

Hellwagner, Timon; Weber, Enzo

**Conference Paper**

## Labour Market Adjustments to Population Decline

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2021: Climate Economics

**Provided in Cooperation with:**

Verein für Socialpolitik / German Economic Association

*Suggested Citation:* Hellwagner, Timon; Weber, Enzo (2021) : Labour Market Adjustments to Population Decline, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2021: Climate Economics, ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/242455>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Labour Market Adjustments to Population Decline: A Historical Macroeconomic Perspective, 1875-2016

Timon Hellwagner\*<sup>†</sup>      Enzo Weber<sup>†‡</sup>

September 10, 2021

## Abstract

According to demographic projections, advanced economies will face population decline in the years and decades to come, particularly among working-age. Despite this impending profound transition, there is little empirical evidence of corresponding labour market implications. Tackling this shortcoming from a historical macroeconomic point of view, we compiled a new dataset for nine advanced economies, covering demographic and labour market variables on an annual basis from 1875 to 2016. By using a panel smooth transition VAR (PSTVAR), we analyze macroeconomic labour market interdependencies while simultaneously exploiting the cross-country variation to identify non-linearities of these interdependencies conditional on the demographic regime, that is periods of population growth or decline. Our results suggest that labour market adjustments to population shocks are, in fact, non-symmetrical. Most notably, labour supply shortages are mitigated by an increased participation rate, largely recovering the pre-shock employment level. Also, investments recover more strongly compared to periods of population growth. We find no clear effects on the unemployment rate and wages.

**Keywords:** population decline, labour market research, historical dataset, panel smooth transition VAR

**JEL Codes:** J11, J21, J31, E24

---

\*Corresponding author: [timon.hellwagner@iab.de](mailto:timon.hellwagner@iab.de)

<sup>†</sup>Institute for Employment Research (IAB), 90478 Nuremberg, Germany

<sup>‡</sup>Department of Economics, University of Regensburg, 93040 Regensburg, Germany

The authors thank Tobias Hartl and Carsten Trenkler for valuable comments, suggestions, and coding support. Financial support through the joint graduate program in labour market research of the Institute for Employment Research (IAB) and the School of Business and Economics at the University of Erlangen-Nuremberg, GradAB, is gratefully acknowledged by Timon Hellwagner.

# 1 Introduction

A hitherto stylized fact, the perpetual growth of the population, is questioned in the short and medium term by a range of demographic forecasts (e.g. [UN 2019](#)). In the years and decades to come, depending on the scenario under consideration, advanced economies will face a stagnation and, sooner or later, a decline of their populations. The latter is expected to be particularly pronounced among those in working-age. In fact, as the solid line in figure 1 illustrates, the aggregate working-age population of advanced economies passed its 'tipping point' in 2018 and has been on decline ever since. Additionally, as the shaded area in figure 1 indicates, also the share of countries facing working-age population decline has seen a surge in the past years and is expected to increase further in the upcoming decades. Thus, in stark contrast to the more recent population history, the impending transformations will be pronounced, widespread and enduring, providing a changed demographic context for most advanced economies in the long run.

Undoubtedly, given the contemporary and future relevance for various countries as well as the importance of demography for economic growth in general and the labour market in particular, questions about the economic implications of population decline emerge. Ultimately, the expected developments may challenge other supposedly stylized facts as well, such as the ever-accelerating growth of GDP (per capita) ([Jones and Romer 2010](#)) or the constant

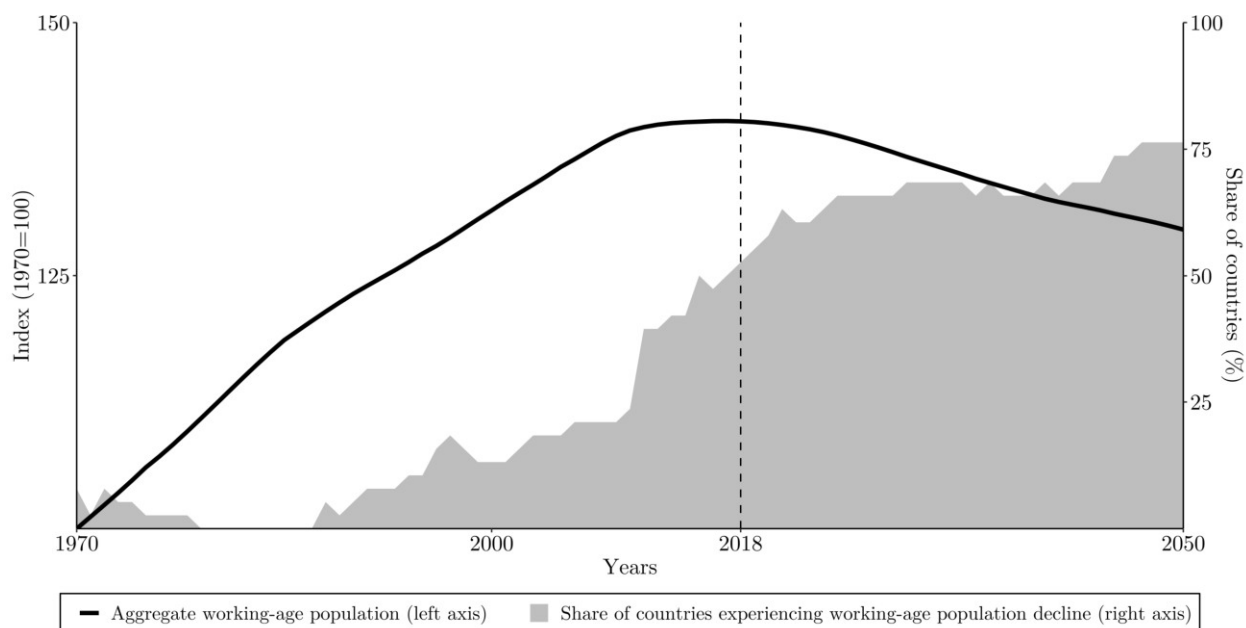


FIGURE 1: VISUALIZATION OF THE IMPENDING OR OCCURRING WORKING-AGE POPULATION DECLINE AMONG ADVANCED ECONOMIES.

Source: [UN 2019](#), medium variant. The definition of advanced economies follows that of [IMF \(2021\)](#).

labour share in national income (Kaldor 1961). However, despite its occurrence or imminence in most advanced economies, yet there is a substantial undercoverage of theoretical and empirical research on the economic implications of population decline, in general as well as with regard to the labour market.

In formal economic modelling most approaches assume a growing or at least stagnant population (Jones 2020). On the contrary, population *decline* and the accompanying implications have hardly been discussed yet. In the literature, some attempts have been made to investigate the effects of demographic changes in Ramsey-type models, for example when population growth is not constant (Kajanovičová et al. 2020) or logistic (Brida and Accinelli 2007). Sasaki (2019) analyzes the consequences of negative population growth on the long run growth rate of per capita output using a Solow-type growth model and demonstrates that, if in such a setting the elasticity of substitution is less than unity, economic growth exclusively depends on the rate of technological progress. Christiaans (2011) as well as Sasaki and Hoshida (2017) use semi-endogenous growth models to investigate the effects of population decline on economic growth. The results suggest varying responses of economic growth to negative population growth, depending on the assumed depreciation rate of capital. More recently, Jones (2020) demonstrates that, in the case of population decline, endogenous and semi-endogenous growth models lead to an “Empty Planet”, that is stagnating living standards and knowledge. By taking one step further and endogenizing fertility, Jones (2020) shows that economic growth can be resumed even under conditions of population decline if the economy switches to an optimal allocation soon enough.

Among empirical studies, the undercoverage is even more distinct and may be explained by the fact that there have been comparatively few periods of population decline among advanced economies in the recent past, hampering the reliable identification of its effects. Consequently, existing research on the demography-economy-nexus focuses on a variety of different issues: a multitude of empirical studies analyze the effects of population growth (see Headey and Hodge 2009 for a comprehensive meta-study), population ageing (e.g. Acemoğlu and Restrepo 2017; Börsch-Supan 2008) or changing mortality, fertility, and human capital patterns (for many: Barro 1991, Barro and Lee 1994, Barro 1998, Bloom and Williamson 1998, Hall and Jones 1999) on economic growth. From a more conceptual perspective, both the secular stagnation debate (e.g. Eggertsson et al. 2019) and the unified growth theory (e.g. Cervellati et al. 2017), among others, have addressed the role of demography for long-term economic development. Conversely, population decline has not yet drawn attention in the literature.

Thus, sparse theoretical contributions such as the more recent by Jones (2020) suggest that the economic effects of growth and decline in the population do not need to follow

symmetrical paths. But, as noted, labour market questions relevant in context of population decline, such as the behaviour of wages, the capital utilization of firms or the elasticity labour supply when the labour force is declining or aggregate demand is decreasing, have so far hardly been addressed, neither in theory nor in empirics. To provide an empirically substantiated starting and orientation point for both policy and future research, such as the incorporation of labour market adjustments to population decline in theoretical growth models, we examine the effects of population decline on the labour market in the very long run in this paper.

Operationalising our analysis consists of two key components. *First*, the occurrence of periods of actual population decline and the availability of labour market data do not necessarily coincide. As noted above, for most advanced economies population shrinkage appears to be a phenomenon of the near future. However, if we take a more historical perspective back to the second half of the 19<sup>th</sup> century, we are able to more identify several periods of decline and low population growth, for example right before the *fin de siècle*, distributed across several countries. On the one hand, this suggests to empirically investigate population decline and its macroeconomic implications in a historical cross-country framework. On the other hand, economic data availability proves to be sparse in the very long run. To this end, we compiled a new historical dataset from a large number national and international sources. We collected information on population, net migration, real GDP, real wages, investment, employment, unemployment, and labour force participation for nine countries, providing an annual coverage from 1875 to 2016. *Second*, the econometric strategy must adequately address possible non-linear responses to population decline. To account for this, we specify a panel smooth transition VAR (PSTVAR). In doing so, we contribute to growing bodies of literature which use, first, cross-country settings (e.g. [Aksoy et al. 2019](#)) and, second, regime-dependent methods (e.g. [Auerbach and Gorodnichenko 2012](#)) to analyze (non-linear) macroeconomic interdependencies. Using the estimation result, we derive regime-dependent impulse response functions to a one percent population shock. We compare the impulse responses both between regimes but also to those obtained from a linear model.

Our results suggest that the effects of population changes on the economy and the labour market are, in fact, regime-dependent. In order to mitigate labour supply shortages, the labour force participation increases and, as a result, employment faces a remarkably less distinct decrease in decline periods compared to the increase in growth periods. Our results do not show an effect on the unemployment rate but indicate that investment is recovering more strongly in times of decline than expected from linear estimation results. Eventually, we do not find clear evidence whether wages increase, as a shortage indicator, or decrease, due to the overall economic downturn coupled with increased labour supply.

The remainder of the paper is structured as follows. In section 2, we introduce our historical dataset. Section 3 outlines the econometric strategy, whereas the corresponding results are presented and discussed in section 4. In section 5, we conduct a series of robustness checks to ensure that our empirical strategy provides reliable results. The last section concludes. The appendix contains additional figures, an extensive list of included data sources, and a detailed explanation of data preparation steps taken to obtain the historical dataset.

## 2 Data

The exploration of historical economic dynamics across countries is a notoriously difficult task, particularly when focused on labour market issues. Well-known data collections such as the International Historical Statistics ([Mitchell 2013](#)) or Maddisons Historical Statistics ([Bolt and van Zanden 2020](#)) and their respective predecessors, among others, have settled the path for comparative historical economic research for decades. In these key sources, however, the annual availability in the very long run remains limited to selected variables. We have seen substantial improvements in recent years by compilations such as the Macrohistory Database ([Jordà et al. 2017](#)) or the Long-Term Productivity Database ([Bergeaud et al. 2016](#)), both starting in the second half of the nineteenth century, covering a variety of advanced economies, and broadening the range of available macroeconomic variables. Nevertheless, the availability of most annual labour market data, such as unemployment and labour force participation, remains limited, hampering corresponding analyzes of the labour market in the very-long run.

Based upon this finding, we compiled a new historical annual labour market dataset as a result of extensive data acquisition efforts, drawing both on existing macroeconomic and demographic databases and, particularly, on a vast number of national data collections. Overall, the compilation combines more than 80 individual sources. The dataset contains information on population (both total and age groups), net migration, real GDP, real wages, investment, employment, unemployment, and labour force participation. Our compilation covers nine countries<sup>1</sup> and 141 years (1875-2016). The dataset is balanced and has no gaps. Population, real GDP, real wages, and employment are prepared as an index allowing us to cope with a multitude of large-scale boundary changes. Investment is given as share of GDP, similarly net migration, unemployment, and labour force participation are given as rates. The latter measures the rate of those active in the labour force among the population 15-64 years. In the section on our econometric strategy below, we outline how these variables enter the estimations. Sources and data preparation steps can be found in the appendix.

---

<sup>1</sup>AUS, DEU, DNK, FRA, GBR, NLD, NOR, SWE, USA

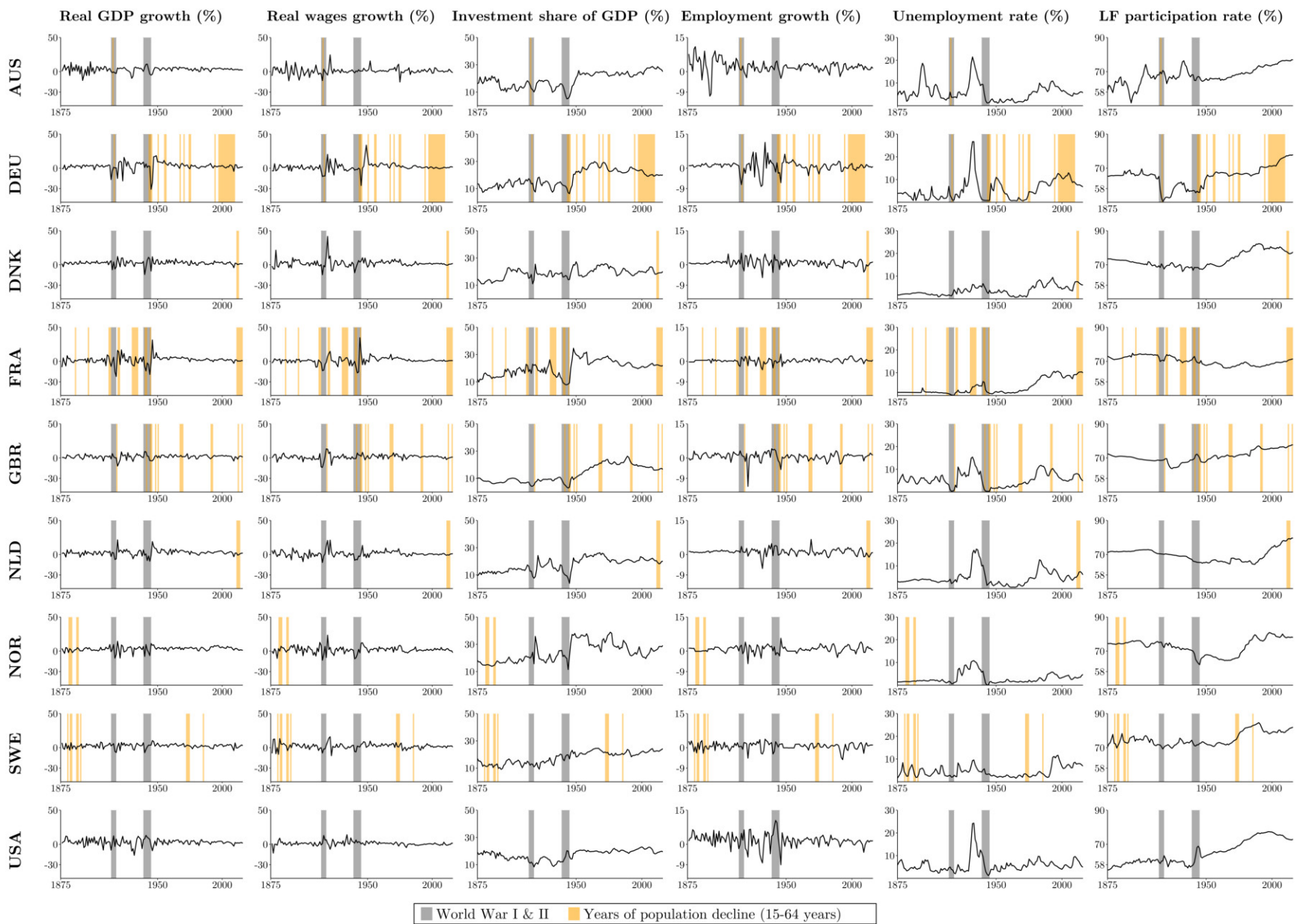


FIGURE 2: OVERVIEW OF THE LABOUR MARKET VARIABLES IN THE HISTORICAL DATASET FROM 1875 TO 2016.

Source: See the appendix.



Name	Definition	Obs.	Mean	SD	Min	Max
POP	Population (aged 15-64), growth rate (%)	1278	0.95	0.81	-5.69	4.99
RGDP	Real GDP, growth rate (%)	1269	2.72	4.34	-30.96	31.42
INVEST	Investment share, percentage	1278	18.45	5.96	2.96	38.89
RWAGES	Real wages, growth rate (%)	1269	1.57	4.74	-25.40	41.73
EMP	Employment, growth rate (%)	1269	1.01	2.07	-12.52	11.32
UNEMP	Unemployment, percentage	1278	4.72	3.46	0.04	26.70
LFPR	LF participation (aged 15-64), percentage	1278	70.05	5.71	50.55	84.67

TABLE 1: DESCRIPTIVE STATISTICS OF THE CORE VARIABLES IN THE HISTORICAL DATASET.

Source: See the appendix.

