Trends and cycles without balanced growth.

Tom D. Holden[1]*, Deutsche Bundesbank and University of Surrey*

Abstract: This paper presents an approach for solving, simulating and estimating models without balanced growth. Key to the simulation method is the addition of an extra unit root variable which acts as a growth switch. We discuss ways of further improving the accuracy of this method, including showing how models may be approximated around a non-steady-state point. We also provide a second order extended Kalman filter for DSGE models solved at second order which tractably propagates a Gaussian approximation to the state distribution without approximating the mean or covariance. In an application, we construct a real business cycle with CES production and preferences, and estimate it on a long span of annual US data. We show that the model is able to explain the decline in the labour share over time.

Keywords: *non balanced growth, CES production, labour share, secular stagnation, non-linear estimation*

JEL Classification: *C32, C63, E17, E2, E3, O4*

[1] Contact address: Wilhelm-Epstein-Straße 14, 60431 Frankfurt am Main, Germany
Contact email: thomas.holden@gmail.com Web-site: http://www.tholden.org/
Financial support provided to the author by the EC is greatly appreciated.
The research leading to these results has received funding from The European Union's Horizon 2020 Framework Programme for Research and Innovation, under grant agreement "Modelling of the Research and Innovation Policies" (MONROE, grant no. 727114).

1. Introduction

The traditional approach to solving and estimating DSGE models requires them to be stationary about a stochastic trend. This limits the model's ability to explain many recent trends, such as the decline of the labour share or the decline of the real rate of interest. In this paper, we present a new approach to the solution, simulation and estimation of models without balanced growth paths. We apply our method to estimate an annual real business cycle model with a flexible productivity structure combined with CES production and preferences.

In a now seminal paper (Kaldor 1961), Kaldor documented the constancy of growth rates, investment returns and labour shares. 60 years on, the evidence for these "facts" appears much weaker. Drawing on the work of Laubach & Williams (2003), Summers (2014) documents the "secular stagnation" hypothesis of declining growth rates and interest rates. Declines in real interest rates are further documented by King & Low (2014), Del Negro et al. (2017; 2018) and Borio et al. (2018), amongst others. Some of the many works documenting the decline in the labour share include Elsby, Hobijn, & Sahin (2013) and Karabarbounis & Neiman (2014a; 2014b).

In order to solve, simulate and estimate a DSGE model by traditional means, one needs to be able to find stochastic trends for each of the model's variables, such that all variables are stationary relative to their stochastic trends. In other words, the model must possess a (possibly stochastic) balanced growth path. In practice, this means that whenever one has a sum of variables exhibiting non-linear growth (e.g. $Y_t = C_t + I_t + G_t + X_t$, where each has an exponential trend), then all of the variables must have the same stochastic trend. This is problematic as many of the leading explanations for the violations of the Kaldor facts involve either CES production or non-homotheticities in preferences.

For CES production functions, the problem arises as the CES production function contains a term of the form $\alpha(A_{K,t}K_t)^{\frac{\sigma-1}{\sigma}} + (1-\alpha)(A_{H,t}H_t)^{\frac{\sigma-1}{\sigma}}$. Thus, traditional solution methods require that $A_{K,t}K_t$ and $A_{H,t}H_t$ have the same stochastic trend. However, since the aggregate market clearing condition in standard models implies that K_t has the same stochastic trend as output, which is growing at least as fast as $A_{H,t}H_t$, the only way there can be a stochastic balanced growth path is if $A_{K,t}$ is stationary. Non-homotheticities lead to terms like $C_t - \bar{C}$ in preferences, which can only be consistent with balanced growth if C_t is not growing.

This paper presents an alternative method for solving, simulating and estimating models without balanced growth. There are two main ideas.

Firstly, consider a model in which some of the driving shock processes are non-stationary but without trend, and assume there are no other trends in the model. Standard perturbation approximations are not valid as there is no steady-state about which to approximate. However, accurate solutions can be produced by approximating each period around the steady-state conditional on the current value of the non-stationary shock processes.

Secondly, observe that a model with one or more trends may be thought of as a special case of a model without trends but with an augmented state containing further random walks. E.g. the model $y_t = g + y_{t-1} + \sigma\varepsilon_t$, is a special case of the model with equations:

$$y_t = gv_t + y_{t-1} + \sigma\varepsilon_t,$$
$$v_t = v_{t-1} + \omega\eta_t.$$

In particular, the original model is the limit of the second model as $\omega \to 0$ with $v_0 = 1$. Since when $v_0 = v$, y_t has no steady state, we cannot approximate around $v_0 = 1$. However, an approximation around $v_0 = 0$, evaluated at $v_0 = 1$ may capture the important dynamics, particularly if a high order approximation is used. In fact, we will show that we can actually do somewhat better than this by taking an approximation around a point other than the steady state.

We also contribute to the broader literature on non-linear estimation by providing a second order extended Kalman filter which exactly propagates the mean and covariance of the state. This is an advance over rival approaches based on approximate numerical integration, such as the unscented (Julier & Uhlmann 1997) or cubature Kalman filter (Arasaratnam & Haykin 2009) which hav been used in macro by Binning and Maih (2015) amongst others.

We motivate our technical contributions with a flexible real business cycle model without balanced growth. This model also serves to guide us in our technical contributions, to ensure that they are truly appropriate for the scale of models in which practitioners are interested. The model contains four CES production functions with unrestricted elasticity of substitution, and unrestricted productivity terms. These functions control the aggregation in preferences of present felicity & future utility, the aggregation in preferences of public & private consumption, the aggregation in preferences of consumption & leisure, and the aggregation in production of capital & labour. Additionally, all relative prices are allowed to follow non-stationary trending processes. This captures the technology available for transforming final goods into either consumption goods, investment goods, government goods or export goods, as well as the technology for transforming investment goods into installed capital goods. As such, the model can capture the observed changes in GDP growth rates, investment returns and labour shares via a number of channels.

The next describes this model in more detail. We then discuss solution and estimation methods. We end with a presentation of the preliminary results from estimating the model.

2. The model

The model is an annual real business cycle model. We work with an annual rather than quarterly model both due to greater availability of annual data, and because we are abstracting from many of the frictions necessary to capture quarterly data.

2.1. Production

We begin by describing the production side of the economy. Since the model contains normalized CES production functions in multiple places, this will aid exposition by providing us an opportunity to introduce the normalized CES production function in its most natural setting. The normalized CES production function (Klump, McAdam & Willman 2012) takes the following form:

$$Y_t = Y_0 \left[\alpha_Y \left(\frac{A_{K,t}\tilde{K}_t}{A_{K,0}\tilde{K}_0} \right)^{\frac{\sigma_Y - 1}{\sigma_Y}} + (1-\alpha_Y) \left(\frac{A_{H,t}H_t}{A_{H,0}H_0} \right)^{\frac{\sigma_Y - 1}{\sigma_Y}} \right]^{\frac{\sigma_Y}{\sigma_Y - 1}}.$$

Here $\tilde{K}_t$ is the utilization adjusted capital stock used in production in period t, and H_t is the number of hours of labour used in production that period. $A_{K,t}$ and $A_{H,t}$ give the productivity of capital and labour respectively. Crucially, all of these terms enter divided by a "period-0" value. This is normalization. Here, "period-0" should not be interpreted too literally as a particular point in time. Rather, it reflects the point of normalization, which we will take as a prolonged time interval, following León-Ledesma, McAdam & Willman (2010). Thus, Y_0, $\tilde{K}_0$ and H_0 will be fixed as geometric means of these quantities from the data, over the normalization period. Given this normalization, α_Y then has a structural interpretation as the capital share over the normalization period.

Perfectly competitive production firms choose capital and hours to maximise their profits:

$$P_{Y,t}Y_t - P_{K,t}\mathcal{R}_t\tilde{K}_t - P_{C,t}W_tH_t,$$

where $P_{Y,t}$ is the price of final goods, $P_{K,t}$ is the price of installed capital, $\mathcal{R}_t$ is the real rental rate of capital in units of the installed capital goods, $P_{C,t}$ is the price of consumption goods, W_t is the real wage in units of consumption goods. Throughout, the use of different price units for different goods will ensure conditions are as simple as possible. Profit maximisation leads to the following first order conditions:

$$\alpha_Y P_{Y,t} Y_0 \left(\frac{Y_t}{Y_0} \right)^{\frac{1}{\sigma_Y}} \left(\frac{A_{K,t}\tilde{K}_t}{A_{K,0}\tilde{K}_0} \right)^{\frac{\sigma_Y - 1}{\sigma_Y}} = P_{K,t}\mathcal{R}_t\tilde{K}_t, \tag{1}$$

$$(1-\alpha_Y) P_{Y,t} Y_0 \left(\frac{Y_t}{Y_0} \right)^{\frac{1}{\sigma_Y}} \left(\frac{A_{H,t}H_t}{A_{H,0}H_0} \right)^{\frac{\sigma_Y - 1}{\sigma_Y}} = P_{C,t}W_tH_t.$$

Of course, substituting these back into the production function gives us that:

$$P_{Y,t}Y_t = P_{K,t}\mathcal{R}_t\tilde{K}_t + P_{C,t}W_tH_t,$$

so profits are zero.

2.2. Households

We now turn to the household side. We assume that the representative household has a population of measure N_t. Throughout, for any variable $\mathcal{U}$, $\Gamma_{\mathcal{U},t}$ will give the gross growth rate of $\mathcal{U}$ at t, i.e. $\frac{\mathcal{U}_t}{\mathcal{U}_{t-1}}$. Thus, the growth rate of population is $\Gamma_{N,t}$.

The representative household's value function is given by:

$$V_t = V_0\left[\alpha_V\left(\frac{A_{X,t}X_t}{A_{X,0}X_0}\right)^{\frac{\sigma_V-1}{\sigma_V}} + (1-\alpha_V)\left(\frac{A_{F,t}F_t}{A_{F,0}F_0}\right)^{\frac{\sigma_V-1}{\sigma_V}}\right]^{\frac{\sigma_V}{\sigma_V-1}},$$

where:

$$F_t := \left(\mathbb{E}_t\Gamma_{N,t+1}^{\eta}V_{t+1}^{\frac{\sigma_V-1}{\sigma_V}}\right)^{\frac{\sigma_V}{\sigma_V-1}},$$

and where X_t is a per capita aggregate of public and private consumption, plus leisure, to be defined. We think of this aggregator between X_t and F_t as representing the technology available to the household to "produce" utility from present and future consumption. As such, having productivity terms $A_{X,t}$ and $A_{F,t}$ is perfectly natural. This gives a way of thinking about the "preference-shocks" common in New Keynesian models such as Smets & Wouters (2007). Non-stationary and/or trending movements in $A_{X,t}$ and $A_{F,t}$ will enable us to generate permanent movements in real interest rates. Note that for simplicity, we are not allowing for the full generality of Epstein-Zin preferences (Epstein & Zin 1989) as the elasticity parameter entering F_t is identical to the one entering V_t. This will mean that the coefficient of relative risk aversion will be equal to the inverse of the elasticity of intertemporal substitution.

