Süssmuth, Bernd; Irmen, Andreas; Heer, Burkhard

Conference Paper

Taxation, Automation Capital, and the Functional Income Distribution

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2020: Gender Economics

Provided in Cooperation with:
Verein für Socialpolitik / German Economic Association

Suggested Citation: Süssmuth, Bernd; Irmen, Andreas; Heer, Burkhard (2020) : Taxation, Automation Capital, and the Functional Income Distribution, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2020: Gender Economics, ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg

This Version is available at:
https://hdl.handle.net/10419/224572

Taxation, Automation Capital, and the Functional Income Distribution

Burkhard Heer[a,b], Andreas Irmen[c,b], and Bernd Süssmuth[d,b]

[a]University of Augsburg, Department of Economics, Universitätsstr. 16, 86013 Augsburg, Germany, Burkhard.Heer@wiwi.uni-augsburg.de

[b]CESifo, Munich, Germany

[c]University of Luxembourg, Luxembourg

[d]University of Leipzig, Germany

Preliminary version: February 27, 2020

Comments welcome!

JEL classification: D33, E62, O41, J11, J20

Key Words: Functional income distribution, labor income share, income taxes, automation capital, demography, growth

Abstract:
The functional income distribution in the US and most OECD countries has been characterized by an increasing capital income share and a declining wage share over the last decades. We present new evidence for the US economy that this fact is not only explained by technical change and globalization, but also by the dynamics of capital and labor income taxation, automation capital, and population growth. In the empirical analysis, we find indications for cointegrating equations for the 1974-2008 period. Permanent effects on factor shares emanate from labor (relative to capital) tax shocks. Changes in relative factor taxation also permanently affect the use of robots. Variance decompositions reveal that taxing accounts for up to 22% and up to 35% of observed changes in the two income shares and in automation capital, respectively. In a second step, we present a standard neoclassical growth model augmented by automation capital and capital adjustment costs that is able to replicate the dynamics of the observed functional income distribution in the US during the 1965-2015 period. In particular, we demonstrate that the fall in the wage share would have been significantly smaller if labor and capital income tax rates had remained at their respective level of the 1960s.

1 Introduction

The functional income distribution in most OECD countries has changed significantly during recent decades. Fig. 1.1 displays the wage share as an annual time series ranging from 1960 to 2018 in the United States (US), Japan, and the Euro-12 area, respectively. Evidently, the wage share in total income has declined by 8% to 15% over the last 50 years in these economies, most markedly in Japan.

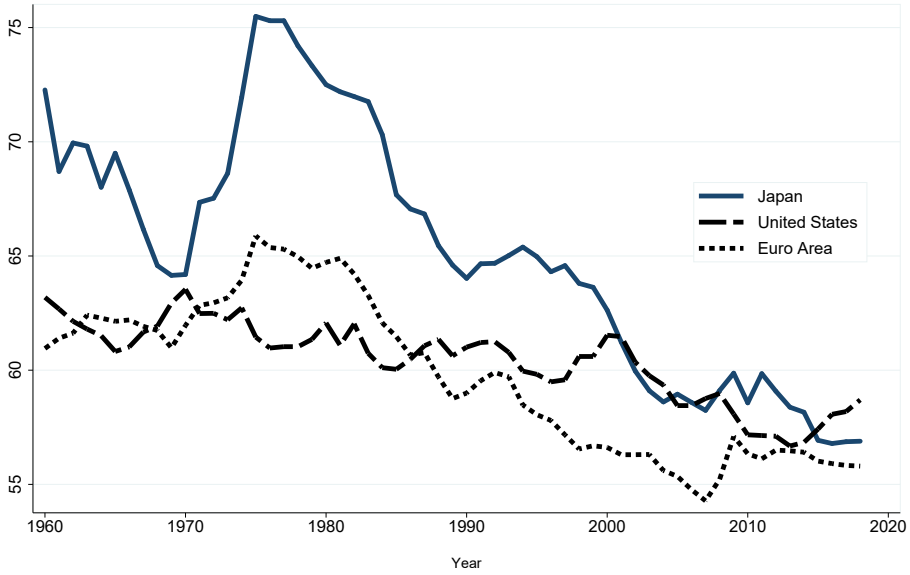
Many hypotheses have been formulated on the factors that contribute to the explanation of the rising capital income share and the declining wage income share. Globalization and skill-biased technological change are among the prime reasons. More recently, Piketty (2014) has brought attention to an additional channel in order to explain the declining share of wage income. In his seminal book on 'Capital in the Twenty-First Century' he points out that the real interest rate is higher than the economic growth rate over recent decades and, accordingly, the income share that is directed towards the capitalists is increasing over time. However, the premises of his hypothesis crucially depend on the assumption that the capitalists have an extremely high and empirically unobserved savings rate. If they consume part of their income, the wealth-income ratio and the capital income might even fall over time for realistic savings rates.[1]

In the US economy, the decrease in the wage share –and the simultaneous (slight) increase in the capital income share– (Fig. 1.2) is mirrored by the shrinking difference in the income taxes on labor and capital income and the decline in the population growth rate (Fig. 1.3). Since labor and capital income are subject to different amounts of allowances and exemptions and have to be paid from households and the corporate sector (differing with regard to their tax treatment), capital and labor income are not burdened equally by the US tax system. Fig. 1.3 (left schedule) presents the effective capital and labor income tax rates for the US during 1954 to 2010 as computed by Gomme et al. (2011). Evidently, capital income was taxed more heavily than labor income in the last century; however, during recent years, the tax rates on capital and labor income have converged. The average capital and labor income rates for this period amount to 41% and 23%, respectively. Another development during recent

[1] To see this, assume a simple economy with two agents, a capitalist and a worker. Assume furthermore that the worker consumes all his income, while the capitalist saves 50% of his income. For a constant growth rate $g = 2\%$ and a real interest rate of $r = 3\%$ and an initial capital-income ratio of 4, it is easy to show that the capital income share falls by half, from 12% to 6%, within 140 years.

decades that coincides with the fall in the wage share is the behavior of population growth rate which has fallen, in annualized terms, from 1.7% in 1950 to 0.7% in recent years. The dynamics of the quarterly population growth rate are displayed in the right schedule of Fig. 1.3.

Figure 1.1: Adjusted wage as a share of GDP [%], 1960-2018: Japan, US, Eurozone

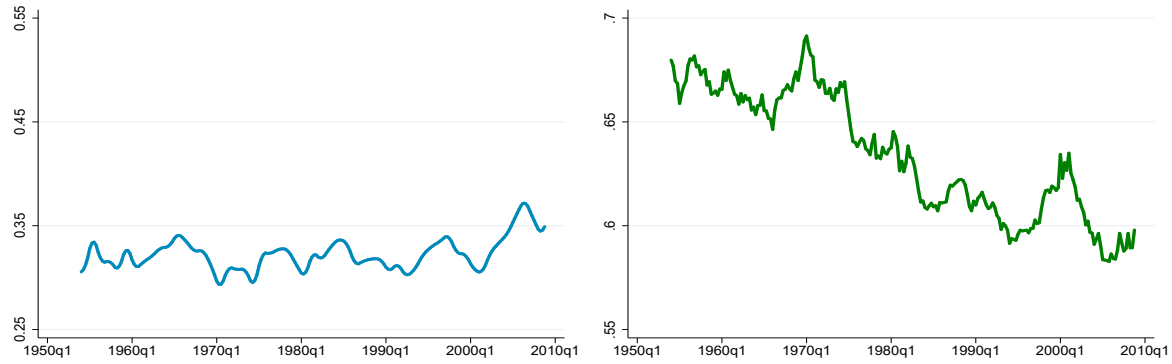


Note: Percentage of GDP at current prices; Source: AMECO database, OECD

Both developments of the income tax rates help to understand why the income of the production factor labor has declined. The downward trend in the capital income tax rate has increased the incentives to build up capital, while labor supply has declined both because of increase in the labor income tax rate and because of lower population growth. As argued by Heer and Irmen (2014), among others, the relative decline in the labor force has increased incentives to invest in labor-saving technology. As we observe empirically, automation capital replaces labor in the production process. Lower population growth and a decline in the difference between capital and labor income taxes precipitate this process. As automation capital replaces labor, the wage share minimaly declines both empirically and in model simulations. A series to proxy automation capital by the (nowcasted) stock of robots per 1,000 (full time) employees in the US is shown in Fig. 1.4.

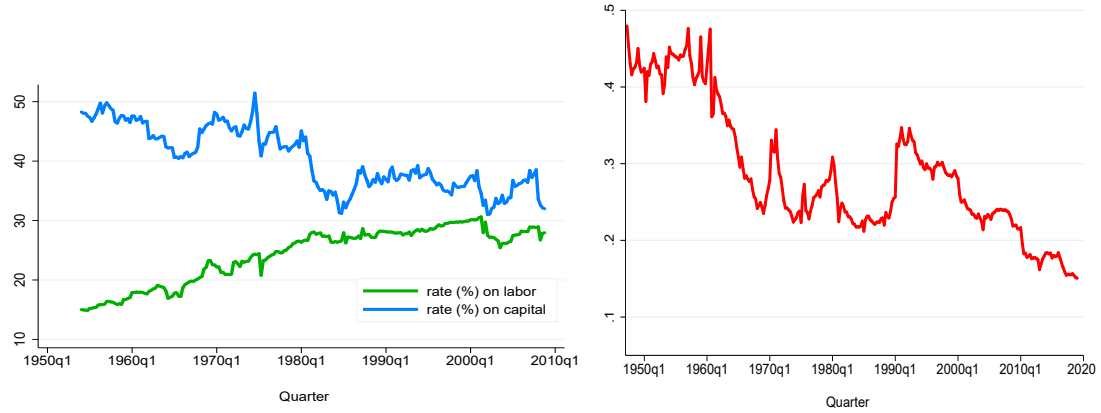
In the first part of the paper, we provide empirical evidence supporting this hypothesis. We find indications for cointegrating relationships for the period from the first

Figure 1.2: Quarterly US capital share (left schedule) and wage share (right schedule)



Sources: Fernald (2014) for capital share; wage share: own calculation based on Bureau of Economic Analysis (BEA) series for GDP and gross domestic income (GDI), the latter has been adjusted for self-employed income; see the Appendix for detail.

Figure 1.3: Quarterly US income tax rates and population growth (SA): each in %

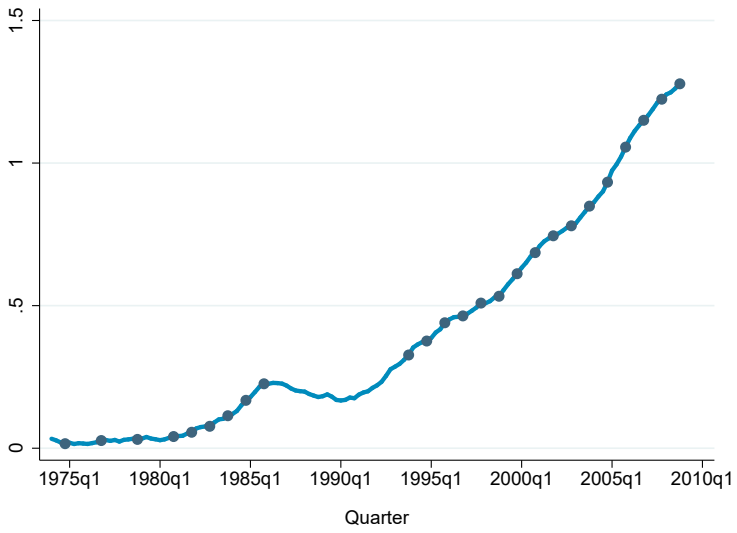


Sources: Gomme et al. (2011) for tax rates; BEA for population growth, seasonal adjustment: X12-ARIMA

quarter of 1974 to the fourth quarter of 2008 (henceforth, 1974:q1-2008:q4). Permanent effects on factor shares emanate from shocks in relative factor taxation. Changes in relative factor taxation also permanently and sizably affect the use of automation capital. The forecast error variance decomposition (FEVD) analysis of fitted vector error correction (VEC) models reveals that taxing policies account for up to roughly 22% of observed changes in the two income shares and for up to about 35% of the dynamics in automation capital.

In the second part of the paper, we formulate a model that incorporates all these explanatory factors of the dynamics for the functional income distribution. Accordingly, we set up a neoclassical growth model with a production function in physical

Figure 1.4: Robots per 1K workers in the US: annual and nowcasted quarterly data



Note: *Dotted* data points up to 1985:q4 are own calculations based on Tani (1989) and CPS, starting 1994:q4 they are obatined from Acemoglu and Restrepo (2019), original source: IFR; *Solid* blue line: nowcasted (pseudo-)quarterly data; detail on nowcasting is given in Appendix

Sources: Tani (1989), Acemoglu and Restrepo (2019), IFR, CPS

capital and augmented labor where the latter production factor is simply the sum of automation capital and labor so that automation capital substitutes linearly for labor. With this model, we compute the dynamics of the wage share, the capital stock, the automation capital stock and endogenous labor over the period 1965-2029. We assume that automation capital was not introduced into production until 1965. As additional elements of the model, we use the time series of labor and capital income tax rates as well as the population growth rate displayed in Fig. 1.3. To render the model even more realistic, we also introduce quadratic adjustment costs of capital in the tradition of Hayashi (1982). As a result, our model is able to replicate the downward trend in the wage share closely; in particular, it replicates the drop in the wage share from 62% to 57% between 1965 and 2015.

The paper is organized as follows. In Section 2, we present time series evidence for the US economy relying on cointegration analysis and VEC models. An extensive range of corresponding test detail and robustness checks is summarized in the Appendix. Section 3 presents a neoclassical growth model augmented by automation capital in production which allows us to study i) the dynamic effects of capital and labor income taxes and ii) replicate the dynamics of the wage share relative to the capital share over

the period 1965-2015. Section 4 presents our results from the transition analysis in the model with automation capital. We demonstrate that all variables population growth, capital income taxes, and labor income taxes have a significant effect on the long-run wage share, while the drop in the wage share has been mainly caused by the decline in the capital income tax rate. Section 5 concludes. The methodological description of the model's transition analysis is delegated to the Appendix.

