

# Assessing the Validity of Four-day Week Pilots

JRC Working Papers Series on  
Social Classes in the Digital Age  
2023/08

Hugo Cuello



This Working Paper is part of a Working paper series on Social Classes in the Digital Age by the Joint Research Centre (JRC) The JRC is the European Commission's science and knowledge service. It aims to provide evidence-based scientific support to the European policymaking process. The scientific output expressed does not imply a policy position of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use that might be made of this publication. The Working paper series on Social Classes in the Digital Age is intended to give visibility to early stage research to stimulate debate, incorporate feedback and engage into further developments of the research. This Working Paper is subject to the Commission Reuse Decision which allows authors to reuse the material without the need of an individual application.

### Contact information

Name: Hugo Cuello  
Contact: h.cuello@outlook.com

EU Science Hub  
<https://ec.europa.eu/jrc>

<https://ec.europa.eu/jrc/en/research/centre-advanced-studies/digclass>



JRC133008

Seville: European Commission, 2023

© European Union, 2023

Credits of the Image in the cover page: kras99, Adobe Stock image n. [175461355](https://www.adobe.com/stock/175461355)



The reuse policy of the European Commission is implemented by the Commission Decision 2011/833/EU of 12 December 2011 on the reuse of Commission documents (OJ L 330, 14.12.2011, p. 39). Except otherwise noted, the reuse of this document is authorised under the Creative Commons Attribution 4.0 International (CC BY 4.0) licence (<https://creativecommons.org/licenses/by/4.0/>). This means that reuse is allowed provided appropriate credit is given and any changes are indicated. For any use or reproduction of photos or other material that is not owned by the EU, permission must be sought directly from the copyright holders.

How to cite this report: Cuello, H., *Assessing the validity of four-day week pilots*, European Commission, Seville, 2023, JRC133008.

## Contents

1	The pilots testing the four-day workweek.....	4
2	Three types of four-day workweek pilots.....	5
2.1	Private Sector pilots.....	6
2.1.1	Multiple-company pilots.....	6
2.1.2	Individual-company pilots.....	7
2.2	Public Sector pilots.....	7
3	Case studies.....	8
3.1	4 Day Week Global (4DWG) pilots.....	9
3.2	Perpetual Guardian pilot.....	10
3.3	Iceland pilots.....	10
4	Overview of performance outcome indicators.....	11
4.1	Health and well-being.....	11
4.2	Job satisfaction and work environment.....	12
4.3	Work-life balance.....	12
4.4	Productivity and public service provision.....	12
4.5	Workers' performance.....	13
4.6	Absenteeism.....	13
4.7	Employment and employee retention.....	13
4.8	Time allocation and life satisfaction.....	14
4.9	Gender equality.....	14
4.10	Environmental footprint.....	14
5	Main limitations of current pilots.....	14
5.1	Addressing causality.....	17
5.2	Collecting accurate data.....	19
5.3	Transparency in reporting results.....	20
5.4	Explaining sample and intervention characteristics.....	22
6	Suggestions to improve upcoming four-day week pilots.....	24
6.1	Plan a robust evaluation method.....	24
6.2	Select a valid comparison group.....	24
6.3	Provide transparent access to data and information.....	25
6.4	Include all the relevant statistical information.....	25
6.5	Pre-commit to a set of outcomes and report them.....	25
6.6	Use larger and more representative samples.....	25
6.7	Find out more about non-compliers.....	26
6.8	Use more administrative data.....	26
6.9	Run longer pilots or increase the time of analysis.....	26
6.10	Consider the scalability of the intervention.....	26
7	References.....	28

# Assessing the Validity of Four-day Week Pilots

Hugo Cuello

## Abstract

The four-day workweek has gained global attention and popularity, with supporters arguing it can improve several conditions, such as productivity, job creation, and the environment. Small-scale pilot programmes are being carried out to better understand the concept, but it is crucial to ensure they are well-designed and implemented for reliable results. This working paper assesses the quantitative outcomes, scope, and methodological limitations of recent pilots and provides insight into future pilot designs for a more comprehensive understanding of the four-day workweek. A framework for categorising the four-day week pilots is presented, followed by an in-depth analysis of three case studies and their approaches to measuring various performance indicators and outcomes. The main section of the working paper provides a comprehensive examination of the quantitative limitations of the four-day week pilots. Drawing on these limitations, the paper offers ten specific recommendations for future researchers and policymakers to enhance the validity of four-day week pilots.

**Keywords:** four-day week; worktime reduction; flexible work arrangements; work-life balance; workers' performance;

**Author:** Hugo Cuello

Acknowledgements: I am very grateful for the constructive comments and valuable suggestions provided by Leire Salazar, Guillem Vidal and José Antonio Noguera. I also thank the JRC for organising a series of workshops on [Working Time Reduction and the Four-Day Week](#) under the Social Classes in the Digital Age (DIGCLASS) project.

Joint Research Centre reference number: JRC130830

Related publications and reports:

Real Utopias for a Social Europe: Working Time Reduction and the 4-Day Week, Summary Report, Salazar Valez, L., Vidal Lorda, G., Noguera, J.A. and Villani, D. editor(s), Publications Office of the European Union, Luxembourg, 2023.

<https://publications.jrc.ec.europa.eu/repository/handle/JRC132852>

Real Utopias for a Social Europe: Universal Benefits, Summary Report, Noguera, J.A., Salazar Valez, L., and Vidal Lorda, G. editor(s), Publications Office of the European Union, Luxembourg, 2023.

<https://publications.jrc.ec.europa.eu/repository/handle/JRC130425>

Moawad, J., Oesch, D. The Myth of the Middle Class Squeeze: Employment and Income by Class in Six Western Countries, 1980-2020, European Commission, Seville, 2023, JRC131515. [https://joint-research-centre.ec.europa.eu/system/files/2023-04/JRC131515\\_Myth\\_of\\_middle\\_class\\_squeeze.pdf](https://joint-research-centre.ec.europa.eu/system/files/2023-04/JRC131515_Myth_of_middle_class_squeeze.pdf)

Giangregorio, L., Villani, D., Income inequality, top shares of income and contemporary social classes in the 21st century, European Commission, Seville, 2023, JRC130501.

[https://joint-research-centre.ec.europa.eu/system/files/2023-01/JRC130501\\_income\\_inequality\\_top\\_shares\\_of\\_income\\_and\\_social\\_classes\\_in\\_the\\_21st\\_cen.pdf](https://joint-research-centre.ec.europa.eu/system/files/2023-01/JRC130501_income_inequality_top_shares_of_income_and_social_classes_in_the_21st_cen.pdf)

Bürgisser, R., Policy Responses to Technological Change in the Workplace, European Commission, Seville, 2023, JRC130830.

[https://joint-research-centre.ec.europa.eu/system/files/2023-01/JRC130830\\_Policy\\_responses\\_technological\\_change\\_workplace.pdf](https://joint-research-centre.ec.europa.eu/system/files/2023-01/JRC130830_Policy_responses_technological_change_workplace.pdf) :

## 1 The pilots testing the four-day workweek

The idea of a four-day workweek, or a reduction in working hours, has gained significant global attention and is increasingly popular. The core of the proposal is to reduce the standard full-time working hours to a four-day workweek, or a shorter number of hours per week, without affecting workers' pay. Supporters of the four-day workweek argue that it can lead to improved productivity, the creation or preservation of jobs, and a positive impact on the environment due to reduced consumption and production.

The impact of reducing working hours has been the subject of research, mainly through examining historical cases, simulations, and modelling. This research has been conducted by analysing the effects of mandatory reductions in working hours that different governments have introduced. For instance, the French government implemented a shorter workweek between 1998 and 2000 to combat high unemployment rates (Estevão & Sa, 2008). This reduced the standard number of weekly working hours from 39 to 35. Similarly, the Portuguese government reduced standard working hours from 44 to 40 per week by law in 1996 (Raposo & Van Ours, 2010). However, these reductions were implemented nationwide and were not preceded by small-scale testing. Consequently, measuring their causal impact has been challenging, as only a few studies have attempted to use quasi-random variations to evaluate the effects of these reductions.<sup>1</sup>

Research examining the effects of reducing working hours has shown that such reductions can positively impact well-being, employment, economics, and productivity outcomes (Barnes, 2019). However, some experts have noted drawbacks as well, as the implementation can be challenging and may not be suited for all types of businesses (Harrington, 2022). For instance, the French 35-hour workweek law increased dissatisfaction amongst workers. This was due to a lack of consideration for diversity in work organisations, causing employees to increase their workload to maintain high productivity (Prunier-Poulmaire & Gadbois, 2001).

The international focus on the four-day workweek is a relatively new development, and limited literature and real-world examples are available. There are simulations of the potential effects of a four-day workweek in countries such as the UK and Ireland (Gilmore, 2019). Some studies have also been conducted to evaluate the desirability, feasibility, and implementation of a four-day workweek (Calvert Jump, 2022). However, those studies don't provide evidence of the impact of the four-day week in the field.

In recent years, interest has resurged in the four-day workweek, with a new wave of small-scale pilot programmes aimed at gaining a better understanding of the concept in practice. Advocacy groups, businesses, and governments are pushing for these pilot programmes to be carried out to understand the effects of a shorter workweek better. Some of these pilots have received significant media attention and have inspired governments in various countries to consider similar plans, with companies showing interest in the USA, Canada and New Zealand, among others (Clark, 2022).

Pilots can provide practical knowledge on the impacts of a four-day working week, which can only be observed in the field. Although these small-scale tests cannot fully capture all the macroeconomic and market effects that would result from a nationwide implementation, they can still provide valuable insight into the potential impacts on productivity, well-being, workers' performance, and gender.

However, it is crucial to ensure that the pilots are correctly designed and implemented to obtain reliable results. Effective pilot design includes elements such as a well-designed recruitment plan, valid measures and data collection strategies, and a suitable evaluation method to answer the

---

<sup>1</sup> Two studies, Leipenteur (2019) and Sanchez (2016) provide quasi-experimental evidence on France and Portugal.

research questions. Monitoring implementation and preparing mitigation plans for compliance or attrition are also important considerations. If the pilots are not well-designed and implemented, the results may not be credible and will not contribute to the existing knowledge on the topic.

