The Employment and Windfall Effects of Short-Time Work: Evidence from Germany  

Christian Kagerl*  
Institute for Employment Research (IAB), FAU Erlangen-Nürnberg  
March 19, 2024

Abstract: I study the German short-time work (STW) scheme and its ramifications, using novel administrative data on STW and drawing on evidence from establishment surveys that are linked to the administrative data. I show that, next to financial reasons, firms value and use STW because it allows them to hoard labor in a tight labor market. During the pandemic, I document a strong negative selection into STW, which I do not observe for the financial crisis, based on measures of firm quality and productivity, a selection pattern that is explained by the differing types of crises. Adjusting for selection, I then investigate the employment effects of STW in the pandemic and find 3-4% higher employment levels for firms utilizing STW. This relationship, however, vanishes quickly after firms exit STW, a result driven by outflows among STW firms being initially lower, but being higher after the end of STW. Partly facilitated by eased access rules, I additionally find that the policy’s windfall effects are large: While back-of-the-envelope calculations suggest that up to half a million jobs were saved by STW in 2020, millions of jobs were supported in total, indicating an insufficient degree of targeting.

Keywords: short-time work, employment, establishments, social insurance

JEL: J20, J65, J68, E24

*Mail: christian.kagerl@iab.de. I thank Bernd Fitzenberger, Anna Houštecká and seminar participants at IAB for valuable comments.
1 Introduction

The pandemic has renewed interest in short-time work (STW). While the US chose to support workers that are laid off by their employers through (increased) unemployment benefits, leading to a sharp jump upwards in the unemployment rate to nearly 15% in early 2020, most European economies instead opted to insure the worker-firm-match explicitly. This insurance took place either through wage subsidies or, more frequently, through short-time work schemes (Giupponi et al., 2022), which allow flexible hours reductions for employees. OECD (2021) estimates that, on average across the OECD, about 20% of employment was supported by some kind of job retention scheme in April 2020. In Germany, Europe’s largest economy and the country whose case I study, roughly 6 out of 33 million eligible employees were in STW at the height of the pandemic’s first wave.

This unprecedented use of STW raises the economic questions as to how valuable these STW schemes are and what kind of effects they produce. Are they worth it, are many job matches (lastingly) saved? Are there unintended consequences of the policy, e.g. windfall gains for firms or reduced reallocation through propping up firms that would otherwise no longer be viable? Using newly available data, for the case of Germany, I study the reasons why firms go into STW, the selection process and estimate employment as well as windfall effects of STW.

The theoretical argument for STW schemes (see e.g. Burdett and Wright, 1989; Boeri and Cahuc, 2023) is that layoffs may be welfare inefficient during a shock or a crisis. Financial constraints due to decreasing revenue and imperfect financial markets might force firms to cut their workforce in the short run. However, in the medium run, after the temporary downturn of demand has reverted, the firm might need these employees again and would have to hire anew. The motive to decrease employment during a shock is further facilitated by the fact that flexibly cutting wages by large amounts is infeasible due to the existence of downward wage rigidity (e.g. Schoefer, 2021). The resulting layoffs incur social costs, for instance arising from unemployment benefits and from a search-and-matching-process with frictions. Under these circumstances, STW is welfare efficient if the costs of the program are lower than the social costs incurred in the absence of STW. Additionally, STW saves jobs, a result that can be described as an employment or insurance effect. This effect constitutes the positive effect of STW.

Simultaneously, the employment effect could be counteracted by the unintended consequences of the policy. First, STW benefits could be going to job matches that would not be lost if there was no STW available, an implication Cahuc et al. (2021) term “windfall effect”. Windfall can consist of outright fraud (receiving STW benefits despite no hours reductions) or, more likely, of a firm’s employees being partially in STW despite there being no acute enough financial
constraints that would lead to layoffs in the scheme’s absence. Another unintended consequence of STW would be a reduction in the labor market’s allocative efficiency coming about because the STW scheme might subsidize low-productivity matches and thus keep firms in business that only survive because of support (‘zombie firms’). From the welfare perspective, however, it would be beneficial to let these firms shut down and have their workforce be more productive in higher-quality firms. Shocks and recessions can have a cleansing effect on the economy (Foster et al., 2016). Ex ante, the magnitude of these effects is unclear and depends on the concrete implementation of the policy.

My analysis relies on administrative data – including novel data on STW – covering the universe of German establishments, further enriched by evidence from establishment surveys that are linked to the administrative data. I have four main findings. First, based on survey information, I show that the financial motive to choose STW (prominent in the theory of STW) is important among establishments, but that the ability of STW to flexibly keep up work processes under work loss and, especially, the motive of labor hoarding are also seen as significant reasons.

Second, based on linked data, I document adverse selection into STW. Firms with lower labor productivity and lower wage premia are significantly more likely to be in STW during the pandemic, even when adjusting for workforce composition, firm size, narrow 5-digit industries and regional differences. Interestingly, this adverse selection is not apparent for the last recession in which STW was used extensively, the financial crisis. The differing nature of the two crises explains this fact. Whereas the financial crisis disproportionately affected big exporting firms, the impact of Covid-19 was broad. With low-quality firms being in a relatively worse state before, the pandemic led to decreasing demand, temporarily shutting down and liquidity shortages, driving the selection into STW.

