European Journal of Political Economy 77 (2023) 102280

Contents lists available at ScienceDirect

European Journal of Political Economy

journal homepage: www.elsevier.com/locate/ejpe

OECD pension reform: The role of demographic trends and the business cycle☆

Ward Romp **, Roel Beetsma *

Amsterdam School of Economics, University of Amsterdam, PO Box 15867, 1001 NJ Amsterdam; Netspar; Tinbergen Institute, Netherlands

ARTICLE INFO

JEL classification:
H55
H62
J11
J26

Keywords:
Pension reform measures
Narrative identification
Expansion
Contraction
Old-age dependency ratio
Business cycle indicators
Adjustment costs

ABSTRACT

Using a new real-time dataset from Beetsma et al. (2020) containing all pension reform measures in 23 OECD countries between 1970 and 2017, we demonstrate that, in contrast to what one might a priori expect, the timing of pension reform measures coincides with business cycle shocks and not with current or projected demographic shocks. The OECD-wide demography only explains the general reform trend. We rationalize this finding using a political-economy model with two-sided adjustment costs to explain a lack of response of pension reform measures to changes in demographic indicators.

1. Introduction

Reform measures to enhance the financial sustainability of pension arrangements are high on the agendas of national policymakers as well as of international organizations. So far, however, systematic empirical investigation of what determines the *timing* of pension reform measures is in short supply. This is unfortunate, because the outcomes of such an analysis may provide insights into the circumstances that are most conducive to the successful implementation of pension reforms.

The contribution of this paper is twofold. In their empirical analysis, Beetsma et al. (2020) demonstrate that changes in the current or projected future old-age dependency ratio do not affect the timing of pension reform measures. By contrast, the current cyclical state of the economy does have a statistically significant effect on reform measures. The demography matters only in the sense that the OECD-wide demography explains the general reform trend, but the trigger for the alleviation of pressure in the pension arrangement is the business cycle: an economic downturn leads to more contractionary measures and an economic upturn to fewer such measures. Here, we demonstrate that if we add measures of accumulating demographic pressure to their empirical framework, the role of the current state of the economy in explaining reform measures is preserved. Second, we develop a theoretical framework capable of rationalising that the timing of pension reforms is determined by the cyclical state of the economy and not by demographic pressure. While its basic set-up is relatively simple, the framework has not been deployed before in the literature.

☆ We thank an anonymous referee for useful comments on an earlier version of the paper.

* Corresponding author.
** Corresponding author.
E-mail addresses: w.e.romp@uva.nl (W. Romp), r.m.w.j.beetsma@uva.nl (R. Beetsma).

https://doi.org/10.1016/j.ejpoleco.2022.102280
Received 29 September 2021; Received in revised form 16 July 2022; Accepted 21 August 2022
Available online 28 August 2022

The theoretical framework features a forward-looking government with a potentially rather high discount rate and fixed costs of pension reform. These costs, typically in terms of credit needed to overcome political resistance and lost popular support for the government, may differ for reforms that increase pension generosity and reforms that reduce it. Through a careful country-specific calibration of these costs we are able to reproduce the numbers of generosity-increasing and contracting reforms. Our simulations demonstrate that the model performs well in predicting the timing of both reform directions. The simulations also allow us to disentangle the predictive contribution associated with fluctuations in the business cycle and changes in current and future demography, confirming that the latter play only a minor role compared to the former.

What is the intuition behind the rationalization of our main empirical finding? Suppose one starts with a pension system that is optimal given the current state of the business cycle and the current demography. With fixed reform costs, a sufficiently large change in these underlying variables is needed to trigger a reform. However, updates of the current and projected demography are typically small on an annual basis, so the additional gain from a pension reform will be small. Indeed, to the extent that demographic changes play a role they would have been incorporated already in the most recent pension reform, because they are largely anticipated, implying that the gain from further reform is generally small. In principle, the cumulation of a large number of small, predictable, changes in the current demography would at some point trigger a reform. However, because the business cycle fluctuates at a higher frequency, it is likely that a recession or boom will trigger a reform before a cumulative change in the demography necessitates one. Also, even though they are unexpected, changes in *projected* demographic conditions are unlikely to affect reform decisions. First, such changes only affect future utility, which makes it hard to overcome the fixed reform cost due to limited government tenure. Second, even if future sustainability issues are foreseen, the government may safely postpone reform, because business cycle movements will trigger a reform before the sustainability problems have had a chance to materialize.

The literature has identified different potential driving forces behind pension reform and its timing.[1] First, there is a strand of the literature exploring the role of demography. Persson and Tabellini (2000) describe two opposing effects of a higher old-age dependency ratio on the size of a pay-*as*-you-go (PAYG) system. In an older society, on the one hand the rate of return on contributions to a PAYG system is lower, making the system less attractive, while on the other hand population ageing enhances the political weight of the elderly, making it harder for politicians to engage in contractionary reform. Other studies (e.g. Gonzales-Eiras and Niepelt, 2008) relate changes in social security to its intergenerational risk-sharing aspects (see also D'Amato and Galasso, 2010). Empirically, however, the role of the demography is not so clear-cut (e.g., Blinder and Krueger, 2004). The empirical evidence for demography as an important determinant of PAYG pension reform is weak. In fact, for the U.S. and Western Europe Razin et al. (2002) even find a negative correlation between the old-age dependency ratio and the generosity of social security transfers. Second, the size of the implicit pension debt is a potential determinant of reform of PAYG defined-benefit pensions (James and Brooks, 2001). Third, external constraints, such as those imposed by Europe's Stability and Growth Pact,[2] may stimulate pension reform. Bertola and Boeri (2002) argue that such constraints could have contributed to a reduction in the generosity of social security after 1997. A fourth motive for reform is the desire to correct the distortions caused by existing arrangements leading employees to work less or retire earlier than under a system with stronger incentives for work – see, e.g., the contributions in Gruber and Wise (2009). Finally, there is the potential role of ideology. Pension privatization in Latin America was stimulated by the paradigm shift towards neo-liberalization inspired by Thatcherism and the promotion of private pensions by international organizations such as the World Bank (see World Bank, 2019; Brooks, 2007, and Orenstein, 2005 and 2013). The new paradigm also emphasized the benefits of capital market deepening, increased private savings and higher economic growth.

Our main empirical finding – the timing of pension reform measures is linked to the business cycle – is related to, but differs in two fundamental ways from, the crisis-induced reform literature (e.g., Rodrik, 1996; Abiad and Mody, 2005; Bonfiglioli and Gancia, 2015; Ranciere and Tornell, 2015; Mahmalat and Curran, 2018).[3] This literature finds that structural reforms are typically legislated during periods of poor economic performance, but it tends to focus on financial liberalization, trade liberalization, and issues of inflation and sovereign indebtedness. We focus on pension reform measures, which are unique in that difficulties with the sustainability of current pension arrangements are known well in advance. By contrast, the "regular" crisis-induced reform literature focuses on contemporaneous rather than anticipated future crises. This may not be surprising: an economic crisis may be a particularly opportune moment for reform, because only then will there be sufficient political awareness of the need to fix structural deficiencies through fundamental reform and can political obstacles to reform be overcome (Tommasi and Velasco, 1996; Tommasi, 2017). A second difference with the regular reform literature is that our data allow us to study both contractionary and expansionary reform measures. While the timing of contractionary reform is linked to a cyclically-weak economy, expansionary reform measures, which include such structural measures as increased coverage of women, tend to be implemented during economic upswings. Such expansionary reform measures are not considered in the crisis-induced literature.[4]

This paper is also related to a broader, mostly empirical, literature on structural reforms. Campos et al. (2017) provide an overview.

[1] An interesting recent paper by Bi and Zubairy (2019), instead of exploring the driving forces of pension reform, explores how narratively-identified news about pension reforms motivated by long-run sustainability concerns affects when people retire and, thereby, old-age spending.

[2] For an analysis of these constraints, see Beetsma and Uhlig (1999) and Beetsma and Debrun (2007).

[3] Campos et al. (2010) find a relatively large role for political crises in determining labour market and trade liberalization reforms.

[4] There is a literature on political and legal constraints that prevent policymakers from pursuing reform in normal times, but that become softer during a crisis. While before the global financial crisis (GFC) the US Treasury was unable to persuade banks to strengthen their capital position, after the onset of the GFC regulators managed to force them to recapitalize – see Swagel (2015).

For example, Chinn and Ito (2006), Quinn and Toyoda (2008), Bumann and Lensink (2016) and Furceri and Loungani (2015) explore the economic consequences of capital account reform; Beck and Levine (2004), Christiansen et al. (2013), Prati et al. (2013), Arcand et al. (2015) and De Haan and Sturm (2017) investigate the economic consequences of domestic financial liberalization; Griffith et al. (2006), Falcetti et al. (2006), Spilimbergo et al. (2009), Bouis and Duval (2011), Fiori et al. (2012), Fatas (2016), Cette et al. (2016) explore the effects of product and labour market reforms; and Demekas et al. (2007), Kneller et al. (2008), Wacziarg and Welch (2008), Campos and Kinoshita (2010) and De Macedo et al. (2014) analyse the effects of trade liberalization.

This paper proceeds as follows. The next section briefly describes the data, drawing on and expanding the exposition in Beetsma et al. (2020), because we supplement their data with variables that capture accumulating demographic pressure. Section 3 presents our empirical analysis, which confirms for a richer specification including changing demographic pressure that the timing of pension reform measures is linked to state of the economy. Sections 4, 5 and 6 are devoted to presenting and analysing the theoretical framework used to rationalize our empirical findings. Section 7 explores the model fit. Finally, Section 8 concludes the main text.

2. Data

In this paper we use the database constructed by Beetsma et al. (2020) and extend it with data that capture accumulating demographic pressure. A narrative approach is deployed to identify all relevant legislated pension reform measures in the 23 first OECD member countries over the period 1970 until 2017. The database covers all pension reform measures, both smaller (parametric) and more fundamental, that affect both the present value of retirement income and the government's intertemporal budget constraint. In each year, and for each country, changes in pension arrangements are listed based on a careful reading of records or documents from four main databases: the NATLEX database of the International Labor Organization (ILO, 2019), the International Social Security Association (ISSA, 2019) database, the OECD, and the European Commission's LABREF (2019), supplemented with data from ad-hoc sources.

These reform measures are classified into two categories: 1) measures with an expansionary effect on the pension arrangements, and 2) measures with a contractionary effect on pension arrangements. Reform measures with a different nature are excluded from our analysis. Examples of the first type are an increase in benefit levels and a weakening of eligibility restrictions. Examples of the second type are a reduction in benefit levels and an increase in the official retirement age (which reduces the present value of expected future pension benefits). Importantly, reform measures are assigned to the year in which they are signed into law. This way of dating is the relevant one if our goal is to explain reform measures based on "real-time" factors, such as the state of the economy or the current or currently projected state of the demography. The time between legislation and implementation can be substantial, in particular when it comes to measures to increase the retirement age. Typically, it is not politically opportune to downscale too much the retirement provisions of those close to retirement. Hence, increases in retirement age tend to be implemented gradually over a long period of time.