Overall, the four-day working week and other proposals for reduced working hours could provide innovative solutions to some of the main challenges facing developed societies. For instance, a recent systematic review (Voglino et al., 2022) explored the impact of reduced working hours on health outcomes. Nevertheless, there is still much to learn about the effects of the four-day week on different fronts, and existing pilots provide only indicative data on the potential positive outcomes. Several companies have reported challenges to implementing it effectively, which point to some significant heterogeneous effects that should be considered (Coleman, 2023). Some companies have even dropped the plans to test it as it would create “significant disruptions and threatening operations”, posing potential harm to some workers (Booth, 2019).

This paper focuses on examining the recent wave of pilots and assessing their outcomes, scope, and limitations to provide a more comprehensive understanding of the quantitative evaluations of the four-day working week. The additional qualitative information in the pilots is considered but not assessed. This paper does not aim to draw any conclusions about the effectiveness of the four-day working week or working time reductions, as it would not conduct a comprehensive meta-evaluation of all the four-day working week pilots. Instead, this paper aims to assess the key quantitative features, performance indicators and methodological problems of the most prominent four-day week pilots through the public information available and provide insights into where future pilots should focus on enhancing existing in-the-field knowledge.

The following section presents a typology for analysing the current wave of pilots. Section 3 will choose one case study for each pilot typology and describe their pilot designs. Section 4 will collect the principal performance indicators used in the case studies. Section 5 highlights the main limitations of the pilots, with a focus on the case studies but applicable to other pilots as well. Finally, Section 6 will draw to a close with recommendations for enhancing the validity of future four-day week pilots.

## **2 Three types of four-day workweek pilots**

Categorising the various four-day week pilots is important for several reasons. It enables policymakers and researchers to compare and contrast similar pilots, what has been studied and the main gaps in the existing evidence. Categorisation also allows for a more targeted analysis of the specific types of target groups. Finally, by identifying commonalities and differences between pilots, relevant stakeholders can identify areas for improvement and optimise the design of future interventions.

One potential categorisation method is based on the ownership and nature of the organisation introducing the intervention - whether it is a public or private organisation. It is essential to recognise that public and private sector dynamics differ, and lessons learned in the public sector may not be easily transferable to private sector employees. Within the private sector, there may also be a distinction between pilots that test the four-day week in one company versus those that include several companies. Table 1 outlines the main characteristics of these pilots.

Private-sector participation is essential for the sustainability of the four-day week policy proposal, as it requires a reduction in working hours without a corresponding reduction in pay. Thus, assessing whether businesses can successfully implement this change is crucial. Piloting can identify potential benefits or challenges that businesses may face in implementing the four-day week, such as changes to scheduling and staffing, and evaluate its impact on crucial business measures like productivity, company outcomes, and workers' well-being.

On the other hand, the public sector is a significant employer and service provider to the community, making it crucial to understand the potential impact of a four-day workweek on its workers and service delivery. Various pilot programmes have been conducted across the public sector, offering valuable insights into the potential impact of reduced working hours on public sector employees and organisations.

Table 1: Types of Four-day Workweek Pilots

	Private Sector Pilots		Public Sector Pilots
	Multiple-company pilots	Individual-company pilots	
<b>Strategy</b>	Coordinate volunteering businesses willing to introduce the four-day week	Volunteering business introducing the four-day week to employees	Select workplaces within a public administration to introduce the four-day week
<b>Unit of intervention</b>	Businesses	Employees	Workplaces
<b>Data</b>	Businesses and employee data	Employee data	Workplace and employee data
<b>Evaluation method</b>	Before-and-after comparisons	Before-and-after comparisons	Enrolled-and-unenrolled comparisons
<b>Promoter</b>	Advocacy organisations and governments	Individual firms	Governments
<b>Countries where it is tested</b>	UK, USA, Ireland, New Zealand, Portugal, Spain	UK, USA, Ireland, New Zealand, Portugal, Spain, Japan	Sweden, Finland, Iceland, Portugal

**2.1 Private Sector pilots**

*2.1.1 Multiple-company pilots*

Some advocacy organisations are coordinating pilot programmes with several companies simultaneously to increase the sample size and representativeness of the effects. This approach enables them to provide training and mentoring to multiple businesses simultaneously and create a network of interested companies willing to participate and learn from one another.

The advocacy organisation 4 Day Week Global (4DWG) Foundation is currently pushing for the largest private-sector pilot to date, which they claim is underway with over 180 companies in several countries. The programme is operated in collaboration with the think tank Autonomy, alongside a data-collection partnership with researchers at Boston College, Cambridge University, and Oxford University (Cohen, 2022).

Private-sector pilot studies continue to be conducted in the UK, US, New Zealand, Australia, Ireland, and Canada under the support of the 4DWG, the 4 Day Week Campaign and other advocacy organisations. For example, over 70 companies in the UK participated in a pilot from June to December 2022, and the first results from 33 companies primarily based in Ireland and the US were recently shared at the end of 2022 and are discussed below as one of the case studies (Schor et al., 2022).



Several governments are also leading private-sector pilots with multiple companies. For instance, the Spanish national and the regional governments of Valencia in Spain are preparing the first pilots in Spain (Kassam, 2021). Additionally, the Scottish and the Portuguese are preparing their pilots with multiple companies (Tzvetozar, 2022).

Other pilots with multiple companies involved are based solely on qualitative research and do not provide comparable quantitative data, such as one with small and medium-sized enterprises (SMEs) in the financial services sector in New Zealand (Delaney et al., 2022).

Most private-sector pilots involve small and medium-sized businesses, which is relevant as they constitute the majority of business ecosystems. For example, one private-sector pilot in Switzerland includes multiple larger Swiss companies (Neubert et al., 2022). However, in this case, the worktime reduction included a pay reduction and should not be considered part of the recent four-day week trend.

### *2.1.2 Individual-company pilots*

In addition to private-sector pilots involving multiple companies, there are also individual company-led pilots in which some employees voluntarily participate in a four-day week. Advocacy organisations and university researchers usually support these pilots, but they are not part of a more extensive pilot programme involving multiple companies. In 2019, the think tank Autonomy identified a dozen of these pilots, most of which were run by small and creative agencies such as Agent Marketing, a marketing firm in Liverpool, and Conversation Creation, a PR and marketing agency in Bristol (Stronge & Harper, 2019). However, there's no centralised database collecting all individual four-day week pilots, and it is difficult to assess how many are currently underway.

The most well-known company introducing the four-day week at the company level was Perpetual Guardian, a New Zealand firm that manages trusts. It conducted a four-day workweek pilot among its employees in 2018, and it is included below as one of the case studies given its relevance.

There have also been working time reduction pilots in larger companies. For example, Volkswagen in Germany reduced the working week from 36 to 28.8 hours (Seifert & Trinczek, 2000). More recently, Microsoft Japan tested a four-day workweek and claimed a productivity boost of 40% (Chappell, 2019).<sup>2</sup> Some large companies, such as Telefónica in Spain, decided to offer a four-day week but with a reduction in pay (Sanchez & Galindo, 2022). Therefore, these are out of scope from the recent four-day week trend.

Although these company-level pilots can be helpful in concrete and specific cases, they have less external validity than private-sector pilots involving multiple firms because the results are likely to be context-driven. In general, these companies have reported successful results, although most share little information publicly, making it difficult to analyse them in detail.

## **2.2 Public Sector pilots**

In the early 2000s, the Nordic countries were the first to run pilot programmes for public sector employees before the recent wave of four-day workweek pilots. For example, a pilot programme in the Swedish city of Gothenburg tested the impact of reduced working hours in a retirement home over 23 months (De Spiegelaere & Piasna, 2017). In this programme, workers' hours were reduced to 30 hours per week, compared to a retirement home that did not have the opportunity to reduce working hours.

These pilot programmes focused mainly on exploring the consequences of reduced work hours in Sweden and Norway, with a particular emphasis on health outcomes (Åkerstedt et al., 2001). Some

---

<sup>2</sup> The main results can be found only in Japanese [here](#).

pilot programmes were also conducted in Finland (Nätti and Anttila, 1999). However, these early-stage pilot programmes had limited sample sizes, non-random comparison groups, and very narrow target groups, such as workers in nursing homes, centres for the elderly, dental healthcare, or kindergartens.<sup>3</sup>

Nonetheless, the largest public-sector pilots in the Nordic countries were run in Iceland from 2015 to 2019, where some public workers experienced a reduced workweek of 35-36 hours without a pay reduction (Haraldsson & Kellam, 2021). Due to this pilot programme's relevance and international attention, it is included as one of the case studies in this working paper.

Some governments have introduced a four-day week for some public workers that did not include a reduction in working hours. As a result, they had to work the same number of working hours condensed into four days, and therefore it would be out of scope for this analysis. For example, Utah's state government pilot in the USA from 2008 to 2011 or, more recently, the Belgium government pilot in 2021, belong to this group (Martens & Pronina, 2022). However, other governments are planning to run four-day week pilots for public sector employees, with reduced working hours and no pay reduction, such as in Scotland and Portugal (Tzvetozar, 2022).

### **3 Case studies**

To provide a more in-depth analysis of the various features, outcomes and limitations of the four-day week pilots, this working paper has chosen three case studies to examine in detail. The goal is to select one example from each category previously described: one multiple-company pilot, one single-company pilot, and one public-sector pilot.

The case study selection criteria are based mainly on the pilots' relevance and the availability of publicly accessible information and results in English. These three pilots have been extensively covered in the media and are highly relevant to the four-day workweek movement. Additionally, these pilots have shared more information than other pilots mentioned earlier. Whilst none of the pilots have given public access to the microdata, they have published quantitative reports that summarise the necessary information to assess the validity of the studies, albeit to varying degrees of detail.

The case studies vary significantly in terms of intervention design, pilot characteristics, and timeframes. Table 2 provides a comparative overview of each case study key characteristics.

---

<sup>3</sup> The main results can be found only in Swedish [here](#).