Third, to identify the repercussions of STW on firms’ employment, I combine (coarsened) exact matching with an event study approach, leveraging the administrative data. I first make STW firms and non-STW firms comparable with respect to a broad range of observable characteristics (including 5-digit industries) and explicitly factor in the negative selection on productivity in the matching step. Controlling for firm and time fixed effects, my event-study approach therefore compares employment in STW firms to those firms without STW, where I also see no differential trends between the two groups prior to the pandemic. I find positive employment effects of STW on the order of 3-4% in the year in which a firm utilizes the policy’s benefits. However, these effects disappear quickly after STW is exited, a result that can ascribed to STW firms having lower outflows during their STW spell, but then having higher outflows (relative to
non-STW firms) afterwards\(^1\). I run a number of econometric specification checks (e.g., inverse probability weighting) and tests (e.g., ruling out establishment death as a relevant margin) to underscore the robustness of my results.

Fourth, I provide evidence of substantial windfall effects. Back-of-the-envelope calculations based on the estimated employment coefficients imply that up to 500,000 jobs were saved by STW in 2020. Yet, 8.75 million jobs (corresponding to 3.7 million employment equivalents) were supported in total, a discrepancy that suggests large windfall effects. These could be lessened by better targeting of shock-affected firms and by introducing an experience rating element (Burdett and Wright, 1989), which entails STW users partially paying back STW costs through increased social security contributions in the future. Corroborating evidence on windfall effects comes from the fact that some establishments collected STW allowances despite reporting no negative pandemic effects or revenue decreases in surveys.

I contribute to several strands of the literature. There exists a growing literature that uses (structural) models grounded in theory to investigate the efficiency of STW and explore counterfactuals without STW or with different rules of the policy. These models often use German data for calibration since the scheme in Germany is available outside of recessions as well (e.g. Peltonen, 2023a; Tilly and Niedermayer, 2017). Cooper et al. (2017), using a search model, find that STW saved jobs in 2009/2010 in Germany, but reduced the allocative efficiency with subsequent output losses. Diaz et al. (2023) report similar results for the case of Spain. Hallmann (2023), again having Germany as the case study, shows that employees at the career peak gain most from being saved by STW, but cautions that such a scheme may be ineffective in recessions caused by structural change. Related, Brey and Hertweck (2020) as well as Gehrke and Hochmuth (2021) suggest that STW is job-saving during strong downturns, but ineffective in normal times. In a general equilibrium setting, Peltonen (2023b) analyzes Germany’s STW during the pandemic, also finding a strong stabilizing effect on unemployment rates. Albertini et al. (2022) show the same for France during the pandemic, but further suggest the existence of strong windfall effects. Balleer et al. (2016) hint at the importance of institutional factors with STW being more effective in countries with less flexible labor markets (i.e. higher firing costs). The positive cross-country correlation between STW take-up and firing costs is empirically corroborated by Cahuc and Carcillo (2011) and Lydon et al. (2019).

Regarding reduced-form evidence (thus also regarding the assumptions that underpin models), the literature is relatively more sparse. Using cross-country regressions for the Great Recession Boeri and Bruecker (2011), Cahuc and Carcillo (2011) and Hijzen and Martin (2013)

\(^1\)In the future, I plan to disentangle whether this is driven by voluntary or involuntary outflows, a distinction with importance for the interpretation.
do not reach a common conclusion on the direction of the employment effects. For Germany, existing micro-level evidence on the firm-level by Speckesser (2010), Bellmann and Gerner (2011), Scholz (2012) and Kruppe and Scholz (2014) does not paint a clear picture of positive employment effects either. I contribute in this case by, first, utilizing newly available and high-quality administrative data on STW that can be linked to other administrative data and survey data – instead of using only limited survey data, which moreover is not always accurate in the context of STW (Kagerl et al., 2022). My second contribution is showing clear evidence of positive employment effects most of which dissipate rapidly. Third, I provide evidence on the extent of windfall effects for the German STW scheme.

More recent studies of STW, all focusing on the financial crisis period, have sought to address the causal identification challenge of estimating employment effects on the firm-level as firms self-select into the programs. Giupponi and Landais (2023) exploit eligibility rules in the Italian scheme to estimate sizeable positive employment effects which vanish quickly after treatment ends, while cautioning that persistent crises can induce negative reallocation effects. For Switzerland, Kopp and Siegenthaler (2021) show positive employment effects based on comparing approved and denied applications for the support scheme. Cahuc et al. (2021) utilize regional approval rate differences to arrive at positive employment effects for the French STW program that are concentrated at firms experiencing big revenue drops, but also establish a substantial degree of windfall effects. My contribution to this literature is twofold: First, I analyze STW in Germany, which has different institutional rules around STW. Second, I do not look at the Great Recession, but consider the Covid-19 pandemic, where STW was used much more extensively and which was a completely different type of shock (e.g. in terms of the shock intensity or the sectoral breadth of the shock). In addition, I show that the STW selection patterns between the two crises differ markedly.