Population growth enters into F_t raised to some power $\eta \in [0,1]$. If $\eta = 1$, then current individuals value the consumption stream of the "new born" just as much as they value their own consumption stream. The $\eta < 1$ case covers lower concern for the future welfare of children or immigrants.

We may derive a more standard additively separable representation of preferences if we define:

$$\widetilde{V}_t := \frac{\sigma_V}{\sigma_V-1}\frac{1}{\alpha_V}\left(\frac{A_{X,0}}{A_{X,t}}\frac{V_t}{V_0}X_0\right)^{\frac{\sigma_V-1}{\sigma_V}},$$

and:

$$\beta := (1-\alpha_V)\left(\frac{V_0}{A_{F,0}F_0}\right)^{\frac{\sigma_V-1}{\sigma_V}}.$$

Then:

$$
\begin{aligned}
\widetilde{V}_t &= \frac{\sigma_V}{\sigma_V - 1} X_t^{\frac{\sigma_V-1}{\sigma_V}} + \frac{\sigma_V}{\sigma_V - 1} \frac{1-\alpha_V}{\alpha_V} \left(\frac{A_{X,0}}{A_{X,t}} X_0 \right)^{\frac{\sigma_V-1}{\sigma_V}} \left(\frac{A_{F,t} F_t}{A_{F,0} F_0} \right)^{\frac{\sigma_V-1}{\sigma_V}} \\
&= \frac{\sigma_V}{\sigma_V - 1} X_t^{\frac{\sigma_V-1}{\sigma_V}} + \beta \frac{\sigma_V}{\sigma_V - 1} \frac{1}{\alpha_V} \left(\frac{A_{X,0}}{A_{X,t}} X_0 \right)^{\frac{\sigma_V-1}{\sigma_V}} \left(\frac{A_{F,t} F_t}{V_0} \right)^{\frac{\sigma_V-1}{\sigma_V}} \\
&= \frac{\sigma_V}{\sigma_V - 1} X_t^{\frac{\sigma_V-1}{\sigma_V}} + \beta \frac{\sigma_V}{\sigma_V - 1} \frac{1}{\alpha_V} \left(\frac{A_{X,0}}{A_{X,t}} \frac{A_{F,t}}{V_0} X_0 \right)^{\frac{\sigma_V-1}{\sigma_V}} \mathbb{E}_t \Gamma_{N,t+1}^{\eta} V_{t+1}^{\frac{\sigma_V-1}{\sigma_V}} \\
&= \frac{\sigma_V}{\sigma_V - 1} X_t^{\frac{\sigma_V-1}{\sigma_V}} + \beta \mathbb{E}_t \Gamma_{N,t+1}^{\eta} \left(\frac{A_{X,t+1}}{A_{X,t}} A_{F,t} \right)^{\frac{\sigma_V-1}{\sigma_V}} \widetilde{V}_{t+1} \\
&= \mathbb{E}_t \sum_{s=0}^{\infty} \beta^s \left(\frac{N_{t+s}}{N_t} \right)^{\eta} \left(\frac{A_{X,t+s}}{A_{X,t}} \right)^{\frac{\sigma_V-1}{\sigma_V}} \left[\prod_{k=0}^{s-1} A_{V,t+k}^{\frac{\sigma_V-1}{\sigma_V}} \right] \frac{\sigma_V}{\sigma_V - 1} X_{t+s}^{\frac{\sigma_V-1}{\sigma_V}}.
\end{aligned}
$$

The representative household maximises their value function subject to several constraints. First, there is the "production" function for X_t, which is given by:

$$
X_t = X_0 \left[\alpha_X \left(\frac{A_{Z,t} Z_t}{A_{Z,0} Z_0} \right)^{\frac{\sigma_X-1}{\sigma_X}} + (1-\alpha_X) \left(\frac{A_{L,t} \left(\bar{h} - \frac{H_t}{N_t} \right)}{A_{L,0} \left(\bar{h} - \frac{H_0}{N_0} \right)} \right)^{\frac{\sigma_X-1}{\sigma_X}} \right]^{\frac{\sigma_X}{\sigma_X-1}}.
$$

Here, $\bar{h}$ is the maximum feasible number of hours per capita, which we treat as a free parameter to estimate, and H_t is the total number of hours supplied. Z_t is the output of the following "production" function:

$$
Z_t = Z_0 \left[\alpha_Z \left(\frac{A_{C,t} C_t N_0}{A_{C,0} N_t C_0} \right)^{\frac{\sigma_Z-1}{\sigma_Z}} + (1-\alpha_Z) \left(\frac{A_{G,t} G_t N_0}{A_{G,0} N_t G_0} \right)^{\frac{\sigma_Z-1}{\sigma_Z}} \right]^{\frac{\sigma_Z}{\sigma_Z-1}},
$$

where, C_t is total consumption of private goods and G_t is total consumption of public goods. Next, there is the law of motion for the (real) capital stock, K_t, which is owned by the household:

$$
K_t = (1-\delta_t) K_{t-1} + \left(1 - \frac{\delta_t}{2} \right) \Omega_t I_t,
$$

where:

$$
\delta_t = \frac{1}{1 + \tilde{\delta}_t^{-1}}.
$$

We allow the process $\tilde{\delta}_t$ driving the depreciation rate δ_t to be stochastic (even non-stationary and trending) to capture fluctuations in the consumption of fixed capital share. In line with the accounting procedures of the BEA's Fixed Assets tables, we assume that half of the stock of new (real) investment, I_t, depreciates within the period (a year). Investment enters multiplied by a relative productivity term Ω_t. This is the number of units of investment goods necessary to increase the capital stock by one unit. Equivalently, $\Omega_t = \frac{P_{I,t}}{P_{K,t}}$, i.e. the price of investment goods divided by the price of capital goods.

2.3. Digression on the price of capital

Ω_t enters the model in the same way as what Justiniano, Primiceri & Tambalotti (2011) call the marginal efficiency of investment. However, whereas Justiniano, Primiceri & Tambalotti (2011) treat it as an unobserved shock, we will use capital price data from the BEA's Fixed Assets tables to pin it down. Ω_t also enters the model in the same way as what Greenwood, Hercowitz & Krusell (1997) call "investment specific technological change", which they identify with the price of consumption in units of investment goods. As Justiniano, Primiceri & Tambalotti (2011) point out, this is a conceptually different quantity, and the two will be distinct in our model. Interestingly, as shown in Figure 1, while the price of investment goods in units of consumption goods is falling, the price of capital goods in units of investment goods is rising so fast that the price of capital goods in units of consumption goods is also rising. Thus, contrary to Greenwood, Hercowitz & Krusell (1997), it is unlikely that changes in the price of installed capital in units of consumption goods are a driver of growth.

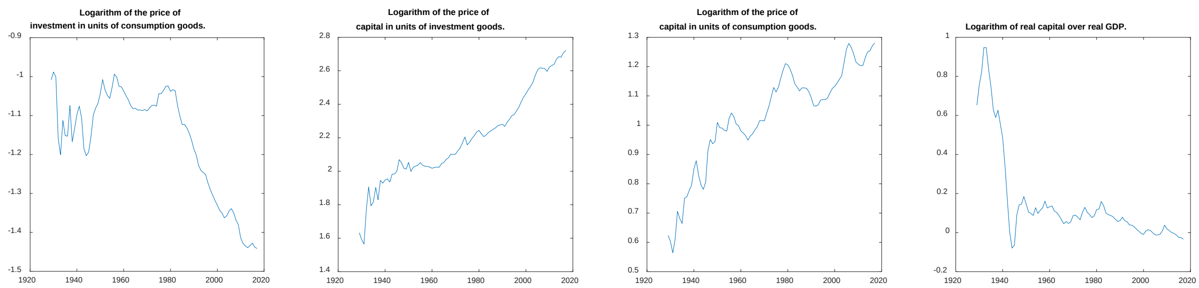

Figure 1: Relative price of capital, investment and consumption (panels one to three). Capital to GDP ratio (panel four).[2]

Indeed, with installed capital becoming more expensive over time, there is an immediate explanation for the decline in the labour share. If, capital and labour are gross complements, in line with the empirical evidence (Antràs 2004; Klump, McAdam & Willman 2007; Chirinko 2008; León-Ledesma, McAdam & Willman 2013), and capital is becoming more expensive over time, then we should expect to see a decline in the labour share. To see this, note that the capital first order condition (1) implies that the capital share is given by:

$$\alpha_Y \left(\frac{A_{K,t}}{A_{K,0}} \frac{Y_0}{\tilde{K}_0} \frac{\tilde{K}_t}{Y_t} \right)^{\frac{\sigma_Y - 1}{\sigma_Y}}.$$

[2] All data is from the BEA. See the appendix for further details.

If capital and labour are gross complements, then $\frac{\sigma_Y-1}{\sigma_Y}<0$, thus, a decline in $\frac{\tilde{K}_t}{Y_t}$ driven by capital's high price will lead to an increase in the capital share, i.e. a fall in the labour share. The decline in $\frac{\tilde{K}_t}{Y_t}$ is clear from the final panel of Figure 1 (though this is without utilization adjustment). This explanation appears to have been missed by the prior literature since people wrongly assumed that since the price of investment in units of consumption goods was falling, the price of capital in units of consumption goods must also have been falling.

2.4. Households, continued

Returning to the model, there are two final constraints faced by the representative household. The first gives the definition of utilisation adjusted capital:

$$\tilde{K}_t := U_t(K_{t-1}+\kappa\Omega_t I_t),$$

where U_t is the utilisation rate and $\kappa\in\left[0,\frac{1}{2}\right]$ is a parameter which controls the strength of time to build frictions. If $\kappa=\frac{1}{2}$, then investment is fully available for production in the year it is installed. ($\kappa>\frac{1}{2}$ would make no sense as roughly half of a year's investment does not happen until the second half of the year, and so cannot contribute fully to the capital stock.) If $\kappa=0$ then we have the one period time to build frictions traditionally used in the quarterly RBC literature.