2 Empirical Analysis

2.1 Time series

To work with a decent sample length, allowing the use of multivariate time series techniques, we throughout consider series in quarterly frequency for the US economy. Our data set entirely spans the period from the first quarter of 1974 to the fourth quarter of 2008. The limiting factor with regard to the length of our sample is the quarterly series for US labor and capital income tax rates taken from Gomme et al. (2011). The corresponding limiting factor regarding the start of the empirical analysis is Tani (1989) who provides the numerator of the early part of actual datapoints (in the 1970s biannually and as of the 1980s up to the middle of the decade in annual frequency) for the nowcast of the automation capital series as shown in Fig. 1.4. That is the industrial robot population in the US, where industrial robot is defined by the still in use Industrial Organization for Standardization (ISO) definiton.[2] The denominator in the construction of datapoints underlying the dots in Fig. 1.4 prior to 1986 is the corresponding annual average of the seasonally adjusted (SA) number of full time employees in the US provided in the Current Population Survey (CPS). The ISO normed definition of robots and the expression in "per thousand workers" units allows us to combine it with corresponding annual data from the International Federation of Robotics (IFR) as provided by Acemoglu and Restrepo (2019) in order to nowcast a quarterly series of automation capital. For detail on the used nowcasting technique, based on the procedure proposed by Shumway and Stoffer (2008) relying on

[2]Accordingly, the "industrial robot is an automatic position-controlled reprogrammable multifunctional manipulator having several degrees of freedom capable of handling materials, parts, tools, or specialized devices through variable programmed motions for the performance of a variety of tasks" (Tani, 1989, p. 192).

the Kalman filter in combination with the expectation maximization (EM) algorithm, see the Appendix.

Fernald (2014) provides the capital share of income series in quarterly frequency. Quarterly series for the US population (in thousand) stems from the BEA and is provided, as not SA series, in the FRED Federal Reserve Bank of St. Louis database. We seasonally adjust this series by means of X12-ARIMA. Also from FRED we retrieve the BEA series of US quarterly gross domestic income, i.e. compensation of employees in the form of paid wages and salaries, and corresponding GDP series, from which we construct our wage share series.

2.2 VEC model analysis

In the following we rely on the popular maximum likelihood (ML) based framework for estimation and inference in cointegrating systems that is also referred to as "Johansen approach." Before going into detail of the VEC model (VECM) analysis, it is advisable to clarify the central variables which are to be included. These series are the two tax rates as well as the two corresponding income shares. While it is, due to natural log transformation, technically feasible to jointly consider the labor and capital income shares, a joint integration of both types of income tax rates, despite taking natural logs, into a particular VECM specification is for reasons of (near perfect) collinearity not viable. An alternative to including both tax rates is the construction and use of the labor-tax-to-capital-tax ratio (LCTR, Fig. 2.1) which serves as a workaround in our VECM specifications. It organically takes care of the historical tax policy mix.

2.2.1 Johansen procedure

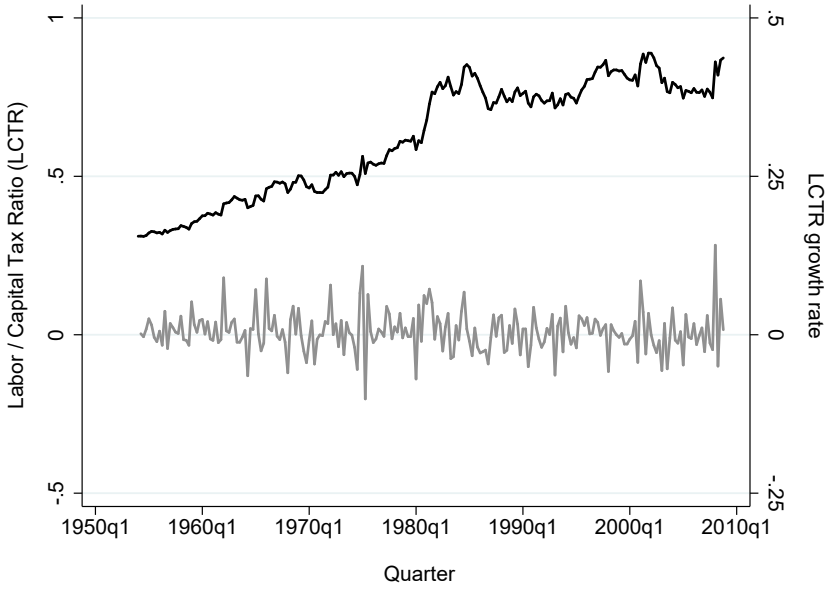
Our reduced form model space consists of three central dimensions: a relatively exogenous variable X_t (population growth), a policy variable Y_t (the factor tax policy mix, that is, the LCTR), and a multivariate group of response variables $\mathbf{W}_t$ (our automation capital proxy and the two factor shares), making it

$$\mathbf{Z}_t = [X_t, Y_t, \mathbf{W}_t] \Rightarrow \mathbf{Z}_t = \mathbf{A}_1\mathbf{Z}_{t-1} + \mathbf{A}_2\mathbf{Z}_{t-2} + ... + \mathbf{A}_p\mathbf{Z}_{t-p} + \mathbf{u}_t, \tag{2.1}$$

which is summarized in standard VEC notation as

$$\Delta\mathbf{Z}_t = \Gamma_1\Delta\mathbf{Z}_{t-1} + \Gamma_2\Delta\mathbf{Z}_{t-2} + ... + \Gamma_{p-1}\Delta\mathbf{Z}_{t-p-1} + \Pi\mathbf{Z}_{t-1} + \mathbf{u}_t, \tag{2.2}$$

Figure 2.1: Labor-tax-to-capital-tax ratio (LCTR), 1954:q1 to 2008:q4



Note: Black – levels (left ordinate); grey – log first differences (right ordinate); data: Gomme et al. (2011)

where $\Gamma_i = (\mathbf{I} - \mathbf{A}_1 - \mathbf{A}_2 - ... - \mathbf{A}_p)$ for all $i = 1, ..., p$. Π can be thought of as being composed of adjustment speed matrix $\mathbf{a}$, and long run coefficients matrix $\mathbf{b}$, such that $\Pi = \mathbf{ab}'$, where $\mathbf{b}'\mathbf{Z}_{t-1}$ is the vectorial analogue to the error correction term in the Engel-Granger approach. For an exemplary unity lag order

$$\Delta \mathbf{Z}_t = \begin{pmatrix} \Delta Y_t \\ \Delta X_t \\ \Delta \mathbf{W}_t \end{pmatrix} = \Gamma_1 \begin{pmatrix} \Delta Y_{t-1} \\ \Delta X_{t-1} \\ \Delta \mathbf{W}_{t-1} \end{pmatrix} + \Pi \begin{pmatrix} Y_{t-1} \\ X_{t-1} \\ \mathbf{W}_{t-1} \end{pmatrix} + \mathbf{e}_t = \Gamma_1 \begin{pmatrix} \Delta Y_{t-1} \\ \Delta X_{t-1} \\ \Delta W_{1\,t-1} \\ \Delta W_{2\,t-1} \\ \Delta W_{3\,t-1} \end{pmatrix} +$$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \\ a_{51} & a_{52} & a_{53} & a_{54} \end{pmatrix} \begin{pmatrix} b_{11} & b_{21} & b_{31} & b_{41} & b_{51} \\ b_{12} & b_{22} & b_{32} & b_{42} & b_{52} \\ b_{13} & b_{23} & b_{33} & b_{43} & b_{53} \\ b_{14} & b_{24} & b_{34} & b_{44} & b_{54} \end{pmatrix} \begin{pmatrix} Y_{t-1} \\ X_{t-1} \\ W_{1\,t-1} \\ W_{2\,t-1} \\ W_{3\,t-1} \end{pmatrix} + \mathbf{e}_t \tag{2.3}$$

;

and the central requirement of cointegration (CI) is a reduced rank of $\Pi = \underset{n \times r}{\mathbf{a}} \left(\underset{n \times r}{\mathbf{b}} \right)'$, that is, $\mathbf{Z}_t \sim I(1) \Rightarrow \Delta \mathbf{Z}_{t-1} \sim I(0) \Rightarrow \Pi \mathbf{Z}_{t-1} \overset{!}{\sim} I(0)$ for $\mathbf{u}_t \sim I(0)$, allowing for

up to $(n-1) = 5-1 = 4$ *CI*-relationships of the form $\mathbf{b}\prime\mathbf{Z}_{t-1} \sim I(0)$, $r \leq (n-1)$ cointegrating vectors $\in \Pi$, i.e. r columns of $\mathbf{b}$ form r linearly independent stationary combinations of variables $\in \mathbf{Z}_t$.

In a first step of the Johansen-procedure, we perform augmented Dickey-Fuller (ADF) tests for unit roots for all series of our dataset as described in the preceding section. In all cases, except the population series, for which the following holds for its log first differences transform, the unit root hypothesis cannot be rejected at a one percent level of significance according to the MacKinnon approximate p-values. For any of the log first differences transforms the null of a unit root is rejected at all conventional levels of significance. The order of integration of variables that are considered is the same, i.e. $I(1)$; see the Appendix for detail. In a second step the choice of the appropriate lag length is made resorting to likelihood ratio (LR) testing, where we set, following Schwert (1989), $p_{\max} = \left[12 \cdot (T/100)^{\frac{1}{4}}\right]$ with rectangular parantheses denoting the nearest integer part of the argument which in our case, i.e. for $T = 140$, equals 13 quarters.

As $\Pi = -(\mathbf{I} - \mathbf{A}_1 - \mathbf{A}_2 - ... - \mathbf{A}_p)$, or equivalently $\Pi = \sum_{i=1}^{p} \mathbf{A}_i - \mathbf{I}$, the Johansen-procedure makes use of Granger's Representation Theorem which states that if matrix Π has reduced rank $r < n$ with n denoting the number of non-stationary variables considered, then there exist $n \times r$ matrices $\mathbf{a}$ and $\mathbf{b}$ each with rank r such that $\Pi = \mathbf{ab}'$ and $\mathbf{b}'\mathbf{Z}_t \sim I(0)$; r then is the number of cointegration relations (cointegrating rank) and each column of $\mathbf{b}$ is a cointegrating vector. However, before estimating Π, deterministic components of general system

$$\begin{aligned} \Delta\mathbf{Z}_t &= \Gamma_1\Delta\mathbf{Z}_{t-1} + ... + \Gamma_{p-1}\Delta\mathbf{Z}_{t-p-1} \\ &\quad + \mathbf{a}\begin{pmatrix} \mathbf{b} \\ \mathbf{m}_1 \\ \mathbf{d}_1 \end{pmatrix}' \begin{pmatrix} \mathbf{Z}_{t-1} & 1 & t \end{pmatrix} + \mathbf{m}_2 + \mathbf{d}_2 t + \mathbf{u}_t \end{aligned} \tag{2.4}$$

need to be chosen. The above system can discriminate four central versions: (v1) No intercept or trend in the cointegrating equation (CE) or VAR part ($\mathbf{d}_1 = \mathbf{d}_2 = \mathbf{m}_1 = \mathbf{m}_2 = 0$); (v2) intercept and no trend in CE part and neither intercept nor trend in VAR part ($\mathbf{d}_1 = \mathbf{d}_2 = \mathbf{m}_2 = 0$), i.e. the no linear trend in data case (first differences have zero mean); (v3) intercept in CE part and VAR part, but no trends ($\mathbf{d}_1 = \mathbf{d}_2 = 0$), i.e. no linear trends in levels of data case; (v4) intercept in CE part and VAR part

paralleled by a linear trend in CE part or VAR part, i.e. the linear trend in the CE case, sometimes also referred to as exogenous growth case. Following the Pantula Principle, we start with the most restrictive model, i.e. $r = 0$ in combination with v1, moving gradually to the least restrictive model or to $r = n - 1$ in combination with v4. For each gradual step, the trace-test statistics is compared with the critical value and the iteration stops when for the first time the null of no cointegration is not rejected. This determines the order of Π, i.e. the number of cointegration vectors. Besides trace-based rank testing, we cross-check and validate our findings with maximum eigenvalue and information criteria-based cointegration rank tests. For detail see the Appendix.

2.2.2 Identification strategy

Our set-up implies an identification scheme for the VAR part of the VECM specification that resembles what has become known as "Slow-r-Fast" scheme –named after order of partitioning of dependent variables' vector by nature of its, partially sub-vectorial, elements– in the literature (Stock and Watson, 2016, p. 455, pp. 477-478). It has been used, in particular, to identify monetary policy shocks (Bernanke et al., 2005). Under this scheme so-called slow-moving variables such as output and prices do not respond to monetary policy rate dynamics or to movements in fast-moving variables, such as expectational variables, within the period. This renders a block recursive scheme, for which the ordering within the respective elements of the partitioned $\mathbf{Z}_t$ vector is not decisive for the system rotation matrix. In our application, population growth is supposed to be "the least" endogenous and does not respond to the policy variable, i.e. to either the labor or the capital tax rate or to the LCTR, and to any of the response variables, which are comprised of our automation capital measure and the two factor shares, i.e. the labor share and capital share of income. As we are only interested in the responses to innovations in the tax policy variables, the ordering within $\mathbf{W}_t$ is uncritical. For the $\Delta\mathbf{Z}$ (VAR-) part of the respective VECM, the following block recursive scheme to identify responses to innovations ε, with orthogonalized analogues η, is implied

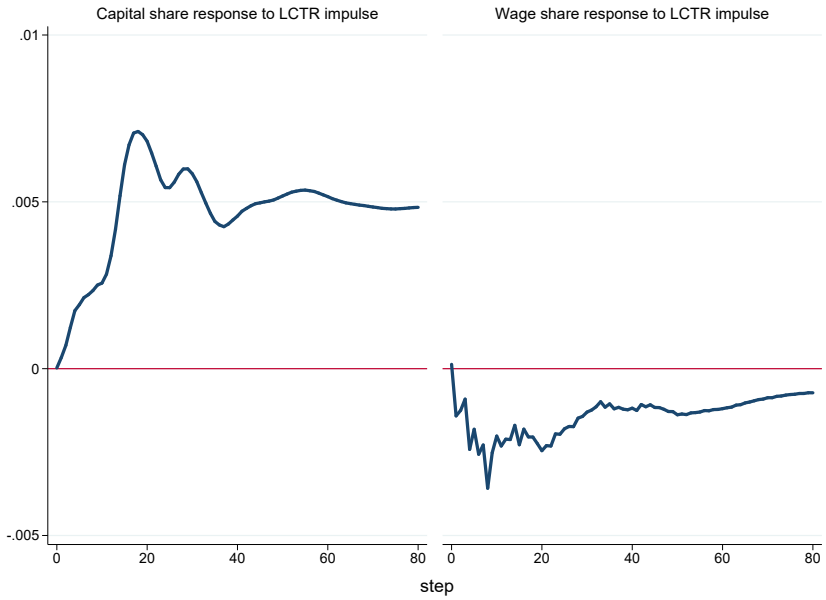
$$\begin{pmatrix} \eta_t^X \\ \eta_t^Y \\ \eta_t^{\mathbf{W}} \end{pmatrix} = \begin{pmatrix} H_{XX} & 0 & 0 \\ H_{YX} & H_{YY} & 0 \\ H_{\mathbf{W}X} & H_{\mathbf{W}Y} & H_{\mathbf{WW}} \end{pmatrix} \begin{pmatrix} \varepsilon_t^X \\ \varepsilon_t^Y \\ \varepsilon_t^{\mathbf{W}} \end{pmatrix} \text{ for } \Phi(L)\Delta\mathbf{Z}_t = \begin{pmatrix} \eta_t^X \\ \eta_t^Y \\ \eta_t^{\mathbf{W}} \end{pmatrix}, \quad (2.5)$$

where $\mathbf{Z}_t$ is partitioned $\mathbf{Z}_t = \begin{pmatrix} X_t & Y_t & \mathbf{W}_t \end{pmatrix}'$, and $H_{\mathbf{WW}}$ is square.

