Abstract

Welfare-state measures often tend to persist when reforms seem indicated. This paper proposes an information-based explanation for this persistence. Since the welfare state ensures against consumption fluctuations and hence the importance of self-insurance, it also reduces incentives to inform oneself about economic fundamentals. When welfare-state arrangements are pronounced, voters remain rationally inattentive for longer, realize changes in economic fundamentals later, and reforms have considerable delays.

Keywords: welfare state, voting, imperfect information
JEL classification: D72, H55, D83

1 Introduction

It is a frequently expressed view that the political process features an asymmetry between the speed of implementations of welfare-state arrangements and the speed of their removals. Reforms enhancing the size of the welfare state seem easily and quickly implemented while opposite reforms face stronger opposition. This paper offers an information-based explanation for such welfare-state persistence, a phenomenon that has been discussed frequently in the literature.

*Bredemeier@wiso.uni-koeln.de. Parts of this research were done while the author was at TU Dortmund University. Financial support from German Science Foundation (DFG) through SFB 823 is gratefully acknowledged.
For example, Lindbeck (2003, p. 20) observes "certain asymmetries between the politics of expansion and retreat" in welfare-state dynamics. Hassler et al. (2003) emphasize that, in the US, the UK, France, and Italy, the great depression led to increased public intervention which did not diminish after the economies had recovered. Brooks and Manza (2006) find similar patterns in welfare-state dynamics of several OECD countries at the end of the twentieth century and summarize that "welfare states within most developed democracies appear quite resilient in the face of profound shifts in their national settings" (p. 1). Welfare-state persistence is further observed by, e.g., Blanchard and Wolfers (2000), Hercowitz and Strawczynski (2004), Beetsma et al. (2009), Ballassone et al. (2010) and Brügemann (2012).

This paper offers an explanation for welfare-state persistence which is based on the effects of the welfare state on attentiveness. Since the welfare state lowers the importance of private decisions, it also reduces incentives to inform oneself about economic fundamentals such as life expectancy or invalidity risk. These fundamentals do not only influence private decisions on savings or insurance but also determine the optimal social choice regarding welfare-state arrangements.

The frequency with which people inform themselves about fundamentals depends on their private financial precaution while the incentives for private precaution depend on welfare-state arrangements. If the degree of social insurance is high, private choices depend only weakly on fundamentals and people inform themselves rarely. Consequently, if initial welfare-state arrangements are pronounced, it takes rather long until a change in fundamentals is noticed by a majority of society and translated into appropriate reforms. By contrast, the political delay is short when initial welfare-state arrangements are weak.

We formalize this argumentation in an OLG model with income uncertainty. In the model, agents are rationally inattentive (Reis 2006a; Reis 2006b) and balance information costs against the benefits of better choices which they can make when being informed.\(^1\) As voters, these inattentive agents decide over social-insurance reforms in direct democ-

\(^1\)The importance of information costs in democratic decision making has been stressed by Downs (1957). Empirical evidence for imperfect information of voters is documented by, e.g., Bartels (1996), Nauneat and Paldam (1997), and Duch et al. (2000). Empirical support for the inattentiveness hypothesis is provided by, e.g., Carroll (2003) and Mankiwi et al. (2003).
racy. Reforms can become necessary by exogenous changes in underlying income risk. But such changes are only noticed with delay due to agents’ rational inattentiveness. Delays tend to be longer when social insurance is strong as private decisions are, then, less important and information is of lower value. This leads to welfare-state persistence.

Strong prudence can potentially counteract the discussed effects as remaining inattentive is a form of risk taking which prudent agents like the less the poorer they are and the efficiency losses associated with the welfare state make agents on average poorer. The model analysis demonstrates that information costs unambiguously generate welfare-state persistence when agents are not too prudent. Further, when the welfare state is pronounced enough, it is persistent, independent of prudence. A numerical evaluation shows that welfare-state persistence robustly arises for standard parameter values. The numerical analysis suggests that inattentiveness leads to political delays of cut-backs in the welfare state which are about one third to two thirds longer than for reforms which expand the size of the welfare state.

Our model generates welfare-state persistence as a consequence of the effect of social insurance on agents’ attentiveness – thus their knowledge about economic fundamentals – and their support for reforms. The negative effect of social insurance on economic knowledge is documented empirically by Jappelli (2010). Boeri and Tabellini (2012) present evidence that economic knowledge increases the support for pension reforms.

Previous explanations have attributed welfare-state persistence to changes in preferences or to changes in the distributional conflict. In the former argument, pro-work attitudes erode when more agents live out of benefits (Lindbeck 1995; Lindbeck and Weibull 1999; Brooks and Manza 2006). The latter argument stresses that the welfare state produces its own support by enforcing distributional conflicts (Bénabou 2000; Hassler et al. 2003; Beetsma et al. 2009) or by generating a group of beneficiaries who would else not exist (Saint-Paul 2002; Brügemann 2012). In our model, preferences are stable and there is no distributional conflict since agents are ex-ante identical. The information effects presented in this paper are thus a complementary source of welfare-state persistence.

The remainder of the paper is organized as follows. Section 2 presents the set-up of the model. In Section 3, the model is solved for individual decisions of agents. Section 4 describes the aggregate dynamics of the model analytically and numerically. Section 5 concludes.
2 Model Set-up

Population. In each period, the economy is populated by two generations. Each generation consists of a continuum of members who live for two periods each. We understand these two periods as the two halves of an agent’s working life. While the agent knows current income with certainty, future income is uncertain. Each agent has exactly one direct ancestor and one direct offspring. We call the infinite stream of an agent’s ancestors and offspring her dynasty. The different generations of a dynasty are linked through the transmission of information: each agent receives all her dynasty’s information at the beginning of her life. This has the reasonable implication that a generation inherits its predecessor’s institutions if not engaging in costly reassessing them. We discuss this assumption in Section 4.3.