Figure 2 gives a visual overview of the economic variables in the dataset, illustrates developments over time, and highlights periods of actual working-age population decline. In the figure, the index variables working-age population (15-64 years), real GDP, real wages, and employment are shown as growth rates (log-differenced indices) since they enter our econometric model as such. Investment share, the unemployment rate, and the labour force participation rate are shown as percentages; again, since they enter our model as percentages. Correspondingly, table 1 shows descriptive statistics of the collected variables.

Additional to a general overview of the dataset, figure 2 illustrates the distribution of population decline observations across time and across countries. Except for Germany in the 2000s and early 2010s, advanced economies have experienced population decline mostly in the far past: Either in the interwar period, such as France, or even earlier, at the beginning of the 20<sup>th</sup> century, as for example in the case of Norway and Sweden.

### 3 Econometric Model

As shown above, periods of population shrinkage have occurred comparatively rarely in the recent past, and, even if one takes a more historical perspective, barely in the same country. To account for this, we draw on a panel dataset and use the cross-country variation to identify the differing effects of population growth and decline. In general, we use a panel VAR (PVAR), and in doing so, we contribute to a growing body of literature making use of panel VARs in macroeconomics (e.g. [Aksoy et al. 2019](#)). Applying a vector autoregressive structure allows to flexibly analyze macroeconomic interdependencies without *a priori* imposing assumptions on the directions of effects. The growing popularity of panel VARs in macroeconomics can be traced back to further appealing features: including both dynamic



and static interdependencies while simultaneously allowing for cross-sectional heterogeneity (Canova and Ciccarelli 2013). Drawing on this literature, we can specify our model in its linear version as

$$Y_{it} = \mu_i + \delta_t + AY_{i,t-j} + u_{it} \quad (1)$$

where  $Y_{it}$  is the vector of endogenous variables,  $\mu_i$  and  $\delta_t$  denote country- and time-fixed effects respectively.  $Y_{it}$  comprises six variables: the growth rate of working-age population<sup>2</sup>,  $POP_{it}$ ; the growth rate of real GDP,  $RGDP_{it}$ ; the growth rate of real wages,  $RWAGES_{it}$ ; the share of investment in GDP,  $INVEST_{it}$ ; the growth rate of employment,  $EMP_{it}$ ; and the unemployment rate,  $UNEMP_{it}$ . All growth rates are log-differenced indices. The investment share, the unemployment rate, and the labour force participation enter the equation in percentages. In the case of nonstationary, the flexible VAR is able to identify the unit root and treat it appropriately (see Sims et al. 1990).

However, since the focus of the present paper is the analysis of possibly different effects of population decline compared to population growth, we take one step further and combine our panel VAR with a non-linear smooth transition approach, thereby following a strand of the literature that uses common vector autoregressive models and non-linear extensions to account for regime-wise interdependencies of macroeconomic variables. Thus, our linear model in (1) is modified as follows

$$Y_{it} = \mu_i + \delta_t + [1 - P(q_{it})]GY_{i,t-j} + [P(q_{it})]DY_{i,t-j} + u_{it} \quad (2)$$

where  $G$  and  $D$  are matrices holding the regime-dependent coefficients of the endogenous variables, and  $P(q_{it})$  refers to the probability of experiencing population decline. This probability is given as

$$P(q_{it}) = \frac{\exp[-\gamma(q_{it} - c)]}{1 + \exp[-\gamma(q_{it} - c)]} \quad (3)$$

where  $q_{it}$  is the transition variable,  $\gamma$  defines the smoothness of the transition, and  $c$  is a location parameter defining the value of  $q_{it}$  at which the regime-switch occurs. Similarly, the error term,  $u_{it} \sim N(0, \Omega_t)$ , is characterized by a regime-dependent variance-covariance matrix  $\Omega$ , such that

$$\Omega_t = [1 - P(q_{it})]\Omega_G + [P(q_{it})]\Omega_D \quad (4)$$

Figure 3 illustrates the principle of the smooth-transition regression approach. The transition

---

<sup>2</sup>Notably, demographic variables such as population growth are sometimes treated exogenously in panel VARs (again see as an example Aksoy et al. 2019). We explicitly model population growth endogenously to investigate the effects of population shocks in the labour market system.

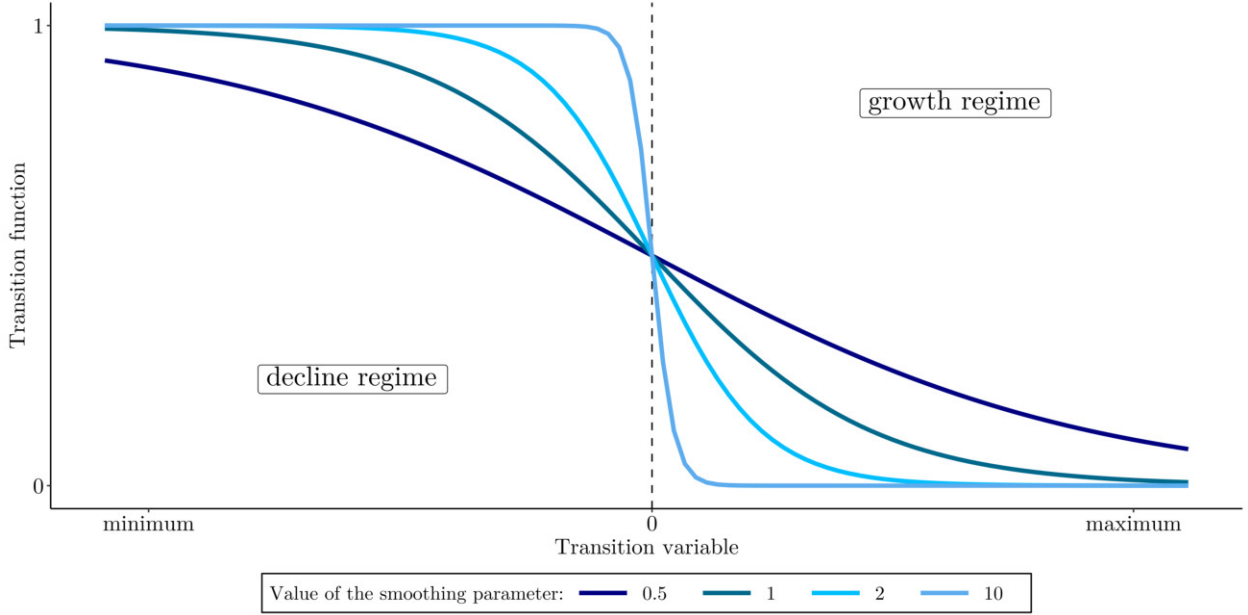


FIGURE 3: STYLIZED STR REGRESSION TECHNIQUE DEPENDING ON A VARYING SMOOTHING PARAMETER  $\gamma$  AND A LOCATION PARAMETER  $c = 0$ .

function in (3) governs the weights on the variables in the growth and the decline regime conditional on the supplied smoothing parameter  $\gamma$ . The larger this value, the faster a change from the growth to the decline and vice versa occurs. Since population changes are usually of low frequency, we may expect a low value of  $\gamma$  in this paper.

Following implementation strategies in the literature (e.g. [Auerbach and Gorodnichenko 2012](#) or [Gehrke and Hochmuth 2020](#)), our transition variable is a standardized index of the working-age population growth rate across all countries with zero mean and unit variance. Using this index,  $\gamma$  is calibrated to match the number of years of population decline in the sample, which is about 6.2 percent of the observations. Put differently, we define population as  $P(q_{it}) > 0.938 = 1 - 0.062$  and calibrate  $\gamma$  to match  $Pr[P(q_{it})] \geq 0.938 \approx 0.062$ . Since we are interested in population decline, we set  $c = 0$ . This setting yields a value of  $\gamma = 2.28$ . Figure 4 displays the decline probabilities, i.e. the 'weights' on the decline regime, in each of the nine countries. To avoid endogenous feedbacks, we use the lagged values of the transition function, i.e.  $P(q_{i,t-1})$ . Thus, the model becomes

$$Y_{it} = \mu_i + \delta_t + [1 - P(q_{i,t-1})]GY_{i,t-j} + [P(q_{i,t-1})]DY_{i,t-j} + u_{it} \quad (5)$$

$$\Omega_t = [1 - P(q_{i,t-1})]\Omega_G + [P(q_{i,t-1})]\Omega_D \quad (6)$$

We address other specifications with country-specific values of  $\gamma$  and differently dated tran-

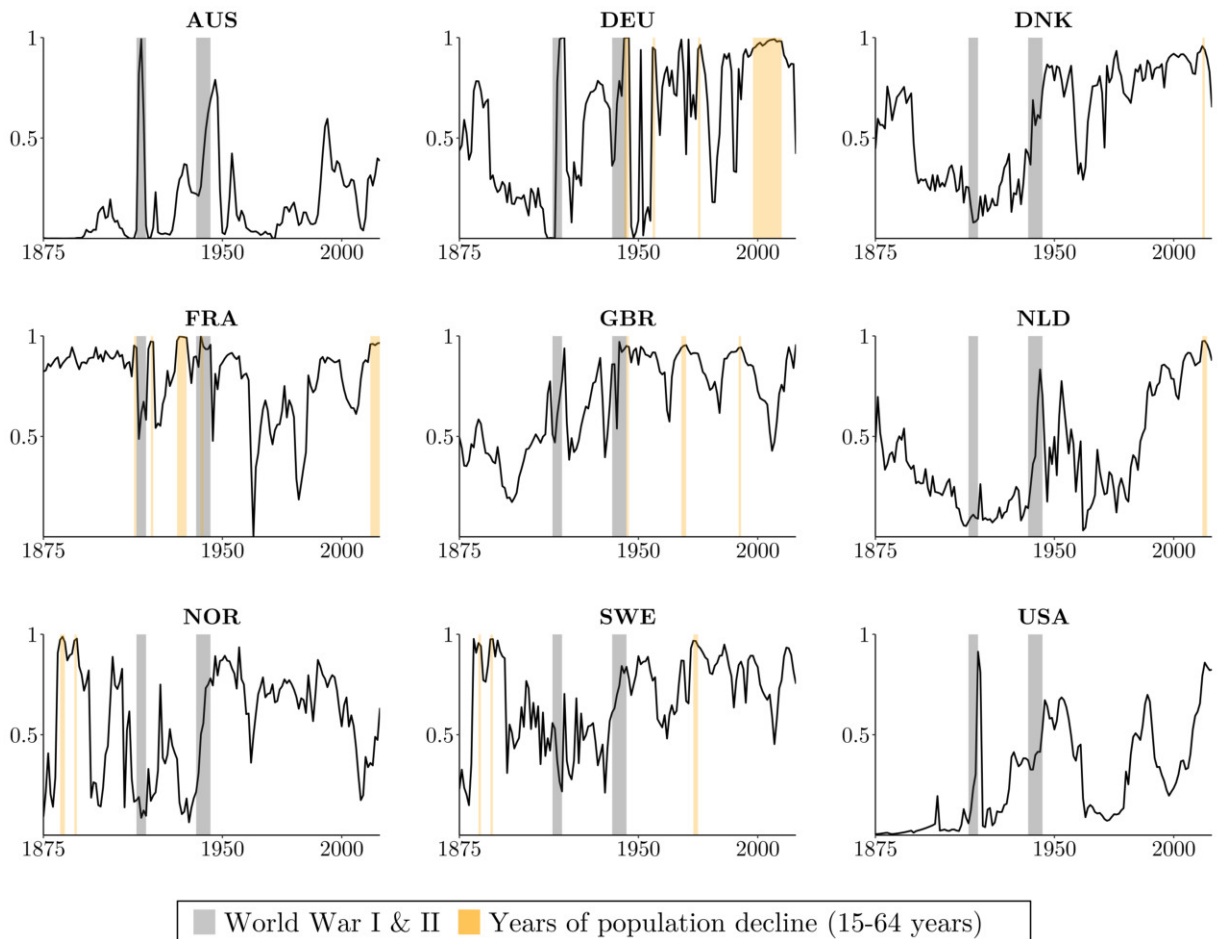


FIGURE 4: WEIGHTS ON THE DECLINE REGIME IN THE BASELINE MODEL.

sition functions in the robustness section to check for inference stemming from the regime-design.

Having found an appropriate estimation strategy, we turn to the model specification and test for the lag length using the Akaike information criterion (AIC). We test for symmetric lag length first. However, since our variable of interest, population growth, is of low frequency, corresponding economic and labour market effects may occur with substantial delay. Thus, we also test for specifications in which population growth possibly take more lags than all other endogenous variables. In fact, the AIC values suggest using five lags of population growth and one lag of all other endogenous variables, showing that our asymmetric specification is superior compared to all other symmetric panel smooth transition VAR specifications.

We use a two-way within-transformation and apply OLS estimation equation-by-equation. As common to empirical approaches using panel data and smooth-transition regression, we investigate whether our non-linear specification is superior to the linear case. To this end,

we apply the  $LM\chi$  linearity test as proposed for panel smooth-transition models by [González et al. \(2017\)](#) equation-wise. The test results strongly reject the null hypotheses, thus our non-linear specification captures the dynamics of the system significantly better than the linear design. This finding also is confirmed by comparing the linear and the non-linear specification using the AIC.

Ordering the population growth equation first and applying a Cholesky decomposition, we calculate impulse responses to obtain the impact of a positive (negative) one percent population shock on all endogenous variables in the growth (decline) regime. In doing so, we assume that the population shock affects all other variables in the same period but, in turn, shocks in other variables do not contemporaneously affect population growth. In the robustness section, we repeat this estimation but remove migration from the population growth series in order to account for possible endogeneity issues. In any case, we calculate the respective median impulse responses and derive 68% as well as 90% confidence intervals using moving-blocks residual resampling with 5,000 draws.

Additionally, in each draw, we obtain the impulse response for the overall labour force participation, which we are able to derive since we use information on population, employment, and unemployment in our model. Thus, the labor force participation rate is included implicitly and can be easily calculated. Put differently, in each draw, we use the IR estimates of population, employment, and unemployment to analyze the change in the labor force participation rate,  $LFPR_{it}$ . Formally, the impulse response of the labor force participation rate at time  $t$ ,  $IR_{LFPR,t}$ , is obtained by

$$IR_{LFPR,t} = \frac{\left( \frac{(emp + emp * IR_{EMP,t})}{(1 - unemp + unemp * IR_{UNEMP,t})} \right)}{(pop + pop * IR_{POP,t})} \quad (7)$$

where  $unemp$  is the average unemployment rate across all countries and years in the dataset,  $emp$  and  $pop$  may take any value since we are using growth rates in our model, and  $IR_{POP,t}$ ,  $IR_{EMP,t}$ , and  $IR_{UNEMP,t}$  are the impulse responses of population, employment, and unemployment respectively at time  $t$ . In the appendix, we demonstrate that the performance of our  $IR_{LFPR,t}$  calculation is robust to estimating the effect of a population shock via explicit inclusion of the labour force participation rate variable,  $LFPR_{it}$ , in our panel VAR. As figure A.1 illustrates, our strategy is already robust when compared to the labour force participation rate derived from a simple linear bivariate PVAR.

## 4 Results

Below, we report the results of the estimated orthogonal impulse response functions to populations shock of one percent, as outlined in the preceding section. First, we analyze the symmetric effects in a linear panel VAR. Subsequently, we investigate the reactions in the non-linear framework, i.e. possibly non-symmetric effects conditional on the regime.

### Symmetric effects to population shocks in the linear case

In figure 5, the symmetric impulse responses of the endogenous variables in the linear panel VAR are shown. Here, the dotted line indicates the median impulse response of the endogenous variables to a positive population shock of one percent. Vice versa, the solid line indicates the median impulse response of the variables to a negative population shock of one percent. By construction, the responses are symmetrical. The gray shaded areas around the median impulse responses represent the 68% and 90% confidence intervals, respectively. The plots for population, real GDP, real wages, and employment show the cumulative impulse responses since the variables entered our model as growth rates; the interpretation is in percent. The plots for investment, unemployment, and labour force participation show the reactions in percentage points because they enter the model as such.

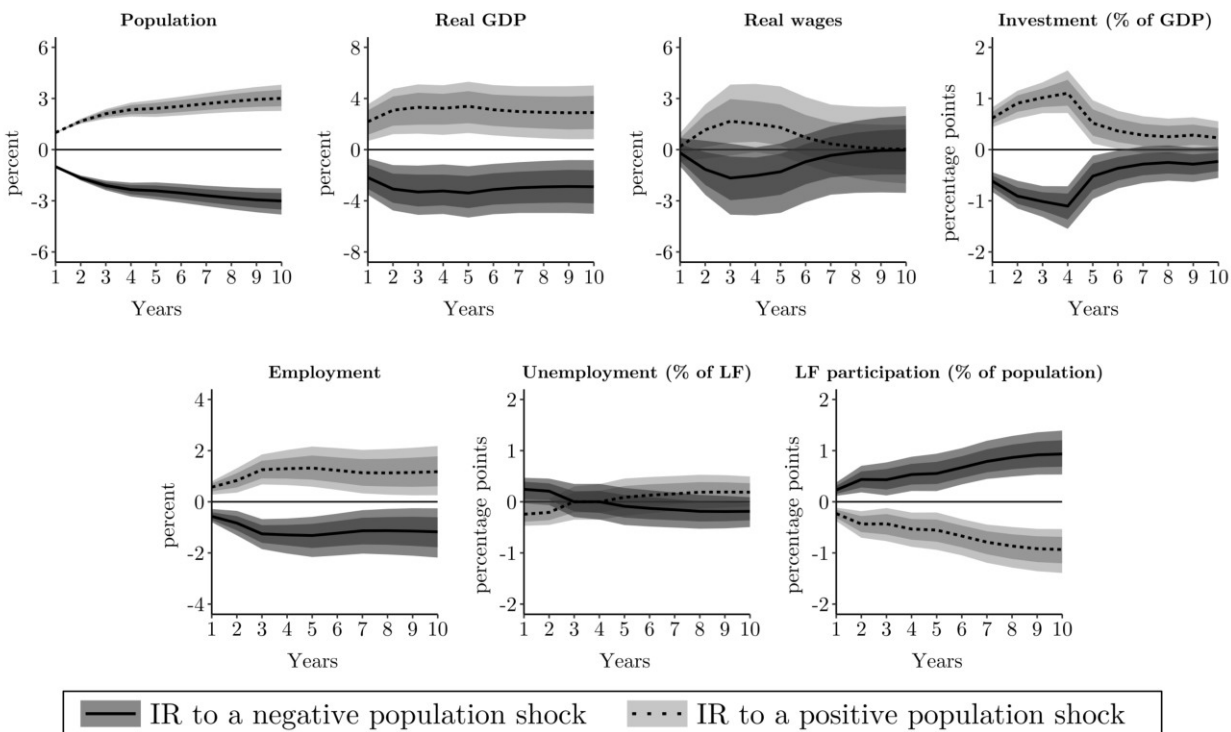


FIGURE 5: SYMMETRIC IMPULSE RESPONSES IN THE LINEAR PVAR.

