implies that households have idiosyncratic levels of wealth and, hence, consumption. However, we follow EHL in supposing that each household has access to perfect consumption insurance. Because of the additive separability of the household utility function, perfect consumption insurance implies equal consumption across households. Note that even though consumption is equal across households, aggregate consumption responds to shocks and is hence not constant over time.}
the steady state gross rate of wage inflation. Households choose the optimal wage by maximizing

$$\max_{W_{j,t}^\text{opt}} \sum_{s=0}^{\infty} (\beta \xi_w)^s E_t \left( \Lambda_{t+s} W_{j,t}^\text{opt} \Pi_s^w N_{j,t+s} - \omega \frac{N_{j,t+s}^{1+\chi}}{1+\chi} \right)$$

subject to labor demand given by equation (7).

2.1.4. Monetary and Fiscal Policy

We assume that the central bank sets the nominal interest rate following a Taylor-type policy rule that is subject to the zero lower bound:

$$1 + i_t = \max \left\{ 1, (1 + \pi_t) \left( \frac{1 + \pi_t}{1 + \bar{\pi}} \right)^{\gamma_p} \left( \frac{Y_t}{Y_t^{\text{pot}}} \right)^{\gamma_z} \right\}$$

where $Y_t^{\text{pot}}$ denotes the level of output that would prevail if prices and wages were flexible.

In terms of fiscal policy, we assume that the government balances its budget each period using lump-sum taxes.

2.1.5. Aggregate Resource Constraint

It is straightforward to show that the aggregate resource constraint is given by

$$C_t = Y_t = (p_t^*)^{-1} (w_t^*)^{-(1-\alpha)} K_t^{\alpha} l_t^{1-\alpha}$$

where

$$p_t^* = \frac{1}{1+\psi_p} \int_0^1 \left( \frac{P_{t,j}}{P_{t,j}^{\text{opt}}} \right)^{-v_p} df + \frac{\psi_p}{1+\psi_p}$$

$$w_t^* = \frac{1}{1+\psi_w} \int_0^1 \left( \frac{W_{t,j}}{W_{t,j}^{\text{opt}}} \right)^{-v_w} dj + \frac{\psi_w}{1+\psi_w}.$$ Aggregate capital $K$ is given by $K = \int K_j df = \int K_j dj$. Aggregate hours per capita $l_t$ supplied by households is given by $l_t = \int N_{j,t} dj = w_t^* N_t$ where $N_t = \int N_{j,t} df$. The variables $p_t^*$ and $w_t^*$ denote the Yun (1996) aggregate price and wage dispersion terms. Both price- and wage-dispersion, ceteris paribus, will lower aggregate output in the economy. The Technical Appendix provides recursive formulations for the sticky price and wage distortion terms $p_t^*$ and $w_t^*$. Note that $p_t^*$ and $w_t^*$ vanish when the model is linearized.

2.2. Parameterization

Time is taken to be quarters. We set the discount factor $\beta = 0.9975$ and the steady state net inflation rate $\pi = 0.005$ implying a steady state nominal interest rate of $i = 0.0075$ (i.e., three
percent at an annualized rate). We set the capital share parameter $\alpha = 0.3$ and the disutility of labor parameter $\chi = 0$. As a compromise between the low estimate of the gross price markup $\phi_p$ in Altig et al. (2011) and the higher estimated value by Smets and Wouters (2007), we set $\phi_p = 1.1$.

Consistent with microeconomic evidence in e.g. Nakamura and Steinsson (2008) we set $\xi_p = 0.66$, i.e. prices change on average every three quarters. To pin down the Kimball parameter $\psi_p$ consider the log-linearized New Keynesian Phillips curve in our model: $\hat{\pi}_t = \beta E_t \hat{\pi}_{t+1} + \kappa_p \delta c_t$, where $\delta c_t$ denotes the log-deviation of real marginal cost from its steady state. $\hat{\pi}_t$ denotes the log-deviation of gross inflation from its steady state. The parameter $\kappa_p$ denotes the slope of the Phillips curve and is given by $\kappa_p \equiv (1-\xi_p)(1-\beta \xi_p)\frac{1}{\xi_p \phi_p}$.

The macroeconomic evidence suggest that the sensitivity of aggregate inflation to variations in real marginal cost is very low, see e.g. Altig et al. (2011). To capture this, we set the Kimball parameter $\psi_p = -12.2$ so that the slope of the Phillips curve is $\kappa_p = 0.012$ given the values for $\beta$, $\xi_p$ and $\phi_p$ discussed above.8

For the parameters pertaining to the nominal wage setting frictions we assume that $\phi_w = 1.1$, $\xi_w = 0.75$, and $\psi_w = -6$. These parameter values correspond roughly to those set and estimated in the medium-sized New Keynesian model discussed below. We use the standard Taylor (1993) rule parameters $\gamma_\pi = 1.5$ and $\gamma_x = 0.125$.

To facilitate comparison between the nonlinear and linearized model, we specify an AR(1) process for the discount factor so that there is no loss in precision due to an approximation:

$$\delta_t - \delta = \rho_\delta (\delta_{t-1} - \delta) + \sigma_\delta \varepsilon_{\delta,t} \quad (13)$$

where $\delta = 1$. Our baseline parameterization adopts a persistence coefficient $\rho_\delta = 0.95$ capturing the persistent drop in output during the Great Recession.

2.3. Solving the Model

Our baseline results for the linearized and nonlinear model are based on the Fair and Taylor (1983) solution algorithm. This solution algorithm is also known as a two-point boundary value solution algorithm or time-stacking algorithm. The Fair-Taylor solution algorithm imposes certainty equivalence on the nonlinear model, just as the linearized model solution does by definition. In other words, the Fair-Taylor solution algorithm allows us to trace out the implications of not linearizing the equilibrium equations which is exactly our objective. All relevant information for solving the

---

8 The median estimates of the Phillips curve slope in pre-Great Recession empirical studies by e.g. Adolfson et al. (2005), Altig et al. (2011), Gali and Gertler (1999), Gali, Gertler and López-Salido (2001), Lindé (2005), and Smets and Wouters (2003, 2007) are in the range of 0.009 – 0.014.
nonlinear and linearized model is captured by the current state of the economy, including the contemporaneous realization of the exogenous discount factor shock.

An alternative approach to solve the model would have been to compute solutions where uncertainty about future shock realizations matters for the dynamics of the economy following e.g. Aruoba, Cuba-Borda, and Frank Schorfheide (2018), Adam and Billi (2006, 2007), Fernández-Villaverde et al. (2015), Gust, Herbst, López-Sávido and Smith (2017) and Nakata (2017). These authors have shown that allowing for future shock uncertainty can have potentially important implications for equilibrium dynamics. Importantly, none of these authors have considered a model with Kimball aggregation. Lindé and Trabandt (2018) solve a simplified version of our model with sticky prices and Kimball aggregation under shock uncertainty using global methods, and show that the effects of future shock uncertainty on the global solution of the nonlinear model are quantitatively negligible lending support for using the Fair-Taylor solution method for our baseline results.9

As a practical matter, we feed the equilibrium equations of the nonlinear and linearized model into Dynare. Dynare is a pre-processor and a collection of MATLAB routines which can solve nonlinear and linearized dynamic models with forward looking variables. The details about the implementation of the algorithm used can be found in Juillard (1996). We use the perfect foresight/deterministic simulation algorithm implemented in Dynare using the ‘simul’ command.10 The algorithm can also easily handle the ZLB constraint: one just writes the Taylor rule including the max operator in the model equations, and the solution algorithm reliably calculates the model solution in fractions of a second. Thus, apart from obtaining intuition about the various mechanisms embedded in dynamic models, there is no need anymore to linearize dynamic models to solve, simulate and work with them.

3. Inflation and Output Dynamics in the Stylized Model

In this section, we report our main results for the linearized and nonlinear solution of the model outlined in the previous section. In section 3.1, we study the joint output-inflation dynamics for large adverse demand shocks, and in section 3.2 we consider the effects of both positive and negative shocks in long simulations of the model.

9 The introduction of wage stickiness and Kimball aggregation in the labor market in the present paper (in addition to price stickiness and Kimball aggregation in the goods market as in Lindé and Trabandt, 2018) should temper the effect of shock uncertainty in the nonlinear model even further. To the extent that allowing for shock uncertainty impacts notably the linearized solution, the differences between the linearized and nonlinear solutions we report in this paper are conservative; they would be even larger if we had allowed for shock uncertainty.