Preferences. Agents maximize expected lifetime utility
\[ E_{i,t}U_{i,t} = E_{i,t}[u(c_{i,t,t}) + \beta \cdot u(c_{i,t,t+1}) - \kappa \cdot a_{i,t}], \]
where the index \(i, t\) refers to the agent belonging to dynasty \(i\) and generation \(t\) (from now: agent \(i, t\)). \(c_{i,t,t}\) and \(c_{i,t,t+1}\) denote the agent’s consumption in periods \(t\) and \(t + 1\), respectively. Through the agent’s life, the first indices \(i, t\) identify her as an individual and do not change while the final index identifies time and moves on with it. The period utility function \(u\) fulfills \(u' > 0, u'' < 0, u''' \geq 0\). We consider general preferences and analyze which properties of \(u\) ensure welfare-state persistence. As we will see, the third derivative \(u'''\) plays an important role.

\(E_{i,t}\) is the statistical expectation operator conditional on information available to agent \(i, t\). \(a_{i,t}\) is an indicator variable describing the agent’s choice whether to be attentive to new information. \(\kappa\) is a fixed utility cost of obtaining, processing, and interpreting information. Reis (2006a) argues that, while some information may be observed at little cost, the costs of understanding it and determining the optimal response can be substantial. Agents rationally balance the cost \(\kappa\) against the benefit of better decisions which lead to a more favorable distribution of consumption. Agent are selfish and do not care about future generations when taking decisions. We discuss this assumption in Section 4.3.

Individual uncertainty. In the first period of their life, agents receive a deterministic net income \(y_{i,t,t} = \bar{y}\) which only depends on the social-
insurance choice $\tau$, described below. Second-period income is stochastic,

$$
y_{i,t,t+1} = \begin{cases} 
y^{\tau,h}, & \text{prob. } 1 - \pi_t \\
y^{\tau,l}, & \text{prob. } \pi_t, 
\end{cases}
$$

(2)

where $y^{\tau,l} < y^{\tau,h}$. Both $y^{\tau,l}$ and $y^{\tau,h}$ depend on the political choice $\tau$.

**Aggregate uncertainty.** The risk $\pi_t$ of receiving the lower income $y^{\tau,l}$ in the second period of life is the same for all members of a generation and follows an exogenous stochastic Markov process. Specifically, $\pi_t$ can take two values, $\pi^l$ and $\pi^h$, $\pi^l < \pi^h$. Thus, there are two aggregate states of the world, a good one with low income risk and a bad one where income risk is high. State changes occur with an exogenous probability $\lambda < \frac{1}{2}$ in any period.

There is no private insurance market.\(^2\) For agents, there are two ways to cope with income risk: private savings and social insurance.

**Private savings.** Agents can save at a (non state-contingent) gross interest rate of $R$ implying the following intertemporal budget constraint:

$$
c_{i,t,t} + \frac{1}{R} \cdot c_{i,t,t+1} = y_{i,t,t} + \frac{1}{R} \cdot y_{i,t,t+1}.
$$

(3)

**Social insurance.** Further, each generation has the choice between two variants of a social-insurance system denoted by $\tau_t \in \{0, 1\}$. $\tau_t = 1$ denotes the variant wherein the degree of social insurance is higher. This manifests in lower income differences between agents. However, this equalization comes at the cost of a lower overall income level. This captures disincentive effects or government inefficiency.

Formally, defining an income level $y^{\tau} = \left( y^{\tau,l} + y^{\tau,h} \right) / 2 + R \cdot \overline{y}$ and income differences $d^{\tau} = y^{\tau,h} - y^{\tau,l}$, the two policy options are characterized by

$$
d^1 < d^0 \quad \text{and} \quad y^1 \leq y^0.
$$

(4)

We denote the income level as a function of income differences, $y^{\tau} = y(d^{\tau})$, with derivative $\theta \geq 0$. No variant is dominant, thus $\theta < \frac{1}{2}$ such that poor agents are better off with stronger social insurance.

---

\(^2\)In the Appendix, we consider a model version with incomplete private insurance and demonstrate that the paper’s results are robust to this model variation.
Politics. Political choices are decided in direct democracy. Each generation $t$ chooses its social-insurance system $\tau_t \in \{0, 1\}$ in a direct vote over the two options. All agents participate, vote truthfully, and support their individual expected-utility maximizing $\tau_t$.\(^3\)

Timing of events and beliefs. The timing of events is illustrated in Figure 1. Prior to period $t$, income risk $\pi_t$ for generation $t$ is determined by nature with $\text{prob}[\pi_t = \pi_{t-1}] = 1 - \lambda > \frac{1}{2}$. In period $t$, an agent of generation $t$ first receives information from the member of her dynasty in generation $t - 1$. Second, she takes part in the vote on $\tau_t$. Third, the agent decides whether to obtain complete information on income risk $\pi_t$. Fourth, the agent decides how much to consume and save, respectively. In the second period of her life, the agent first bequeaths information to a member of generation $t + 1$. After this, she observes her net income $y_{i,t,t+1}$, and consumes it together with savings and interests.

Due to potential updates of beliefs, one has to distinguish between prior and posterior beliefs. Beliefs are described by the dynasty’s perceived probability of the bad state and are denoted by $p_{i,t} = \text{prob}_{i,t}[\pi_t = \pi^b]$ and the analogously defined $\tilde{p}_{i,t}$ which refer to the point in time before ($p_{i,t}$) and after ($\tilde{p}_{i,t}$) the updating decision, respectively.

\(^3\)With two generations alive in each period, there are two potentially different social-insurance systems operating at the same time, one for each generation. When Germany increased the age of retirement in 2010, the reform fully applied only to rather young workers while older workers continued to fall under the old regulation.
3 Individual Decisions

Since income in the second period of life realizes after all decisions are taken, agents with identical beliefs make identical decisions. Starting with the same initial beliefs of all dynasties, this implies that beliefs remain homogenous as updating is an active choice and, within a period, every updater receives the same (perfect) information. Thus, all agents within one generation take the same decisions and we can drop the index \( i \).\(^4\) We solve an agent’s problem by backward induction.