The symmetric impulse responses from the linear PVAR suggest that a negative 1% population growth shock has permanent negative effects on population as well as on economic and labour market variables. After the shock, the cumulative effect on the population further expands to -3.3% in the long term ( $t_{10}$ ). The median response of real GDP shows a significant decrease of 2.1% in the contemporaneous period ( $t_1$ ) and 2.9% in the long term, indicating a permanently lower GDP as a consequence. Contrarily, real wages show more delayed and less permanent effects, with an insignificant contemporary reaction and a subsequent decline of 1.5% until  $t_3$ , being significant on the 68% level, followed by a return to quantitatively and qualitatively non-relevant differences from zero. Investment shows a substantial and significant decline at the beginning, -0.6 percentage points, indicating that investment, in relative terms, declines even more than the overall economy. This decrease turns out to be permanent, as the median impulse response partially returns back to zero in the long term but remains significantly lower on both the 68% and 90% level.

Notably, employment faces a statistically significant decrease, not only as a contemporaneous (-0.6%) reaction but also in the long term (-1.3%). Conversely, the unemployment rate hardly reacts in response to a population shock. Initially, there is an increase of 0.2 percentage points, being significant on the 68% level in the contemporaneous as well as the subsequent period, but showing insignificant effects afterwards and reaching -0.3 percentage points in  $t_{10}$ . Lastly, the labour force participation rate shows a remarkable and statistically significant surge, with median 0.2 percentage points in the beginning and 1.0 percentage points in the long term.

Now, turning to the core of the empirical analysis, that is the investigation of labour market adjustments in times of population decline, we analyze in which way the results outlined above change when incorporating the non-linear structure of our baseline model. In the subsequent sections, we first inspect the impulse responses in the decline regime, compare them to the linear model, and then contrast these findings with the impulse responses in the growth regime.

### **Effects of a negative population shock in times of population decline**

Thus, focusing on the analysis of regime-dependent effects, figure 6 shows the impulse responses of the endogenous variables to a negative one percent population growth shock in the decline regime as resulting from our baseline specification. As before, the solid line represents the median impulse responses, the shaded areas indicate 68% and 90% confidence intervals, and the interpretation is in percent respectively percentage points, depending on the variable under consideration. The solid red line illustrates the median impulse of the linear model for comparison.

Notably, there are some substantial differences in economic reactions when allowing for regime-dependent interdependencies. First, compared to the linear model, the reaction of population quickly flattens out after two periods and reaching  $-1.4\%$  in the long term, contrarily to the ongoing decline in the linear model. Second, while real GDP has seen a permanent and significant decline in the linear specification, the impulse response to a negative population shock now tends to suggest a return back to a less severe cumulative effect in the long term (median impulse response in  $t_{10}$  is at  $-2.1\%$ ; significant only on the 68% level). Again, real wages react with a one-period delay and show an effect of  $-1.7\%$  in  $t_2$ . Compared to the linear PVAR, the return to insignificance is even faster. Nevertheless, in the long term, the quantitative effect is greater in the non-linear case, showing a cumulative effect on real wages of  $-0.9\%$ . The impulse response of the investment share shows similar reactions as above but, again, the return to insignificance turns out to be faster, returning to a higher share in the long term compared to the results of the linear model. Importantly, the impulse response of employment differs from the previous findings. While the contemporaneous reaction is comparable ( $-0.6\%$ ), the employment level now increases again, reaching a median impulse response of  $-0.3\%$  in the long term, being statistically insignificant on both the 68% and the 90% level. As before, the unemployment rate shows minor reactions (0.2 percentage points), statistically significant only on the 68% level in

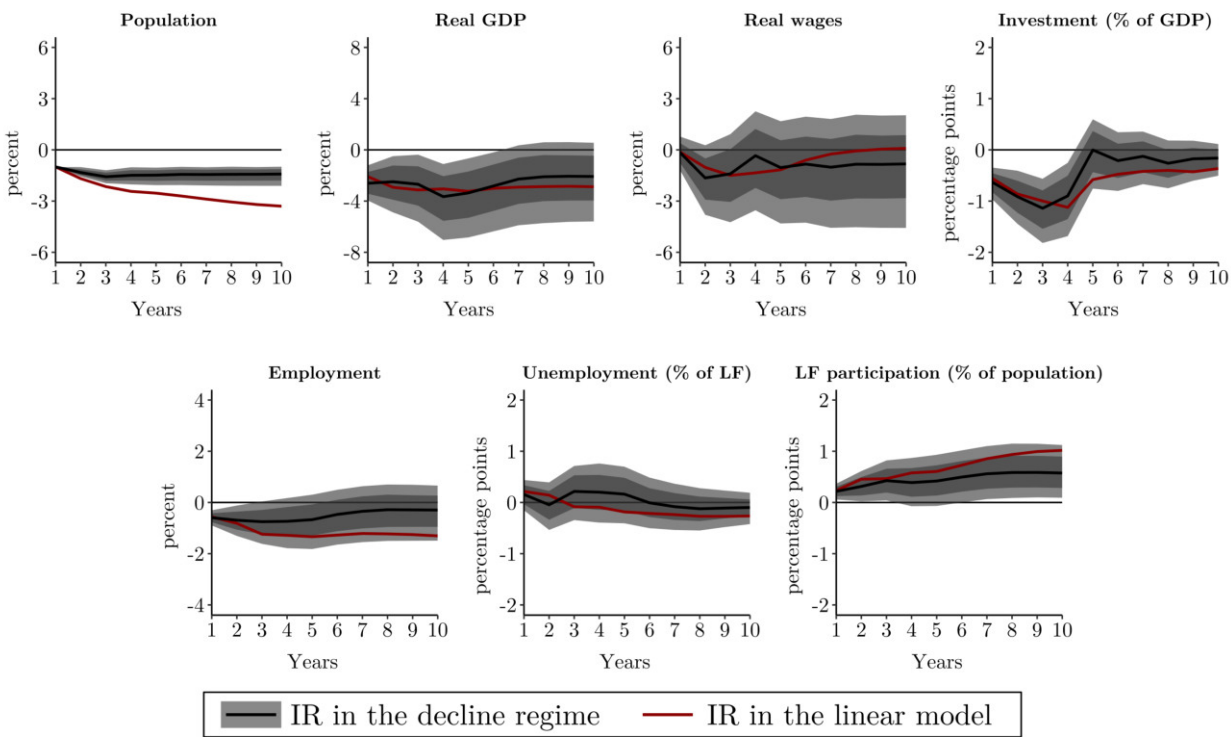


FIGURE 6: IMPULSE RESPONSES IN THE DECLINE REGIME OF THE NON-LINEAR PSTVAR.



the contemporaneous case. The labour force participation shows an initial increase (0.2 percentage points), growing to 0.6 percentage points in the long term. Now, as the reaction of population has flattened out more quickly, the labour force participation rate after  $t_3$  is only driven by dynamics in employment and unemployment.

### **Non-symmetrical effects in times of population growth and decline**

In figure 7, the impulse responses in the decline regime as analyzed above are compared to those in the growth regime. However, as intended to uncover by our estimation strategy and demonstrated in the plots, the effects of population shocks do not unfold in a strictly symmetrical manner across regimes. In both cases, the negative (positive) population shock causes an immediate decline (growth) of GDP, employment, and investment as well as an increase (decrease) of overall labour force participation and the unemployment rate. However, our results indicate that a negative population shock has quantitatively and qualitatively more distinct effects.

The impulse response of the real GDP is more pronounced in the decline regime compared to the upswing in the growth regime (1.3% in the contemporaneous period, 1.1% in  $t_{10}$ ), with the latter being insignificant both on the 68% and 90% level in the long run. In accordance with the economic downturn, employment faces a significant reduction even years after a negative population shock. However, in the long term, it shows a return close to its pre-shock level. These effects are not fully mirrored in the employment impulse response of the growth regime. Here, only the contemporaneous effect is significant on the 90% level, with 0.6% in the median. Additionally, employment sees a permanent, though not significant, increase of 0.8% in the long run – a more distinct impact compared to times of population decline.

The fact that employment largely returns to its previous level may be well explained with increased labour force participation. Put differently, as a reaction to worker shortages, the participation rate surges. Interestingly, in the growth regime, the participation rate decreases more strongly (-0.2 percentage points contemporaneously; -0.9 percentage points in  $t_{10}$ ; statistically significant). This may be explained by the population dynamics following the respective shocks as population continues to increase after a positive population shock and the corresponding economic upswing, but quickly flattens out after a negative shock.

Furthermore, the results suggest that the reduction of unemployment has played a subordinated role in the recovery of employment. The contemporaneous effect on the unemployment rate is symmetric across regimes but only the long term increase in the growth regime is significant (0.2 percentage points; 68% level). The non-significant and quantitatively small effect on the unemployment rate in the decline regime from year one after the shock onwards

is striking. It might be the result of opposing and thus offsetting effects, that is a increase due to the economic downturn and a decrease due to labour force shortage. However, even tough unemployment does not decrease in relative terms, in does so in absolute.

In times of decline, the reduction in investment is contemporaneously more distinct but less permanent in the long run: In the growth regime, the median impulse response reaches a value of 0.4% percentage points in  $t_{10}$ , representing a significant permanent increase compared to a quantitatively smaller and statistically insignificant long-term effect in the decline regime. These results may indicate that firms initially react to labour force shortage by reducing investments but largely return to (relative) pre-shock levels after overall participation, and thus employment, has picked up again.

Following a positive population shock, the economic upswing causes real wages to grow both longer, up to three years after the shock, and to a greater extent, peaking at 2.6% in  $t_4$ , while in the decline regime, the decrease of real wages reverses already two years after shock. Nevertheless, in the long run, the decline (growth) regime impulse responses remain below (above) the respective pre-shock level in a largely symmetrical order. The finding that, in response to a negative shock, wages do not grow but fall opposes economic intuition. Rather, as a shortage indicator, wages would be expected to increase in times of labour force decline.

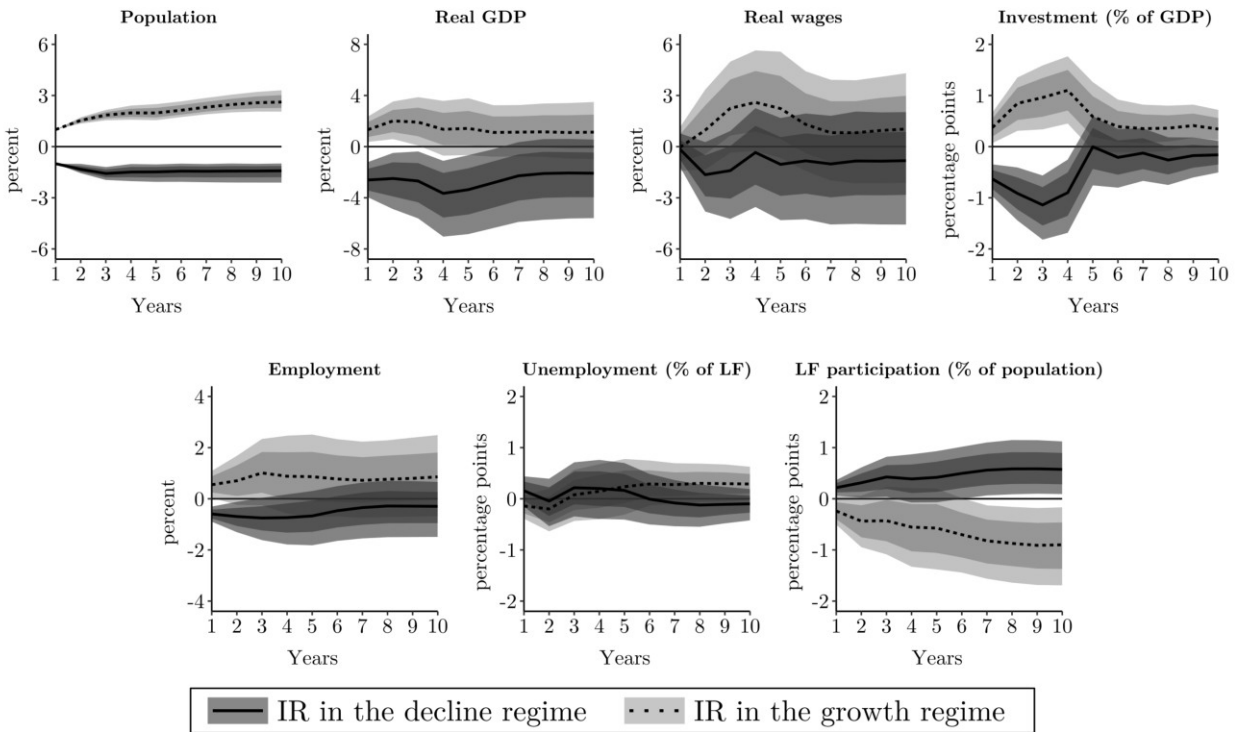


FIGURE 7: IMPULSE RESPONSES IN THE GROWTH AND THE DECLINE REGIME OF THE NON-LINEAR PSTVAR.

However, this effect could have been offset by increased labour force participation as well as adverse effects caused by the economic downturn.

Overall, the visually small but statistically pronounced differences between the impulse responses in both regimes as well as the contrasts to the linear case are crucial and justify an interpretation of regime-dependent adjustments. In other words, both quantity and quality of the contemporaneous reactions as well as the dynamics following the initial downturn respectively upswing unveil that from the short to the medium and long term the labour market reacts differently to population shocks, that is adjustment processes are in fact regime-dependent.

## 5 Robustness

We check the robustness of our estimates using a series of alternative specifications. In particular, we analyze to what extent changes in the identification strategy as well as in the regime-setting may alter the results found in the previous section.

### Migration-adjusted population data

First, as its well-known and widely discussed, demographic variables such as population growth are usually prone to endogeneity. In order to avoid such issues, we remove the possibly endogenous part, migration, from the population growth data beforehand. Figure 8 illustrates the effect of removing positive (green areas) and negative (red areas) net migration from the population growth data (solid line) across countries. Historically, emigration-intensive countries such as the UK, Norway, or Sweden faced long periods of negative net migration. This pattern has reversed in recent decades. Vice versa, Australia and the United States have seen continued immigration from the late 19<sup>th</sup> century until today. By removing the migration effects, we yield an exogenous working-age population growth series being dependent on fertility processes in the far past and, conversely, being independent of short-term economic and labour market effects.

Now, as a robustness check, we use the migration-adjusted population growth data (dotted line; figure 8) instead of the original population growth data. Using different population growth data implies changes in the regime-specification as well. Figure A.2 illustrates the weights on the decline regime using the migration-adjusted data. As before, we derive the impulse responses to a migration-adjusted population shock (figure A.3).

The results largely confirm our findings in the baseline model. Real GDP, investment, and employment face an initial decline but return closely or fully to the pre-shock level in the long term. As in the baseline model, the labour force participation sees a permanent increase.

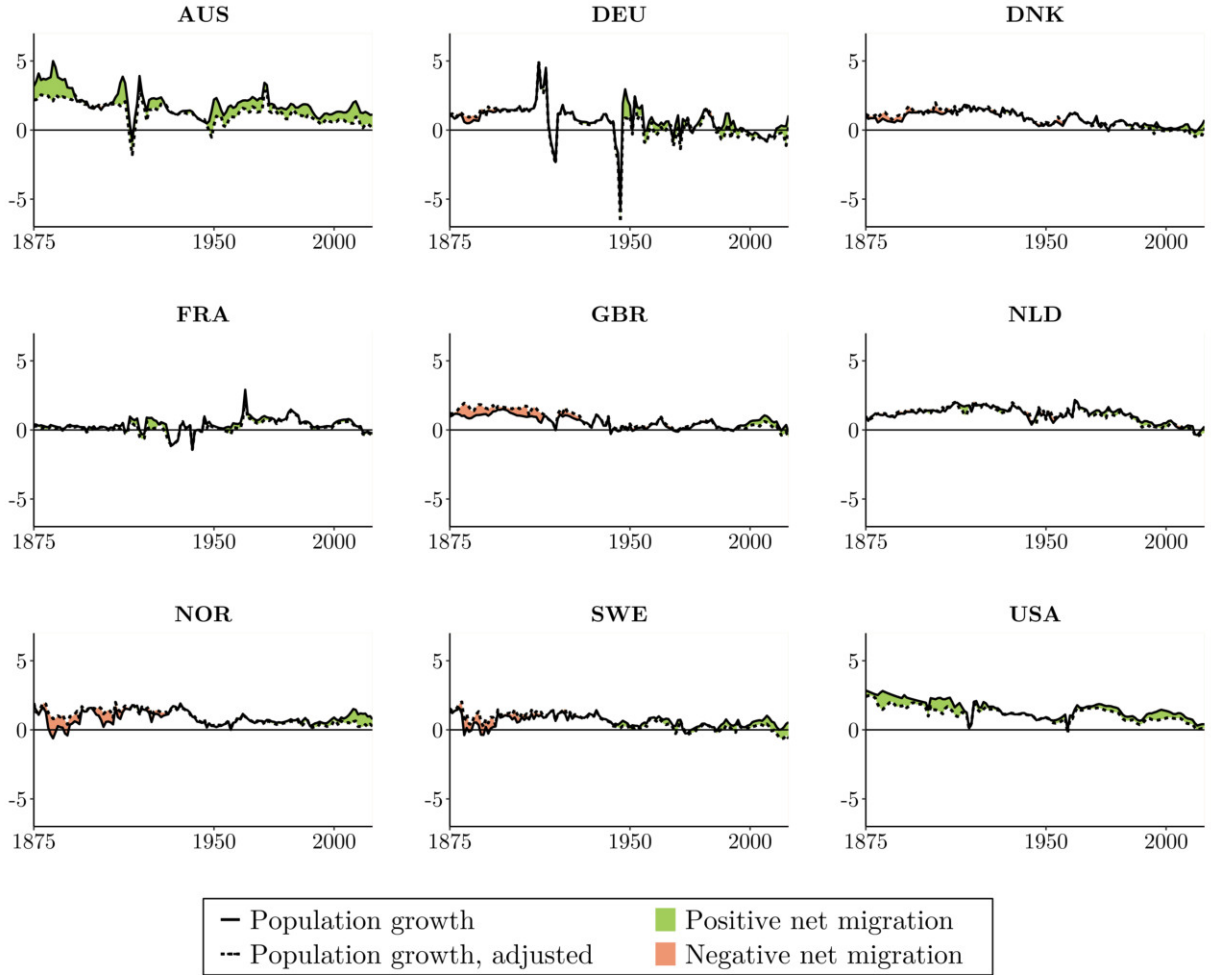


FIGURE 8: POPULATION GROWTH ADJUSTED FOR MIGRATION.