Table 2: Case Studies of Four-day Week Pilots

	<b>Multiple-company Pilot</b>	<b>Single-company Pilot</b>	<b>Public Sector Pilot</b>
	<u>4 Day Week Global</u>	<u>Perpetual Guardian</u>	<u>Iceland</u>
<b>Relevance</b>	First two pilots of a longer list of pilots run during 2022 and 2023	First pilot run by an individual company	Two largest pilots run by the Reykjavík City Council and the Icelandic National Government
<b>Locations</b>	Remotely, Ireland, US, Australia, New Zealand	New Zealand	Iceland
<b>Intervention length</b>	6 months	2 months	Approximately 4 years and 6 months
<b>Period of intervention</b>	First pilot from February to August 2022, and second pilot from April to October 2022	From March 2018 to May 2018	First pilot from March 2015 to September 2019 and second pilot from May 2017 to 2021
<b>Original sample</b>	33 companies and 903 employees	1 company and 240 employees	66 public workplaces and 2,500 employees
<b>Expected change in hours worked</b>	From 40 to 32 hours	From 37.5 to 30 hours	From 40 to 35 or 36 hours on average
<b>Adopted a four-day workweek?</b>	Yes	Unclear	No
<b>Supporting organisations</b>	Boston College, University College Dublin, and Cambridge University	The University of Auckland and Auckland University of Technology	The think-tank Autonomy

Note: The information provided is based on the reports published, which in some cases is unclear or contradictory and may not be exact.

### 3.1 4 Day Week Global (4DWG) pilots

In 2022, the 4DWG Foundation supported companies and non-profit organisations interested in implementing a four-day work week with no reduction in pay, leading to the world's first coordinated six-month pilots and a large-scale research effort on the impacts of a four-day week (Cooban, 2022). The 4DWG recently presented research results from the first two pilots, which included 33 volunteer companies and a total of 903 employees (Schor et al., 2022).

Most of the companies were based fully remotely (12 companies) in Ireland (11 companies) and the US (6 companies), with the remaining four companies from the UK, Australia, and New Zealand. The pilots were conducted from February to August 2022 and from April to October 2022, with two months of preparation time for the companies, including attending workshops, coaching and mentoring, and participating in a peer support network. Even if there are two separate pilots, they have been studied and analysed as a single project.

Whilst companies were not required to implement a specific work time reduction strategy or a four-day week to participate in the pilots, 29 of the 33 companies adopted a four-day week for all employees, with the remaining four companies reducing daily hours for a subset of employees who remained on a five-day schedule. Of the 33 companies, 27 completed a final survey in which they were asked about their overall experience with the four-day week and whether they would continue. The final matched sample size for employee data was 495 workers who completed baseline and endpoint surveys.

Alongside facilitating the pilots, the 4DWG also provided research conducted by academic researchers from Boston College, University College Dublin, and Cambridge University. The research comprised two parts, with one part focusing on administrative data from the participating companies and the other part gathering self-reported survey data from employees with ranges applied to different outcomes. In both cases, a before-and-after comparison was conducted as the evaluation method.

### **3.2 Perpetual Guardian pilot**

In March 2018, Perpetual Guardian, a New Zealand company specialising in trusts, wills and estate planning, initiated an eight-week trial of a four-day workweek. The company aimed to motivate employees to take a day off instead of compressing their working hours into shorter days. Employees received the same remuneration and were expected to deliver the same output level. Overall, employees worked for 30 hours but were paid for 37.5 hours. However, it is unclear how many employees worked exactly four days as opposed to five.

The Perpetual Guardian pilot was the first of its kind at the company level and generated significant international attention. In the spring of 2018, the pilot was the subject of global news coverage after it claimed a 20% increase in employee productivity and a 45% improvement in work-life balance (Chappell, 2019). Following the eight-week trial, the company was so impressed that it allowed workers to opt-in to the four-day week across its operations. Some reports suggest that the policy has since been made permanent.

Official records show that the pilot involved all 240 employees. For the analysis, respondents were matched based on the baseline and endpoint surveys, resulting in a final sample of 122 employees. The analysis also included a survey of supervisors, with a matched sample of 27 respondents. The analysis relied on self-reported data and used a set of ranges to assess various outcomes. The evaluation was conducted through a before-and-after comparison using the baseline and endpoint surveys.

To evaluate the level of employee engagement, Perpetual Guardian partnered with academic institutions, The University of Auckland and Auckland University of Technology. The main researcher provided some pertinent quantitative information that can be used to gauge the effectiveness of the pilot (Haar, 2018). However, as noted in the report, this is only a summary of preliminary data published in July 2018. As of early 2023, no additional reports with more conclusive data have been publicly released.

### **3.3 Iceland pilots**

In 2015 and 2017, two major pilots of a shorter working week were initiated in Iceland by the Reykjavík City Council and the Icelandic National Government in response to campaigning by trade unions and civil society organisations. The pilots involved a different reduction in working hours, depending on the particular workplace, with no reduction in pay. According to complementary information, 47% of workplaces reduced 1 or 2 hours of working time, whilst 45% reduced 3 hours

and 8% reduced 4 hours. On average, workplaces reduced 2.4 hours of working time.<sup>4</sup> Even if there are two separate pilots, they have been studied and analysed as a single project.

The first Reykjavík City pilot started in March 2015, with two workplaces and a few dozen workers, and it expanded to over 2,500 staff during the following years, although it is unclear how the sample developed over time. According to the additional information provided, a total of 61 workplaces participated in the pilot. An additional workspace was included as a comparison group at the beginning of the pilot, but it is not specified how it was selected or if more workspaces were included in the comparison group.

The pilot was implemented in very diverse workplaces, including offices, playschools, city maintenance facilities, and care homes. The pilot ended on September 1, 2019, with staff reverting to their previous working hours. However, agreements were signed a few months later, guaranteeing shorter hours.

Another pilot at the national level began in April-May 2017, with 17 workplaces applying and 4 selected, including a police station and the Directorate of Internal Revenue. A total of 5 workplaces participated in the pilot. Four other workplaces were selected as comparison groups for the pilot that would not receive the four-day week mandate. The criteria for group selection are not specified.

According to official figures, around 440 staff members were part of the intervention group, with 364 in comparison group workplaces. The pilot ended in 2021, but it is not specified when the workplaces ended and if they all did so simultaneously.

Quantitative studies based on administrative and self-reported data collected on different surveys for treatment and comparison groups were routinely conducted by the Reykjavík City Council and the Icelandic Government, both as part of their usual operations and specifically for the pilots. Unfortunately, these government reports are in Icelandic and cannot be analysed in this assessment due to language barriers.<sup>5</sup> However, a think-tank, Autonomy, published a final report summarising all the relevant quantitative information that can be used for a general assessment of the pilots (Haraldsson & Kellam, 2021). It should be noted that more detailed information may be available in the government reports that are not included in this assessment.

## **4 Overview of performance outcome indicators**

This section compares the outcome measures used in the selected case studies of four-day week pilots (Schor et al., 2022; Haar, 2018; Haraldsson & Kellam, 2021). The objective is to identify similarities and differences in measuring the data. By comparing the outcome measures used in the case studies, it should be easier to identify best practices and potential areas for improvement in measuring the success of the four-day week interventions.

### **4.1 Health and well-being**

The pilots used various survey-based indicators to assess the impacts on the health and well-being of participants. The Perpetual pilot aimed to measure health satisfaction and job stress before and after the intervention, which refers to workers' assessments of how stressful their jobs are. The Iceland pilot aimed to capture self-reported well-being at work and symptoms of stress for the treatment and the comparison groups.

---

<sup>4</sup> These are the author's calculations based on the additional information provided in Appendix I of Haraldsson & Kellam (2021). The main information provided in the same report states that hours per week were shortened "from 40 hours to 35 or 36".

<sup>5</sup> Additional information in Icelandic can be found [here](#).

In contrast, the 4DWG pilots focused more on self-perception of mental health and positive and negative emotions, such as cheerful and vigorous or lonely and tense. The Perpetual pilot also asked workers to rate how well the company cares for their physiological health and well-being and compared them before and after the intervention.

In addition, the 4DWG pilots collected data on self-reported anxiety levels, self-perception of physical health, and weekly physical exercise frequency. The Iceland pilot also reported that some participants claimed to exercise more than before, although it is unclear whether this information was obtained through qualitative or quantitative data.

## **4.2 Job satisfaction and work environment**

In order to measure changes in job satisfaction among pilots, various survey questions were used. Both the 4DWG and Perpetual surveys included a direct question about job satisfaction and compared them before and after the intervention, with the former asking respondents to rate their level of satisfaction on a scale of 0 to 10. The survey of the 4DWG also included additional questions on burnout and work stress. In the case of Perpetual, questions related to workload and work nature, such as work demand, team strength, and team cohesion.

The 4DWG survey also included questions about turnover intentions, job security, and second job involvement. Additionally, they asked questions about work intensity and complexity, such as the opportunity to learn new things or have more autonomy. The 4DWG survey also included retrospective questions that asked workers to reflect on the beginning of the pilot and provide feedback on changes in work pace and second job intentions.

In contrast, the Icelandic pilots focused their questions on work operationalisation and asked survey questions regarding team support, role confusion, fair and encouraging management, and worker control over work pace, and compared these with workplaces in the comparison group.

## **4.3 Work-life balance**

All three groups of pilots collected information on self-perceived work-life balance, though using slightly different survey questions. In the Perpetual pilot, workers reported how well they balanced their work and non-work roles. The 4DWG asked about the ease of combining paid work with social life on a 1-5 scale. The Icelandic pilot reportedly collected information in work-life balance for the treatment and the comparison groups.

In addition to the primary work-life balance question, the 4DWG included additional questions to explore the conditions of workers, such as sleep problems and fatigue. They also asked questions related to family dynamics and compared them before and after the intervention, including whether workers felt too tired from work to do household tasks and had difficulty concentrating on work due to family responsibilities. The 4DWG survey also collected information on combining paid work with care responsibilities.

## **4.4 Productivity and public service provision**

This is the only question where the pilots aimed to capture data from administrative sources instead of using survey-based questions. However, the Perpetual pilot did not report any measure of business productivity.

The 4DWG pilots aimed to capture changes before and after the intervention in the revenue of the companies by weighting the data by company size. They also retrospectively asked for revenue from the same six-month period for the previous year before the pilot began, which they used as a comparison period. They then calculated the weighted average percentage change across the companies that supplied sufficient data and compared it with that period.

