Aging, Secular Stagnation and the Business Cycle

Prepared by Callum Jones∗

March 2018

Abstract

As of 2015, U.S. log output per capita was 12 percent below what its pre-2008 linear trend would predict. To understand why, I develop and estimate a model of the US with demographics, real and monetary shocks, and the occasionally binding ZLB on nominal rates. Demographic changes generate slow-moving trends in the real interest rate, employment, and productivity. I find that demographics alone can explain one-third of the gap between log output per capita and its linear trend in 2015. Demographics also lowered real rates, causing the ZLB to bind between 2009 and 2015, contributing to the slow recovery after the Great Recession.

JEL Classification Numbers: E3, E4, E5, J11.

Keywords: Great Recession, Demographics, Zero Lower Bound, Forward Guidance.

Author’s E-Mail Address: CJones3@imf.org.

∗ I am particularly thankful to Virgiliu Midrigan for valuable advice. I thank Saki Bigio, Jaroslav Borovicka, Katarina Borovickova, Tom Cooley, Tim Cogley, Mark Gertler, Dan Greenwald, Mariano Kulish, Julian Kozlowski, Steven Pennings, and Thomas Philippson for helpful discussions. I also thank participants of a number of seminars for helpful suggestions. The Appendix is available online at https://callumjones.github.io. All remaining errors are mine alone.
1 Introduction

There has been much attention on the *secular stagnation* trends that have emerged in the US economy since the Great Recession. As of 2015, log output per capita was 12% lower than its 1950 to 2007 linear trend would predict. The Federal Reserve has held the Fed Funds rate at or near its lower bound since 2009, while the real interest rate declined from about 4% in 1995 to 1% in 2015, and the employment-population ratio has fallen from 64% in 2000 to just under 60% in 2016. Figure 1 illustrates these series.

Economists have proposed several competing theories for these observations. Summers (2014), for example, argues that the US has entered a period of secular stagnation characterized by low real interest rates and constrained monetary policy. Gordon (2016) suggests that supply-side factors are key, as productivity growth has slowed over time. Rogoff (2015) argues that slow output growth is driven by financial factors and a protracted period of deleveraging, while Jones, Midrigan and Philippon (2017) show that a protracted period of deleveraging can be highly contractionary when monetary policy is constrained.

The aging of the US population can explain, in theory, why real interest rates are low and therefore why the zero lower bound (ZLB) binds – an individual’s savings changes with her age, and an economy with a higher fraction of older people has more savings and lower interest rates (Eggertsson, Mehrotra and Robbins, 2017). The aging of the US population can also explain why productivity growth and the employment-population ratio are low — younger workers have faster productivity growth and older workers work fewer hours.

This paper develops a tractable model of the US economy to quantitatively account for the gap in observed log output from its simple pre-crisis linear trend. My model has a rich demographic structure where individuals in overlapping generations can live for up to 80 years. They choose their labor supply and savings each period, borrow when young, and consume their savings when old. These demographic features generate slow-moving trends over many decades. I model business cycle fluctuations around the demographic trend with a set of aggregate shocks typically featured in models of the business cycle (for example, as in Smets and Wouters, 2007). Monetary policy has a role because of nominal rigidities, and the Fed Funds rate is set by a monetary policy rule but is subject to the ZLB.

It is important to model the aging of the population alongside the occasionally binding ZLB constraint because the aging of the population causes real and nominal interest rates to decline, so that the ZLB is more likely to bind following a shock. There are therefore highly nonlinear interactions between demographic trends and constrained monetary policy.
that I am able to capture in my estimated model. To handle foreseen demographic changes
together with the nonlinearities associated with the ZLB, I combine the piecewise-linear
solution methods recently developed by Kulish and Pagan (2016) and Jones (2017).

My key contribution is to compare the importance, for output growth, of long-run trends
driven by demographic changes vis-à-vis business cycle shocks. By doing so, I am able to
assess the importance of demographic trends for output growth along two dimensions: first,
directly, through the effect that demographic trends have had on aggregate labor and capital
over time, and second, indirectly, through the effect that demographics have had in causing
the real interest rate to fall and the ZLB to bind, thereby constraining the response of
monetary policy to adverse shocks that arose from 2008 onwards.

My second contribution is methodological. I show how to parsimoniously model anticipated
demographic trends alongside the ZLB and a standard set of business cycle shocks such
that it becomes feasible to apply likelihood methods to estimate the model’s parameters. In
particular, I show that the model has a convenient representation whereby demographics
can be approximated by anticipated changes to the parameters of the model. I can exploit
the resulting computational advantages to filter quarterly output and consumption for the
shocks that generate those series, while accounting for the trends caused by demographics.
Furthermore, with my estimation method, I am able to account for the stimulatory effects
of forward guidance by disciplining, with survey data, expectations of how long the ZLB is
expected to last each quarter between 2009 and 2015.

I calibrate the demographic trends from 1940 and projected out to 2070, and use my
methodology to estimate the model’s business cycle shocks from 1984 to 2015. I find that
between 1990 and 2015, changes in the composition of the workforce caused by the aging of
the population explain a decline of around 2 percentage points in the labor force participation
rate, with a further decline of 4 percentage points expected by 2035. This contraction of the
labor force together with a decline in capital accumulation caused by fewer workers saving
for retirement drag down output growth. I therefore estimate that demographic changes
alone can explain about 4 percentage points of the 12% gap in log output per capita in 2015,
relative to its long-run linear trend. Furthermore, demographic changes are responsible for
about a 1 percentage point decline in real interest rates and about a 2 percentage point
decline in the nominal interest rate between 1990 and 2015. Demographic trends, therefore,
reproduce well the long-run trends observed in the US economy.

The slow decline in real interest rates caused by demographics are also key for the

\[1\] These methods are related to Eggertsson and Woodford, (2013); and Guerrieri and Iacoviello, (2015).
propagation of shocks. I estimate that, holding the demographic profile constant from 1984, the ZLB would not have been a binding constraint at any time between 2009 and 2015. However, I also find that the contractionary effects of the ZLB were largely mitigated with forward guidance policy and that, absent forward guidance, US output per capita would have declined by a further 2 percentage points between 2011 and mid-2013.

The nonlinearities associated with the ZLB and forward guidance policy mean that it is a nontrivial problem to assess which shocks, other than demographics, were responsible for the decline in output following the Great Recession. To shed light on this, I hold fixed the expected duration of the ZLB each quarter between 2009 and 2015, and find that the contribution of aggregate shocks that proxy for financial distress both depress output, and are partly offset by positive government spending shocks between 2009 and 2012.

My model has a rich demographic structure. Workers accumulate human capital over the lifecycle and gradually withdraw from the labor market as they age. Workers also have age-specific labor productivities, which I calibrate using US Census and survey data. There are two exogenous demographic trends in the model. First, workers have age-specific and time-varying mortality rates, which are matched to observed and projected mortality profiles. Second, the size of the initial cohort each period is chosen to match observed changes in the relative size of the the 16-year-old cohort. The model’s lifecycle parameters are calibrated to individual level data, and the demographic trends are fully anticipated. I show that the demographics in my model closely track those that are observed in the US.

The unintentional bequests of those who die between periods are redistributed to the remaining living members of each generation, following Yaari (1965) and Blanchard (1985). I show that this insurance against mortality risk means that the consequences of demographic changes are approximated very well by slow-moving, anticipated changes to parameters of the preferences of a representative agent and to aggregate productivity. These parameters vary over time with changes in the composition of the population and can be readily constructed from population data and lifecycle parameters – in particular, the age-productivity profile and the disutility of labor supply by age. Approximating the model with these time-varying parameters reduces the model’s state-space substantially, and makes it feasible to conduct a Bayesian estimation with quarterly data and the ZLB.

Using this efficient representation, I estimate the parameters of the six aggregate shocks to the discount factor, productivity, markups, investment costs, monetary policy, and government spending, that drive fluctuations around the model’s demographic trend, using quarterly

\[ \text{See also } \text{Jones, Midrigan and Philippon (2017); Wu and Xia (2016); Swanson and Williams (2014).} \]
data on output, consumption, investment, the Fed Funds rate, and inflation between 1984 and 2015. During the ZLB period, I discipline the durations that the Fed Funds rate are expected to remain at zero with survey data from Blue Chip and the New York Fed (as in Jones, Midrigan and Philippon, 2017; Kulish, Morley and Robinson, 2017).

Estimating the model with demographic trends affects the posterior estimates of the model’s structural parameters and the interpretation of which shocks generate business cycles. The estimate of annual trend growth is 0.4% higher when demographic trends are not accounted for, as the changing composition of the workforce implied by those trends is favorable to growth across the estimation sample. Not accounting for slow-moving demographic trends also raises the importance (and persistence) of productivity shocks, and lowers the importance of markup shocks, in driving output and consumption.

Using the model, I explore in detail the extent to which demographic trends alone are responsible for the observed trends in aggregate variables. First, I find that the aging population has caused the real interest rate to decline by 1 percentage point from 1985 to 2015. This compares with the decline in the real interest rate implied by the observed capital-output ratio in the US of 1.6 percentage points between 1980 to 2015, and is in line with the conclusions reached by others (Gagnon et al., 2016; Eggertsson et al., 2017). The gradual fall in mortality rates causes a compositional shift towards an older population: longer expected lifetimes means an increased need for savings and capital for retirement consumption, contributing about a 1 percentage point decline in the real interest rate. Fertility shocks cause a slow-moving increase and then decrease in the real interest rate around the path implied by the fall in mortality rates because they change, over time, the relative size of labor and capital, affecting the marginal product of capital.