Source: See the appendix.

Slightly different results can be observed for real wages for which the median impulse response also declines in the year after the shock and increases subsequently, however, compared to the estimation results above, the median impulse response nearly recovers to its pre-shock level.

### Regime-specification: transition variable timing and individual smoothing parameter

Second, we evaluate the consequences of a differently timed transition variable. In our baseline specification, we use the decline probability at  $t - 1$ . This procedure follows examples from the literature such as the well-known application in [Auerbach and Gorodnichenko \(2012\)](#). However, as we are including more than one lag of population growth in our estimation, this gives rise to the question of whether  $t - j$  would be a more appropriate timing for

the transition variable. Put differently, our baseline specification assumes that the effects of lagged population growth, as of all other lagged endogenous variables, in each equation is only based on the 'current' demographic regime. Conversely, using  $t - j$  to weight the regressors suggests that the impact is dependent on the demographic regime at lag  $j$ . Thus, we re-estimate our baseline model using this different setup. Formally, we can re-write the model as

$$Y_{it} = \mu_i + \delta_t + [1 - P(q_{i,t-j})]GY_{i,t-j} + [P(q_{i,t-j})]DY_{i,t-j} + u_{it} \quad (8)$$

The weights on the respective regimes remain the same as shown in figure 4, and so do all other specifications from the baseline case. Again, the results show that our model is robust to this alternative strategy (figure A.4). As in the baseline estimation, real GDP, investment, and employment face a decline with a subsequent recovery. Again, employment shows a much smaller long-term decline in response to a negative shock than growth in response to a positive shock. This confirms our finding that the labour market is, in particular, adjusting to population decline by increased labour force participation. In this case, however, the investment share is, in the long run, significantly lower on the 68% level. Notably, in contrast to our baseline results, the median impulse response of real wages strongly increase in the long term, suggesting that real wages are reacting as shortage indicator. However, the result is only statistically significant in the medium but not in the long term.

Third, not only the transition variable but also the parameters governing the transition function as a whole may intervene in the results stemming from our empirical strategy. While the location parameter  $c = 0$  is fixed by design due to our research question focused on population decline, the speed of the transition indicated by  $\gamma$  is more flexible. In our baseline specification, we have found  $\gamma$  across all countries and time periods. However, a different approach might draw on a country-specific values of  $\gamma$ . Such an approach may account for quite different demographic developments and patterns among the sample countries in the past. Thus, as another robustness check, we use the same specification as in the baseline version but find  $\gamma$  individually for each country. In doing so, we exclude the United States from our sample since there is no observation of working-age population decline in the period under investigation.

In re-defining the transition function for country-specific z-standardized indices,  $\gamma$ , being equal to 2.28 in the baseline specification, now spans from 1.58 in the Netherlands to 2.50 in Denmark. With these varying smoothing parameters, also the weights on the decline and growth regimes change. Figure A.5 illustrates the resulting probabilities. Applying the changed regime-specification, keeping the remaining baseline strategy as it is, and deriving impulse response functions demonstrates that our the main findings are robust to this

alternative specification as well (see figure A.6), that is the effects on employment, unemployment, and labour force participation as well as real GDP and investment are confirmed. However, as in the preceding robustness check, the median impulse response of real wages increases in the long run, indicating that wages grow both in response to a positive as well as to a negative shock.

As mentioned before, to account for the reduced sample size and to obtain adequate interpretation, we do not only compare the results of this robustness check to those of the baseline specification as it is but also to the baseline specification re-estimated without the United States. This delivers further insights whether changes in the robustness check emerge due to the exclusion of the United States. As can be seen, the impulse responses are very similar to those obtained from the baseline model (see figures A.7 for slightly changed decline probabilities and A.8 for the estimation results). Thus, the changed results in figure A.6 stem from the changed probabilities as shown by figure A.5.

Overall, the robustness checks confirm our results derived in the initial estimation. In particular, the most distinct regime-dependent adjustment pattern – the labour force participation increases in times of population decline in order to maintain the employment level – is supported by the results in this section. Similarly, the finding that the unemployment rate hardly reacts in response to population shocks is underpinned by the robustness checks as well. As noted, this non-reaction may be explained by offsetting effects – economic decline (increase) and outflow into employment due to shortage (decrease). In any case, the consistent findings on employment, unemployment, and labour force participation show that the latter is driving the recovery of the overall employment level in response to a negative population shock. Similarly, the finding that the investment share nearly or fully recovers after employment has picked up again can be confirmed by the estimation results above as well.

Contrarily to our baseline model, however, all three robustness check show only a minimal decline or even an increase of wages in the long run. Given these apparent differences to the initial results as well as between the individual robustness checks, our empirical assessment does not allow to conclude whether wages increase, as a shortage indicator, or decrease, due the economic downturn, possibly coupled with increased labour supply as a consequence of increased participation, in times of population decline.

## 6 Conclusion

According to recent population projections, most advanced economies will face population decline in the years and decades to come, providing the demographic context in these coun-

tries in the long term. Notably, this decline is expected to be particularly pronounced among those in working-age. Although a decreasing population may have profound economic implications, above all in the labour market, there is still little theoretical and empirical evidence on this issue. We contribute to this sparse body of literature by focusing on the latter. We compiled a new historical dataset using more than 80 different sources, containing information on demographics (population, net migration) and labour market variables (real GDP, investment, real wages, employment, unemployment, labour force participation). Drawing on this extensive dataset, we estimate a panel smooth-transition VAR (PSTVAR) to account for possible non-linear effects of population changes on the economy and the labour market depending on the demographic regime. We derive regime-dependent impulse response functions to trace the effects of positive (negative) population shocks in the labour market in times of population growth (decline).

Our findings suggest that effects of population shocks on the economy and the labour market are regime-dependent, that is reactions to population changes in times of decline differ from those in times of growth. More precisely, our results indicate that, initially, a negative population shock causes an economic downturn and, vice versa, a positive population shock causes an upswing. But, however, these effects are quantitatively and qualitatively more distinct in the decline compared to the growth regime, and tend to differ in the long run. Most notably, regime-dependent adjustment processes take place in overall employment and labour force participation. More precisely, in order to cope with labour supply shortages and maintain the employment level, the labour force participation increases. Thus, the negative effects on employment are less severe in times of population decline compared to the increase of employment in periods of growth. Contrarily, we neither find evidence of a decline in the unemployment rate as a response of the labour supply shortage nor for an increase as response to the economic downturn. However, we find slightly asymmetric effects on investment, that is the share in overall GDP declines less strongly in the aftermath of a negative population shock than it increases following a positive population shock. Put differently, based on our non-linear framework, investments recover more strongly in the aftermath of a negative population shock than suggested by linear model results, likely after employment figures have picked up again. Eventually, we do not find clear evidence on effects on wages. More specifically, our results point towards offsetting effects, that is wages increase as a shortage indicator and decrease due the economic downturn, possibly also coupled with increased labour supply as a consequence of increased participation.

Overall, our findings provide the starting point for different future research avenues. Improvements might encompass a more precise identification strategy, such as instrumental variable shock identification. However, such attempts are, inter alia, confronted with prob-



lems of data availability, as historical macroeconomic data is sparse in general and that of suitable instruments for population growth in particular. Moreover, results derived by empirical studies assessing the effects of population decline and corresponding adjustment processes as conducted in this paper might contribute to future incorporation into macroeconomic growth models.

## References

- Acemoglu, D. & Restrepo, P. (2017). Secular stagnation? The effect of aging on economic growth in the age of automation. *American Economic Review*, 107(5), 174–179.
- Aksoy, Y., Basso, H. S., Smith, R. P., & Grasl, T. (2019). Demographic structure and macroeconomic trends. *American Economic Journal: Macroeconomics*, 11(1), 193–222.
- Auerbach, A. J. & Gorodnichenko, Y. (2012). Measuring the Output Responses to Fiscal Policy. *American Economic Journal: Economic Policy*, 4(2), 1–27.
- Barro, R. J. (1991). Economic growth in a cross section of countries. *Quarterly Journal of Economics*, 106(2), 407–443.
- Barro, R. J. (1998). *Determinants of Economic Growth: A Cross-Country Empirical Study*. Cambridge, MA.
- Barro, R. J. & Lee, J. W. (1994). Sources of economic growth. *Carnegie-Rochester Confer. Series on Public Policy*, 40, 1–46.
- Bergeaud, A., Clette, G., & Lecat, R. (2016). Productivity Trends in Advanced Countries between 1890 and 2012. *Review of Income and Wealth*, 62(3), 420–444.
- Bloom, D. E. & Williamson, J. G. (1998). Demographic transitions and economic miracles in emerging Asia. *World Bank Economic Review*, 12(3), 419–455.
- Bolt, J. & van Zanden, J. L. (2020). The Maddison Project: Maddison style estimates of the evolution of the world economy. A new 2020 update. Maddison-Project Working Paper WP-15: <https://www.rug.nl/ggdc/historicaldevelopment/maddison/publications/wp15.pdf>.
- Börsch-Supan, A. (2008). The Impact of Global Aging on Labor, Product and Capital Markets. *Demographic Change in Germany: The Economic and Fiscal Consequences*, (pp. 65–88).
- Brida, J. G. & Accinelli, E. (2007). The Ramsey model with logistic population growth. *Economics Bulletin*, 3(15), 1–8.
- Canova, F. & Ciccarelli, M. (2013). Panel vector autoregressive models: A survey. *Advances in Econometrics*, 32, 205–246.
- Cervellati, M., Sunde, U., & Zimmermann, K. F. (2017). Demographic dynamics and long-run development: insights for the secular stagnation debate. *Journal of Population Economics*, 30(2), 401–432.
- Christiaans, T. (2011). Semi-endogenous growth when population is decreasing. *Economics Bulletin*, 31(3), 2667–2673.
- Eggertsson, G. B., Lantier, M., & Summers, L. H. (2019). Aging , Output Per Capita , and Secular Stagnation. *American Economic Review: Insights*, 1(3), 325–342.
- Gehrke, B. & Hochmuth, B. (2020). Counteracting Unemployment in Crises: Non-Linear Effects of Short-Time Work Policy. *The Scandinavian Journal of Economics*, 123(1), 144–183.
- González, A., Teräsvirta, T., van Dijk, D., & Yang, Y. (2017). Panel Smooth Transition Regression Models.
- Hall, R. E. & Jones, C. I. (1999). Why Do Some Countries Produce So Much More Output Per Worker Than Others? *The Quarterly Journal of Economics*, 114(1), 83–116.
- Headey, D. D. & Hodge, A. (2009). The effect of population growth on economic growth: A meta-regression analysis of the macroeconomic literature. *Population and Development*

- Review*, 35(2), 221–248.
- IMF (2021). WEO Database. Online (last checked 2021-08-11): <https://www.imf.org/en/Publications/WEO/weo-database/2021/April>.
- Jones, C. I. (2020). The End of Economic Growth? Unintended Consequences of a Declining Population. NBER Working Paper 26651: [https://www.nber.org/system/files/working\\_papers/w26651/w26651.pdf](https://www.nber.org/system/files/working_papers/w26651/w26651.pdf).
- Jones, C. I. & Romer, P. M. (2010). The new Kaldor facts: Ideas, institutions, population, and human capital. *American Economic Journal: Macroeconomics*, 2(1), 224–245.
- Jordà, Ò., Schularick, M., & Taylor, A. M. (2017). Macroeconomic history and the new business cycle facts. *NBER Macroeconomics Annual*, 31, 213–263.
- Kajanovičová, V., Novotný, B., & Pospíšil, M. (2020). Ramsey model with non-constant population growth. *Mathematical Social Sciences*, 104, 40–46.
- Kaldor, N. (1961). Capital Accumulation and Economic Growth. In F. A. Lutz & D. Hague (Eds.), *The Theory of Capital. Proceedings held by the International Economic Association* (pp. 177–222). London: Palgrave Macmillan UK.
- Mitchell, B. R. (2013). *International Historical Statistics. 1750-2010*. London: Palgrave Macmillan Ltd.
- Sasaki, H. (2019). The Solow growth model with a CES production function and declining population. *Economics Bulletin*, 39(3), 1979–1988.
- Sasaki, H. & Hoshida, K. (2017). The effects of negative population growth: An analysis using a semiendogenous R&D growth model. *Macroeconomic Dynamics*, 21(7), 1545–1560.
- Sims, C. A., Stock, J. H., & Watson, M. W. (1990). Inference in linear time series models with some unit roots. *Econometrica*, 58(1), 113–144.
- UN (2019). World Population Prospect 2019. Online (last checked 2020-05-20): <https://population.un.org/wpp/>.

## Appendix A Additional figures

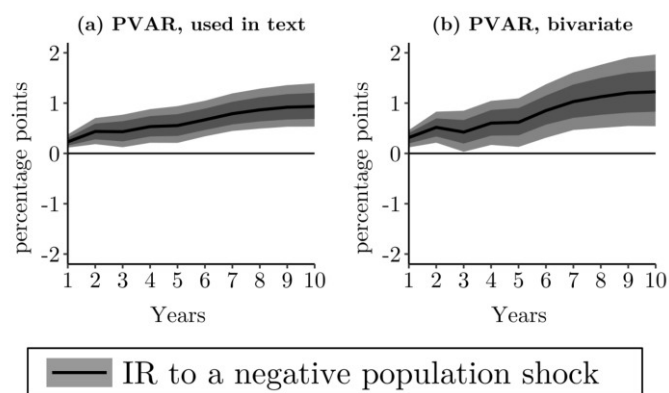


FIGURE A.1: COMPARISON OF (A) THE LABOUR FORCE PARTICIPATION RATE AS INDIRECTLY DERIVED FROM THE LINEAR PVAR TO (B) THE LABOUR FORCE PARTICIPATION RATE DERIVED IN A SIMPLE BIVARIATE PVAR.

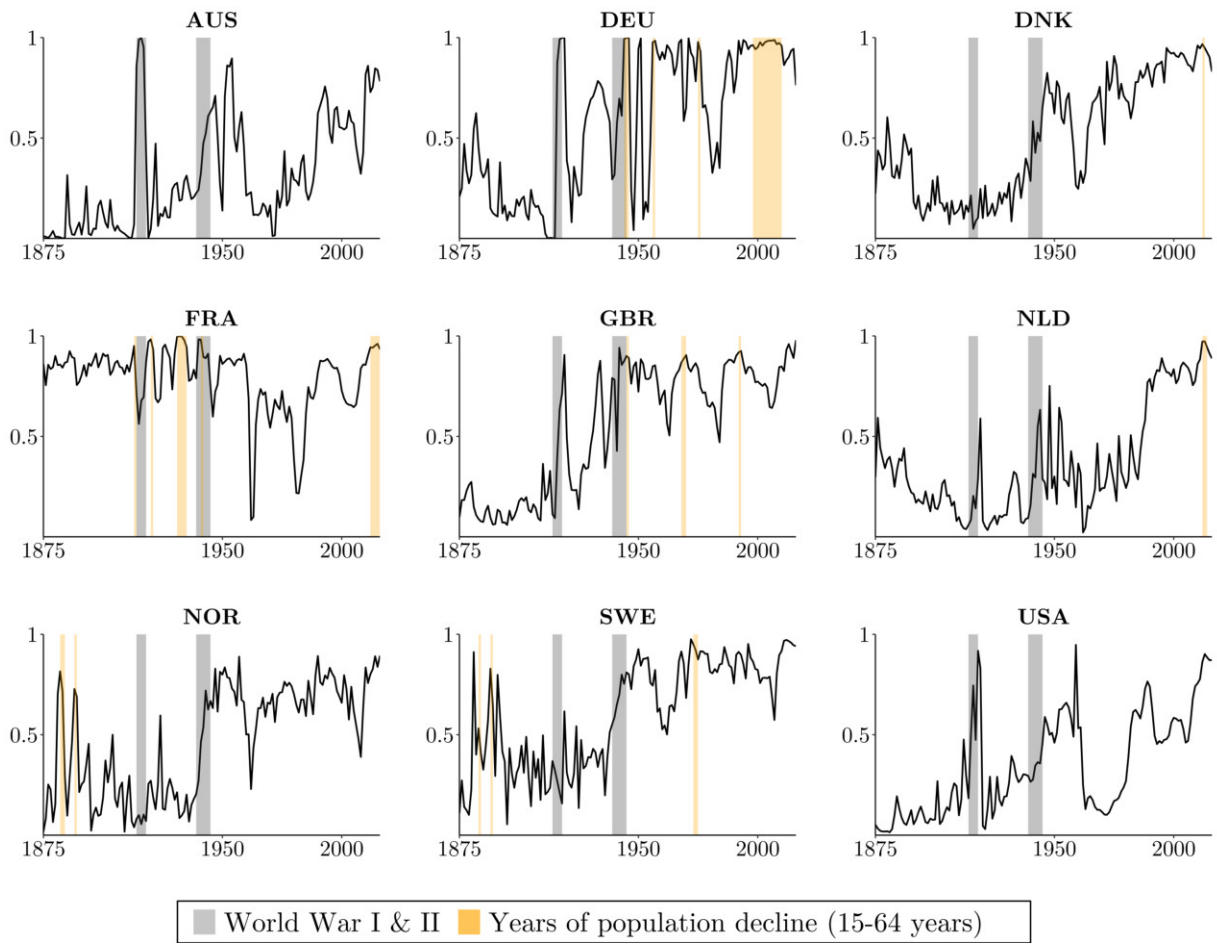


FIGURE A.2: WEIGHTS ON THE DECLINE REGIME IN THE MIGRATION-ADJUSTED MODEL (POINT ESTIMATION).