3.1 Consumption and savings decision

When deciding on individual consumption, an agent knows the degree of social insurance, the relevant belief is the posterior belief \( \tilde{p}_t \), and updating costs are sunk. Together with (1) to (3), this gives the problem

\[
\max_{c_{t,t}} \tilde{U}_c (c_{t,t}, \tilde{p}_t) = u(c_{t,t}) + \tilde{\pi}_t^e \cdot \beta \cdot u \left( y(d^\tau) - \frac{d^\tau}{2} - R \cdot c_{t,t} \right) + (1 - \tilde{\pi}_t^e) \cdot \beta \cdot u \left( y(d^\tau) + \frac{d^\tau}{2} - R \cdot c_{t,t} \right) - \kappa \cdot a_t,
\]

where \( \tilde{\pi}_t^e = \pi^l + \tilde{p}_t \cdot (\pi^h - \pi^l) \) is the perceived probability of the low income realization. In the solution to this problem, it holds that

\[
\tilde{U}_c^\tau = 0 \quad \text{and} \quad \tilde{U}_{cc}^\tau < 0,
\]

where \( \tilde{U}_c^\tau \) and \( \tilde{U}_{cc}^\tau \) denote the first and second derivative of \( \tilde{U}^\tau \) in \( c_{t,t} \) direction. Further derivatives are denoted analogously. Condition (6) implicitly defines a consumption function \( c^\tau (\tilde{p}_t) \) with derivative

\[
c_p^\tau = -\tilde{U}_{cp}/\tilde{U}_{cc}^\tau < 0.
\]

When \( \tilde{p}_t \) is high, the agent perceives to be rather poor and, consequently, consumes rather little. Further, it is important that social insurance impacts on how strongly consumption depends on the belief \( \tilde{p}_t \). Generally, shrinking differences between income realizations make their probabilities less important (\( \tilde{U}_{c\pi d}^\tau < 0 \)). However, the agent’s risk aversion may

\(^4\)Starting with initially heterogeneous beliefs, the first policy reform is a free signal and subsequently synchronizes updates and equalizes decisions. This implies that rational abstentions are not relevant as all agents have the same information.
be increasing, constant, or decreasing in \( d \), depending on \( u'' \). Nevertheless, we can state a local result about the effect of income differences on the steepness of the consumption function:

**Proposition 1** The absolute steepness of the consumption function \( c^\tau(p_t) \) increases in income differences \( d \) if (but not only if)

1. agents are not too prudent, i.e. \( u''' \) is not too big, or
2. income differences \( d^\tau \) are not too pronounced.

**Proof.** Take the derivative of (7) in \( d^\tau \) direction to obtain \( c_{pd}^\tau = -(\bar{U}_{cpd}\bar{U}_{cc} - \bar{U}_{cp}\bar{U}_{cc})^2 \). Divide this expression by \( c_p^\tau \) as in (7) to obtain

\[
\frac{c_{pd}^\tau}{c_p^\tau} = \frac{\bar{U}_{cpd}}{\bar{U}_{cp}} - \frac{\bar{U}_{ccd}}{\bar{U}_{cc}}. \tag{8}
\]

\( \bar{U}_{cpd}/\bar{U}_{cp} \) is unambiguously positive as \( \bar{U}_{cpd} < 0 \) and \( \bar{U}_{cp} < 0 \). Now consider the limits:

1. \( u''' \to 0; \lim_{u'' \to 0} \bar{U}_{cc}^\tau = 0 \Rightarrow \lim_{u'' \to 0} c_{pd}^\tau/c_p^\tau = \bar{U}_{cpd}/\bar{U}_{cp} > 0. \)

2. \( d^\tau \to 0; \lim_{d^\tau \to 0} \bar{U}_{cp}^\tau = 0 \) and \( \lim_{d^\tau \to 0} \bar{U}_{cc}^\tau < 0 \Rightarrow \lim_{d^\tau \to 0} c_p^\tau = 0. \) With \( c_p^\tau < 0 \) for \( d^\tau > 0 \), it follows that, close to \( d^\tau = 0, c_{pd}^\tau/c_p^\tau > 0. \)

The proposition follows by continuity. ■

Proposition 1 contains the main reason for why social insurance reduces the value of information. When income differences are low, choices depends less strongly on the perceived state of the world, provided that the mentioned conditions are met. This implies that information has a lower impact on behavior which tends to reduce its value.

Throughout the analysis, our results depend on prudence, i.e. \( u''' \), and on the degree of social insurance, i.e. \( d^\tau \) as in Proposition 1. The model’s key mechanism relies on the fact that, when income differences are small, probabilities of income realizations matter less and the value of information is lower. However, prudence can lead to a counteracting effect. Prudent agents tend to be more risk averse (\( u'' \) higher in absolute terms) when being poorer, i.e., when social insurance is strong. However, the first effect always dominates when the social-insurance system equalizes incomes sufficiently such there is sufficiently little effective risk. In Section 4.2, we evaluate the model numerically and find that the discussed conditions are robustly met under standard calibrations.
Expected indirect utility. At previous decisions, the agent takes into account optimal savings behavior. Expected indirect lifetime utility net of updating costs is determined by the value function to (5),

\[ \tilde{V}^\tau (\tilde{p}_t) := \tilde{U}^\tau (c^\tau (\tilde{p}_t), \tilde{p}_t). \]