In Iceland, they measured public service provision by observing the number of open cases and active cases as a proxy for public sector productivity. They observed changes in both open and active cases as well as the number of applications, although authors highlighted that it may be influenced by seasonal factors. However, each workplace is different, and they collected various metrics, such as the time of each process and the number of visitors or calls received. They also observed other metrics, such as the number of immigration applications or customer satisfaction.

#### **4.5 Workers' performance**

Private-sector pilots have aimed to capture changes in worker performance through various means. In the Iceland pilots, however, no specific questions were reported about worker performance. In the Perpetual pilot, workers were asked to self-report their perceived service performance and compare them before and after the intervention, specifically how they engage with customers and whether there has been an improvement. Additionally, work engagement was measured to determine if workers were more engaged and willing to go the extra mile. The 4DWG pilot did not include similar questions but rather asked about workers' schedule control and compared their current work ability to their lifetime best as a proxy for productivity and performance.

In addition to the primary questions, the Perpetual pilot also included additional inquiries into team dynamics, such as team creativity, reflecting workers' perceptions of how innovative and creative their team is, and team behaviours, capturing employees' perceptions of how well their team engages in helpful work behaviours.

To provide a contrasting perspective based on workers' and supervisors' viewpoints, the Perpetual pilot also aimed to capture supervisor data. Supervisors were asked to reflect on how they perceived their employee teams to perform. Supervisors were also asked about their team's attendance behaviours, specifically related to being on time for their job positions, as well as their team's creativity behaviours and additional positive behaviours outside of those required in the job description.

#### **4.6 Absenteeism**

Among the three case studies, only the 4DWG pilot included an analysis of absenteeism. The study employed two methods to capture this data. Firstly, workers were asked to report the number of days they were absent from work due to sickness or other health-related reasons in the weeks preceding the survey and compared them before and after the intervention.

Secondly, administrative data was used to collect information on the number of sick and personal days taken by each employee per month and was compared with the comparison period a year before the pilot began.

#### **4.7 Employment and employee retention**

Private-sector pilots aimed to assess the impact on employment, although the Icelandic pilot did not report on this aspect, as the employment dynamics in the public sector differ from those in the private sector. The Perpetual pilot did not incorporate employment measures but instead relied on survey-based questions to gauge workers' attitudes toward remaining with the organisation.

In contrast, the 4DWG pilot aimed to obtain administrative data on employment. Specifically, it collected information on the number of new employees, as well as resignation and hiring rates per 100 employees. The data was compared from the start of the pilot to the endpoint and adjusted for company size.

In addition to tracking these changes, the study also used the comparison period of one year prior to the start of the pilot to compare hiring and resignation rates for the companies that provided that information.

## **4.8 Time allocation and life satisfaction**

The three groups of pilots collected various information regarding time allocation. Both Perpetual and the 4DWG pilots used survey-based questions to investigate life satisfaction and compared them before and after the intervention, although they used different metrics. The 4DWG pilots used a 1-10 scale to determine satisfaction with life, whereas Perpetual collected data on the percentage of workers who reported being satisfied with their lives.

In addition to life satisfaction, Perpetual also inquired about leisure satisfaction, reflecting overall satisfaction with leisure time, and community satisfaction, pertaining to involvement with the local community. The 4DWG pilots aimed to capture satisfaction with time, attempting to collect data on where the extra time was allocated, as well as observing any changes in time dedicated to hobbies, volunteering, housework, and other caregiving activities. The study also collected information on household finance satisfaction and relationship satisfaction.

The Iceland pilots also included questions regarding participants' satisfaction with the time they allocated for themselves and their families. They also reported whether workers felt more energised and less stressed, resulting in more energy for other leisure activities. The pilots were particularly interested in the impact on single parents, although it is unclear whether this data was collected through interviews or other qualitative means instead of surveys.

## **4.9 Gender equality**

Efforts were made to measure the impact of the three groups of pilots on gender equality, but limited information is available. The Perpetual pilot tried to explore gender differences across the examined variables, although they did not specify which ones were explored. Similarly, the Iceland pilots observed a division of household labour. However, they did not provide details on the source of this effect.

The 4DWG pilots attempted to capture self-reported changes before and after the intervention in the division of labour by gender in couples, as well as collecting survey-based data on the share of time spent on childcare and housework by gender. They also collected information on childcare costs.

## **4.10 Environmental footprint**

The private-sector pilots aimed to assess the environmental impact of the pilots, although they used different approaches. The Iceland pilot did not mention this matter. In the Perpetual pilot, participants were asked a series of retrospective questions about changes in energy use, leisure travel, and other pro-environmental behaviours such as recycling, purchasing eco-friendly products, and choosing to walk or cycle instead of driving. The pilot also asked about environmental education and volunteering to help care for the environment.

In contrast, the 4DWG pilots attempted to collect self-reported data on commuting time, to observe changes in hours spent commuting over the study period. They also observed changes in the percentage of workers commuting by car. The pilots also collected self-reported data on the number of domestic and international trips taken in the past weeks. The pilot reported collecting information on the carbon footprint, but these results have not been shared yet.

## **5 Main limitations of current pilots**

This section is the central part of the working paper, focusing on the limitations of the four-day week pilots. Whilst the three case studies analysed are the primary source of information, the discussion extends to other pilots as well.



When the OECD aims to verify the validity of business support program evaluations, it uses a quality assessment metric. A low rating is attributed to evaluations with a limited sample size, simplistic and poorly executed evaluation methods, misaligned impact variables with program aims, and an absence of survival analysis (OECD, 2023).

Whilst the OECD score may not be directly applicable, the evaluation quality assessment framework is valuable for analysing the validity of the four-day week pilot programmes. By using it as a basis, this section will delve deeper into the methodological limitations, making a conscious effort to use accessible language wherever possible and focusing more specifically on the inherent problems of four-day week interventions. Table 3 presents a general summary of the main issues encountered in each set of problems for the case studies. The table includes a tick mark in each box where the specific issue was found in the corresponding case study.

The limitations found in the case studies can be divided into four sets of problems: concerns with addressing causality, issues with collecting accurate data, limited transparency when reporting results, and incomplete information on sample and intervention characteristics. The categorisation of limitations into these four sections is designed to offer an outline for understanding some of the main methodological challenges associated with a quantitative impact evaluation. For instance, the use of the evaluation method, the choice of impact variables, the implementation of the evaluation method and the selection of the sample.

By outlining these categories, this section provides a comprehensive overview of the methodological issues that impact the validity and generalisability of the pilots' findings. Identifying these challenges will help researchers develop strategies to improve the quality of evidence and increase the value of the intervention.

Table 3: Main Quantitative Limitations Presented on the Four-day Week Case Studies

	<b>4DWG Pilots</b>	<b>Perpetual Guardian</b>	<b>Iceland Pilots</b>
<b>Problems addressing causality</b>			
Results affected by omitted variable bias	✓	✓	
Results affected by selection bias			✓
<b>Collecting accurate data</b>			
Overreliance on self-reported data		✓	✓
Low-quality administrative data	✓	✓	✓
Unreliable survey questions	✓		✓
<b>Transparency in reporting results</b>			
Incomplete statistical information	✓	✓	✓
Lack of pre-registration plans	✓	✓	✓
Limited analysis of non-positive responses		✓	✓
Weak robustness checks	✓	✓	✓
Misuse of accessible summaries with digestible formats		✓	✓
<b>Explaining sample and intervention characteristics</b>			
Problems with the sample size	✓	✓	
Lack of representativeness of the sample	✓	✓	✓
Sample affected by low compliance	✓	✓	
Constrained ability to observe heterogeneous effects	✓	✓	✓
Missing analysis of non-compliers	✓	✓	✓
Limited information on costs and scalability of the intervention	✓	✓	✓

## 5.1 Addressing causality

Pilots are designed to investigate cause-and-effect relationships by posing a causal question, which is "What is the impact of the four-day week on various outcomes, such as productivity, work-life balance, and workers' performance?" However, establishing causality is one of the major challenges in social science research. If there are other factors that contribute to a correlation between a treatment and an outcome, the overall correlation cannot be interpreted as a causal effect. This issue is often referred to as the endogeneity problem, and it requires the creation of a valid counterfactual, which demonstrates what would have happened in the absence of the programme (Hill et al., 2021).

There are a variety of experimental and quasi-experimental strategies available to assess causality in research, each with its own set of strengths and limitations. Whilst these evaluation methods require certain assumptions in order to generate valid estimates, non-experimental estimations are generally considered weaker for addressing causality due to the tendency of their underlying assumptions to be violated (Angrist & Pischke, 2010). Of particular concern are two common non-experimental evaluation methods: pre-post (or before-and-after) comparisons and self-selected (or enrolled-and-unenrolled) comparisons, which have been shown to yield unreliable estimates of the counterfactual due to the presence of selection bias and omitted variable bias (Gertler et al., 2016). As a result, most evaluations and academic experts consider the use of these methods to be highly risky.<sup>6</sup>

All private-sector pilots mentioned in this working paper, including those that are not yet finalised, employ before-and-after comparisons, with only one exception that used a more robust evaluation method.<sup>7</sup> The general approach taken involves comparing participant businesses before and after the introduction of the four-day week intervention and tracking how it has changed over time. For example, the 4DWG pilots mostly use the baseline surveys collected six months earlier as a comparison period, as well as data collected a year before the intervention was introduced. Some researchers may have tried to introduce a comparison group with non-participant companies but didn't seem to work as it was not included in the final report.<sup>8</sup>

The validity of the before-and-after comparison approach hinges on the assumption that other changes in the outcome of interest over time are not expected to be influential, which is highly unlikely in this scenario. Consequently, changes that are unrelated to the program's impact may be erroneously included as part of the estimated effect of the four-day week. For example, the 4DWG pilot captured a statistically significant reduction in childcare costs, which has nothing to do with the four-day week pilot. Various economy-wide effects, such as earnings or employment, or external shocks affecting outcomes, such as productivity or environmental footprint, could be influenced by other relevant factors as well. Additionally, external conditions like a pandemic or an economic crisis can impact employees' perceptions and well-being.