2.2.3 Results

Our lag order selection of $p = 13$ for (2.1), (2.2), corresponding to $p_{\max}$, is supported by the adequate Likelihood ratio (LR) test. The Johansen testing procedure fails to reject the null of at most three cointegrating equations in (2.1), (2.2) for all versions, (v1) to (v4), of VECM representation (2.4). The second cointegration equation in the standardly used Johansen-normalization identification clearly indicates a statistically significant equilibrium relationship between the LCTR series and the two factor shares. Additionally, our specification is stable adhering to the implied eigenvalue stability condition. Further detail on these findings confirming the acceptability of our specification is summarized in the Appendix.

Figure 2.2: Orthogonalized IR functions of factor shares to LCTR shock

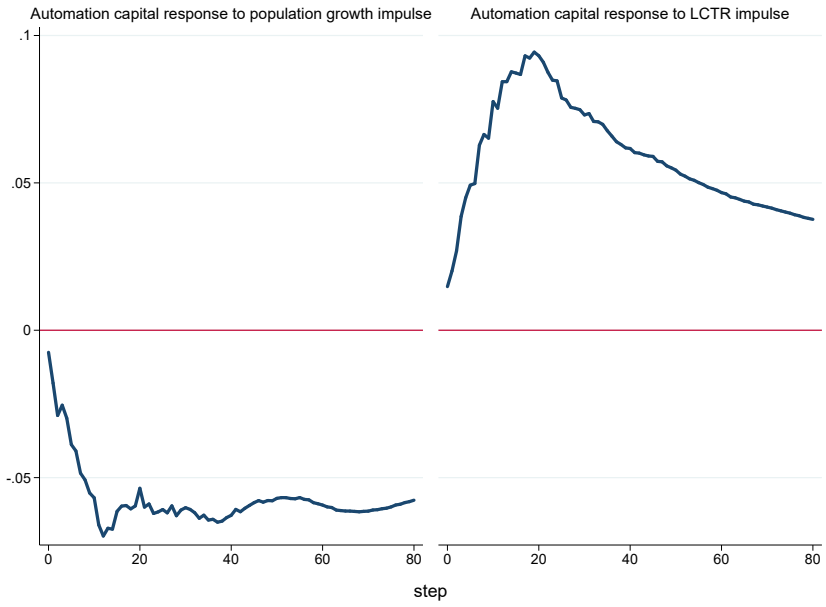


Impulse response (IR) functions of factor shares in response to a relative increase of taxing labor vis-à-vis capital, i.e. to a positve LCTR innovation, are given in Fig. 2.2. A positive one percent LCTR shock lets the capital share of production permanently increase by five percent. The wage share response (right schedule of Fig. 2.2) is negative. Though also permanent in nature, it is quantitatively less pronounced.

The use of robots per 1,000 workers as a response variable reasonably shows a permanent negative reaction to a positive shock of population growth (left schedule of Fig. 2.3) with an orthogonalized one percent shock implying a long-run decrease of

five percent. In contrast, the automation capital use in the US responds to relative factor taxation in a more pronounced fashion, peaks after 25 quarters, and lies slightly below 0.5 percent after two decades (right schedule of Fig. 2.3).

Figure 2.3: Orthogonalized IR functions of automation capital use series



Note: Left (Right) diagram shows response to population growth (LCTR) impulse.

Table 2.1: FEVD values for an LCTR shock by response variable

Step	Capital share	Wage share	Automation capital
10	0.010460	0.126208	0.296361
20	0.050732	0.118803	0.354056
30	0.080220	0.125973	0.347809
40	0.085780	0.127680	0.332902
50	0.086906	0.129162	0.325100
60	0.090602	0.128036	0.313584
70	0.091387	0.123696	0.299503
80	0.091690	0.117447	0.289272

Table 2.1 shows some FEVD statistics for the three response variables, capital share, wage share and automation capital, given a positive LCTR innovation. After about 30 to 40 quarters, relative factor taxation accounts for up to 9% (13%) of observed

changes in the capital (labor) income share and for up to 35% of the dynamics in the automation capital series. While the effect starts to die out after 20 years (falling below 30%) for the use of robots, it only slightly diminishes for the two factor shares.

3 The Neoclassical Growth Model with Dynamic Taxes and Automation Capital

In this section, we present our neoclassical growth model that is augmented by automation capital. In addition, we integrate the dynamics of the tax rates on capital and labor income.

3.1 Households

The number of households is denoted by N_t. Population grows at rate n_t:

$$N_{t+1} = (1 + n_t)N_t. \tag{3.1}$$

Households maximize inter-temporal utility

$$U_0 = \sum_{t=0}^{\infty} \beta^t \left[u(c_t, 1 - l_t) + \nu(g_t)\right], \tag{3.2}$$

where c_t, $1 - l_t$, and g_t denote consumption, leisure, and government consumption in period t. The household is endowed with one unit of time in each period t and supplies labor at the amount of l_t. The utility from government consumption g_t is additively separate so that g_t does not affect household optimization with respect to consumption and labor.

Instantaneous utility $u(.,.)$ is specified as a Cobb-Douglas function:

$$u(c, l) = \frac{1}{1 - \eta} \left(c^{\theta} l^{1-\theta}\right)^{1-\eta}, \tag{3.3}$$

where θ and $1 - \theta$ denote the weights of consumption and leisure in utility.

Households own two kind of assets, traditional capital k_t and automation capital p_t (in per capita terms), which both depreciate at the common rate δ:

$$(1 + n_t)k_{t+1} = (1 - \delta)k_t + i_t^k, \tag{3.4a}$$

$$(1 + n_t)p_{t+1} = (1 - \delta)p_t + i_t^p. \tag{3.4b}$$

Investment on physical and automation capital is denoted by i_t^k and i_t^p, respectively. Following Hayashi (1982), we introduce quadratic adjustment costs for both kinds of capital k_t and p_t with $x \in \{k, p\}$ symmetrically:

$$\Phi^x(i_t^x, x_t) = i_t^x + a_1 \frac{(i_t^x)^2}{2(a_2 + x_t)}, \tag{3.5}$$

where, in addition to the standard specification of $\Phi(.)$ as, for example, in Heer and Scharrer (2018), a small constant a_2 is introduced to handle the case with zero automation capital, $p_t = 0$.

The household receives income from labor $w_t l_t$, which is taxed at the rate τ_t^l, and interest income on both physical capital, $r_t^k k_t$, and automation capital, $r_t^p p_t$, which are both taxed at rate τ_t^k, respectively. In addition, the household receives government transfers tr_t in period t. The household spends his income on consumption c_t, which is taxed at the constant rate τ^c, and on investment in both forms of capital. Accordingly, the household budget constraint is presented by:

$$(1+\tau^c)c_t + \Phi^k(i_t^k, k_t) + \Phi^p(i_t^p, p_t) = (1-\tau_t^l)w_t l_t + (1-\tau_t^k)r_t^k k_t + (1-\tau_t^k)r_t^p p_t + tr_t. \tag{3.6}$$

The first-order conditions of the household are presented by:

$$\lambda_t(1+\tau^c) = \theta c_t^{\theta(1-\eta)-1}(1-l_t)^{(1-\theta)(1-\eta)}, \tag{3.7a}$$

$$\lambda_t(1-\tau_t^l)w_t = (1-\theta)c_t^{\theta(1-\eta)}(1-l_t)^{(1-\theta)(1-\eta)-1}, \tag{3.7b}$$

$$q_t^k = \lambda_t \left(1 + a_1 \frac{i_t^k}{a_2 + k_t}\right), \tag{3.7c}$$

$$q_t^p = \lambda_t \left(1 + a_1 \frac{i_t^p}{a_2 + p_t}\right), \tag{3.7d}$$

$$(1+n_t)q_t^k = \beta \left\{ \lambda_{t+1} \left[(1-\tau_{t+1}^k) r_{t+1}^k - a_1 \frac{\left(i_{t+1}^k\right)^2}{2(a_2 + k_{t+1})^2} \right] + q_{t+1}^k (1-\delta) \right\}, \tag{3.7e}$$

$$(1+n_t)q_t^p = \beta \left\{ \lambda_{t+1} \left[(1-\tau_{t+1}^k) r_{t+1}^p - a_1 \frac{\left(i_{t+1}^p\right)^2}{2(a_2 + p_{t+1})^2} \right] + q_{t+1}^p (1-\delta) \right\}, \tag{3.7f}$$

where q_t and q_t^p denote Tobin's q on pyhsical and automation capital, respectively, and are equal to the Lagrange multipliers on the capital accumulation equations.[3]

[3]The first-order conditions are derived in the Appendix.

3.2 Production

We consider the production function following Steigum (2011) and Prettner (2017):[4]

$$Y_t = A\left[L_t + \kappa P_t\right]^{1-\alpha} K_t^{\alpha}. \tag{3.8}$$

with the aggregate labor supply $L_t \equiv N_t l_t$, aggregate automation capital $P_t \equiv N_t p_t$, and aggregate traditional capital $K_t \equiv N_t k_t$.

Firms maximize profits

$$\Pi_t = Y_t - w_t L_t - r_t^k K_t - r_t^p P_t,$$

implying the first-order conditions

$$w_t = (1-\alpha) A \left[\frac{K_t}{L_t + \kappa P_t}\right]^{\alpha}, \tag{3.9a}$$

$$r_t^p = (1-\alpha)\kappa A \left[\frac{K_t}{L_t + \kappa P_t}\right]^{\alpha}, \tag{3.9b}$$

$$r_t^k = \alpha A \left[\frac{L_t + \kappa P_t}{K_t}\right]^{1-\alpha}. \tag{3.9c}$$

In equilibrium, profits are zero due to the assumption of constant returns to scale and competitive goods and factor markets.

3.3 Government

The government spends its tax revenue on government consumption and transfers:

$$N_t g_t + N_t tr_t = N_t tax_t, \tag{3.10}$$

with per capita taxes tax_t presented by

$$tax_t = \tau^c c_t + \tau_t^l w_t l_t + \tau_t^k r_t^p p_t + \tau_t^k r_t^k k_t. \tag{3.11}$$

[4] In addition to these authors, we introduce a productivity factor κ of automation capital P_t. For the specification with $\kappa = 1.0$, we find that the wage share drops to 10% in the first year when the households build up automation capital $P_t > 0$.

3.4 Equilibrium conditions

In equilibrium, the resource constraint (in per capita terms) of the economy is presented by:

$$y_t = c_t + g_t + \Phi^k(i_t^k, k_t) + \Phi^p(i_t^p, p_t). \tag{3.12}$$

Accordingly, total production is spent on private and public consumption and investment in both capital goods.

3.5 Initial steady state without automation capital

In the initial steady state corresponding to the time period $t = 0$ (which we set to 1965), the stock of automation capital is set equal to zero, $p_0 = 0$. We also assume that the population growth rate n, the tax rates τ^c, τ^l, and τ^k are all constant and equal to the values prevailing in 1965. The economic growth rate of the economy is also equal to zero.

Therefore, the initial steady state is represented by the following 8 equations in the 8 endogenous variables k, l, i^k, c, w, r^k, q^k, and λ:

$$Al^{1-\alpha}k^{\alpha} = c + g + \Phi^k(i^k, k), \tag{3.13a}$$

$$w = (1-\alpha)A\left[\frac{k}{l}\right]^{\alpha}, \tag{3.13b}$$

$$r^k = \alpha A\left[\frac{l}{k}\right]^{1-\alpha}, \tag{3.13c}$$

$$\frac{1-\tau^l}{1+\tau^c}w = \frac{1-\theta}{\theta}\frac{c}{1-l}, \tag{3.13d}$$

$$i^k = (n+\delta)k, \tag{3.13e}$$

$$q^k = \lambda\left(1 + a_1\frac{i}{a_2+k}\right), \tag{3.13f}$$

$$\left(\frac{1+n}{\beta} - (1-\delta)\right)\frac{q^k}{\lambda} = (1-\tau^k)r^k - a_1\frac{(i^k)^2}{2(a_2+k)^2}, \tag{3.13g}$$

$$(1+\tau^c)\lambda = \theta c^{\theta(1-\eta)-1}(1-l)^{(1-\theta)(1-\eta)}. \tag{3.13h}$$

3.6 Automation Capital and Asymptotic Steady State

Assume that population growth is constant in the long run, $n_t = n$. In addition, the tax rates are constant, too, with $\tau_t^k = \tau^k$ and $\tau_t^l = \tau^l$. The factor prices are presented

by

$$w_t = (1-\alpha)A\left[\frac{\tilde{k}_t}{1+\kappa\tilde{p}_t}\right]^{\alpha}, \tag{3.14a}$$

$$r_t^p = (1-\alpha)\kappa A\left[\frac{\tilde{k}_t}{1+\kappa\tilde{p}_t}\right]^{\alpha}, \tag{3.14b}$$

$$r_t^k = \alpha A\left[\frac{1+\kappa\tilde{p}_t}{\tilde{k}_t}\right]^{1-\alpha}. \tag{3.14c}$$

with $\tilde{k}_t \equiv \frac{k_t}{l_t}$ and $\tilde{p}_t \equiv \frac{p_t}{l_t}$.