2.1. Reform regimes

Based on the above classification, every country-year combination is assigned one of three possible "reform regimes". The first regime is "Expanding only". We assign this regime to a country-year combination if at least one reform of expansionary nature was legislated in that country and in that year and no reforms of contractionary nature were legislated. The second regime is "Contracting", which occurs when at least one contractionary reform is legislated in that country and in that year. If both reforms of an expansionary and a contractionary nature are legislated, then this is also classified as "Contracting". A careful reading of all reform measures clearly shows that large systemic reforms of a contractionary nature fall into this regime.[5] Often these reforms are so substantial that governments have to offset part of the contracting reform by also accepting one or more (smaller) expansionary measures to either buy political support or to compensate some targeted groups. The third and final regime is the default regime of "No reform". This makes the three regimes mutually exclusive. Table 1 reports the number of occurrences of each regime both for the full sample period 1970–2017 and for a split in two subsample periods of equal length. The number of "Expanding only" regimes is approximately equal to "Contracting". While the incidence of "Expanding only" regimes is roughly equally spread over the two periods, the number of "Contracting" regimes is almost four times larger in the second-than in the first subperiod. This may not be surprising in view of rising projections of old-age dependency ratios and growing awareness of the future ageing costs.

One concern is that our method gives all reforms the same weight, regardless of how many people are affected. To tackle this concern, we use an additional classification with which we try to capture to number of people affected by a reform. Based on a careful reading of all reforms, we split them into "Many affected" and "Few affected". The introduction of a wife's pension, which replaces a less generous wife's allowance (like Australia did in 1972) clearly falls into the "Many affected" category, just like lowering the general retirement age in Germany in 1972. Lowering the retirement age of only severely handicapped persons in the Germany in 1978 is an example of a reform that we classified as "Few affected". If a country in a specific year implemented one or more reforms, then we classify this country-year as "Many affected" if at least one of the reforms classifies as "Many affected". If not, then we classify this country-year as "Few affected". This does not capture how severely those people are affected by the reform. For that we need a more detailed quantitative analysis of each reform, which is not possible with the available data. Table 2 shows the number of regimes in

[5] In one sensitivity analysis, we split the "Contracting" regime into "Contracting only" and "Contracting and Expanding". When significant, the estimated coefficients always have the same sign as before and are of comparable magnitude; the standard deviation of the coefficients is larger since we split the number of observations in this category into two.

Table 1
Number of occurrences of each policy regime by time period.

	1970–2017	1970–1993	1994–2017
"Expanding only" regime	223	121	102
"Contracting" regime	235	49	186
"No reform" regime	646	382	264
Total	1104	552	552

Note: this table reports the number of (country, year)-combinations for which the different regimes apply.

each category. This table clearly shows that conditional on a reform action by the government, usually many people are affected.

As explanatory variables, we use the same baseline variables as in Beetsma et al. (2020). These are year-on-year growth of GDP per capita ($GROWTH_{it}$), the unemployment rate ($UNEMPL_{it}$), and the government's public deficit (DEF_{it}), all intended to capture elements of the state of the economy of country i in year t, and the current old-age dependency ratio (OAD_{it}) and its 25-year ahead forecast ($OAD25_{it}$), as well as transformations of these basic variables. The old-age dependency ratio is measured as the number of people of 65 years and older divided by the number of people in the age group 15–64 years. Finally, we use a dummy variable $MAASTRICHT_{it}$, which is one for all country-year combinations as of 1992 (the year of signing the Maastricht Treaty) if country i is a member of the EU, and zero otherwise.[6] The motivation is that the need to meet the budgetary criteria for entry into the eurozone and, later, the requirements imposed by the Stability and Growth Pact have forced countries to take pension reform measures that ease the public budget constraint.

2.2. Accumulating demographic pressure

Not included in Beetsma et al. (2020), but potentially important, is the role of accumulating demographic pressure on the pension system. Hence, in this paper we test whether such accumulating demographic pressure increases the probability of reforms of a contracting nature and decreases the probability of reforms of an expansionary nature. We capture accumulating demographic pressure using two indicators. The first is the Time (in years) since the Last Reform

$$TsLR_{i,t} = \begin{cases} TsLR_{i,t-1} + 1, & REF_{i,t-1} = 0, \\ 0, & REF_{i,t-1} = 1, \end{cases}$$

where $REF_{i,t}$ is 1 if at least one reform measure was legislated in that country-year, so the dummy indicator of either the "Expanding only" or "Contracting" regime is 1. Note that this indicator is independent of the direction of the last reform. Such indicator could be relevant in a setting where the pension system is re-evaluated on a periodic basis. The more time has passed since the last reform, the more likely it is that the system is no longer fully suited and thus needs to undergo some change.

The second indicator measures the change of the Old-Age dependency Ratio since the Last Reform

$$OAD25sLR_{i,t} = \begin{cases} OAD25sLR_{i,t-1} + OAD25_{i,t} - OAD25_{i,t-1}, & REF_{i,t-1} = 0 \\ 0, & REF_{i,t-1} = 1 \end{cases}$$

If pension reforms are driven by accumulating demographic pressure then a high value of this measure should coincide with a high probability of the "Contracting" regime and a low probability of the "Expanding only" regime.

Both indicators are reset when a pension reform measure is legislated, regardless of the nature of the reform measure. The idea is that implementing a pension reform takes all recent information into account, so it resets the pressure indicator back to zero.

3. Empirical results

Our reform regimes are mutually exclusive, which suggests the use of a multinomial logit regression to estimate the relationship between our baseline variables and the probability of observing one of our reform regimes. The probability $p_{it,r}$ of country i being in reform regime r in year t is:

$$p_{it,r} = \frac{\exp(z_{it,r})}{1 + \sum_{h=1}^{R} \exp(z_{it,h})},$$

where h counts over the set of possible reform regimes "Expanding only" and "Contracting"[7] and where $z_{it,r}$ is a reform-regime specific

[6] Beetsma et al. (2020) perform an extensive sensitivity analysis by including other economic, budgetary, political and crisis indicators. The inclusion of these other variables does not change the estimated coefficients of the baseline variables and the estimated coefficients of these other variables are only occasionally significant. The role of political indicators is further discussed below.

[7] Hence, $R = 2$ and the likelihood of ending up in the "No reform" regime is

$$1 - \sum_{h=1}^{R} p_{it,h}.$$

Table 2
Number of occurrences of each policy regime split by the scope of the reforms.

	Many affected	Few affected	Total
"Expanding only" regime	133	90	223
"Contracting" regime	200	35	235
"No reform" regime			646

Note: this table reports the number of (country, year)-combinations for which the different regimes apply.

linear function of the vector $EXVAR_{it}$ of explanatory variables:

$$z_{it,r} = \alpha_{0i,r} + \alpha_r{}' EXVAR_{it},$$

where $\alpha_{0i,r}$ captures country-fixed effects and α_r is a coefficient vector of appropriate dimensions. When we include time-fixed effects, a Wald test that these effects are jointly zero fails to reject for each of the regimes, if we normalize one of these time-fixed effects to zero, so as to keep $\overline{OAD25_t}$, the cross-country average of $OAD25_{it}$, as one of the explanatory variables in the model (results available upon request). Because we normalize these time-fixed effects such that they capture any potential omitted time-specific determinants after accounting for $\overline{OAD25_t}$, their insignificance also implies there is no indication of such omitted determinants. This supports our conclusion that it is OECD-wide ageing projections that affect reforms. Moreover, when time-fixed effects are added, the coefficients on the business cycle indicators remain essentially unchanged. Henceforth, we do not include time-fixed effects in our estimations.

Table 3 presents the baseline multinomial logit regression with the three regimes and including the baseline variables of Beetsma et al. (2020). These are the demographic variables $\overline{OAD25_t}$, $OADDEV25_{it} \equiv OAD25_{it} - \overline{OAD25_t}$, the country-specific deviation of the projected old-age dependency ratio, $\Delta OAD25_{it} \equiv OAD25_{it} - OAD25_{i,t-1}$, the change in the projected old-age dependency ratio, the business cycle indicators ($GROWTH_{it}$, DEF_{it} and $UNEMPL_{it}$) and $MAASTRICHT_{it}$. Different from Beetsma et al. (2020) we do not include an interaction term between economic growth and the first decade in our sample, i.e. the 1970s. The inclusion of this interaction term does not change the main results.

The columns labelled "Coefficient" report the variables' coefficient estimates and the column labelled "Marg. Eff." shows the mean over countries of that variable's marginal effect evaluated at the countries' mean of the variable over time; the so-called "average marginal effect".[8] Table 3 confirms the main findings of Beetsma et al. (2020): the world-wide ageing process as measured by the mean of the projected old-age dependency ratio lowers the probability of expansionary reform measures and raises the probability of the "Contracting" reform regime. The coefficients of the three business cycle indicators, $GROWTH_{it}$, DEF_{it} and $UNEMPL_{it}$, all have the expected signs. The coefficient on $GROWTH_{it}$ is positive and highly significantly different from zero for the "Expanding only" regime, while it is negative and highly significantly different from zero for the "Contracting" regime. A one percentage point increase in the GDP growth rate raises the likelihood of "Expanding only" by 1.4 percentage points and lowers the likelihood of "Contracting" by 1.4 percentage points. The coefficient on $UNEMPL_{it}$ is positive and highly significantly different from zero for "Contracting". A one percentage point higher unemployment rate raises the likelihood of "Contracting" by 1.5 percentage points. A higher unemployment rate does not affect the likelihood of "Expanding only". The public deficit DEF_{it} exerts a significantly positive effect on the likelihood of "Contracting". A one percentage point higher deficit raises the likelihood of this regime by 0.9 percentage points. A likely explanation for the role of the state of the business is that an adverse state undermines the affordability of current arrangements. Finally, the Maastricht dummy is positive and significant for both reform regimes. This could be indicative of lower adjustment costs of pension reforms if the country has signed the Maastricht Treaty. Being a signatory increases the probability of a reform by 13%–15%.[9]

In Table 4 we add to our baseline regression the change of the 25 year-ahead forecast since the last reform, $OAD25sLR_{it}$, and in Table 5 also the interaction of this variable with economic growth. We include these variables since mounting demographic pressure may induce pension reforms, while after a pension reform the pressure for further reform is at least temporarily taken away. The interaction of the pressure indicator with the main business cycle indicator could be relevant if both a business cycle fluctuation and mounting demographic pressure are required to trigger a pension reform. For completeness, Tables 6 and 7 report the same regressions as the preceding two tables, but with the time since the last reform, $TsLR_{it}$, replacing $OAD25sLR_{it}$ as the pressure indicator.