Second, I find that demographic trends can explain a decline in the employment-population ratio of 1.9 percentage points from 1990 to 2015. This compares to the observed decline of 3.5 percentage points from 1990 to 2015. In the model, the trend decline is mainly caused by changes in the composition of the workforce as workers live for longer and the labor force participation rate declines with age.

Third, I find that my model predicts that the growth rate of labor productivity – output per hour worked – has declined by about 0.5 percentage points from 1990 to 2011. This compares to the observed average decline of about 1 percentage point over the same period (Fernald, 2015). While much of the change is driven by variation in aggregate labor and capital, I find that changes in labor quality have been an important factor since the 1990s. In particular, when the baby boomer generation enters the workforce, there is an increase, and
subsequent decrease, in the growth rates of human capital and thus productivity.

I model the US as a closed economy. However, conclusions about the fall in real interest rates can, in theory, be sensitive to international factors. As in Gagnon et al. (2016) the relationship between an aging population and a falling real interest rate is arguably robust to open-economy considerations since, first, most advanced economies are experiencing similar demographic changes, and second, the observed magnitude and direction of international capital flows were opposite to the outflows that the aging of the US population should imply. Furthermore, Eggertsson et al. (2017) consider open-economy factors in their study of the real interest rate from the 1970s, finding that a decline in the real interest rate caused by capital inflows were roughly the same as an increase caused by the growth in US public debt – both factors are absent in my model.

Related literature My paper relates to a number of studies on the secular stagnation debate (Hamilton et al., 2015; Rogoff 2015; Bernanke 2015; Summers 2014). My emphasis on the importance of demographics in explaining the decline in the real interest rate is shared by the quantitative work of Carvalho et al. (2015), Gagnon et al. (2016), and Eggertsson et al. (2017). Other papers tie to the secular stagnation literature by studying the decline in long-run growth and its consequences (Gordon 2016; Antolin-Diaz et al. 2014; Fernahl 2015; Christiano et al. 2015). I complement this body of work by highlighting a distributional channel from demographics to output, by which the age-composition of the workforce interacts with the age-productivity curve. Furthermore, to the best of my knowledge, my approach is the first to jointly model demographics, the ZLB, and the business cycle, emphasizing their nonlinear interactions, to study the stagnation of output growth.

The paper also relates to research that studies the macroeconomic implications of demographic trends alone (see, for example, Aksy et al. 2015; Fujita and Fujiwara 2014; Backus et al. 2014; Bloom et al. 2001; Auerbach and Kotlikoff 1987) and empirical studies of the macroeconomic implications of demographic changes (see, for example, Bloom et al. 2011). My paper builds on the insights of the empirical work connecting demographic trends to output growth (Maestas et al. 2016), productivity growth (Feyer 2007), and the labor force participation rate (Aarsonson et al. 2014).

Finally, my paper relates to studies of how the propagation of transitory shocks can change as an economy undergoes structural changes, either anticipated or unanticipated (Kulish and Pagan 2016; Canova et al. 2015; Wong 2015; Jaimovich and Sin 2012; Fernandez-Villaverde et al. 2007). I build on this literature by outlining a methodology for explicitly accounting
for demographics in the estimation of a business cycle model, as anticipated shocks to the parameters of a business cycle model. I also focus on the nonlinear interactions that take place between the decline in the real interest rate and the ZLB.

2 Model

In this section, I develop a model in which trends in key macroeconomic variables are driven by population aging. The model features households of different ages, firms that face price adjustment costs, aggregate shocks, monetary policy, and the ZLB on nominal interest rates.

2.1 Households

Demographics Individuals are organized into overlapping generations. Each generation lives for a maximum of \( T \) periods, so that the age range of an individual is 0 to \( T - 1 \). Generation \( s \) is of mass \( n_t^s \) and comprised of a continuum of identical members of age \( s \). The size \( n_t^s \) is measured at the start of period \( t \). I abstract from trend population growth, so that, in the absence of exogenous changes in the size of the incoming population, the initial population size is normalized to \( n_t^0 = 1 \). The total size of the population at time \( t \) is:

\[
n_t = \sum_{s=0}^{T-1} n_t^s.
\]

(1)

Each period, a fraction of each cohort dies with the exogenous age-specific mortality rate \( \gamma_t^s \):

\[
n_{t+1}^s = (1 - \gamma_t^s) n_t^s.
\]

(2)

These mortality rates vary over time, for example, permanent decreases in mortality rates imply increases in longevity. A maximum lifespan of \( T \) implies \( \gamma_t^{T-1} = 1 \).

Household problem An individual of age \( s \) has the period utility function \( u(c_t^s, \ell_t^s) \) and chooses consumption \( c_t^s \), labor supply \( \ell_t^s \), and savings \( a_t^s \)—made up of capital \( k_t^s \) and one-period risk-free bonds \( b_t^s \)—to maximize lifetime utility. The value function of an individual of age \( s \) at period \( t \) is therefore:

\[
V_t^s = \max_{\{c_t^s, \ell_t^s, a_t^s\}} \left\{ \chi_t u(c_t^s, \ell_t^s) + \beta (1 - \gamma_t^s) \mathbb{E} V_{t+1}^{s+1} \right\},
\]

(3)

where the expectation is taken with respect to the aggregate stochastic shocks, \( \beta \) is the discount factor common to all individuals, and \( \chi_t \) is an autoregressive preference process subject to shocks and common to all individuals:

\[
\ln \chi_t = (1 - \rho_\chi) \ln \chi + \rho_\chi \ln \chi_{t-1} + \sigma_\chi \varepsilon_{\chi,t}.
\]

(4)
An individual’s age-specific mortality rate $1 - \gamma_s^t$ features in their discounting of the future.

The unintentional bequests made by individuals of a household who die between periods are aggregated and redistributed as an annuity to the remaining living households of the same generation, following Yaari (1965) and Blanchard (1985), eliminating mortality risk.

Individuals have age-specific productivities $z^s$, receive a transfer from the government $\xi^s_t$ which are described in more detail below, earn a return $R_t$ on last period’s bond holdings $b^s_{t-1}$, earn the rental rate on capital $r_t$ on last period’s capital holdings $k^s_{t-1}$, and receive $d^s_t$ for the redistributed unintentional bequest. Individuals pay a quadratic capital adjustment cost parameterized by $\phi_k$. These adjustment costs are subject to an aggregate, exogenous autoregressive process $\kappa_t$:

$$\ln \kappa_t = (1 - \rho_{\kappa}) \ln \kappa + \rho_{\kappa} \ln \kappa_{t-1} + \sigma_{\kappa} \varepsilon_{\kappa,t}. \quad (5)$$

The period budget constraint of the individual is therefore:

$$c^s_t + k^s_t + \frac{\phi_k}{2} \left( \frac{k^s_t}{k^s_{t-1}} - 1 \right)^2 k^{s-1}_{t-1} + \frac{b^s_t}{p_t R_t} \leq z^s w_t k^s_t (1 - \tau_t) + \xi^s_t + \frac{b^{s-1}_{t-1}}{p_t} + (r_t + 1 - \delta) k^{s-1}_{t-1} + d^s_t. \quad (6)$$

Consumption in the last period of life equals the return on remaining assets:

$$c^T_t \leq (r_t + 1 - \delta) k^{T-1}_{t-1} + \frac{R_{t-1} b^{T-1}_{t-1}}{p_t}. \quad (7)$$

By assumption, an individual retires fully from the labor market in her last period of life. Individuals are born with zero wealth, so that $k^0_t = 0$ and $b^0_t = 0$ for all $t$, and nominal bonds are in net zero supply $b_t = \sum_s n^s_t b^s_t = 0$. Substituting in for the unintentional bequests $d^s_t$, and denoting the marginal utility of wealth of an individual of age $s$ in time $t$ by $\lambda^s_t$, the optimal choice of risk-free bonds implies the standard Euler equation:

$$\mathbb{E}_t \frac{\lambda^{s+1}_{t+1}}{\lambda^s_t} = \frac{1}{\beta} \mathbb{E}_{t+1} \frac{\Pi_{t+1}}{R_t}, \quad (8)$$

where $\Pi_t = p_t/p_{t-1}$ is the rate of inflation.