It is straightforward to show that, asymptotically, the two returns on the assets must be equal, $r^k = r^p$,[5] implying

$$(1-\alpha)\kappa A\left[\frac{\tilde{k}_t}{1+\kappa\tilde{p}_t}\right]^{\alpha} = \alpha A\left[\frac{1+\kappa\tilde{p}_t}{\tilde{k}_t}\right]^{1-\alpha},$$

or

$$\tilde{k}_t = \frac{\alpha}{1-\alpha}\frac{1+\kappa\tilde{p}_t}{\kappa}. \tag{3.15}$$

Therefore, the asymptotic real interest rate on capital is equal to

$$r^k = r^p = \alpha^{\alpha}(1-\alpha)^{1-\alpha}\kappa^{1-\alpha}A. \tag{3.16}$$

In the long-run equilibrium with positive growth, $\gamma > 0$, the coefficient $i_t^k/(a_2+k_t)$ converges to $i_t^k/k_t = i^k/k$. From the equation for the capital stock accumulation, (3.4a), we get

$$\frac{i^k}{k} = (1+n)(1+\gamma) - 1 + \delta, \tag{3.17}$$

with $k_{t+1}/k_t \equiv 1+\gamma$.

The first-order condition with respect to i_t^k, (3.7c), implies

$$\frac{q}{\lambda} = 1 + a_1\frac{i^k}{k}. \tag{3.18}$$

From first-order condition with respect to k_{t+1}, (3.7e), we get

$$\frac{1+n}{\beta}\frac{q}{\lambda} = (1+\gamma)^{\theta(1-\eta)-1}\left[(1-\tau^k)r^k - \frac{a_1}{2}\left(\frac{i^k}{k}\right)^2 + \frac{q}{\lambda}(1-\delta)\right]. \tag{3.19}$$

In this derivation, we have used the fact that leisure converges to one in the long run so that the growth factor of the Lagrange multiplier λ_t converges to $\lambda_{t+1}/\lambda_t = (1+\gamma)^{\theta(1-\eta)-1}$.

[5] See the Appendix.

3.7 Calibration

We calibrate the model as follows. Assume that we have initially picked two values for A and κ. Next, we assume that the economy is in steady state without automation capital in the year 1965. Therefore, the income tax rates τ^l and τ^k as well as the population growth rate are equal to the values prevailing in 1965.

In the initial steady state, capital k_t is constant implying that i_t^k is constant, too:

$$(n+\delta)k = i^k.$$

The first-order conditions with respect to i_t^k and k_{t+1} in steady state without automation capital imply the following two equations

$$\frac{q^k}{\lambda} = 1 + a_1 \frac{n+\delta}{a_2+k} k, \tag{3.20a}$$

$$\frac{1+n}{\beta}\frac{q^k}{\lambda} = (1-\tau^k)\alpha A k^{\alpha-1} l^{1-\alpha} - a_1 \frac{(n+\delta)^2 k^2}{2(a_2+k)^2} + \frac{q^k}{\lambda}(1-\delta), \tag{3.20b}$$

which can be solved for q^k/λ and k for given calibration $\bar{l} = 0.3$ and standard parameterization for the discount factor $\beta = 0.96$ and the depreciation rate $\delta = 7\%$ as, for example, applied by Trabandt and Uhlig (2011). The parameters of the adjustment cost function, $a_1 = 12$, is taken from Heer and Schubert (2012). The small constant a_2 which allows the function evaluation for zero automation capital is set equal to 0.1. In order to have an initial wage share of 62%, we set $\alpha = 38\%$.

With the help of k and l, we can compute y, i^k, w, and r^k for given value of A. Following Trabandt and Uhlig (2011), we set the government consumption share in GDP equal to 18%. We keep the assumption that the government-GDP share remains constant during the transition and in the final steady state. The resource constraint, $c = y - \Phi^k(i^k, k) - g$ implies steady-state consumption c. Given the first-order conditions with respect to labor and consumption, we can calibrate $\theta = 0.3938$. Finally, we can derive the steady-state value of λ and, hence, q using the value of 1/2 for the intertemporal elasticity of substitution, $1/\eta$. To compute the asymptotic steady state growth rate γ, we solve (3.16)-(3.19), for r^k, q/λ, γ and i^k/k. We iterate this procedure by trying different values of A and κ until the long-run growth rate is equal to 2.0% and the decline in the wage share is approximately equal to that prevailing during 1965-2010. Therefore, we choose the values of $A = 24.9$ and $\kappa = 0.0073$.

In the computation of the transition, we assume that, after the final period of transition, k, p, i^k, i^p, and c all grow at the common rate γ.

4 Results from the Model's Transition Analysis

We first describe the benchmark economy with automation capital and the dynamic tax rates prevailing during 1965-2010. We show that the model explains the drop in the wage share from 62% to 57% that took place between 1965 and 2015. In the second part of this section, we perform an analysis where we study the individual effects of the two tax rates and the population growth rate. We demonstrate that the wage share would have been several percentage points higher if the tax rates had not changed since 1965.

4.1 Benchmark

Figs. 4.1 and 4.2 illustrate the transition dynamics in our model economy during 1965-2030. The transition is computed assuming the following:

1. The economy is in steady state prior to 1965. The wage share amounts to 62.0% during this period. The steady state values of the tax rates (17.9% and 47.3% on labor and capital income) and the population growth rate (1.7%) are equal to the values prevailing in 1965.

2. In 1965, the transition starts. The initial capital stock is given by the initial steady-state capital stock. Households build up savings and supply labor according to their Euler equations for both forms of capital and their optimal first-order condition with respect to labor supply.[6]

3. The tax rates and the population growth rates during 1965-2015 are identical to their empirical counterparts.

4. In 1965, the households start to invest in automation capital p_t as well.

5. After 2010, the tax rates on both labor (28.3%) and capital income (37.0%) as well as the population growth rate (0.9%) remain constant.

As is evident from the first row of Fig. 4.1, both traditional capital k_t (upper left panel) and automation capital p_t (upper right panel) grow over time. As we argue in the Appendix, our model behaves like the 'Ak'-model of Romer (1986) asymptotically

[6] In the Appendix, we describe the computation of the transitional dynamics in more detail.

Figure 4.1: Transition Dynamics in the Model with Automation Capital and Adjustment Costs, Part I

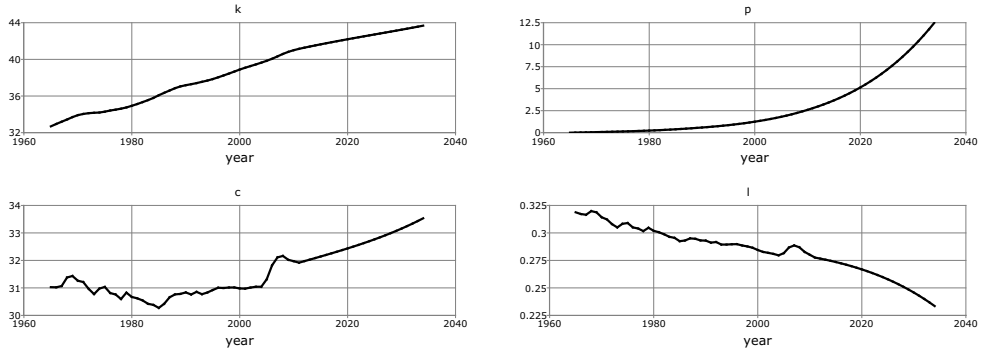
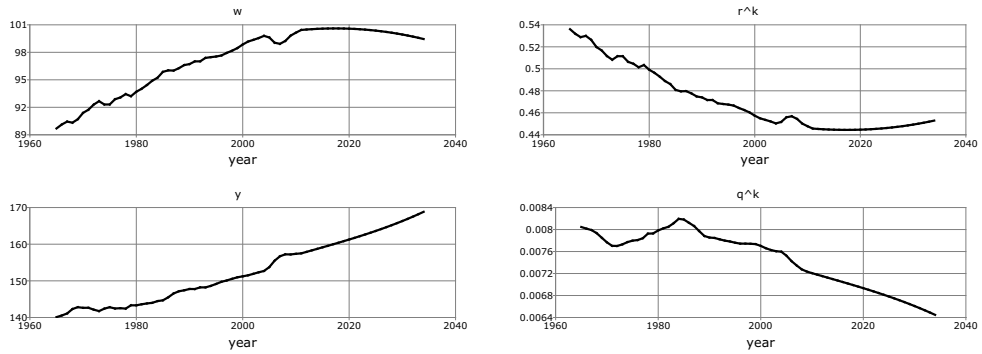


Figure 4.2: Transition Dynamics in the Model with Automation Capital and Adjustment Costs, Part II



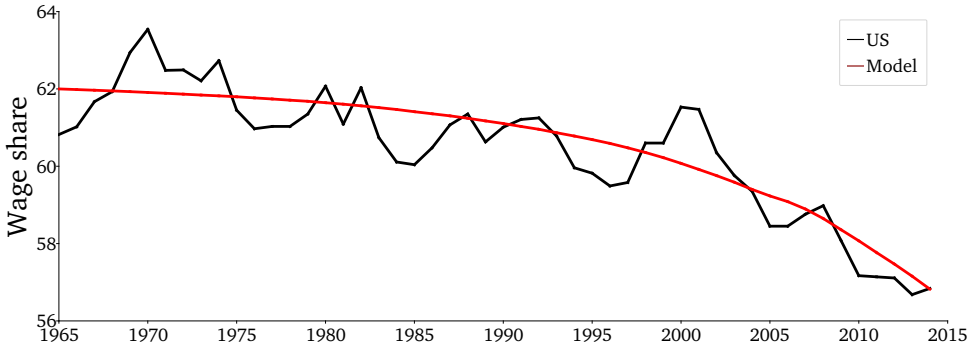
and converges to an economy that is displaying a growth rate of 2.0%. In the first year of transition, labor increases from 0.300 in the old steady state to 0.319 in 1966. The household substitutes labor intertemporally and increases its labor supply during the initial years of the transition when labor income taxes are low. Between 1966 and 2015, labor supply drops from 0.319 to 0.274 (lower right panel).[7] This is the effect of two channels which are simultaneously at work in our model. First, labor is replaced by automation capital in production. Second, the incentives to supply labor are reduced because of the increase in labor income taxes τ_t^l from 17.9% to 28.3% in the US during 1965-2010. The dominant factor on the labor supply is the change in tax rates, while

[7]Eventually, labor supply is completely replaced by automation capital and declines to zero. In our benchmark case, the lower limit $l_t = 0$ is reached by the year 2069.

the substitution effect of labor by automation capital accounts for approximately one third. For example, with constant labor income tax rates $\tau^l = 17.9\%$, the labor supply in 2015 would have only dropped to 0.302. Consumption (lower left panel) falls between 1970 and 1985 because of two reasons. First, labor supply decreases while labor income taxes increase so that the net wage income declines. Second, the household increases savings as the capital income taxes falls. In the long run, consumption grows at the endogenous growth rate.

The dynamics of the factor prices w_t and r_t^k during 1965-2030 are displayed in the upper row of Fig. 4.2. Since traditional capital k_t, on the one hand, increases during 1965-2005, while the initial increase in the automation capital, on the other hand, is relatively moderate, the wage rate w_t increases over time. Consequently, the interest rate on traditional capital, r_t^k, also falls. Production y_t (lower left panel) increases over time and the growth rate converges to its asymptotic rate, while Tobin's q on physical capital, q_t^k, declines over time. As is evident from the first-order condition (3.7c), the dynamics of Tobin's q mimic that of marginal utility from consumption, λ_t. As consumption c_t continues to increase, λ_t falls.

Figure 4.3: Wage Share Dynamics



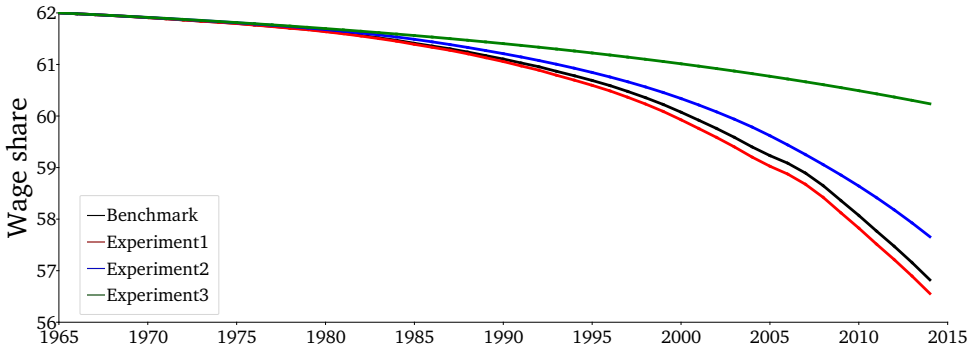
The dynamics of the wage share, $w_t l_t / y_t$, are illustrated in Fig. 4.3. Our model with automation capital (red line) is able to replicate the downward trend of the wage share in the US economy (black line) during 1965-2015. In fact, the model is able to generate a drop in the wages by five percentage points between 1965 and 2015 as observed empirically. This effect is explained by the substitution of labor for automation capital. The income share accruing to these two production factors is constant and amounts to $1 - \alpha = 62\%$, while the relative income share of automation capital is increasing over

time at the expense of the residual share for labor.

4.2 Experiments

In this section, we conduct three counterfactual experiments. In the first experiment, we keep the population growth rate constant during the transition setting it equal to the value in 1965. The effect on the wage share is illustrated by the red line in Fig. 4.4. In comparison to the benchmark case (black line), the wage share decreases. Therefore, if population growth had not declined during the last 55 years, the wage share would be lower by approximately 0.26 percentage points in 2015. This seems small, but keep in mind that the population growth rate has only declined from 1.7% to 0.9% in the US economy during 1965-2015. The fall in the population growth rate is associated with an increase in the endogenous growth rate as the capital deepening intensifies.

Figure 4.4: Wage Share, Income Taxes, and Population Growth



In our second experiment (blue line), we also set the labor income tax rate constant and equal to its initial steady state value of 17.9%. As a consequence, the incentives to supply labor are higher than in the benchmark economy and, therefore, less labor is substituted by automation capital. The wage share in 2015 would increase by 1.01 percentage points so that the wage share only would have dropped to 57.7%.