The final constraint to consider is the household's budget constraint:

$$\begin{aligned}\frac{1}{1-\tau_{C,t}}P_{C,t}C_t+\frac{1}{1-\tau_{I,t}}P_{I,t}I_t+P_{E,t}E_t+\tau_{Y,t}P_{Y,t}Y_t+\frac{\theta_{K,t}}{2}\left(\log\frac{K_t}{K_{t-1}}\right)^2P_{I,t}K_{t-1}\\+\frac{\theta_{I,t}}{2}\left(\log\frac{I_t}{I_{t-1}}\right)^2P_{I,t}K_{t-1}-\xi_t\log(1-U_t^\nu)\,P_{I,t}K_{t-1}+B_t\\=R_{t-1}B_{t-1}+(1-\tau_{K,t})P_{I,t}\mathcal{R}_t\tilde{K}_t+(1-\tau_{H,t})P_{C,t}W_tH_t.\end{aligned}$$

Here, $\tau_{\mho,t}$ is a tax (subsidy if negative) on the good $\mho_t$, and $P_{\mho,t}$ is the price of the good $\mho_t$. All $\tau_{\mho,t}$ variables are constrained to lie in $[-1,1]$ and the way they enter is transformed to make this restriction reasonable. For example, as $\tau_{C,t}$, the consumption tax goes to one, the cost of a unit of consumption goes to infinity. As $\tau_{C,t}$ goes to minus one, the cost of a unit of consumption becomes half what it would have been otherwise. Note that $\tau_{Y,t}$ is effectively a lump-sum tax/subsidy.

Other terms in the budget constraint include real net exports, E_t, which for simplicity we model as a gift to the rest of the world, $\theta_{K,t}$, which controls the strength of capital growth adjustment costs, $\theta_{I,t}$, which controls the strength of investment growth adjustment costs, ξ_t, which controls the cost of utilisation, ν, which determines the rate of increase in this cost as utilisation goes to one, B_t, holdings of nominal government debt and R_t, the nominal interest rate.

To solve for household optimal behaviour, we introduce the following Lagrange multipliers. $\lambda_{X,t}$ is the multiplier on the definition of X_t. $\lambda_{Z,t}$ is the multiplier on the definition of Z_t. $\lambda_{K,t}$ is proportional to the Lagrange multiplier on the law of motion of the capital stock (the actual multiplier is $\lambda_{K,t}\frac{\lambda_{B,t}P_{K,t}}{P_{C,t}N_t}$). $\lambda_{B,t}$ is proportional to the Lagrange multiplier on the budget constraint (the actual multiplier is $\lambda_{B,t}\frac{1}{P_{C,t}N_t}$). It is also helpful to define the stochastic discount factor Ξ_{t+1}:

$$\Xi_{t+1} := \beta A_{V,t}^{\frac{\sigma_V-1}{\sigma_V}}\left(\frac{N_{t+1}}{N_t}\right)^{\eta-1}\left(\frac{A_{X,t+1}}{A_{X,t}}\right)^{\frac{\sigma_V-1}{\sigma_V}}\frac{\lambda_{B,t+1}P_{C,t}}{\lambda_{B,t}P_{C,t+1}}$$

In terms of these multipliers and Ξ_{t+1}, we have the following first order conditions:

$$X_t^{-\frac{1}{\sigma_V}} = \lambda_{X,t},$$

$$\alpha_X\lambda_{X,t}X_0\left(\frac{X_t}{X_0}\right)^{\frac{1}{\sigma_X}}\left(\frac{A_{Z,t}Z_t}{A_{Z,0}Z_0}\right)^{\frac{\sigma_X-1}{\sigma_X}}\frac{1}{Z_t} = \lambda_{Z,t},$$

$$\frac{1-\alpha_X}{1-\tau_{H,t}}\frac{\lambda_{X,t}}{\lambda_{B,t}}X_0\left(\frac{X_t}{X_0}\right)^{\frac{1}{\sigma_X}}\left(\frac{A_{L,t}\left(\bar{h}-\frac{H_t}{N_t}\right)}{A_{L,0}\left(\bar{h}-\frac{H_0}{N_0}\right)}\right)^{\frac{\sigma_X-1}{\sigma_X}}\frac{1}{\bar{h}-\frac{H_t}{N_t}} = W_t,$$

$$\alpha_Z(1-\tau_{C,t})\lambda_{Z,t}Z_0\left(\frac{Z_t}{Z_0}\right)^{\frac{1}{\sigma_Z}}\left(\frac{A_{C,t}C_tN_0}{A_{C,0}N_tC_0}\right)^{\frac{\sigma_Z-1}{\sigma_Z}}\frac{N_t}{C_t} = \lambda_{B,t},$$

$$\begin{aligned}\lambda_{K,t} &+ \theta_{K,t}\left(\log\frac{K_t}{K_{t-1}}\right)\frac{K_{t-1}}{K_t}\\ &+ \mathbb{E}_t\Xi_{t+1}\frac{P_{K,t+1}}{P_{K,t}}\left[\frac{\theta_{K,t+1}}{2}\left(\log\frac{K_{t+1}}{K_t}\right)^2 + \frac{\theta_{I,t+1}}{2}\left(\log\frac{I_{t+1}}{I_t}\right)^2\right.\\ &\left.- \xi_{t+1}\log(1-U_{t+1}^{\nu})\right]\\ &= \mathbb{E}_t\Xi_{t+1}\frac{P_{K,t+1}}{P_{K,t}}\Big[\lambda_{K,t+1}(1-\delta_{t+1}) + (1-\tau_{K,t+1})\mathcal{R}_{t+1}U_{t+1}\\ &+ \theta_{K,t+1}\left(\log\frac{K_{t+1}}{K_t}\right)\Big],\end{aligned}$$

$$\begin{aligned}\lambda_{K,t}\left(1-\frac{\delta_t}{2}\right)\Omega_t &+ (1-\tau_{K,t})\mathcal{R}_tU_t\kappa\Omega_t + \mathbb{E}_t\Xi_{t+1}\frac{P_{K,t+1}}{P_{K,t}}\theta_{I,t+1}\left(\log\frac{I_{t+1}}{I_t}\right)\frac{K_t}{I_t}\\ &= \frac{1}{1-\tau_{I,t}}\Omega_t + \theta_{I,t}\left(\log\frac{I_t}{I_{t-1}}\right)\frac{K_{t-1}}{I_t},\end{aligned}$$

$$1 = R_t\mathbb{E}_t\Xi_{t+1},$$

$$(1-\tau_{K,t})\mathcal{R}_tU_t(K_{t-1}+\kappa\Omega_tI_t)(1-U_t^{\nu}) = \nu\xi_tK_{t-1}U_t^{\nu}.$$

2.5. Government

The government in the model sets government expenditure as:

$$P_{G,t}G_t = \frac{1+\tau_{GE,t}}{2}(1-\tau_{E,t})P_{Y,t}Y_t.$$

Of course, $\tau_{GE,t}$ and $\tau_{E,t}$ are not actually taxes, but this recycling of notation will ease presentation. Analogously, we assume that:

$$P_{E,t}E_t = \frac{1+\tau_{GE,t}}{2}\tau_{E,t}P_{Y,t}Y_t.$$

We allow the government to set these taxes as a function of the business cycle. In particular, we assume that for each taxed variable $\mho$:

$$\tau_{\mho,t} = \frac{2}{1 + \exp\left[-\left(\phi_{\mho} \log \frac{\mathrm{GDP}_t}{\mathrm{GDP}_{t-1}} + \psi_{\mho} \log \frac{H_t N_{t-1}}{N_t H_{t-1}} + \chi_{\mho} \log \frac{U_t}{U_{t-1}} + \tilde{\tau}_{\mho,t}\right)\right]} - 1,$$

where $\tilde{\tau}_{\mho,t}$ is an exogenous driving process and GDP_t is real gross domestic product in period t, which is defined following the national accounts as a Fisher aggregate of each of the components of GDP:

$$\begin{aligned}\frac{\mathrm{GDP}_t}{\mathrm{GDP}_{t-1}} := {} & \sqrt{\frac{P_{C,t}C_t + P_{I,t}I_t + P_{G,t}G_t + P_{E,t}E_t}{P_{C,t}C_{t-1} + P_{I,t}I_{t-1} + P_{G,t}G_{t-1} + P_{E,t}E_{t-1}}} \\ & \cdot \sqrt{\frac{P_{C,t-1}C_t + P_{I,t-1}I_t + P_{G,t-1}G_t + P_{E,t-1}E_t}{P_{C,t-1}C_{t-1} + P_{I,t-1}I_{t-1} + P_{G,t-1}G_{t-1} + P_{E,t-1}E_{t-1}}}.\end{aligned}$$

GDP_t will not be proportional to Y_t, as Y_t includes assorted adjustment costs.

The government's budget constraint is that:

$$R_{t-1}B_{t-1} + P_{G,t}G_t = T_t + B_t$$

where T_t is total nominal government revenue, which is given by:

$$T_t = \frac{\tau_{C,t}}{1 - \tau_{C,t}} P_{C,t}C_t + \frac{\tau_{I,t}}{1 - \tau_{I,t}} P_{I,t}I_t + \tau_{Y,t}P_{Y,t}Y_t + \tau_{K,t}P_{K,t}\mathcal{R}_t\tilde{K}_t + \tau_{H,t}P_{C,t}W_tH_t.$$

For simplicity, we do not impose any constraints on the path of government debt, B_t, (e.g. we do not impose that $\frac{B_t}{P_tY_t}$ is stationary). In order then to be able to solve the model imposing the standard Blanchard & Kahn (1980) conditions, we will not include either the government budget constraint or B_t in the model that we simulate. Since the only other place that B_t appears is the household's budget constraint, which we may drop due to Walras's law, this has no effect on the rest of the model.