### 2.2 Firms

There are a continuum of intermediate goods firms who hire capital and labor from households to supply a substitutable good $y_t(i)$ to final goods producers, who in turn sell the final goods...
good to consumers at price $p_t$. Aggregate capital is the sum of each cohort’s capital savings: $k_t = \sum s n_t^s k_t^s$, while aggregate labor hired by the firm is in efficiency units of labor $\ell_t = \sum s z^n_t \ell_t^s$. I assume that firms use the capital of the deceased in production. The elasticity of substitution between intermediate goods $\xi_t$ is subject to stochastic shocks, which generates time-varying markups over marginal costs. The production function is:

$$y_t(i) = \mu_t^{1-\alpha} \left( k_{t-1}(i) \right)^\alpha \left( Z_t \ell_t(i) \right)^{1-\alpha},$$

(9)

where $\frac{Z_t}{Z_{t-1}} = z$ governs trend growth, and $\mu_t$ is an aggregate autoregressive TFP process:

$$\ln \mu_t = (1 - \rho_\mu) \ln \mu + \rho_\mu \ln \mu_{t-1} + \sigma_\mu \varepsilon_{\mu,t},$$

(10)

where $\varepsilon_{\mu,t}$ is a standard normal innovation. Intermediate goods-producing firms face a Rotemberg quadratic cost of adjusting prices, parameterized by $\phi_p$. Denoting $mc_t$ as the Lagrange multiplier on the firm’s cost minimization problem, the rental rate on capital is:

$$r_t = \alpha mc_t \frac{y_t}{k_{t-1}},$$

(11)

and the wage is:

$$w_t = (1 - \alpha) mc_t \frac{y_t}{\ell_t}.$$  

(12)

The first order condition on the optimal choice of price reseting is:

$$\beta \phi_p E_t \frac{\lambda_{t+1}}{\lambda_t} \frac{y_{t+1}}{y_t} \left[ \frac{\Pi_{t+1}}{\Pi_t} - 1 \right] = \xi_t - \xi_t mc_t + \phi_p \left[ \frac{\Pi_t}{\Pi_{t-1}} - 1 \right] \left[ \frac{\Pi_t}{\Pi_{t-1}} \right],$$

(13)

where $\xi_t$ is an autoregressive process for the elasticity of substitution between intermediate goods:

$$\ln \xi_t = (1 - \rho_\xi) \ln \xi + \rho_\xi \ln \xi_{t-1} + \sigma_\xi \varepsilon_{\xi,t}.$$  

(14)


2.3 Monetary Policy

Monetary policy operates in one of two possible regimes. In the first regime, the nominal interest rate can be set according to a Taylor rule as in Smets and Wouters (2007):

$$\frac{R_t}{R} = \left( \frac{R_{t-1}}{R} \right) \phi_R \left( \frac{\Pi_t}{\Pi_t^*} \right)^{(1-\phi_R)\phi_\Pi} \left( \frac{y_t}{y_t^F} \right)^{(1-\phi_R)\phi_y} \left( \frac{y_t/y_{t-1}}{y_t^F/y_{t-1}^F} \right) \phi_y \exp(\sigma_R \varepsilon_{R,t}).$$

(15)
The nominal interest rate $R_t$ responds to deviations in inflation $\Pi_t$ from a target rate $\Pi^*$, deviations in output $y_t$ from its flexible-price level $y^*_t$, and the growth rate of output relative to the growth rate of potential output, and is subject to stochastic shocks $\varepsilon_{R,t}$.

In the second regime, the nominal interest rate is at zero:

$$\log(R_t) = 0. \quad (16)$$

Monetary policy can enter the ZLB regime in two ways: first, if the Taylor rule calls for negative nominal interest rates, and second, if the Federal Reserve has announced, or has previously announced, an extension of the ZLB beyond that implied by the constraint and the Taylor Rule. I assume therefore that the Fed can manipulate expectations of how the path of interest rates evolves, as in Eggertsson and Woodford (2003) and Werning (2015). In estimation, I use survey data from the New York Federal Reserve to discipline the expected duration of the zero interest rate regime during the 2009 to 2015 period.

### 2.4 Government

The government taxes labor income at the rate $\tau_t$ to fund a pay-as-you-go social security system. I follow Attanasio, Kitao and Violante (2007) in specifying that the benefit paid each period above an eligibility age $T^*$ depends on three elements: the expected remaining life of the recipient, the accumulated pre-tax labor income of the worker, and a parameter $\omega$ governing the replacement rate of past earnings. Denote by $W^s_t$ accumulated gross lifetime earnings, defined recursively as:

$$W^s_t = \begin{cases} w_t z^s \ell^s_t + W^{s-1}_{t-1}, & \text{if } s < T^* \\ W^{s-1}_{t-1}, & \text{if } s \geq T^*. \end{cases} \quad (17)$$

The amount $\xi^s_t$ redistributed to an agent of age $s \geq T^*$ depends on $W^s_t$:

$$\xi^s_t = \omega \frac{W^s_t}{(T^* - 1)}, \quad (18)$$

where the denominator reflects the amount of time that $W^s_t$ is accumulated over. For those younger than the eligibility age $T^*$, the transfer $\xi^s_t = 0$. The government budget constraint is:

$$\sum_s n^s_t \xi^s_t = \sum_s n^s_t z^s w_t \ell^s_t \tau_t. \quad (19)$$

The tax rate $\tau_t$ adjusts to equalize social security outlays and tax revenues. 

---

5 I abstract from questions about the sustainability of pension systems in an aging society: see Attanasio, Kitao and Violante (2007).
Finally, the government levies a lump-sum tax on households to pay for government expenditures $g_t$, which are assumed to be autoregressive and subject to stochastic shocks:

$$\ln g_t = (1 - \rho_g) \ln g + \rho_g \ln g_{t-1} + \sigma_g \varepsilon_{g,t}. \quad (20)$$

### 3 Approximation

To study the interactions between demographics and monetary policy, I set the model at a quarterly frequency and estimate the shocks driving the business cycle. To solve for a rational expectations equilibrium in this model would require knowing the distribution of wealth across individuals. To deal with the computational challenges this raises, I argue in this section that the model can be very well approximated by a representative agent framework with exogenous time-varying parameters that affect the preferences of a representative household and aggregate TFP.

#### 3.1 Derivation

I derive the approximation by considering a model identical to the decentralized one described in Section 2 but with one exception: instead of being born with zero wealth and making period-by-period asset trades, each generation is born with positive wealth and can trade claims to future consumption before they are born. The redistribution scheme for unintentional bequests means that individuals are insured against the idiosyncratic uncertainty associated with mortality risk. As a result, in a time-zero trading setup, they trade in a complete markets environment so that the ratio of marginal utilities of wealth across individuals is the same at any point in time.

This alternative setup can be solved using the equivalent problem of a social planner who maximizes a weighted sum of each individual’s utility function. To solve this problem, I first determine how a planner distributes aggregate consumption and aggregate labor supply between individuals alive each period. Given these allocations, I then solve the problem where the planner maximizes aggregate consumption, capital, and labor supply subject to the economy’s resource constraint. This approach reflects the aggregation arguments made in Constantinides (1982) and Maliar and Maliar (2003) and, while the solution to this alternative time-zero approach and the full model in Section 2 can have very different implications at the individual level, I will argue that the aggregate dynamics are indistinguishable.

In the Appendix, I follow this approach to show that the aggregate outcomes of the following two models are the same: (i) a model where individuals are born with positive
wealth and can trade claims to future consumption before they are born, and who have the period utility function over consumption \( c_t \) and hours \( \ell_t \) of the type \( \frac{(c_t^\sigma)}{1-\sigma} - v_s(\ell_t)^{1+\varphi} \), and (ii) a model with a representative agent who has preferences over aggregate consumption \( c_t \) and aggregate **efficiency units** of labor \( \ell_t \) with the period utility function taking the form:

\[
U(c_t, \ell_t) = \phi_t \frac{c_t^{1-\sigma}}{1-\sigma} - v_t \frac{\ell_t^{1+\varphi}}{1+\varphi}.
\] (21)

The representative agent’s problem will be to maximize this utility function over time with discount factor \( \beta_s t \) by choosing aggregate consumption \( c_t \), efficiency units of labor \( \ell_t \) and capital \( k_t \) subject to the economy’s resource constraint and its production function \( y_t = \theta_t^{1-\alpha} k_t^\alpha \ell_t^{1-\alpha} \), and where the relationship between the efficiency units of labor and aggregate hours is \( \ell_t = A_t h_t \). In the time-zero trading problem demographics therefore affects the aggregate economy through five time-varying and exogenous parameters: \( \phi_t, v_t, \theta_t, A_t, \) and \( s_t \). First, demographics affects aggregate output through productivity \( \theta_t \), and second, through the labor-input relationship by \( A_t \). The third is the parameter multiplying the marginal utility of consumption \( \phi_t \). The fourth affects the aggregate disutility of providing labor \( v_t \), which, together with the preference shock \( \phi_t \), attaches to the labor wedge, affecting the incentive to supply labor. The fifth, \( s_t \), is a adjustment to the discount factor caused by changing longevity. Next, I discuss the microfoundations of \( \phi_t, v_t, \theta_t, \) and \( A_t \).

The reason for why the approximation works well is similar to the near aggregation arguments made by [Krusell and Smith (1998)](Krusell1998) for why average aggregate quantities are sufficient, in practice, to forecast prices: individuals’ policy functions are sufficiently linear so that, in practice, it does not matter how the planner distributes wealth across individuals. So, in my problem, while the allocations across individuals can be very different depending on the prevailing distribution of wealth, the aggregate implications – my primary focus – are not. This is discussed in more detail in the Appendix and is explored with examples.