The results in Tables 4–7 are consistent: the coefficients on the individually entered pressure variables generally have the expected negative sign for the "Expanding only" regime and the expected positive sign for the "Contracting" regime, but the coefficient on the change in the projected old-age dependency ratio since the last reform, $OAD25sLR_{it}$, is significantly different from zero only for the "Contracting" regime in Table 4 and only so at the 5% level, while the coefficient on the number of years since the last reform, $TsLR_{it}$, is never significantly different from zero at the 5% level. The coefficients on the interaction terms of these variables with $GROWTH_{it}$ do not differ significantly from zero in all instances. The estimates of the coefficients of the baseline variables are robust throughout and show little change.

One may ask whether the estimated coefficient of the change of the forecast since the last reform, $OAD25sLR_{it}$ is affected by the

[8] Concretely, we first calculate for each country the mean of each variable and then evaluate for each country the marginal effect at that point. After this, we take the mean of those marginal effects, which is the "average marginal effect".

[9] A regression with the interaction of having signed the Maastricht Treaty with unemployment shows that signatories of the Treaty have a higher probability of implementing a pension reform, but respond less to fluctuations in unemployment than other countries.

Table 3
Baseline multinomial logit estimations.

Independent variables	Expanding only		Contracting	
	Coefficient	Marg. Eff.	Coefficient	Marg. Eff.
$\overline{OAD25_t}$	−4.15***	−0.81***	4.73***	0.79***
	(1.53)	(0.22)	(1.42)	(0.18)
$OADDEV25_{it}$	3.05	0.43	0.77	−0.011
	(2.96)	(0.43)	(2.78)	(0.36)
$\Delta OAD25_{it}$	1.32	−0.14	9.01	1.16
	(7.21)	(1.05)	(6.81)	(0.87)
$GROWTH_{it}$	7.36*	1.46***	−9.13**	−1.50***
	(3.59)	(0.53)	(3.98)	(0.51)
DEF_{it}	−0.50	−0.33	6.81*	0.93**
	(3.17)	(0.46)	(3.16)	(0.41)
$UNEMPL_{it}$	0.87	−0.28	11.1***	1.45***
	(3.57)	(0.52)	(3.85)	(0.49)
$MAASTRICHT_{it}$	1.11***	0.13***	1.25***	0.13***
	(0.29)	(0.046)	(0.33)	(0.045)
N	1081			
McFadden R2	0.14			

Note: standard errors in parentheses, $^{*}p < 0.05$, $^{**}p < 0.025$, $^{***}p < 0.01$. N = number of observations.

Table 4
Multinomial logit estimation with pressure indicator based on OAD25.

Independent variables	Expanding only		Contracting	
	Coefficient	Marg. Eff.	Coefficient	Marg. Eff.
$\overline{OAD25_t}$	−4.15***	−0.81***	4.83***	0.80***
	(1.54)	(0.22)	(1.43)	(0.18)
$OADDEV25_{it}$	3.05	0.44	0.61	−0.033
	(2.96)	(0.43)	(2.80)	(0.36)
$\Delta OAD25_{it}$	1.84	0.18	2.74	0.30
	(7.95)	(1.15)	(7.40)	(0.95)
$GROWTH_{it}$	7.26*	1.43***	−8.85*	−1.45***
	(3.60)	(0.53)	(4.00)	(0.51)
DEF_{it}	−0.55	−0.34	6.98*	0.95**
	(3.17)	(0.46)	(3.15)	(0.40)
$UNEMPL_{it}$	0.93	−0.29	11.5***	1.49***
	(3.56)	(0.52)	(3.86)	(0.49)
$MAASTRICHT_{it}$	1.10***	0.12***	1.30***	0.14***
	(0.29)	(0.047)	(0.33)	(0.045)
$OAD25sLR_{it}$	−1.35	−0.64	11.6*	1.60*
	(6.28)	(0.92)	(5.79)	(0.74)
N	1081			
McFadden R2	0.15			

Note: standard errors in parentheses, $^{*}p < 0.05$, $^{**}p < 0.025$, $^{***}p < 0.01$.

presence of the change in the projection $\Delta OAD25_{it}$. However, dropping this last variable leaves the coefficients of $OAD25sLR_{it}$, as well as those of the other variables, virtually unaffected. In particular, the role of $OAD25sLR_{it}$ remains very minor. The results are available upon request.

A legitimate question concerns the completeness of the empirical framework. Beetsma et al. (2020) conduct an elaborate analysis of the potential role of political variables taken from the Comparative Political Data Set I (Armingeon et al., 2018a, 2018b). However, they find no evidence that these variables help to explain reform regimes. Another concern is that of possible feedback effects from reforms to the business cycle variables, which could bias their coefficient estimates. Such effects are unlikely to play a role, because reforms, once signed into law, are often only gradually implemented, implying that their effect on disposable income of retirees (which in any case constitutes only part of national disposable income) is spread over years. Indeed, instrumental variables estimation in Beetsma et al. (2020) suggests that feedback effects play no role. Another potential concern could be a structural break in the late 80's. The 1970s and 1980s were a time when some European governments legislated early retirement schemes which would open up jobs for the younger cohorts who suffered from high unemployment rates. However, a separate regression only for 1970–1992 suggests that the high unemployment rates caused by the oil crises only increased the probability of contracting reforms (results available upon request). As a final sensitivity analysis, we use our classification of how many people are affected by the reforms legislated in a specific country-year combination. This gives a multinominal logit regression with five categories; the two regimes "Expanding only" and "Contracting" are each split into two sub-regimes indicating whether many (a broad group) were affected or only few (a specific/small group) were affected. The results are documented in Table 8. Due to space limitations, we only present the coefficients; the marginal

Table 5
Multinomial logit estimation with pressure indicator based on OAD25 and interaction with growth.

Independent variables	Expanding only		Contracting	
	Coefficient	Marg. Eff.	Coefficient	Marg. Eff.
$\overline{OAD25_t}$	−4.13***	−0.80***	4.81***	0.80***
	(1.54)	(0.22)	(1.43)	(0.18)
$OADDEV25_{it}$	2.95	0.42	0.71	−0.015
	(2.96)	(0.43)	(2.80)	(0.36)
$\Delta OAD25_{it}$	1.71	0.16	2.56	0.28
	(7.91)	(1.15)	(7.44)	(0.95)
$GROWTH_{it}$	8.78**	1.45***	−10.2*	−1.49***
	(3.90)	(0.53)	(4.54)	(0.52)
DEF_{it}	−0.87	−0.39	7.01*	0.97**
	(3.19)	(0.46)	(3.15)	(0.40)
$UNEMPL_{it}$	1.07	−0.26	11.4***	1.48***
	(3.56)	(0.52)	(3.86)	(0.49)
$MAASTRICHT_{it}$	1.10***	0.12***	1.29***	0.14***
	(0.29)	(0.047)	(0.33)	(0.045)
$OAD25sLR_{i,t}$	3.03	−0.80	10.4	1.63*
	(7.23)	(0.93)	(6.54)	(0.74)
$GROWTH_{it} \times OAD25sLR_{i,t}$	−227.4		125.7	
	(206.2)		(226.9)	
N	1081			
McFadden R2	0.15			

Note: standard errors in parentheses, $^*p < 0.05$, $^{**}p < 0.025$, $^{***}p < 0.01$.

Table 6
Multinomial logit estimation with pressure indicator based on time since last reform.

Independent variables	Expanding only		Contracting	
	Coefficient	Marg. Eff.	Coefficient	Marg. Eff.
$\overline{OAD25_t}$	−4.17***	−0.82***	5.04***	0.83***
	(1.54)	(0.22)	(1.44)	(0.18)
$OADDEV25_{it}$	3.08	0.43	0.88	0.0019
	(2.96)	(0.44)	(2.80)	(0.36)
$\Delta OAD25_{it}$	1.29	−0.12	8.44	1.08
	(7.21)	(1.05)	(6.84)	(0.88)
$GROWTH_{it}$	7.27*	1.43***	−8.86*	−1.45***
	(3.60)	(0.53)	(4.00)	(0.51)
DEF_{it}	−0.49	−0.33	6.83*	0.93**
	(3.16)	(0.46)	(3.17)	(0.41)
$UNEMPL_{it}$	0.87	−0.29	11.3***	1.47***
	(3.57)	(0.52)	(3.86)	(0.49)
$MAASTRICHT_{it}$	1.10***	0.12***	1.32***	0.14***
	(0.29)	(0.047)	(0.33)	(0.046)
$TsLR_{i,t}$	−0.0092	−0.0043	0.077	0.011
	(0.040)	(0.0060)	(0.045)	(0.0058)
N	1081			
McFadden R2	0.14			

Note: standard errors in parentheses, $^*p < 0.05$, $^{**}p < 0.025$, $^{***}p < 0.01$.

effects are available upon request. Table 8 shows that the results are mostly driven by reform regimes in which many are affected, whereas the coefficients of the regimes in which few people are affected are mostly insignificant. It, however, is unclear if this is due to the (very) small number of observations in these regimes or due to a truly non-existent relationship.

4. The theoretical framework

This section develops a theoretical framework that can simultaneously explain why the cyclical state of the economy can trigger adjustments in pension generosity, while changes in the current and projected old-age dependency ratio are less capable in doing so. To highlight the intuition behind the mechanisms, we deliberately endow the model with the minimal set of features that allow us to replicate our main empirical findings. Although the basic set-up of the framework is relatively simple, it has not been deployed before in the literature. The model features some fixed cost of implementing reforms. Typically, the economic literature has worked with fixed adjustment costs of prices ("menu costs", see e.g. Mankiw, 1985) and capital adjustment costs. Here, we introduce a fixed cost associated with a change to the pension arrangement. In our context, the fixed reform cost is interpreted as a price to pay for overcoming political resistance to reform and the loss of popular support.

Table 7
Multinomial logit estimation with pressure indicator based on time since last reform and interaction with growth.

Independent variables	Expanding only		Contracting	
	Coefficient	Marg. Eff.	Coefficient	Marg. Eff.
$\overline{OAD25_t}$	−4.18***	−0.82***	5.04***	0.83***
	(1.54)	(0.22)	(1.44)	(0.18)
$OADDEV25_{it}$	3.00	0.42	0.86	0.0029
	(2.97)	(0.44)	(2.80)	(0.36)
$\Delta OAD25_{it}$	1.57	−0.077	8.41	1.07
	(7.24)	(1.05)	(6.84)	(0.88)
$GROWTH_{it}$	6.08	1.43***	−8.26	−1.45***
	(4.29)	(0.53)	(4.85)	(0.51)
DEF_{it}	−0.43	−0.32	6.86*	0.93**
	(3.16)	(0.46)	(3.17)	(0.41)
$UNEMPL_{it}$	0.86	−0.29	11.3***	1.48***
	(3.56)	(0.52)	(3.86)	(0.49)
$MAASTRICHT_{it}$	1.10***	0.12***	1.33***	0.14***
	(0.29)	(0.047)	(0.33)	(0.046)
$TsLR_{i,t}$	−0.028	−0.0042	0.083	0.011
	(0.055)	(0.0060)	(0.053)	(0.0059)
$GROWTH_{it} \times TsLR_{i,t}$	0.80		−0.47	
	(1.57)		(1.95)	
N	1081			
McFadden R2	0.15			

Note: standard errors in parentheses, *$p < 0.05$, **$p < 0.025$, ***$p < 0.01$.