10 The solution algorithm implemented in Dynare’s simul command is the method developed in Fair and Taylor (1983).


3.1. A Recession Scenario

We first study the effects of a large adverse demand shock. Following the literature on fiscal multipliers (e.g. Christiano, Eichenbaum and Rebelo, 2011), the particular shock we consider is a large positive shock to the discount factor $\delta_t$. Specifically, we assume that $\varepsilon_{\delta,1} = 0.01$ in (13) so that $\delta_t$ increases from 1 to 1.01 in the first period and then gradually reverts back to steady state.

Figure 2 reports the linear and nonlinear solutions for a selected set of variables, assuming that the economy is in the deterministic steady state in period 0, and then the shock hits the economy in period 1. In the Technical Appendix we report results for an extended set of variables. The left column of Figure 2 shows results when the ZLB is, hypothetically, not assumed to be binding, whereas the right column shows the effects when the ZLB binds. As is evident from the left column, the same-sized shock has a rather different impact on the economy depending on whether the model is linearized or solved in its original nonlinear form. For instance, we see that while output falls more in the nonlinear model, price inflation falls notably less than in the linearized solution.

In the right column in Figure 2, we report the effects of the same shock, but now assume that the central bank is constrained by the ZLB on the policy rate. Important insights about the differences between the linearized and nonlinear solutions can be gained. First, although the drop in the potential real rate (not shown) is about the same in both models, the linearized model generates a much longer liquidity trap because inflation and expected inflation fall much more, which in turn causes the actual real interest rate (not shown) to rise much more initially. The larger initial rise in the actual real interest rate — and thus the rise in gap between actual and potential real interest rates — triggers a larger fall in the output gap.\(^\text{11}\) Even so, and perhaps most important, we see that price inflation falls substantially less in the nonlinear model compared to the linearized model. After running counterfactual experiments when we linearize parts of the model only, we find that the bulk of the differences between the linearized and nonlinear solutions is driven by the linearization of the price- and wage-setting block of the model.

It is also instructive to compare the solutions with and without imposing the ZLB. Comparing the linearized solutions, we find that imposing the ZLB results in a notably larger fall in output (from -3.5 to almost -7 percent) and deflation in prices and wages (not shown). For the nonlinear solution, we find that imposing the ZLB (albeit admittedly so with a shorter duration compared to the linearized solution) does not affect the price and wage inflation paths much — they are essentially unaffected. The main impact of imposing the ZLB in the nonlinear model is — apart

\(^{11}\) Real GDP also falls more in the linearized model than in the nonlinear model because the discount factor shock does not impact potential real GDP.
from the interest rate path — a somewhat deeper output contraction. According to United States congressional budget office (CBO), the output gap fell by roughly 6 percent during the Great Recession but PCE price inflation (4-quarter change) never fell below 1 percent. Our nonlinear solution is roughly consistent with these facts. By contrast, the linearized model is associated with a notably more depressed path for inflation which is counterfactual compared to the data.¹²

The Kimball aggregator is key for shrinking the sensitivity of inflation to the large adverse shock in economic activity in the nonlinear model. To show this, we solve our model under the assumption that final good and labor services are aggregated with the standard Dixit-Stiglitz constant elasticity demand schedule ($\psi_p = \psi_w = 0$). Because this adjustment changes the slopes of the linearized price and wage Phillips curves, we adjusted $\xi_p$ and $\xi_w$ so that the Phillips curve slopes of the Dixit-Stiglitz specification are identical to the benchmark calibration with the Kimball aggregator. In practice, this requires increasing $\xi_p$ and $\xi_w$ to about 0.9, respectively. With the Dixit-Stiglitz aggregator in both wage- and price-setting, Figure A.4 in the Technical Appendix shows that price inflation would fall even more in the nonlinear model compared to the linearized model. So, the strategic complementarities (or real rigidities) introduced by the Kimball aggregator are essential for a muted inflation response.

3.2. Phillips Curves

To understand the unconditional differences in the dynamics implied by the linearized and nonlinear solutions, we now undertake stochastic simulations of the model for shocks to the stochastic discount factor $\delta_t$. We solve and simulate the linearized and nonlinear model for a long sample of 50,000 periods contingent on exactly the same sequence of shocks $\{\varepsilon_{\delta,t}\}_{t=1}^{50,000}$ in equation (13). However, we use somewhat different standard deviations for the linearized model ($\sigma_\delta = 0.0014$) and the nonlinear model ($\sigma_\delta = 0.0017$), to ensure that the probability of hitting the ZLB is 10 percent in both model solutions. The left column in Figure 3 shows the paths of the first 10,000 periods with simulated data in the nonlinear model, whereas the right column shows the simulated data in the linearized model. In the Technical Appendix, we report results for an extended set of variables.

Figure 3 shows noticeable differences between the linearized and nonlinear model for the behavior of price and wage inflation. The simulated data from the linearized model shows several episodes with substantial deflation, whereas the nonlinear model does not feature any spells with deflation.

¹² Figure A.3 in the Technical Appendix shows the paths of variables when the size of the discount factor shock is set to generate an identical fall in the output gap on impact in the linearized and nonlinear model. Our main conclusion remains unaffected: inflation falls less in the nonlinear model compared to the linearized model.
in prices and wages. There are several episodes when price and wage inflation is persistently low, but there are no spells with deflationary outcomes because the Kimball aggregator implies that firms (unions) are reluctant to change prices (wages) much when relative demand is low. On the other hand, in periods when relative demand is high, firms (unions) are more willing to change their prices. As a result, the nonlinear model produces episodes with more elevated price and wage inflation than the linearized solution in which households and firms are – except for ZLB periods – equally sensitive to changes in desired price and wage markups in recessions and booms.

While the results in Figure 3 are instructive to understand many features of the nonlinear model, it is not straightforward to relate the behavior of price and wage inflation to the state of the business cycle. The relationship between actual price/wage inflation and some measure of resource utilization is traditionally referred to as the Phillips curve. Phillips (1958) drew this original downward-sloping relationship between the rate of wage inflation and the unemployment rate. Subsequently, influential work by Roberts (1995), Clarida, Galí and Gertler (1999) and Blanchard (2000) has extended Phillips’ approach to the relationship between price inflation and the output gap. Thus, we use the simulated data in Figure 3 to produce bivariate scatter plots between price (and wage) inflation on the y-axis and the negative of the output gap on the x-axis. The negative output gap is the difference between potential output and actual output. By using the negative of the output gap, we derive a downward-sloping relationship as in Phillips (1958).

Figure 4 shows the results, with the price- and wage-inflation Phillips curves in the upper- and lower panel, respectively. Given that the x-axis plots the negative output gap, i.e. potential minus actual real GDP, a positive number on the x-axis corresponds to recessions and a negative number on the x-axis corresponds to booms. In the linearized model (blue circles), we see that the relationship between wages and prices and the negative output gap is characterized by a constant negative slope coefficient around the steady state rate of inflation of two percent. However, when the economy hits the ZLB, which tends to happen when the negative of the output gap is around or above 2.5 percent, then the slopes of the linear price and wage Phillips curves flatten somewhat, implying that the output gap is more strongly affected by discount factor shocks than price and wage inflation. At the ZLB, the nominal interest rate is constant and further declines in inflation elevate the real interest rate gap which amplifies the decline in economic activity even more – hence a somewhat flatter slope. Even so, the implied flattening of the linearized Phillips curves due to the ZLB is not nearly enough to resolve the missing inflation puzzle. The linearized model erroneously

\[13\] In the EHL model, we do not allow for permanent productivity growth which would raise nominal wage inflation above price inflation in the steady state. We relax this simplifying assumption in the estimated model in Section 4.
predicts a much larger fall in inflation by extrapolating local decision rules around the steady state to deep recessions far away from the steady state.

By contrast, and consistent with the intuition developed in section 2.1.2, the nonlinear model with the Kimball aggregator (red crosses) implies nonlinear price and wage Phillips curves, i.e. nonlinear relationships between price and wage inflation and the negative output gap. The slopes of our price and wage Phillips curves are notably flatter in recessions than in booms, i.e. our nonlinear Phillips curves have a banana- or boomerang-shape as in the seminal paper by Phillips (1958). Consequently and consistent with the data, inflation falls relatively little in deep recessions in our nonlinear model. Through the lens of our model, the missing deflation puzzle arises when working with the linearized version of the underlying nonlinear model.