### 3.2 Productivity and Labor-Input Parameters

The first time-varying parameters \( \theta_t \) and \( A_t \) affect productivity, shifting total output and shifting labor-inputs. Substituting the definition of efficiency units of labor \( \ell_t = A_t h_t \) into the production function gives:

\[
y_t = \theta_t^{1-\alpha} k_t^\alpha (A_t h_t)^{1-\alpha}.
\] (22)

In the Appendix, I show that \( \theta_t \) and \( A_t \) in the time-zero trading problem are:

\[
\theta_t = \sum_s n_t^s z^s, \quad \text{and} \quad A_t = \frac{\sum_s n_t^s (z^s)^{1+1/\varphi} (v^s \lambda^s)^{-1/\varphi}}{\sum_s n_t^s (z^s)^{1/\varphi} (v^s \lambda^s)^{-1/\varphi}},
\] (23)
where the value $\hat{z}^s = z^s / \theta_t$ denotes individual $s$’s skill level relative to the average skill level in the economy $\theta_t$, and the $\lambda^s$ parameters are the Pareto weights attached to an individual of age $s$. The shock $\theta_t$ encodes changes in output caused by changes in the size of the workforce and its composition over idiosyncratic skill levels. In particular, larger populations have higher levels of $\theta_t$, as do populations in which the composition of more productive workers in the labor market is higher.

The time-varying parameter $A_t$ affects the hours needed to obtain an effective unit of labor, and is a population-weighted average of relative productivity and the disutility of providing labor. When labor supply is inelastic, $\theta_t$ and $A_t$ together affect productivity by $\sum_s n_t^s \hat{z}^s / \sum_s n_t^s$: the productivity adjustment in this case reflects only the age-composition of the population. In principle, $\theta_t$ and $A_t$ are straightforward to compute, requiring only hours by age and assumptions about the age-productivity profile, the age-disutility of work profile, Pareto weights for each generation, and the inverse Frisch labor supply elasticity $\varphi$.

### 3.3 Consumption Preference

The second time-varying parameter resulting from demographic trends attaches to the marginal utility of consumption:

$$\phi_t c_t^{-\sigma} = \lambda_t. \quad (24)$$

where $\lambda_t$ is the Lagrange multiplier on the economy’s resource constraint. In the Appendix, I show that $\phi_t$ has a simple expression:

$$\phi_t = \left[ \sum_s n_t^s (\lambda^s)^{\frac{1}{\varphi}} \right]^{\sigma}. \quad (25)$$

The parameter $\phi_t$ simply maps to the size of the population at each point in time, and is increasing in the curvature of the utility function $\sigma$, so that it scales the marginal value of wealth in the economy.

### 3.4 Marginal Disutility of Labor

The third time-varying parameter, $v_t$, is an aggregate trend in the marginal disutility of labor, shown in the Appendix to be:

$$v_t = \left[ \sum_s n_t^s (\hat{z}^s)^{\frac{1}{\varphi} + 1} (v^s \lambda^s)^{-\frac{1}{\varphi}} \right]^{-\varphi}, \quad (26)$$

so that $v_t$ is a population-weighted average of age-specific disutilities of providing labor. Equating the marginal utility of consumption and the marginal disutility of labor, and
substituting in for hours worked gives:

\[ \frac{w_t}{\ell_t^{\sigma}} = \frac{v_t}{\phi_t}. \tag{27} \]

This says that a component of the labor wedge measured with efficiency units of labor is the ratio of the two demographic time-varying parameters \(v_t/\phi_t\). An increase in the average disutility of labor increases the labor wedge, so that a level increase in \(v^s\) scales linearly \(v_t\). Furthermore, changes in the distribution of the population towards younger workers who have lower disutility of providing labor reduces \(v_t\).

### 3.5 Remaining Equilibrium Conditions

In addition to the firms’ equilibrium conditions and monetary policy, the representative household has an optimal choice over consumption:

\[ \lambda_t = \phi_t \chi_t c_t^{1-\sigma}, \tag{28} \]

and an optimal choice of efficiency units of labor:

\[ \lambda_t = \frac{v_t \chi_t (A_t h_t)^{\mu}}{w_t (1 - \tau_t)}. \tag{29} \]

There is an Euler equation associated with the choice of nominal bonds:

\[ \mathbb{E}_t \frac{\lambda_{t+1}}{\lambda_t} = \frac{1}{\beta} \mathbb{E}_t \frac{\Pi_{t+1}}{R_t}, \tag{30} \]

and an Euler equation associated with the choice of capital:

\[ \mathbb{E}_t \frac{\lambda_{t+1}}{\lambda_t} = \mathbb{E}_t \frac{1}{\beta} \left[ 1 + \phi_k \kappa_t \left( \frac{k_t}{k_{t-1}} \kappa_t - 1 \right) \right] \left[ \frac{\phi_k}{2} \left( \frac{k_t^2}{k_{t-1}^2} \kappa_t^2 - 1 \right) + 1 - \delta + r_{t+1} \right] \lambda_t, \tag{31} \]

The goods market clears:

\[ y_t = c_t + k_t + g_t - (1 - \delta) k_{t-1} + \Phi_k \left( \frac{k_t}{k_{t-1}} - 1 \right)^2 k_{t-1} + \frac{\phi_p}{2} \left[ \frac{\Pi_t}{\Pi_{t-1}} - 1 \right]^2 y_t, \tag{32} \]

where \(g_t\) is exogenous government spending. The full set of equations characterizing the model’s equilibrium are in the Appendix. I assume all demographic changes, and the path of the time-varying parameters associated with those demographic changes, are perfectly foreseen. I linearize each equation around the steady-state associated with the model’s parameters at each point in time, and use the solution method described in the next section.
4 Estimation

Next, I discuss how the parameters of the model are set, including the calibration of the lifecycle parameters of the model, the transitory and permanent demographic changes which drive the trends in the model, and the estimation of the transitory shocks that govern the business cycle around the model’s demographic trends.

4.1 Assigned and Calibrated Parameters

Before estimation, I calibrate a subset of the parameters and the demographic shocks.

4.1.1 Lifecycle Parameters

The model is quarterly. Individuals begin life at 16 years of age and live for at most 80 more years, up to age 95. Full retirement is only imposed in the last period of life.

I calibrate the disutility of providing labor $v^s$ with a scaled cumulative density function of a normal distribution, so that $v^s$ increasing in $s$. This specification is motivated by studies which link the disutility of work to deteriorating health. The parameters of the function for $v^s$ are chosen so that the labor force participation rates by age broadly match those observed in 2000. For the social security system, I set the replacement ratio of accumulated earnings $\lambda$ to 46.7%, the same value that is used in Attanasio et al. (2007). Retirement benefits are received only from age 65 on ($T^* = 49$).

I calibrate the age-productivity parameters $z^s$ to the age-experience earnings profile. I follow Elsby and Shapiro (2012) in constructing the log experience-earnings profile using deflated data on full-time, full-year workers. The data is decennial Census data from 1960 to 2000, and annual American Community Survey data from 2001 to 2007. To minimize cohort effects, I pool, across years, high school dropouts, high school graduates, those with some college education, and those who have completed college or higher education.

---

6During the transition, the proportion of people who live past 95 is small – about 0.5% of the population. 7Given the low choice of labor supply at older ages, this choice is not too important. 8Kulish et al. (2010) show how unanticipated changes in life expectancy can change labor supply decisions deep into the period of retirement, as older workers with very few assets remaining return to the workforce to fund consumption during their unanticipated increase in the lifespan. 9Kulish et al. (2010) use this specification and also choose to make the function time-varying with increases in life-expectancy, with the result that the disutility of labor from employment declines in the gap between age and life-expectancy (see also, Bloom et al. 2011). I keep it constant to ensure age-participation rates do not vary significantly over time. 10Computed off IPUMS-USA extracts. A full description is given in the Appendix. 11Clearly, a finer calibration would distinguish between workers of different education levels. In robustness exercises reported in the appendix, I distinguish between these education groups and analyze how anticipated changes in the earnings profile map to labor supply decisions. In particular, Kong et al. (2016) find flattening
Figure 2 plots the earnings-profile over age. The estimates imply a peak increase in wages of about 134% at age 45, before gradually declining around the age of 50. This profile is in line with the estimate of Guvenen et al. (2015) who find an increase in the earnings of the mean worker of 127%. There is less reliable data on the earnings of older workers, so after age 65, I calibrate the productivity of workers to decay by 20% a year.

I assess the calibration of the lifecycle parameters in Figure 2 by plotting, in Panel B, the labor force participation rate by age, in 2000, in the model and the data, and in Panel C, the age-profile of assets normalized to asset holdings at age 60, in 2013, in the model and in the Survey of Consumer Finances. The calibrated age-productivity profile and the disutility of labor parameter at each age implies labor force participation rates that rise when young, flatten out during an individual’s prime working life, and decline rapidly around retirement ages. The lifecycle asset-profile is hump-shaped and peaks around 60 years old. Individuals borrow when young in the model and begin accumulating assets around the age of 35.