Table 8
Multinomial logit estimation with "Many affected" and "Few affected".

Independent variables	Expanding only		Contracting	
	Many affected	Few affected	Many affected	Many affected
$\overline{OAD25_t}$	−3.26	−5.83**	5.32***	1.64
	(1.83)	(2.33)	(1.51)	(3.00)
$OADDEV25_{it}$	3.13	2.67	0.67	−0.36
	(3.45)	(4.94)	(2.93)	(6.49)
$\Delta OAD25_{it}$	−5.21	8.77	9.01	11.1
	(9.12)	(10.1)	(7.19)	(14.3)
$GROWTH_{it}$	11.5***	0.70	−10.4**	−1.40
	(4.38)	(5.31)	(4.19)	(9.00)
DEF_{it}	−1.82	0.20	7.59**	2.99
	(3.95)	(4.53)	(3.30)	(7.03)
$UNEMPL_{it}$	−1.59	5.34	9.03**	23.2***
	(4.35)	(5.26)	(4.02)	(8.43)
$MAASTRICHT_{it}$	1.31***	0.88*	1.24***	1.25
	(0.35)	(0.42)	(0.35)	(0.71)
N	1081			
McFadden R2	0.15			

Note: table reports coefficients. Marginal effects are available upon request. Standard errors in parentheses, *$p < 0.05$, **$p < 0.025$, ***$p < 0.01$.

4.1. Model setup

We consider an economy with a political party that runs the current government and discounts the future with a factor $0 < \pi < 1$. This factor captures both the government's innate time preference as well as a potential reduction in its effective discount factor resulting from the possibility of losing office to a political competitor. The current government cares – among other things – about the income position of the elderly as measured by the pension pay-out per retiree $P > 0$.[10] This pay-out is chosen taking into account current and future economic and demographic conditions. Current economic conditions are fully summarized by a business cycle indicator $Y > 0$. These economic conditions follow a Markov process with a stationary distribution. Hence, all shocks to Y are of a temporary nature.

The demography in our theoretical framework is captured by the old-age dependency ratio, again defined as the number of people of 65 years and older divided by the number of people in the age group 15–64 years. Ideally, forecasts about the old-age dependency ratio would be based on the population pyramid and the fundamental demographic forces: fertility, mortality and migration. In most countries, the current size of each cohort is well known. Levels of fertility and mortality change only slowly over time. Thus, when

[10] The positive value of P implies that transfers from the old to the young are excluded.

fertility or mortality is high in one year, this is likely also the case the next year. This implies a strong, but not perfect, serial correlation in these two variables. International migration is more volatile, but some degree of serial correlation should be expected, because the driving economic, legal, political and social conditions tend to change only slowly over time (e.g., Preston et al., 2000). In addition, in most developed countries, migrant flows in the various age categories are small relative to the existing cohorts. All these factors render the old-age dependency ratio relatively predictable in the short to medium term.[11]

Modelling the three fundamental sources of demographic change explicitly and keeping track of all the cohorts is well beyond the scope of this paper. Instead, we try to capture the current and projected future demographic situation with only the current old-age dependency ratio B and the long-run level b to which B is expected to converge. Such a long-run level exists if the demographic composition of the population is stable, which is the case when migration flows, age-specific fertility and age-specific mortality rates are constant. In view of the continuing medical progress, especially this last assumption seems unrealistic. However, if the length of the working life and the average retirement phase change proportionally with life expectancy this would continue to produce a roughly constant long-run old-age dependency ratio even when mortality rates are falling. We allow for fluctuations in migration flows, fertility and mortality reduction by assuming that b follows a (highly-persistent) time-stationary Markov process and that B slowly gravitates towards this long-run value. We also assume that b and B are independent of Y.

The government's instantaneous utility is fully determined by current economic and demographic conditions. Hence, it can be written as $U(P, Y, B)$. For simplicity, we assume that U is continuous and twice differentiable in all its arguments, strictly concave in P, and increasing and concave in Y. We exclude financing of the pension system with government debt, so there exists an upper-bound $\overline{P}(Y, B) < \infty$ on P, which is determined by Y and B. Strict concavity in P arises, for example, in a situation in which an endowment needs to be divided over various generations (a higher pension pay-out is at the cost of working generations) or when a higher pension pay-out reduces the resources available for other public spending. We denote the pension benefit that maximizes instantaneous utility by

$$\widetilde{P}(Y, B) = \underset{P}{\operatorname{argmax}} U(P, Y, B).$$

The literature (see, e.g., Gonzales-Eiras and Niepelt, 2008; and Ciurila and Romp, 2015) suggests that there are two opposing forces of the old-age dependency ratio B on the optimal pension pay-out chosen by a politician. A higher old-age dependency ratio raises the cost of the pension system, putting downward pressure on the individual pension pay-out, but it also increases the electoral weight of the group of retirees, which causes upward pressure on the individual pay-out. Hence, the overall effect on the individual pay-out could go into both directions. Ciurila and Romp (2015) conclude that in a probabilistic voting setting an office-seeking politician will divide the financial burden of a higher old-age dependency ratio over all generations. The retirees contribute with a reduction of their individual pension benefit.[12] This suggests a downward sloping relationship between B and the optimal individual pension pay-out $\widetilde{P}$, so $\widetilde{P}_B(Y, B) \leq 0$, which requires $U_{PB} \leq 0$. Finally, we assume that windfall gains are divided over both the active and retired part of the population, hence $U_{PY} \geq 0$, so $\widetilde{P}_Y(Y, B) \geq 0$.

A special case that satisfies these assumptions is when the government maximizes $U = f(G) + B\, v(P)$ subject to $G + BP = Y$, where the functions $f(\cdot)$ and $v(\cdot)$ are continuous, increasing and strictly concave and G is the amount of a public good. Hence, the government aims at optimally allocating a given endowment Y over the public good and pension provision. Another special case, which we will use later, is where $\widetilde{P}$ is proportional to Y and instantaneous utility only depends on the optimal pay-out relative to the current pay-out, so instantaneous utility can be written as $U = u(Y\widetilde{P}(1, B)/P)$, with $u(\cdot)$ strictly concave and $\widetilde{P}(1, B)$ positive, but decreasing in B. For now, we continue with the more general setup.

Increasing the pension benefit P comes at a fixed utility cost K^E, while decreasing the pension benefit comes with a fixed cost K^C. This cost may be incurred immediately or stretch over multiple periods. What matters is that, once the decision has been taken, the full cost must be incurred. Therefore, the relevant cost K^C may be the present value of a stream of period costs associated with a given change in the pension benefit. There is no cost attached to keeping the existing benefit unchanged. To keep the model tractable, we assume that the costs K^E and K^C are constant over time. We consider a discrete time setting, hence the government is free to implement a change at fixed moments, as long as it pays the corresponding fixed cost. The government's optimization problem is described by

$$V(P, Y, B, b) = \max_{P'} U(P', Y, B) - I\left(P' - P, K^E, K^C\right) + \pi E[V(P', Y', B', b')|Y, B, b] \tag{1}$$

where $I(\cdot)$ is the selection function

[11] Short-to medium-term estimates of mortality and fertility rates are very accurate; immigration does fluctuate, but net migration flows are small. Longevity and fertility mostly have long-term effects, but even the 25-year ahead average absolute forecast error for the old-age dependency ratio in our sample is only 2.5 percentage points.

[12] Various contributions, such as McDonald and Budge (2005), Disney (2007) and Sanderson and Scherbov (2007), point to the projected future increase in old-age dependency ratios and the fact that countries with older populations feature higher government social spending. However, empirical evidence on how an older population affects *individual* benefit levels is relatively sparse. Hollanders and Koster (2012) find that "… a higher age of the median voter leads to lower benefits per retiree", thereby providing some support for the above theoretical result.

$$I\left(P^{'}-P,K^{E},K^{C}\right)=\begin{cases}K^{E}, & P^{'}-P>0\\ 0, & P^{'}-P=0\\ K^{C}, & P^{'}-P<0\end{cases} \tag{2}$$

Primes denote values in the next period. The government's problem of selecting the optimal pension benefit $P^{'}$ is comparable to a standard optimal pricing problem with menu costs or any other (s,S)-type of model, but with one additional complexity: the government has additional information concerning the next period's situation via the long-term old-age dependency ratio b.

4.2. The relevance of current economic conditions

We first derive an analytical result that provides intuition why reform measures commonly coincide with business cycle fluctuations and are unrelated to (projected) demographic changes. For simplicity we focus on the symmetric case with $K^E = K^C = K$.

Each period the government compares the value of changing the pension system now (V^N) with the value of postponing a change (V^P). The government sets a new level $P^{'} \neq P$ for the pension benefit if and only if $V^N > V^P$. We use the convention that in the case of equality, the pension benefit remains unchanged. The value of changing the system now is

$$V^{N}(Y,B,b)=\max_{P^{'}}U(P^{'},Y,B)-K+\pi E[V(P^{'},Y^{'},B^{'},b^{'})|Y,B,b], \tag{3}$$

with $V(\cdot,\cdot,\cdot,\cdot)$ as defined in (1). The value of retaining the current pension benefit is given by

$$V^{P}(P,Y,B,b)=U(P,Y,B)+\pi E[V(P,Y^{'},B^{'},b^{'})|Y,B,b]. \tag{4}$$

Using the maximum instantaneous gain defined as

$$F(P,Y,B)\equiv U(\widetilde{P},Y,B)-U(P,Y,B)\geq 0, \tag{5}$$

we can formulate sufficient and necessary conditions for a change.

Proposition 1. If $F(P,Y,B) \leq (1-\pi)K$, then $V^P(P,Y,B,b) \geq V^N(Y,B,b)$, so postponing a benefit change is optimal. Hence, $P^{'} = P$. If $F(P,Y,B) > (1+\pi)K$, then $V^N(Y,B,b) > V^P(P,Y,B,b)$, so implementing a benefit change now is optimal. Hence, $P^{'} \neq P$.

Proof. see Appendix A.

Note that both conditions only depend on the current business cycle situation (Y), the current old-age dependency ratio (B) and the current pension payout (P). The demographic forecast b is potentially only relevant in the region not covered by the two inequalities in the above proposition, i.e. the region for which $(1-\pi)K < F(P,Y,B) \leq (1+\pi)K$.