This finding is very interesting as it — in addition to explaining why inflation fell so little during the Great Recession — offers a possible explanation why prices and wages have risen so little when advanced economies recovered from the Great Recession. As evidenced by Figure 4, our nonlinear model implies that price and wage inflation remain subdued until the level of economic activity has recovered sufficiently relative to its potential for our demand shock. Put differently, even though economic growth may resume after a deep recession, price and wage inflation will only increase noticeably when economic slack has subsided sufficiently. Therefore, our nonlinear model does not only resolve the missing deflation puzzle during the Great Recession, it also offers an explanation for the so-called missing inflation puzzle observed in the aftermath of the Great Recession (see IMF, 2016 and IMF, 2017). According to our model, subdued price and wage inflationary pressures as observed during the recovery from the Great Recession are indicative of the presence of substantial economic slack. While the linearized model suggests that price and wage inflation should rise sharply during a recovery from a deep recession, the nonlinear model suggests otherwise: price and wage inflation will not rise until economic slack has narrowed sufficiently from its elevated recession level.

Moreover, the nonlinear model also features stronger responses of price and wage inflation in booms than the linearized model, lending support for inflation scares in booms (see e.g. Goodfriend, 1993). Another difference between the linearized and nonlinear solutions is that the output gap is more volatile in the linearized solution, mainly due to strong propagation of the ZLB constraint. Finally, it is imperative to understand that the relationships in Figure 4 are contingent on the assumption that the discount factor shock is the single driver of business cycles. No other shocks are assumed to affect the economy. This is why we obtain a tight negative relationship between
price and wage inflation and economic slack. As we will see in the estimated model that we study next, this tight negative relationship ceases to exist in both the linearized and nonlinear model when different types of shocks affect the economy simultaneously.

4. An Estimated Medium-Sized New Keynesian Model

The benchmark EHL model studied so far is useful to understand how nonlinearities in real rigidities in price- and wage-setting affect inflation dynamics in deep recessions. Specifically, we used the EHL model to demonstrate some of the benefits of taking nonlinearities into account as opposed to the traditional approach which entails log-linearizing the key model equations apart from the monetary policy rule.

However, this analysis was done in a stylized model with one shock and without allowing for e.g. endogenous capital accumulation. In this section we move on to a substantive analysis with the aim of examining the importance real rigidities in a nonlinear setting in a more quantitatively realistic model environment. We specify and estimate a medium-sized New Keynesian model with endogenous capital accumulation that follows closely the seminal model of Christiano, Eichenbaum and Evans (2005) but allows for variety of shocks as in Smets and Wouters (2003, 2007).

Next, we use the estimated model to examine the properties of the nonlinear and linearized solutions of it along several dimensions. First, the paths of model variables of the linearized and nonlinear models are compared with the actual outcomes during the Great Recession following Christiano, Eichenbaum and Trabandt (2015). Second, we redo the Phillips curve analysis in Section 3.2 using the estimated model. Third and finally, we examine the ability of the estimated model to capture the unconditional skewness of price and wage inflation as well as the skewness in other macroeconomic data.

4.1. Medium-Sized Model

Following Christiano, Eichenbaum and Evans (2005) and Smets and Wouters (2003, 2007), the medium-sized model includes both sticky nominal wages and prices. As in the EHL model, the Kimball (1995) aggregator is used to aggregate intermediate goods and labor to final output goods and effective labor input. The medium-sized model also features internal habit persistence in consumption, and embeds a \(Q\)–theory investment specification where changing the level of investment is costly.\(^{14}\) We use the same seven shocks as in Smets and Wouters (2007), i.e. shocks to money-

\(^{14}\) A difference with respect to Christiano, Eichenbaum and Evans (2005) and Smets and Wouters (2003, 2007) is that we do not allow for indexation to past price- and wage-inflation of non-optimizing firms and wage-setters. By
tary policy, government consumption, stationary neutral technology, stationary investment-specific
technology, risk premium, price- and wage-markups. We relegate a more detailed description of the
model to Appendix A.

4.2. Estimation

We now proceed to discuss how the model is estimated using U.S. data from 1965Q1-2007Q4. We
estimate the linearized model with Bayesian maximum likelihood techniques. To solve and estimate
the model we use Dynare. We estimate a similar set of parameters as Smets and Wouters (2007).

4.2.1. Data

We use seven key macroeconomic quarterly U.S. time series as observable variables: the log dif-
ferences of real per capita GDP, consumption and investment; log-differences of compensation per
hour and the PCE deflator; log deviations of hours per capita from its average as well as the Fed-
eral Funds rate. Further details about the data and the measurement equations linking the model
variables to their data counterparts are provided in the Technical Appendix.

4.2.2. Estimation Methodology

Following Smets and Wouters (2003, 2007), we use full information Bayesian techniques to estimate
the model. Bayesian inference starts out from a prior distribution that describes the available
information prior to observing the data used in the estimation. The observed data is subsequently
used to update the prior, via Bayes’ theorem, to a posterior distribution of the model’s parameters
which can be summarized in the usual measures of location (e.g. mode or mean) and spread (e.g.
standard deviation and probability intervals).\(^{15}\)

Some of the parameters in the model are kept fixed throughout the estimation procedure (i.e.,
are subject to priors with a degenerate distribution). We choose to calibrate the parameters that
are weakly identified by the data that we use in the estimation. Table 1 provides the calibrated
parameters which are set to standard values from the literature, whereas Table 2 contains steady
states and implied parameters of the estimated model. Finally, Table 3 contains information about
the parameter prior and posterior distributions.

\(^{15}\)We refer the reader to Smets and Wouters (2003, 2007) for a more detailed description of the estimation
procedure.
4.3. Role of Nonlinearities During the Great Recession

To illustrate the scope of nonlinearities for accounting for the missing deflation puzzle, we use the estimated parameters and subject the nonlinear and linearized model to a positive risk premium shock. We compare the resulting model paths with the outcomes in the data following the methodology of Christiano, Eichenbaum and Trabandt (2015), CET henceforth. The risk premium shock $\epsilon_{RP,t}$ enters the model in the optimality condition for bondholdings:

$$\Lambda_t = \beta E_t \Lambda_{t+1} R_t \epsilon_{RP,t}$$

(14)

where $\Lambda_t$ denotes the Lagrange multiplier on the household budget constraint and $R_t$ is the gross nominal interest rate. $\epsilon_{RP,t}$ denotes the risk premium shock used in Smets and Wouters (2003, 2007) which follows an estimated AR(1) process. Via its effect on the Lagrange multiplier $\lambda_t$, the risk premium shock also affects the optimality conditions for investment and capital and enables the model to reproduce the strong positive correlation between consumption and investment observed in the data. Both CET and Lindé, Smets and Wouters (2016) argue that the risk premium shock was a key shock driving the Great Recession.

The risk premium $\epsilon_{RP,t}$ in eq. (14) is assumed to rise in a uniform fashion for 16 quarters before gradually receding. Its size is set so that both the linearized and nonlinear models’ output path roughly matches the “actual outcome” (discussed in detail below) during the crisis. Figure 5 depicts the results of the nonlinear and linearized model for a variety of macroeconomic variables together with actual outcomes in the data.

The gray area and black solid lines are computed using the methodology in CET and imply that pre-crisis trends are deducted from the actual data 2008Q3-2015Q2. The idea behind this procedure is an assessment about how the economy would have evolved absent the Great Recession. For each variable shown in Figure 5, we fit linear trends from date $x$ to 2008Q2, where $x = \{1985Q1, 2003Q1\}$. To characterize what the data would have looked like absent the Great Recession, we follow CET and extrapolate a trend line for each variable for the period 2008Q3-2015Q2. Following CET we calculate the so-called “target gaps” as the difference between the actual data and the fitted pre-crisis trends. The min-max ranges of the target gaps correspond to the gray intervals displayed in Figure 5 and the black solid line depicts the mean of the target gaps.

The purpose of our analysis is to assess whether the nonlinear and/or the linearized model is able to generate dynamics for the endogenous model variables that lie within the target gap ranges for the post 2008Q2-period. As can be seen from Figure 5, the elevated risk premium shock exerts
a significant adverse impact on the economy in which economic activity dampens and inflation falls. As a result, the policy rate is driven towards a prolonged episode at the zero lower bound. The model matches well the decline in consumption, but the fall of investment is somewhat smaller than in the data, presumably because the model lacks financial friction amplification mechanisms. Importantly, for the same-sized output response, price and wage inflation in the nonlinear model fall by about one percentage point less than in the linearized model, confirming our results in the benchmark EHL model.

4.4. Phillips Curves in the Estimated Model

To provide intuition for the muted inflation response in the nonlinear model following a positive risk premium shock, Figure 6 shows a scatter plot of price inflation and the negative output gap in both the linearized and the nonlinear variant of the estimated medium-sized New Keynesian model. Following the procedure in Section 3.2 to simulate data from the model, the upper left scatter plot is generated by sampling only risk premium innovations from a normal distribution using the estimated posterior mode and then simulating a long sample of 50,000 periods. Notice that the simulations are initiated at the steady state, and that the negative of the output gap is plotted on the x-axis, which means that a large positive number is associated with a deep recession.