Only one pilot provided additional information to test the before-and-after assumption, albeit unintentionally, indicating that the assumption was not met. To observe if the data was "potentially inflated" (Haar, 2018; p:11), the Perpetual pilot collected information four months before the baseline and compared it with the baseline and the endpoint results separately. As illustrated in Chart 1 below,

---

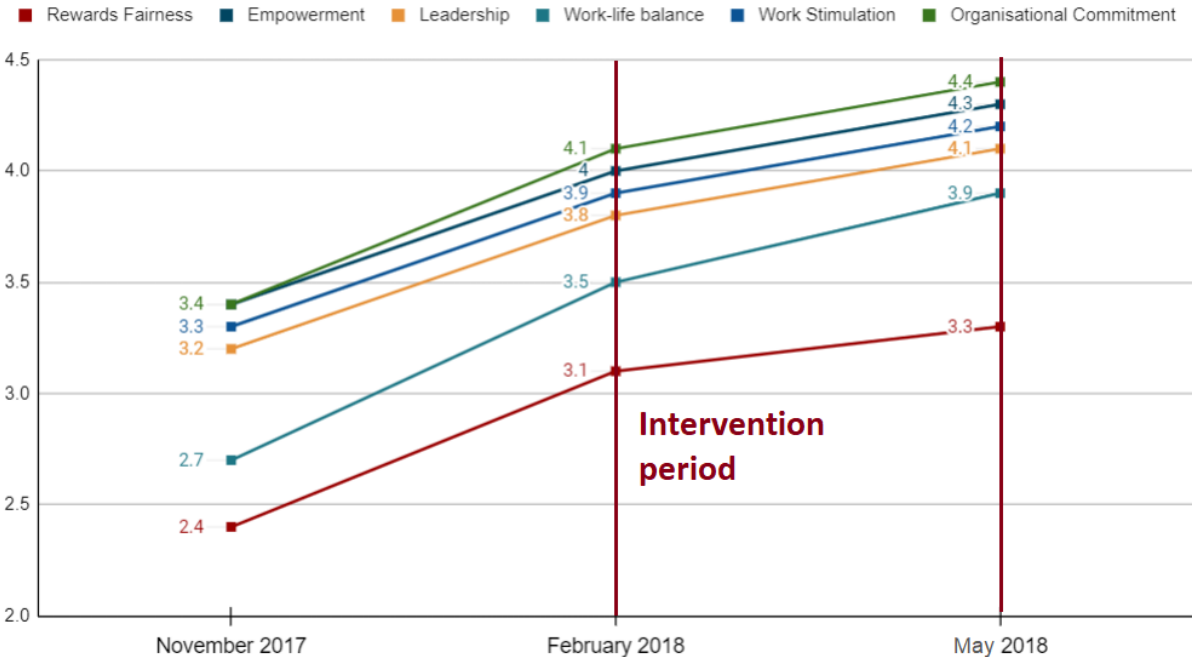
<sup>6</sup> To learn more about the limitations of these impact evaluation methods and the use of experimental and quasi-experimental methods, an introductory assessment for public policy by the World Bank can be found [here](#).

<sup>7</sup> Neubert et al. (2022) present a pilot program in Switzerland that reduced working hours while also reducing pay that employed a quasi-experimental variation to estimate its effects. Nonetheless, the authors also discuss various methodological limitations associated with this approach.

<sup>8</sup> A detailed report of the Irish companies within the 4DWG pilots stated that they aimed to compare the results against two companies that decided not to participate in the pilot (Kelly et al., 2022). However, the results weren't shared in detail, and only a small sample of employees from the non-participant companies responded. Regardless of these limitations, it would still be affected by selection bias.

all variables experienced statistically significant increases during the four months before the pilot began, from November 2017 to February 2018, when the baseline data was collected. Some changes are even greater than the increase between February and May 2018 during the intervention period. This unusual scenario likely points to other factors captured in the analysis and biasing of the before-and-after results.<sup>9</sup>

*Chart 1: Time Trends of Employee Engagement Data in Perpetual Guardian Before and During the Four-day Workweek Intervention*



Source: Composed by Cuello, H. based on data from Haar, J. (2018), *Overview of the Perpetual Guardian 4-day (paid 5) Work Trial*.

On the other hand, the public sector pilots have also used before-and-after comparisons with certain variables. However, they have generally tried to incorporate a comparison group as a counterfactual. This approach was adopted in most evaluations of the Iceland and Sweden pilots, where other offices and public centres that did not experience a reduction in working hours were also included in the analysis. Nonetheless, the comparison groups were not randomly assigned, and no quasi-experimental identification strategy was employed for non-random group selection, except for one pilot where randomisation was effectively applied at the workplace level.<sup>10</sup> Thus, the majority of evaluation methods employed by public-sector pilots fall under the category of traditionally weak enrolled-and-unenrolled comparisons.

The Iceland pilots involved several workplaces at different time points as part of the treatment and comparison groups. The Reykjavik pilot included at least one workplace as a comparison group, and

<sup>9</sup> The author of the report suggests that the positive impact of the intervention can be observed in the baseline results because employees were already aware of the upcoming change. However, there is no additional evidence presented to support this claim. From a theoretical standpoint, it is unlikely that the improvement can be fully attributed to the four-day workweek since, in practice, it had not yet been implemented, and employees were still working for five days. For instance, it is difficult to explain why the work-life balance appeared to have improved more when employees were still working for five days than when the four-day week was implemented, based solely on an earlier shift in their perceptions.

<sup>10</sup> The only randomised controlled trial found among the Nordic pilots was Schiller et al. (2017). It was run in Sweden in 2005 and tested the effects on sleep and perceived stress among public sector employees.

the Icelandic government pilot had at least four workplaces as a comparison group. However, when group selection is not randomly allocated, both groups are likely and significantly different from each other at baseline in both observable and unobservable characteristics.

This presents significant endogeneity problems and a major threat to the internal validity of the Iceland pilots, as the comparison may be highly skewed due to pre-intervention differences across the groups. For instance, treated workplaces may comprise more motivated public sector employees with higher team engagement and better management systems that would bias the comparison.

In essence, both private and public sector pilots have significant limitations in establishing causality, and their evaluations may require revision. Therefore, it is crucial to approach the quantitative results of most four-day week pilots with caution and consider them solely as indicative data. Qualitative reports have provided greater value and can aid in comprehending the potential direction of the impacts. However, given the limited internal validity of the quantitative results, it should not be considered robust evidence of the positive impact of the four-day week.

## **5.2 Collecting accurate data**

The present private and public sector pilots rely mostly on self-reported survey data, which concentrate on employee perceptions. Whilst this can provide useful insights, it is vital to acknowledge its limitations, which are not regularly discussed in detail in the reports. Self-reported data can be vulnerable to biases and reporting errors.

It is possible that individuals may not accurately report changes or present a biased view of their experiences due to various factors. For instance, employees may report a positive impact on productivity or work performance to encourage their workplace to continue with the four-day week, as they place greater importance on the benefits to their well-being. Additionally, the Hawthorne effect may pose a potential limitation, as participants may alter their behaviour simply because they are aware of being part of a pilot study. Moreover, participants may be more likely to respond positively due to social desirability bias, particularly in companies with a progressive and inclusive management style that are willing to implement the four-day week.

Even more problematic is the use of retrospective questions, which ask participants to recall how they felt at the beginning of the pilot and provide feedback on changes. This approach has been used in 4DWG pilots to measure job satisfaction and productivity. The problem with retrospective data is that it is subject to memory biases and inaccuracies. Participants may have difficulty accurately recalling their initial feelings and experiences, especially if a significant amount of time has passed since the start of the pilot. Additionally, participants' current feelings and experiences may influence their recollection of past events, leading to distorted or unreliable data.

To address some of the limitations associated with workers' self-reported data, some pilots have included separate perceptions from managers or supervisors to compare with those of employees. For instance, Perpetual Guardian gathered insights from 23 managers, which provided valuable qualitative data, but the small sample size limits its quantitative validity.

Ideally, any survey-based data should be supplemented with administrative data to provide a more comprehensive picture. The 4DWG pilots have incorporated company data, such as revenue, average and total hours worked, resignations, new hires, and sick and personal days taken, as part of their evaluation. However, this approach was also limited by a small sample, as only 50% of the participating companies provided sufficient data. Furthermore, as will be explained below, the data was weighted, and statistical significance was not always reported.

When it comes to performance indicators, there are notable discrepancies in the level of detail, relevance and comprehensiveness of the information provided. Whilst some areas have extensive complementary data, others, such as absenteeism and employment or gender equality, are missing critical information. Overall, the Icelandic trials lack sufficient information on most of the outcomes, with a mix of quantitative and qualitative data presented without clear reporting of the origin of the

data and the effect sizes. It is unclear whether outcomes were observed and surveyed or mentioned in interviews but not necessarily representative of the sample.

Matching impact measures with programme objectives is also essential. Some pilots adopt more valid and reliable survey questions in certain areas than others. For instance, the Perpetual Guardian pilot included more relevant questions to evaluate workers' performance than the 4DWG pilots, which almost only reported work ability compared to the workers' lifetime best. This is a highly ambiguous way of measuring workers' performance because it does not provide a clear understanding of their current performance level or the change in performance over time. Nor it reported any questions observing changes in team creativity, ideas generated or reduction of wasted time.

On the other hand, the Perpetual pilot falls short and provides less relevant information than the 4DWG pilots on areas like employee retention. The Perpetual Guardian pilot did not incorporate any administrative data on employment, unlike 4DWG, which provided data on resignations, new hires, and sick and personal days taken. The Perpetual pilot only relied on survey-based questions to assess workers' attitudes toward remaining with the organisation.

When analysing the productivity impact of private-sector pilots, there are certain limitations that need to be considered. First, it can be difficult to identify valid productivity indicators that are comparable across different businesses and time periods. Additionally, seasonal variations can significantly influence productivity data, which can be problematic when conducting short-term pilots and making before-and-after comparisons. To address these issues, it is important to encourage participating businesses to collect productivity data in a standardised and comparable manner from the beginning of the pilot.

The 4DWG pilots attempted to mitigate the effects of seasonal variation by retrospectively collecting revenue data from the same period a year earlier, but this approach has limitations as it can be subject to selection bias and doesn't account for natural business changes. Similar limitations can be observed in the public-sector pilots conducted in Iceland, where the number of open cases is highly affected by seasonal factors, and before-and-after comparisons may not provide reliable results. For example, in one police department, the number of investigative cases closed during the trial period increased after the pilot started, but it is unclear if this increase can be attributed to the pilot itself, as caseloads can fluctuate between years and months.