Finally, in our third experiment (green line), we set all three exogenous variables, the two income tax rates on labor and capital income, τ_t^l and τ_t^k, as well as the population growth rate n_t constant and equal to their respective values in 1965. As a consequence of the lower capital income tax rate, the incentives to build up capital increase and, therefore, the ratio of traditional capital to augmented labor, $k_t/(l_t+\kappa p_t)$, increases. As

a consequence, the marginal product of labor is higher than in the benchmark case and the incentives to supply labor rise. Therefore, less automation capital p_t is substituted for labor l_t. In addition, the growth rate of the economy is smaller so that the speed at which labor is replaced by automation slows down. As a consequence, the wage share increases by an additional 2.5 percentage points in comparison to the experiment 2. Combining all three effects — the changes in the two tax rates and in the population growth rate — account for the larger part of the decline in the wage share in the US. In fact, for the values of these variables prevailing in 1965, the US wage share would have only dropped from 62% in 1965 to 60.2% rather than 57% in 2015.

Figure 4.5: Transition Dynamics in Experiments 1-3

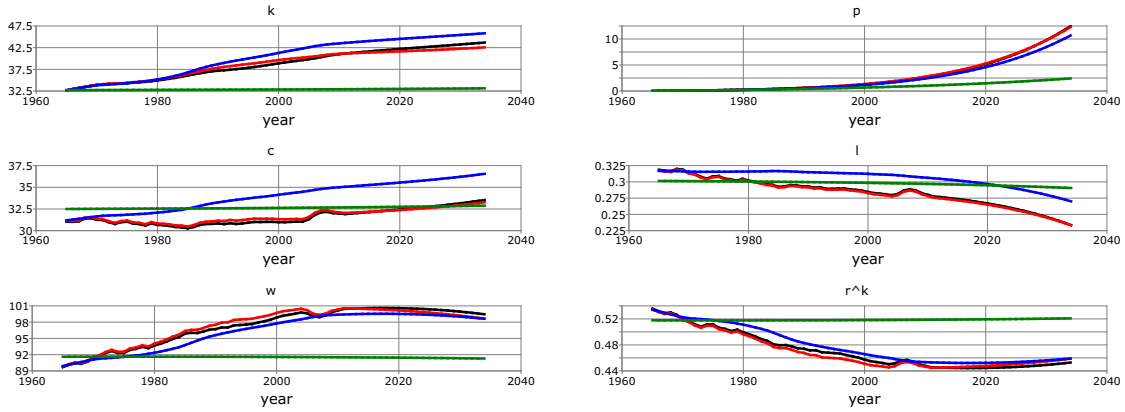


Fig. 4.5 illustrates the dynamics of the traditional capital stock k_t, automation capital p_t, consumption c_t, labor l_t, the wage rate w_t, and the interest rate on traditional capital, r_t^k, (from upper left to lower right panel) in these experiments and compares them with those in the benchmark economy. The dynamics confirm our arguments above. The lower population growth rate (red line) has little effects on the dynamics of capital and labor in comparison with the benchmark (black line) so that the two lines almost coincide. The lower labor income tax rate (blue line) increases both labor and physical capital while it decreases automation capital. It increases labor because of the substitution effect in labor supply. It also increases capital because the marginal product of capital rises. Finally, the higher capital income tax rate under experiment 3 (green line) decreases the accumulation of traditional capital stock k_t, and, hence, the interest rate would be much higher, while wages decline.

Notice further that the two tax rates on labor and capital income affect the asymp-

totic growth rate (not illustrated). In experiment 2, the growth rate declines by 0.1 percentage point, from 2.0% to 1.9%. In experiment 3, the decline is even much more pronounced and the growth rate even drops to 0.86%. Therefore, our model is also a model of endogenous growth where, as in the 'Ak'-model of Romer (1986), the capital income tax rate has significant negative growth effects.

5 Conclusion

We find that tax rates on both labor and capital income as well as population growth have a significant effect on the wage share. For the US economy, we estimate the effect to amount to approximately 3 percentage points over the period 1965-2015. We demonstrate that this results holds both empirically and in a neoclassical growth model with automation capital. Our growth model also predicts a continuing fall in the wage share over the coming decades. However, we would like to interpret this latter finding in a cautious way because 4we neglect other aggravating factors like artificial intelligence (AI). In our model, the productivity of the automation capital does not increase over time. In future research, we plan to endogenize the investment in AI and its effects on the functional income distribution.

References

Acemoglu, D. and P. Restrepo (2019). Robots and Jobs: Evidence from US Labor Markets. *Journal of Political Economy (forthcoming)*.

Bernanke, B. S., J. Boivin, and P. Eliasz (2005). Measuring the effects of monetary policy: A factor-augmented vector autoregressive (favar) approach. *Quarterly Journal of Economics 120*, 387 – 422.

Fernald, J. G. (2014). A quarterly, utilization-adjusted series on total factor productivity. *FRBSF Working Paper 2012-19*.

Glover, A. and J. Short (2020, January). Can Capital Deepening Explain the Global Decline in Labor's Share? *Review of Economic Dynamics 35*, 35–53.

Gomme, P., B. Rvaikumar, and P. Rupert (2011). The Return to Capital and the Business Cycle. *Review of Economic Dynamics 14*(2), 262 – 278.

Graetz, G. and G. Michaels (2018). Robots at work. *Review of Economics and Statistics 100*(5), 753 – 768.

Hayashi, F. (1982). Tobin's marginal q and average q: A neoclassical interpretation. *Econometrica 50*(1), 213 – 224.

Heer, B. and A. Irmen (2014). Population, pensions and endogenous economic growth. *Journal of Economic Dynamics and Control 46*, 50–72.

Heer, B. and C. Scharrer (2018). The age-specific burdens of short-run fluctuations in government spending. *Journal of Economic Dynamics and Control 90*, 45–75.

Heer, B. and S. F. Schubert (2012). Unemployment and debt dynamics in a highly indebted small open economy. *Journal of International Money and Finance 31*, 1392 – 1413.

Piketty, T. (2014). *Capital in the Twenty-First Century*. Cambridge, Massachusetts: Harvard University Press.

Prettner, K. (2017). The implications of automation for economy growth and the labor share. *Macroeconomic Dynamics 12*, 1–8.

Romer, P. M. (1986). Increasing returns and long-run growth. *Journal of Political Economy 94*, 1002 – 1037.

Schwert, G. W. (1989). Tests for unit roots: A monte carlo investigation. *Journal of Business and Economic Statistics 3*, 147 – 159.

Shumway, R. H. and D. S. Stoffer (2008). An approach to time series smoothing and forecasting using the em algorithm. *Journal of Time Series Analysis 3*, 253 – 264.

Steigum, E. (2011). Robots and growth. In *Frontiers of Economics and Globalization: Economic Growth and Development*, pp. 543–557. Emerald Group.

Stock, J. and M. W. Watson (2016). Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics. In J. B. Taylor and H. Uhlig (Eds.), *Handbook of the Macroeconomics*, pp. 415–525. Amsterdam: Elsevier.

Tani, A. (1989). International comparisons of industrial robot penetration. *Technological Forecasting and Social Change 34*, 191 – 210.

Trabandt, M. and H. Uhlig (2011). The laffer curve revisited. *Journal of Monetary Economics 58*, 305 – 327.

6 Appendix

In the Appendix, we describe some methodological and mathematical details along with summary statistics of time series and time series tests.

First, accompanying the empirical part of the paper, we provide detail on the used nowcasting technique based on the procedure proposed Shumway and Stoffer (2008) relying on the Kalman filter in combination with the expectation maximization (EM) algorithm. It is used to generate a quarterly nowcast of the (bi-)annual, shortfall plagued automation capital series based on robots per 1,000 workers figures by Tani (1989), Acemoglu and Restrepo (2019), the IFR, and the CPS. Proceedingly, detailed results of the Johansen testing procedure (adhering to the Pantula Principle) as a necessary prerequisite for our VECM estimation and analysis is given.

Secondly, accompanying the theoretical part of the paper, we derive some properties of the model with automation capital. In addition, we present some details on the computation of the transition dynamics. In essence, we describe the solution of a large-scale system of difference equations in four endogenous variables over a time horizon of 120 periods.

A.1 Nowcasting (bi-)annual automation capital to obtain a quarterly series

A.1.1 Time series

To-be-nowcasted time series. The series we seek to nowcast –or to generate (pseudo-)quarterly data for– is the use of industrial robots per 1,000 workers in the US as provided in Acemoglu and Restrepo (2019). The original source is the IFR. This series is of annual frequency and starts in 1993. It ends in 2014. However, due to data limitations with regard to other series both of the nowcasting procedure and of our VECM analysis (discussed in the empirical part of the paper) we end it in 2008. We merge these data with observations that we construct in the following way. Tani (1989) in his Tab. 1 (col. 3, p. 193) provides data for the industrial robot population in the US for the years 1974, 1976, 1978, and in annual frequency from 1980 to 1985. As Tani (1989) in his Tab. 3 standardizes these data to workers in the manufacturing sector only, we refrain from using his standardized series, but divide the non-standardized robots figures through the respective annual averages of the US full time employees,

SA (in 1K), data that we obtain from the CPS. The result is a series of mixed bi-annual, annual frequency with missing values for 1986 to 1992. Hence, covered years are 1974, 1976, 1978, 1980-1985, 1993-2008. Graphically this series is made of the dots shown in Fig. 1.4. As this series represents a stock variable (robots per 1,000 employees) for particular years, we might interpret it as end-of-year or q4-values.

General strategy and information set series. Given the (quasi-)q4-data of robots per 1,000 workers, the procedure runs in two main steps. In the first step, the missing (quasi-)q4-, or annual, values are generated using information from other use of automation capital related variables, for which we have data over the entire period and, at best, at a quarterly observation frequency. In a second step, for the obtained complete annual frequency series, running from 1974 to 2008, an analogue nowcasting approach is followed to generate a (pseudo-)quarterly series. Our baseline information set essentially uses variables from the Fernald (2014) database in contemporaneous and first lag expression that recently have been shown by Graetz and Michaels (2018) to be profoundly and significantly associated with robots input: hours worked, labor productivity, different estimates of labor quality (i.e. labor composition), total factor productivity (TFP), and utilitization-adjusted TFP. Additionally, we also consider the US tax rate on labor income provided by Gomme et al. (2011) as firms adopt robots mainly for saving on labor costs (besides ensuring uniform quality). Generally, labor costs depend on labor productivity and taxation.

A.1.2 Method

Starting point of the procedure is the notion of a general state space model for an n-dimensional time series $\boldsymbol{y}_t$ consisting of a measurement equation that relates the observed data to an m-dimensional state vector $\boldsymbol{\alpha}_t$. The generation of the state vector $\boldsymbol{\alpha}_t$ from the past state $\boldsymbol{\alpha}_{t-1}$, for $t = 1, ..., T$, is determined by the state equation. The measurement equation has the form

$$\boldsymbol{y}_t = \boldsymbol{Z}_t\boldsymbol{\alpha}_t + \boldsymbol{d}_t + \boldsymbol{u}_t, \quad t = 1, ..., T. \tag{A.1.1}$$

In (A.1.1), $\boldsymbol{Z}_t$ is an $n \times m$ matrix called measurement or observation matrix, $\boldsymbol{d}_t$ is an $n \times 1$ vector and $\boldsymbol{u}_t \sim$ iid $N(\boldsymbol{0}, \boldsymbol{H}_t)$ is an error vector. The state equation is given by

$$\boldsymbol{\alpha}_t = \boldsymbol{T}_t\boldsymbol{\alpha}_{t-1} + \boldsymbol{c}_t + \boldsymbol{R}_t\boldsymbol{\nu}_t, \quad t = 1, ..., T. \tag{A.1.2}$$

In (A.1.2), $\boldsymbol{T}_t$ is an $m \times m$ matrix called transition matrix, $\boldsymbol{c}_t$ is an $m \times 1$ vector, $\boldsymbol{R}_t$ is an $m \times g$ matrix and $\boldsymbol{\nu}_t \sim$ iid $N(\boldsymbol{0}, \boldsymbol{Q}_t)$ is a $g \times 1$ error vector. The matrices $\boldsymbol{Z}_t$, $\boldsymbol{d}_t$, $\boldsymbol{H}_t$, $\boldsymbol{T}_t$, $\boldsymbol{c}_t$, $\boldsymbol{R}_t$ and $\boldsymbol{Q}_t$ are referred to as system matrices. Usually, it is assumed that the errors of the measurement and the transition equation are uncorrelated, i.e.

$$E[\boldsymbol{u}_t \boldsymbol{\nu}_t'] = \boldsymbol{0} \quad \forall s, t = 1, ..., T.$$

Furthermore, it is assumed that the initial state is given by a normal vector

$$\boldsymbol{\alpha}_0 \sim N(\boldsymbol{a}_0, \boldsymbol{P}_0); \quad E[\boldsymbol{u}_t \boldsymbol{a}_0'] = \boldsymbol{0}, \; E[\boldsymbol{\nu}_t \boldsymbol{a}_0'] = 0, \quad t = 1, ..., T.$$

In our application of a state-space model, as defined by (A.1.1) and (A.1.2), we seek to generate estimators for the underlying unobserved signal $\boldsymbol{\alpha}_t$ given the data $\boldsymbol{y}_s$, for $s = 1, ..., S$. Whenever $s = t$ this problem is called filtering, while we speak of smoothing if $s > t$ and forecasting in case $s < t$. The problem of finding such estimators is solved by the Kalman Filter (KF), Kalman Smoother (KS) and forecasting recursions, respectively. The KF is a set of recursion equations (prediction equations and updating equations) that determine the optimal estimates for the state vector $\boldsymbol{\alpha}_t$ given the information available at t (henceforth, I_t). The following definitions are used

$$\boldsymbol{a}_t \coloneqq E[\boldsymbol{\alpha}_t | I_t] \tag{A.1.3}$$

and

$$\boldsymbol{P}_t \coloneqq E[(\boldsymbol{\alpha}_t - \boldsymbol{a}_t)(\boldsymbol{\alpha}_t - \boldsymbol{a}_t)' | I_t]. \tag{A.1.4}$$

That is, $\boldsymbol{a}_t$ is the optimal estimator of $\boldsymbol{\alpha}_t$ based on I_t and $\boldsymbol{P}_t$ is the mean square error (MSE) matrix of $\boldsymbol{a}_t$.