4.1.2 Mortality Profiles

Next, I calibrate the mortality probabilities of each generation during the 80 years they could possibly live, $\gamma_s^t$, to the actuarial probabilities reported by the Social Security Administration. By calibrating to these probabilities, I also match changes in the life expectancy of each generation over time, conditional on an individual reaching 16 years of age. The values used are the cohort-specific survival rates computed for the cohort year of birth. These profiles include both observed survival rates of cohorts up to their current age, and extrapolated survival rates based on the Social Security Administration’s forecasts of life expectancy. I assume that all changes to these actuarial probabilities are exogenous and perfectly foreseen. For the initial $\gamma_s^t$ profile, I use the survival probabilities reported for those born in 1900 onwards. For those cohorts born before 1900 but who are alive in 1940, I use extrapolated values of the survival probabilities. Under the calibration, between 1950 and 2020, life expectancy for 16 year olds increases from about 77 years to about 85 years.
4.1.3 Incoming Cohort Size Shocks

I choose anticipated shocks to the size of the incoming cohort so that the change in the observed cohort share is the same as the change in the model cohort share. This ensures that the model captures the baby-boomer generation and imperfectly captures changes in the population distribution due to, for example, immigration. I assume that changes to the incoming population beyond 2015 decay to zero, so that the population distribution converges to the steady-state implied by the mortality profile that is constant from 2070.

I plot in Panel D of Figure 2 the median age of the population above 16 years of age implied by the calibrated mortality profiles and the incoming cohort size shocks. The profile tracks well the corresponding median age of those above 16 years of age in the data, declining from around the 1960s to around the 1980s to about 37 years of age, before steadily increasing as the baby-boomer population ages and longevity continues to rise.

4.1.4 Preference and Nominal Parameters

I calibrate the remaining parameters to values which imply steady-state capital-output ratios that align with those in the Bureau of Labor Studies’ Multifactor Productivity (BLS-MFP) program (see, Fernald, 2012). I set capital to depreciate by $\delta = 10.6\%$ a year. The quarterly discount factor $\beta$ is set to 0.99875. The capital share $\alpha$ is set to $1/3$, the average of the capital share reported by Fernald (2012) over 1948 to 2015. The intertemporal elasticity of substitution $\sigma$ is set to 1, and the inverse Frisch elasticity of labor supply $\phi$ is set to 2, in line with the estimates of Reichling and Whalen (2012) and with the analysis of Rios-Rull et al. (2012). Both values are also in the range considered by Auerbach and Kotlikoff (1987) in a computational overlapping generations model. These parameters imply a capital-output ratio in 2000 of about 2.7, which is the observed capital-output ratio in the BLS-MFP in 2000.

I calibrate a small set of the parameters describing the nominal side of the economy to values commonly used in the literature. The steady-state value of $\xi$ is set at 8, which implies a steady-state markup over marginal costs of $\xi/(\xi - 1) = 14\%$. The parameter governing the quadratic cost of capital adjustment $\phi_k$ is set to 40. The slope of the Philips curve, $\epsilon_p$,

---

15Choosing initial population shocks to matching the changes is necessary because the model is initialized at the 1940 steady-state and matching the actual cohort sizes would imply very large and counterfactual initial population shocks.

16To understand how important migration is to changing the population structure, I considered statistics on the characteristics of immigrants, reported by The Migration Policy Institute, at www.migrationpolicy.org. The median age of immigrants in the US has increased from the 1980s/1990s to 2010. In 2014, the median age of immigrants was 43 years, compared to the 36 years in the general population.

17These observed capital-output ratios vary between 2 and 2.7 over the period 1950 to 2013. A description of the BLS-MFP dataset used is given in the Appendix.
and therefore the quadratic cost of price adjustment $\phi_p$ is estimated in the next section. I estimate this parameter because it is key for determining the behavior of inflation and therefore the real interest rate, particularly in response to changes in the expected ZLB duration which are due to forward guidance announcements. The annual inflation target $\Pi^*$ is set to 2.2%. Finally, I use the values that Smets and Wouters (2007) estimate for the Taylor rule parameters, which at their posterior mode are: $\phi_r = 0.81$, $\phi_\pi = 2.03$, $\phi_y = 0.08$, and $\phi_g = 0.22$.

4.2 Bayesian Estimation

I next use Bayesian likelihood techniques to estimate the slope of the Philips curve, trend growth, and the model’s shocks that drive business cycle fluctuations around the demographic trend. I first discuss the solution method used to efficiently incorporate demographic trends as anticipated changes to the model’s parameters, and also to impose the ZLB.

4.2.1 Solution Method

I adapt a methodology based on anticipated structural changes to the parameters of the model, where the sequences of demographic wedges are taken as an anticipated path of the structural parameters of the model and where the ZLB is handled with a regime-switching procedure. The standard linear, time-invariant approximation cannot be used for two reasons: first, because demographic changes are approximated as changes to the parameters of the model which are fully anticipated by agents in the economy, and second, because the ZLB is an occasionally binding, nonlinear constraint. The final structure of the economy is the one that arises at the expected completion of the demographic transition and under Taylor-rule policy. Under my calibration, this final demographic structure applies from the year 2070 onwards.

To fix notation, I first describe the time-invariant approximation of a rational-expectations model of the form $\mathbf{x}_t = \Psi(\mathbf{x}_{t-1}, \mathbb{E}_t \mathbf{x}_{t+1}, \mathbf{\varepsilon}_t)$ where $\mathbf{x}_t$ is the vector of model variables (state and jump), and $\mathbf{\varepsilon}_t$ is a vector of exogenous unanticipated shocks. The linearized rational-expectations approximation of the model is:

$$
\mathbf{A} \mathbf{x}_t = \mathbf{C} + \mathbf{B} \mathbf{x}_{t-1} + \mathbf{D} \mathbb{E}_t \mathbf{x}_{t+1} + \mathbf{F} \mathbf{\varepsilon}_t,
$$

(33)

---

18See also Canova et al. (2015), Kulish and Pagan (2016), and Fernández-Villaverde et al. (2007). More generally, Jones (2017) discusses how the ZLB is a change in the structural parameters of the monetary policy rule that applies for a state-contingent period, or to stimulate the economy with forward guidance. Since the structure of the economy changes at the ZLB, the interpretation of transitory structural shocks can differ when the ZLB binds, as discussed in Section 5.3.
where \( A, B, C, D, \) and \( F \) are matrices that encode the structural equations of the model. A solution to \((33)\), following \( \text{Binder and Pesaran (1995)} \) or \( \text{Sims (2002)} \), is:

\[
x_t = J + Q x_{t-1} + G \epsilon_t,
\]

(34)

where \( J, Q, \) and \( G \) are conformable matrices which are functions of \( A, B, C, D, \) and \( F \).

In my model, agents have time-varying beliefs about the evolution of the economy’s structural parameters. Denote the corresponding structural matrices for the model linearized at each point in time around the steady-state corresponding to the time \( t \) structural parameters by \( A_t, B_t, C_t, D_t, \) and \( F_t \).

\[\text{A solution to the problem with time-varying structural matrices exists if agents expect the structural matrices to be fixed at some point in the future at values which are consistent with a time-invariant equilibrium (Kulish and Pagan, 2016).} \]

In this case, the solution has a time-varying VAR representation:

\[
x_t = J_t + Q_t x_{t-1} + G_t \epsilon_t,
\]

(35)

where \( J_t, Q_t, \) and \( G_t \) are conformable matrices which are functions of the evolution of beliefs about the time-varying structural matrices \( A_t, B_t, C_t, D_t, \) and \( F_t \), as detailed in the Appendix. The law of motion for the model’s state variables at a time period \( t \) therefore depends on the full anticipated path of the structural matrices.

### 4.2.2 The Zero Lower Bound

I implement the occasionally binding ZLB constraint using the solution \((35)\) and a regime-switching algorithm, where the two regimes are the ZLB regime and a Taylor-rule policy regime (see \( \text{Jones (2017); Guerrieri and Iacoviello (2015)} \)). Agents have rational expectations over which of the two regimes apply in each quarter. The algorithm iterates on the periods that the ZLB regime applies until it accords with agents’ expectations. To obtain the time-varying representation \((35)\) which is consistent with an expected duration of the ZLB at each point in time, I iterate backwards through the model’s structural equations using, as the initial point of the iteration, the system \((33)\) that arises at the expected exit from the ZLB regime. The ZLB duration that agents expect is not required to be the same duration as that implied by structural shocks. In this case, the central bank has actively extended the ZLB duration through a policy of calendar-based forward guidance (see Jones, 2017).

\[\text{One can instead linearize the model around its original steady-state, the steady-state associated with the time-varying system’s final structure, or the steady-state implied by the structure at each point in time. Given the somewhat large movements in the steady-state induced by demographic changes, I use the latter approach, linearizing each set of structural matrices around the steady-state implied by that structure.}\]
4.2.3 Quarterly Data

The solution expressed in equation (35) has a state-space representation, allowing me to adapt likelihood methods to estimate the remaining parameters of the model, such as Bayesian methods typically applied in the New Keynesian literature (see An and Schorfheide, 2007). The quarterly data series used is:

$$\text{Data} = \left\{ \log R_t, \log \Pi_t, \log \left( \frac{y_t}{y_{t-1}} \right), \log \left( \frac{c_t}{c_{t-1}} \right), \log \left( \frac{i_t}{i_{t-1}} \right), T_t \right\},$$

(36)

over the time period 1984Q1 to 2015Q1. I use, as observables, the growth rate of output per capita, of consumption per capita, of investment per capita, the GDP deflator, and the Fed Funds rate, and follow Smets and Wouters (2007) in constructing these series. The nominal interest rate is no longer an observable when the ZLB binds between 2009Q1 and 2015Q1. I implement this with a time-varying observation equation in the state-space representation of the model. The sequence of expected durations of the ZLB, $T_t$, between 2009 and 2015 are taken from the Blue Chip Financial Forecasts survey from 2009 to 2010 and the New York Federal Reserve’s Survey of Primary Dealers from 2011 to 2015.