For the voting and updating decisions, intersections of the value functions \( \tilde{V}^0 \) and \( \tilde{V}^1 \) as well as their curvatures are relevant. Especially, gains from updating will depend on the convexity of the value function. The following derivatives of the problem (5) are key for the results:

\[
\begin{align*}
\tilde{U}^\tau_{\text{cc}} &= u''(c_{t,t}) + R^2 \cdot \beta \cdot \left( \pi_t^e \cdot u''(y(d^\tau) - \frac{d^\tau}{2} - R \cdot c_{t,t}) \right) \\
&\quad + (1 - \pi_t^e) \cdot u''\left(y(d^\tau) + \frac{d^\tau}{2} - R \cdot c_{t,t}\right) < 0, \\
\tilde{U}^\tau_{\text{cp}} &= R \cdot (\pi^h - \pi^l) \cdot \beta \cdot \left( u'(y(d^\tau) + \frac{d^\tau}{2} - R \cdot c_{t,t}) \right) \\
&\quad - u'\left(y(d^\tau) - \frac{d^\tau}{2} - R \cdot c_{t,t}\right) \leq 0, \\
\tilde{U}^\tau_{\text{ccd}} &= u'''(c_{t,t}) - R^3 \cdot \beta \cdot \left( \pi_t^e \cdot u'''(y(d^\tau) - \frac{d^\tau}{2} - R \cdot c_{t,t}) \right) \\
&\quad + (1 - \pi_t^e) \cdot u'''\left(y(d^\tau) + \frac{d^\tau}{2} - R \cdot c_{t,t}\right), \\
\tilde{U}^\tau_{\text{cp}} &= -R^3 \cdot \beta \cdot \left( \pi_t^e \cdot u'''(y(d^\tau) - \frac{d^\tau}{2} - R \cdot c_{t,t}) \right) \\
&\quad + (1 - \pi_t^e) \cdot u'''\left(y(d^\tau) + \frac{d^\tau}{2} - R \cdot c_{t,t}\right) \leq 0,
\end{align*}
\]

and \( \tilde{U}^\tau_{pp} = 0 \). Now consider the following two limiting cases:

1. No prudence: \( u''' \to 0 \): \( \lim_{u''' \to 0} \tilde{U}^\tau_{\text{ccc}} = 0 \) and \( \lim_{u''' \to 0} \tilde{U}^\tau_{\text{ccd}} = 0 \).

By contrast, \( \lim_{u''' \to 0} \tilde{U}^\tau_{\text{cc}} < 0 \) and \( \lim_{u''' \to 0} \tilde{U}^\tau_{\text{cp}} < 0 \).

2. Complete insurance: \( d^\tau \to 0 \): \( \lim_{d^\tau \to 0} \tilde{U}^\tau_{\text{cp}} = 0 \). By contrast, \( \lim_{d^\tau \to 0} \tilde{U}^\tau_{\text{cc}} < 0 \).
We now summarize important results on the value functions $\tilde{V}^\tau$:

**Proposition 2** If the value functions $\tilde{V}^0$ and $\tilde{V}^1$ cross, then agents prefer strong social insurance when they perceive income risk to be high.

**Proof.** Both value functions are decreasing in $\hat{p}_t$, $\tilde{V}^\tau_p = (\pi^h - \pi^l) \cdot \beta \cdot (u(y(d^\tau)) - \frac{d^\tau}{2} - R_{c,t}) - u(y(d^\tau)) + \frac{d^\tau}{2} - R_{c,t}) + \tilde{U}_c^\tau \cdot c^*_p < 0$ as $\tilde{U}_c^\tau = 0$, which implies $\tilde{V}^\tau (1) < \tilde{V}^\tau (0)$. Further, $\tilde{V}^\tau_{pd} = (\pi^h - \pi^l) \cdot \beta \cdot (\theta - \frac{1}{2}) \cdot u'(y(d^\tau)) - \frac{d^\tau}{2} - R_{c,t}) - (\theta + \frac{1}{2}) \cdot u'(y(d^\tau)) + \frac{d^\tau}{2} - R_{c,t})) > 0$ such that $\tilde{V}^1_p < \tilde{V}^0_p$ and $\tilde{V}^0$ crosses $\tilde{V}^1$ from above, $\tilde{V}^0 (1) < \tilde{V}^1 (1) < \tilde{V}^1 (0) < \tilde{V}^0 (0)$. ■

The intuition behind Proposition 2 builds on the fact that agents’ support for (strong) social insurance increases in the probability of becoming a beneficiary of the system.

**Proposition 3** The value function of the savings problem (5) is convex.

**Proof.** Generally, $\tilde{V}^\tau_{pp} = c^*_pp \cdot \tilde{U}_c^\tau + (c^*_p)^2 \cdot \tilde{U}_{cc}^\tau + 2 \cdot c^*_p \cdot \tilde{U}_{cp}^\tau + \tilde{U}_{pp}^\tau$. With $\tilde{U}_{pp}^\tau = 0$, applying (6) and (7) gives

$$\tilde{V}^\tau_{pp} = -(\tilde{U}_{cp}^\tau)^2 / \tilde{U}_{cc}^\tau > 0.$$ (9)

which is positive because $\tilde{U}_{cc}^\tau < 0$. ■

As we will see below, the convexity of the value function implies that there are potential gains from updating information. For the analysis of welfare-state dynamics, it is important to compare gains from updating in the two social-insurance systems. While the effect of the income level $y$ is ambiguous, the effect of income differences $d$ is unambiguously positive. Higher income differences lead to a higher curvature of the value function. Proposition 4 contains a local result on the total effect:

**Proposition 4** The curvature of the value function of the savings problem (5) increases in income differences $d$ if (but not only if)

1. agents are not too prudent, i.e. $u''$ is not too big, or
2. income differences $d^\tau$ are not too pronounced.
Proof. Differentiate (9) in $d^\tau$ direction to obtain

$$
\bar{V}_{ppd}^\tau = -\frac{2 \cdot \bar{U}_{cp}^\tau \cdot \bar{U}_{cc}^\tau \cdot \bar{U}_{cpd}^\tau + (\bar{U}_{cp}^\tau)^2 \cdot \bar{U}_{ccd}^\tau}{(\bar{U}_{cc}^\tau)^2} > 0.
$$

(10)

The functional form of (5) implies $\bar{U}_{cp}^\tau < 0$, $\bar{U}_{cc}^\tau < 0$, $\bar{U}_{cpd}^\tau < 0$. Now consider the limits:

1. $u'' \to 0$: $\lim_{u'' \to 0} \bar{U}_{ccd}^\tau = 0 \Rightarrow \bar{V}_{ppd}^\tau \to -2 \cdot \bar{U}_{cp}^\tau \cdot \bar{U}_{cc}^\tau \cdot \bar{U}_{cpd}^\tau / (\bar{U}_{cc}^\tau)^2 > 0$.