2.6. Market clearing and shock processes

The model is closed with the market clearing condition:

$$\begin{aligned}P_{Y,t}Y_t = \mathrm{NGDP}_t + \frac{\theta_{K,t}}{2}\left(\log \frac{K_t}{K_{t-1}}\right)^2 P_{K,t}K_{t-1} + \frac{\theta_{I,t}}{2}\left(\log \frac{I_t}{I_{t-1}}\right)^2 P_{K,t}K_{t-1} \\ - \xi_t \log(1 - U_t^{\nu})\, P_{K,t}K_{t-1},\end{aligned}$$

where:

$$\mathrm{NGDP}_t = P_{C,t}C_t + P_{I,t}I_t + P_{G,t}G_t + P_{E,t}E_t.$$

We assume that:

$$P_{C,t} = \frac{\tilde{P}_{C,t}}{P_{Y,t}}, \qquad P_{I,t} = \frac{\tilde{P}_{I,t}}{P_{Y,t}}, \qquad P_{G,t} = \frac{\tilde{P}_{G,t}}{P_{Y,t}}, \qquad P_{E,t} = \frac{\tilde{P}_{E,t}}{P_{Y,t}}.$$

We allow all of these relative price variables ($\tilde{P}_{C,t}$, $\tilde{P}_{I,t}$, $\tilde{P}_{G,t}$, $\tilde{P}_{E,t}$) to follow non-stationary trending stochastic processes. This is equivalent to assuming that there is a technology for converting final goods into consumption goods with productivity $\tilde{P}_{C,t}^{-1}$, and similarly for investment, government expenditure and net exports.

For $\mathcal{V} \in \{N, \tilde{\delta}, \xi, \theta_K, \theta_I, \Omega, A_K, A_H, A_X, A_F, A_Z, A_L, A_C, A_G, A_H, P_Y, \tilde{P}_C, \tilde{P}_I, \tilde{P}_G, \tilde{P}_E\}$ we assume that:

$$\Gamma_{\mathcal{V},t} = \frac{\mathcal{V}_t}{\mathcal{V}_{t-1}} = \Gamma_{\mathcal{V}}^{v_t} \frac{S_{\mathcal{V},t}}{S_{\mathcal{V},t-1}^{\gamma_{\mathcal{V}}}} \frac{A_{0,t}^{\iota_{0,\mathcal{V}}}}{A_{0,t-1}^{\iota_{0,\mathcal{V}}}} A_{1,t}^{\iota_{1,\mathcal{V}}}.$$

Here, $\Gamma_{\mathcal{V}}$ is the long-run growth rate of $\mathcal{V}_t$. v_t switches growth on and off, and will always equal one in simulation. $S_{\mathcal{V},t}$ is a shock process with law of motion:

$$S_{\mathcal{V},t} = S_{\mathcal{V},t-1}^{\rho_{\mathcal{V}}} \exp\left(\sigma_{\mathcal{V}} \varepsilon_{\mathcal{V},t}\right), \qquad \varepsilon_{\mathcal{V},t} \sim \text{NIID}(0,1).$$

$A_{1,t}$ and $A_{0,t}$ are common shocks with law of motion:

$$A_{i,t} = A_{i,t-1}^{\rho_{A_i}} \exp\left(\varepsilon_{A_i,t}\right), \qquad \varepsilon_{A_i,t} \sim \text{NIID}(0,1),$$

for $i \in \{0,1\}$. $\gamma_{\mathcal{V}}$ controls whether the shock $\tilde{\Gamma}_{\mathcal{V},t}$ has a permanent ($\gamma_{\mathcal{V}} = 0$) or transitory impact ($\gamma_{\mathcal{V}} = 1$). ι_0 and ι_1 control the strength of the response to the permanent shocks. As one would expect, we simulate the logarithm of $\mathcal{V}_t$ and $\Gamma_{\mathcal{V},t}$, not their levels. Indeed, all variables which are positive an unbounded above are simulated in logarithms. Those with an upper bound, such as U_t and $h_t := \frac{H_t}{N_t}$ are simulated in logits.

Likewise, for $\mathcal{V} \in \{GE, E, H, C, K, I, Y\}$, we assume:

$$\tilde{\tau}_{\mathcal{V},t} - \tilde{\tau}_{\mathcal{V},t-1} = S_{\tilde{\tau}_{\mathcal{V}},t} - \gamma_{\tilde{\tau}_{\mathcal{V}},t} S_{\tilde{\tau}_{\mathcal{V}},t-1} + \iota_{0,\tilde{\tau}_{\mathcal{V}}} \tau_{0,t} - \iota_{0,\tilde{\tau}_{\mathcal{V}}} \tau_{0,t-1} + \iota_{1,\tilde{\tau}_{\mathcal{V}}} \tau_{1,t},$$

where parameters have the same interpretation as before, and:

$$S_{\tilde{\tau}_{\mathcal{V}},t} = \rho_{\tilde{\tau}_{\mathcal{V}}} S_{\tilde{\tau}_{\mathcal{V}},t-1} + \sigma_{\tilde{\tau}_{\mathcal{V}}} \varepsilon_{\tilde{\tau}_{\mathcal{V}},t}, \qquad \varepsilon_{\tilde{\tau}_{\mathcal{V}},t} \sim \text{NIID}(0,1),$$

$$\tau_{i,t} = \rho_{\tau_i} \tau_{i,t-1} + \varepsilon_{\tau_i,t}, \qquad \varepsilon_{\tau_i} \sim \text{NIID}(0,1), \qquad i \in \{0,1\}.$$

These flexible laws of motions permit the economy to undergo substantial structural change over the estimation period.

2.7. Fixed parameters and steady-state computation

The parameters V_0, F_0, X_0, Z_0 and Y_0 will not be identified from our dataset, so we fix them to 1. However, in light of this movements in V_t must be interpreted with caution. Additionally, without any loss of generality, we may set $A_{\mathcal{V},0} = 1$ for $\mathcal{V} \in \{K, H, X, F, Z, L, C, G, H\}$.

We set other "period 0" values to a geometric average from our dataset, over the period in which we have data on all variables, 1954 to 2017.[3] This gives $N_0 = 0.24$ (measured in billions of people), $H_0 = 173.27$ (measured in billions of hours worked per year across the whole population) and $U_0 = 0.81$ (with 1 being full utilization). Additionally, for comparability with the national accounts we set the following "period 0" values to their geometric average over the same period measured in the

[3] Utilisation data is from the Federal Reserve. All other data used in this construction is from the BEA. See the appendix for further details.

index numbers of the national accounts: $\Omega_0 = 0.10$, $C_0 = 45.39$, $G_0 = 62.95$, and $\frac{\widetilde{K}_0}{U_0} = 49.61 + 4.10\kappa + \mathrm{O}(\kappa^2)$.[4]

Given these fixed parameters, it may be shown that the steady-state can be computed by solving a single non-linear equation to determine the level of output.[5] As such, it may be solved quite robustly by first searching for a sign change over progressively wider grids. There was always at most one such sign change for all parameters we tried. However, for some parameter values, we found the residual was complex for all values of output tried, meaning that the model may possess no steady-state for certain parameters. This seems to stem from the upper limit $\bar{h}$ on $\frac{H_t}{N_t}$.

Providing there is a steady state, we would expect the model to be determinate around it.[6] However, given that the model contains unit roots, we need to treat eigenvalues of roughly one as stable. Numerical issues in solving a model on this scale invariably mean that eigenvalues which ought to be exactly one are measured as being greater than one, e.g. $1 + 10^{-5}$. We thus adopt a relaxed criterion for assessing the stability of eigenvalues. In particular, we always just select the number of eigenvalues that ought to be in the unit circle according to the Blanchard & Kahn (1980) conditions, i.e. the number of non-forward-looking variables.

[4] The truncated approximation to $\frac{\widetilde{K}_0}{U_0}$ is only for presentational purposes. We do not make any approximation in the code.

[5] Strictly, calculating the residual of this non-linear equation requires inverting $z \mapsto ze^z$, i.e. evaluating the Lambert W function. Since the Lambert W function has two real branches for some values, there may be two solutions. However, for all of the parameters we examined, only the "−1" branch produced real results, so we always select that one.

[6] While the model contains some potential sources of either indeterminacy or explosive behaviour (tax responses, capacity utilisation, investment being available within the period, complementarity in production) these are unlikely to be quantitatively important.

3. Solving, simulating and estimating models without balanced growth

In this section we develop our method for solving, simulating and estimating models without balanced growth paths. We begin with a discussion of two simulation approaches we do not pursue, before moving on to our preferred approach. We then discuss our method for estimation. We conclude this section with documentation of the changes we made to Dynare (Adjemian et al. 2011) to support our solution, simulation and estimation approaches.

3.1. Asymptotic approximation

Consider a variable v_t, possibly with a linear trend. In the model just presented, this might be $\log A_{K,t}$, for example. Whether the trend in v_t is positive, negative or zero, $x_t := \frac{1}{1+\exp(-v_t)}$ (the logistic transformation of v_t), is guaranteed to have a finite steady-state. This steady-state will be 1 if v_t has a positive trend, 0 if v_t has a negative trend, and some value in $(0,1)$ if it has a zero trend. Thus, if we replace v_t everywhere with $\log\left(\frac{x_t}{1-x_t}\right)$, and likewise for all other trending variables, then we will produce a model with a well-defined steady-state, about which we may approximate. However, there are two problems with this approach. Firstly, it may not be possible to differentiate the model at this steady-state, and, if it is possible, the resulting derivatives may be trivial. For example, suppose that capital and labour are gross complements and capital specific productivity is falling. Then asymptotically, the labour share is zero, and changes in labour specific productivity will have no effect. This leads on to the second problem. Even if the asymptotic derivatives are non-trivial, the asymptotic dynamics may still not be particularly informative for the non-asymptotic dynamics, as the current state is always far from infinity.

3.2. Approximation about a path

Maliar et al. (2015) show that models without balanced growth may be accurately approximated by assuming that after some large number of periods, all growth will stop in the economy. In the language of our model, this is equivalent to assuming that there is some T such that $v_t = 1$ for $t = 0, \dots, T$, and $v_t = 0$ for $t > T$. (Recall that v_t switches growth on and off in our model.) Maliar et al. (2015) prove a turnpike type result that implies that as $T \to \infty$, the approximation error stemming from having a finite T goes to zero. For models with trends but no non-stationary shock processes, this lends itself to a potentially quite efficient solution and estimation procedure. First, solve for the (exact) path of the economy under perfect-foresight, with no shocks, starting from some fixed (or estimated) initial state, and imposing the aforementioned dynamics of v_t. Then, take an approximation to the model's behaviour about this path, e.g. by using the method of Ajevskis (2017). Simulating and estimating the model is then not much harder than if an approximation about the steady-state had been taken.