4.2.4 Parameter Estimates

Table 1 reports moments of the prior and posterior distributions. The priors are diffuse and are common to the Bayesian estimation literature. The prior distribution over the slope of the Philips curve is wide and allows for a high degree of price flexibility. Translating the quadratic price adjustment cost into a Calvo price-reset probability, the 10th percentile of the prior distribution for the slope implies a quarterly Calvo reset probability of 13%, while the 90th percentile of the prior distribution of the slope implies a Calvo reset probability of 34%. The prior distribution for the trend rate of growth is also wide, with the 10th (90th) percentiles implying annual trend growth of about 1% (3%). I compute two independent chains of 150,000 draws. The Appendix provides full details of the estimation procedure and an analysis of the convergence of the chains.

First, I find that prices are quite inflexible, with the posterior estimate of the slope of the Philips curve $\epsilon_p$ centered tightly around 0.01. At the posterior mode, this value translates into a Calvo probability of price adjustment every quarter of about 10%, consistent with the estimates in Smets and Wouters (2007), but a little more flexible than the recent estimates of Del Negro, Giannone and Schorfheide (2015) and Aruoba et al. (2017).

Footnote: The Appendix provides more details of the data series used in estimation.
Next, I find that the modal estimate of trend growth implies an annual growth rate of around 1.35%. This is about 0.4% annual percentage points less than the estimate implied by Smets and Wouters (2007), in part because my estimation sample includes a period of lower average growth, but also because demographic trends account for some of the growth over this period, as discussed in the next section.

The remaining parameters are those governing the persistence and size of the shock processes. Instead of interpreting these values directly, I report, in Table 2, the forecast error variance decomposition of the observable variables (and wages) at both the infinite (unconditional) horizon and the one-year ahead horizon. These decompositions reveal how important each shock is in driving the observable variables around the demographic trend. At the one-year horizon, monetary policy shocks, exogenous government spending shocks, and investment shocks together make up about half of the forecast error variance for output, while at a longer horizon, about 70 percent of the forecast error variance is caused by markup shocks. About one-third of the forecast error variance of the Fed Funds rate is driven by preference shocks. As discussed in the next section, these shocks are largely accommodated by monetary policy, but can be very contractionary when monetary policy is constrained.

Monetary policy shocks do not affect output, consumption, and investment much, comprising around 5% of the forecast error variance at the one-year horizon, and less than 1% at the infinite horizon, which is about the same as the fraction reported by Smets and Wouters (2007). Monetary policy shocks instead account for about 23% of the forecast error variance of the Fed Funds rate at short horizons. Finally, consistent with the results in Justiniano, Primiceri and Tambalotti (2010), investment shocks are very important for output and investment at both the one-year and infinite horizons, explaining 27% and 50% of the forecast error variance of output and investment at the one-year horizon, and 18% and 30% of the forecast error variance of output and investment at the infinite horizon.

I gauge the importance of including demographic trends in the estimation by reporting, in the Appendix, the posterior distributions of the estimated parameters and the forecast error variance decomposition of the observables in an estimation where demographics are held constant from 1984 on. I find that the estimate of annual trend growth is higher by 0.4%, since demographic trends contribute positively to growth, as discussed in more detail in the next section. I also find a much more important role for technology shocks in explaining output and consumption when demographics are not explicitly taken into account. This suggests that demographic trends manifest themselves in ways that reflect highly persistent changes in productivity.
5 Demographic Trends and the Business Cycle

I next study, using the estimated model, the role that demographic trends have had in explaining the decline in log output relative to its pre-crisis trend. I first show that, absent all other shocks, demographic trends alone are responsible for about a third of the decline in output. On top of this, demographic changes have an additional effect by causing the ZLB to be a binding constraint between 2009 and 2015. To show this, I extract the model’s shocks that drive variables around demographic trends. Using those shocks, I show that the Fed Funds rate would not have hit the ZLB had the economy’s demographics been held constant from 1984. As a result, there are strong nonlinearities that arise because of the interaction of demographic changes and the ZLB. After removing the forward guidance response of the Fed to a binding ZLB, I find that output would have fallen by an additional 2% relative to trend, primarily between 2011 and mid-2013.

5.1 Effect of Demographics

Effect of Demographics Alone First, I discuss how demographics affects the economy’s key variables, plotted in Figure 3. Starting in 1984, I turn off all of the economy’s shocks except for the expected demographic trends. Panel A shows that, between 1990 and 2015, the Fed Funds rate declines by about 2 percentage points. Panel B plots an index of log output and illustrates how demographic changes are expected to cause a slowdown in output growth relative to log output’s 1984 to 2007 trend. In Panel C, I show that the real interest rate is expected to fall by almost 1 percentage point over the period, while in Panel D, demographic trends alone cause a decline in employment of about 2.5% between 1990 and 2015.

In Figure 4, I explore these trends in more detail by plotting first, in Panel A, the employment-population ratio in the model and the data. Demographics alone capture the dynamics of the aggregate employment-population ratio well under the calibration of the lifecycle parameters, which generated age-specific labor force participation rates that are consistent with those observed. The labor force participation rate declines in the model at a pace that is roughly as fast as that observed and is predicted to continue to fall by a further 4 percentage points from 2020 to 2040. This result is driven by the compositional changes in the workforce towards workers with lower participation rates, plotted in Panel B of Figure 2 (see also Aaronson et al, 2014).

Demographic trends have important implications for the capital-output ratio, plotted in the second panel of Figure 4. As longevity increases with declining mortality rates, aggregate
savings increases to fund longer expected retirements. As a result, the capital-output ratio increases and the real interest rate falls as the marginal product of capital declines (see also Carvalho et al., 2015). In addition, the aging of the baby boomer cohorts generates an increase, and then decrease, in the path for the real interest rate around the secular decline implied by increasing longevity, as in Gagnon et al. (2016). The oscillation is driven by changes in the relative size and composition of the workforce. The workforce is relatively young as the baby-boomers enter the labor market in the 1960s to 1980s, so that aggregate hours supplied is high relative to capital, thereby increasing the marginal return to capital. As the baby-boomer cohort ages and accumulate savings for retirement, the marginal return to capital and the real interest rate decline. This decline is then reinforced by the withdrawal of the baby-boomer cohort from the labor market, rapidly decreasing the marginal return to capital and staying low beyond 2030.

As illustrated in the second panel of Figure 3 demographic changes are a substantial drag on total output growth. I explore this in more detail in the Appendix, and show that the contribution to overall growth from demographics remains a drag throughout the forecast horizon to 2070. In total, output growth declines from peak-to-trough by about 1.5 percentage points. From 1980 to 2015, I find that demographic trends cause output growth to decline by about 1.25 percentage points, which is about the same decline reported by Gagnon et al. (2016) in their study of the macroeconomic consequences of demographics.

There are three main ways that output growth can change over time because of changing demographics. Individuals can supply more hours, affecting both output and aggregate labor. There are also changes in physical capital, as individuals save and consume out of accumulated savings in retirement. Consumption smoothing motives ensure that the level of savings changes at a different pace than the supply of labor. Third, the quality of labor can change. In particular, changes in the distribution of workers resulting from demographic changes alters the average skill-level of the workforce, which shows up in a decomposition of productivity growth as fluctuations in the average productivity of labor.

In the Appendix, I decompose the model’s predictions for output growth and labor productivity growth into their component parts and show that accelerating capital accumulation increases the growth rate of both labor productivity and total output up to 1995, after which the growth rate starts to decline. The change in labor supply has a large negative effect on productivity growth, but a positive effect on total growth, when the baby boomer cohorts enter the labor force around 1960. A key component of both labor productivity and total growth is the change in the quality of the workforce which arises as the composition of the
workforce interacts with the age-productivity profile. The decomposition implies that the contribution of the change in average labor quality to the growth rate of output and output per worker peaks around 1990, adding roughly 0.3 percentage points to total growth and productivity growth. The contribution of labor quality becomes a drag on productivity growth in 2000 as a large fraction of workers reach the peak of the age-productivity profile, exhausting the potential for further growth in average human capital across the workforce. This force is forecast to reduce productivity growth until 2030.

**Decline in Real Interest Rate and the ZLB** Next, I examine the nonlinear interactions between the decline in the real interest rate and the ZLB on the Fed Funds rate. In Figure 5 in the first panel I plot the response of the Fed Funds rate, and in the second panel I plot output, to a large negative investment shock, for two different initial states – one initialized at the steady-state associated with the 1990 demographic profile, and the other, for the steady-state associated with the 2008 demographic profile. The Fed Funds rate is higher in steady-state for the 1990 demographics, when the population is younger the supply of savings is relatively lower. The two responses illustrate how the same shock can have very different implications for the economy. Initialized at the 2008 steady-state, the shock is large enough to cause the ZLB to bind for about two years, with output falling by an additional 2 percentage points on impact. These impulse responses show that the decline in the real interest rate can be quantitatively important in the presence of the ZLB, as discussed next.