Prediction equations. Given $\boldsymbol{a}_{t-1}$ and $\boldsymbol{P}_{t-1}$,

$$\begin{aligned} \boldsymbol{a}_{t|t-1} &= E[\boldsymbol{\alpha}_t | I_{t-1}] \\ &= \boldsymbol{T}_t \boldsymbol{a}_{t-1} + \boldsymbol{c}_t \end{aligned} \tag{A.1.5}$$

$$\begin{aligned} \boldsymbol{P}_{t|t-1} &= E[(\boldsymbol{\alpha}_t - \boldsymbol{a}_{t-1})(\boldsymbol{\alpha}_t - \boldsymbol{a}_{t-1})' | I_{t-1}] \\ &= \boldsymbol{T}_t \boldsymbol{P}_{t-1} \boldsymbol{T}_t' + \boldsymbol{R}_t \boldsymbol{Q}_t \boldsymbol{R}_t. \end{aligned} \tag{A.1.6}$$

And the optimal predictor of $\boldsymbol{y}_t$ is obtained from

$$\begin{aligned} \boldsymbol{y}_{t|t-1} &= \boldsymbol{Z}_t \boldsymbol{a}_{t|t-1} + \boldsymbol{d}_t \\ &= \boldsymbol{Z}_t \boldsymbol{T}_t \boldsymbol{a}_{t-1} + \boldsymbol{Z}_t \boldsymbol{c}_t + \boldsymbol{d}_t \\ &= \boldsymbol{Z}_t (\boldsymbol{T}_t \boldsymbol{a}_{t-1} + \boldsymbol{c}_t) + \boldsymbol{d}_t. \end{aligned} \tag{A.1.7}$$

The corresponding prediction error and its MSE matrix are

$$\begin{aligned}\boldsymbol{e}_t &= \boldsymbol{y}_t - \boldsymbol{y}_{t|t-1} \\ &= \boldsymbol{y}_t - \boldsymbol{Z}_t\boldsymbol{a}_{t|t-1} - \boldsymbol{d}_t \\ &= \boldsymbol{Z}_t\boldsymbol{\alpha}_t + \boldsymbol{d}_t + \boldsymbol{u}_t - \boldsymbol{Z}_t\boldsymbol{a}_{t|t-1} - \boldsymbol{d}_t \\ &= \boldsymbol{Z}_t(\boldsymbol{\alpha}_t - \boldsymbol{a}_{t|t-1}) + \boldsymbol{u}_t \end{aligned} \tag{A.1.8}$$

and

$$E[\boldsymbol{e}_t\boldsymbol{e}_t'] \coloneqq \boldsymbol{F}_t = \boldsymbol{Z}_t\boldsymbol{P}_{t|t-1}\boldsymbol{Z}_t' + \boldsymbol{H}_t. \tag{A.1.9}$$

Updating equations. The moment $\boldsymbol{y}_t$ is observed the optimal predictor and its MSE matrix are updated according to

$$\begin{aligned}\boldsymbol{a}_t &= \boldsymbol{a}_{t|t-1} + \boldsymbol{P}_{t|t-1}\boldsymbol{Z}_t'\boldsymbol{F}_t^{-1}(\boldsymbol{y}_t - \boldsymbol{y}_{t|t-1}) \\ &= \boldsymbol{a}_{t|t-1} + \boldsymbol{P}_{t|t-1}\boldsymbol{Z}_t'\boldsymbol{F}_t^{-1}(\boldsymbol{y}_t - \boldsymbol{Z}_t\boldsymbol{\alpha}_{t|t-1} - \boldsymbol{d}_t) \\ &= \boldsymbol{a}_{t|t-1} + \boldsymbol{P}_{t|t-1}\boldsymbol{Z}_t'\boldsymbol{F}_t^{-1}\boldsymbol{e}_t \end{aligned} \tag{A.1.10}$$

$$\boldsymbol{P}_t = \boldsymbol{P}_{t|t-1} - \underbrace{\boldsymbol{P}_{t|t-1}\boldsymbol{Z}_t\boldsymbol{F}_t^{-1}}_{\text{Kalman Gain}}\boldsymbol{Z}_t\boldsymbol{P}_{t|t-1}. \tag{A.1.11}$$

Filter derivation. The KF-derivation makes use of the following properties of a bivariate normal distribution. Given y, the distribution of x is normal with

$$E[x|y] = \boldsymbol{\mu}_{x|y} = \boldsymbol{\mu}_x + \boldsymbol{\Sigma}_{xy}\boldsymbol{\Sigma}_{yy}^{-1}(\boldsymbol{y} - \boldsymbol{\mu}_y) \tag{A.1.12}$$

$$Var(x|y) = \boldsymbol{\Sigma}_{xx} - \boldsymbol{\Sigma}_{xy}\boldsymbol{\Sigma}_{yy}^{-1}\boldsymbol{\Sigma}_{yx}. \tag{A.1.13}$$

For the state vector at $t = 1$,

$$\boldsymbol{\alpha}_1 = \boldsymbol{T}_1\boldsymbol{\alpha}_0 + \boldsymbol{c}_1 + \boldsymbol{R}_1\boldsymbol{\nu}_1,$$

with $\boldsymbol{\alpha}_0 \sim N(\boldsymbol{a}_0, \boldsymbol{P}_0)$, $\boldsymbol{\nu}_1 \sim N(\boldsymbol{0}, \boldsymbol{Q}_1)$ and $E[\boldsymbol{\alpha}_0\boldsymbol{\nu}_1'] = \boldsymbol{0}$. In a linear Gaussian state-space model the initial state vector is normally distributed with

$$\begin{aligned}\boldsymbol{a}_{1|0} &\coloneqq E[\boldsymbol{\alpha}_1] = \boldsymbol{T}_1\boldsymbol{a}_0 + \boldsymbol{c}_1 \\ \boldsymbol{P}_{1|0} &\coloneqq Var(\boldsymbol{\alpha}_1) = \boldsymbol{T}_1\boldsymbol{P}_{1|0}\boldsymbol{T}_1' + \boldsymbol{R}_1\boldsymbol{Q}_1\boldsymbol{R}_1'. \end{aligned} \tag{A.1.14}$$

From the measurement equation, it follows that

$$\boldsymbol{y}_1 = \boldsymbol{Z}_1\boldsymbol{\alpha}_1 + \boldsymbol{d}_1 + \boldsymbol{u}_1,$$

with $\boldsymbol{u}_1 \sim N(\boldsymbol{0}, \boldsymbol{H}_1)$ s.t.

$$\begin{aligned} \boldsymbol{y}_{1|0} &\coloneqq E[\boldsymbol{y}_1] = \boldsymbol{Z}_1\boldsymbol{a}_{1|0} + \boldsymbol{d}_1 \qquad (A.1.15)\\ Var(\boldsymbol{y}_1) &= E[(\boldsymbol{y}_1 - \boldsymbol{y}_{1|0})(\boldsymbol{y}_1 - \boldsymbol{y}_{1|0})'] \\ &= E[(\boldsymbol{Z}_1\{\boldsymbol{\alpha}_1 - \boldsymbol{a}_{1|0}\} + \boldsymbol{u}_1)(\boldsymbol{Z}_1\{\boldsymbol{\alpha}_1 - \boldsymbol{a}_{1|0}\} + \boldsymbol{u}_1)'] \\ &= \boldsymbol{Z}_1\boldsymbol{P}_{1|0}\boldsymbol{Z}_1' + \boldsymbol{H}_1. \end{aligned}$$

Equations (A.1.14) and (A.1.15) are the predictions equations for $\boldsymbol{\alpha}_1$ and $\boldsymbol{y}_1$ at $t = 0$.

In a next crucial step one has to find the distribution of $\boldsymbol{\alpha}_1$ conditional on $\boldsymbol{y}_1$ being observed (updating). For this purpose the joint normal distribution of $(\boldsymbol{\alpha}_1', \boldsymbol{y}_1')$ must be determined. In finding the joint normal distribution we use

$$\begin{aligned} \boldsymbol{\alpha}_1 &= \boldsymbol{a}_{1|0} + (\boldsymbol{\alpha}_1 - \boldsymbol{a}_{1|0}) \\ \boldsymbol{y}_1 &= \boldsymbol{y}_{1|0} + \boldsymbol{y}_1 - \boldsymbol{y}_{1|0} \\ &= \boldsymbol{Z}_1\boldsymbol{a}_{1|0} + \boldsymbol{d}_1 + \boldsymbol{Z}_1(\boldsymbol{\alpha}_1 - \boldsymbol{a}_{1|0}) + \boldsymbol{u}_1. \end{aligned}$$

Note that as

$$\begin{aligned} Cov(\boldsymbol{\alpha}_1, \boldsymbol{y}_1) &= E[(\boldsymbol{\alpha}_1 - \boldsymbol{a}_{1|0})(\boldsymbol{y}_1 - \boldsymbol{y}_{1|0})'] \\ &= E[(\boldsymbol{\alpha}_1 - \boldsymbol{a}_{1|0})(\boldsymbol{Z}_1\{\boldsymbol{\alpha}_1 - \boldsymbol{a}_{1|0}\} + \boldsymbol{u}_1)'] \\ &= E[(\boldsymbol{\alpha}_1 - \boldsymbol{a}_{1|0})(\{\boldsymbol{\alpha}_1 - \boldsymbol{a}_{1|0}\}\boldsymbol{Z}_1' + \boldsymbol{u}_1')] \\ &= E[(\boldsymbol{\alpha}_1 - \boldsymbol{a}_{1|0})(\boldsymbol{\alpha}_1 - \boldsymbol{a}_{1|0})\boldsymbol{Z}_1'] + E[(\boldsymbol{\alpha}_1 - \boldsymbol{a}_{1|0})\boldsymbol{u}_1'] \\ &= \boldsymbol{P}_{1|0}\boldsymbol{Z}_1', \end{aligned}$$

$$\begin{pmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{y}_1 \end{pmatrix} \sim N\left(\begin{pmatrix} \boldsymbol{a}_{1|0} \\ \boldsymbol{Z}_1\boldsymbol{a}_{1|0} + \boldsymbol{d}_1 \end{pmatrix}, \begin{pmatrix} \boldsymbol{P}_{1|0} & \boldsymbol{P}_{1|0}\boldsymbol{Z}_1' \\ \boldsymbol{Z}_1\boldsymbol{P}_{1|0} & \boldsymbol{Z}_1\boldsymbol{P}_{1|0}\boldsymbol{Z}_1' + \boldsymbol{H}_1 \end{pmatrix} \right).$$

In combination with (A.1.12) and (A.1.13), $(\boldsymbol{\alpha}_1|\boldsymbol{y}_1) \sim N(\boldsymbol{a}_1, \boldsymbol{P}_1)$ follows with

$$\begin{aligned} \boldsymbol{a}_1 &= \boldsymbol{a}_{1|0} + \boldsymbol{P}_{1|0}\boldsymbol{Z}_1'(\boldsymbol{Z}_1\boldsymbol{P}_{1|0}\boldsymbol{Z}_1' + \boldsymbol{H}_1)^{-1}(\boldsymbol{y}_1 - \boldsymbol{Z}_1\boldsymbol{a}_{1|0} - \boldsymbol{d}_1) \\ &= \boldsymbol{a}_{1|0} + \boldsymbol{P}_{1|0}\boldsymbol{Z}_1'\boldsymbol{F}_1^{-1}\boldsymbol{e}_1 \qquad (A.1.16)\\ \boldsymbol{P}_1 &= \boldsymbol{P}_{1|0} - \boldsymbol{P}_{1|0}\boldsymbol{Z}_1'(\boldsymbol{Z}_1\boldsymbol{P}_{1|0}\boldsymbol{Z}_1' + \boldsymbol{H}_1)^{-1}\boldsymbol{Z}_1\boldsymbol{P}_{1|0} \\ &= \boldsymbol{P}_{1|0} - \boldsymbol{P}_{1|0}\boldsymbol{Z}_1'\boldsymbol{F}_1^{-1}\boldsymbol{Z}_1\boldsymbol{P}_{1|0}. \qquad (A.1.17) \end{aligned}$$

Note (A.1.16) and (A.1.17) are the Kalman Filter updating equations for $t = 1$.

ML-estimation and EM algorithm. Let $\boldsymbol{\theta}$ denote the parameters of the state-space model. These parameters are embodied in the system matrices. The likelihood of the state-space model is calculated based on the prediction errors $\boldsymbol{e}_t$ with $t = 1, ..., T$. The prediction error decomposition of the (negative) log-likelihood looks as follows

$$-2 \ln L(\boldsymbol{\theta}|\boldsymbol{y}) = nT \ln(2\pi) + \sum_{t=1}^{T} \ln |\boldsymbol{F}_t(\boldsymbol{\theta})| + \sum_{t=1}^{T} \boldsymbol{e}_t'(\boldsymbol{\theta}) \boldsymbol{F}_t^{-1}(\boldsymbol{\theta}) \boldsymbol{e}_t(\boldsymbol{\theta}). \tag{A.1.18}$$

Shumway and Stoffer (2008) proposed a procedure based on the EM algorithm that is conceptually simpler and more efficient than alternative procedures such as the New-Raphson algorithm. The basic idea is that if all the states $\boldsymbol{\alpha}_T = \{\boldsymbol{\alpha}_t\}_{t=0}^T$ together with the observations $\boldsymbol{y}_T = \{\boldsymbol{y}_t\}_{t=1}^T$ could be observed, one could consider the entire data space $\{\boldsymbol{\alpha}_T, \boldsymbol{y}_T\}$. The complete data likelihood might, thus, be written as

$$\begin{aligned} -2 \ln L(\boldsymbol{\theta}|\boldsymbol{\alpha}, \boldsymbol{y}) = \ln |\boldsymbol{F}_0| + (\boldsymbol{\alpha}_0 - \boldsymbol{a}_0)' \boldsymbol{F}_0^{-1} (\boldsymbol{\alpha}_0 - \boldsymbol{a}_0) \\ + n \ln |\boldsymbol{Q}_t| + \sum_{t=1}^{T} (\boldsymbol{\alpha}_t - \boldsymbol{T}_t \boldsymbol{\alpha}_{t-1})' \boldsymbol{Q}_t^{-1} (\boldsymbol{\alpha}_t - \boldsymbol{T}_t \boldsymbol{\alpha}_{t-1}) \\ + n \ln |\boldsymbol{H}_t| + \sum_{t=1}^{T} (\boldsymbol{y}_t - \boldsymbol{Z}_t \boldsymbol{\alpha}_t)' \boldsymbol{H}_t^{-1} (\boldsymbol{y}_t - \boldsymbol{Z}_t \boldsymbol{\alpha}_t). \end{aligned} \tag{A.1.19}$$