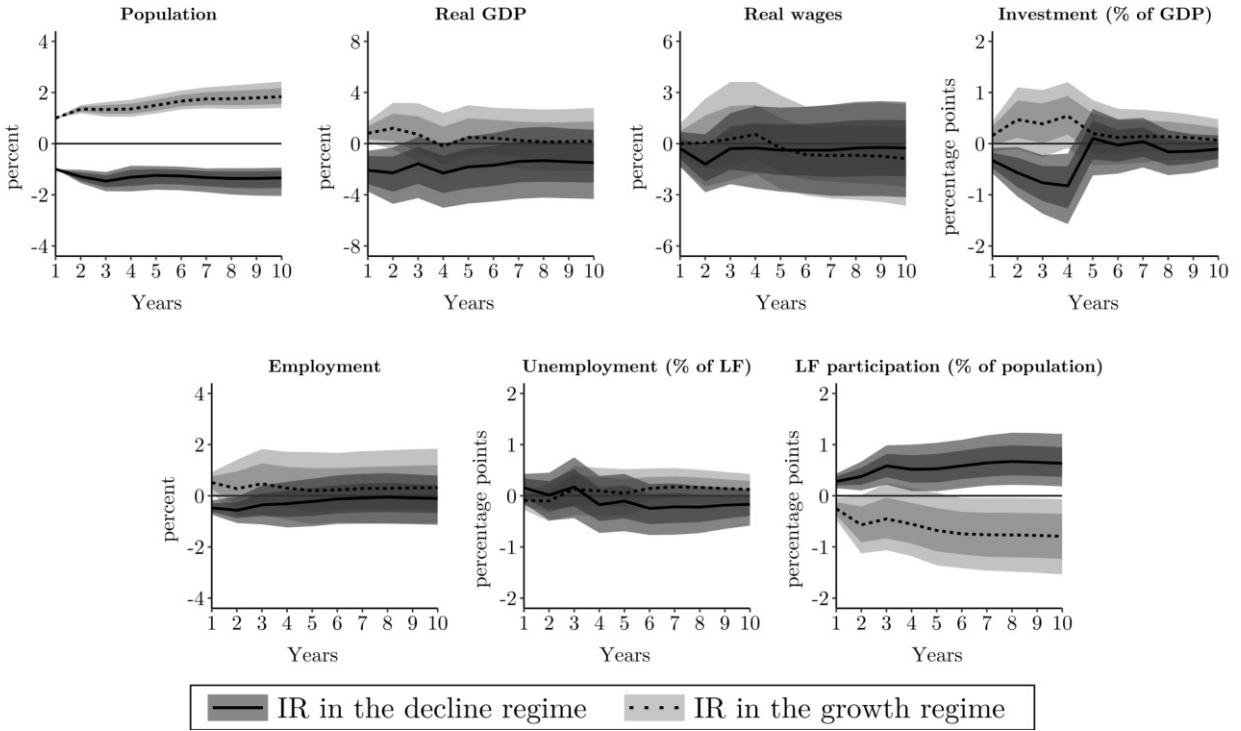


FIGURE A.3: REGIME-DEPENDENT IMPULSE RESPONSE FUNCTIONS TO A ONE-PERCENT POPULATION SHOCK IN THE MIGRATION-ADJUSTED MODEL.

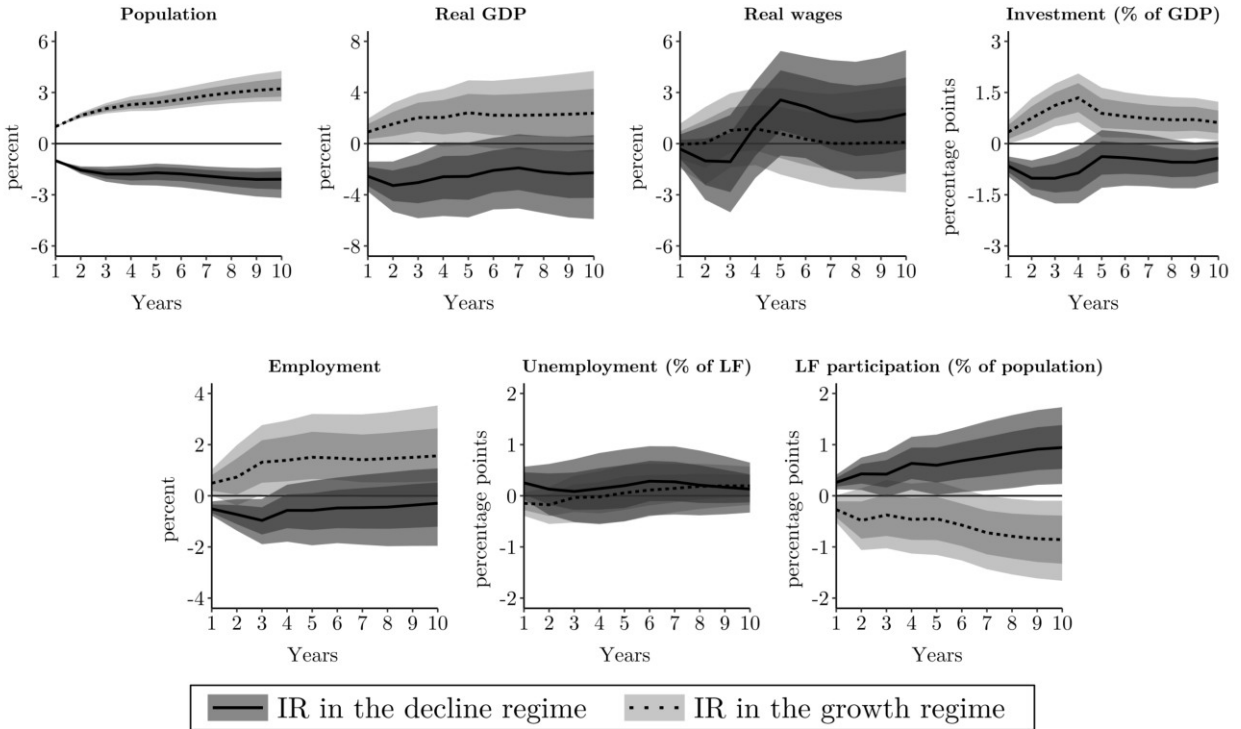


FIGURE A.4: REGIME-DEPENDENT IMPULSE RESPONSE FUNCTIONS TO A ONE-PERCENT POPULATION SHOCK IN THE TIMING-ADJUSTED MODEL.



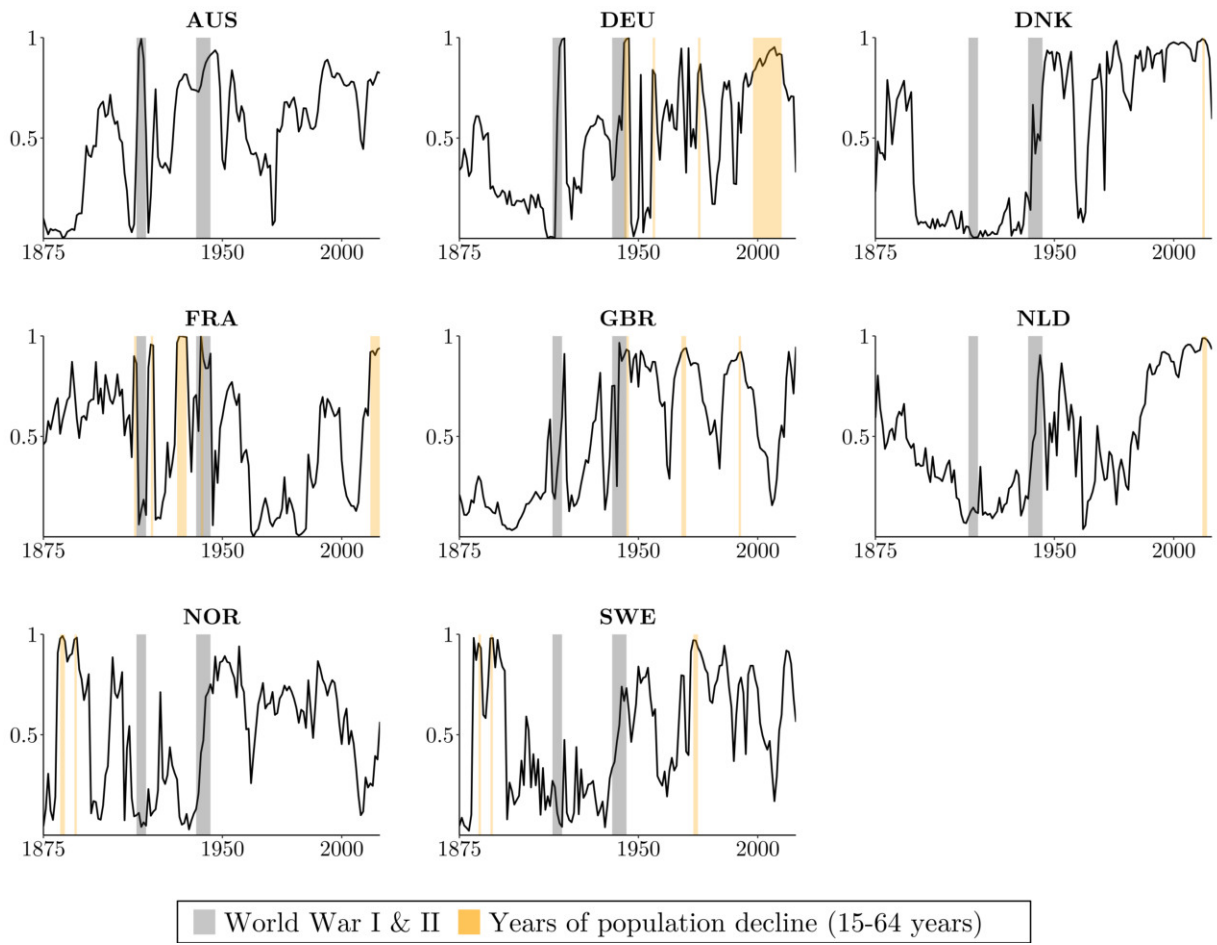


FIGURE A.5: WEIGHTS ON THE DECLINE REGIME IN THE MODEL WITH ADJUSTED, I.E. COUNTRY-SPECIFIC, SMOOTHING PARAMETER  $\gamma$  (POINT ESTIMATION; USA EXCLUDED).

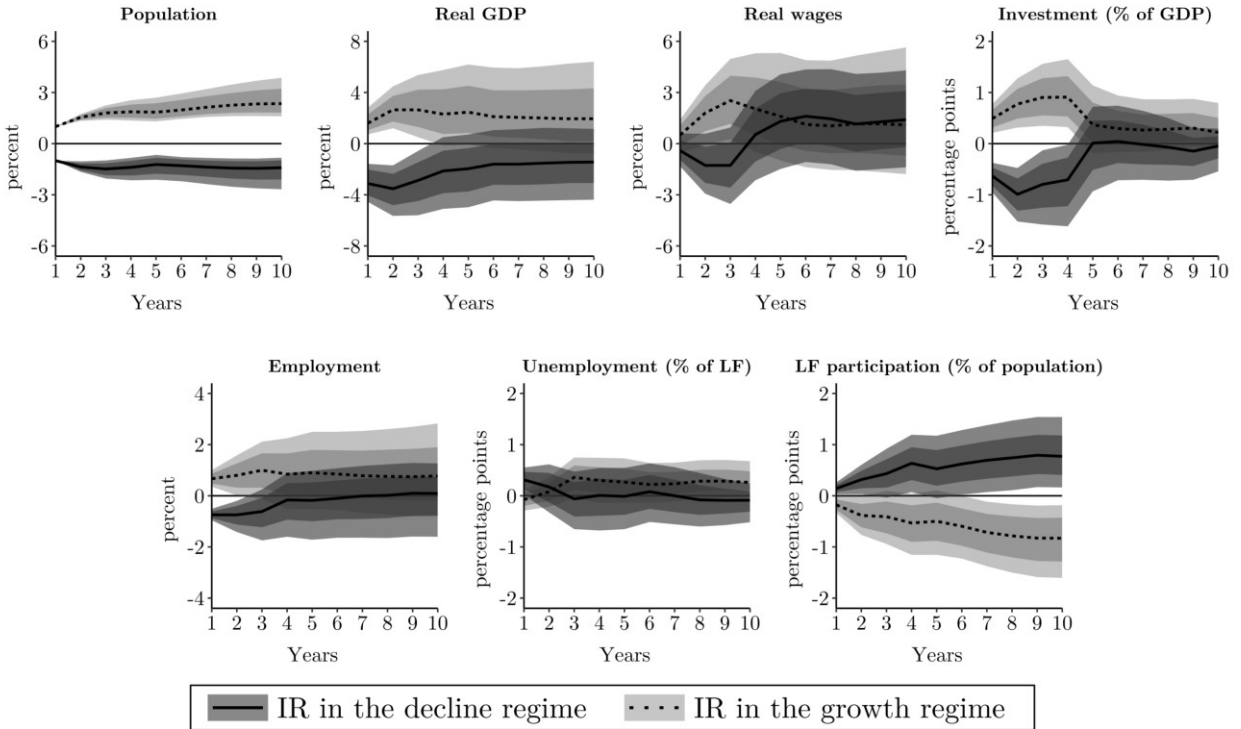


FIGURE A.6: REGIME-DEPENDENT IMPULSE RESPONSE FUNCTIONS TO A ONE-PERCENT POPULATION SHOCK IN THE MODEL WITH ADJUSTED, I.E. COUNTRY-SPECIFIC, SMOOTHING PARAMETER  $\gamma$  (USA EXCLUDED)

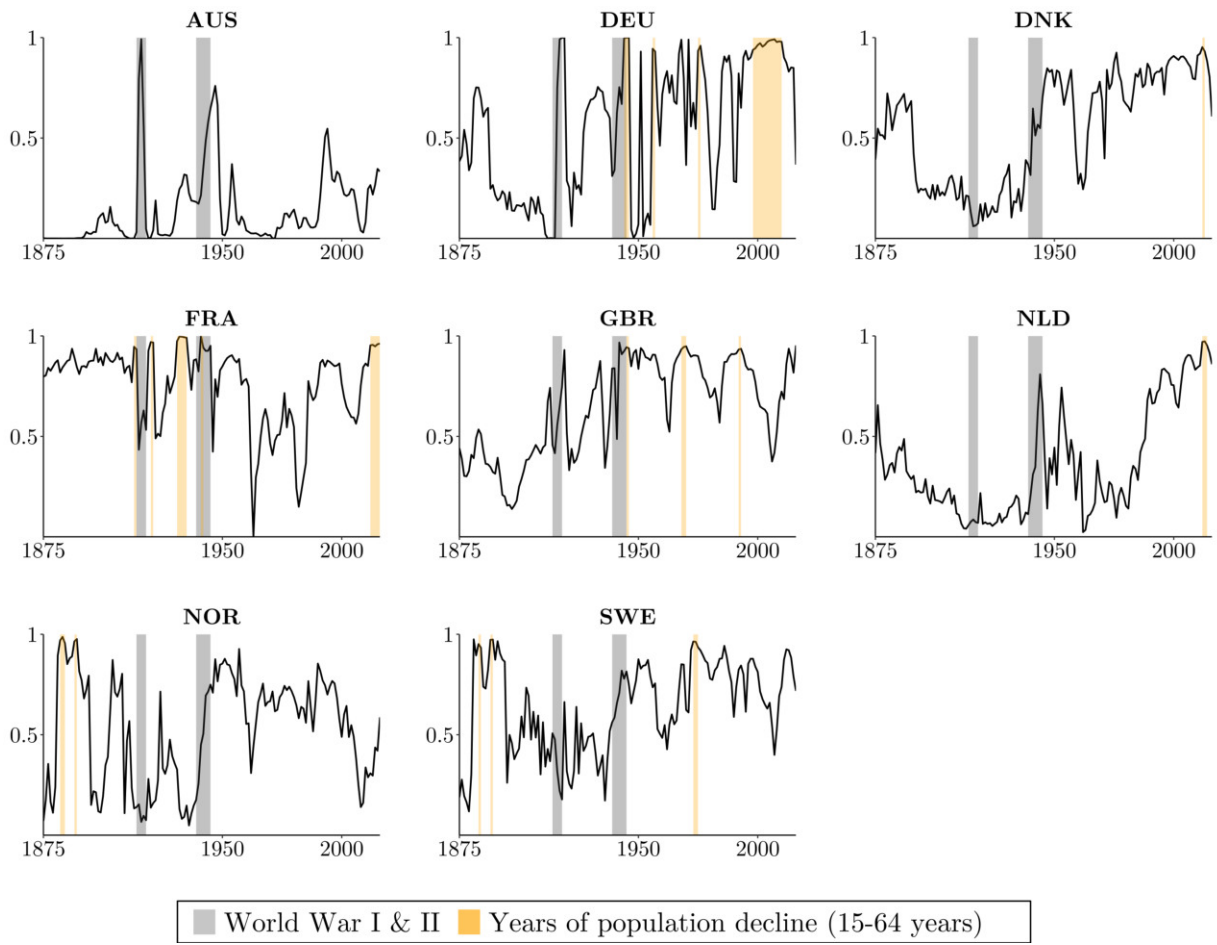


FIGURE A.7: WEIGHTS ON THE DECLINE REGIME IN THE BASELINE MODEL (POINT ESTIMATION; USA EXCLUDED).