**Holding Demographic Trends Constant from 1984** Here, I examine whether demographic trends were responsible for the Fed Funds rate hitting the ZLB between 2009 and 2015. To answer this, I hold the demographic profile constant at its 1984 value, which was the first year that I use quarterly data in the Bayesian estimation, and construct a counterfactual using the estimated structural shocks. As illustrated in the first Panel of Figure 6, had the population not aged between 1984 and 2015, the Fed Funds rate would have remained above zero in the aftermath of the Great Recession, falling to, at its lowest point, 0.5% in annual terms in 2012Q4. Furthermore, consistent with the way that demographic trends have caused a secular decline in the employment-population ratio, Panel D shows that holding the demographic profile fixed at its 1984 value, the employment gap would have been closed just prior to the recession, and in 2009 would have been about 2 percentage points smaller.

Panel B shows the effect of fixed demographics on output. Interestingly, the counterfactual

---

21 This period corresponds to the time of the lowest 10-year Treasury yields, which were around 1.6%.
response shows that output growth would have been lower. The reason for this is that the composition of the workforce in 1984 is skewed towards younger workers and is therefore favorable to productivity growth between 1984 and the mid-2000s, since younger workers move up the age-productivity profile, resulting in economy-wide productivity growth.

5.2 Effect of Forward Guidance

The previous section showed how demographic trends over the past 25 years caused a decline in real and nominal interest rates, thereby making the ZLB a constraint on monetary policy. I explore how much of a constraint the ZLB was, quantitatively, by constructing a counterfactual simulation of the economy in which the Fed acts passively in response to shocks that cause the ZLB to bind. In this simulation, the expected ZLB duration adjusts in response to the shock only. By contrast, in the estimation, I fix the expected ZLB duration to those observed in survey data, which allows for the Fed to extend the ZLB duration beyond the duration implied by the shocks themselves (see Jones 2017, and Jones, Midrigan, and Philippon, 2017). The counterfactual simulation therefore provides a measure of the degree to which the ZLB was a binding constraint, absent explicit forward guidance policies.

Figure 7 illustrates the counterfactual path of the economy without forward guidance. Panel C shows that inflation would have been lower by about 1 to 2 percentage points between 2009 and 2015, and Panel D illustrates that employment would have been lower by a further 0.7 percentage points in 2012Q2. Cumulatively, over the 2009 to 2015 period, employment would have fallen by 7.6%.

Figure 8 plots the ZLB durations implied by survey data against the durations implied by the shocks themselves, with the difference between the two series approximately measuring the degree of forward guidance stimulus. I find some degree of forward guidance stimulus every quarter between 2009 and 2015, but the strongest forward guidance announcements occur between 2011Q3 and 2013Q3, when the forward guidance component of the total duration is estimated to be between 8 and 9 quarters. This period corresponds to low long-rate yields and the explicit calendar-based targets announced by the Fed. These results are also consistent with the findings in Swanson and Williams (2014), who show that between 2009 and 2011, long-term yields were relatively unconstrained, and that after 2011, long-term yields tightened significantly towards their lower bounds and is consistent with the Fed announcing

\footnote{I calculate the ZLB durations implied by the structural shocks, using a method I describe in Jones (2017). The difference between those computed endogenous durations and the durations used in the estimation is the contribution of forward guidance, or the extension of the ZLB regime that, together with the structural shocks, will generate the observed series.}
expansive unconventional monetary policies. In particular, around 2011, the Fed Reserve’s FOMC announced its “to mid-2013” guidance announcement, the first of many subsequent calendar-based extensions of the lower bound regime.\footnote{For example, in the FOMC press release, August 9 2011, the FOMC announced: “The Committee currently anticipates that economic conditions – including low rates of resource utilization and a subdued outlook for inflation over the medium run – are likely to warrant exceptionally low levels for the federal funds rate at least through mid-2013.”}

5.3 Output Since the Great Recession

In this section, I put together the results from the previous two sections to study how the model decomposes the decline in log output relative to its long-run linear trend. First, Figure\textsuperscript{9} plots the difference between output and its long-run trend relative to 2008. The data show the severity of the slowdown in output, with the gap between the data and its long-run trend widening between 2008 and 2015 to 12% by 2015. Demographics alone account for a third – 4 percentage points – of the gap in 2015. Without expansionary forward guidance policy, the gap would have been larger by, at most, 2 percentage points, primarily between 2011 and 2013.

Because of forward guidance policy and the nonlinearities associated with the ZLB, there is no clean decomposition of the remaining gap between output and its demographic path into the contribution of individual shocks.\footnote{The shocks are presented and discussed in the Appendix.} For this reason, I explore how each estimated shock affected output in two cases: first, where the ZLB duration adjusts in response to changes in shocks, and second, where the sequence of ZLB durations are held constant at their observed values.

Figure\textsuperscript{10} plots the contribution of each shock to output, computed in the following way. First, I compute counterfactual paths of output when each shock is set to zero, respectively, and the Fed Funds rate responds endogenously. I then evaluate the contribution of each shock as the difference between what is observed and the counterfactual series. The results suggest that investment and markup shocks primarily explain why output is low relative to trend and that government spending shocks kept output from falling even further between mid-2009 and 2012. Discount factor shocks have a small effect on output because they are primarily responsible for the ZLB to bind, and thereby allow monetary policy to offset other shocks. This observation is consistent with the variance decompositions of Table\textsuperscript{2} where about a third of the forecast error variance of the Fed Funds rate is explained by preference shocks, and a much smaller fraction of the forecast error variance of output, consumption,
and investment.

Finally, in Figure 11 I compute the contribution of each shock holding fixed the sequence of ZLB durations between 2009 and 2015. This exercise magnifies the role of each shock, since the Fed does not endogenously react when that shock is removed. Government spending shocks have a larger marginal impact because these shocks stimulate inflation and lower real interest rates more when nominal interest rates are fixed (see also Christiano et al. 2011). Similarly, markup shocks have a positive effect on output when the interest rate is fixed in 2012, because they raise inflation and lower real interest rates (see also, Wieland 2017). The estimated preference shocks have a substantial contractionary effect on output when the interest rate is fixed. These shocks, which capture to some extent difficult financial conditions, have a large effect on the real interest rate when they are not accommodated by monetary policy, thereby affecting consumption and output.

5.4 Robustness and Alternative Calibrations

I conduct a number of experiments to verify that the baseline predictions of the model’s demographic trends are robust to alternative calibrations. The full results of each experiment are presented in the Appendix but discussed briefly here.

Borrowing Constraints I first check that the model’s predictions hold when individuals face a constraint restricting their borrowing early in life. With borrowing constraints, there are more savings, pushing up the capital-output ratio. As a consequence, the real interest rate is lower than in the baseline model. The magnitude of the fluctuations of the real interest rate, the participation rate and output growth are very similar to the baseline model.

Time-varying Productivity Profile The second robustness check is to adopt time-varying productivity profiles to account for a possible flattening of productivity profiles over time. Such a flattening can affect the accumulation of human capital and can impact aggregate productivity measures in two ways: first, by a growth effect, by lowering the potential for new workers to accumulate human capital, and second, by a level effect, by affecting the productivity level that individuals enter the workforce on. I calibrate the age-productivity profiles by recomputing for each cross-sectional sample, the profile and then interpolating between those points in time. The overall pattern of aggregate labor productivity is much the same as the baseline model, although the magnitude of the amplitude of the change in labor productivity growth is smaller, with demographics contributing the most to labor
productivity growth in 1980 rather than in 1990 (as in the baseline results).

**Female Labor Force Participation and Multiple Skill Types** From 1985 on, the baseline predictions for the participation rate, aggregate labor productivity growth and the real interest rate are largely unaffected when the age-productivity and labor disutility profiles are calibrated to match female age-earnings profiles and female labor force participation rates from the 1940s to 1990s, after which female labor force participation is roughly constant. As a final point of comparison, I verify that the directions of the aggregate predictions are robust to a calibration where an additional source of heterogeneity is modeled—where there are two types of workers, low or high skilled, where low skill workers are those with less than college education. These robustness exercises emphasize how the important demographic dynamics are captured primarily through changes in the size of the population.

6 Conclusions

This paper studies why the level of US output per capita was 12% below its pre-crisis trend in 2015. I use a New Keynesian model with demographic trends and the ZLB to show that declining mortality rates and changes to the age population composition can generate trends that match the low frequency movement of output growth, productivity, the real interest rate, and the employment-population ratio.

I estimate the transitory shocks of the model using Bayesian likelihood methods and accounting for ZLB, forward guidance, and the demographic transition. With the estimated model, I find that the ZLB would not have been a binding constraint between 2009 and 2015 if the population not aged. I find that the aging of the population alone is responsible for a sizeable proportion of the decline in output per capita relative to its pre-crisis trend by 2015—about one-third. Furthermore, using the estimated shocks, my results suggest that absent any forward guidance policy used by the Fed, the ZLB would have caused output to fall by an additional 2 percentage points between 2011 and 2013.