However, our experiments revealed this method to be impractical on our motivating example. Firstly, finding the economy's path under perfect-foresight proved to be effectively impossible. While the initial and terminal steady-states may be derived by solving only univariate non-linear systems, finding the transition path between these two states requires solving a non-linear system with tens of thousands of variables. (Note that T must be chosen to be quite large, and that we must solve the system past T as even after growth stops it will take some time to reach the asymptotic steady-state.) Even with initial conditions taken from a first order approximation, and with homotopy used to gradually increase growth rates, we still struggled to reliably solve the non-linear system.

Secondly, even when we could solve the non-linear system, doing so was too slow to be practical in our context. While in the absence of non-stationary shock processes, one only needs to solve this non-linear system once per likelihood evaluation, with non-stationary shocks it must be re-solved at least once for each period of the data. Doing so would have taken many minutes per likelihood evaluation.

3.3. Our simulation approach

At its core, our simulation approach is very simple. In our model, all growth rates were raised to the power of v_t. If $v_t = 0$, then the model is not growing, and it has a steady-state about which we can approximate the model. But there is no reason that v_t must be treated as a constant. Instead, we add the equation:

$$v_t = v_{t-1}$$

to the model, and only fix $v_t = 0$ in steady state. At first order, the resulting approximation to the model will take the form:

$$\begin{bmatrix} \hat{x}_t \\ v_t \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{x}_{t-1} \\ v_{t-1} \end{bmatrix} + \begin{bmatrix} B_1 \\ 0 \end{bmatrix} \varepsilon_t,$$

where $\hat{x}_t$ contains all of the model's other endogenous variables, relative to their steady state. Thus, if in simulation or estimation we fix $v_{t-1} = 1$, then we will produce a system with the law of motion:

$$\hat{x}_t = A_{12} + A_{11}\hat{x}_{t-1} + B_1\varepsilon_t.$$

This system is potentially growing over time, due to the constant term, A_{12}.

There are two sources of inaccuracy with this approach. The first source is that given the non-stationarity or trends of the model, an approximation around the initial state of a simulation or likelihood evaluation is not likely to produce good results for the model's behaviour in later periods. Thus, we should re-approximate the model every period. The natural point to approximate around is the point at which the economy would converge to with no future shocks, given $v_{t-1} = 0$ and the rest of the current state. In practice, this means that the routine for finding the model's steady state must accept the current level of the model's non-stationary shock processes as inputs.

Re-approximating the model every period comes with additional complications. Firstly, there may be no steady-state at the new point. If this happens during estimation, then the likelihood of those parameters must be set to minus infinity. To avoid this, it is very important that the model's non-stationary shock processes are all defined in such a way that they can reasonably take values over the entire real line. For example, in a previous version of this paper, we assumed that the logarithms of the growth rate of δ_t, not $\tilde{\delta}_t$, was an auto-regressive process. As a result, we discovered that during estimation, δ_t was going above one, inevitably leading to steady-state non-existence. Secondly, while the model may satisfy the Blanchard & Kahn (1980) conditions at the initial point, it may not satisfy it at later points, at least with a fixed criterion for assessing the stability of eigenvalues. This makes our relaxed approach to assessing eigenvalue stability particularly important.

The second source of inaccuracy with our simulation approach comes from the fact that we are approximating around $v_{t-1} = 0$, but evaluating at $v_{t-1} = 1$. While with perturbation approximations one always needs to be concerned about the radius of convergence, here the problem is more severe. At least when shocks are small enough we can be confident in a standard perturbation approximation, but here we also need growth rates to be small enough for the approximation to be valid. One way to mitigate the approximation errors coming from evaluating far from the steady-state is to use a high order of approximation. This is particularly helpful as the effects of time variation in some processes drops out entirely at first order without growth. For example, in our model, $\theta_{K,t}$ only enters multiplied by $\log \frac{K_t}{K_{t-1}}$, which is zero in a steady-state without growth. Pursuing this line, we developed a new efficient approach to estimating models at second order, discussed below. However, high orders of approximation will not help if $v_{t-1} = 1$ is outside of the radius of convergence, and they will also substantially slow down the code.

As an alternative, we propose approximating around a non-steady-state point. Let us start by defining $y_t := \begin{bmatrix} x_t \\ v_t \end{bmatrix}$, where $x_t = x + \hat{x}_t$ and x is the steady state of x_t. Thus, $y := \begin{bmatrix} x \\ 0 \end{bmatrix}$ is the steady-state of the complete model. We further define $\hat{y}_t := y_t - y = \begin{bmatrix} \hat{x}_t \\ v_t \end{bmatrix}$. Next, note that a general non-linear rational expectations model takes the form:

$$0 = \mathbb{E}_t f(y_{t-1}, y_t, y_{t+1}, \varepsilon_t),$$

for some function f. Since y is the steady-state, we must have:

$$0 = f(y, y, y, 0).$$

Now, suppose we have a first order approximation around this steady-state, say:

$$\hat{y}_t \approx c + A\hat{y}_{t-1} + B\varepsilon_t. \tag{2}$$

Since we are approximating around the steady-state, c will equal 0, but it will help our future purposes to leave it in. Note that our results will readily generalise to higher order approximations.

Furthermore, let y^* be some arbitrary, non-steady-state point. In practice, we will choose $y^* = \begin{bmatrix} x \\ 1 \end{bmatrix}$, so that growth is switched on in the model. Define $\hat{y}^* := y^* - y$. Note that according to the first order approximation, if the economy started at y^*, the next period it would be at $y + c + A\hat{y}^*$, and the period after it would be at $y + c + Ac + A^2\hat{y}^*$. This suggests that we should apply Taylor's theorem to approximate f around the point $(y + \hat{y}^*, y + c + A\hat{y}^*, y + c + Ac + A^2\hat{y}^*, 0)$. Note that this will restore the effects of shocks to variables like our $\theta_{K,t}$ as $\log\frac{K_t}{K_{t-1}}$ will not be zero at the point of approximation. This gives:

$$\begin{aligned}
0 &\approx f^* + f_{-1}(y_{t-1} - y - \hat{y}^*) + f_0(y_t - y - c - A\hat{y}^*) + f_1\mathbb{E}_t(y_{t+1} - y - c - Ac - A^2\hat{y}^*) \\
&\qquad + f_\varepsilon\varepsilon_t \\
&= f^* + f_{-1}(\hat{y}_{t-1} - \hat{y}^*) + f_0(\hat{y}_t - c - A\hat{y}^*) + f_1\mathbb{E}_t(\hat{y}_{t+1} - c - Ac - A^2\hat{y}^*) + f_\varepsilon\varepsilon_t,
\end{aligned}$$

where:

$$\begin{aligned}
f^* &:= f(y + \hat{y}^*, y + c + A\hat{y}^*, y + c + Ac + A^2\hat{y}^*, 0), \\
f_{-1} &:= \left.\frac{\partial f(y_{t-1}, y_t, y_{t+1}, \varepsilon_t)}{\partial y_{t-1}}\right|_{y_{t-1}=y+\hat{y}^*, y_t=y+c+A\hat{y}^*, y_{t+1}=y+c+Ac+A^2\hat{y}^*, \varepsilon_t=0}, \\
f_0 &:= \left.\frac{\partial f(y_{t-1}, y_t, y_{t+1}, \varepsilon_t)}{\partial y_t}\right|_{y_{t-1}=y+\hat{y}^*, y_t=y+c+A\hat{y}^*, y_{t+1}=y+c+Ac+A^2\hat{y}^*, \varepsilon_t=0}, \\
f_1 &:= \left.\frac{\partial f(y_{t-1}, y_t, y_{t+1}, \varepsilon_t)}{\partial y_{t+1}}\right|_{y_{t-1}=y+\hat{y}^*, y_t=y+c+A\hat{y}^*, y_{t+1}=y+c+Ac+A^2\hat{y}^*, \varepsilon_t=0}, \\
f_\varepsilon &:= \left.\frac{\partial f(y_{t-1}, y_t, y_{t+1}, \varepsilon_t)}{\partial y_\varepsilon}\right|_{y_{t-1}=y+\hat{y}^*, y_t=y+c+A\hat{y}^*, y_{t+1}=y+c+Ac+A^2\hat{y}^*, \varepsilon_t=0}.
\end{aligned}$$

As usual, we guess a linear solution to this linear difference equation, of the form:

$$\hat{y}_t \approx c' + A'\hat{y}_{t-1} + B'\varepsilon_t.$$

Substituting in gives:

$$\begin{aligned}
0 &\approx f^* + f_{-1}(\hat{y}_{t-1} - \hat{y}^*) + f_0(c' + A'\hat{y}_{t-1} + B'\varepsilon_t - c - A\hat{y}^*) \\
&\qquad + f_1\mathbb{E}_t(c' + A'(c' + A'\hat{y}_{t-1} + B'\varepsilon_t) + B'\varepsilon_{t+1} - c - Ac - A^2\hat{y}^*) + f_\varepsilon\varepsilon_t \\
&= f^* + f_{-1}(\hat{y}_{t-1} - \hat{y}^*) + f_0(c' - c + A'\hat{y}_{t-1} - A\hat{y}^* + B'\varepsilon_t) \\
&\qquad + f_1\left(c' - c + A'c' - Ac + A'^2\hat{y}_{t-1} - A^2\hat{y}^* + A'B'\varepsilon_t\right) + f_\varepsilon\varepsilon_t. \\
&= f^* + f_{-1}(\hat{y}_{t-1} - \hat{y}^*) + f_0(c' - c + (A' - A)\hat{y}^* + A'(\hat{y}_{t-1} - \hat{y}^*) + B'\varepsilon_t) \\
&\qquad + f_1\left(c' - c + A'c' - Ac + \left(A'^2 - A^2\right)\hat{y}^* + A'^2(\hat{y}_{t-1} - \hat{y}^*) + A'B'\varepsilon_t\right) + f_\varepsilon\varepsilon_t.
\end{aligned}$$

Finally, equating like terms gives:

$$0 = f^* + f_0(c' - c + (A' - A)\hat{y}^*) + f_1\left(c' - c + A'c' - Ac + \left(A'^2 - A^2\right)\hat{y}^*\right), \tag{3}$$

$$0 = f_{-1} + f_0A' + f_1A'^2,$$

$$0 = f_0B' + f_1A'B'\varepsilon_t + f_\varepsilon.$$

The second and third equations are the standard equations characterising the transition matrices of a DSGE model and are independent of c'. Since the model's transversality conditions imply they are subject to the usual Blanchard & Kahn (1980) conditions, they may be solved by routines already existing in Dynare (Adjemian et al. 2011).[7] Having solved for A', then from the first equation, (3), we have that:

$$c' = (f_0 + f_1 + A')^{-1}\left[f_0(c + (A - A')\hat{y}^*) + f_1\left(c + Ac + \left(A^2 - A'^2\right)\hat{y}^*\right) - f^*\right].$$

We have now solved for a new linear approximation to the dynamics of y_t. We can now repeat the whole procedure. I.e. we set $c := c'$, $A := A'$ and $B := B'$ in equation (2) and go on to derive a new c', A' and B'. We iterate this fixed-point procedure until convergence. In practice, this occurs very quickly, certainly within no more than twenty iterations. At the fixed point, $c = c'$ and $A = A'$, thus from equation (3), $f^* = 0$. Hence, the procedure converges to a kind of "pseudo-steady-state" of the original system. This also means that it is less important that y is actually the true steady-state, so relaxed numerical tolerances may be used in solving for y, speeding up and robustifying computation.