Given the *complete* data without any missing values and mostly in the desired (quarterly) frequency, one could easily obtain the ML-estimates of $\boldsymbol{\theta}$. However, as this is not the case, we may find the ML-estimates based on the *incomplete* data with short-fall by successively maximizing the conditional expectation of the complete data likelihood. This is done in the following steps:

1. Find some initial values for parameters $\boldsymbol{\theta}^{(0)}$,
2. Calculate the incomplete data likelihood $-\ln L(\boldsymbol{\theta}^{(j-1)}|\boldsymbol{y})$; see equation (A.1.18),
3. At iteration $j = 1, 2, ...$ use the KF and KS to obtain smoothed values for $\boldsymbol{\alpha}_t^{(S)}$, $\boldsymbol{P}_t^{(S)}$ and $\boldsymbol{P}_{t|t-1}^{(S)}$ for $t = 1, ..., T$ based on the parameters $\boldsymbol{\theta}^{(j-1)}$. Use the smoothed

values to calculate the conditional expectation of the complete data likelihood

$$\begin{aligned}
Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(j-1)}) =& E\left\{-2 \ln L(\boldsymbol{\theta}|\boldsymbol{\alpha}, \boldsymbol{y})|\boldsymbol{y}_n, \boldsymbol{\theta}^{(j-1)}\right\} \\
&= \ln |\boldsymbol{F}_0| + \text{ tr } \left\{\boldsymbol{F}_0^{-1}\left[\boldsymbol{P}_0^{(S)} + \left(\boldsymbol{\alpha}_0^{(S)} - \boldsymbol{a}_0\right)\left(\boldsymbol{\alpha}_0^{(S)} - \boldsymbol{a}_0\right)'\right]\right\} \\
&\quad + n \ln |\boldsymbol{Q}_t| + \text{ tr } \left\{\boldsymbol{Q}^{-1}\left[S_{11} - S_{10}\boldsymbol{Z}_t' - \boldsymbol{Z}_t S_{10} + \boldsymbol{Z}_t S_{00}\boldsymbol{Z}_t'\right]\right\} \\
&\quad + n \ln \boldsymbol{H} \\
&\quad + \text{ tr } \left\{\boldsymbol{H}^{-1}\sum_{t=1}^{T}\left[\left(\boldsymbol{y}_t - \boldsymbol{Z}_t\boldsymbol{\alpha}_t^{(S)}\right)\left(\boldsymbol{y}_t - \boldsymbol{Z}_t\boldsymbol{\alpha}_t^{(S)}\right)' + \boldsymbol{Z}_t\boldsymbol{P}_t^{(S)}\boldsymbol{Z}_t'\right]\right\},
\end{aligned}$$

where

$$\begin{aligned}
S_{11} &= \sum_{t=1}^{T}\left(\boldsymbol{\alpha}_t^{(S)}\boldsymbol{\alpha}_t^{(S)'} + \boldsymbol{P}_t^{(S)}\right), \\
S_{10} &= \sum_{t=1}^{T}\left(\boldsymbol{\alpha}_t^{(S)}\boldsymbol{\alpha}_{t|t-1}^{(S)'} + \boldsymbol{P}_{t|t-1}^{(S)}\right) \text{ and} \\
S_{00} &= \sum_{t=1}^{T}\left(\boldsymbol{\alpha}_{t|t-1}^{(S)}\boldsymbol{\alpha}_{t|t-1}^{(S)'} + \boldsymbol{P}_{t|t-1}^{(S)}\right).
\end{aligned}$$

4. Update $\boldsymbol{\theta}_0$ according to

$$\begin{aligned}
\boldsymbol{T}_t^{(j)} &= S_{10}S_{00}^{-1}, \\
\boldsymbol{Q}_t^{(j)} &= n^{-1}\left(S_{11} - S_{10}S_{00}^{-1}S_{10}'\right) \text{ and} \\
\boldsymbol{H}_t^{(j)} &= n^{-1}\sum_{t=1}^{T}\left[\left(\boldsymbol{y}_t - \boldsymbol{Z}_t\boldsymbol{\alpha}_t^{(S)}\right)\left(\boldsymbol{y}_t - \boldsymbol{Z}_t\boldsymbol{\alpha}_t^{(S)}\right)' + \boldsymbol{Z}_t\boldsymbol{P}_t^{(S)}\boldsymbol{Z}_t'\right]
\end{aligned}$$

to obtain $\boldsymbol{\theta}^{(j)}$.

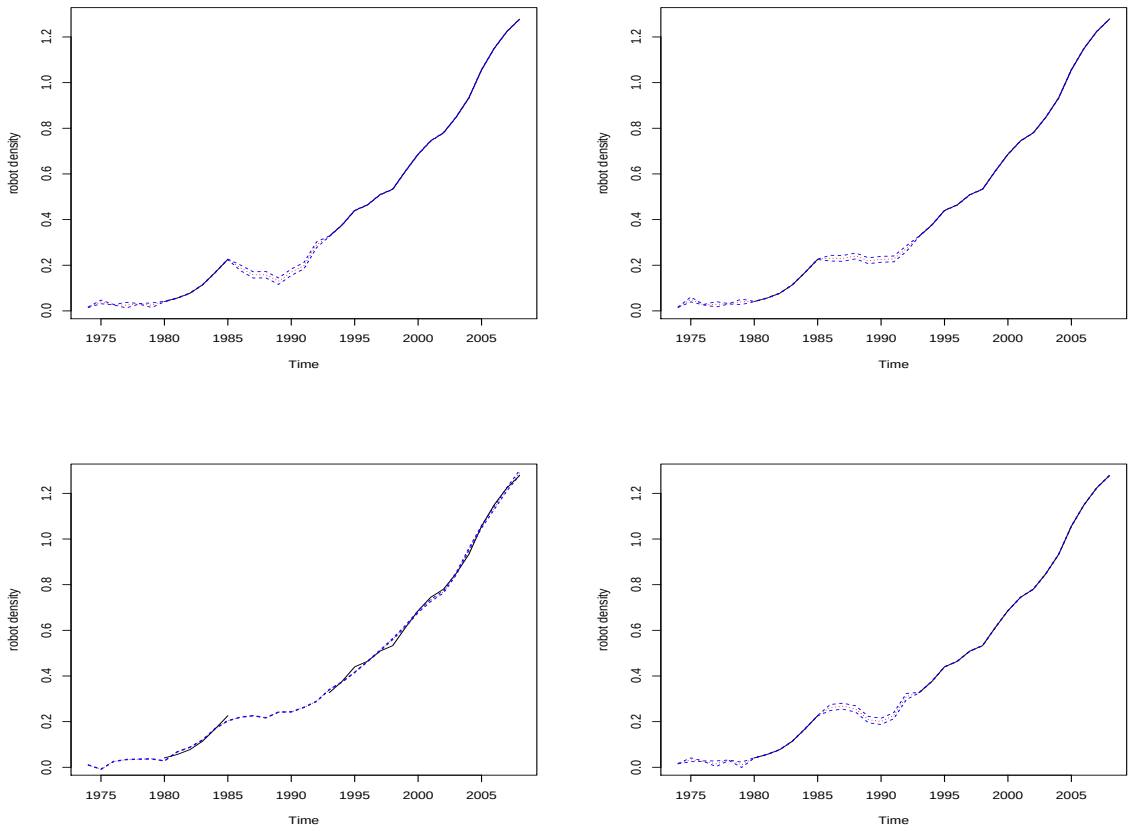
5. Repeat steps 2 to 4 until convergence is achieved (i.e.until parameters or likelihood values stabilize in the sense of differing from their predecessor values by some predetermined small amount κ only).

A.1.3 Application

Annual series nowcast. Having sketched the nowcasting procedure and indicators set, referred to as information set I_t above, we run the annual series (or q4-value) nowcast for four differently sized set of indicators: I_1 (index t dropped for notational ease) considers 16 series, i.e. in contemporaneous and first lag expression: hours worked,

labor productivity, actually used labor composition/quality, labor composition/quality as reported by the Bureau of Labor Statistics (BLS), TFP, utilitization-adjusted TFP, and the tax rate on labor income. The nex two considered sets, I_2 and I_3, are similar in size and nowcasting performance: Set I_3 comprises ten series by dropping the two labor composition/quality indicator series and the utilization-adjusted TFP series. Using just one (irrespectively which one) or both TFP series does not alter the nowcasted values. It merely changes fourth or higher decimal places. The same applies to the two different labor quality indicators. I_2 is of same size as I_3 but considers the actually used labor composition/quality indicator and leaves out the labor tax series. Information set I_4 compared to I_3 includes the labor tax indicator and drops labor productivity.

Figure A.1: Annual nowcast of robots per 1K workers for different information sets

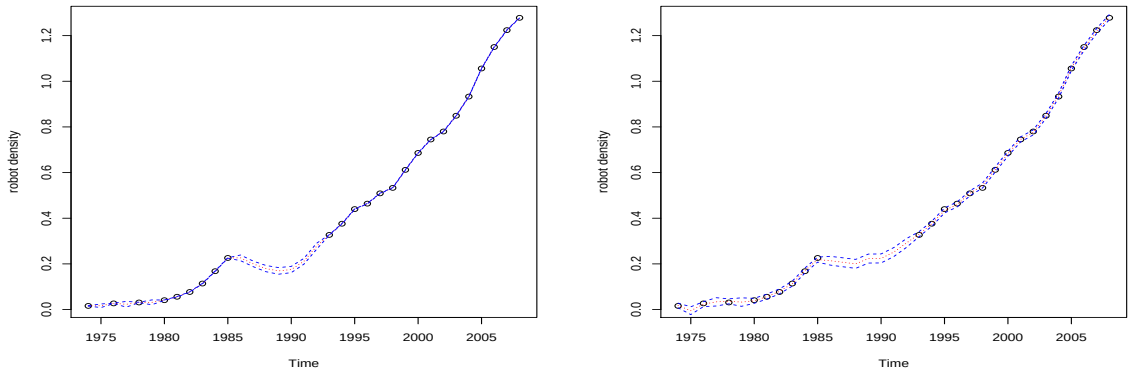


Note: Underlying indicator sets, first row: I_1, I_2, second row: I_3, I_4 (from left to right); black line and dots: empirical values, red dots: smoother values, blue dashed lines: 95 % C.I. of prediction errors

In line with our intuition that –both and primarily– productivity and costs matter with

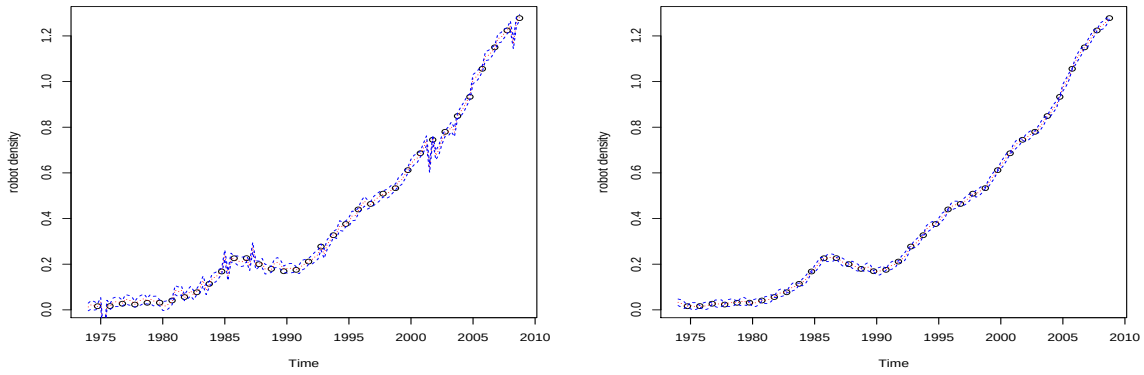
regard to automization, information set I_3 (lower left schedule in Fig. A.1) generates the most accurate and reasonable annual series nowast.

Figure A.2: Observation-equation-variance sensitivity of I_3-based nowcast



Note: Variances in observation equation matrix ($\boldsymbol{H}_t$) doubled from 1 (left) to 2 (right schedule); black circles: empirical values, red dots: smoother values, blue dashed lines: 95 % C.I. of prediction errors

Figure A.3: Observation/state-equation-variance sensitivity of quarterly nowcast



Note: Quarterly I_3-based (filling up I_3-based annual nowcast); variances in observation (state) equation matrix $\boldsymbol{H}_t$ ($\boldsymbol{Q}_t$) are set to 0.01 (1) for left schedule and 0.01 (0.01) for right schedule, respectively; black circles: empirical and nowcasted (q4-/fourth quarter) annual values; remaining legend as for Fig. A.2

As can be seen from Fig. A.2 our nowcasts are slightly sensitive with regard to observation equation variances and produce more accurate predictions of the empirical observations for lower values. Thus, we proceed with the I_3-based annual nowcast with the lower observation equation variance values.

Quarterly series nowcast. For our quarterly series nowcast, we fill up the I_3-based annual projections and empirical values but now rely on an I_3 analogue using all quarterly (instead of just q4 or fourth quarter) information. The result is shown for different observation and state equation variance values, $\boldsymbol{H}_t$ and $\boldsymbol{Q}_t$, in Fig. A.3. The right schedule circle and red dot values correspond to our nowcasted quarterly series of choice and corresponds to the time series (solid blue line) displayed in Fig. 1.4.

A.2 VEC model analysis

A.2.1 Johansen procedure

Table A.1: Unit root (UR) and stationarity test statistics

	ADF I	ADF II	PP	KPSS
Log levels:				
Population growth	–1.732	–1.679	–1.840	0.399***
Factor tax policy mix (LCTR)	–1.717	–2.003	–5.379	0.988***
Robots per 1K workers	–2.841*	–2.792	–0.481	0.504***
Wage share	–1.253	–2.861	–2.338	0.290***
Capital share	–1.377	–1.937	–1.876	0.361***
Log first differences:				
Population growth	–11.36***	–11.39***	–11.36***	0.0631
Factor tax policy mix (LCTR)	–7.887***	–17.13***	–14.08***	0.0359
Robots per 1K workers	–5.086***	–5.468***	–10.61***	0.1030
Wage share	–15.81***	–9.181***	–12.52***	0.0415
Capital share	–6.547***	–6.618***	–4.111***	0.0381

Note: ADF – Augmented Dickey-Fuller (UR under null; I/II = without/with linear trend component); PP – Phillips-Perron (UR under null); KPSS – Kwiatkowski-Phillips-Schmidt-Shin (stationarity under null; with automatic bandwidth selection and autocovariances weighted by quadratic spectral kernel); * $p < .10$, ** $p < .05$, *** $p < .01$.