I further contribute by showing which establishments select into STW and what firms find valuable about the policy instrument. German firms’ motives for using STW are not restricted to the improved liquidity such schemes can provide, a main argument in the theoretical literature. The ability to keep up processes under work loss and the motive of labor hoarding (retaining employees that would be hard to replace later) are rated as just as important and as even more important, respectively. On the matter of selection into STW, I document strong adverse selection of establishments based on firm quality and labor productivity in the pandemic; a pattern of selection that is not apparent for the Great Recession, the prior recession with extensive use

---

2Bermudez and Cockx (2023) and Benkovskis et al. (2023) are two further examples documenting positive employment effects for Belgium and Latvia, respectively. Aiyar and Dao (2021) use sector-state-level data from Germany to suggest STW being preventive of unemployment increases. Kato and Kodama (2019) even document a positive impact of the Japanese STW scheme on later firm performance measures like return on assets.
This paper is organized as follows: Section 2 describes the institutional background of STW in Germany. The establishment-level data on which my analyses build is described in Section 3. Section 4 examines what establishments find attractive about STW and which establishments select into STW and if lower-quality and less productive establishments are more likely to receive benefits. Subsequently, Section 5 investigates the employment effects of STW in the pandemic. Section 6 analyzes the extent of windfall effects. Finally, Section 7 concludes by summarizing the findings and discussing their (policy) implications.

2 Institutional Background

Germany’s STW scheme nominally has three components: Benefits for companies that are restructuring (Transferkurzarbeitergeld), seasonal STW and konjunkturelles Kurzarbeitergeld, which is intended for cushioning (business cycle) shocks. The last part dwarfs the other two in size as well as importance, especially in crises, and is thus the sole focus of this study. Establishments can apply for STW through a form (Anzeige) at the Federal Employment Agency (FEA) when they are expecting work loss, but this initial application is flexible. For instance, while only workers listed in this application can receive the benefits, an establishment has an incentive to list all of its workers there, even if not all of them will need it, just to be on the safe side. After the initial notice is granted and there was a work loss in a given month, the establishment has three months to submit a second notification (Abrechnung) that details and tabulates the work loss of its employees in that month. The establishment has to give a reason for the work loss and, after a check, is reimbursed by the FEA according to the details of the second form. The data I use comes from the granted second notification, i.e. it captures the actual work loss of the establishment for which it was reimbursed.

Crucially, only the workforce in employment subject to social security contributions (‘regular employment’ or ‘contributory employment’) can receive STW benefits, while ‘marginal employment’ may not. Generally, a third of an establishments contributory workforce needs to have a reduction in working hours to be eligible for a payout; this constraint was eased to 10% (with at least 10% work loss) at the beginning of the pandemic. A firm’s STW may last for 12 months at a time (interruptions of fewer than three months are not considered for this calculation), but this threshold was extended to up to 28 months during Covid (until June 2022). Section 4 provides statistics on the extent of how STW is used. Further eased access rules during the pandemic specified that apprentices and temp agency workers were also eligible for STW and there were fewer restrictions on taking other jobs while on STW. For employees, being in STW implies that
60% of the lost hours’ wages are covered by the FEA (also referred to as the replacement rate of STW), 67% for people with children in the household. During the pandemic up until June 2022, this rate increased to 70/77% (80/87%) if an employee was in STW for 4 (7) months and the work loss was at least 50%. Generally, the share of hours lost is flexible and may be up to 100%, i.e. effectively not working. In addition, the FEA covered the social security contributions of workers in STW until December 2021 (and 50% coverage in the first three months of 2022), a part of the wage costs not usually covered. Taken together, this implementation of the German scheme unburdens employers as it requires very little (in 2020 and 2021 essentially zero) co-financing of the lost hours’ wage costs; a fact where some other countries’ STW schemes differ (OECD, 2021). Furthermore, Germany has no experience rating element built into its version of STW.

3 Data Sources

Until the pandemic, there was no administrative micro-level data available on the usage of short-time work in Germany. Previous studies (e.g. Kruppe and Scholz, 2014) relied on small samples from surveys of establishments or individuals and on a sample of granted notifications (Abrechnungen) collected in the district around Nuremberg. This study uses newly available data from the Federal Employment Agency on the take-up of STW among all establishments in Germany. More specifically, the data on STW (a data set named BTR-KuG) contain which establishment was in STW in a given month, how many of its employees were in STW, the preceding number split by the gender of workers and the work loss/employment equivalent in STW in the establishment. Establishment-level data are available from January 2009 onwards, currently until the end of 2022.