### **5.3 Transparency in reporting results**

Openness and transparency have long been considered crucial tenets of the scientific community, as the integrity of research is predicated on the ability to replicate results. For this reason, reporting should follow established academic conventions, widely recognised within the scientific community, to enable others to reproduce the outcomes of an intervention. In the social sciences, adhering to these standards is important as it helps to determine the value and quality of the research presented. Nevertheless, recent studies indicate that current research practices often deviate from this ideal and can occasionally produce misleading findings (Christensen & Miguel, 2018).

Four-day week pilot studies are largely undertaken by governments and advocacy organisations, and the results are typically not published in academic journals but rather accessed through their websites. As a result, these studies are not held to the same rigorous and transparent reporting standards, such as peer review and high-quality robustness checks, which are typically required in academic research. This lack of transparency in reporting may result in issues such as p-hacking, cherry-picking, and data dredging, ultimately undermining the credibility of the analysis.

Some pilots, including those conducted in Iceland and the company data of 4DWG, do not consistently report basic statistical information on the results, such as the statistical significance of the estimation, as they only show the p-values for some coefficients but not for others. Consequently, some positive results reported may be caused by random sampling error, which is not always acknowledged or addressed. For instance, in the 4DWG pilots, when measuring the weighed resignation rate per 100

employees, they observed a small positive effect that is highly likely to be statistically insignificant, as the unweighted data points to a small negative effect and the sample size is small.

In general, statistical significance needs to be better reported and discussed. For instance, the Perpetual Guardian pilot has addressed the issue of statistical significance by showing only the p-values. However, there is still a need for more comprehensive reporting, such as providing confidence intervals, standard errors and standard deviations to give a more complete and precise picture of the magnitude and direction of the effect.

In some cases, the limited statistical information provided may underestimate the credibility of the evaluations. For example, the Iceland pilots did not provide any balance check or descriptive statistics in the report that compared the baseline results of the treatment and comparison groups. Therefore, it is essential to clarify whether the assumption is met or if there were significant imbalances that would correlate with the outcomes and bias the results of the group comparison.

Since the organisations and governments coordinating and running the pilots did not share any pre-analysis or pre-registration plan, it is unclear what they intended to measure in the first place. This might have led to reporting only those results that had positive effects, leaving several unreported outcomes of interest that did not show the expected results. This becomes more concerning as the pilots analysed a very large number of outcomes, and the likelihood of finding a positive result by chance rises due to the multiple comparison problems.

The 4DWG provides valuable analysis of non-positive responses. However, in general, pilots only report a few non-positive results, and when they do, authors tend to claim that the result was expected. However, this cannot be verified, and arguments sometimes indicate confirmation bias. For example, on the Perpetual Guardian pilot, the supervisor's data show no statistically significant impact on job performance, likely driven by a small sample size. The author states that this is an expected result: "given they have 20% less time to do their jobs, I think it is likely there will be no increase in performance" (Haar, 2018; p:8). However, this directly contradicts an earlier result where workers' data found an improvement in self-reported performance and the author stated that it was expected as "employees are likely to work harder" (Haar, 2018; p:4). This is an apparent contradiction that undermines the credibility of the interpretation of the results.

In general, pilots only include a few robustness checks. The Perpetual pilot includes estimating two variables that, according to their model, should not be influenced by the four-day week (proactive personality and job complexity) to "test for the confidence of the data" (Haar, 2018; p:5). Researchers do not find an effect on the selected variables, leading them to claim that the results are not biased. However, it would be crucial to include higher-quality checks, as this does not rule out the possibility that more relevant estimations would be affected by omitted variable bias.

Finally, when reporting the summarised findings in a digestible format to the outside world, it is crucial to be transparent and avoid cherry-picking results and changing estimations. For instance, on the main website providing general information about the Perpetual Guardian pilot results, some pieces of information are confusing, such as the actual sample size or the before-and-after evaluations, which are compared with the November 2017 pre-pilot results instead of using the baseline survey. This inflation of results misleads the real impact of the four-day week. As explained above, there was a significant increase in many outcomes before the beginning of the pilot that should not be entirely attributed to the four-day week. This has misled most media to report a 45% increase in work-life balance when the results reported using the baseline is closer to an 11% increase.<sup>11</sup>

---

<sup>11</sup> These are the author's calculations based on a summary report that presents the work-life balance outcomes by comparing the pre-pilot survey of November 2017, conducted four months before the intervention. The results are presented in percentage format (54%) and compared with the endpoint survey (78%), indicating an increase of approximately 44.4% (24 percentage points). However, when using the baseline results of February

## 5.4 Explaining sample and intervention characteristics

When analysing any pilot, sample size and characteristics are essential elements to consider. A small sample size can challenge the pilot's internal validity, limiting the credibility of the results. For instance, it can lead to increased data variability, making it more difficult to detect meaningful differences.

However, the most significant problem with small sample sizes is lower statistical power, making it more difficult to detect significant differences between groups or establish statistical significance. This is especially important when effect sizes are small, as in some four-day week pilots. It also limits the ability to control for confounding variables or include a bigger comparison group. For example, the Reykjavik trial only included one workspace as a comparison.

Smaller samples also threaten external validity, especially if the sample is not diverse and representative enough. If the sample does not include a diverse range of participants, it can limit the ability to generalise the findings to different groups within the business or employee population.

Most pilots have been transparent about the lack of representativeness of their sample. So far, all private-sector pilots have relied on self-selected volunteers to test the four-day week, and no pilot has yet aimed to use a representative sample of the business population. This is not necessarily a limitation on all fronts. It could still be valuable to run pilots with more motivated businesses and observe the impact of the four-day week in the best-case scenario. Although it would be essential to state and acknowledge that the pilot should not be the only information to use in an informed judgement on scaling up the four-day week, as its generalisability is very low.

The Perpetual Guardian pilot provided helpful information about the representativeness of the sample. They compared some indicators with the overall New Zealand business population and found that the participants were overall more satisfied with their job, more engaged, and more likely to stay in their job than the average New Zealand business before the pilot began. Additionally, the participants in the pilot were more inclined to embrace change, which indicates a greater willingness to take risks compared to the average employee in New Zealand. This difference in participants' characteristics is a noteworthy aspect to consider.

The 4DWG pilots have provided detailed information about participating companies, including their nationality and industry. However, as they come from different countries, their representativeness is more limited. The sampled businesses are mainly from the IT, telecoms, and professional services sectors, and the employees are not representative, with 71% holding bachelor's degrees. This is an important consideration, as more educated employees and innovative firms may be better positioned to adopt the four-day week, potentially increasing social inequality if the programme is scaled-up.

In Iceland, the first workplaces selected for the pilot were chosen based on their high levels of stress, making them unrepresentative of the overall public sector workplaces in the Icelandic public service. Additionally, the participants were volunteers, likely indicating higher motivation levels than average, which may not be replicated in more widespread changes to working arrangements.

Small sample sizes also limit the opportunity to conduct a robust analysis of the four-day week's heterogeneous effects, with pilots mainly relying on qualitative information to do so. For instance, in the Iceland pilots, managers felt they could not work shorter days than their staff, but the quantitative evidence did not support this claim due to the need for more statistical power to capture these effects.

Actual sample sizes are also affected by take-up, completion rates, and attrition levels, which pilots typically report. However, a detailed analysis and discussion of non-respondents and their implications

---

2018 (70%), the actual increase would be 11.4% (8 percentage points). The summary report of the Perpetual Guardian results can be accessed [here](#).



is often missing. As survey responses are often around 50% of the expected sample, sampling bias can occur, affecting the generalisability of the results even more. In the 4DWG pilots, 55% of all employees responded to the surveys, whilst Perpetual Guardian had a 50% response rate.

In essence, the sample size is crucial when designing a study, and the final sample size must be sufficient to provide reliable and meaningful results. It is also important to study companies that decided not to participate in the pilot or fail to provide company data, as well as those workers who dropped out and returned to a five-day week. These groups might be less inclined to respond naturally to the surveys.

Furthermore, other characteristics, such as intervention and pilot design, should also be considered when assessing the generalisability of results. Although they all fall under the "four-day week" pilots trend, there are significant differences between the interventions, such as the reduction in working hours in Iceland from approximately 40 to 37 hours, which is only a three-hour change starting from high working hours. The 4DWG intended to reduce it to four days or 32 hours, but in practice, workers reported working 34.8 hours.

When discussing working reduction programmes, it is essential to keep in mind that the effect of a four-day week with a day off may not be the same as working 32 hours. Many pilots do not differentiate between the two and group them together. Additionally, the starting point and the number of hours worked are critical factors, as the impact of each hour worked may differ. For example, the marginal effects of reducing the workweek from 48 to 40 hours in Japan may be different to the reduction from 40 to 35 hours in France.

To ensure pilot design is effective, it is recommended to follow the typology outlined in this working paper based on the nature of the organisation. This approach can help avoid making the assumption that public sector pilot results can be automatically applied to the private sector. Private companies have different dynamics, and the importance of productivity may significantly affect the effectiveness of the four-day week, which may not be captured in a public sector pilot. Even within the private sector, it is unlikely that results from one sector or company size can be generalised to other firms, given their diverse operating environments, organisational structures, and resource constraints.

The information on the costs linked with the adoption of the four-day week is highly variable among pilots as well and has not been adequately discussed. Most pilots have provided training and mentoring for participating businesses for several months before introducing the four-day week, but they have not provided information on the economic costs associated with the training or how much the business is covering in the process of adopting the four-day week.

From a policy perspective, this limits the possibility of conducting cost-effectiveness analyses and comparing the cost of implementing the four-day week against other flexible work arrangement policies. Whilst some pilots acknowledge some associated costs for businesses, which need to be covered by the participating companies, in other pilots, such as the future government-led Spanish pilot, 70 businesses would receive 150.000 euros each to cover the costs of adopting the four-day week (Europa Press, 2022). Also, when governments provide funding to businesses, the statistical analysis may be compromised as it may be difficult to discern the impact of this funding from the impact of implementing a four-day workweek.