Table A.2: Johansen test procedure and test statistics for cointegrating equations

		Trace		Max EV		Info Criterion
Model	Max Rank	Stats	5% c.v.	Stats	5% c.v.	HQIC
(v1)	0	101.04	59.46	43.72	30.04	–35.52362
	1	57.31	39.89	30.31	23.8	–35.64426
	2	27.00	24.31	17.77	17.89	–35.70900
	3	9.23*	12.53	9.21	11.44	–35.72472*
	4	0.01	3.84	0.01	3.84	–35.72275
(v2)	0	145.46	76.07	68.08	34.4	–35.52362
	1	77.38	53.12	30.36	28.14	–35.81121
	2	47.02	34.91	23.44	22.00	–35.85150
	3	23.58	19.96	17.56	15.67	–35.88699
	4	6.01*	9.42	6.01	9.24	–35.92591*
(v3)	0	133.92	68.52	68.07	33.46	–35.49021
	1	65.85	47.21	30.23	27.07	–35.80259
	2	35.62	29.68	21.39	20.97	–35.86672
	3	14.22*	15.41	12.56	14.07	–35.91094
	4	1.66	3.76	1.66	3.76	–35.93535*
(v4)	0	207.72	87.31	116.61	37.52	–35.49021
	1	91.10	62.99	36.70	31.46	–36.15999
	2	54.39	42.44	25.86	25.54	–36.25025
	3	28.53	25.32	17.42	18.96	–36.30482
	4	11.11*	12.25	11.11	12.52	–36.3426*
(v5)	0	188.16	77.74	111.36	36.41	–35.51998
	1	76.79	54.64	36.56	30.33	–36.17327
	2	40.23	34.55	25.80	23.78	–36.28723
	3	14.43*	18.17	13.46	16.87	–36.36616
	4	0.97	3.74	0.97	3.74	–36.39762*

Note: Trace – trace test; Max EV – maximum eigenvalue test; Info Criterion – information criterion (IC) with IC of choice: HQIC – Hannan-Quinn information criterion; procedure starts with test for zero cointegrating equations (CE), i.e. a maximum rank of zero, and then accepts the first null that is not rejected (indicated by '$*$'); the Pantula Principle sequence is: (v1) no intercept or trend in CE or in VAR part; (v2) intercept and no trend in CE part and neither intercept nor trend in VAR part; (v3) intercept in CE part and VAR part, but no trends; (v4) intercept in CE part and in VAR part paralleled by linear trend in CE part or in VAR part.

As can be seen from Table A.1 and Table A.2 , all series in log levels used in the fitted VECM, as described in the empirical part of the paper, are $I(1)$ and the result of the Johansen test procedure is that there are, at least, three cointegrating relationships. The latter particularly concern population growth and the wage share, the factor tax policy mix (LCTR) and the capital share, and robots density and capital share, respectively. However, the last of these CE relationships is statistically significant at a 68% level of significance only according to the z-statistics of the Johansen normalized restriction test.

A post-estimation stability check confirms three (two) imposed unit moduli of eigenvalues of the companion matrix of our fitted VECM with two excact unit eigenvalues and one very close to one, i.e. with a value of .97. All remaining moduli of eigenvalues of the companion matrix are strictly less than one indicating stability. Serial correlation of residuals is clearly rejected by appropriate lagrange multiplier (LM) tests.

Against the backdrop of the performed tests (with further detail available on request) we assess our VECM specification as being, all in all, acceptable.

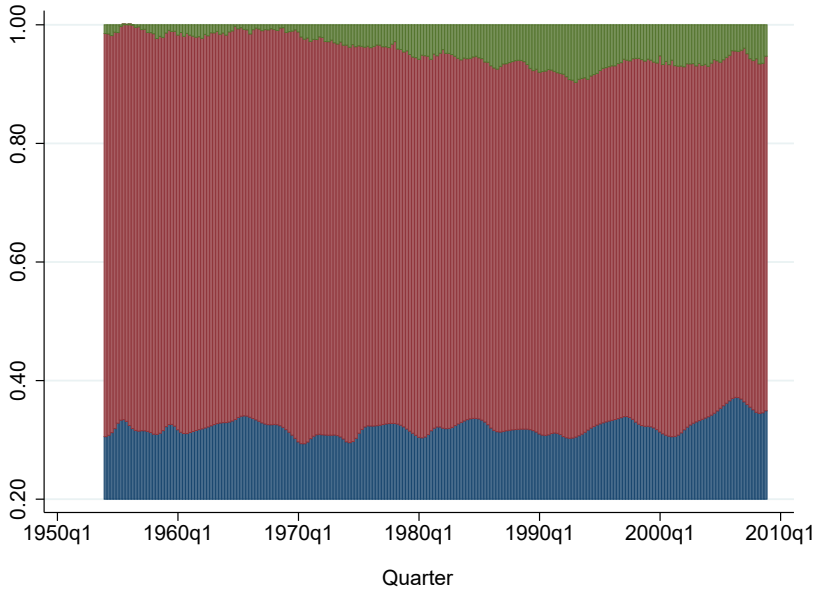
A.3 Construction of adjusted labor income share

Our adjusted labor income share is based on BEA time series for GDP and GDI, that is, compensation of employees, paid wages and salaries. The share has been adjusted assuming that a one third add-on to GDI is attributable to the self-employed. Just like any adjustment of the empirical labor's share, it represents a crude approximation. Contrary to alternative approaches, it precludes double counting.

A mean difference test for the two factor shares, capital income share and labor income share, summing to one (with mean rounded based on the digit at fourth decimal place) fails to reject the null of unity at any conventional level of significance.

An implied aggregate elasticity of substitution between capital and labor σ that is close to but slightly less than a unitary elasticity, $\sigma \leq 1$, is in line with recent evidence based on longitudinal data and consistent estimates by Glover and Short (2020). A $\sigma \lesssim 1$ (see the sum of the red and blue area in the following figure), as used in the empirical part of our paper, excludes the simple capital deepening explanation of the global decline in the labor's share. For $\sigma \lesssim 1$, the fall in the labor's share cannot be rationalized by "rising effective capital ratios through physical investment in response to the fall in investment prices" (Glover and Short, 2020, p. 35).

Figure A.4: Capital income share, labor income share, and elasticity of substitution



Note: Blue area: capital income share, red area: labor income share; green area: rest to unity

A.4 Summary statistics of series used in VECM

Table A.3: Summary statistics of series used (as log levels) in VECM

Series	Mean	Std. dev.	Min	Max	Skewness	Kurtosis
Population growth	0.0026	0.0004	0.0021	0.0035	0.7574	2.0393
Factor tax policy mix (LCTR)	0.7417	0.0985	0.4733	0.8891	–1.0011	3.0182
Robots per 1K workers	0.3936	0.3695	0.0149	1.0278	0.9128	2.0670
Wage share	0.6341	0.0299	0.5827	0.6914	0.0355	2.0890
Capital share	0.3247	0.0164	0.2953	0.3717	0.9221	3.0684

Note: Summarized series were transformed to log levels before using them in the VECM as outlined in the empirical part of the paper; robots per 1K workers is the pseudo-quarterly series nowcasted as described above; throughout the observation period ranges from 1974:q1 to 2008:q4 (N obs = 140). All sources of data are given in the text (empirical part of the paper).

A.5 Model with Automation Capital and Quadratic Capital Adjustment Costs

A.5.1 Household Optimization Problem

The Lagrangean of the household optimization problem is presented by

$$\begin{aligned}\mathscr{L} =\sum_{t=0}^{\infty}\beta^t\Bigg\{ & \frac{\left(c_t^{\theta}(1-l_t)^{1-\theta}\right)^{1-\eta}}{1-\eta} \\ & + \lambda_t\Big[(1-\tau_t^l)w_t l_t + (1-\tau_t^k)r_t^k k_t + (1-\tau_t^k)r_t^p p_t + tr_t - (1+\tau^c)c_t \\ & - \Phi^k(i_t^k, k_t) - \Phi^p(i_t^p, p_t)\Big] \\ & + q_t^k\left[i_t^k + (1-\delta)k_t - (1+n_t)k_{t+1}\right] \\ & + q_t^p\left[i_t^p + (1-\delta)p_t - (1+n_t)p_{t+1}\right]\Bigg\}.\end{aligned}$$

The first-order conditions of the household are presented by:

$$\lambda_t(1+\tau^c) = \theta c_t^{\theta(1-\eta)-1}(1-l_t)^{(1-\theta)(1-\eta)}, \tag{A.5.1a}$$

$$\lambda_t(1-\tau_t^l)w_t = (1-\theta)c_t^{\theta(1-\eta)}(1-l_t)^{(1-\theta)(1-\eta)-1}, \tag{A.5.1b}$$

$$q_t^k = \lambda_t\Phi_i^k(i_t^k, k_t), \tag{A.5.1c}$$

$$q_t^p = \lambda_t\Phi_i^p(i_t^p, p_t), \tag{A.5.1d}$$

$$(1+n_t)q_t^k = \beta\left\{\lambda_{t+1}\left[(1-\tau_{t+1}^k)r_{t+1}^k - \Phi_k^k(i_t^k, k_t)\right] + q_{t+1}^k(1-\delta)\right\}, \tag{A.5.1e}$$

$$(1+n_t)q_t^p = \beta\left\{\lambda_{t+1}\left[(1-\tau_{t+1}^p)r_{t+1}^p - \Phi_p^k(i_t^p, p_t)\right] + q_{t+1}^p(1-\delta)\right\}, \tag{A.5.1f}$$

where $\Phi_i^x(i^x, x)$, $x \in \{k, p\}$, denotes the first derivative of the adjustment cost function $\Phi^x(i^x, x)$ with respect to investment i^x. Similarly, Φ_x^x denotes the first derivative of this function with respect to the second argument $x \in \{k, p\}$. For the specification of the adjustment cost function Φ in (3.5), the first-order conditions are equivalent to those in (3.7).

A.5.2 Asymptotic steady state

If κ is sufficiently large, the households have incentives to invest in automation capital as well so that the economy displays economic growth $\gamma > 0$ in the long run with $p_t > 0$. Eventually, the economy converges to an 'Ak'-economy as in Romer (1986). Let γ define the growth rate of the capital stock in the economy with $k_{t+1}/k_t = 1+\gamma$.

Furthermore, from the dynamic equations of the capital stocks k_t in (3.4a), $i_t^k/k_t = i^k/k = (1+\gamma)(1+n) - 1 + \delta$ must be constant asymptotically, too, and therefore, from (3.7c), $q_t^k/\lambda = q^k/\lambda$. Asymptotically, the labor supply converges to zero.[8] This observation follows from the inspection of (3.7a) and (3.7b). For ever-increasing k and p, y and c are increasing as well, while the replacement of l by p and the build-up of augmented labor, $l + \kappa p_t$, implies that the wage rate grows at a lower rate (which will converge to zero in the long run). Therefore, the first-order conditions of the households with respect to his labor supply imply an ever-decreasing labor supply due to the income effect until the lower bound of zero starts binding.

When labor supply is zero in the asymptotic steady state, the marginal utility grows at the factor $(1+\gamma)^{\theta(1-\eta)-1}$ as implied by (3.7a). Therefore, (3.7e) implies that r^k converges to a constant. A constant asymptotic real interest rate r^k in turn implies a constant ratio $k_t/p_t = k/p$ so that k_t and p_t must grow at the same rate. From the dynamics of the automation capital stock p_t, (3.4b), this implies that $i^p/p = i^k/k$. Therefore, q_t^p is equal to q_t^k asymptotically (which follows from (3.7d)) and the two interest rates on both forms of capital must be equal, $r_t^k = r_t^p$. From this observation, (3.16) follows.

A.6 Computation of the transition dynamics

To compute the transition dynamics, we need to solve a difference equation system in the state variable $\{k_t, p_t\}_{t=1965}^{2085}$. We choose a time horizon of 120 years (=periods) so that the growth rates of the variables have stabilized and are numerically close to the asymptotic ones. By this choice of transition periods, we also observe that the wage share in 2015 has stabilized.

As endogenous variables, we use consumption c_t, investment in both capital stocks, i_t^k and i_t^p, and labor l_t. The endogenous equation consists of the household first-order conditions (3.7e), (3.7f), (3.7c), and the resource constraint (3.12). With the help of the endogenous variables, it is straightforward to compute the dynamics of the two capital stock, k_t and p_t, from (3.4) for given initial capital stock k_{1965} and $p_{1965} = 0$ as computed from the initial steady state without automation capital.

With the help of k_t, p_t, and l_t, we can compute the factor prices w_t, r_t^k, and p_t. From the first-order conditions (3.7a), (3.7c) and (3.7d) we can compute λ_t, q_t^k, and q_t^p.

[8] As noted above, this occurs in the year 2069 in our benchmark economy.

Therefore, we have specified the values of all variables that show up in the endogenous equations. For the endogenous variables in the year 2086 which we also need for the computation of the endogenous variables, we assume that c_t, i_t^k, and i_t^p grow at the asymptotic growth rate, while l_t falls at the asymptotic growth rate. We find the effect of this assumption to be negligible for the transition during 1965-2030 which is the focus of the analysis.

In essence, we have to solve a non-linear equations problem in $4 \times 120 = 480$ variables. This is a non-trivial task and the problem is to find a good initial value. We, therefore, proceed as following:[9]

Algorithm: Computation of the transition in the benchmark model with automation capital

Step 1: Compute the initial steady state of the neoclassical growth model.

Step 2: Compute the final steady state of the neoclassical growth model (with $\kappa = 0$) for the tax rates and the population growth rate prevailing in 2015.

Step 3: Project a transition path for the neoclassical growth model (with $\kappa = 0$) for $\{i_t^k, c_t, l_t\}_{t=1965}^{2015}$ in the form of a linear adjustment.

Step 4: Solve the simple neoclassical growth model (with $\kappa = 0$).

Step 5: Use the transition path from the standard neoclassical growth model with $p_t \equiv 0$ for $t = 1965, \ldots, 2015$ as initial guess for the computation of the transition in the model with automation capital.

Step 6: Iterate over the time horizon of the transition T by incremental steps of one year. Use the transition path of the previous iteration as an initial guess assuming that in the period $T + 1$, the variables $\{i_t^k, i_t^p, c_t\}_{t=1965}^{1965+T}$ grow at the rate γ, while l_t falls at the rate γ.

Step 7: Stop when the dynamics of the model during the period 1965-2030 do not change any more.

[9]The Gauss computer code is available from the authors upon request.