The big advantage of this data set is that the information on STW can be linked to other data sets on the establishment-level from the Federal Employment Agency that are collected at the Institute for Employment Research: First, to the longitudinal administrative data on establishments – the Establishment History Panel (Ganzer et al., 2023) – which, in turn, emanates from aggregating employment biographies of individual workers. From the Establishment History Panel, I obtain a plethora of information on the universe of establishments and their complete histories, including the size and constitution of their workforce (e.g., information on occupations, gender, skill, age etc.), their employment flows, their wage structure, their 5-digit-

---

3This latter measure essentially captures the employees’ intensive margin of STW as well. For example, if an establishment has 100 full-time employees and in one month 50 of these workers are in STW for 50% of their respective contractual working hours, then the employment equivalent in STW (or total work loss) amounts to $50 \times 0.5 = 25\%$. 

industry code and their region. Moreover, I merge AKM fixed effects (Abowd et al., 1999) of establishments (Bellmann et al., 2020) to the data, representing establishment wage premia. From these data (STW data and Establishment History Panel with AKM-FE), I construct a data set containing all the administrative information on establishments on a yearly basis.

Second, I link this administrative data set to establishment surveys that draw their samples from the administrative establishment data. For more specific questions and for information not covered in the administrative data (e.g. measuring labor productivity), I link the data to the Establishment Panel, a yearly survey of roughly 16,000 establishments (for more details see Ellguth et al., 2014). Furthermore, I use data from the BeCovid-survey (for example, data on which aspects 2,000 establishments value about STW), which covered how establishments adapted to the economic shock of the pandemic (for more details see Bellmann et al., 2022).

One peculiarity of the German scheme concerns companies that consist of multiple establishments. Establishments, whom I refer to interchangeably as firms, are the focus of the scheme and also of the analysis. However, it is possible that, e.g., the headquarters of a company (which is an establishment in itself) reports STW for subordinate establishments within the company structure. Survey evidence suggests that this happens occasionally (Kagerl et al., 2022) and company networks (i.e., which establishments belong to the same company) are not directly evident from the administrative data. To address this idiosyncrasy, I present the robustness of my results to only including single-plant establishments, demonstrating that this peculiarity is not an issue. The data on which establishments are single-plant comes from a (so far, preliminary and incomplete) link of the Mannheim Enterprise Panel (Bersch et al., 2014) to the administrative data.

4 How is STW Used and Which Establishments Go into STW?

I begin by providing background numbers detailing the overall scope and patterns of STW since January 2009, the start date of the STW data. Figure 1 plots the shares of establishments, employees and employment equivalent in STW. At the height of the first lockdown in Germany in April 2020, roughly six million individuals were in STW, representing roughly 18% of the eligible workforce. Expressed in employment equivalents, which take into account the intensive margin of STW, the share of overall work that was lost/subsidized was 9% in April 2020.

Given that the decision to go into STW is made by the firm, it is also relevant to consider what share of establishments was in STW. Defining the population of establishments as all establishments that have at least one employee subject to social contributions (since only contributory employment can receive benefits), 34% of establishments had at least one month
Notes: The graph shows shares in STW by different categories. For establishments, the number of establishments with at least one employee subject to social security contributions is used for the calculation of the shares. In the case of employees, the number of workers subject to social security contributions is used. The share of the employment equivalent in STW is obtained by setting the total employment equivalent in STW in relation to the total number of employees.

of STW in 2020 (nearly 30% in April 2020) and 22% in 2021. For comparison, the share was merely 0.4% in 2019 and well below one percent for every year since 2012. The pandemic-era usage of STW additionally dwarfs the take-up in the financial crisis, which peaked at 5.2% in 2009. One noticeable pattern in Figure 1 when comparing the two crises is that the share of employees in STW was higher than the share of establishments in STW in 2009, but the reverse always holds during the pandemic.

On the level of the establishment, both crises exhibit a high degree of persistence, i.e. firms in STW use it for a relatively long time. Figure 2 plots histograms of establishment-level measures of the intensive margin of STW, under the condition of utilizing STW at once, comparing the period January 2009 until April 2011 with the period March 2020 to June 2022. Panel (a) of Figure 2 considers the total number of months an establishment is in STW over the two respective 28-month-periods. The two distributions overall are quite similar to each other. On the far end of the distribution, the regulations explain the difference: While the maximum number of consecutive months of STW was 24 in the financial crisis, it was 28 in the pandemic.

As the population total is about 2.1 million in 2020, this implies that more than 700,000 establishments received STW benefits in 2020 and roughly 600,000 in April 2020. These numbers illustrate the strong persistence of establishments in STW. Moreover, of the about 100,000 establishments in STW at the end of 2021, more than 90% had seven or more months of STW usage since the start of the pandemic.
Panel (b) of Figure 2 shows histograms for the average share of employment equivalent in STW in an establishment across the entire period – including months without STW as zeros in the computation. It is evident that the intensity of work loss within an establishment was meaningfully bigger in the pandemic compared to the financial crisis.

Figure 2: Establishment-Level Intensive Margin of STW

Notes: Panel (a): The graph shows histograms for the number of STW months in an establishment for 28-month-periods in the two crises (only including establishments that had at least one STW spell in the respective time period). Each bar represents one month; the population of STW-using establishments is used. Panel (b): The graph shows histograms for average share of employment equivalent in STW in an establishment for 28-month-periods in the two crises (only including establishments that had at least one STW spell in the respective time period). Each bar has a width of five percentage points; the population of STW-using establishments is used.