2. $d^\tau \to 0$: $\lim_{d^\tau \to 0} \bar{U}_{cp}^\tau = 0$ and $\lim_{d^\tau \to 0} \bar{U}_{cc}^\tau < 0 \Rightarrow \lim_{d^\tau \to 0} \bar{V}_{pp}^\tau = 0$.

With $\bar{V}_{pp}^\tau > 0$ for $d^\tau > 0$, it follows that, close to $d^\tau = 0$, $\bar{V}_{ppd}^\tau > 0$.

The proposition follows by continuity. ■

Proposition 4 implies that, when the discussed conditions are fulfilled, gains from updating are higher when social insurance is weak. The intuition for this effect lies in the fact that consumption decisions then depend more heavily on the belief $\bar{p}_t$, as stated in Proposition 1, such that information has a stronger impact on behavior when social insurance is weak. The conditions in Proposition 4 resemble those in Proposition 1 but the specific thresholds can be potentially be different.

3.2 Updating decision

An agent updates her information if her expected utility is higher when doing so. She takes this decision with the prior belief $p_t$ about income risk and knowledge about social insurance $\tau_t$. When the agent decides not to update, she will choose consumption according to her prior belief and expects a lifetime utility of $\bar{V}^\tau (p_t)$ since $\bar{p}_t = p_t$ and $a_t = 0$.

When the agent decides to be attentive, she will know $\pi_t$ after updating, i.e. $\bar{p}_t = 0$ or $\bar{p}_t = 1$ but she has to bear the information cost $\kappa$. Lifetime utility is then either $\bar{V}^\tau (0) - \kappa$ or $\bar{V}^\tau (1) - \kappa$. Prior to updating, she expects to observe $\pi_t = \pi^h$ with probability $p_t$ and $\pi_t = \pi^l$ with probability $1 - p_t$. The agent will decide to update if

$$
(1 - p_t) \cdot \bar{V}^\tau (0) + p_t \cdot \bar{V}^\tau (1) - \bar{V}^\tau (p_t) > \kappa.
$$

(11)

Since $\bar{V}^\tau$ is convex in $\bar{p}_t$, there are potential gains from updating, see Figure 2. If $\kappa$ is sufficiently low, there is a unique updating range between
When $p_t \in (p^\tau, \bar{p}^\tau)$, the agent decides to obtain perfect information about income risk. By contrast, when $p_t \notin (p^\tau, \bar{p}^\tau)$, the agent remains inattentive. We now compare the updating ranges in the two social-insurance systems.

**Proposition 5** If the conditions in Proposition 4 are met, the updating range is larger with weak social insurance.

**Proof.** The difference function $\tilde{V}^0 - \tilde{V}^1$ is convex in $p_t$ since $\tilde{V}^0_{pp} > \tilde{V}^1_{pp}$, by Proposition 4. Thus, the left hand side of (11) is always larger for $\tau = 0$ than for $\tau = 1$. Therefore, whenever $p_t$ fulfills condition (11) for $\tau = 1$, it also does so for $\tau = 0$ but not vice versa. ■

Proposition 5 contains the paper’s key mechanism: the stronger agents are socially insured, the less attentive they are, provided that the discussed conditions are fulfilled. This result reflects that, when social insurance is strong ($\tau_t = 1$), decisions depend less strongly on the state of the economy, see Proposition 1, and, thus, better information has a lower influence on lifetime utility.

The inattentiveness ranges $(0, p^\tau)$ and $(\bar{p}^\tau, 1)$ do not need to be symmetric, i.e. it does not need to hold that $1 - \bar{p}^\tau = p^\tau$. This potential asymmetry implies that agents might, for a given level of uncertainty,
take different updating decisions depending on which state they perceive as relatively more likely. As one can already anticipate, we will be particularly interested in comparing \(1 - \bar{p}^1\) (the length of the inattentiveness range that is relevant when agents prefer strong social insurance) and \(p^0\) (its counterpart when agents prefer weak social insurance). If agents are rather attentive when being pessimistic (i.e., when \(p_t\) is high), this counteracts the effects of social insurance on attentiveness stated in Proposition 5. In Proposition 6, we summarize sufficient conditions for the latter effects to be dominant.

**Proposition 6** The relevant inattentiveness range in the strong social-insurance system is larger than its counterpart in the weak social-insurance system \((1 - \bar{p}^1 > p^0)\) if (but not only if)

1. agents are not too prudent, i.e. \(u''\) is not too big, or
2. income differences with strong social insurance, \(d^1\) are not too pronounced.

**Proof.** The length of the relevant inattentiveness ranges are \(p^0\) and \(1 - \bar{p}^1\). Consider condition (11), which implies \(1 - \bar{p}^\tau \gtrless \rho_\tau \iff \vec{V}_{\rho p p} \gtrless 0\), together with Proposition 5, to see that, only for strongly decreasing curvature of \(\vec{V}_\tau\), it is possible that \(1 - \bar{p}^1 < p^0\). So, \(1 - \bar{p}^1 > p^0\) if \(V_{\rho pp}^1\) is not too strongly increasing in \(p_t\). To see that conditions 1 and 2 in the proposition are sufficient, consider

\[
\vec{V}_{\rho pp}^1 = \frac{U_{\rho p}}{U_{\rho c}} \cdot (3U_{\rho c}^1 - U_{\rho p}^1 \cdot U_{\rho c c}^1 / U_{\rho c c}^1) \text{ with } U_{\rho c c}^1 < 0, U_{\rho p}^1 < 0, \text{ and } U_{\rho c c p} < 0, \text{ and the limits}
\]

1. \(u'' \to 0 : U_{\rho c c}^1 \to 0 \implies \vec{V}_{\rho pp}^1 < 0,\)
2. \(d^1 \to 0 : U_{\rho c}^1 \to 0 \implies \vec{V}_{\rho pp}^1 \to 0,\)

The Proposition follows by continuity. ■

Proposition 6 contains an important result. When one of the conditions in the Proposition is fulfilled, agents will, with strong social insurance, remain inattentive for a larger set of prior beliefs than they would with weak social insurance.