The intuition behind Proposition 1, and in particular the dominant role of the current state variables and the current pension payout, is the following. Consider a government who inherits a pension benefit that is "currently fairly optimal" given the current state of the business cycle and the current old-age dependency ratio. That is, the maximum instantaneous gain from changing the benefit level does not exceed $(1-\pi)K$. At the same time this government is aware of the fact that the demographic forecasts are such that the inherited pension benefit is unsustainable in the future. Hence, it faces two options: change the pension benefit now to make the pension system future proof or keep the pension benefit at the inherited level and change it in the future. Postponing the change in pension benefit has three advantages over the first option. First, the government also postpones paying the fixed cost. Second, the government does not have to set a pension benefit that potentially lowers the current instantaneous utility. Third, in the next period, the government can freely choose the pension benefit that is optimal from that period onwards. Clearly, provided that the inherited benefit is "fairly optimal", it will never be optimal to change the pension benefit now in order to ensure sustainability of the system in the future. In other words, information about the future is irrelevant in this case.

A similar argument holds when the current economic situation is such that the inherited pension benefit is "currently far from optimal". That is, the maximum instantaneous gain exceeds $(1+\pi)K$. In this case, the government will always want to change the pension benefit, irrespective of the demographic forecasts. In taking a decision about the new benefit level, it may take knowledge about the future into account. On the one hand, if both the current state of the business cycle and the demographic forecast are such that the pension benefit should be cut, the government may want to set a new level that reduces the chance of another cut in the future, so as to avoid paying the adjustment cost again. On the other hand, if the current situation asks for an unsustainable increase in the generosity of the pension benefit, the government knows that it will have to change the pension benefit again in the next period. Hence, it might now just as well set the benefit to $\widetilde{P}$.

Rolling Equation (1) forward shows that the current value $V(P,Y,B,b)$ is the present value of all expected future instantaneous utilities minus the sum of all future adjustment costs. To trigger a reform measure, a change in one of underlying state variables (Y,B,b) must be large enough and, ideally, unexpected with a significant effect on current instantaneous utility. Business cycle fluctuations satisfy these requirements; changes in current and projected demographic conditions do not. Changes in the current demographic conditions are typically small, so they only result in small changes in the possible gain from resetting the benefit, which on an annual basis are not enough to offset the fixed adjustment costs. Second, these changes are expected, and should already largely have been incorporated in the continuation value, the last part of (1). Hence, the gain from further resetting the benefit would generally be small. The cumulative effect of small, predictable, changes in the current demographic conditions could at some point trigger a reform, but business cycle fluctuations are more frequent. Well before the cumulative effect of current demographic conditions has become large enough, likely a recession or boom will already have caused a reform. During this recession or boom, the pension benefit is reset, taking into account the current state of the economy and all the cumulative changes of current and projected future demographic conditions since the last reset.

Why are changes in the (long-term) *projected* demographic conditions unlikely to affect the decision to change the pension benefit? These changes are unexpected. However, they merely change instantaneous utility in the distant future. Future utility is discounted for two reasons: pure time discounting and political uncertainty. Especially political uncertainty results in a high discount rate, so the present value of the utility gains in the distant future is not sufficient to trigger a reform now. Further, a small discount factor shrinks the region between $(1-\pi)K$ and $(1+\pi)K$ where changes of future demographic conditions may trigger a reform. Finally, even if the government foresees sustainability problems due to averse ageing shocks, it can safely postpone a reform since it is highly likely that a business cycle fluctuation will trigger a reform well before sustainability problems have had a chance to materialize. This also explains our empirical finding why even cumulative changes in the projected old-age dependency ratio since the last reform or the time elapsed since the last reform do not or hardly affect the likelihood of reform. Before the evolution of these variables could have a significant bite, it is likely that a reform induced by the state of the economy has already eased the pressure to reform.

Contrary to the other demographic variables in our regressions, the OECD-wide average projected old-age dependency ratio did turn out to be empirically relevant. A potential explanation for this finding could be that the ongoing global ageing trend is raising awareness of its costs, causing the fixed costs of expansionary and contractionary measures to diverge. In particular, it becomes more difficult to implement expanding measures and easier to implement contracting measures. In a more general formulation of (1), this would be captured by introducing measure-specific adjustment costs that also depend on the old-age dependency ratio. The state-dependent adjustment cost of an expansionary measure would then rise with the old-age dependency ratio and that of a contractionary measure would gradually fall. Yet, the timing of reform measures would still be determined by the business cycle. However, a downturn of given size would make a contractionary measure more likely and a given upturn would make an expansionary measure less likely when the OECD-wide average projected old-age dependency ratio rises. Hence, a simultaneous trend increase in the cost of expansionary reform and trend decrease in the cost of contractionary reform could rationalize the increasing frequency of contractionary and decreasing frequency of expansionary pension reform measures reported in Section 2.

5. Implementation

We will now implement our theoretical framework in order to rationalize the role of business cycle fluctuations in explaining reform measures, while changes in demographic projections and accumulation of demographic pressure are unable to explain the timing of those measures. To do so, we need to further specify the model and calibrate it. To practically implement our model, we make five assumptions:

1. Instantaneous utility of the government only depends on *the ratio* of the instantaneously optimal pension benefit $\widetilde{P}$ and the current level of the pension benefit P.
2. This optimal benefit $\widetilde{P}$ is *proportional* to the business cycle indicator Y.
3. The logarithm of the business cycle indicator follows an AR(1) process:

$$\log(Y') = \varphi_y \log(Y) + \varepsilon, \varepsilon \sim N(0, \sigma_\varepsilon^2).$$

4. The current old-age dependency ratio gravitates geometrically to its long-run value:

$$B' = \lambda B + (1-\lambda)b, 0 < \lambda < 1.$$

5. The long-run old-age dependency ratio takes on discrete values $b^i \in (0,1)$ and features constant transition probabilities p_b^{ij}:

$$Prob(b' = b^j | b = b^i) = p_b^{ij}.$$

Assumption 2 is crucial to ensure that pensions vary with the business cycle,[13] in line with our empirical finding. However, this assumption does not explain why demographic conditions do not determine when pension reform measures are taken. The strongest assumption is Assumption 4. This clearly violates the non-monotonicity over time in the forecasts of the old-age dependency ratio. In most countries, the old age-dependency ratio is expected to peak between 2030 and 2050, and then gravitate towards a lower value. However, to model such a more realistic forecast one would need a higher-dimensional demographic model, which would complicate the model significantly, without adding additional insights. Our assumption is in line, though, with official predictions for the very long run.

Our dataset does not contain values of the pension pay-out, only the years when pension reform measures were enacted. This

[13] Of course, fluctuations in the state of the business cycle are not the only source of change in the pension pay-out. Generally, the past wage history helps to determine the pension entitlement. However, the entitlement tends to be indexed with inflation or nominal wage, which depends on the state of the business cycle.

implies that we can freely transform P as long as the jumps associated with reform of the benefit level are preserved. Moreover, $V(\cdot,\cdot,\cdot,\cdot)$, $U(\cdot,\cdot,\cdot)$ and both adjustment costs K^E and K^C are in utility terms, so we can also freely use affine transformations, without changing the optimal timing of reforms in the model. Under Assumption 2, we can write the optimal benefit $\widetilde{P}(Y,B)$ as $\widetilde{P}(Y,B) = Y\widetilde{P}(1,B)$. Now, use Assumption 1 to write instantaneous utility as

$$U(P,Y,B) = u(\widetilde{P}(Y,B)/P) = u(Y\widetilde{P}(1,B)/P).$$

By construction, $u(\cdot)$ has a maximum at 1. $\widetilde{P}(1,B)$ is positive and decreasing in B, so a natural first-order approximation is $\widetilde{P}(1,B) \approx Ae^{-cB}$, where $A > 0$ is a scaling factor and $c > 0$ measures the semi-elasticity of the optimal pay-out $\widetilde{P}$ with respect to B. Under Assumption 3, the unconditional variance of $\log Y$ is $\sigma_y^2 \equiv \sigma_\varepsilon^2/(1-\varphi_y^2)$. Now, define $y \equiv \frac{1}{\sigma_y}\log(Y)$, $p \equiv \frac{1}{\sigma_y}\log[P/A]$. $k^E \equiv -\frac{1}{2}u''(1)\sigma_y^2 K^E$, $k^C \equiv -\frac{1}{2}u''(1)\sigma_y^2 K^C$, and $\gamma \equiv c/\sigma_y$. Then, we can rewrite the full optimization problem (1) up to a second-order approximation as the following minimization problem which has the same timing of reform measures:

$$v(p,y,B,b) = \min_{p'}(y-\gamma B-p')^2 + I\left(p'-p,k^E,k^C\right) + \pi E[v(p',y',B',b')|y,B,b] \tag{6}$$

$$y' = \varphi_y y + \xi_t,\ \xi_t \sim N\left(0, 1-\varphi_y^2\right) \tag{7}$$

$$B' = \lambda B + (1-\lambda)b \tag{8}$$

$$Prob\left(b' = b^j \middle| b = b^i\right) = p_b^{ij} \tag{9}$$

Despite the apparent simplicity, we are unable to obtain an analytic solution to this model. The literature (in particular, the standard reference Stokey, 2009) does offer some leads, but in the final steps we always have to rely on a numerical solution.[14]

6. Calibration

We discretize the business cycle indicator y to mimic the AR(1) process as closely as possible using 51 grid points. These grid points and the corresponding transition probabilities are determined using the Rouwenhorst (1995) method (see Kopecky and Suen, 2010; for a formal analysis).[15] For the AR-coefficient φ_y, we use the typical value from the business cycle literature of 0.8 per quarter, which translates to 0.41 per year.

The discount factor π captures time discounting and the probability of the current government losing power. One period corresponds to one year, so we set time discounting to 0.98, corresponding to a time preference rate of about 2% a year, which is in the ballpark range of the literature. The probability of losing power is highly country specific. We use the data of Armingeon et al. (2018a) to calculate the yearly probability of an ideological change.[16] This probability varies between 0 and 50 percent. The probability of losing power clearly dominates the traditional time discounting. In our baseline simulations we use a value of 0.75 for π. Sensitivity analysis shows that our results are robust, even for extreme values for π of 0.5 and 0.9.

To calibrate the demographic processes, we choose the minimum and maximum values of the projected 25-year ahead old-age dependency ratios. The minimum value in our data is 13% (Japan, 1970) and the maximum value is 63% (also Japan, 2014). To capture all the fluctuations of this projected old-age dependency ratio, we divide the assumed 10–70% range into 6 equally sized bins of 10 percentage points each. This yields 6 central points ranging from 15% to 65%, which we use as the grid points for our long-run old-age dependency ratio b.[17] In our dataset, the forecast of the old-age dependency ratio changes bins in 8% of the years, so we set the diagonal values of the transition matrix to 92% and all off-diagonal values to (8/5)%, given that there are 5 other intervals the ratio could jump to. Our results are robust to alternative divisions of the remaining 8% over the off-diagonal values. For the current old-age dependency ratio we use an evenly-spaced grid of 21 points ranging from 15% to 65%. The adjustment parameter λ of the old-age dependency ratio is set to generate a half time of 25 years, so $\lambda^{25} = 0.5$.