However, the differences in STW usage between the two crises are not surprising as the two recessions constitute quite different types of negative shocks. The pandemic shock was also a health shock that, together with the containment measures, disrupted how work was done, thus particularly affecting employment. Correspondingly, the usage of STW was lower during the financial crisis, as was the average work loss when in STW. Third, the crisis in the late aughts was concentrated in manufacturing while the pandemic hit the economy much more broadly, severely affecting the service sector as well. This last observation helps to explain the share of establishments in STW always being larger than the share of employees in 2020 and 2021, since the average establishment size in manufacturing is much larger than in the service sector.

4.1 Why Establishments Take up STW

The literature reviewed in Section 1 portends some reasons why establishments would find it attractive to use STW. For one, wage costs can be greatly reduced with STW when an economic shock hits, easing financial constraints. Not having to lay off workers keeps the workforce in place and means avoiding a potentially costly and protracted process of rehiring employment.
afterwards. Depending on the STW scheme, firms might also value the option of reducing hours flexibly for each individual employee (as, e.g., in the German case), leading to less disruption of work processes that are still running.

Yet, there has been no evidence on what firms themselves find valuable about the instrument of STW. To test which of these reasons are most relevant to establishments when thinking about STW, I leverage data from the twelfth wave of the BeCovid establishment survey (from April 2021), which presented establishments with four statements regarding why STW is a good instrument and to what degree establishments agreed with those statements (Kuhn et al., 2021). Three statements relate directly to the literature: keeping up work processes under work loss, keeping employees that would be hard to replace later and improving liquidity through reduced wage costs. The fourth statement considers an equality argument within the firm, namely that STW can help split losses evenly among employees. Figure 3 depicts the shares of establishments that agree with each of the four statements about STW.

Figure 3: What Firms Value about STW

![Figure 3: What Firms Value about STW](image)

Notes: The figure shows the share of establishments that agree with each of the four statements about why STW is a good policy instrument plus their respective 95% confidence intervals. Establishments could respond on a scale of one to five, where one corresponds to ‘do not agree at all’ and five to ‘agree completely’. Depicted are the shares of establishments that responded with four or five. The data are from the twelfth wave of the BeCovid-survey which took place in April 2021. Survey weights are applied to make the sample representative of the population of establishments with at least one employee subject to social security contributions. N = 2,000 establishments.

5In Kuhn et al. (2021), some of the results in this subsection 4.1 have been reported in German for a policy audience.

6Establishments could respond on a scale of one to five, where one corresponds to ‘do not agree at all’ and five to ‘agree completely’. I measure ‘agreeing’ as responding with four or five.
Figure 3 attests that, in general, establishments respond positively to the statements about the usefulness of STW. The statement on equally splitting losses among employees receives the lowest support, with just under half of all establishments agreeing. Agreement rates for the three other statements, however, are between 75% and 85%.

An interesting finding is that the canonical main argument that STW eases liquidity constraints only comes in third with 75% agreeing. A share that is significantly lower than the 85% agreeing that STW allows firms to keep employees that would be hard to replace later. Hence, it seems that the motive of labor hoarding plays a bigger role for firms when thinking about the benefits of STW than does improved liquidity. This could reflect that Germany’s labor market has been tight for some time and filling vacancies poses a considerable challenge for many firms (Bossler and Popp, 2024).

The flexibility that the German STW scheme offers for adjusting work processes is also prized highly, by four out of five establishments. In firms with a high level of coordination in the production process, shocks can present a bigger challenge when workers are missing, a setting where the flexibility of STW is valuable (Kuhn et al., 2023).

4.2 Selection Into Short-Time Work

The crucial question regarding selection into STW is, however, whether the quality of establishments before the pandemic recession is a strong predictor of STW take-up during it. This requires eliminating observable differences between establishments that are correlated with establishment quality or productivity, most notably e.g. the industry in which a firm is active, how big it is and in which region it is located. To more carefully assess the issue of possible adverse selection into STW, I therefore estimate models based on the following selection equation:

\[ STW_e = \alpha + \beta \text{Quality}_e + \sum_{occ=1}^{Occ} \gamma_o \text{OccShare}_{eo} + \delta_i + \psi_s + \phi_r + \varepsilon_e \]  

(1)

Establishments are denoted by the subscript \( e \). \( STW \) is a measure for the usage of STW by the establishment. On the extensive margin, for instance a dummy taking the value one if the establishment has any STW spell in a given year. On the intensive margin, for example an establishment’s share of employees in STW averaged across a given year. Shares of occupations within an establishment are denoted \( o \) and consist of twelve occupational categories based on the Blossfeld classification (for details see Ganzer et al., 2023). \( \delta_i \) are 5-digit industry fixed effects (838 categories), \( \psi_s \) denote size class fixed effects\(^7\) and \( \phi_r \) are regional fixed effects on the level of the 401 German counties. \( \varepsilon \) is an error term. The coefficient of interest here is \( \beta \), indicating