As can be seen from the upper left plot in Figure 6, the linearized model is associated with a linear relationship between inflation and the output gap when only risk premium shocks are used in the simulation, whereas the nonlinear model suggests a banana- or boomerang-shape relationship. The relationship in the linearized model is entirely linear because the risk premium shocks are by themselves not large enough in the estimated model to drive the economy into a liquidity trap. Even so, the risk premium shock is a key driver of business cycle dynamics in the estimated model and the difference between the linearized and nonlinear solution has important implications for the dynamics of inflation and output.

Figure 6 also contains the implied price Phillips curves for the other six shocks of the model: we run stochastic simulations for each of the shocks – one at a time – in both the linearized and nonlinear model and then use the simulated data to construct the corresponding scatter plots. To do this, we use the estimated shock processes and parameters reported in Table 3.

Interestingly, equally-sized markup shocks have notably larger effects on inflation in the nonlinear model compared to the linearized model. We shrink the size of price and wage markup shocks in the nonlinear model such that the unconditional volatility of price and wage inflation is the same in the nonlinear and linearized model. In other words, the nonlinear model does not depend as much on large markup shocks as the linearized model to explain the volatility of price and wage inflation. In this sense, the nonlinear model is less prone to the critique against New Keynesian models by Chari, Kehoe and McGrattan (2010).
right panel of Figure 6, we plot the results when all shocks are used in the model simulation.

As can been seen from Figure 6, risk premium and wage markup shocks account for the bulk of the volatility of inflation and output in the estimated model. For these two shocks, we obtain the most sizeable fluctuations in inflation and the output gap. Because these two shocks move the equilibrium far away from the steady state, the nonlinearities are also most evident for these two shocks. Given the estimated parameters, the wage markup shocks are the only source of fluctuations which have the potential to generate close to zero inflation or very mild deflationary episodes. It is striking how the dynamics of the price and wage markup shocks differ between the linearized and nonlinear solution. Other shocks which only cause moderate fluctuations in the output gap and inflation result in small differences between the linearized and nonlinear solutions because they do not generate any substantial deviations from the steady state.

Moreover, it is evident that there are no relevant trade-offs in stabilizing output and inflation to fluctuations in risk-premium, government spending, and neutral technology shocks. By contrast, investment-specific and markup shocks – especially wage markup – create noticeable trade-offs between inflation and output gap stabilization. Furthermore, when using all shocks in the simulation, the importance of the wage markup shock renders the Phillips curve completely flat or even upward-sloping, consistent with the empirical observation of no clear unconditional Phillips curve pattern in post-war U.S. data.

It is important to understand that many shocks do not generate any noticeable differences between the linearized and nonlinear solutions when simulated one at a time since they do not drive the economy far away from the steady state. Note, however, that this does not imply that these shocks cannot propagate differently in a recession (or boom) in the nonlinear model relative to the linearized solution. Figure 7 shows the results corresponding to those in Figure 6 conditional on a deep recession with an expected 8-quarter liquidity trap.

To construct Figure 7, we first simulate a baseline using an adverse risk premium shock which generates a recession with an anticipated 8-quarter liquidity trap. Relative to the baseline, we run 50,000 stochastic simulations of a counterfactual scenario in which other shocks are drawn from their estimated stochastic processes. We then compute the differences between the 50,000 counterfactual scenarios and the baseline after one year.\footnote{We plot the observations after one year since habit formation and investment adjustment costs imply that for most shocks, output and inflation attain their peak effects with some delay.}

Figure 7 shows notable differences between the linearized and nonlinear solutions for many of the shocks for which Figure 6 reported no differences, e.g. monetary policy, government spending,
neutral technology and investment-specific technology shocks. According to the results in Figure 7, monetary policy is much less potent to affect inflation in a prolonged liquidity trap than in normal times. The effects of monetary policy shocks on output are also somewhat moderated in the nonlinear model compared to the linearized solution. Negative investment-specific shocks, on the other hand, have notably more negative effects on the output gap in a liquidity trap. In contrast to wage markup shocks, price markup shocks have essentially no effect in a liquidity trap in the nonlinear model: firms are unwilling to strongly respond to desired markup variations with Kimball aggregation.

It is important to point out that the results in Figure 7 generally imply notably flatter Phillips curves for many shocks in a liquidity trap. Hence, the nonlinear model offers a possible explanation to the empirical observation that the sensitivity of inflation to economic activity has become smaller in linearized models since the onset of the crisis (see e.g. Lindé, Smets and Wouters, 2016). In the nonlinear model, however, the smaller sensitivity is only temporary since it is contingent on a persistently negative output gap. The sensitivity will increase when economic slack has diminished sufficiently after the recession.

4.5. Accounting for Skewness in the Data

We have documented that the nonlinear model can account for the dynamics of inflation during the Great Recession and thereafter, including the flatter slope of the Phillips curve. We now turn to stochastic simulations to examine whether the nonlinear model may also help to explain unconditional moments of inflation and other macroeconomic data. We do this to show that the nonlinear model can explain the dynamics of inflation also prior to the Great Recession better than the linearized model.

It is well known that inflation has a positive skew during the postwar period, i.e. there are episodes with inflation bursts and then there are episodes with very low and moderate rates of inflation but no long-lived deflationary episodes. This positive skew is a robust finding for different measures of inflation. The gray area in Figure 8 shows the kernel smoothed distribution for inflation measured by the core PCE deflator for the sample period 1965Q1-2007Q4.

The blue and red lines show the corresponding kernel smoothed distributions based on the stochastic simulations of the linearized and nonlinear model when all shocks are used in the simulation, i.e. the observations plotted in the bottom right panel in Figure 6. As can be seen from Figure 8, the nonlinear model fits the unconditional statistical properties of inflation remarkably well –
much better than the linearized model. The linearized model features a completely symmetric normal distribution for inflation, with significant mass in deflationary territory. By contrast, the nonlinear solution features zero density in deflationary territory for inflation, in line with the data. It is striking how much better the nonlinear model captures unconditional U.S. inflation dynamics during the post-war period compared to the linearized model.  

Figure 8 also depicts density plots for wage inflation, real GDP growth and the Federal Funds rate. The nonlinear model captures the skewness of wage inflation and the Federal Funds rate better than the linearized model. The Technical Appendix also reports results for an extended set of variables.

5. Related Literature and Evidence

In this section, we discuss literature that is related to our work. We also discuss micro evidence that sheds light on the empirical relevance of real rigidities with a special emphasis on Kimball versus Dixit-Stiglitz aggregation. Finally, we also discuss macroeconomic evidence related to the missing deflation puzzle.

5.1. Micro Evidence

A large body of literature has documented patterns of nominal price stickiness at a disaggregated good level, see e.g. Nakamura and Steinsson (2008), Klenow and Malin (2010) and the references therein. The empirical findings suggest that prices change on average about every three to four quarters, i.e. nominal stickiness is relatively short-lived. Christiano, Eichenbaum, and Evans (1999) and Gertler and Karadi (2015) have documented persistent real effects of monetary policy shocks on output. Short-lived price stickiness and persistent real effects of monetary shocks can be reconciled with each other in modeling frameworks with sufficient real rigidities, i.e. features that make firms reluctant to change their prices by large amounts.

Amid this background, an empirical literature has started to emerge recently to examine whether there are quantitatively important real rigidities present in the data, see e.g. Beck and Lein (2015),

---

18 Keep in mind that the model estimation has not used information about the skewness of inflation. In this sense, the results in Figure 8 can be interpreted as posterior predictive checks.
19 Note that the sample in Figure 8 only spans data prior to the Great Recession. Had we included data until 2018, i.e. including the U.S. ZLB period, the fit between the model and the data would improve markedly for the Federal Funds rate. For the other variables the effect of including more recent data is less pronounced.
20 There is a long-standing I/O literature which consistent with the implications of our nonlinear model argues that prices tend to respond faster in booms than in recessions, see for instance Peltzman (2000) and the references therein.
Dossche, Heylen and Van den Poel (2010), Klenow and Willis (2016), and Levin and Yun (2008). In general, the empirical literature appears to provide evidence in favor of Kimball aggregation rather than for Dixit-Stiglitz aggregation. That is, the empirical literature has estimated demand curves which exhibit state-dependent demand elasticities rather than constant demand elasticities. Put differently, there appears to be more support for quasi-kinked demand curves as implied by Kimball aggregation than for log-linear demand as implied by Dixit-Stiglitz aggregation (see the upper left panel in Figure 1 for an illustration).