3.4. Our estimation approach

Whether we take a first or second order approximation to the model, the model will be non-linear, since we have to re-solve the model in order to step forward one period. While the standard Kalman filter does handle models with time varying linear transition equations, it does not handle cases in which the coefficients of these linear transition equations are themselves functions of the state, which is our situation. Furthermore, given the computational cost of repeatedly re-solving the model, estimation methods based on simulation, such as the particle filter (see e.g. Fernández-Villaverde & Rubio-Ramírez (2007)) or the cubature or unscented Kalman filters (see e.g. Binning and Maih (2015)) are unlikely to be viable.

Instead, we take the simpler approach of using the extended Kalman filter (Smith et al. 1962). The standard extended Kalman filter (EKF) is based upon repeated first order approximation to the model's law of motion about the current best estimate of the state. Note that for us, since we cannot solve the model without approximating it first, the EKF will be a first order approximation to a first/second order approximation to the solution. If we are solving the model at first order, then this is particularly simple. Given the estimate of the state produced by the "update" step in the previous period, we can re-solve the model about the level of the non-stationary shock processes from that state estimate. This provides the required linear transition equations. The "predict" and "update" steps are then identical to those of the standard Kalman filter.

[7] We again choose the eigenvalues with smallest magnitude to enter A'.

The chief downside of this approach has already been hinted at. While smoothed estimates (i.e. full sample) of the state may not be particularly volatile, the filtered estimates (i.e. conditional on information up to that point) can move around substantially. As a result, we end up solving the model at extreme corners of the state-space. Even with a robust steady-state routine, careful modelling and our relaxed approach to eigenvalue stability, we still frequently encountered problems in solving or approximating the model at points within the likelihood evaluation. To ameliorate this, in estimation we decided to truncate movements in the levels of the non-stationary shock processes about which the model is solved to ± 0.1 relative to their previous value. With all non-stationary shocks except taxes in logarithms, this corresponds to $\pm 10\%$ moves. For taxes, this corresponds to at most a five percentage-point move in the $\tau_{\mho,t}$ variables. Additionally, to help the solver move towards a sensible region, when we failed to solve the model even with these constraints, we repeatedly re-attempt the step with progressively smaller bounds. In these cases we add a large penalty to the likelihood, to ensure that our final estimates solve each period with the original ± 0.1 bounds.

If we are taking a second order approximation to solve the model, then we can actually do better than the standard EKF. Gustafsson & Hendeby (2012) show that if $z \sim \mathrm{N}(\mu, \Sigma)$, then for an arbitrary twice differentiable function g, to a second order approximation:

$$g(z) \overset{\text{approx}}{\sim} \mathrm{N}\left(g(\mu) + \frac{1}{2}\left[\operatorname{tr}\left(\Sigma g_i''(\mu)\right)\right]_i, g'(\mu)\Sigma g'(\mu)^\top + \frac{1}{2}\left[\operatorname{tr}\left(\Sigma g_i''(\mu)\Sigma g_j''(\mu)\right)\right]_{i,j}\right),$$

where g' is the Jacobian of g, g_i'' is the Hessian of the i^{th} component of g, $[v_i]_i$ is the vector with i^{th} element v_i and $\left[m_{i,j}\right]_{i,j}$ is the matrix with $(i,j)^{\text{th}}$ component $m_{i,j}$. Furthermore, if g is actually quadratic, then the approximation to the mean and covariance is exact. Thus, for stationary models using this as a basis for a filter enables us to exactly propagate a Gaussian approximation to the state. This is in contrast to the cubature or unscented Kalman filters (see e.g. Binning and Maih (2015)) which further approximate the propagated covariance.

The complete second order extended Kalman filter proceeds as follows. Our presentation here extends Gustafsson & Hendeby (2012) by allowing for shocks to enter non-linearly. We are given a non-linear model with law of motion:

$$y_t = g_{t-1}(y_{t-1}, \varepsilon_t), \qquad \varepsilon_t \sim \mathrm{N}(0, Q_{t-1}),$$

and measurement equation:

$$m_t = Z_{t-1} y_t + \eta_t, \qquad \eta_t \sim \mathrm{N}(0, R_{t-1}).$$

We suppose that we begin with an approximation to the current state as:

$$y_{t-1|t-1} \sim \mathrm{N}(a_{t-1|t-1}, P_{t-1|t-1}).$$

Now define:

$$\tilde{P}_{t-1} := \begin{bmatrix} P_{t-1|t-1} & 0 \\ 0 & Q_{t-1} \end{bmatrix}.$$

Then:

$$\begin{bmatrix} y_{t-1|t-1} \\ \varepsilon_t \end{bmatrix} \sim \mathrm{N}\left(\begin{bmatrix} a_{t-1|t-1} \\ 0 \end{bmatrix}, \tilde{P}_{t-1} \right).$$

Further define the steady-state Jacobian and Hessian by:

$$J_{t-1}^{*} := \left[\begin{matrix} \dfrac{\partial g_{t-1}(y_{t-1}, \varepsilon_t)}{\partial y_{t-1}} & \dfrac{\partial g_{t-1}(y_{t-1}, \varepsilon_t)}{\partial \varepsilon_t} \end{matrix} \right] \Bigg|_{y_{t-1}=a_{t-1|t-1}, \varepsilon_t=0},$$

$$H_{i,t-1} := \left[\begin{matrix} \dfrac{\partial^2 g_{i,t-1}(y_{t-1}, \varepsilon_t)}{\partial y_{t-1} \partial y_{t-1}} & \dfrac{\partial^2 g_{i,t-1}(y_{t-1}, \varepsilon_t)}{\partial y_{t-1} \partial \varepsilon_t} \\ \dfrac{\partial^2 g_{i,t-1}(y_{t-1}, \varepsilon_t)}{\partial \varepsilon_t \partial y_{t-1}} & \dfrac{\partial^2 g_{i,t-1}(y_{t-1}, \varepsilon_t)}{\partial \varepsilon_t \partial \varepsilon_t} \end{matrix} \right] \Bigg|_{y_{t-1}=a_{t-1|t-1}, \varepsilon_t=0}.$$

These quantities are straight-forward to compute from the coefficients of the solution for the model's second order decision rules, whether or not "pruning" (Kim et al. 2008) is used. Out of steady-state, the Jacobian will then be given by:

$$J_{t-1} = J_{t-1}^{*} + \left[\begin{bmatrix} y_{t-1}^{\top} & 0 \end{bmatrix} H_{i,t-1} \right]_{i,\cdot}.$$

Using the aforementioned result from Gustafsson & Hendeby (2012), we then have that if:

$$a_{t|t-1} := g_{t-1}(a_{t-1|t-1}, 0) + \frac{1}{2} \left[\operatorname{tr}\left(\tilde{P}_{t-1} H_{i,t-1} \right) \right]_i,$$

$$P_{t|t-1} := J_{t-1} \tilde{P}_{t-1} J_{t-1}^{\top} + \frac{1}{2} \left[\operatorname{tr}\left(\tilde{P}_{t-1} H_{i,t-1} \tilde{P}_{t-1} H_{j,t-1} \right) \right]_{i,j},$$

then:

$$y_{t|t-1} \overset{\text{approx}}{\sim} \mathrm{N}(a_{t|t-1}, P_{t|t-1}).$$

This completes the "predict" step of the filter. Now note that:

$$\begin{bmatrix} y_t \\ m_t \end{bmatrix} = \begin{bmatrix} I & 0 \\ Z_{t-1} & I \end{bmatrix} \begin{bmatrix} y_t \\ \eta_t \end{bmatrix}.$$

Hence:

$$\begin{aligned} \begin{bmatrix} y_{t|t-1} \\ m_t \end{bmatrix} &\overset{\text{approx}}{\sim} \mathrm{N}\left(\begin{bmatrix} I & 0 \\ Z_{t-1} & I \end{bmatrix} \begin{bmatrix} a_{t|t-1} \\ 0 \end{bmatrix}, \begin{bmatrix} I & 0 \\ Z_{t-1} & I \end{bmatrix} \begin{bmatrix} P_{t|t-1} & 0 \\ 0 & R_{t-1} \end{bmatrix} \begin{bmatrix} I & Z_{t-1}^{\top} \\ 0 & I \end{bmatrix} \right) \\ &\overset{d}{=} \mathrm{N}\left(\begin{bmatrix} a_{t|t-1} \\ Z_{t-1} a_{t|t-1} \end{bmatrix}, \begin{bmatrix} P_{t|t-1} & P_{t|t-1} Z_{t-1}^{\top} \\ Z_{t-1} P_{t|t-1} & F_t \end{bmatrix} \right). \end{aligned}$$

Finally, define:

$$F_t := Z_{t-1} P_{t|t-1} Z_{t-1}^{\top} + R_{t-1},$$

$$K_t := P_{t|t-1} Z_{t-1}^{\top} F_t^{-1},$$
$$a_{t|t} := a_{t|t-1} + K_t(m_t - Z_{t-1} a_{t|t-1}),$$
$$P_{t|t} := P_{t|t-1} - P_{t|t-1} Z_{t-1}^{\top} F_t^{-1} Z_{t-1} P_{t|t-1},$$

then by standard results on conditional distributions of multivariate Gaussians:

$$y_{t|t} \overset{\text{approx}}{\sim} \mathrm{N}\left(a_{t|t}, P_{t|t-1} - P_{t|t-1} Z_{t-1}^{\top} F_t^{-1} Z_{t-1} P_{t|t-1}\right).$$

This completes the "update" step of the filter. In practice, we use a square root form of the update step based on a QR decomposition. This increases robustness when the various covariance matrices are close to being rank deficient.