---

\(^7\)Size classes are: 1 to 4 employees, 5 to 9, 10 to 49, 50 to 99, 100 to 499, 500+
the relationship between establishment quality and STW take-up in a given year, conditional on
the set of fixed effects and the occupational structure. I use two quality indicators, both being
computed from prior years’ data to avoid contamination: First, the AKM-FE of an establishment
– available for a majority (≈60%) of all establishments, but for more than 90% of establishments
with 50 or more employees. Second, log revenue per worker (‘labor productivity’) – available for
establishments answering the revenue question in the establishment panel. To do so, I construct
quintiles of labor productivity for establishments surveyed in the Establishment Panel. I focus
on establishments that consist of a single plant and are not part of a multi-establishment firm,
due to two reasons. First, different establishments in a network that are part of the same firm
can in theory report STW for each other and Kagerl et al. (2022) find that some headquarters
report STW for establishments that are subordinate to them. Second, and more importantly,
the revenue amount surveyed might not reflect the establishment’s revenue for multi-plant firms
if it is not determined in the firms’ accounting procedures. Thus, the restriction to single-
plant establishments insures that the calculated revenue per worker is truly specific to the
establishment in question. Moreover, all regressions with labor productivity use survey weights
to be representative of the population of establishments with at least one employee subject to
social security contributions.

Figure 4 shows results from equation (1) for the extensive margin. Coefficients from 20 five-
percent-intervals (ventiles) of the AKM-FE distribution and an indicator for a missing AKM-FE
are depicted in Panel (a), serving as the first measure of establishment quality which uses the
universe of establishments. The eleventh ventile of the distribution (51%-55%) is the reference
category (also shown). In Panel (b), quintiles of log labor productivity as computed from the
Establishment Panel proxy firm quality.

Figure 4 reveals a strong adverse selection based on the AKM-FE measure: Establishments
in the lower part of the distribution are considerably more likely to have experienced STW than
those in the upper part. Comparing the highest to the lowest decile of the AKM-FE distribution
when adjusting for occupational structure and for the set of fixed effects, establishments in the
uppermost decile are roughly ten percentage points less likely to be in STW than in the lowest
decile. Relative to the baseline share of about one third in 2020, this is a sizeable difference.
Establishments with missing AKM-FE are about three percentage points less likely to be in
STW than the middle of the distribution. Figure A1 in the Appendix shows that this result of

---

8Establishments without an AKM-FE are coded as having a missing AKM-FE and are included as a separate category in (1). Essentially, there are two big reasons why the AKM-FE might be missing. First, as the AKM-FE is computed over the time period 2010 to 2017, younger establishments have not existed yet and hence have no wage premium. Second, establishments are not part of the connected set because of limited mobility of employees that is the requisite for estimating AKM-FE. This disproportionately applies to very small establishments which make up the overwhelming mass of all establishments, explaining the 60% number above.
Figure 4: Selection: Firm Quality and the Extensive Margin of STW in 2020

Notes: The figure shows coefficients for establishment quality based on specification (1) with the full set of fixed effects and occupational controls: panel (a) for the year 2020 using the position in the distribution of the AKM-FE which were calculated from the period 2010 to 2017, panel (b) using the position in the distribution of labor productivity which was calculated for 2019. The dependent variable is a dummy taking the value one if an establishment has at least one STW spell in the given year, and zero otherwise. The reference category is the eleventh of the twenty ventile, i.e. between the 51st and the 55th percentile for panel (a). The reference category in panel (b) is the first quintile and survey weights are applied. $N_{AKM}^{2020} = 2,097,889$ establishments and $N_{LP}^{2020} = 6,913$ establishments.

Selection in the pandemic also holds when additionally adjusting for the ventiles of the median wage distribution across establishments in specification (1). Hence, the observed patterns are not driven by differing wage levels, but by the wage premia. The other panel of Appendix Figure A1 further shows the results’ robustness when only considering single-plant establishment.

Results for the intensive margin of STW are shown in Figure A2 in the Appendix, only considering establishments with at least one STW month in the respective year. The measure I use is the average work loss in STW across the entire year (months without STW thus have value zero). For the intensive margin, 2020 again exhibits the strong pattern of adverse selection.

Panel (b) of Figure 4 affirms the selection result when using the labor productivity measure. Compared to the first quintile, establishments in the fifth quintile are ten percentage points less likely to have an STW spell in 2020. Because of the much-reduced sample size and the application of survey weights, the estimates are less precise, but the just-mentioned point estimate is very similar to the interpretation of Panel (a), see above. Yet, the two measures are not fully congruent with a correlation coefficient of 0.45 between the position in the AKM distribution and the position in the labor productivity distribution. The partial correlation coefficient between the two measures is 0.2, when holding the other factors from (1) constant. This provides evidence that, while being somewhat different, both measures capture underlying establishment quality.
So far, the selection analysis has been confined to the initial pandemic recession in 2020. The German STW scheme, however, is always active and be utilized by firms, even in boom times during the business cycle. Therefore, I now consider all years with STW data availability, which comprise the time period 2009 through 2022. Figure 5 plots the results for both proxies of firm quality. To summarize the establishment quality measures in one coefficient for every year, I show the difference between the fifth and the first quintile for labor productivity. For the AKM fixed effects, I show the estimated value for increasing the AKM-FE by one standard deviation within a 5-digit industry. Both of these quality proxies are always computed in the year prior, except for crisis years: 2019 is also used for 2021/2022 and 2008 is also used for 2010/2011.