However, although the evidence provided by the empirical literature supports Kimball over Dixit-Stiglitz aggregation in general, a balanced assessment of the evidence provided by the above papers is that it does not seemingly support the degree of real rigidities needed to fit the macro evidence. That said, the empirical literature has typically used micro data which does not cover the Great Recession period which is at the core of our study. In addition, real rigidities are difficult to identify and difficult to measure in the data as noted by Gopinath and Itskhoki (2010). Aggregate real rigidities imply a sluggish response of firms’ marginal costs to aggregate shocks. Firm-level real rigidities imply a muted response of firms’ prices conditional on price adjustment to marginal cost shocks. Since data on marginal costs is usually unavailable, it is hard to test these mechanisms directly. So, the evidence provided in e.g. Klenow and Willis (2016) is merely indicative, and their implied large idiosyncratic shock variance needed for their findings to be compatible with our estimate of the demand elasticity may be reconciled by allowing for firm-specific customer relationships as in Levin and Yun (2008) and Kleshchelski and Vincent (2009).

All told, we interpret the micro evidence as providing at least partial support for our findings. However, we believe that further empirical analysis based on joint micro data of goods prices, firms’ marginal costs and sales in a sample which also covers the Great Recession are needed before decisive conclusions can be drawn. We leave this analysis for future research.

5.2. Macro Evidence

Recent research has examined possible resolutions of the missing deflation puzzle. Using standard time methods on U.S. data, an emerging body of literature finds support for the presence of non-linearities in the Phillips curve. Doser et al. (2018) estimate a two-state regime Phillips curve and report a lower slope of the Phillips curve in a state in which the unemployment rate is elevated.

There has also been considerable theoretical work that explores the effects of real rigidities in firms’ price-setting behavior for small shocks, see e.g. Dotsey, King and Wolman (1999), Dotsey and King (2005), Dupraz (2017), Kocherlakota (2017) and Kleshchelski and Vincent (2009). In this section, however, we focus on micro evidence on real rigidities.
Gagnon and Collins (2019) argue – based on time series evidence for a wide range of inflation time series – that the Phillips curve is likely to be highly nonlinear when inflation is low. The evidence provided in both papers is consistent with our structural model.

Another body of literature has used structural models to examine possible explanations for the missing deflation puzzle. Lindé, Smets and Wouters (2016) and Fratto and Uhlig (2018) find that the Smets and Wouters (2007) model relies on large offsetting positive price markup shocks to cope with the small drop in inflation in the face of a persistent fall in output observed during the Great Recession.

Recent research has also emphasized that financial frictions can be helpful to account for a small elasticity between output and inflation witnessed during the Great Recession. Christiano, Eichenbaum and Trabandt (2015) use a model to show that the observed fall in total factor productivity and the rise in firms’ cost to borrow funds for working capital played critical roles in accounting for the small drop in inflation that occurred during the Great Recession. Del Negro, Giannoni and Schorfheide (2015) show that the introduction of a financial accelerator together with a flattening of the Phillips curve can account for the small drop in inflation in the Great Recession. Gilchrist, Schoenle, Sim and Zakrajsek (2017) develop a model in which firms face financial frictions when setting prices in an environment with customer markets. Financial distortions create an incentive for financially constrained firms to raise prices in response to adverse financial or demand shocks in order to preserve internal liquidity and avoid accessing external finance. While financially unconstrained firms cut prices in response to these adverse shocks, the share of financially constrained firms is sufficiently large in their model to attenuate the fall in inflation in response to a large contraction in GDP. Gilchrist, Schoenle, Sim and Zakrajsek (2017) examine a micro data set which supports the implications of their model.

Our work is also related to Aruoba, Boccola and Schorfheide (2017). Using asymmetric price and wage adjustment costs, the authors show that their model can produce skewness in inflation and output growth as observed in the data. However, their model cannot account for the skewness of the Federal Funds rate data while ours does. In contrast to these authors, our model does not rely on asymmetric adjustment costs but generates the skewness observed in macro data due to the Kimball (1995) aggregator. Moreover, Aruoba, Boccola and Schorfheide (2017) do not study the implications of the zero lower bound while our paper focuses on the interplay of large shocks and the zero lower bound to characterize nonlinearities in price and wage Phillips curves.

More generally, our work offers an alternative and perhaps complementary explanation to un-
derstand the missing deflation puzzle. Our resolution of the puzzle stresses the nonlinear influence of strategic complementarities and real rigidities in price- and wage-setting. We find it attractive due to its simplicity and due to its ability to address additional issues beyond the missing deflation puzzle which we will discuss in the next section.

6. Conclusions

We have formulated a nonlinear macroeconomic model which goes a long way towards accounting for the missing deflation puzzle, i.e. the empirical fact that inflation fell little against the backdrop of a large and persistent fall in output during the Great Recession. Our resolution of the puzzle stresses the nonlinear influence of strategic complementarities and real rigidities in price- and wage-setting. Our proposed nonlinear framework is also attractive along the following additional dimensions: (i) it mitigates the tension between the macroeconomic evidence of a low Phillips curve slope and the microeconomic evidence of frequent price changes; (ii) it explains inflation dynamics prior to the crisis and the empirical positive skew in inflation without relying on a similarly positive skew in output (which would be counterfactual); and (iii) it sheds light on the missing inflation puzzle, i.e. why price and wage inflation remained persistently low in the aftermath of the Great Recession.

While our nonlinear model allows us to simultaneously explain the quantitative inflation performance prior to the Great Recession, the missing deflation during the Great Recession, as well as the missing inflation in recent years, a linearized version of the underlying nonlinear model fails to do so. All told, our results caution against the common practice of using linearized models to study inflation and output dynamics when the economy is exposed to large shocks.

References


Christiano, Lawrence, Martin Eichenbaum and Benjamin K. Johannsen (2016), “Does the New Keynesian Model Have a Uniqueness Problem?”, manuscript, Northwestern University.


Klenow, Peter J. and Benjamin A. Malin (2010), “Microeconomic Evidence on Price-Setting”, Chapter 6 in Benjamin M. Friedman and Michael Woodford (Eds.), Handbook of Monetary Economics, Elsevier.


Appendix A. The Estimated Medium-Sized New Keynesian Model

This appendix describes the estimated medium-sized model. The model closely follows Christiano, Eichenbaum and Evans (2005), Smets and Wouters (2007) and Christiano, Eichenbaum and Trabandt (2016).

Like the canonical sticky price and wage model described in Section 2, the medium-sized model includes monopolistic competition in the goods and labor markets and nominal rigidities in prices and wages with Kimball (1995) aggregation. In addition to saving in risk-free bonds, households can also save in physical capital. It takes one period before new investment turns into productive capital. In addition to the canonical model, the medium-sized model also features several real rigidities in the form of habit formation in consumption, investment adjustment costs, variable capital utilization, and fixed costs in production.

The model dynamics are driven by the same seven structural shocks as in Smets and Wouters (2007). The following four shocks affect the actual economy but not the potential (flexible price, flexible wage) economy: price- and wage markup shocks, risk premium shocks and monetary policy shocks. The following three shocks affect both the actual and the potential economy: neutral technology shocks, investment-specific technology shocks and government spending shocks. All shocks follow AR(1) processes in logs, except for the markup shocks which follow ARMA(1,1) processes in logs and monetary policy shocks which are assumed to be white noise.

Below, we describe the households' and firms' problems in the model, and state the market clearing conditions. The Technical Appendix which is available online provides a list of nonlinear and log-linearized equilibrium equations, measurement equations and further technical information.
A.1. Households

There is a continuum of households \( j \in [0, 1] \) in the economy. Each household supplies a specialized type of labor \( j \) to the labor market. The \( j^{th} \) household is the monopoly supplier of the \( j^{th} \) type of labor service.