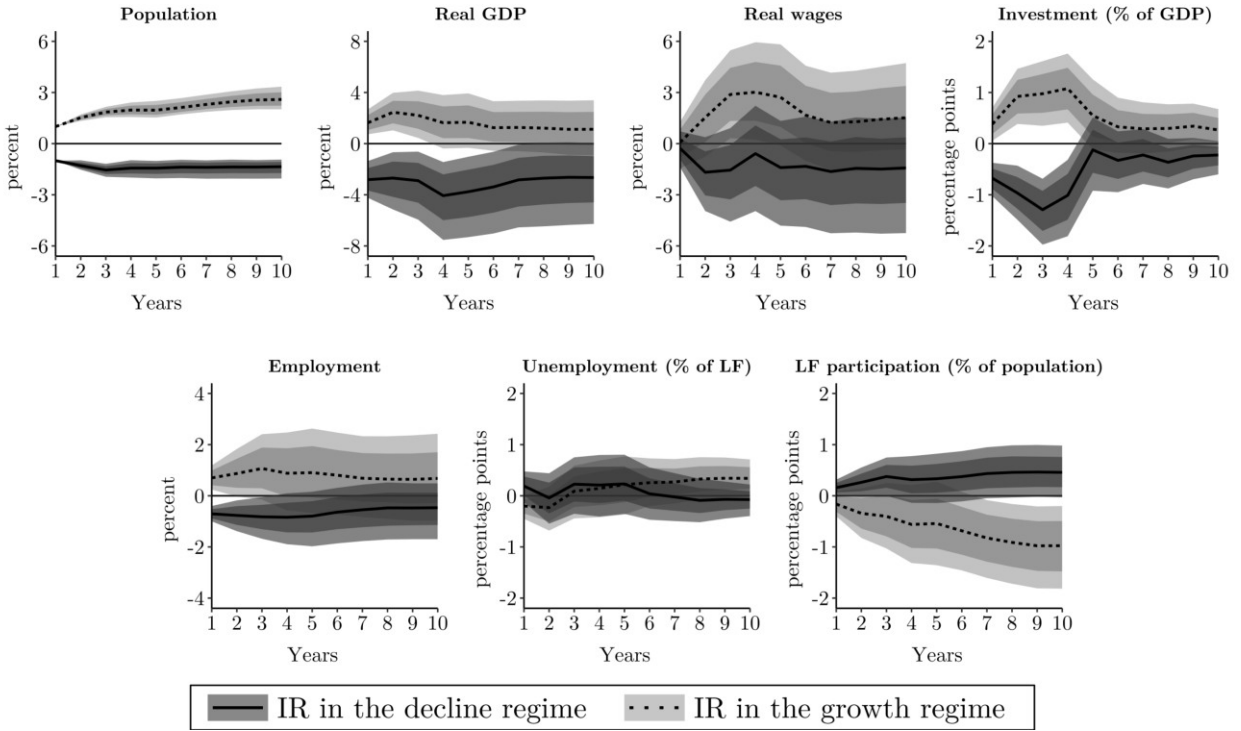


FIGURE A.8: REGIME-DEPENDENT IMPULSE RESPONSE FUNCTIONS TO A ONE-PERCENT POPULATION SHOCK IN THE BASELINE MODEL (USA EXCLUDED).

## Appendix B Dataset

This section gives a detailed description of the variables and the respective data sources.

### Population (total and age groups), index

Total population and population in the respective age group as of 1<sup>st</sup> January. Population refers to civilian and military wherever possible. If population data was available on another date (e.g. mid-year), we followed the Human Mortality Database (HMD hereafter) methodology and estimated the population as of 1<sup>st</sup> January by linear interpolation. Boundary changes were accounted for by chain-linking the indices. See footnotes 4 and 5 in appendix B.

**Australia** 1860 – 1861: used relative changes from [Bolt and van Zanden \(2020\)](#) to backwards extend the subsequent source; 1862 – 1921: annual figures estimated, see appendix B; 1921 – 2016: taken from [HMD \(2020a\)](#)

**Denmark** 1860 – 2016: [HMD \(2020a\)](#)

**France** 1860 – 1911: [HMD \(2020a\)](#), total national population (civilian and military); 1911 – 1921: annual figures estimated, see appendix B; 1921 – 2016: [HMD \(2020a\)](#), total national population (civilian and military)

**Germany** 1860 – 1871: used relative changes from [Bolt and van Zanden \(2020\)](#) to backwards extend the subsequent source; 1872 – 1910: original source is [Lösch \(1936\)](#), data largely taken from the Gesis database, some adjustments (error correction) using the original source; 1911 – 1949: annual figures estimated, see appendix B; 1950 – 1955: [Sensch \(2007\)](#); 1956 – 2016: [HMD \(2020a\)](#); up to 1990 data for West Germany, overall Germany subsequently

**Netherlands** 1860 – 2016: [HMD \(2020a\)](#)

**Norway** 1860 – 2016: [HMD \(2020a\)](#)

**Sweden** 1860 – 2016: [HMD \(2020a\)](#)

**United Kingdom** 1860 – 1922: data for Great Britain from [HMD \(2020a\)](#), additionally estimated Southern and Northern Ireland figures, see appendix B; 1922 – 2016: [HMD \(2020a\)](#), refers to United Kingdom without Southern Ireland from here onwards

**United States** 1860 – 1900: annual figures estimated, see appendix B; 1901 – 1979: [U.S. Bureau of the Census \(2000\)](#); 1979 – 2016: [HMD \(2020a\)](#)

### Net migration, rate

Net migration as a rate in the respective year.

**Australia** 1860 – 2015: [ABS \(2019\)](#), linear interpolation of emigration figures for 1914 to 1919 to account for departures of military personnel; 2016: [ABS \(2021\)](#)

**Denmark** 1860 – 1944: [Flora et al. \(1987\)](#), linear interpolation of missing data in 1921; 1945 – 1963: [DEMIG \(2015\)](#); 1964: [Flora et al. \(1987\)](#); 1965 – 1979: [DEMIG \(2015\)](#); 1980 – 2016: [Statistics Denmark \(2021\)](#)

**France** 1860 – 1945: [Flora et al. \(1987\)](#), linear interpolation of missing data for 1861, 1871, 1914-1918, 1940-1945; 1946 – 1961: [INSEE \(2020\)](#); 1962: [Mitchell \(2013\)](#); 1963 – 2016: [INSEE \(2020\)](#)

**Germany** 1860 – 1938: [Flora et al. \(1987\)](#), linear interpolation of missing data for 1919 and 1935-1936; 1939 – 1949: linear extrapolation of missing data 1939-1943 based on change of 1937-1938; subsequently, using the obtained value for 1943, linear interpolation to obtain missing data for 1944-1946; this adjustment mirrors the large scale inflow of migrants from 1944 onwards; 1947 – 1949: [Flora et al. \(1987\)](#); 1950 – 1969: difference of inflows and outflows from [DEMIG \(2015\)](#); 1970 – 2016: [StBA \(2019\)](#), table 1.1

**Netherlands** 1860 – 1863: [Flora et al. \(1987\)](#); 1864: linear interpolation; 1865 – 1928: [Mitchell \(2013\)](#); 1929 – 1994: [DEMIG \(2015\)](#); 1995 – 2016: [CBS \(2019\)](#)

**Norway** 1860 – 1957: [Flora et al. \(1987\)](#); 1958 – 2016: [SSB \(2021\)](#)

**Sweden** 1860 – 1874: [Flora et al. \(1987\)](#); 1875 – 2016: [SCB \(2021\)](#)

**United Kingdom** 1860 – 2016: [Bank of England \(2017\)](#)

**United States** 1860 – 1867: used constant 1868 value; 1868 – 1949: first, we used emigration numbers among departures of foreigners presented in [Axelrod \(1972\)](#), page 41 for 1908 to 1925; we chain-linked the 1908 figures with the time series on departures among foreigners as also presented in [Axelrod \(1972\)](#), page 39 for 1899-1908; similarly, we linked chain-linked the emigrant data from 1925 until 1950 with foreigner departures as presented in [U.S. Bureau of the Census \(1975\)](#), page 119, series 301; in both cases, we assume that the share of emigrants among total departures of foreigners remains the same as in 1908 respectively 1925; 1950 – 2016: [UN \(2019\)](#)

## Real GDP, index

Real GDP in PPP per capita as included in the Maddison database ([Bolt and van Zanden 2020](#)). Obviously, the underlying population here differs from that used in the demography variables above. In general, the RGDP per capita data was taken from Maddison due to its internal consistency, which enables using growth rates without caring about breaks due to boundary changes. Based upon the Maddison data, we construct a total real GDP series using the population indices above. However, compared to the Maddison data, we replace data in 1938-1948 in Belgium, Denmark, France, Netherlands, and Norway from [Klemann and Kudryashov \(2012\)](#) to account for common underestimation in GDP statistics during German occupation. For Germany, to achieve a consistent underlying territory across

variables and account more realistically for economic consequences at the end as well as in the aftermath of the Second World War, we use Maddison data up to 1945, GDP data by [Ritschl and Spoerer \(1997\)](#) for 1946-1950 (excluding Saarland and Western Berlin), GDP data by [Deutsche Bundesbank \(2020\)](#) for 1950-1991, and again Maddison for 1991-2016. We chain linked the data in 1945 with the index by [Barro and Ursúa \(2010\)](#).

### **Investment, share of GDP**

Basic source is the Macrohistory database ([Jordà et al. 2017](#)). Replaced data and filled missing years using the following additional sources

**Australia** 1947 – 1948: filled missing ratio in [Jordà et al. \(2017\)](#) with sum of investment figures and gross domestic product, both in current prices, from [Butlin \(1977\)](#), page 79; preceding and subsequent values in [Jordà et al. \(2017\)](#) are consistent with this source

**Denmark** 1914 – 1922: filled missing values by total gross investment in current prices and gross domestic product in current prices from [Abildgren \(2017\)](#), series S006A and S046A

**France** 1919: filled missing data in [Jordà et al. \(2017\)](#) by using investment (series IE, IM, IZG) and GDP (series PIBVAL) data from [Villa \(1993\)](#), following the methodology of [van Meerten \(2003\)](#); 1920: corrected entry using original source ([van Meerten 2003](#)); 1924: apparently, the data in the MHD does not refer to [van Meerten 2003](#); however, to avoid the break between the series used in the MHD, we chain-link the real investment index obtained as outlined above by using the 1924 value from [van Meerten \(2003\)](#); 1945: filled missing data in [Jordà et al. \(2017\)](#) by using investment (series IE, IM, IZG) and GDP (series PIBVAL) data from [Villa \(1993\)](#), following the methodology of [van Meerten \(2003\)](#); due to missing IG data, we assumed the same share of IG in investment as in 1939

**Germany** 1913 – 1920: indexed real total investment (construction/‘Bauinvestitionen’ and equipment/‘Ausrüstungsinvestitionen’; equipment contains only selected categories) from [Kirner \(1968\)](#) to fill missing years; refers to the territory of FRG including Saarland and Western Berlin; since the availability of data on equipment investment increases over time (more categories are included), we weight the figures in each year with the shares of the respective sector in 1930, which is the first year in which all sector of relevant size are listed by [Kirner \(1968\)](#); 1920 – 1959: indexed real total gross non-residential investment from [Maddison \(1994\)](#), table 8b, to fill missing years; calculated as per capita values using the population sources as indicated above; relies on [Kirner \(1968\)](#)

**Netherlands** 1914 – 1920: indexed real total investment from [Groote et al. \(1996\)](#) to the fill the missing years; 1940 – 1947: same as for the previous period

**Norway** 1940 – 1945: investment share from [Grytten \(2004\)](#); retrieved as updated series from [Norges Bank \(2018\)](#), table p1\_c6\_table\_6

**United States** 1929 – 1945: investment share calculated as non-residential investment less defence expenditures as given by [BEA \(2020\)](#)

### **Real wages, index**

Builds strongly on [Williamson \(1995\)](#) and his well-known compilation of indexed real wages for urban unskilled workers until the second half of the 20<sup>th</sup> century. For countries and years not included in [Williamson \(1995\)](#), the data sources and definitions are stated in detail below.

**Australia** 1875 – 1984: [Williamson \(1995\)](#); 1985 – 2016: hourly earnings in manufacturing, taken as nominal wages from [OECD \(2019\)](#) and deflated using CPI data from [OECD \(2020\)](#)

**Denmark** 1875 – 1913: [Williamson \(1995\)](#); 1913 – 1917: calculated using average hourly earnings in manufacturing, nominal from [Abildgren \(2017\)](#), series S172A, and deflated using CPI data from [Jordà et al. \(2017\)](#); 1917 – 1988: [Williamson \(1995\)](#); 1988 – 2016: hourly earnings in manufacturing, taken as nominal wages from [OECD \(2019\)](#) and deflated using CPI data from [OECD \(2020\)](#)

**France** 1875 – 1988: [Williamson \(1995\)](#); 1988 – 2016: hourly earnings in manufacturing, taken as nominal wages from [OECD \(2019\)](#) and deflated using CPI data from [OECD \(2020\)](#)

**Germany** 1875 – 1918: [Williamson \(1995\)](#); 1918 – 1924: real weekly wages, original source is [Pierenkemper \(1984\)](#), data taken from [Sensch \(2012\)](#); 1924 – 1943: [Williamson \(1995\)](#); 1944 – 1948: used annual growth rates of gross earnings after social security payments to calculate real wages, original source is [Schewe and Nordhorn \(1967\)](#), data taken from [Müller and Skiba \(2008\)](#); 1948 – 1988: [Williamson \(1995\)](#); 1988 – 2016: hourly earnings in manufacturing, taken as nominal wages from [OECD \(2019\)](#) and deflated using CPI data from [OECD \(2020\)](#)

**Netherlands** 1875 – 1988: [Williamson \(1995\)](#); 1988 – 2016: hourly earnings in manufacturing, taken as nominal wages from [OECD \(2019\)](#) and deflated using CPI data from [OECD \(2020\)](#)

**Norway** 1875 – 1941: [Williamson \(1995\)](#); 1941 – 1944: annual real wages in manufacturing, taken as nominal wages from [Grytten \(2007\)](#), table 6.A.1, page 370, and deflated using CPI data from [Jordà et al. \(2017\)](#); 1945 – 1988: [Williamson \(1995\)](#); 1988 – 2016: hourly earnings in manufacturing, taken as nominal wages from [OECD \(2019\)](#) and deflated using CPI data from [OECD \(2020\)](#)

**Sweden** 1875 – 1988: [Williamson \(1995\)](#); 1988 – 2016: hourly earnings in manufacturing, taken as nominal wages from [OECD \(2019\)](#) and deflated using CPI data from [OECD \(2020\)](#)

**United Kingdom** 1875 – 1988: [Williamson \(1995\)](#); earlier data refers to Great Britain; 1988 – 2016: hourly earnings in manufacturing, taken as nominal wages from [OECD \(2019\)](#)



and deflated using CPI data from [OECD \(2020\)](#)

**United States** 1875 – 1988: [Williamson \(1995\)](#); 1988 – 2016: hourly earnings in manufacturing, taken as nominal wages from [OECD \(2019\)](#) and deflated using CPI data from [OECD \(2020\)](#)

### **Employment (total), index**

Since not all sources offer coverage up to 2016, consistent employment data from Penn World Table ([Feenstra et al. 2015](#)) from 1950 onwards except otherwise stated.