Figure 5: Selection Over Time: The Extensive Margin of STW

Notes: The figure shows coefficients for establishment quality based on specification (1) with the full set of fixed effects and occupational controls. The dependent variable is a dummy taking the value one if an establishment has at least one STW spell in the given year, and zero otherwise. In panel (a) the coefficient values represent increasing the SD of the AKM-FE (within a 5-digit-industry) by one. Panel (b) shows the difference obtained between the fifth and first quintile of labor productivity while applying survey weights. Both of these quality proxies are computed in the year prior, except 2021 and 2022 where 2019 is used and 2010 and 2011 where 2008 is used.

There are two key takeaways from Figure 5. First, that in the financial crisis in 2009 and 2010 – in contrast to the pandemic recession – there is no apparent pattern of adverse selection into STW; the AKM measure in Panel (a) even suggests a slight positive selection. However, the labor productivity measure does not confirm this pattern, rather it seems that there is no selection on establishment productivity in the Great Recession. Second, there is no pronounced positive or negative selection on establishment quality into STW in the intervening non-crisis years. Figure A3 in the Appendix documents the robustness of this pattern: First, to including median wages as a control, and, second, to only looking at establishments in manufacturing (since manufacturing constituted the vast majority of STW establishments in the financial crisis). Figure A4 in the Appendix shows, that, at the intensive margin, there also only is adverse
selection in the years of the pandemic. Additionally, Figure 5 reveals that the degree of selection was strongest in 2020 and decreases in absolute terms the two years thereafter. The labor productivity measure suggests further that 2022 already returns to the pattern seen before the pandemic.

Reviewing the results on selection into STW, it appears that a negative selection on establishment quality into STW – when conditioning on sector, size and other factors – is evident for the pandemic, but not for the last big recession, the financial crisis. Such a pattern hints at the fact that the two crises affected firms along the quality distribution differently. I affirm this by looking at the relationship between labor productivity and reported crisis effects in the Establishment Panel, drawing on similar questions in 2010 and 2020 on whether the respective crisis negatively affected establishments. Figure 6 plots differences in the fractions of establishments self-reporting negative crisis effects, relative to the first labor productivity quintile. The raw differences show the unadjusted differences, only applying survey weights, while adjusted differences are obtained from running the selection equation (1) and taking a dummy of reported negative crisis ramifications as the dependent variable. The baseline fraction of negatively affected establishments amounts to 0.35 in the financial crisis and to 0.74 in the Covid-19 pandemic’s initial phase. Reporting negative crisis effects is, naturally, a very strong predictor of STW take-up, even conditionally: Including the indicator in specification (1) with STW as the dependent variable yields that crisis-hit establishments are 10 percentage points more likely to be in STW in 2010 (overall STW share 7%) and 33 percentage points more likely in 2020 (overall STW share 40%).

Panel (a) of Figure 6 shows differences obtained for 2010, where establishments were surveyed about whether the financial crisis had any negative effects on them in 2009/2010. Raw differences show the upper two quintiles being about 8 percentage points more likely to report negative effects. Conditionally, however, there is no discernible relationship between labor productivity and reported crisis ramifications, just as the right-hand side panel of Figure 5 indicates for STW. One possible explanation for this result, even when only considering manufacturing, could be in the fact that the crisis strongly decreased global demand. To illustrate: German exports decreased by more than 18% in 2009 relative to the year before, but the corresponding decrease in 2020 amounted to only 9%. This means that firms with more export exposure suffered greater consequences and were more likely to go into STW. As exporters are typically positively selected on size and productivity (Melitz, 2003), this could constitute one explanation for the

---

9Revenue changes could be an alternative, less subjective, measure. However, available data only measure year-on-year changes, where it is unclear whether and how much of a potential decrease is due to the crisis itself or due to being in STW. Hence, I rely on establishments’ assessments of the crisis effects.
Notes: The figure shows coefficients and 95%-CIs for labor productivity on the probability of reporting negative crisis effects taken from the Establishment Panel; panel (a) for the financial crisis in 2010, panel (b) for the pandemic in 2020. The raw difference is the difference in means for the negative effects dummy relative to the first labor productivity quintile, the adjusted difference is based on specification (1) with the full set of fixed effects and occupational controls. The dependent variable is a dummy taking the value one if an establishment reports negative effects of a respective crisis, and zero otherwise. The reference category in both panels is the first quintile and survey weights are applied. The baseline fractions of reported negative crisis effects for the first quintile are 0.35 in 2010 and 0.74 in 2020. Labor productivity is measured before the respective crisis. \(N_{LP}^{2010} = 5,728\) establishments and \(N_{LP}^{2020} = 6,900\) establishments.