The lack of information on costs and the failure to compare with relevant policy alternatives also limit the analysis of the pilots' scalability. Knowing the cost of an intervention is essential for policymakers to determine its affordability and sustainability. Without accurate cost information, policymakers may not be able to make informed decisions about the intervention's value for money. For example, suppose the intervention requires extensive training and months of changing the organisational structure. In that case, it can be challenging to implement on a larger scale, limiting the number of people who can effectively implement it. This can lead to higher implementation costs and may make it difficult to reach larger audiences successfully.

## **6 Suggestions to improve upcoming four-day week pilots**

This final section presents ten recommendations for future researchers and policymakers to improve the quantitative analysis of four-day week pilots. The recommendations are based on the limitations identified in the previous section, and it aims to address the main methodological challenges associated with the quantitative evaluation of these programmes.

The purpose of these recommendations is to provide a general guideline that can be used to design and evaluate future four-day week pilots more effectively. By following these recommendations, researchers and policymakers can ensure that future pilots provide more useful information, which can help to extend the evidence base and enable better decision-making regarding the impact of four-day week policies.

However, it would be unlikely and inadequate to rely solely on a single pilot study to acquire the necessary information to make informed decisions about the potential impact of the four-day workweek. It would still be necessary to conduct multiple pilots, each focusing on distinct aspects in different contexts. These pilots should aim to refine the intervention over time and identify areas of potential improvement. By conducting multiple pilots, policymakers can gather a more comprehensive understanding of the intervention and make informed decisions about its scalability and how best to maximise its effectiveness and impact.

### **6.1 Plan a robust evaluation method**

When evaluating a programme's effectiveness, relying solely on before-and-after comparisons can weaken the validity of results and reduce their credibility. Such comparisons have limited internal validity and can be easily biased by other factors not captured in the analysis. To avoid this, it is important to use a more robust counterfactual that can effectively address causality by comparing the programme group with a similar group that did not participate in the programme.

It is crucial to plan the evaluation method during the design stage to avoid problems of omitted variable bias. This is best achieved by allowing enough time to prepare, ideally through a trial protocol that researchers and implementers can follow. Such a protocol can help to establish clear criteria for the selection of participants and data collection, reducing the likelihood of bias and ensuring that the data collected is of high quality.

### **6.2 Select a valid comparison group**

When selecting comparison groups for a pilot programme, rolling basis selection methods that involve volunteering workspaces or businesses based on performance or other selection criteria may lead to biased results. Simply comparing an enrolled group with a non-enrolled group of self-selected participants is not an effective way to measure the impact of a four-day week, as other factors may influence the results. Randomising the selection of workplaces or businesses at the outset of the programme is the most straightforward and reliable method to avoid selection bias. This ensures that the groups have similar baseline characteristics, and any differences observed in the outcomes are due to the four-day week.

If randomisation is not possible, a quasi-experimental strategy should be implemented. Ideally, it should account for unobservable characteristics and include enough information to validate the assumptions. It is crucial to carefully consider the selection of comparison groups in the design stage to ensure that the chosen method is robust and reliable.

### **6.3 Provide transparent access to data and information**

To ensure the credibility of four-day week pilot studies, it is crucial to ensure that all relevant information is made publicly available. This should include extensive reports providing a more granular data analysis and interpretation of the results, as well as reporting all outcomes and conditions of the programme implementation. The use of summaries and policy briefs in more digestible formats should not compromise the quality of the information. These should especially avoid any changes in estimations and cherry-picking results.

Furthermore, providing open access to the data and results of the study would allow other researchers to replicate the study, thereby enhancing the credibility and robustness of the research. Replication is a critical aspect of scientific research, as it allows other researchers to validate or refute the findings, contributing to a more comprehensive understanding of the effects of the four-day week.

### **6.4 Include all the relevant statistical information**

In order to enhance the transparency of four-day week pilot studies, it is essential to comply with academic reporting standards and conventions. This means that statistical significance and confidence intervals of the estimation must be included to determine whether the results are due to random sampling error or a real effect. Furthermore, the sample's descriptive statistics, such as sample size, mean, and standard deviation, should be reported on all outcomes to provide context for the results. A balance check of the groups should also be conducted to ensure that the groups being compared are similar, and any differences should be acknowledged and addressed.

Moreover, it is necessary to provide a clear and thorough explanation of the evaluation method and estimations used to analyse the data. This will aid other researchers in interpreting the results and replicating the study, which is essential for validating the findings. Additionally, it is important to discuss the limitations and uncertainties of the research, including any potential sources of bias, to enable the reader to draw their own conclusions about the reliability of the findings.

### **6.5 Pre-commit to a set of outcomes and report them**

To ensure the reliability and credibility of four-day week pilot studies, researchers must pre-register the research design and specific outcomes that will be measured. Outcomes should be timely, relevant, clear and unambiguous. This guarantees that the outcomes match the programme objectives, reducing the risk of cherry-picking results. The pre-analysis plan should be made publicly available, and all deviations from the plan should be stated and explained.

In addition, it is important to avoid a long list of outcomes as it increases the likelihood of finding positive results by chance due to the multiple comparison problems. The appropriate statistical strategies should be put in place to avoid confusion. Pre-registration of research design and specific outcomes can also increase the clarity of the research and make it easier to replicate and compare results across different pilot studies.

### **6.6 Use larger and more representative samples**

Small sample sizes can limit the internal and external validity of pilot studies, making it difficult to detect significant differences and establish statistical significance. Whilst self-selected samples can offer valuable insights into the impact of a four-day week on specific groups, businesses, and sectors, it is important to acknowledge their limitations regarding external validity. For greater credibility, pilots should attempt to use a sample that is comparable to the overall target population, and measures should be taken to increase take-up and decrease attrition levels.

To increase the generalisability of the studies, pilots should aim to include more representative groups in the long run. This will provide a more comprehensive understanding of the effects of a four-day week and improve the intervention's reliability before scaling up. By including a more diverse range

of participants, pilots can better identify the factors that impact the success of a four-day week and provide insights that can be applied to different workplaces and industries.

### **6.7 Find out more about non-compliers**

To ensure the accuracy of the results of four-day week pilot studies, it is crucial to address non-compliance. A thorough analysis of non-responders can help to avoid bias and provide a more comprehensive understanding of the heterogeneous effects of the four-day week. Factors such as take-up rate, completion rate, and attrition levels should be considered when analysing the data, as they can significantly affect the final results. Identifying the reasons for non-compliance can also shed light on potential barriers to the implementation of the four-day week and help inform strategies to increase participation in future studies.

It is also important to recognise that non-compliance can be multifaceted and may be influenced by various factors. Therefore, an in-depth analysis of non-compliance can provide insights into the complexities of implementing the four-day week in some contexts and help tailor strategies to address these heterogeneous factors. By doing so, future studies can increase the likelihood of achieving higher levels of compliance and obtaining more representative samples, ultimately improving the generalisability and applicability of the findings.

### **6.8 Use more administrative data**

Self-reported surveys are an important tool for collecting data on the impact of the four-day week, but they are subject to biases and errors. To gain a more comprehensive understanding of the impact on productivity, it is crucial to supplement these surveys with additional administrative data, such as company data on revenue, average hours worked, resignations, new hires, and sick and personal days taken. However, it is important to acknowledge that participating businesses may not collect this information or use very different productivity and performance metrics that are not always comparable. Therefore, project planners should encourage participating businesses to collect data in a similar fashion from the moment they decide to participate in the pilot.

Moreover, it is important to recognise that productivity data often takes some time to manifest, which means that the focus for project planners should be on developing medium-term indicators of productivity that can be tracked throughout the study period and even after the pilot has concluded. For instance, operational management practices can be identified within a specific time frame of the pilot. By tracking these indicators, a more comprehensive evaluation of the impact of the intervention on productivity can be achieved.

### **6.9 Run longer pilots or increase the time of analysis**

Short-term pilots can have limitations in capturing the full effects of management changes and business outcomes due to substantial seasonal variations. To overcome this limitation, extending the analysis period can help ensure that the impacts are long-lasting. Running more extended pilots would allow a more comprehensive analysis of heterogeneous effects, providing more detailed information on the impact of the four-day week on productivity, work performance and other key outcomes for different groups.

Furthermore, longer pilots would help mitigate the limitations of self-reported data, as they provide a more accurate understanding of the long-term effects of the four-day workweek. Analysing the company data in the medium and long term, even after the pilot is finalised, would provide even more insights into the effectiveness and sustainability of the intervention in key performance indicators.

### **6.10 Consider the scalability of the intervention**

It is essential to discuss the costs associated with implementing a four-day workweek, including the costs of training and mentoring businesses during the adoption process, as well as the extent to which

participating companies are required to bear these costs. Running a cost-benefit analysis would help policymakers to determine the intervention's affordability and sustainability in the long term and on a larger scale.

Policymakers need to consider the scalability of the intervention beyond the pilot phase. This involves providing a careful assessment of the characteristics of the target group and the intervention. More than just assessing whether the intervention works on average, policymakers should aim to understand for whom it works and under what conditions.