**Australia** 1875 – 1950: total employment from [Whiters et al. \(1985\)](#), pages 203-205, column C; 1950 – 2016: [Feenstra et al. \(2015\)](#)

**Denmark** 1875 – 1900: total employment obtained by using labour force from [Abildgren \(2017\)](#), series S062A, and unemployment rate as outlined below; 1900 – 1950: total employment from [Abildgren \(2017\)](#), series S079A; 1950 – 2016: [Feenstra et al. \(2015\)](#)

**France** 1875 – 1890: we took the decadal figures by [Clark \(1957\)](#), page 125, and linearly interpolated; 1890 – 1950: total employment from [Villa \(1993\)](#), series EMP, online source; unclear whether annual average; for 1914-1918 and 1940-1945, series EMPE, which contains total firm employment, has been used instead of linear interpolation to account for movement in employment to during war years; 1950 – 2016: [Feenstra et al. \(2015\)](#)

**Germany** 1875 – 1880: total employment, original source is [Hoffmann \(1965\)](#), data obtained from [Ralf \(2015\)](#), series x0457, online source; 1880 – 1939: total employment, original source is [Hoffmann \(1965\)](#), data (with filled gaps) obtained from [Sommariva and Tullio \(1987\)](#), pages 234-236; 1939 – 1944: total employment from [Hohls \(1991\)](#), data retrieved from [Hohls and Kaelbe \(1989\)](#); 1944 – 1950: total employment, original source is [Hoffmann \(1965\)](#), data (with filled gaps) obtained from [Sommariva and Tullio \(1987\)](#), pages 234-236; to account for the structural break 1945/1946, we assumed that the total employment developed just like the population aged 15 to 64 years; 1950 – 1991: total employment from [Deutsche Bundesbank \(2020\)](#); 1991 – 2016: [Feenstra et al. \(2015\)](#)

**Netherlands** 1875 – 1950: total employment from [CBS \(2014\)](#), series 71882eng; 1950 – 2016: [Feenstra et al. \(2015\)](#)

**Norway** 1875 – 1900: we obtain total employment as the residual figures from labour force and unemployment rate as outlined below; 1900 – 1950: total employment from [SSB \(2020\)](#); 1950 – 2016: [Feenstra et al. \(2015\)](#)

**Sweden** 1875 – 1950: total employment from [Schön and Krantz \(2012\)](#); 1950 – 2016: [Feenstra et al. \(2015\)](#)

**United Kingdom** 1875 – 1950: total employment from [Bank of England \(2017\)](#); break-adjusted series of original values referring to United Kingdom including the later Republic

of Ireland (percentual change in the years prior to 1920 identical for both U.K. with and without Republic of Ireland; used series for both countries since unemployment refers to the same); 1950 – 2016: [Feenstra et al. \(2015\)](#)

**United States** 1875 – 1899: total employment from [Vernon \(1994\)](#), page 710; 1900 – 1947: total employment from [U.S. Bureau of the Census \(1975\)](#), calculated by subtracting unemployed persons from total labour force, series D1-10, columns 1 and 8, page 126; 1947 – 1950: total employment from [U.S. Bureau of the Census \(1975\)](#), calculated by subtracting unemployed persons from total labour force, series D11-25, columns 12 and 18, page 127; 1950 – 2016: [Feenstra et al. \(2015\)](#)

### **Unemployment, rate**

Although desirable, only few sources of employment figures above also offered information on unemployment. Wherever possible, the sources refer to unemployment among total labour force. However, particularly earlier sources often contain trade union figures. Whenever applicable and appropriate, we follow methodologies of well-known sources to ensure the highest degree of consistency of historical unemployment data within countries. Details are given in every case below. More recent and consistent data is taken from [OECD \(2016\)](#), in some cases also [OECD \(2005\)](#).

**Australia** 1875 – 1963: unemployment rate from [Whiters et al. \(1985\)](#), pages 203-205, column A; 1964 – 2016: [OECD \(2016\)](#)

**Denmark** 1875 – 1899: unemployment rate from [Abildgren \(2005\)](#); we used the relative change to extend the subsequent series; 1900 – 1959: unemployment rate from [Abildgren \(2017\)](#), S041A; 1960 – 1968: [OECD \(2005\)](#); 1969 – 2016: [OECD \(2016\)](#)

**France** 1875 – 1894: we took the decadal figures by [Clark \(1957\)](#), page 125, and linearly interpolated; 1894 – 1950 unemployment rate, calculated using unemployment numbers and total employment from [Villa \(1993\)](#); 1950 – 1960: unemployment rate, original source is OECD Labour Force Statistics, data taken from [Maddison \(1982\)](#), page 207; 1960 – 2016: [OECD \(2016\)](#)

**Germany** 1875 – 1886: we linearly interpolate the labour force between 1882 and 1895 by [Stockmann et al. \(1982\)](#) and extrapolate backwards to 1875 by using the annual growth rate of the same period; then, we obtain unemployment figures by subtracting the employment as outlined above; 1887 – 1913: unemployment rate, original sources are [Kuczynski \(1962\)](#) and [Kuczynski \(1967\)](#), modified by [Pierenkemper \(1987\)](#) and tabulated by [Hohls \(1991\)](#), data taken from [Sensch \(2016\)](#); for 1887-1902, the data refers to Kuczynskis estimations of unemployment among industry workers; for 1903-1913, the data refers to unemployment in trade unions; 1913 – 1921: unemployment rate by [Mitchell \(2013\)](#), data refers to unemployment

in trade unions; 1922 – 1940: unemployment rate by [Petzina et al. \(1978\)](#), data refers to unemployment rate among dependent employees, data taken from [Sensch \(2016\)](#); 1941 – 1945: for this period, there is no information on unemployment rates available; the unemployment rate by [Petzina et al. \(1978\)](#) shows a value of 0.2 for 1940; we follow the analysis of [Kosche and Bach \(1991\)](#), who suggests, that the labour shortage intensified after 1940, additionally to large numbers of forced labour and recruited foreign workers, and assume that there was negligible unemployment during these times; therefore, we set the unemployment rate to 0 throughout this period; 1946 – 1950: unemployment rate from [Galenson and Zellner \(1957\)](#), pages 455-456; figures are based upon employment exchange statistics; the figure for 1948 only refers to U.S. and U.K. occupation zone (according to [Mitchell 2013](#)); from 1949 onwards, the figures refer to West-Germany; we took the relative changes in this original series and backward extended the subsequent series; 1950 – 1990: unemployment rate from [BA \(2020\)](#); 1991 – 2016: unemployment rate from [BA \(2020\)](#)

**Netherlands** 1875 – 1960: unemployment rate from [CBS \(2014\)](#), series 71882eng; 1960 – 2016: [OECD \(2016\)](#)

**Norway** 1875 – 1887: we prepare the unemployment rate as follows: first, we obtain labour force figures from [Central Bureau of Statistics of Norway \(1966\)](#), page 29, for 1870, 1875, 1880, 1885 and from the same source, page 39, for 1871, 1877, 1887 by multiplying total GDP with GDP per person of labour force; we linearly interpolated the remaining missing years; then, we calculate the ratio of public seeking relief from [Central Bureau of Statistics of Norway \(1966\)](#), page 578, to the total labour force; we use the relative changes of this ratio to backward extend the subsequent series; 1887 – 1900: we obtain the unemployment rate as follows: first, we obtain labour force figures from [Central Bureau of Statistics of Norway \(1966\)](#), page 29, for 1890, 1895, 1900, and for 1887, 1899 from the same source, page 39, by multiplying total GDP with GDP per person of labour force; we linearly interpolated missing years; then we calculated the share of people receiving public assistance for the first time among all people receiving public assistance from the same source, page 578; we use the relative changes of this ratio to backward extend the unemployment rate from the subsequent series; 1900 – 1903: we use employment figures as outlined above and the same procedure to backward extend the subsequent unemployment rate series as for the period 1887-1900; 1903: estimated by multiplying 1904 unemployment rate as below using the 1903-to-1904-ratio of seven-months-average of union unemployment from [NOS \(1949\)](#), page 363; 1904 – 1920: unemployment rate by [Grytten \(1994\)](#), page 325; 1920 – 1939: unemployment rate by [Grytten \(2008\)](#); 1940 – 1949: unemployment rate, calculated as: (1) using absolute numbers of unemployed persons (official unemployment statistics as listed in [SSB \(1945\)](#), pages 231 and 277, value for 1940 refers to average of October to December) for 1935 – 1943, and

SSB (1950) for 1944 – 1949, page 117); (2) total employment figures as indicated above; (3) calculating unemployment rate from (1) and (2); (4) extrapolating the relation of the rate calculated in (3) and Grytten (1994) series from 1935 – 1939, and multiplying the results with (3); 1950 – 1960: unemployment rate, original source is OECD Labour Force Statistics, data taken from Maddison (1982), page 207; 1960 – 2016: OECD (2016)

**Sweden** 1875 – 1911: unemployment rate from Clark (1957); 1911 – 1920: backward extending the unemployment rate in 1920 from Grytten (2008) by using index number of jobseekers per 100 vacancies at the public labour exchanges from Molinder (2018) and Bengtsson and Molinder (2017); 1920 – 1938: unemployment rate by Grytten (2008); 1939 – 1960: unemployment rate as given by unemployment insurance from Molinder (2018) and Bengtsson and Molinder (2017); we took the relative changes in this original series and backward extended the subsequent series; 1960 – 2016: unemployment rate from OECD (2016)

**United Kingdom** 1875 – 1970: unemployment rate from Bank of England (2017), per cent of total workforce in the United Kingdom, including the Republic of Ireland; 1971 – 2016: OECD (2016)

**United States** 1875 – 1898: unemployment rate from Vernon (1994), page 710; 1899 – 1928: unemployment rate from Romer (1986), page 31; used the relative changes of this source to backward extend the subsequent series; 1929 – 1943: unemployment rate from Darby (1976), page 8; 1943 – 1960: unemployment from the U.S. Bureau of the Census (1975); 1960 – 2016: unemployment rate from OECD (2016)

### **Labour force participation among population aged 15-64 years, rate**

The labour force participation rate consists of three major building blocks. First, we use all available information concerning the labour force participation rate as given by the OECD (2021), except for the Netherlands; here, due to a break in the data we use OECD data only from 1987 onwards. In the OECD source, the starting year of coverage varies by country. From 1950 to whatever the starting year is, we use the labour force 5-year-interval estimates of ILO (1977), which are divided by age groups. We linearly interpolate the labour force shares of those younger than 15 as well 65 and above. Before 1950, we assume the shares of 1950 to be constant. We apply these shares to our total labour force to obtain the labour force aged 15-64. Then, we calculate the labour force participation rate 15-64 using population index as above.

## Appendix C State Space Model for Age Groups

Population growth data is of central importance in our application. In general, the coverage can be described as good. Due to historical databases such as Bolt and van Zanden (2020) total population data is available over the whole time period under investigation. Contrarily, more detailed age-group data, needed to obtain the development of working-age population, does not exist for every observation, although extensive collections such as HMD (2020a) provide great coverage. More precisely, five countries show missing or contradictory (due to e.g. boundary changes) data on annual population by age groups: Australia 1875-1920, France 1911-1921, Germany 1911-1949, the United Kingdom 1901-1921, and the United States 1875-1899. To address this issue and obtain data for these age-groups, we use state-space models of the standard form in the Durbin and Koopman (2012) notation:

$$y_t = \alpha_t + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \quad (9)$$

$$\alpha_{t+1} = \alpha_t + \eta_t, \quad \eta_t \sim N(0, \sigma_\eta^2) \quad (10)$$

We use information from available annual data on total population as well as information on age groups before, after, and, if available, within (census years) the respective period. There are two aspects to consider in this setting: First, the territorial coverage of the historical data used may vary considerably. Second, the data often refers to different dates (days) during the respective year. We account for territorial changes by estimating a state-space model for each ‘territorial period’ separately. This is only the case for Germany. To address different dates, we explicitly model the component DEVI, which is the difference between the annual total population and the summarized age groups in the respective year. Hence, we assume that the total population can be fully explained by summarized age groups, the only difference between both is the day of documentation, captured by DEVI, and there are no other idiosyncratic error components.

For a more intuitive understanding, the model can be written in matrix notation. Our observation and state equations can be written as:

$$Y_n = Z\alpha + \varepsilon, \quad \varepsilon \sim N(0, H) \quad (11)$$

$$\alpha = T(\alpha_1^* + R\eta), \quad \varepsilon \sim N(0, Q) \quad (12)$$

where  $Y_n$  is a  $(g + 2) \times 1$  observation vector comprising the total population,  $g$  age groups and the deviation component;  $\alpha$  and  $\alpha_1^*$  are  $(g + 1) \times 1$  (initial) state vectors comprising  $g$  age groups as well as the DEVI component;  $\varepsilon$  is a  $(g + 2) \times 1$  vector of errors;  $H$  is a zero matrix since  $\varepsilon$  is excluded from the model as described above. The variance-covariance-matrix  $Q$  is

taken from the initial available observations. The observation and transition matrices  $Z$  and  $T$  enter the models as follows:

$$Z = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 0 & & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \\ 0 & 0 & & 1 \end{bmatrix}, \quad T = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \\ 0 & 0 & & 1 \end{bmatrix}$$

The model is estimated by maximum likelihood using a Kalman smoother. The inclusion of a variety of data sources and additional data preparation steps were necessary:

**Australia, 1875 – 1920** The data on total population was taken from the Maddison database ([Bolt and van Zanden 2020](#)), age group data for the census years 1871, 1881 and 1891 were taken from [ABS \(2019\)](#), for 1901, 1911, and 1921 from [Mitchell \(2013\)](#). Before 1901, the Federation of Australia did not exist. Therefore, we aggregate the data of the separate colonies in 1871, 1881 and 1891. For Queensland, only aggregated age group data for 60-69 years and 70-79 years exist. We assume the age structure to be the same is in the remaining colonies and derive more fine-grained 5-years-age-groups.

**France, 1911 – 1921** In general, the Human Mortality Database offers information for this period. However, due to the upheavals of the First World War, the population data in the HMD does not cover all regions of France and, consequently, the development of the total population after 1918 differs significantly from other data sources. To achieve consistency, we re-estimate the period from 1911 to 1921 using population data by age groups at census dates 1911 and 1921 from [Mitchell \(2013\)](#) and total population from [Bolt and van Zanden \(2020\)](#).

**Germany, 1911 – 1949** In the period under investigation, Germany underwent several major territorial changes. In order avoid distortions caused by different territorial coverages in different data sources, we estimate the model for Germany for three separate periods.

Period 1, 1911 – 1918: The data on total population was taken from the Maddison database, age group data for the census years 1910 and 1919 were taken from [Franzmann \(2015\)](#). Since the 1919 total population and census data refer to a different territory than the population structure in the previous years, we weight those data according to the 1910 census proportion of the areas no longer covered in 1919 ([StRA 1925](#)).

Period 2, 1920 – 1945: The data on total population was taken from the Maddison database, age group data for the census years 1919, 1925, 1933, and 1946 were taken from [Franzmann \(2015\)](#), for 1939 from [Mitchell \(2013\)](#). Since the census in 1925 and 1933 excluded the ‘Saargebiet’, we weight the age group data of [Franzmann \(2015\)](#) accordingly.

Period 3, 1947 – 1949: The data on total population was taken from [Sensch \(2004\)](#), original data source is [Rinne \(1996\)](#). Age group data for the census year 1946 were taken from Franzmann, for 1950 from [Mitchell \(2013\)](#). Since the 1946 data refers to whole of Germany, we subtract, using 1946 census data from Mitchell, the age group data for East Germany and Berlin, the latter proportionally to the occupation zones ([AVBZ 1950](#)).

**United Kingdom, 1855 – 1921** To achieve consistency with other UK statistics prior to 1922, the population data for Ireland is calculated and added to the population for Great Britain supplied by the [HMD \(2020a\)](#). To achieve this, we use total population for Ireland (Southern and Northern) for 1870-1926 from [Bank of England \(2017\)](#) and population data by age groups for the census dates 1851 and 1861 from [HMSO \(1863\)](#), for 1871 and 1881 from [HMSO \(1882\)](#), for 1891 and 1901 from [HMSO \(1902\)](#), for 1911 from [HMSO \(1913\)](#), and for 1926 from [Mitchell \(2013\)](#).

**United States, 1851-1899** The data on total population and the data by age groups for the census years 1850, 1860, 1870, 1880, and 1890 was taken from [U.S. Bureau of the Census \(1975\)](#), the data by age groups for 1900 was taken from [U.S. Bureau of the Census \(2000\)](#). Since parts of data for 1850 and 1860 is available for aggregated age groups, we assume that these aggregates split up according to the shares of age groups in Canadian census in 1851 and 1861 (shares taken from [Mitchell 2013](#)).

Finally, we interpolate the obtained population data by age groups to January 1<sup>st</sup> wherever possible following the methodology<sup>3</sup> of the [HMD \(2020b\)](#) to achieve consistent population figures for further usage, e.g. in case of labor force participation. The persistence of census dates or other available population by age group sources within a country varies considerably. In Germany, nearly each census in the estimated time period was taken on another day in

---

<sup>3</sup>As explained in the HMD methods protocol, the population of January 1<sup>st</sup> can be linearly interpolated using the simple formula (in the case of estimates by 1<sup>st</sup> October, non-leap year):

$$P(x, 01.01.YYYY) = \frac{273}{365}P(x, 01.01.YYYY - 1) + \frac{92}{365}P(x, 01.01.YYYY)$$



the year. If there is only slight variation in the census dates (single days), we choose a reasonable date in between. The dates from which we interpolated our estimates are as follows: Australia: 1<sup>st</sup> April; France: 5<sup>th</sup> March; Germany: 1<sup>st</sup> December in periods 1 and 2 (the estimated figures of both periods were chain-linked using the 1919-1920 growth rate of period 2 before interpolation), due to highly heterogenous data in terms of survey dates and territory covered, we do not interpolate our estimates for period 3 but assume that our model approximates the respective population change; UK/Ireland 3<sup>rd</sup> April; USA: 1<sup>st</sup> July for 1875-1890 and 1<sup>st</sup> June for 1890-1900.