The shocks themselves can have highly nonlinear effects on output and consumption, depending on whether the ZLB adjusts in response to those shocks. I assess the contribution of each of the estimated shocks to the decline in output since the Great Recession, and find important roles for investment and markup shocks in causing output to fall, counteracted by positive government spending shocks between 2009 and 2011. Preference shocks alone have little effect unless the expected ZLB duration is held fixed, in which case they cannot be offset by monetary policy and are thus highly contractionary.
The results clearly show the importance of demographic trends in explaining macroeconomic trends over time. Further research could focus on how demographic trends interact with the housing market in a macroeconomic framework. The methodology combining long-run trends and the ZLB in an estimated framework could be applied more generally to study empirically relevant and nonlinear interactions between trends, shocks, and the ZLB.
References


Antolin-Diaz, Juan, Thomas Drechsel, and Ivan Petrella, “Following the Trend: Tracking GDP when Long-Run Growth is Uncertain,” 2014.


Carvalho, Carlos, Andrea Ferrero, and Fernanda Nechio, “Demographics and Real Interest Rates: Inspecting the Mechanism,” 2015.

Casanova, Maria, “Revisiting the Hump-Shaped Wage Profile,” 2013.


Jaimovich, Nir and Henry Siu, “The Trend is the Cycle: Job Polarization and Jobless Recoveries,” 2012.


Wu, Jing Cynthia and Fan Dora Xia, “Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound,” 2016. manuscript.

### Table 1: Estimated Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Prior Dist</th>
<th>Median</th>
<th>10%</th>
<th>90%</th>
<th>Posterior Mode</th>
<th>Median</th>
<th>10%</th>
<th>90%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\epsilon_p$</td>
<td>U</td>
<td>0.1</td>
<td>0.0</td>
<td>0.2</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>$400 \times (z - 1)$</td>
<td>N</td>
<td>2.0</td>
<td>1.1</td>
<td>3.0</td>
<td>1.34</td>
<td>1.35</td>
<td>1.24</td>
<td>1.46</td>
</tr>
<tr>
<td>$\rho_\chi$</td>
<td>B</td>
<td>0.5</td>
<td>0.3</td>
<td>0.7</td>
<td>0.94</td>
<td>0.94</td>
<td>0.93</td>
<td>0.95</td>
</tr>
<tr>
<td>$\rho_\mu$</td>
<td>B</td>
<td>0.5</td>
<td>0.3</td>
<td>0.7</td>
<td>0.96</td>
<td>0.96</td>
<td>0.95</td>
<td>0.97</td>
</tr>
<tr>
<td>$\rho_\theta$</td>
<td>B</td>
<td>0.5</td>
<td>0.3</td>
<td>0.7</td>
<td>0.95</td>
<td>0.95</td>
<td>0.94</td>
<td>0.97</td>
</tr>
<tr>
<td>$\rho_g$</td>
<td>B</td>
<td>0.5</td>
<td>0.3</td>
<td>0.7</td>
<td>0.95</td>
<td>0.95</td>
<td>0.94</td>
<td>0.97</td>
</tr>
<tr>
<td>$\rho_\kappa$</td>
<td>B</td>
<td>0.5</td>
<td>0.3</td>
<td>0.7</td>
<td>0.94</td>
<td>0.94</td>
<td>0.93</td>
<td>0.95</td>
</tr>
<tr>
<td>$100 \times \sigma_\chi$</td>
<td>IG</td>
<td>1.2</td>
<td>0.5</td>
<td>3.7</td>
<td>2.15</td>
<td>2.18</td>
<td>1.95</td>
<td>2.47</td>
</tr>
<tr>
<td>$100 \times \sigma_\mu$</td>
<td>IG</td>
<td>1.2</td>
<td>0.5</td>
<td>3.7</td>
<td>0.48</td>
<td>0.52</td>
<td>0.33</td>
<td>0.73</td>
</tr>
<tr>
<td>$100 \times \sigma_\theta$</td>
<td>IG</td>
<td>1.2</td>
<td>0.5</td>
<td>3.7</td>
<td>3.51</td>
<td>3.56</td>
<td>3.20</td>
<td>4.03</td>
</tr>
<tr>
<td>$100 \times \sigma_i$</td>
<td>IG</td>
<td>1.2</td>
<td>0.5</td>
<td>3.7</td>
<td>0.16</td>
<td>0.16</td>
<td>0.15</td>
<td>0.18</td>
</tr>
<tr>
<td>$100 \times \sigma_g$</td>
<td>IG</td>
<td>1.2</td>
<td>0.5</td>
<td>3.7</td>
<td>1.05</td>
<td>1.06</td>
<td>0.98</td>
<td>1.15</td>
</tr>
<tr>
<td>$100 \times \sigma_\kappa$</td>
<td>IG</td>
<td>1.2</td>
<td>0.5</td>
<td>3.7</td>
<td>1.00</td>
<td>1.04</td>
<td>0.86</td>
<td>1.25</td>
</tr>
</tbody>
</table>

### Table 2: Variance Decomposition Due to Shocks, %

#### A. Conditional, 4 Quarter Ahead

<table>
<thead>
<tr>
<th>Variable</th>
<th>Preference</th>
<th>Technology</th>
<th>Markup</th>
<th>Policy</th>
<th>Government</th>
<th>Investment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fed Funds Rate</td>
<td>28.5</td>
<td>7.8</td>
<td>5.8</td>
<td>22.9</td>
<td>8.3</td>
<td>26.8</td>
</tr>
<tr>
<td>Inflation</td>
<td>21.9</td>
<td>2.4</td>
<td>42.1</td>
<td>8.9</td>
<td>6.4</td>
<td>18.4</td>
</tr>
<tr>
<td>Wages</td>
<td>6.8</td>
<td>2.6</td>
<td>74.8</td>
<td>10.5</td>
<td>0.6</td>
<td>4.8</td>
</tr>
<tr>
<td>Output</td>
<td>0.8</td>
<td>0.3</td>
<td>47.7</td>
<td>4.3</td>
<td>19.6</td>
<td>27.4</td>
</tr>
<tr>
<td>Consumption</td>
<td>28.0</td>
<td>0.7</td>
<td>26.7</td>
<td>6.8</td>
<td>26.7</td>
<td>11.2</td>
</tr>
<tr>
<td>Investment</td>
<td>7.2</td>
<td>0.1</td>
<td>38.8</td>
<td>2.3</td>
<td>0.8</td>
<td>50.8</td>
</tr>
</tbody>
</table>

#### B. Unconditional

<table>
<thead>
<tr>
<th>Variable</th>
<th>Preference</th>
<th>Technology</th>
<th>Markup</th>
<th>Policy</th>
<th>Government</th>
<th>Investment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fed Funds Rate</td>
<td>34.6</td>
<td>1.8</td>
<td>28.7</td>
<td>5.6</td>
<td>9.7</td>
<td>19.6</td>
</tr>
<tr>
<td>Inflation</td>
<td>20.5</td>
<td>1.1</td>
<td>55.7</td>
<td>4.2</td>
<td>5.8</td>
<td>12.7</td>
</tr>
<tr>
<td>Wages</td>
<td>1.7</td>
<td>0.4</td>
<td>93.0</td>
<td>2.0</td>
<td>0.2</td>
<td>2.7</td>
</tr>
<tr>
<td>Output</td>
<td>4.2</td>
<td>0.0</td>
<td>70.9</td>
<td>0.7</td>
<td>6.2</td>
<td>18.0</td>
</tr>
<tr>
<td>Consumption</td>
<td>7.0</td>
<td>0.1</td>
<td>70.0</td>
<td>1.0</td>
<td>13.6</td>
<td>8.4</td>
</tr>
<tr>
<td>Investment</td>
<td>8.6</td>
<td>0.0</td>
<td>59.1</td>
<td>0.5</td>
<td>1.2</td>
<td>30.5</td>
</tr>
</tbody>
</table>
Figure 1: Macroeconomic Trends

A. Fed Funds Rate

B. Output (2000 = 1)

C. Real Interest Rate

D. Employment (1990 = 0)
Figure 2: Individual Profiles

A. Productivity (Age 16 = 1)

B. Participation Rate in 2000

C. Assets in 2013 (Age 60 = 1)

D. Median Age, 16+

Notes: The data for the asset profile is taken from the Survey of Consumer Finances for 2013, HP-filtered, and normalized to the values of assets when 60. The productivity profile is computed from pooled Census and American Community Survey datasets. All details are in the Appendix.
Figure 3: Path of Variables, Demographics Only

A. Fed Funds Rate

B. Output (2000 = 1)

C. Real Interest Rate

D. Employment (1990 = 0)
Notes: The data for the capital-output ratio is extracted from the BLS’s Multifactor Productivity program. All details are in the Appendix.
Figure 5: Impulse Response to Large Investment Shock

A. Fed Funds Rate

B. Output

[Graph showing the impulse response of Fed Funds Rate and Output to large investment shock for 1990 and 2008 steady-states]
Figure 6: Path of Variables, Demographics Fixed from 1984

A. Fed Funds Rate

B. Output (2000 = 1)

C. Inflation

D. Employment (1990 = 0)
Figure 7: Path of Variables, No Forward Guidance

A. Fed Funds Rate

B. Output (2000 = 1)

C. Inflation

D. Employment (1990 = 0)
Figure 8: Decomposition of ZLB Durations, Quarters
Figure 10: Output Relative to Trend, Contribution of Shocks
Figure 11: Output Relative to Trend, Contribution of Shocks Fixing ZLB Durations