The intuition behind the two conditions is the following. Remaining inattentive is a form of risk taking. Prudence tends to lead to stronger
risk aversion when agents are poor. This tends to make attentiveness more attractive in the strong social-insurance system (wherein agents are poorer). If prudence is not too strong, attentiveness is lower with strong social insurance, as stated in Condition 1. Condition 2 implies that information is of relatively small value in the strong social-insurance regime because the two exogenous states of the world are relatively similar from the viewpoint of agents. The numerical evaluations in Section 4.2 indicate that the conditions in Proposition 6 are robustly fulfilled under standard calibrations of the main parameters.

**Expected indirect utility.** For the voting decision, we determine the expected indirect utility which arises from optimal savings and optimal updating. We denote this function as $V^\tau (p_t) := E_t [U_t | \tau_t = \tau]$. Taking into account the optimal updating choice, see Figure 2, this function is

$$V^\tau (p_t) = \begin{cases} \tilde{V}^\tau (p_t), & p_t \notin (p^\tau, \bar{p}^\tau) \\ (1 - p_t) \cdot \tilde{V}^\tau (0) + p_t \cdot \tilde{V}^\tau (1) - \kappa, & p_t \in (p^\tau, \bar{p}^\tau). \end{cases} \quad (12)$$

$V^0$ and $V^1$ have a unique intersection $p^*$ on $(0, 1)$ since they are weakly convex. The notion of a political delay implies that a reform is actually caused by a change in fundamentals. This is ensured when the expected indirect utility functions $V^0$ and $V^1$ intersect in the updating ranges. Then, a policy reform only takes place when agents actually observe a different state of the world than at the last update.

### 3.3 Voting decision

An agent decides how to vote such as to maximize expected indirect utility, $V^\tau (p_t)$. The agent votes for stronger social insurance whenever $V^1 (p_t) > V^0 (p_t)$ and votes against it otherwise. Revisiting the expected indirect utility functions $V^\tau$, it follows that there is a unique prior for which the agent is indifferent between the two political regimes, $p^*$.

### 3.4 Belief formation

By the transmission of information through generations, an agent knows the time of her dynasty’s last update and what was observed at that time. In period $t$, the probability that income risk is still the same as at the last update in $t - j$, $j > 0$, equals the probability that the number of

---

5This is not necessary but possible as demonstrated in Section 4.2.
state changes between $t - j$ and $t$ was even, given by

$$prob[\pi_t = \pi_{t-j}] = \begin{cases} j! (1 - \lambda)^j \sum_{n=0}^{j/2} \frac{(\lambda^2)^n((1-\lambda)^{-2})^n}{(j-2n)![(2n)!]^j}, & j \text{ even} \\ j! (1 - \lambda)^j \sum_{n=0}^{(j-1)/2} \frac{(\lambda^2)^n((1-\lambda)^{-2})^n}{(j-2n)![(2n)!]^j}, & j \text{ odd.} \end{cases}$$ \tag{13}

Since $\lambda < \frac{1}{2}$, beliefs converge monotonically (from above or below, depending on $\pi_{t-j}$) towards $1/2$. The speed of convergence is independent of $\pi_{t-j}$.

4 Aggregate Dynamics

4.1 Duration of inattentiveness and political delay

The duration of inattentiveness $I^r$ is the time between two updates and depends on the current social-insurance system, $\tau$. In the system with weak social insurance, the posterior belief after the last update was $\tilde{p}_{t-j} = 0$ and beliefs evolve according to $p_t = 1 - prob[\pi_t = \pi_{t-j}]$. The duration of inattentiveness $I^0$ is the time between the last update and the first period in which $p_t > p^0_0$,

$$I^0 = \min \{ j \in \mathbb{N} \mid prob[\pi_t = \pi_{t-j}] < 1 - p^0 \} . \tag{14}$$

With strong social insurance, the duration of inattentiveness is

$$I^1 = \min \{ j \in \mathbb{N} \mid prob[\pi_t = \pi_{t-j}] < p^{1} \} . \tag{15}$$

**Proposition 7** If the conditions in Proposition 6 are met, the duration of inattentiveness is never longer with weak social insurance than it is with strong social insurance.

**Proof.** By Proposition 6, $\tilde{p}^1 < 1 - p^0$. As $prob[\pi_t = \pi_{t-j}]$ is a decreasing function of $j$, it falls below $1 - p^0$ not after it falls below $\tilde{p}^1$ such that $I^0 \leq I^1$.

The political delay is the time between a change in income risk and the appropriate policy reform. The maximum delay is the duration of inattentiveness $I^r$ and occurs when the change in income risk happens right after an update. The minimum delay is one period and occurs when income risk changes right before an update. Since state changes occur with equal probability each period, the expected political delay is

$$D^r = \frac{1}{2} \cdot (I^r + 1).$$
Proposition 8  If the conditions in Proposition 6 are met, the expected political delay is never longer with weak social insurance than it is with strong social insurance.