This leaves us with the choices of the adjustment cost parameters $k^E, k^C > 0$ and the effect of the old-age dependency ratio on the optimal pension benefit, as summarized by $\gamma > 0$. We can use these three parameters without a clear empirical counterpart to match

[14] We explored a continuous-time version of the model, which features three exogenous state variables: the business cycle and the two-variable demographic framework. The Hamilton-Jacobi-Bellman (HJB) equation is a three-dimensional linear non-homogeneous stochastic differential equation, but with non-constant coefficients since the process of the current old-age dependency (OAD) ratio depends on the long-run OAD ratio. We can only solve this HJB equation with the given boundary conditions using numerical methods, which renders it impossible to derive interesting analytical properties. Hence, since we would in any case have to resort to numerical methods to solve our model, we have opted for a discrete-time version of the model here, which is preferable to the continuous-time version as it allows for a simpler presentation.

[15] With this number of grid points, the specific choice of the discretization method is immaterial; indeed, the Tauchen-method gives similar results.

[16] We take the variable that measures the ideological composition of the government through the percentages of total cabinet posts held by right-, centre- and left-wing parties. The probability of an ideological change is the fraction of years that this variable changes by at least 10 percent. It ranges from zero percent for Switzerland to around 50 percent for Italy.

[17] The alternative would be to use country-specific grids. However, the number of grid points is sufficiently high for such refinements not to matter. Hence, we stick to the use of a common grid across countries.

the model to the observed pension reform regimes.

The adjustment cost parameters are closely related to the number and nature of reform measures in each country, which varies widely across the countries. In our sample period 1970–2017, France implemented at least one reform measure in 30 out of the 48 years (63% of all years), while Iceland merely implemented one reform measure in on average 5 years (10% of all years). Moreover, the share of reforms with an expanding nature also varies wildly. In most countries, the share of "Expanding only" regimes as a fraction of all years with at least one reform measure falls between 40 percent and 60 percent. There are, however, clear outliers like Japan which has only four years in which it implemented only expanding reforms, compared to 11 years in which it implemented only contracting reforms or a combination of expanding and contracting. Iceland is also a clear outlier; policymakers only legislated reforms with an expanding nature according to our database. In view of the different reform frequencies and different nature of these reforms across our sample countries, we allow for country-specific, two-sided adjustment costs. Differences in the adjustment costs are likely related to differences in the political resistance to reform.

For each country-year, the model predicts one of three possible outcomes: no reform (NONE), an expanding reform (EXP), or a contracting reform (CON). The data is classified accordingly, with the "Expanding only" regime as EXP and the "Contracting" regime as CON, and if none of these regimes applies, the data is NONE. To determine the predicted reforms in the model for a given set of parameter values, we simulate each country using year-to-year economic growth as the business cycle indicator y.[18] The current and 25-year ahead old-age dependency ratios are used for the current old-age dependency ratio B and long-run old-age dependency ratio b, respectively. We start the simulation setting the pension pay-out equal to the instantaneous optimal one implied by the actual economic and demographic situation in the first sample year. Then, for every year, we check whether the economic situation, the current old-age dependency ratio and the 25-year ahead old-age dependency ratio in the actual data have changed enough to make it optimal to change the pension pay-out. If this is the case, then we choose the new optimal pension pay-out which leads to either a predicted contracting reform or a predicted expanding reform.

7. Model fit

We fix the transition probabilities and the discount factor π to the values given in the previous subsection, but the three remaining parameters – the effect of the old-age dependency ratio on the optimal pension benefit, γ, and the two adjustment cost parameters – have no empirical counterparts. These three parameters are used to fit the model to reality.

We choose the three free parameters such that we maximise a measure of fit similar to the likelihood in the multinomial logit regression used in the empirical part of this paper. In the multinomial logit regressions, the coefficients generate a probability of a reform regime (NONE, EXP, and CON) for each country-year. In the theoretical model, we lack the concept of a probability of a reform since the model simply predicts one of the three possible outcomes. To arrive at a measure similar to the likelihood used in the multinomial logit regressions, we assign some high predicted probability $0 < \beta < 1$ to the regime prediction by the model, and some low (but not zero) probability $0 < \alpha < \beta$ to the other outcomes.[19] That is, for each country i in year t we define the "probability of regime r" as

$$p_{r,it} = \begin{cases} \beta & \textit{if model predicts regime } r \\ \alpha & \textit{otherwise} \end{cases}$$

Using this definition for probabilities, we can define the likelihood of the model for a country i, given a set of parameters, similarly to the usual likelihood of a multinomial logit regression. In Appendix B, we show that maximising this likelihood of the model is equivalent to maximising the number of correctly predicted regimes.

Following this procedure to the letter creates a bias towards the NONE regime because it is very attractive to predict many NONE regimes since this is by far the largest category, and it is trivial to predict many NONE regimes; simply set the adjustment cost parameters high. Predicting the EXP and CON regimes correctly is much harder, because the legislation procedures often lead to a one-year delay, or even longer, where the model predicts an instantaneous response. To correct for this, we count a prediction as correct if the predicted regime occurs in the current year or one year later, and we put additional weight on correctly predicted expansionary and contractionary regimes. In choosing this additional weight we have to make a trade off. A higher weight on these two regimes pushes the calibrated adjustment costs down since this generates more predicted EXP and CON regimes, which increases the probability of correctly predicting one of those regimes. A numerical sensitivity analysis shows that assigning twice the weight of the NONE regimes in the EXP and CON regimes improves the probability of correctly predicting those regimes, without a significant loss of the probability of a correctly predicted NONE regime.

Table 9 presents the values of the calibrated parameters. The parameters in Table 9 should be interpreted with some care: first, the sample is merely 48 years, which is rather small in the presence of a slow-moving demographic process. Second, the parameters that maximise the likelihood or equalize the predicted regimes with the observed regimes are not unique. We performed a three-dimensional grid search with in total 4 different values for γ, and 17 values for each of the two adjustment cost parameters, giving a grid of 1156 points. If multiple grid points gave the same weighted number of correct predictions, then we chose those parameter values that correctly predicted the largest number of expanding and contracting regimes. In those cases, in which there was still a tie,

[18] We experimented using the first principal component of growth, unemployment and the government's deficit as the business cycle indicator. The various changes in the definition of especially unemployment pollute these results, however.

[19] The probabilities for each country-year should obviously sum to 1, so for our three possible outcomes, we have $\alpha + \alpha + \beta = 1$.

Table 9
Calibration targets and calibrated parameters.

Country	Total # correct predictions	#Predictions/#Correct predictions/#Observed regimes					
		EXP	CON	NONE	γ	k^E	k^C
Australia	32	18/10/14	16/11/15	13/11/18	5	0	0.4
Austria	30	9/6/8	21/10/12	17/14/27	3	0.75	0
Belgium	32	15/7/12	15/8/10	17/17/25	1	0.3	0.5
Canada	35	9/6/13	18/8/9	20/21/25	1	1	0
Denmark	34	11/3/7	9/5/11	27/26/29	5	0.3	0.75
Finland	34	18/10/14	6/6/12	23/18/21	3	0	2
France	25	18/9/14	24/14/16	5/2/17	5	0	0
Germany	27	14/7/14	14/3/10	19/17/23	1	0.3	0.5
Greece	32	16/8/9	16/9/11	15/15/27	2	0	0.1
Iceland	43	13/5/5	4/0/0	30/38/42	5	0.1	3
Ireland	29	21/12/14	13/2/6	13/15/27	3	0	0.3
Italy	32	6/2/6	12/6/14	29/24/27	2	2	0.4
Japan	37	8/3/4	9/6/11	30/28/32	1	1	0.5
Luxembourg	40	7/5/8	7/2/4	33/33/35	3	1.5	1.25
Netherlands	33	12/9/16	7/2/6	28/22/25	5	0.1	1.5
New Zealand	39	8/2/5	4/4/8	35/33/34	5	1	5
Norway	38	4/1/7	3/1/4	40/36/36	5	2.5	5
Portugal	30	6/4/9	12/8/17	29/18/21	2	0.75	0.3
Spain	30	9/4/9	8/3/10	30/23/28	2	0.4	0.75
Sweden	32	15/3/5	9/4/10	23/25/32	5	0	1.5
Switzerland	35	9/4/8	12/7/13	26/24/26	1	1.25	0.3
United Kingdom	34	4/0/7	11/7/11	32/27/29	5	3	0
United States	37	5/6/11	14/7/10	28/24/26	5	5	0
Average	33.5	11.1/5.5/9.5	11.5/5.8/10	24.4/22.2/27.5	3.3	0.92	1.05

Note: A prediction is correct if the predicted regime (EXP, CON, or NONE) is observed in the current year, or one year later.

we chose those parameter values predicting the highest number of expanding regimes.

Three interesting observations can be made. First, a high adjustment cost for reform into one direction also reduces reforms into the other direction. The reason is that the policymaker will be hesitant to move the arrangement into this latter direction realizing that a potential future move into the other direction is costly. Finland, for example, has 12 years with a contractionary regime and 14 years with an expansionary regime. The calibration procedure assigns a fixed cost of 2 to a contractionary regime, leading to six correctly

Table 10
Effect of country-specific economic and demographic conditions.

	Country specific parameters						Common parameters (equal to average values)		
	Parameters			#Predictions			#Predictions		
Country	γ	k^E	k^C	EXP	CON	NONE	EXP	CON	NONE
Australia	5	0	0.4	18	16	13	8	9	30
Austria	3	0.75	0	9	21	17	6	9	32
Belgium	1	0.3	0.5	15	15	17	8	11	28
Canada	1	1	0	9	18	20	9	10	28
Denmark	5	0.3	0.75	11	9	27	6	8	33
Finland	3	0	2	18	6	23	7	8	32
France	5	0	0	18	24	5	5	8	34
Germany	1	0.3	0.5	14	14	19	6	9	32
Greece	2	0	0.1	16	16	15	5	6	36
Iceland	5	0.1	3	13	4	30	8	8	31
Ireland	3	0	0.3	21	13	13	7	6	34
Italy	2	2	0.4	6	12	29	7	9	31
Japan	1	1	0.5	8	9	30	6	6	35
Luxembourg	3	1.5	1.25	7	7	33	9	8	30
Netherlands	5	0.1	1.5	12	7	28	8	10	29
New Zealand	5	1	5	8	4	35	11	10	26
Norway	5	2.5	5	4	3	40	6	6	35
Portugal	2	0.75	0.3	6	12	29	5	7	35
Spain	2	0.4	0.75	9	8	30	8	7	32
Sweden	5	0	1.5	15	9	23	8	11	28
Switzerland	1	1.25	0.3	9	12	26	9	10	28
United Kingdom	5	3	0	4	11	32	7	8	32
United States	5	5	0	5	14	28	8	9	30
Average	3.3	0.92	1.05	11.09	11.48	24.43	7.26	8.39	31.35

Note: the last three columns fix the cost parameters for each country at the country averages reported in the final line of the table.

predicted contractions. Despite the zero fixed cost of an expansionary regime, the model only generates 18 years with a predicted expansionary regime. Similar results are found for e.g. Iceland, and Canada.