## References (appendix)

- Abildgren, K. (2005). Estimates of the Danish general government budget balance and the cyclical budget volatility 1875-2003. Danmarks Nationalbank Working Papers No. 30: [https://www.nationalbanken.dk/en/publications/Documents/1987/DNWP\\_30\\_2005.pdf](https://www.nationalbanken.dk/en/publications/Documents/1987/DNWP_30_2005.pdf).
- Abildgren, K. (2017). A Chart & Data Book on the Monetary and Financial History of Denmark. Working Paper: <https://sites.google.com/view/kim-abildgren/historical-statistics>.
- ABS (2019). Australian Historical Population Statistics. Online (last checked 2021-06-22): <https://www.abs.gov.au/statistics/people/population/historical-population/latest-release>.
- ABS (2021). National, state and territory population. Online (last checked 2021-06-22): <https://www.abs.gov.au/statistics/people/population/national-state-and-territory-population/latest-release>.
- AVBZ (1950). *Volkszählung. Tabellenteil*. Berlin.
- Axelrod, B. (1972). Historical Studies of Emigration from the United States. *International Migration Review*, 6(1), 32.
- BA (2020). Arbeitslosigkeit im Zeitverlauf: Entwicklung der Arbeitslosenquote (Jahreszahlen). Online (last checked 2021-06-22): [statistik.arbeitsagentur.de](http://statistik.arbeitsagentur.de).
- Bank of England (2017). A millenium of macroeconomic data. The Bank of England's collection of historical macroeconomic and financial statistics. Online (last checked 2020-06-12): <https://www.bankofengland.co.uk/statistics/research-datasets>.
- Barro, R. J. & Ursúa, J. F. (2010). Barro-Ursua Macroeconomic Dataset. Online (last checked 2020-07-17): <https://scholar.harvard.edu/barro/publications/barro-ursua-macroeconomic-data>.
- BEA (2020). National Income and Product Accounts. Table 5.2.5. Gross and Net Domestic Investment by Major Type, Quantity Indexes (A). Online (last checked 2021-02-08): [www.bea.gov](http://www.bea.gov).
- Bengtsson, E. & Molinder, J. (2017). The economic effects of the 1920 eight-hour working day reform in Sweden. *Scandinavian Economic History Review*, 65(2), 149–168.
- Bolt, J. & van Zanden, J. L. (2020). The Maddison Project: Maddison style estimates of the evolution of the world economy. A new 2020 update. Maddison-Project Working Paper WP-15: <https://www.rug.nl/ggdc/historicaldevelopment/maddison/publications/wp15.pdf>.
- Butlin, M. (1977). A Preliminary Annual Database 1900/01 to 1973/74. Source Papers in Economic History No. 22: <https://rse.anu.edu.au/research/centre-economic-history/source-papers-economic-history>.
- CBS (2014). Labour force; 1800-2013 (12-hours threshold), series 71882eng. Online (last checked 2020-07-07): [https://opendata.cbs.nl/statline/portal.html?%5C\\_la=%0Aen%5C%&%5C\\_catalog=CBS%5C&tableId=71882eng%5C%&%5C\\_theme=1136](https://opendata.cbs.nl/statline/portal.html?%5C_la=%0Aen%5C%&%5C_catalog=CBS%5C&tableId=71882eng%5C%&%5C_theme=1136).
- CBS (2019). Population and population dynamics by month; 1995-2018. Online (last checked 2021-06-22): <https://opendata.cbs.nl/#/CBS/en/dataset/37943eng/table>.
- Central Bureau of Statistics of Norway (1966). *Langtidslinjer i norsk økonomi 1865-1960 (Trends in Norwegian Economy 1865-1060)*. Oslo.

- Clark, C. (1957). *The Conditions of Economic Progress*. London: Macmillan, 3rd edition.
- Darby, M. R. (1976). Three-and-a-Half Million U.S. Employees Have Been Misaid: Or, an Explanation of Unemployment, 1934-1941. *Journal of Political Economy*, 84(1), 1–16.
- DEMIG (2015). DEMIG TOTAL, version 1.5. Oxford: International Migration Institute, University of Oxford. Online (last checked 2021-06-20): [www.imi-n.org](http://www.imi-n.org).
- Deutsche Bundesbank (2020). Lange Zeitreihen: Zur Wirtschaftsentwicklung in Deutschland. Online (last checked 2020-06-23): <https://www.bundesbank.de/de/statistiken/indikatorensetze/lange-zeitreihen/lange-zeitreihen-zur-wirtschaftsentwicklung-in-deutschland-832406>.
- Durbin, J. & Koopman, S. J. (2012). *Time Series Analysis by State Space Methods*. Oxford: Oxford University Press, 2nd edition.
- Feenstra, B. R. C., Inklaar, R., & Timmer, M. P. (2015). The Next Generation of the Penn World Table. *American Economic Review*, 105(10), 3150–3182.
- Flora, P., Kraus, F., & Henning, W. (1987). *State, Economy and Society in Western Europe, 1815-1975. A Data Handbook, Vol. II: The Growth of Industrial Societies and Capitalist Economies*. Frankfurt: Campus-Verlag.
- Franzmann, G. (2015). Histat-Datenkompilation: Bevölkerung nach Alter in Jahren und nach Geschlecht für das Deutsche Reich, die frühere Bundesrepublik und Deutschland, 1871-2010. GESIS Datenarchiv, Köln. Histat. Studiennummer 8617. Datenfile Version 1.0.0. Online (last checked 2020-11-23): <https://doi.org/10.4232/1.12313>.
- Galenson, W. & Zellner, A. (1957). International Comparison of Unemployment Rates. In Universities-National Bureau (Ed.), *The Measurement and Behavior of Unemployment*, volume 1 (pp. 439–584). Princeton: Princeton University Press.
- Groote, P., Albers, R., & de Jong, H. (1996). A standardised time series of the stock of fixed capital in the Netherlands, 1900-1995. Working paper: <https://research.rug.nl/en/publications/a-standardised-time-series-of-the-stock-of-fixed-capital-in-the-n>.
- Grytten, O. H. (1994). *En empirisk analyse av det norske arbeidsmarked 1918-1939: arbeidsledigheten i Norge i internasjonalt perspektiv*. Bergen: Norges handelshøyskole.
- Grytten, O. H. (2004). The gross domestic product for Norway 1830 – 2003. In Ø. E. Eitheim, J. T. Klovland, & J. F. Qvigstad (Eds.), *Historical Monetary Statistics for Norway 1819-2003* chapter 6, (pp. 241–288). Oslo.
- Grytten, O. H. (2007). Norwegian wages 1726-2006 classified by industry. In Ø. Eitheim, J. T. Klovland, & J. F. Qvigstad (Eds.), *Historical Monetary Statistics for Norway - Part II. Norges Bank Occasional Papers No. 382* chapter 6, (pp. 343–384). Oslo.
- Grytten, O. H. (2008). Why was the Great Depression not so Great in the Nordic Countries? Economic Policy and Unemployment. *The Journal of European Economic History*, 37(2/3), 369–384.
- HMD (2020a). All statistics for HMD. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Online (last checked 2020-07-08): [www.mortality.org](http://www.mortality.org).
- HMD (2020b). Methods Protocol for the Human Mortality Database. Online: <https://www.mortality.org/Public/Docs/MethodsProtocol.pdf>.
- HMSO (1863). *Census of Ireland. For the Year 1861. General Report*. London.
- HMSO (1882). *Census of Ireland, 1881. Part II. General Report, with Illustrative Maps and*

- Diagrams, Tables, and Appendix.* Dublin.
- HMSO (1902). *Census of Ireland, 1901. Part II. General Report, with Illustrative Maps and Diagrams, Tables, and Appendix.* Dublin.
- HMSO (1913). *Census of Ireland, 1911. General Report, with Tables and Appendix.* Dublin.
- Hoffmann, W. G. (1965). *Das Wachstum der deutschen Wirtschaft seit der Mitte des 19. Jahrhunderts.* Berlin, Heidelberg: Springer.
- Hohls, R. (1991). *Arbeit und Verdienst. Entwicklung und Struktur der Arbeitseinkommen im Deutschen Reich und in der Bundesrepublik.* Dissertation, Berlin.
- Hohls, R. & Kaelbe, H. (1989). *Die regionale Erwerbsstruktur im Deutschen Reich und in der Bundesrepublik 1895-1970. Quellen und Forschung zur Statistik von Deutschland [Data taken from GESIS Datenarchiv, Köln. Histat. Studiennummer 8212. Datenfile Version 1.0.0].* St. Katharinen: Scripta Mercaturae.
- ILO (1977). *Labour force estimates and projections, 1950-2000.* Geneva: ILO, 2nd edition.
- INSEE (2020). Components of population changes, metropolitan France. Online (last checked 2021-06-22): <https://www.insee.fr/en/statistiques/2382601?sommaire=2382613#titre-bloc-3>.
- Jordà, Ò., Schularick, M., & Taylor, A. M. (2017). Macrofinancial history and the new business cycle facts. *NBER Macroeconomics Annual*, 31, 213–263.
- Kirner, W. (1968). *Zeitreihen für das Anlagevermögen der Wirtschaftsbereiche in der Bundesrepublik Deutschland. DIW-Beiträge zur Strukturforchung. Heft 5.* Berlin: DIW.
- Klemann, H. A. & Kudryashov, S. (2012). *Occupied Economies. An Economic History of Nazi-occupied Europe, 1939-1945.* London, New York: Bloomsburg Publishing.
- Kosche, E. & Bach, H. W. (1991). Beschäftigung, Arbeitslosigkeit und Arbeitsmarktpolitik in Deutschland seit Mitte des 19. Jahrhunderts: Ein Überblick. *Sozialer Fortschritt*, 40(5), 122–132.
- Kuczynski, J. (1962). *Die Geschichte der Lage der Arbeiter unter dem Kapitalismus. Bd. 3: 1871-1900.* Berlin: Akademie-Verlag.
- Kuczynski, J. (1967). *Die Geschichte der Lage der Arbeiter unter dem Kapitalismus. Bd. 4: 1900-17/18.* Berlin: Akademie-Verlag.
- Lösch, A. (1936). *Bevölkerungswellen und Wechsellagen im Deutschen Reich von 1871 bis 1910. Data taken from GESIS Datenarchiv, Köln. Histat. Studiennummer 8270. Datenfile Version 1.0.0.* Jena: Fischer.
- Maddison, A. (1982). *Phases of capitalist development.* Oxford: Oxford University Press.
- Maddison, A. (1994). Standardised Estimates of Fixed Capital Stock: A Six Country Comparison. Working paper: <https://research.rug.nl/en/publications/standardised-estimates-of-fixed-capital-stock-a-six-country-compa>.
- Mitchell, B. R. (2013). *International Historical Statistics. 1750-2010.* London: Palgrave Macmillan Ltd.
- Molinder, J. (2018). Swedish National Unemployment Rate According to Available Sources, 1911-2010, version 1.0, SHLM Database. Online (last checked 2020-12-18): <https://jakobmolinder.com/data/>.
- Müller, J. H. & Skiba, R. (2008). Das Wachstum des durchschnittliche Lohneinkommen in Deutschland 1850-1967. GESIS Datenarchiv, Köln. ZA8285 Datenfile Version 1.0.0. Online (last checked 2020-11-23): <https://doi.org/10.4232/1.8285>.
- Norges Bank (2018). The gross domestic product for Norway. Online (last

- checked 2020-12-23): <https://www.norges-bank.no/en/topics/Statistics/Historical-monetary-statistics/Gross-domestic-product/>.
- NOS (1949). *Statistiske Oversikter 1948*. Oslo: NOS.
- OECD (2005). Economic Outlook No. 77, January 2005. Online (last checked 2020-07-08): <https://doi.org/10.1787/data-00089-en>.
- OECD (2016). Economic Outlook No. 100, November 2016. Online (last checked 2020-07-08): <https://doi.org/10.1787/7fa317bf-en>.
- OECD (2019). Main Economic Indicators: Hourly Earnings: Manufacturing for Australia. Online (last checked 2020-01-15): <https://doi.org/10.1787/data-00052-en>.
- OECD (2020). Main Economic Indicators: Consumer price indices (CPIs). Online (last checked 2020-03-27): <https://doi.org/10.1787/data-00900-en>.
- OECD (2021). Labour force statistics by sex and age: indicators. Online (last checked 2021-02-01): <https://doi.org/10.1787/data-00310-en>.
- Petzina, D., Abelshausen, W., & Faust, A. (1978). *Sozialgeschichtliches Arbeitsbuch: Bd. III: Materialien zur Statistik des Deutschen Reiches 1914-1945*. Munich: Beck.
- Pierenkemper, T. (1984). Die Vermarktung von Arbeitskraft in Deutschland. Begriffe, Tendenzen, Forschungsprobleme. In K. J. Bade (Ed.), *Auswanderer - Wanderarbeiter - Gastarbeiter* (pp. 139–178). Ostfildern: Scripta Mercaturae.
- Pierenkemper, T. (1987). The Standard of Living and Employment in Germany, 1850-1980: An Overview. *The Journal of European Economic History*, 16(1), 51–73.
- Ralf, T. (2015). Zeitreihendatensatz für Deutschland, 1834-2012. Studiennummer 8603. GESIS Datenarchiv, histat. Köln. Online (last checked 2020-07-07): <https://doi.org/10.4232/1.12202>.
- Rinne, H. (1996). *Wirtschafts- und Bevölkerungsstatistik*. Munich, Vienna: Oldenbourg, 2nd edition.
- Ritschl, A. & Spoerer, M. (1997). Das Bruttosozialprodukt in Deutschland nach den amtlichen Volkseinkommens- und Sozialproduktsstatistiken 1901-1995. *Jahrbuch für Wirtschaftsgeschichte (Economic History Yearbook)*, 38(2), 27–54.
- Romer, C. (1986). Spurious Volatility in Historical Unemployment Data. *Journal of Political Economy*, 94(1), 1–37.
- SCB (2021). Population and Population Changes 1749-2020. Online (last checked 2021-06-22): <https://www.scb.se/en/finding-statistics/statistics-by-subject-area/population/population-composition/population-statistics/pong/tables-and-graphs/yearly-statistics--the-whole-country/population-and-population-changes/>.
- Schewe, D. & Nordhorn, K. (1967). *Übersicht über die soziale Sicherung in Deutschland*. Bonn: Bundesministerium für Arbeit und Sozialordnung, 6th edition.
- Schön, L. & Krantz, O. (2012). *Swedish Historical National Accounts 1800-2000*. Lund *Papers in Economic History* 123. Lund: Lund University.
- Sensch, J. (2004). histat-Datenkompilation online: Geschichte der deutschen Bevölkerung seit 1815. Daten taken from GESIS Datenarchiv, Köln. Histat. Studiennummer 8171. Datenfile Version 1.0.0. Online (last checked 2020-09-01): <https://histat.gesis.org/histat/za8171>.
- Sensch, J. (2007). histat-Datenkompilation online: Bevölkerungsstand, Bevölkerungsbewegung, Haushalte und Familien in der Bundesrepublik Deutschland 1947

- bis 1999. GESIS Datenarchiv, Köln. histat. Studiennummer 8200. Datenfile Version 2.0.0. Online (last checked 2020-09-01): <https://histat.gesis.org/histat/za8200>.
- Sensch, J. (2012). histat-Datenkompilation online: Ausgewählte Zeitreihen aus Studien zur Entwicklung der Löhne/Gehälter und des Einkommens aus unselbständiger Arbeit in Deutschland, 1844-1990. GESIS Datenarchiv, Köln. Histat. Studiennummer 8177. Datenfile Version 2.0.0. Online (last checked 2020-09-01): <https://histat.gesis.org/histat/za8177>.
- Sensch, J. (2016). histat-Datenkomilation online: Entwicklung der Arbeitslosigkeit in Deutschland 1887-2000. GESIS Datenarchiv, Köln. Histat. Studiennummer 8218. Online (last checked 2020-09-01): <https://histat.gesis.org/histat/za8218>.
- Sommariva, A. & Tullio, G. (1987). *German Macroeconomic History, 1880-1979: A Study of the Effects of Economic Policy on Inflation, Currency Depreciation and Growth*. London: Palgrave Macmillan UK.
- SSB (1945). *Statistisk-økonomisk utsyn over krigsårene*. Oslo: SSB.
- SSB (1950). *Økonomisk Utsyn over året 1949*. Oslo: SSB.
- SSB (2020). Sysselsatte personer etter næring 1900-1969. 1000. Online (last checked 2020-11-01): [https://www.ssb.no/a/kortnavn/hist%5C\\_tab/ht-0901-sys.html](https://www.ssb.no/a/kortnavn/hist%5C_tab/ht-0901-sys.html).
- SSB (2021). Migrations. Online (last checked 2021-06-22): <https://www.ssb.no/en/befolkning/flytting/statistikk/flyttinger>.
- Statistics Denmark (2021). INDVAN: Immigration by sex, age, country of origin and citizenship. Online (last checked 2021-06-22): <https://www.statbank.dk/statbank5a/selectvarval/define.asp?PLanguage=1&subword=tabssel&MainTable=INDVAN&PXSid=206362&tablestyle=&ST=SD&buttons=0>.
- StBA (2019). Wanderungen - Fachserie 1 Reihe 1.2 - 2019. Online (last checked 2021-06-22): [https://www.destatis.de/DE/Service/Bibliothek/\\_publikationen-fachserienliste-1.html?nn=206136](https://www.destatis.de/DE/Service/Bibliothek/_publikationen-fachserienliste-1.html?nn=206136).
- Stockmann, R., Kleber, W., & Wilms-Herget, A. (1982). Entwicklung der Erwerbspersonen und Branchenstrukturen in Deutschland 1875 bis 1980. Data taken from GESIS Datenarchiv, Köln. Histat. Studiennummer 8181. Datenfile Version 1.0.0. Online (last checked 2020-09-01): <https://histat.gesis.org/histat/za8181>.
- StRA (1925). Vorläufige Ergebnisse der Volkszählung im Deutschen Reich vom 16. Juni 1925. *Sonderhefte zu Wirtschaft und Statistik*, 5(2).
- UN (2019). World Population Prospect 2019. Online (last checked 2020-05-20): <https://population.un.org/wpp/>.
- U.S. Bureau of the Census (1975). *Historical Statistics of the United States: Colonial Times to 1970, Bicentennial Edition, Part 1*. Washington D.C.: U.S. Dept. of Commerce, Bureau of the Census.
- U.S. Bureau of the Census (2000). National Intercensal Tables: 1900-1990. Online (last checked 2020-07-14): <https://www.census.gov/data/tables/time-series/demo/popest/pre-1980-national.html>.
- van Meerten, M. (2003). *Capital formation in Belgium, 1900-1945*. Leuven: Leuven University Press.
- Vernon, J. R. (1994). Unemployment rates in postbellum America: 1869-1899. *Journal of Macroeconomics*, 16(4), 701-714.
- Villa, P. (1993). *Une analyse macroéconomique de la France au XXe siècle*. Paris: CNRS

éditions.

Whiters, G., Endres, T., & Perry, L. (1985). Australian Historical Statistics: Labour Statistics. Source Papers in Economic History No. 7: <https://rse.anu.edu.au/research/centre-economic-history/source-papers-economic-history>.

Williamson, J. G. (1995). The Evolution of Global Labor Markets since 1830: Background Evidence and Hypotheses. *Explorations in Economic History*, 32(2), 141–196.