Proof. \( D^* \) is increasing in \( I^* \) and \( I^0 \leq I^1 \) by Proposition 7 such that \( D^0 \leq D^1 \). □

Proposition 8 is the main result of the paper. When the discussed conditions are met, the model generates welfare-state persistence as a result of rational inattentiveness of voters: the expected delay of a reduction in social insurance is longer than the expected delay of an increase in social insurance. Thus, reforms enhancing the size of the welfare state are relatively quickly implemented while opposite reforms take longer.

This result relies on the disincentive effects of social insurance. With strong social insurance, if the discussed conditions are met, agents’ behavior depends less strongly on beliefs and they can thus gain less from information. As a consequence, agents remain inattentive for longer periods of time. Changes in income risk are then, in expectations, realized later and reforms have longer delays. In Section 4.2, we evaluate numerically in which situations the model does generate welfare-state persistence. We find that it does under standard parameter values. Additionally, in order to generate \( D^1 \leq D^0 \), strong deviations from standard parameters are needed.

4.2 A numerical evaluation

Our baseline calibration is as follows. We use common CRRA utility, \( u = \frac{c^{1-\sigma}}{1-\sigma} \). As a starting point, we choose a standard value of \( \sigma = 1.75 \), implying a relative risk aversion of 1.75 and a relative prudence of 2.75. Further, we choose \( \beta = 0.67 \) and \( R = 1.5 \). Interpreting a period as 20 years, this gives annual discount and interest rates of about 2% each. Further, we calibrate possible income realizations and their probabilities to match specific moments of the income distributions in the United States and the Euro Area, respectively. The US is our prototype of a country with weak social insurance while the Euro Area serves as an example for a strong social-insurance system. We calibrate \( \pi^l = 0.267 \) and \( \pi^h = 0.357 \) to match economic dependency ratios. We set \( y^0 = 4.84 \) and \( y^1 = 3.75 \) to match the ratio of per-capita GDPs in 2012. In order to match the empirical 80/20 percentile ratios of about 3 and 2,
respectively, we set \( d^0 = 5 \) and \( d^1 = 2.4 \). For the information cost, we use the two values \( \kappa = 0.0005 \) and \( \kappa = 0.0001 \) which correspond to utility losses associated with reductions of 0.04% and 0.01% of income, respectively. With the interpretation of a period as 20 years, this is in line with Reis (2006a) where updating costs range from 0.2% to 0.8% of annual income. Regarding the probability of state changes, we report results for \( \lambda = 0.01 \) and for \( \lambda = 0.05 \).

Table 1 summarizes the results under the baseline calibration. In all considered combinations of \( \kappa \) and \( \lambda \) (shown in lines 1 and 2), the updating range with strong social insurance \((\bar{p}^0, \bar{p}^0)\) is a strict subset of its counterpart with weak social insurance \((\bar{p}^1, \bar{p}^1)\), so the conditions in Proposition 5 are met and agents are less attentive with stronger social insurance, see lines 3-6 of Table 1. Most importantly, the length of the relevant inattentiveness range when social insurance is strong, \( 1 - \bar{p}^1 \), exceeds its counterpart with weak social insurance, \( \bar{p}^0 \). Thus, the conditions in Proposition 6 are met and the model generates welfare-state persistence. Further, value functions intersect within both updating ranges, see line 7 of Table 1, which justifies the notion of political delays.

How persistent the welfare state is, depends on \( \kappa \) and \( \lambda \). Generally, the higher information costs and the more unlikely state changes are, the more pronounced is the persistence of either system. However, this does

\[\text{Table 1: Results under baseline calibration.}\]

| (1) | information costs \( \kappa \) | 0.0005 | 0.0002 | 0.0005 | 0.0002 |
| (2) | state change prob. \( \lambda \) | 0.0500 | 0.0500 | 0.0100 | 0.0100 |
| (3) | updating range w. \( \bar{p}^0 \) | 0.1115 | 0.0356 | 0.1115 | 0.0356 |
| (4) | strong social ins. \( \bar{p}^0 \) | 0.8737 | 0.9593 | 0.8737 | 0.9593 |
| (5) | updating range w. \( \bar{p}^1 \) | 0.1560 | 0.0476 | 0.1560 | 0.0476 |
| (6) | weak social ins. \( \bar{p}^1 \) | 0.8249 | 0.9458 | 0.8249 | 0.9458 |
| (7) | value fct. intersect \( p^* \) | 0.3745 | 0.3745 | 0.3745 | 0.3745 |
| (8) | duration of \( I^0 \) | 3 | 1 | 13 | 4 |
| (9) | inattentiveness \( I^1 \) | 5 | 2 | 22 | 6 |
| (10) | rel. welf.-st. pers. \( D^1/D^0-1 \) | 33% | 50% | 65% | 40% |

\[\text{6We take dependency ratios from the Labor Force Projections of the BLS and the Ageing Report of the European commission, GDP per capita from the IMF (US) and Eurostat (EA), and 80/20 ratios from the key figures of the Luxembourg Income Study.}\]
not translate into unambiguous effects on relative persistence. Reforms reducing the size of the welfare state are predicted to have about one third to two thirds longer delays than opposite reforms, see the final line of Table 1.

**Robustness of welfare-state persistence.** We now evaluate how robust welfare-state persistence is. In particular, we vary the key parameters $\sigma$ (under CRRA, the coefficient of relative prudence is $\sigma + 1$) and $d^1$ mentioned in Proposition 6. All other parameters are held at their baseline values as described above. Figure 3 shows the results.

The concave, downward-sloping, and dotted line marks parameter combinations that imply $p^0 = 1 - \overline{p}^1$. Note that, even for $\sigma \to 0$, there is positive prudence, $\sigma + 1$, here such that the threshold for $p^0 < 1 - \overline{p}^1$ is finite. There is a range of parameter combinations around this line where both systems persist equally long because the duration of inattentiveness is an integer. These combinations lie between the two solid lines. Constellations below the lower solid line lead to welfare-
state persistence. For constellations above the upper solid line, the weak social-insurance system is more persistent. The discrete steps in both lines are due to the discreteness in $I^r$, see equations (14) and (15). The vertical dotted line marks the level of $d^1$ above which the strong social-insurance system is dominated as $y^0 - d^0/2 > y^1 - d^1/2$.