A second observation is that the average adjustment cost of an expansionary regime is slightly lower than that of a contractionary regime. This is as expected since an expansion would usually be met with less political resistance. The difference, however, is small, and there is significant heterogeneity between countries. For 9 out of 23 countries in our sample, the cost of expanding the pension arrangement exceeds the cost of contracting the arrangement.

Finally, for some countries, economic and demographic conditions seem to explain most of the reform regimes. Australia, for example, has 14 observed expanding regimes and 15 contracting regimes. Surprisingly, the adjustment cost of the expanding regime is negligible, while that of the contracting regime is 0.4. Despite this higher cost of contraction, the state of the economy and the demography warrant more frequent contracting than expanding.

Differences among countries are due to different fundamental parameters and due to different economic and demographic developments. In Table 10 we disentangle these two effects by setting the model parameters to their average values for all countries. Doing so, differences among countries are solely driven by different economic and demographic developments. The distributions of the adjustment parameters are positively skewed, so setting these parameters to their averages results in relatively high adjustments costs. This explains the low number of predicted expansionary and contractionary regimes in Table 10 compared to Table 9. On average, our model now predicts 7.3 expansionary regimes per country and 8.4 contractionary regimes.

Somewhat surprisingly, some countries with many observed expansionary and contractionary regimes, like France and Portugal, are now left with only few predicted expansionary and contractionary regimes. For these countries, the model predicts many small changes to the pension pay-out, because of low adjustment costs. A careful reading of the reforms in our dataset confirms this: these countries spread out their pension reforms over adjacent years, where certainly in the 2000s, later reforms offset or reinforce reforms earlier in the wave. In Belgium and Canada, on the other hand, the many reforms seem to be driven by its economic and demographic fluctuations, since the number of predicted reform regimes hardly responds to higher adjustment costs.

Table 11 shows the total score for each country, keeping the two adjustment costs at their optimal level, while varying the demographic parameter between 1 and 5. This table shows that the demographic parameter has a minor role in the explanatory power of the model, again indicative of the minor role of demographic fluctuations in explaining the timing of pension reforms. The average difference between the best and the worst score in terms of correct predictions of expanding and contracting regimes is 2.1, with zero effect on the score for five countries.

8. Conclusion and extensions

This paper has shown that the timing of pension reform measures is linked to the state of the business cycle, while changes in

Table 11
Effect of varying the demographic parameter keeping the cost parameters at their original values.

			γ				
Country	k^E	k^C	1.0	2.0	3.0	5.0	Min - Max
Australia	0	0.4	30	32	30	32	2
Austria	0.75	0	26	26	30	26	4
Belgium	0.3	0.5	32	30	30	30	2
Canada	1	0	35	34	34	33	2
Denmark	0.3	0.75	33	33	34	34	1
Finland	0	2	27	29	34	27	7
France	0	0	26	26	26	27	1
Germany	0.3	0.5	27	25	24	25	3
Greece	0	0.1	28	32	30	30	4
Iceland	0.1	3	43	43	43	43	0
Ireland	0	0.3	29	29	29	28	1
Italy	2	0.4	29	32	31	30	3
Japan	1	0.5	37	36	36	36	1
Luxembourg	1.5	1.25	35	35	40	35	5
Netherlands	0.1	1.5	31	31	32	33	2
New Zealand	1	5	39	39	39	39	0
Norway	2.5	5	38	38	38	38	0
Portugal	0.75	0.3	30	30	27	27	3
Spain	0.4	0.75	30	30	28	28	2
Sweden	0	1.5	31	31	31	32	1
Switzerland	1.25	0.3	35	32	32	31	4
United Kingdom	3	0	33	33	34	34	1
United States	5	0	37	37	37	37	0
Average	0.92	1.05	32.2	32.3	32.6	32.0	2.1

Note: we keep the cost parameters fixed at the originally estimated values reported in Table 7 and vary the demographic parameter over a common set of values for all countries. The columns under the different values of γ give the number of correctly predicted regimes.

current and projected old-age dependency ratios play no role in explaining when reforms take place. The demography matters only in the sense that the OECD-wide demography explains the general reform trend. However, the trigger to release the pressure in the pension system is the business cycle. We find this to be the case also when we include in our regressions demographic pressure, as for example captured by the change in the (projected) old-age dependency ratio since the last pension reform. We have rationalized these empirical findings with a theoretical framework that features fixed costs of pension reform, both expansionary and contractionary. Simulation of the model demonstrates that it is optimal for a government to postpone adjustments after demographic shocks until a business cycle shock triggers an adjustment; the adjustment triggered by a business cycle movement responds optimally to all information about current and future demographic forecasts. The model does a good job in predicting both the timing and direction of a reform based on the actual values of the business cycle indicator and the demographic data. Future research could explore refinements of the theoretical framework allowing for state-dependent fixed costs of expanding and contracting reform measures. Such a refinement might help to rationalize the observed trend increase in contractionary reforms and the trend fall in expanding reforms.

Our analysis also points to potential policy implications. First, while annual changes in old-age dependency ratios have no effect on the timing of reforms, we do see a relationship between the trend in projected old-age dependency ratios and both expanding and contracting reforms, consistent with the advice of international institutions and increasing public and political awareness of the need to reform. These world-wide trends could make it easier for individual countries that feature political opposition to reform to sell such reform in the domestic political arena and to borrow from best practices in reform elsewhere. Second, apparently, a weak state of the business cycle is politically speaking the most opportune moment to introduce contractionary reforms. At the same time, it is the moment when reductions in disposable income are least welcome. Therefore, the question is whether there are ways to encourage such contractionary reforms at moments when the economy is in a better shape to handle them. A possibility would be to try to introduce some form of automaticity in measures to enhance the sustainability of pension arrangements, for example by introducing an automatic link between increases in life expectancy and the retirement age, such as was done in the Netherlands in 2012. Ideally, such link should be introduced sufficiently far in advance to ride on the effect of the discounting of future changes and reduce the political resistance to signing such a link into law. In addition, it seems desirable to make the link as "smooth" as possible, so it would be best to have frequent but small changes in the retirement age, since each small change may be insufficient to muster enough political resistance to actually implementing it. Of course, this argument abstracts from the potential administrative costs of implementing frequent changes in the retirement age.

Declaration of competing interest

Both authors, Ward Romp and Roel Beetsma, declare absence of any conflict of interest in this research.

Data availability

Data will be made available on request.

Appendix A. Proof of Proposition 1

Define V^N and V^P as in equations (3) and (4). An upper bound for V^N is given by a situation in which the pension benefit can be changed in the next period without adjustment cost, so

$$V^N(Y,B,b) \leq U(\widetilde{P},Y,B) - K + \pi E\left[\max_{P'} V(P',Y',B',b')\middle| Y,B,b\right]. \tag{10}$$

Analogously, a lower bound for V^N is attained when the government changes the pension benefit both now *and* in the next period, paying the adjustment costs:

$$V^N(Y,B,b) \geq U(\widetilde{P},Y,B) - (1+\pi)K + \pi E\left[\max_{P'} V(P',Y',B',b')\middle| Y,B,b\right]. \tag{11}$$

A lower bound for V^P is given by a situation in which the government changes the benefit in the next period:

$$V^P(P,Y,B,b) \geq U(P,Y,B) - \pi K + \pi E\left[\max_{P'} V(P',Y',B',b')\middle| Y,B,b\right], \tag{12}$$

while an upper bound is given by a situation in which the government can freely change the pension benefit in the next period without adjustment cost:

$$V^P(P,Y,B,b) \leq U(P,Y,B) + \pi E\left[\max_{P'} V(P',Y',B',b')\middle| Y,B,b\right] \tag{13}$$

Subtracting (12) from (10) and combining with the assumption in Proposition 1 yields

$$V^N(Y,B,b) - V^P(P,Y,B,b) \leq U(\widetilde{P},Y,B) - U(P,Y,B) - (1-\pi)K = F(P,Y,B) - (1-\pi)K \leq 0,$$

hence the first result follows immediately. Subtracting (13) from (11) and combining with the assumption in Proposition 1 yields

$$V^N(Y,B,b) - V^P(P,Y,B,b) \geq U(\widetilde{P},Y,B) - U(P,Y,B) - (1+\pi)K = F(P,Y,B) - (1+\pi)K > 0,$$

hence the second result follows immediately as well.

Appendix B. maximising the likelihood of the theoretical model

The likelihood function of a multinomial logit model for country i is defined as

$$L_i = \prod_t \left(\prod_r {p_{r,it}}^{I_r(y_{it})} \right)$$

where the probability in the multinomial logit is determined by the standard logit function. In our model, we define the probabilities as

$$p_{r,it} = \begin{cases} \beta & \textit{if model predicts regime } r \\ \alpha & \textit{otherwise} \end{cases}$$

and the usual indicator function

$$I_r(y_{it}) = \begin{cases} 1 & \textit{if } y_{it} = r \\ 0 & \textit{otherwise} \end{cases}$$

This gives four possible outcomes for ${p_{r,it}}^{I_r(y_{it})}$ in the likelihood function

$${p_{r,it}}^{I_r(y_{it})} = \begin{cases} \alpha^0 = 1 & \textit{if } p_{r,it} = \alpha,\ I_r(y_{it}) = 0\ (\textit{correct}) \\ \alpha^1 = \alpha & \textit{if } p_{r,it} = \alpha,\ I_r(y_{it}) = 1\ (\textit{incorrect}) \\ \beta^0 = 1 & \textit{if } p_{r,it} = \beta,\ I_r(y_{it}) = 0\ (\textit{incorrect}) \\ \beta^1 = \beta & \textit{if } p_{r,it} = \beta,\ I_r(y_{it}) = 1\ (\textit{correct}) \end{cases}$$

For each country-year, the prediction by the model generates exactly one $p_{r,it} = \beta$; all others are α. Similarly, the indicator function $I_r(y_{it})$ is one for one regime, and zero for all others. For each country-year combination, the model prediction is either correct or incorrect, so for each country-year combination we have for the product over all outcomes in the likelihood function

$$\prod_r {p_{r,it}}^{I_r(y_{it})} = \begin{cases} \beta & \textit{if model correctly predicted } y_{it} \\ \alpha & \textit{if model is incorrect} \end{cases}$$

and the likelihood reduces to

$$L_i = \prod_t \left(\prod_r {p_{r,it}}^{I_r(y_{it})} \right) = \beta^{\#correct} \alpha^{T-\#correct} = \alpha^T \left(\frac{\beta}{\alpha} \right)^{\#correct}$$

with #correct the number of correct predictions and T the (fixed) number of years. This shows that for $0 < \alpha < \beta < 1$ maximising the "likelihood" is equivalent to maximising the number of correct predictions.