There is one final novelty in our estimation approach. Traditionally, the initial distribution of the state is either given by its stationary distribution (in stationary models) or by a highly diffuse distribution (in non-stationary models). Since without balanced growth there is no stationary distribution, the former option is not a possibility. Additionally, the latter option is inappropriate when the EKF is being used, both because the initial value of the state will be essentially uninformative, and thus a poor point to approximate around, and because with uninformative initial state, the state estimate will be extremely volatile in the early periods of the dataset. Instead, we propose to treat the initial value of the non-stationary shock processes as a set of additional parameters to estimate. This is akin to estimating an AR(1) process by conditioning on the initial value (i.e. via OLS) rather than by full unconditional maximum likelihood. We set the initial distribution of other elements of the state to the stationary distribution they would have were the non-stationary shock processes in fact constant. Finally, we may get a better approximation to the stationary distribution of other variables, given the model's non-linearities, if we insert a run of empty observations at the start of the dataset. This gives the model a few periods to transition to the appropriate distribution. We use ten years of empty observations in our estimation exercise.

3.5. Documentation of new Dynare features

All of our new solution, simulation and estimation methods are implemented in the "4.5" branch of the custom version of Dynare maintained by the author. A package containing the latest version is always available from:

https://github.com/tholden/dynare/releases

and the full source is available to browse here:

https://github.com/tholden/dynare/tree/4.5 .

By way of documentation for the new features introduced for the purpose of this paper, we list the new options below, and give a brief discussion of their potential use cases. All statements settings elements of the "`options_`" structure should be included directly in the MOD file, before any call to "`stoch_simul`" or "`estimation`".

- **Square root Kalman filters.** No option setting is required as these are used by default in the author's version of Dynare providing the user does not explicitly request a diffuse filter. The square root Kalman filters avoid many problems coming from measurement or state covariance matrices having small eigenvalues. They are necessary, for example, if one wants to use "`lik_init = 2`" in estimation, combined with a very broad initial state distribution (achieved by setting e.g. "`options_.Harvey_scale_factor = 1e6;`").
- **"`options_.non_bgp = 1;`".** This option ensures that $v_t = 1$ during simulation. To use it, one must define an endogenous variable named "`GrowthSwitch`" within the MOD file (i.e. via "`var GrowthSwitch;`"), which should have the law of motion "`GrowthSwitch = GrowthSwitch(-1);`". In steady-state, the variable "`GrowthSwitch`" should be set to zero, however in simulation and estimation it will be fixed at one. This variable thus plays the role of v_t.
- **"`options_.non_bgp_growth_iterations = 20;`".** This option turns on our algorithm for solving the model around the non-steady-state point with $v_t = 1$. Here "`20`" determines the maximum number of iterations, not counting the initial solution around the steady state. It can be replaced by any integer. Setting it to zero returns to approximating around the steady state. If the user wants to combine this with a non-exact original steady state, then the user can set e.g. "`options_.dynatol.f = 1e300;`" to disable Dynare's checks that the provided steady state actually solves the model.
- **"`options_.accurate_nonstationarity = 1;`".** This option instructs Dynare to re-solve the model at each period of a simulation or estimation. For each of the model's state variables, Dynare looks to see if there is a parameter with name "`Initial_STATE_NAME`", where "`STATE_NAME`" is the name of the state variable. If there is, then Dynare sets "`Initial_STATE_NAME`" to the value produced in the previous step. It is then up to user's code for finding the model's steady state to read in the value of these parameters in order to fix the steady state at the correct point. Alternatively, if the user is using Dynare's ability to find the steady state numerically, then the user could add a "`[dynamic]`" tag to the current law of motion for "`STATE_NAME`", and then add an additional line to the model block with "`[static] STATE_NAME = Initial_STATE_NAME;`". Note that the "`options_.accurate_nonstationarity`" option is independent of the

"`options_.non_bgp`" one, so this can be used for any model with non-stationary shocks, even ones without growth.

- **"`options_.accurate_nonstationarity_step_width = 0.1;`".** This option controls the maximum allowable change in the parameters which are updated due to the previous option, during estimation. The value "`0.1`" can be replaced by any number. Using a value that is not too large prevents Dynare from attempting to solve the model in extreme areas of the parameter space during estimation. The default is infinity.
- **"`options_.endogenous_qz_criterium = 1;`".** Setting this option makes Dynare select a number of the smallest eigenvalues from the matrix quadratic equation equal to the number of non-forward-looking variables, rather than just selecting all eigenvalues that are below "`options_.qz_criterium`". This will enable the simulation of the fundamental solution of an indeterminate model, or the least explosive solution of an unstable model. Internally this works by first solving and selecting all eigenvalues below "`options_.qz_criterium`", then, if necessary, re-solving with a relaxed tolerance. As a result, for the sake of speed it is a good idea if the level of "`options_.qz_criterium`" is sensible. We use "`options_.qz_criterium = 1 + 1e-4;`" in the results of the next section.
- **"`options_.extended_kalman_filter = 1;`".** Setting this option turns on the second order EKF. It is not necessary (or desirable) to set this option to use the first order EKF with a model being solved with "`options_.accurate_nonstationarity`". To use this option, somewhere in the MOD file, there should be a call to "`stoch_simul`" with the "`order = 2`" option. Additionally, if the user want to estimate with "pruning", then the call to "`estimation`" must be preceded with "`options_.pruning = 1;`". This option is independent of all of the others previously mentioned, so it permits the computationally tractable estimation of any DSGE model at second order.
- **"`options_.add_empty_presamples = 10;`".** Setting this option makes Dynare add "`10`" periods of empty observations to the start of the sample, to give the model time to transition to its stationary distribution. Obviously, "`10`" may be replaced by any integer.

4. Results

To speed up computation, for now, we estimate at first order. However, we use our algorithm for approximating around a non-steady-state point, which allows us to capture e.g. the effects of shocks to $\theta_{K,t}$. We stress that the second order estimation approach introduced here is highly computationally efficient for models with a balanced growth path. However, the necessity of repeatedly re-solving in the absence of a balanced growth path makes its cost a little high for preliminary work. (The majority of computation time is actually spent finding the decision rules when we estimate at second order. Relative to this, the costs of the second order EKF steps are trivial.)

We estimate via maximum likelihood, to obtain an undistorted view on the data. However, since all of our estimation methods interface with standard Dynare methods, estimating with a prior would have been no harder. The full dataset is described in the appendix, and the complete code for estimating this model is available from this paper's GitHub repository, here:

https://github.com/tholden/NonBalancedGrowth

We start by discussing the key estimated parameters, which are given in the table below. This table excludes most parameters pertaining to shock processes, for conciseness.[8] Since these are preliminary estimates,[9] we will not report standard errors.

Variable	**Value**	**Variable**	**Value**	**Variable**	**Value**	**Variable**	**Value**
α_V		α_X		α_Z		α_Y	
σ_V		σ_X		σ_Z		σ_Y	
η		$\bar{h}$		κ		ν	
$\Gamma_{\tilde{\delta}}$		Γ_{ξ}		Γ_{θ_K}		Γ_{θ_I}	
Γ_N		Γ_{Ω}		Γ_{A_K}		Γ_{A_H}	
Γ_{A_F}		Γ_{A_Z}		Γ_{A_L}		Γ_{A_C}	
Γ_{A_G}		Γ_{A_H}		Γ_{P_Y}		$\Gamma_{\widetilde{P}_C}$	
$\Gamma_{\widetilde{P}_I}$		$\Gamma_{\widetilde{P}_G}$		$\Gamma_{\widetilde{P}_E}$			

The largest departure from the Cobb Douglas special case appears to be in σ_X, the elasticity between consumption goods and leisure, where we find strong

[8] In total there are 212 estimated parameters, so reporting all of them would be overwhelming. The full set of parameter estimates is contained in the file "InitParams.txt" within the repository linked above.

[9] As ever, the likelihood is quite flat in the vicinity of the optimum. Thus, while points with likelihood close to that of the optimum may be obtained quite quickly, actually finding the optimum can take an arbitrarily large amount of time. This is of course one of the reasons for the predominance of Bayesian estimation methods in DSGE macro.

complementarity. Otherwise the estimated departures from Cobb Douglas are small, though there appears to be moderate complementarity between capital and labour.

We conclude by showing some simulations with the estimated parameters. These start from the estimated initial state of the non-stationary shock processes, which corresponds to the year 1915. We switch off all of the model's shocks, thus in a model with a balanced growth path, these plots would just be straight lines. Consequently, any dynamics we see are coming from the interaction between the model's various trends. These simulations are carried out at second order to accurately capture the effects of growth.

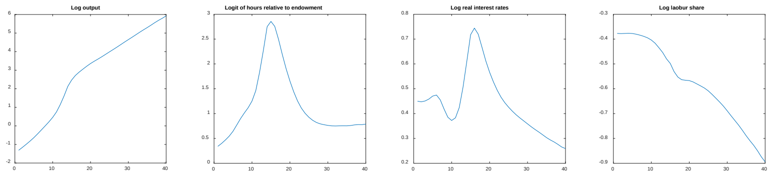

Figure 2: Simulation of the model with the estimated parameters.

We see that the model is generating both a decline in the labour share, and a decline in real interest rates. However, the levels of these variables are somewhat off. One explanation is that the unit root shock processes are leading to weak identification of the initial levels. This will be explored in future work.

5. Conclusion

This paper has put forward a new approach to the solution, simulation and estimation of models without a balanced growth path. All of the associated methods have been implemented in a custom version of Dynare which is available for public use.

For solving such models, we presented a simple approach based on adding an additional endogenous variable which functions as a growth switch. We also showed that performance could be improved by approximating around a non-steady-state point. For simulating these models, we introduced an approach based on repeatedly solving the model given the current values of the non-stationary shock processes. Finally, for estimating these models, we showed that the extended Kalman filter is particularly appropriate, and we developed a second order variant of it.

We have also built and estimated a real business cycle model without a balanced growth path, to guide us in the search for practical solutions to the difficulties raised by such models. The model was able to explain the decline in the US labour share, perhaps in part due to complementarities in production.