The results shown in Figure 3 demonstrate that the generated welfare-state persistence is very robust around the baseline calibration. Given the baseline income differences, $d^1 = 2.4$, the model generates welfare-state persistence for every level of prudence. Holding prudence at $\sigma + 1 = 2.75$, we can generate equal persistences of both system only for $d^1 > 3$ and a stronger persistence of the system with weak social insurance only for $d^1 > 4$. Above both thresholds, the strong social-insurance systems is dominated such that the strong social-insurance system would never be chosen in the first place. With levels of income differences which fulfill non-dominance, the model does not generate welfare-state persistence only in very small parameter ranges with rather high levels of income differences in the strong social insurance system, $d^1 > 2.7$. But even in this range, we do not generate a weaker persistence of the strong social-insurance system but both systems persist equally long.

4.3 Discussion

Here, we discuss two simplifying model assumptions.

Public/private information. In the model, information is a private good across dynasties but, within a dynasty, it inevitably passes from one generation to another. The latter assumption does not necessarily imply that information is a public good within dynasties. It might well be that receiving this information is associated with costs but the younger generation cannot avoid it, e.g. because it is part of their breeding. If agents learn only a part of their parents’ information inevitably, beliefs would erode quicker, equivalently to an increase in $\lambda$. That information is modeled as private across dynasties reflects the importance of its processing. If information gathering is public or allows freeriding, processing and understanding information remains a private problem but total information costs might fall. Both, quicker movements of beliefs and lower information costs would make both systems less persistent, have an ambiguous effect on relative persistence, but would not alter the qualitative asymmetry between the two systems, see Table 1.
Internalization of welfare-state persistence. We have assumed that agents do not care about their offspring when taking decisions. If agents felt some altruism for future generations, they would take into account that the probability that future generations live in suboptimal social-insurance systems is higher when they choose $\tau_t = 1$ which makes this system relatively less attractive. This would, however, impact only on the relative levels of the value functions but not on their curvatures which are key for the discussed persistence results. The same argument holds for the case that agents live longer than for two periods.

5 Conclusion

This paper has offered an information-based explanation for welfare-state persistence. It is based on the incentive effects of the welfare state on attentiveness. Under reasonable assumptions on prudence, the welfare state reduces incentives to be attentive to developments in economic fundamentals. When the degree of social insurance is high, people remain inattentive for longer periods of time and it takes long until a change in fundamentals is noticed by a majority of society and until it is translated into appropriate political reforms.

References


Appendix

A model version with (incomplete) private insurance

In this web appendix, we consider a model version wherein agents have access to a private insurance market. The insurance market is incomplete in the sense that prices do not reflect the true event probabilities perfectly. Specifically, for every unit of insurance payment in period \( t \), \( a_{i,t} \), agent \( i, t \) receives \( \mu \cdot a_{i,t} \) units of the consumption good in period \( t+1 \) if and only if she receives the low net income \( y^{r,l} \). Consequently, the agent’s final-stage problem is described by

\[
\max_{c_{i,t}} \tilde{U}^{r} (c_{i,t}, \tilde{h}_{t}) = u(c_{i,t}) + \tilde{\pi}^{e}_{t} \cdot \beta \cdot u \left( \hat{y} (d^{r}) - \frac{d^{r}}{2} + \mu \tilde{y}^{r} - \mu c_{i,t} \right) \quad (16)
\]

\[
+ (1 - \tilde{\pi}^{e}_{t}) \cdot \beta \cdot u \left( \hat{y} (d^{r}) + \frac{d^{r}}{2} \right),
\]

where \( \hat{y} (d^{r}) = (y^{r,l} + y^{r,h}) / 2 \), instead of (5). As before we assume \( 0 < \partial \tilde{y}/\partial d^{r} = \theta < \frac{1}{2} \).

The problem has the following relevant derivatives:

\[
\tilde{U}^{r}_{cc} = u'' (c_{i,t}) + \mu^{2} \cdot \tilde{\pi}^{e}_{t} \cdot \beta \cdot u'' \left( \hat{y} (d^{r}) - \frac{d^{r}}{2} + \mu \tilde{y}^{r} - \mu c_{i,t} \right) < 0
\]

\[
\tilde{U}^{r}_{cp} = -\mu \cdot (\pi^{h} - \pi^{l}) \cdot \beta \cdot u' \left( \hat{y} (d^{r}) - \frac{d^{r}}{2} + \mu \tilde{y}^{r} - \mu c_{i,t} \right) < 0
\]

\[
\tilde{U}^{r}_{ccc} = u'''(c_{i,t}) - \mu^{3} \cdot \tilde{\pi}^{e}_{t} \cdot \beta \cdot u''' \left( \hat{y} (d^{r}) - \frac{d^{r}}{2} + \mu \tilde{y}^{r} - \mu c_{i,t} \right)
\]

\[
\tilde{U}^{r}_{ccd} = -\left( \frac{1}{2} - \theta \right) \cdot \mu^{2} \cdot \tilde{\pi}^{e}_{t} \cdot \beta \cdot u''' \left( \hat{y} (d^{r}) - \frac{d^{r}}{2} + \mu \tilde{y}^{r} - \mu c_{i,t} \right) \leq 0
\]

Note that these derivatives have the same signs as in the baseline model (expressed on page 19). Consequently, the derivatives have the same signs for the limiting case of linear-quadratic utility \( (u''' \to 0) \) but not for the limiting case of complete social insurance \( (d^{l} \to 0) \). In this model variant, the results expressed in Propositions 1, 4, and 6 consequently hold when agents are not too prudent as expressed in the first condition of the propositions. Further, Propositions 3, 4, 5, 7, and 8 can be shown in an equivalent way as on pages 20-21 of the main paper.