A.1.1. Preferences and Budget Constraint

The \( j^{th} \) household maximizes

\[
E_0 \sum_{t=0}^{\infty} \beta^t \left\{ \ln (C_{j,t} - b C_{j,t-1}) - \frac{\omega}{1 + \chi} \zeta_{h,t}^{1/\kappa_w} N_{j,t}^{1+\chi} \right\}
\]

subject to

\[
P_t C_{j,t} + P_{I,t} I_{j,t} + \frac{B_{j,t}}{\epsilon_{RP,t}} = (R_{K,t} u_{j,t} - a(u_{j,t}) P_{I,t}) K_{j,t-1} + W_{j,t} N_{j,t} + R_{t-1} B_{j,t-1} - T_{j,t} + A_{j,t}
\]

Here, \( T_{j,t} \) denotes lump-sum taxes net of profits, \( P_t \) denotes the price of consumption goods \( C_{j,t} \), \( P_{I,t} \) denotes the price of investment goods \( I_{j,t} \), \( B_{j,t} \) denotes one period risk-free bonds purchased in period \( t \) with gross return, \( R_t \). The object \( R_{K,t} \), denotes the rental rate of capital services, \( K_{j,t-1} \) denotes the household’s beginning of period \( t \) stock of capital, \( a(u_{j,t}) \) denotes the cost, in units of investment goods, of the capital utilization rate, \( u_{j,t} \) and \( u_{j,t} K_{j,t-1} \) denotes the household’s period \( t \) supply of capital services. The functional form for the increasing and convex function, \( a(\cdot) \), is described below. All prices, taxes and profits in the budget constraint are in nominal terms.

\( \epsilon_{RP,t} \) denotes the risk premium shock as used by Smets and Wouters (2003, 2007) which follows an AR(1) process.

The choice variables of the household are consumption, investment, risk-free bonds, capital utilization and the wage which is subject to Calvo wage setting frictions following EHL. The household understands that when choosing \( W_{j,t} \) it must supply the amount of hours \( N_{j,t} \) demanded. \( A_{j,t} \) represents net proceeds of an asset that provides insurance against the idiosyncratic uncertainty associated with the Calvo wage-setting friction.

The parameters \( b, \omega \) and \( \chi \) control the degree of habit formation, the level of disutility from labor and the labor supply elasticity, respectively.

\( \zeta_{h,t}^{1/\kappa_w} \) denotes an exogenous preference shifter for labor, scaled by \( 1/\kappa_w \). The scaling factor is the inverse slope of the linearized wage Phillips curve in terms of the marginal rate of substitution minus the real wage, i.e. \( 1/\kappa_w = 1/ \left( \frac{(1-\xi_w)(1-\beta \xi_w)}{\xi_w} \frac{1}{1-(1+\theta_w)\psi_w} \right) \). The scaling implies that \( \zeta_{h,t} \) enters the linearized wage Phillips curve with a unit coefficient. The same scaling is used in Smets
and Wouters (2007) and in many other empirical papers. Note that $\zeta_{h,t}$ equals unity in steady state so that the scaling does not affect the steady state. The scaling gets rid of the high negative correlation between the estimated standard deviations of wage markup shocks and the estimated parameter of the slope of the wage Phillips curve. Economically, $\zeta_{h,t}$ can be interpreted as a wage markup shock. We assume that $\zeta_{h,t}$ follows an ARMA(1,1) process.

**A.1.2. Physical Capital**

The representative household’s stock of capital evolves as follows:

$$K_{j,t} = (1 - \delta) K_{j,t-1} + \Upsilon_t [1 - S (I_{j,t}/I_{j,t-1})] I_{j,t}. $$

The functional form for the increasing and convex adjustment cost function, $S(\cdot)$, is described below.

$\Upsilon_t$ denotes a stationary investment-specific technology shock which follows an AR(1) process. We scale the investment-specific technology shock by the factor $\eta = k/i = 1/(1 - (1 - \delta)/(\mu_\mu_i))$ such that the investment-specific technology shock enters the linearized law of motion for capital with a unit coefficient. The same scaling was used in Adolfson, Laséen, Lindé and Villani (2007).

**A.1.3. Wage Setting**

Following EHL, we assume wage setting is subject to Calvo sticky wages. The wage setting problem is nearly identical to the one described in the stylized model, see section 2.1.3. The only differences are i) dropping the discount factor shock and ii) the inclusion of the labor preference shock in our medium-sized model. Specifically, households choose the optimal wage by maximizing

$$\max_{W_{opt}j,t} \mathbb{E}_t \sum_{s=0}^{\infty} (\beta \xi_w)^s \left( \Lambda_{t+s} W_{opt}^{j,t} \Pi_{w}^{s} N_{j,t+s} - \omega \zeta_{h,t+s}^{1/\kappa_w} N_{j,t+s}^{1+\chi} \right)$$

subject to labor demand given by equation (7).

**A.2. Firms and Price Setting**

*Final Goods Production.* Like in the stylized model, the single final output good $Y_t$ is produced using a continuum of differentiated intermediate goods $Y_{f,t}$. Firms who produce the final output good are perfectly competitive in product and factor markets. The optimization problem is exactly the same as described in section 2.1.1. The resulting optimality conditions are (2), (3) and (4).
**Intermediate Goods Production.** Like in the stylized model, a continuum of intermediate goods $Y_{f,t}$ for $f \in [0,1]$ is produced by monopolistically competitive firms, each of which produces a single differentiated good. Each intermediate goods producer faces the demand schedule in eq. (2) and rents capital services $K_{f,t}$ and labor services $N_{f,t}$ for production. The form of the production function is Cobb-Douglas:

$$Y_{f,t} = \epsilon_t K_{f,t}^\alpha (z_t N_{f,t})^{1-\alpha} - \phi_t.$$  

Firms face perfectly competitive factor markets for renting capital and hiring labor. Thus, each firm chooses $K_{f,t}$ and $N_{f,t}$ taking as given both the rental price of capital $R_{K,t}$ and the aggregate wage rate $W_t$. Firms can costlessly adjust either factor of production. Thus, the standard static first-order conditions for cost minimization imply that all firms have identical marginal cost per unit of output. Also, $z_t$ is unit-root neutral technology which we assume grows deterministically over time. $\epsilon_t$ denotes a stationary neutral technology shock which we assume to follow an AR(1) process. $\phi_t$ represents a fixed cost of production.

We assume that price setting is subject to Calvo sticky prices. The price setting problem is identical to the one described in section 2.1.1, except that we drop the discount factor shock in the medium-sized model. Specifically:

$$\max_{p^\text{opt}_{f,t}} \mathbb{E}_t \sum_{s=0}^{\infty} (\beta \xi_p)^s \Lambda_{t+s} \left( \Pi_s p^\text{opt}_{f,t} - MC_{t+s} \right) Y_{f,t+s}$$

where $\Lambda_{t+j}$ is household marginal utility and demand $Y_{f,t+j}$ is given by equation (2). Similar to the canonical model, nominal marginal cost is given by:

$$MC_t = \tau_t^{1/\kappa_p} W_t / \epsilon_t (1-\alpha) (u_t K_{t-1}/N_t)^\alpha.$$  

We add $\tau_t$ to the model-consistent expression for marginal cost. It captures a variety of variations in marginal costs that are exogenous to the model. We assume that $\tau_t$ follows an ARMA(1,1) process. Further, we scale $\tau_t$ by the parameter $1/\kappa_p$. The scaling factor is the inverse slope of the linearized price Phillips curve in terms of real marginal cost, i.e. $1/\kappa_p = 1 / \left( \frac{(1-\xi_p)(1-\beta \xi_p)}{\xi_p} \right)^{1/(1-(1+\theta_p)\psi_p)}$.  

The scaling implies that $\tau_t$ enters the linearized price Phillips curve with a unit coefficient. The same scaling is used in Smets and Wouters (2007) and in many other empirical papers. Note that $\tau_t$ equals unity in steady state so that the scaling does not affect the steady state. The scaling gets rid of the high negative correlation between the estimated standard deviations of price markup shocks and the estimated parameter of the slope of the price Phillips curve.
A.3. Market Clearing and Functional Forms

Market clearing in the markets for labor and capital services require:

\[ \int_0^1 N_{j,t} dj = l_t \quad \text{and} \quad \int_0^1 N_{f,t} df = N_t \]
\[ \int_0^1 u_{j,t} K_{j,t-1} dj = \int_0^1 K_{f,t} df = u_t K_{t-1} \]

Homogeneous output, \( Y_t \), can be used to produce private or government consumption goods or investment goods. The production of the latter uses a linear technology in which one unit of \( Y_t \) is transformed into \( \Psi_t \) units of \( I_t \).

Market clearing for final goods requires:

\[ C_t + (I_t + a(u_t) K_{t-1})/\Psi_t + G_t = Y_t \]

where \( C_t = \int_0^1 C_{j,t} dj \) and \( I_t = \int_0^1 I_{j,t} dj \). \( G_t \) denotes government consumption.