References

Abiad, A., Mody, A., 2005. Financial reforms: what shakes it? What shapes it? Am. Econ. Rev. 95 (1), 66–88.
Arcand, J.L., Berkes, E., Panizza, U., 2015. Too much finance? J. Econ. Growth 20 (2), 105–148.
Armingeon, K., Wenger, V., Wiedemeier, F., Isler, C., Knöpfel, L., Weisstanner, D., Engler, S., 2018a. *Comparative Political Data Set 1960-2016*. Bern. Institute of Political Science, University of Berne.
Armingeon, K., Wenger, V., Wiedemeier, F., Isler, C., Knöpfel, L., Weisstanner, D., 2018b. Supplement to the Comparative Political Data Set – Government Composition 1960-2016. Bern: Institute of Political Science, University of Berne.
Beck, T., Levine, R., 2004. Stock markets, banks, and growth: panel evidence. J. Bank. Finance 28 (3), 423–442.
Beetsma, R., Debrun, X., 2007. The new stability and growth Pact: a first assessment. Eur. Econ. Rev. 51 (2), 453-478.
Beetsma, R., Uhlig, H., 1999. An analysis of the stability and growth Pact. Econ. J. 109 (458), 546–571.
Beetsma, R., Klaassen, F., Romp, W., van Maurik, R., 2020. What drives pension reforms in the OECD? Econ. Pol. 35 (102), 357–402.
Bertola, G., Boeri, T., 2002. EMU labour markets two years on: microeconomic tensions and institutional evolution. In: Buti, M., Sapir, A. (Eds.), EMU and Economic Policy in Europe: the Challenges of the Early Years. Edward Elgar, Aldershot.
Bi, H., Zubairy, S., 2019. Public Pension Reforms and Fiscal Foresight: Narrative Evidence and Aggregate Implications. *Mimeo*, Federal Reserve Bank of Kansas City and Texas A&M University.
Blinder, S., Krueger, A., 2004. What Does the Public Know about Economic Policy, and How Does it Know it? NBER Working Paper, No. 10787.
Bonfiglioli, A., Gancia, G., 2015. Economic Uncertainty and Structural Reforms. CEPR Discussion Paper, No.10937.
Bouis, R., Duval, R., 2011. Raising Potential Growth after the Crisis: a Quantitative Assessment of the Potential Gains from Various Structural Reforms in the OECD Area and beyond. OECD Discussion Paper, No. 835.
Brooks, S., 2007. Globalization and pension reforms in Latin America. Lat. Am. Polit. Soc. 49 (4), 31–62.

Bumann, S., Lensink, R., 2016. Capital account liberalization and income inequality. J. Int. Money Finance 61, 143–162.
Campos, N.F., Kinoshita, Y., 2010. Foreign direct investment and structural reforms: evidence from Eastern Europe and Latin America. IMF Staff Pap. 57 (2), 326–365.
Campos, N., Hsiao, C., Nugent, J., 2010. Crises, what crises? New evidence on the relative roles of political and economic crises in begetting reforms. J. Dev. Stud. 46 (10), 1670–1691.
Campos, N.F., De Grauwe, P., Ji, Y., 2017. Structural Reforms, Growth and Inequality: an Overview of Theory, Measurement and Evidence. IZA Discussion Paper, No. 11159.
Cette, G., Fernald, J., Mojon, B., 2016. The pre-Great Recession slowdown in productivity. Eur. Econ. Rev. 88 (C), 3–20.
Chinn, M.D., Ito, H., 2006. What matters for financial development? Capital controls, institutions, and interactions. J. Dev. Econ. 81 (1), 163–192.
Christiansen, L., Schindler, M., Tressel, T., 2013. Growth and structural reforms: a new assessment. J. Int. Econ. 89 (2), 347–356.
Ciurila, N., Romp, W., 2015. The Political Arrangement of Pay-As-You-Go Pension Systems in the Presence of Financial and Demographic Shocks. *Netspar Discussion Paper*. No. 12/2015-037.
D'Amato, M., Galasso, V., 2010. Political intergenerational risk sharing. J. Publ. Econ. 94 (9–10), 628–637.
De Haan, J., Sturm, J.E., 2017. Finance and income inequality: a review and new evidence. Eur. J. Polit. Econ. 50 (C), 171–195.
De Macedo, J.B., Martins, J.O., Rocha, B., 2014. Are complementary reforms a "luxury" for developing countries? J. Comp. Econ. 42 (2), 417–435.
Demekas, D.G., Horváth, B., Ribakova, E., Wu, Y., 2007. Foreign direct investment in European transition economies — the role of policies. Journal of Comparative conomics 35 (2), 369–386.
Disney, R., 2007. Population ageing and the size of the welfare state: is there a puzzle to explain? Eur. J. Polit. Econ. 23 (2), 542–553.
Falcetti, E., Lysenko, T., Sanfey, P., 2006. Reforms and growth in transition: Re-examining the evidence. J. Comp. Econ. 34 (3), 421–445.
Fatas, A., 2016. The agenda for structural reform in Europe. In: Caselli, F., Centeno, M., Tavares, José (Eds.), After the Crisis: Reform, Recovery, and Growth in Europe. Oxford University Press, Oxford, U.K.
Fiori, G., Nicoletti, G., Scarpetta, S., Schiantarelli, F., 2012. Employment effects of product and labour market reforms: are there synergies? Econ. J. 122 (558), 79–104.
Furceri, D., Loungani, P., 2015. The distributional effects of capital account liberalization. J. Dev. Econ. 130 (C), 127–144.
Gonzales-Eiras, M., Niepelt, D., 2008. The future of social security. J. Monetary Econ. 55, 197–218.
Griffith, R., Huergo, E., Mairesse, J., Peters, B., 2006. Innovation and productivity across four European countries. Oxf. Rev. Econ. Pol. 22 (4), 483–498.
Gruber, J., Wise, D.A., 2009. Social Security Programs and Retirement Around the World: Fiscal Implications of Reform. University of Chicago Press, Chicago, IL.
Hollanders, D., Koster, F., 2012. The Greying of the Median Voter. CentER Discussion Paper Series, No. 2012-061.
ILO, 2019. International Labour Organization: NATLEX: Database of National Labour, Social Security and Related Human Rights Legislations. http://www.ilo.org/dyn/natlex. access date: March 1 and March 20, 2019.
ISSA, 2019. https://www.issa.int/en_GB/country-profiles. (Accessed 8 March 2019).
James, E., Brooks, S., 2001. The political economy of structural pension reform. In: Holzmann, R., Stiglitz, J. (Eds.), New Ideas about Old Age Security. World Bank, Washington DC.
Kneller, R., Morgan, C.W., Kanchanahatakij, S., 2008. Trade liberalisation and economic growth. World Econ. 31 (6), 701–719.
Kopecky, K.A., Suen, R.M.H., 2010. Finite state markov-chain approximations to highly persistent processes. Rev. Econ. Dynam. 13, 701–714.
LABREF, 2019. https://webgate.ec.europa.eu/labref/public/. (Accessed 17 April 2019).
Mahmalat, M., Curran, D., 2018. Do crises induce reform? A critical review of conception, methodology and empirical evidence of the 'crisis hypothesis. J. Econ. Surv. 32 (3), 613–648.
Mankiw, N.G., 1985. Small menu costs and large business cycles: a macroeconomic model of monopoly. Q. J. Econ. 100 (2), 529–538.
McDonald, M.D., Budge, I., 2005. Elections, Parties, Democracy: Conferring the Median Mandate. Oxford University Press, Oxford.
Orenstein, M., 2005. The new pension reform as global policy. Global Soc. Pol. 5 (2), 195–202.
Orenstein, M., 2013. Pension privatization: evolution of a paradigm. Governance: An International Journey of Policy, Administration, and Institutions 26 (2), 259–281.
Persson, T., Tabellini, G., 2000. Political Economics: Explaining Economic Policy. MIT Press, Cambridge, MA.
Prati, A., Onorato, M.G., Papageorgiou, C., 2013. Which reforms work and under what institutional environment? Evidence from a new data set on structural reforms. Rev. Econ. Stat. 95 (3), 946–968.
Preston, S., Heuveline, P., Guillot, M., 2000. Demography: Measuring and Modeling Population Processes. John Wiley and Sons, Hoboken, NJ.
Quinn, D.P., Toyoda, A.M., 2008. Does capital account liberalization lead to growth? Rev. Financ. Stud. 21 (3), 1403–1449.
Ranciere, R., Tornell, A., 2015. Why Do Reforms Occur in Crises Times? *Mimeo*. IMF and UCLA.
Razin, A., Sadka, E., Swagel, P., 2002. The aging population and the size of the welfare state. J. Polit. Econ. 110 (4), 900–918.
Rodrik, D., 1996. Understanding economic policy reform. J. Econ. Lit. 34, 9–41.
Rouwenhorst, K.G., 1995. Asset pricing implications of equilibrium business cycle models. In: Cooley, T.F. (Ed.), Frontiers of Business Cycle Research. Princeton University Press, Princeton, NJ, pp. 294–330.
Sanderson, W.C., Scherbov, S., 2007. A near electoral majority of pensioners: prospects and policies. Popul. Dev. Rev. 33 (3), 543–554.
Spilimbergo, A., Prati, A., Ostry, J.D., 2009. Structural Reforms and Economic Performance in Advanced and Developing Countries. IMF Working Paper, No. 268.
Stokey, N.L., 2009. The Economics of Inaction: Stochastic Control Models with Fixed Costs. Princeton University Press, Princeton, NJ.
Swagel, P., 2015. Legal, political, and institutional constraints on the financial crisis policy response. J. Econ. Perspect. 29 (2), 107–122.
Tommasi, M., 2017. Crisis, Political Institutions, and Policy Reform: the Good, the Bad, and the Ugly. *Mimeo*, Universidad de San Andrés and Yale University.
Tommasi, M., Velasco, A., 1996. Where are we in the political economy of reform? J. Pol. Reform 1, 187–238.
Wacziarg, R., Welch, K.H., 2008. Trade liberalization and growth: new evidence. World Bank Econ. Rev. 22 (2), 187–231.
World Bank, 2019. World Databank: World Development Indicators. Retrieved from. http://databank.worldbank.org/data/reports.aspx?source=2&Topic=3#.