## 7 References

- Åkerstedt, T. et al. (2001), *A 6-Hour Working Day-Effects on Health and Well-Being*. Journal of Human Ergology, vol. 30, no. 1/2, pp. 197-202.
- Angrist, J. & Pischke J. (2010), *The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics*. Journal of Economic Perspectives, vol. 24, n.o 2, pp. 3-30. DOI.org <https://doi.org/10.1257/jep.24.2.3>.
- Barnes, C. (2019), *Guidelines for an Outcome-Based Trial – Raising Productivity and Engagement*. White Paper – The Four-day Week. Perpetual Guardian, New Zealand. Retrieved from <https://apo.org.au/node/220571>.
- Booth, R. (2019), *Wellcome Trust drops plans to trial four-day working week*. The Guardian. Retrieved from <https://www.theguardian.com/uk-news/2019/apr/12/wellcome-trust-drops-plans-to-trial-four-day-working-week>
- Calvert Jump, R. et al. (2022), *A Future Fit For Wales: The roadmap to a shorter working week*. Autonomy Research Ltd. Retrieved from <https://www.futuregenerations.wales/wp-content/uploads/2022/02/SWW-Full-report-English.pdf>
- Chappell, B. (2019), *4-day work week boosted workers' productivity by 40%, Microsoft Japan says*. Retrieved from <https://www.npr.org/2019/11/04/776163853/microsoft-japan-says-4-day-work-weekboosted-workers-productivity-by-40>
- Christensen, G. & Miguel E. (2018), *Transparency, Reproducibility, and the Credibility of Economics Research*. Journal of Economic Literature, vol. 56, n.o 3, pp. 920-80. [www.aeaweb.org](http://www.aeaweb.org), <https://doi.org/10.1257/jel.20171350>.
- Clark, P. (2022), *Get ready for the four-day working week*. Financial Times. Retrieved from <https://www.ft.com/content/c5d83853-682e-4076-81c1-813b246309f8>
- Cohen A. (2022), *Four-Day Week Pilot Findings: Successful for Most Firms, But Not All*. Bloomberg. Retrieved from <https://www.bloomberg.com/news/articles/2022-09-20/four-day-week-pilot-findings-successful-for-most-firms-but-not-all>
- Coleman, A. (2023), *Four-Day Workweeks May Be The Future For Startups, One Pilot Program Suggests*. Forbes. Retrieved from <https://www.forbes.com/sites/alisoncoleman/2023/02/15/four-day-workweeks-may-be-the-future-for-startups-one-pilot-program-suggests/?sh=6c47e55d4488>
- Cooban, A. (2022), *The world's biggest four-day work week pilot begins*. CNN Business. Retrieved from <https://edition.cnn.com/2022/06/06/business/four-day-week-trial-uk/index.html>
- De Spiegelaere, S. & Piasna, A. (2017), *The why and how of working time reduction*. European Trade Union Institute.
- Delaney, H. & Casey, C. (2021), *The promise of a four-day week? A critical appraisal of a management-led initiative*. Employee Relations: The International Journal, vol. 44, n.o 1, pp. 176-90. Emerald Insight, <https://doi.org/10.1108/ER-02-2021-0056>
- Estevão, Marcello & Sá F. (2008), *The 35-hour workweek in France: Straightjacket or welfare improvement?* Economic Policy, vol. 23, n.o 55, pp. 418-63. RePEc - Econpapers, [https://econpapers.repec.org/article/oupecpoli/v\\_3a23\\_3ay\\_3a2008\\_3ai\\_3a55\\_3ap\\_3a418-463..htm](https://econpapers.repec.org/article/oupecpoli/v_3a23_3ay_3a2008_3ai_3a55_3ap_3a418-463..htm)
- Europa Press (2022), *El Gobierno dará hasta 150.000 euros a las pymes que implanten la semana laboral de cuatro días*. Europa Press Economía Finanzas. Retrieved from

<https://www.europapress.es/economia/laboral-00346/noticia-gobierno-dara-150000-euros-pymes-incienc-semana-laboral-cuatro-dias-20221212130744.html>

Gertler, P. et al (2016), *Impact Evaluation in Practice, Second Edition*. Inter-American Development Bank and World Bank. <https://openknowledge.worldbank.org/handle/10986/25030> License: CC BY 3.0 IGO.

Gilmore, O. (2019), *A 4-day week for Ireland? A report on the feasibility of a four-day working week in Ireland*. Four Day Week Ireland. Retrieved from <http://fourdayweek.ie/wp-content/uploads/2019/09/4DayWeek.pdf>

Haar, J. (2018), *Overview of the Perpetual Guardian 4-day (paid 5) Work Trial*. Auckland University of Technology. Retrieved from [https://news.aut.ac.nz/\\_\\_data/assets/pdf\\_file/0019/188002/Final-Perpetual-Guardian-report\\_Professor-Jarrood-Haar\\_July-2018.pdf](https://news.aut.ac.nz/__data/assets/pdf_file/0019/188002/Final-Perpetual-Guardian-report_Professor-Jarrood-Haar_July-2018.pdf)

Haraldsson, G. & Kellam, J. (2021), *Going Public: Iceland's journey to a shorter working week*. Alda and Autonomy Research Ltd. Retrieved from [https://autonomy.work/wp-content/uploads/2021/06/ICELAND\\_4DW.pdf](https://autonomy.work/wp-content/uploads/2021/06/ICELAND_4DW.pdf)

Harrington C. (2022), *The 4-Day Week Is Flawed. Workers Still Want It*. Wired. Retrieved from <https://www.wired.com/story/four-day-week-burnout/>

Hill, A. et al. (2021), *Endogeneity: A Review and Agenda for the Methodology-Practice Divide Affecting Micro and Macro Research*. *Journal of Management* 47:1, 105-143.

Kassam, A. (2021), *Spain to launch trial of four-day working week*. The Guardian. <https://www.theguardian.com/world/2021/mar/15/spain-to-launch-trial-of-four-day-working-week>

Kelly, O. et al. (2022), *The Four Day Week: Assessing global trials of reduced work time with no reduction in pay: Evidence from Ireland*. University College Dublin, Press. Retrieved from <http://fourdayweek.ie/wp-content/uploads/2022/11/4DayWeekAssessingGlobalTrials.pdf>

Lepinteur, Anthony (2019), *The Shorter Workweek and Worker Wellbeing: Evidence from Portugal and France*. *Labour Economics*, vol. 58, pp. 204-20. ScienceDirect, <https://doi.org/10.1016/j.labeco.2018.05.010>.

Martens, J. & Pronina L. (2022), *Belgium becomes the latest country to adopt a 4-day workweek*. Fortune. Retrieved from <https://fortune.com/2022/02/15/belgium-4-day-week-employees-work-life-balance/>

Nätti, J. & Anttila T. (1999), *Experiments of Reduced Working Hours in Finnish Municipalities*. *Journal of Human Resource Costing & Accounting*, vol. 4, n.o 2, pp. 45-61. Emerald Insight, <https://doi.org/10.1108/eb029057>.

Neubert, S. et al. (2022), *Free Days for Future? Longitudinal Effects of Working Time Reductions on Individual Well-Being and Environmental Behaviour*. *Journal of Environmental Psychology*, vol. 82, p. 101849. ScienceDirect, <https://doi.org/10.1016/j.jenvp.2022.101849>.

OECD (2023), *Framework for the Evaluation of SME and Entrepreneurship Policies and Programmes 2023*. OECD Studies on SMEs and Entrepreneurship, OECD Publishing, Paris, <https://doi.org/10.1787/a4c818d1-en>.

Prunier-Poulmaire, S. & Gadbois, C. (2001), *The French 35-Hour Workweek: A Wide-ranging Social Change*. *Journal of Human Ergology*, vol. 30, no. 1/2, pp. 41-46.

Raposo, P. & Van Ours, J. (2010), *How a Reduction of Standard Working Hours Affects Employment Dynamics*. *De Economist* 158, 193-207 <https://doi.org/10.1007/s10645-010-9142-5>

Sanchez, R. (2017), *Does a Mandatory Reduction of Standard Working Hours Improve Employees' Health Status?* *Industrial Relations: A Journal of Economy and Society*, vol. 56, n.o 1, pp. 3-39. <https://doi.org/10.1111/irel.12163>.

Sanchez E. & Galindo C. (2022), *The four-day working week: 'They'd have to pay me double to go back to the old schedule'*. El Pais. Retrieved from <https://english.elpais.com/economy-and-business/2022-06-20/the-four-day-working-week-theyd-have-to-pay-me-double-to-go-back-to-the-old-schedule.html>

Schiller, H., et al. (2017), *The impact of reduced worktime on sleep and perceived stress – a group randomized intervention study using diary data*. Scandinavian Journal of Work, Environment & Health, vol. 43, n.o 2, 2017, pp. 109-16, <https://doi.org/10.5271/sjweh.3610>.

Schor, J. et al. (2022), *The Four Day Week: Assessing Global Trials of Reduced Work Time with No Reduction in Pay*. Four Day Week Global, Auckland, NZ. Retrieved from <https://static1.squarespace.com/static/60b956cbe7bf6f2efd86b04e/t/6387be703530a824fc3adf58/1669840498593/The+Four+Day+Week-+Assessing+Global+Trials+of+Reduced+Work+Time+with+No+Reduction+in+Pay+%E2%80%93+F+%E2%80%93+30112022.pdf>

Seifert H. and Trinczek R. (2000), *New approaches to working time policy in Germany: the 28,8 hour working week at Volkswagen Company*. WSI Discussion Papers 80, Düsseldorf, Wirtschaftsund Sozialwissenschaftliches Institut. Retrieved from [https://www.boeckler.de/pdf/p\\_wsi\\_diskp\\_080.pdf](https://www.boeckler.de/pdf/p_wsi_diskp_080.pdf)

Stronge W. & Harper A. (2019), *The Shorter Working Week: A Radical And Pragmatic Proposal*. Autonomy Research Ltd. Retrieved from <https://autonomy.work/wp-content/uploads/2019/03/Shorter-working-week-docV6.pdf>

Tzvetozar V. (2022), *Portugal proposes pilot for 4-day work week* <https://www.themayor.eu/en/a/view/portugal-proposes-pilot-for-4-day-work-week-11183>

Voglino, G. et al. (2022), *How the Reduction of Working Hours Could Influence Health Outcomes: A Systematic Review of Published Studies*. BMJ Open, vol. 12, n.o 4, p. e051131. PubMed, <https://doi.org/10.1136/bmjopen-2021-051131>.



## **GETTING IN TOUCH WITH THE EU**

### **In person**

All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at: [https://europa.eu/european-union/contact\\_en](https://europa.eu/european-union/contact_en)

### **On the phone or by email**

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),

- at the following standard number: +32 22999696, or

- by electronic mail via: [https://europa.eu/european-union/contact\\_en](https://europa.eu/european-union/contact_en)

## **FINDING INFORMATION ABOUT THE EU**

### **Online**

Information about the European Union in all the official languages of the EU is available on the Europa website at: [https://europa.eu/european-union/index\\_en](https://europa.eu/european-union/index_en)

### **EU publications**

You can download or order free and priced EU publications from EU Bookshop at: <https://publications.europa.eu/en/publications>. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see [https://europa.eu/european-union/contact\\_en](https://europa.eu/european-union/contact_en)).

## The European Commission's science and knowledge service

Joint Research Centre

### JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.



**EU Science Hub**

[ec.europa.eu/jrc](https://ec.europa.eu/jrc)



@EU\_ScienceHub



EU Science Hub - Joint Research Centre



EU Science, Research and Innovation



EU Science Hub