Perfect competition in the production of investment goods implies that the nominal price of investment goods equals the corresponding marginal cost:

\[ P_{I,t} = P_t / \Psi_t. \]

Aggregate output is given by:

\[ Y_t = (p^*_t)^{-1} \left( \epsilon_t (u_t K_{t-1})^\alpha (z_t N_t)^{1-\alpha} - \phi_t \right) \]

where \( N_t = (w^*_t)^{-1} l_t \). The variables \( p^*_t \) and \( w^*_t \) denote the price and wage dispersion terms defined in the main text.

The sources of growth in our model are unit-root neutral and investment-specific technological progress. Let:

\[ \Phi_t = \Psi_t^{1-\alpha} z_t \]

be the composite level to technology. In our model, \( Y_t/\Phi_t, C_t/\Phi_t, (W_t/P_t)/\Phi_t \) and \( I_t/(\Psi_t \Phi_t) \) converge to constants in nonstochastic steady state. We also assume that \( \ln \mu_z \equiv \ln (z_t/z_{t-1}) \) and \( \ln \mu_{\Psi} \equiv \ln (\Psi_t/\Psi_{t-1}) \), i.e. growth is deterministic in our model. Accordingly, let \( \ln \mu \equiv \ln (\Phi_t/\Phi_{t-1}) \).

In order to ensure a well-defined balanced growth path, we assume that the fixed cost of production follows \( \phi_t = \Phi_t \phi \) where \( \phi \) is set such that profits are zero along a balanced growth path.
We assume that the cost of adjusting investment takes the form:

\[ S(I_t/I_{t-1}) = \frac{1}{2} \left( \exp \left[ \sqrt{S''} \left( I_t/I_{t-1} - \mu \times \mu_{\Psi} \right) \right] + \exp \left[ -\sqrt{S''} \left( I_t/I_{t-1} - \mu \times \mu_{\Psi} \right) \right] \right) - 1. \]

Here, \( \mu \) and \( \mu_{\Psi} \) denote the unconditional growth rates of \( \Phi_t \) and \( \Psi_t \). The value of \( I_t/I_{t-1} \) in nonstochastic steady state is \((\mu \times \mu_{\Psi})\). In addition, \( S'' \) denotes the second derivative of \( S(\cdot) \), evaluated at steady state, and is a parameter to be estimated. It is straightforward to verify that \( S(\mu \times \mu_{\Psi}) = S'(\mu \times \mu_{\Psi}) = 0 \).

We assume that the cost associated with setting capacity utilization is given by:

\[ a(u_t) = \sigma_a \sigma_b (u_t)^2 / 2 + \sigma_b (1 - \sigma_a) u_t + \sigma_b (\sigma_a / 2 - 1) \]

where \( \sigma_a \) and \( \sigma_b \) are positive scalars. For a given value of \( \sigma_a \) we select \( \sigma_b \) so that the steady state value of \( u_t \) is unity. The object \( \sigma_a \) is a parameter to be estimated.

### A.4. Fiscal and Monetary Policy

Government consumption \( G_t \) is scaled by the level of composite unit-root technology \( \Phi_t \), i.e. \( g_t = G_t / \Phi_t \). We assume that \( g_t \) follows an exogenous AR(1) process. The government is assumed to balance its budget each period by adjusting lump-sum taxes.

When the central bank is unconstrained by the ZLB, it sets the notional gross nominal interest rate \( R_t^{\text{not}} \) according to:

\[
\ln R_t^{\text{not}} = \rho_R \ln R_{t-1}^{\text{not}} + (1 - \rho_R) \left[ \ln R + r_\pi \ln \left( \frac{\pi_t^4}{\bar{\pi}} \right) + r_{\Delta_y} \ln \left( \frac{\mathcal{Y}_t / \mathcal{Y}_t^f}{\mathcal{Y}_{t-4} / \mathcal{Y}_{t-4}^f} \right) \right] + \epsilon_{r,t}
\]

where \( \pi_t^4 = \pi_t \pi_{t-1} \pi_{t-2} \pi_{t-3} \) denotes annual inflation, \( \mathcal{Y}_t \) denotes real GDP defined as

\[ \mathcal{Y}_t = C_t + I_t / \Psi_t + G_t \]

and \( \mathcal{Y}_t^f \) denotes potential real GDP, i.e. real GDP if prices and wages are flexible. \( \bar{\pi} \) denotes the inflation target of the central bank. \( \epsilon_{r,t} \) denotes the monetary policy shock. Finally, we impose the zero lower bound on the nominal interest rate by

\[ R_t = \max(1, R_t^{\text{not}}). \]
Table 1: Non-Estimated Parameters and Calibrated Variables in Medium-Sized Model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: Parameters</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\delta$</td>
<td>0.025</td>
<td>Depreciation rate of physical capital</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.999</td>
<td>Discount factor</td>
</tr>
<tr>
<td>$\phi_w$</td>
<td>1.1</td>
<td>Gross wage markup</td>
</tr>
<tr>
<td>$\xi_w$</td>
<td>0.75</td>
<td>Calvo wage stickiness</td>
</tr>
<tr>
<td>$\chi$</td>
<td>0</td>
<td>Inverse labor supply elasticity</td>
</tr>
<tr>
<td>$\phi_p$</td>
<td>1.1</td>
<td>Gross price markup</td>
</tr>
<tr>
<td>$\xi_p$</td>
<td>0.66</td>
<td>Calvo wage stickiness</td>
</tr>
<tr>
<td>Panel B: Steady State Values</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$400(\pi - 1)$</td>
<td>2.0</td>
<td>Annual net inflation rate</td>
</tr>
<tr>
<td>$profits$</td>
<td>0</td>
<td>Intermediate goods producers profits</td>
</tr>
<tr>
<td>$l$</td>
<td>0.63</td>
<td>Employment to population ratio</td>
</tr>
<tr>
<td>$G/Y$</td>
<td>0.2</td>
<td>Government consumption to gross output ratio</td>
</tr>
</tbody>
</table>

Table 2: Steady States and Implied Parameters at Estimated Posterior Mode in Medium-Sized Model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K/Y$</td>
<td>6.03</td>
<td>Capital to gross output ratio (quarterly)</td>
</tr>
<tr>
<td>$C/Y$</td>
<td>0.61</td>
<td>Consumption to gross output ratio</td>
</tr>
<tr>
<td>$I/Y$</td>
<td>0.19</td>
<td>Investment to gross output ratio</td>
</tr>
<tr>
<td>$R$</td>
<td>1.01</td>
<td>Gross nominal interest rate (quarterly)</td>
</tr>
<tr>
<td>$R_{real}$</td>
<td>1.005</td>
<td>Gross real interest rate (quarterly)</td>
</tr>
<tr>
<td>$mc$</td>
<td>0.91</td>
<td>Marginal cost (inverse markup)</td>
</tr>
<tr>
<td>$\sigma_b$</td>
<td>0.03</td>
<td>Capacity utilization cost parameter</td>
</tr>
<tr>
<td>$Y$</td>
<td>0.87</td>
<td>Gross output</td>
</tr>
<tr>
<td>$w$</td>
<td>1.11</td>
<td>Real wage</td>
</tr>
<tr>
<td>$\omega$</td>
<td>1.91</td>
<td>Disutility of labor parameter</td>
</tr>
<tr>
<td>$\phi/Y$</td>
<td>0.10</td>
<td>Fixed cost to gross output ratio</td>
</tr>
</tbody>
</table>
Table 3: Estimated Parameters of the Medium-Sized Model