We have also made a contribution to the understanding of the sources of the decline in the labour share, as the prior literature has somehow overlooked the increase in the relative price of capital. Together with complementarities in production, this generates a gradual decline in the labour share.

6. References

Adjemian, Stéphane, Houtan Bastani, Fréderic Karamé, Michel Juillard, Junior Maih, Ferhat Mihoubi, George Perendia, Johannes Pfeifer, Marco Ratto & Sébastien Villemot. 2011. *Dynare: Reference Manual Version 4*. CEPREMAP.

Ajevskis, Viktors. 2017. "Semi-Global Solutions to DSGE Models: Perturbation around a Deterministic Path." *Studies in Nonlinear Dynamics & Econometrics* 21 (2). Studies in Nonlinear Dynamics & Econometrics: 1–28.

Antràs, Pol. 2004. "Is the U.S. Aggregate Production Function Cobb-Douglas? New Estimates of the Elasticity of Substitution." *Contributions in Macroeconomics* 4 (1).

Arasaratnam, Ienkaran & Simon Haykin. 2009. "Cubature Kalman Filters." *Automatic Control, IEEE Transactions On* 54 (6): 1254–1269.

Binning, Andrew & Junior Maih. 2015. *Sigma Point Filters for Dynamic Nonlinear Regime Switching Models*. Norges Bank.

Blanchard, Olivier Jean & Charles M. Kahn. 1980. "The Solution of Linear Difference Models under Rational Expectations." *Econometrica* 48 (5): 1305–1311.

Borio, Claudio, Piti Disyatat, Mikael Juselius & Phurichai Rungcharoenkitkul. 2018. *Why So Low for So Long? A Long-Term View of Real Interest Rates*. Puey Ungphakorn Institute for Economic Research.

Chirinko, Robert S. 2008. "σ: The Long and Short of It." *The CES Production Function in the Theory and Empirics of Economic Growth* 30 (2) (June 1): 671–686.

Del Negro, Marco, Domenico Giannone, Marc P. Giannoni & Andrea Tambalotti. 2017. "Safety, Liquidity, and the Natural Rate of Interest." *Brookings Papers on Economic Activity* 48 (1 (Spring). Brookings Papers on Economic Activity: 235–316.

———. 2018. "Global Trends in Interest Rates." Federal Reserve Bank of New York.

Elsby, Michael, Bart Hobijn & Ayseful Sahin. 2013. "The Decline of the U.S. Labor Share." *Brookings Papers on Economic Activity* 44 (2 (Fall)). Brookings Papers on Economic Activity: 1–63.

Epstein, Larry G. & Stanley E. Zin. 1989. "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework." *Econometrica* 57 (4) (July 1): 937–969.

Fernández-Villaverde, Jesús & Juan F. Rubio-Ramírez. 2007. "Estimating Macroeconomic Models: A Likelihood Approach." *Review of Economic Studies* 74 (4). Review of Economic Studies: 1059–1087.

Greenwood, Jeremy, Zvi Hercowitz & Per Krusell. 1997. "Long-Run Implications of Investment-Specific Technological Change." *The American Economic Review* 87 (3): 342–362.

Gustafsson, Frederik & Gustaf Hendeby. 2012. "Some Relations Between Extended and Unscented Kalman Filters." *IEEE Transactions on Signal Processing* 60 (2) (February): 545–555.

Julier, Simon J. & Jeffrey K. Uhlmann. 1997. "New Extension of the Kalman Filter to Nonlinear Systems." In , 3068:3068–12.

Justiniano, Alejandro, Giorgio E. Primiceri & Andrea Tambalotti. 2011. "Investment Shocks and the Relative Price of Investment." *Review of Economic Dynamics* 14 (1) (January): 102–121.

Kaldor, Nicholas. 1961. "Capital Accumulation and Economic Growth." In *The Theory of Capital: Proceedings of a Conference Held by the International Economic Association*, edited by F. A. Lutz & D. C. Hague, 177–222. London: Palgrave Macmillan UK.

Karabarbounis, Loukas & Brent Neiman. 2014a. "Capital Depreciation and Labor Shares Around the World: Measurement and Implications." *National Bureau of Economic Research Working Paper Series* No. 20606.

———. 2014b. "The Global Decline of the Labor Share*." *The Quarterly Journal of Economics* 129 (1) (February 1): 61–103.

Kim, Sunghyun Henry, Jinill Kim, Ernst Schaumburg & Christopher A. Sims. 2008. "Calculating and Using Second Order Accurate Solutions of Discrete Time Dynamic Equilibrium Models." *Journal of Economic Dynamics and Control* 32 (11): 3397–3414.

King, Mervyn & David Low. 2014. "Measuring the '"World"' Real Interest Rate." *National Bureau of Economic Research Working Paper Series* No. 19887.

Klump, Rainer, Peter McAdam & Alpo Willman. 2007. "Factor Substitution and Factor-Augmenting Technical Progress in the United States: A Normalized Supply-Side System Approach." *The Review of Economics and Statistics* 89 (1) (February 1): 183–192.

———. 2012. "The Normalized Ces Production Function: Theory And Empirics." *Journal of Economic Surveys* 26 (5) (December): 769–799.

Laubach, Thomas & John C. Williams. 2003. "Measuring the Natural Rate of Interest." *The Review of Economics and Statistics* 85 (4) (November 1): 1063–1070.

León-Ledesma, Miguel A., Peter McAdam & Alpo Willman. 2010. "Identifying the Elasticity of Substitution with Biased Technical Change." *American Economic Review* 100 (4): 1330–57.

———. 2013. "Production Technology Estimates and Balanced Growth." *Oxford Bulletin of Economics and Statistics* 77 (1) (October 11): 40–65.

Maliar, Lilia, Serguei Maliar, John Taylor & Inna Tsener. 2015. *A Tractable Framework for Analyzing a Class of Nonstationary Markov Models*. National Bureau of Economic Research, Inc.

Smets, Frank & Rafael Wouters. 2007. "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach." *American Economic Review* 97 (3) (June): 586–606.

Smith, Gerald L., Stanley F. Schmidt, Leonard A. McGee, United States. & National Aeronautics and Space Administration. 1962. *Application of Statistical Filter Theory to the Optimal Estimation of Position and Velocity on Board a Circumlunar Vehicle*. Washington, D.C.; Springfield, Va.: National Aeronautics and Space Administration ; For sale by the Clearinghouse for Federal Scientific and Technical Information.

Summers, Lawrence H. 2014. "U.S. Economic Prospects: Secular Stagnation, Hysteresis, and the Zero Lower Bound." *Business Economics* 49 (2) (April 1): 65–73.

7. Appendices

7.1. Data sources

Variables and their data sources are given in the table below. All data series run until 2017. Complete code to download and process the data is provided within the "SourceData" folder of the repository for this paper, https://github.com/tholden/NonBalancedGrowth. Note that we do not directly include $P_{K,t}$ in the data used for estimation. Instead, we include Ω_t, where $\Omega_t = \frac{P_{I,t}}{P_{K,t}}$.

Model variable	Data source	Table	Data variable	Start year
K_t	BEA Fixed Assets	2.2	Real private fixed assets.	1925
$P_{K,t}$	BEA Fixed Assets	2.1 & 2.2	Nominal over real private fixed assets.	1925
N_t	BEA NIPA	2.1	Population.	1929
H_t	BEA NIPA	6.9B-D	Hours worked by full-time and part-time employees.	1948
C_t	BEA NIPA	1.1.3	Real personal consumption expenditures.	1929
I_t	BEA NIPA	1.1.3	Real gross private domestic investment.	1929
G_t	BEA NIPA	1.1.3	Real government consumption expenditures and gross investment.	1929
$\frac{\mathrm{GDP}_t}{\mathrm{GDP}_{t-1}}$	BEA NIPA	1.1.3	Current over lagged real gross domestic product.	1930
NGDP_t	BEA NIPA	1.1.5	Nominal gross domestic product.	1929
T_t	BEA NIPA	3.1	Nominal current tax receipts.	1929
$\frac{P_{C,t}W_tH_t}{P_{Y,t}Y_t}$	BEA NIPA	2.1 & 1.1.5	Compensation of employees over nominal gross domestic product minus proprietor's income with inventory valuation and capital consumption adjustment.	1929
$P_{C,t}$	BEA NIPA	1.1.3 & 1.1.5	Nominal over real personal consumption expenditures.	1929
$P_{I,t}$	BEA NIPA	1.1.3 & 1.1.5	Nominal over real gross private domestic investment.	1929
$P_{G,t}$	BEA NIPA	1.1.3 & 1.1.5	Nominal over real government consumption expenditures and gross investment.	1929

R_t	FRB	H.15	1-Year Treasury Constant Maturity Rate.	1954
U_t	FRB	G.17	Total index of capacity utilisation, back-cast from 1967 as described below.	1948

7.2. Back-casting utilisation data

The FRB's "Total Index" of capacity utilisation (from table G.17) is only available from January 1967. However, data on capacity utilisation in (1) manufacturing, (2) primary & semifinished processing, and (3) finished processing is available from January 1948. We use these three other series in order to back-cast the "Total Index" measure to January 1948.

To do this, we first scale all variables so they are between 0 and 1, rather than 0 and 100, then we take logit transformations to produce transformed variables defined on the entire real line.

As a first step, we estimate a regression of the *time reversed* "Total Index" variable on the *time reversed* values of the three indicators, allowing for ARMA errors with t-distributed shocks. Using a t-distribution was strongly preferred over normal shocks with the AIC criterion. Additionally, on the basis of the AIC criterion we selected AR lags 1 and 12 and MA lag 1. Having time reversed the original data, we can back-cast the original data by forecasting the time reversed data, which is straightforward. Were the shocks normally distributed, this would be the optimal back-cast, however with t-distributed shocks it will not be.

To obtain optimal back casts, we then estimate a regression of the original "Total Index" variable on the three indicators, allowing for the same error structure as before. Given the non-Gaussianity, these estimates will not in general agree with those on the time-reversed data. Holding fixed these parameters, for any hypothetical back-cast, we can calculate the likelihood of the complete time-span, treating the back-cast as if it was observed. We can then obtain the maximum likelihood back cast by maximising this likelihood over the space of possible back-casts. This is relatively straight-forward despite the high dimension due to the existence of a good starting point, namely the back-cast produced from the time-reversed data.

We finish by taking logistic transforms to map back to the unit interval, then averaging the monthly data to obtain annual data.