<table>
<thead>
<tr>
<th>Parameter Description</th>
<th>Prior Distribution</th>
<th>Posterior Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Price Setting Parameters</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kimball Price-Setting, $-\psi_p$</td>
<td>$\mathcal{N}(6.00,4.00)$</td>
<td>12.55 [7.58;17.7]</td>
</tr>
<tr>
<td>Kimball Wage-Setting, $-\psi_w$</td>
<td>$\mathcal{N}(6.00,4.00)$</td>
<td>8.31 [1.97;14.8]</td>
</tr>
<tr>
<td><strong>Monetary Authority Parameters</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Taylor Rule: Smoothing, $\rho_R$</td>
<td>$\mathcal{B}(0.80,0.10)$</td>
<td>0.73 [0.68;0.79]</td>
</tr>
<tr>
<td>Taylor Rule: Inflation, $r_{\pi}$</td>
<td>$\mathcal{N}(1.50,0.25)$</td>
<td>1.44 [1.22;1.67]</td>
</tr>
<tr>
<td>Taylor Rule: Output Gap, $r_y$</td>
<td>$\mathcal{N}(0.13,0.05)$</td>
<td>0.07 [0.03;0.11]</td>
</tr>
<tr>
<td>Taylor Rule: Change Output Gap, $r_{\Delta y}$</td>
<td>$\mathcal{N}(0.25,0.1)$</td>
<td>0.31 [0.22;0.40]</td>
</tr>
<tr>
<td><strong>Preferences and Technology Parameters</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consumption Habit, $b$</td>
<td>$\mathcal{B}(0.70,0.15)$</td>
<td>0.67 [0.60;0.73]</td>
</tr>
<tr>
<td>Capacity Utilization Adj. Cost, $\sigma_a$</td>
<td>$\mathcal{N}(1.00,0.75)$</td>
<td>0.76 [0.22;1.46]</td>
</tr>
<tr>
<td>Investment Adjustment Cost, $S''$</td>
<td>$\mathcal{N}(4.00,1.00)$</td>
<td>4.89 [3.53;6.29]</td>
</tr>
<tr>
<td>Capital Share, $\alpha$</td>
<td>$\mathcal{N}(0.30,0.07)$</td>
<td>0.20 [0.18;0.21]</td>
</tr>
<tr>
<td>Steady State Real Investment Growth, $\Delta i$</td>
<td>$\mathcal{N}(3.00,0.25)$</td>
<td>2.71 [2.43;3.00]</td>
</tr>
<tr>
<td>Steady State Real GDP Growth, $\Delta i$</td>
<td>$\mathcal{N}(1.75,0.25)$</td>
<td>1.54 [1.45;1.63]</td>
</tr>
<tr>
<td><strong>Exogenous Processes Parameters: Standard Deviations</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investment Technology, $\sigma_I$</td>
<td>$\mathcal{IG}(0.20,\text{Inf})$</td>
<td>0.36 [0.25;0.48]</td>
</tr>
<tr>
<td>Neutral Technology, $\sigma_e$</td>
<td>$\mathcal{IG}(1.00,\text{Inf})$</td>
<td>0.59 [0.53;0.66]</td>
</tr>
<tr>
<td>Government Consumption, $\sigma_G$</td>
<td>$\mathcal{IG}(1.00,\text{Inf})$</td>
<td>1.73 [1.54;1.91]</td>
</tr>
<tr>
<td>Wage Markup, $\sigma_{\zeta_h}$</td>
<td>$\mathcal{IG}(0.20,\text{Inf})$</td>
<td>0.48 [0.40;0.57]</td>
</tr>
<tr>
<td>Price Markup, $\sigma_\tau$</td>
<td>$\mathcal{IG}(0.20,\text{Inf})$</td>
<td>0.17 [0.13;0.21]</td>
</tr>
<tr>
<td>Monetary Policy, $\sigma_R$</td>
<td>$\mathcal{IG}(0.20,\text{Inf})$</td>
<td>0.23 [0.21;0.26]</td>
</tr>
<tr>
<td>Risk Premium, $\sigma_{rp}$</td>
<td>$\mathcal{IG}(0.20,\text{Inf})$</td>
<td>0.68 [0.29;1.14]</td>
</tr>
<tr>
<td><strong>Exogenous Processes Parameters: AR(1) and MA(1) Coefficients</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investment Technology, $\rho_I$</td>
<td>$\mathcal{B}(0.70,0.15)$</td>
<td>0.40 [0.26;0.55]</td>
</tr>
<tr>
<td>Neutral Technology, $\rho_e$</td>
<td>$\mathcal{B}(0.70,0.15)$</td>
<td>0.92 [0.88;0.95]</td>
</tr>
<tr>
<td>Government Consumption, $\rho_G$</td>
<td>$\mathcal{B}(0.70,0.15)$</td>
<td>0.95 [0.91;0.98]</td>
</tr>
<tr>
<td>Wage Markup Shock, $\rho_{\zeta_h}$</td>
<td>$\mathcal{B}(0.70,0.15)$</td>
<td>0.94 [0.91;0.97]</td>
</tr>
<tr>
<td>Price Markup, $\rho_\tau$</td>
<td>$\mathcal{B}(0.70,0.15)$</td>
<td>0.79 [0.68;0.89]</td>
</tr>
<tr>
<td>Risk Premium, $\rho_{rp}$</td>
<td>$\mathcal{B}(0.70,0.15)$</td>
<td>0.78 [0.66;0.89]</td>
</tr>
<tr>
<td>Wage Markup, MA(1), $\rho_{\zeta_h,MA}$</td>
<td>$\mathcal{B}(0.50,0.15)$</td>
<td>0.89 [0.83;0.95]</td>
</tr>
<tr>
<td>Price Markup, MA(1), $\rho_{\tau,MA}$</td>
<td>$\mathcal{B}(0.50,0.15)$</td>
<td>0.51 [0.28;0.72]</td>
</tr>
</tbody>
</table>

Notes: Sample: 1965Q1 to 2007Q4. Observables: real GDP growth (per capita), real investment growth (per capita), real consumption growth (per capita), federal funds rate, wage inflation (compensation), price inflation (PCE). $\mathcal{B}$, $\mathcal{IG}$ and $\mathcal{N}$ denote beta, inverse gamma and normal distributions, respectively. Posterior distributions based on 15 MCMCs each with 500,000 draws (150,000 used for burn-in). Acceptance rates around 23 percent.
Figure 1: Kimball vs. Dixit-Stiglitz -- Demand Curves, Profits and Optimal Prices

Demand Curves

- **Dixit-Stiglitz** ($\psi_p = 0$)
- **Kimball** ($\psi_p = -12$)

Profit Functions

- **Dixit-Stiglitz** ($\psi_p = 0$)
- **Kimball** ($\psi_p = -12$)

Kimball: Profit Functions

- $p_f^{opt} = 0.9970$
- $p_f^{opt} = 1.0000$
- $p_f^{opt} = 1.0045$

Dixit-Stiglitz: Profit Functions

- $p_f^{opt} = 0.950$
- $p_f^{opt} = 1.000$
- $p_f^{opt} = 1.050$

Kimball: Optimal Relative Price

Dixit-Stiglitz: Optimal Relative Price

- Nonlinear Model
- Linearized Model

% Change in $p_f^{opt}$ vs. % Change in Marginal Cost

Nonlinear Model: % Change in $p_f^{opt}$ vs. % Change in Marginal Cost
Figure 2: Impulse Responses to a 1% Discount Factor Shock

Panel A: ZLB Not Imposed

Panel B: ZLB Imposed

Nonlinear Model
Linearized Model
Figure 3: Stochastic Simulation of Nonlinear and Linearized Model

Panel A: Nonlinear Model

Panel B: Linearized Model

- Inflation
- Output Gap
- Nominal Interest Rate
- Wage Inflation

Quarters

Annualized Percent

Percent

Nominal Interest Rate

Wage Inflation

-5 0 5 10
-5 0 5 10
-5 0 5 10
-5 0 5 10

2000 4000 6000 8000 10000
Figure 4: Price and Wage Phillips Curves

Price Phillips Curve (Kimball)

Wage Phillips Curve (Kimball)
Figure 5: The U.S. Great Recession: Data vs. Estimated Medium-Sized Model

Notes: Data and model variables expressed in deviation from no-Great Recession baseline. Data from Christiano, Eichenbaum and Trabandt (2015)
Figure 6: Phillips Curves in Estimated Medium-Sized Model

Risk Premium Shocks Only

Monetary Policy Shocks Only

Gov. Cons. Shocks Only

Technology Shocks Only

Investment Shocks Only

Price Markup Shocks Only

Wage Markup Shocks Only

All Shocks

- Inflation vs. Negative Output Gap (% of potential actual GDP)
- Linearized Model (○)
- Nonlinear Model (+)
Figure 7: Phillips Curves in Estimated Model Conditional on an 8-Quarter Liquidity Trap

- Risk Premium Shocks Only
- Monetary Policy Shocks Only
- Gov. Cons. Shocks Only
- Technology Shocks Only
- Investment Shocks Only
- Price Markup Shocks Only
- Wage Markup Shocks Only
- All Shocks

Inflation and output gap in deviation from baseline -- annualized percentage points and percentage points, respectively. Baseline is a demand shock driven deep recession that triggers a liquidity trap where the ZLB is expected to bind for 8 quarters. Random shocks from estimated model hit in the first quarter when ZLB binds in the baseline. Inflation and output gap shown one year after random shocks have hit.
Figure 8: Densities of Data vs. Stochastic Model Simulations

- **Core PCE Inflation**
- **Wage Inflation (Hourly Earnings)**
- **Real GDP Growth**
- **Federal Funds Rate**

Data (1965Q1-2007Q4)  | Linearized Medium-Sized Model  | Nonlinear Medium-Sized Model