Panel (b) of Figure 6 shows differences obtained for 2020, where establishments were surveyed about whether the pandemic had any negative effects on them in the first half of 2020. There is a negative relationship between quality and reported crisis effects. For instance, the highest quintile is more than 10 percentage points less likely to be affected than the lowest, in raw as well as in adjusted differences. Therefore, even conditioning on narrow industries and other factors, lower-productivity establishments seem to be more adversely impacted by the pandemic.

The survey questions in 2020 allow for further analyzing some specific types of negative effects – in contrast to the financial crisis. Establishments reporting negative pandemic effects were further asked about several types of negative ramifications and whether the establishment experienced them. Specifically, those are decreasing demand, (temporarily) shutting down, supply chain problems, staff shortages and liquidity shortages. Table 1 investigates the relationship between STW, negative pandemic ramifications and labor productivity.

Panel A of Table 1 considers whether the reported pandemic effects can explain the adverse selection into STW. Column (1) reproduces the conditional difference between the fifth and first labor productivity quintile, shown for instance in the right panel of Figure 5. In column (2), the sample is restricted to the set of establishments for which data on all types of negative

<table>
<thead>
<tr>
<th>Quintile</th>
<th>Raw Difference</th>
<th>Adjusted Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>-0.25</td>
<td>-0.2</td>
</tr>
<tr>
<td>2nd</td>
<td>-0.2</td>
<td>-0.2</td>
</tr>
<tr>
<td>3rd</td>
<td>-0.15</td>
<td>-0.15</td>
</tr>
<tr>
<td>4th</td>
<td>-0.1</td>
<td>-0.1</td>
</tr>
<tr>
<td>5th</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

| Notes: | The figure shows coefficients and 95%-CIs for labor productivity on the probability of reporting negative crisis effects taken from the Establishment Panel; panel (a) for the financial crisis in 2010, panel (b) for the pandemic in 2020. The raw difference is the difference in means for the negative effects dummy relative to the first labor productivity quintile, the adjusted difference is based on specification (1) with the full set of fixed effects and occupational controls. The dependent variable is a dummy taking the value one if an establishment reports negative effects of a respective crisis, and zero otherwise. The reference category in both panels is the first quintile and survey weights are applied. The baseline fractions of reported negative crisis effects for the first quintile are 0.35 in 2010 and 0.74 in 2020. Labor productivity is measured before the respective crisis. \(N_{LP}^{2010} = 5,728\) establishments and \(N_{LP}^{2020} = 6,900\) establishments.

Panel (b) of Figure 6 shows differences obtained for 2020, where establishments were surveyed about whether the pandemic had any negative effects on them in the first half of 2020. There is a negative relationship between quality and reported crisis effects. For instance, the highest quintile is more than 10 percentage points less likely to be affected than the lowest, in raw as well as in adjusted differences. Therefore, even conditioning on narrow industries and other factors, lower-productivity establishments seem to be more adversely impacted by the pandemic.

The survey questions in 2020 allow for further analyzing some specific types of negative effects – in contrast to the financial crisis. Establishments reporting negative pandemic effects were further asked about several types of negative ramifications and whether the establishment experienced them. Specifically, those are decreasing demand, (temporarily) shutting down, supply chain problems, staff shortages and liquidity shortages. Table 1 investigates the relationship between STW, negative pandemic ramifications and labor productivity.

Panel A of Table 1 considers whether the reported pandemic effects can explain the adverse selection into STW. Column (1) reproduces the conditional difference between the fifth and first labor productivity quintile, shown for instance in the right panel of Figure 5. In column (2), the sample is restricted to the set of establishments for which data on all types of negative
**Table 1: Explaining Adverse Selection into STW during the Pandemic**

### Panel A: STW and Negative Pandemic Effects

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent Variable: STW Indicator for 2020</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Labor Productivity</td>
<td>-0.102***</td>
<td>-0.096***</td>
<td>-0.050</td>
<td>-0.048</td>
</tr>
<tr>
<td>5th vs. 1st quintile</td>
<td>(0.034)</td>
<td>(0.034)</td>
<td>(0.032)</td>
<td>(0.032)</td>
</tr>
<tr>
<td>Negative Pandemic Effects</td>
<td>0.334***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decreasing Demand</td>
<td></td>
<td></td>
<td></td>
<td>0.273***</td>
</tr>
<tr>
<td>Shuttering Down</td>
<td></td>
<td></td>
<td></td>
<td>0.087***</td>
</tr>
<tr>
<td>Supply Chain Problems</td>
<td>-0.003</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Staff Shortage</td>
<td></td>
<td></td>
<td></td>
<td>0.018</td>
</tr>
<tr>
<td>Liquidity Shortage</td>
<td></td>
<td></td>
<td></td>
<td>0.108***</td>
</tr>
<tr>
<td>Full Set of Controls</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Establishments</td>
<td>6913</td>
<td>6525</td>
<td>6525</td>
<td>6525</td>
</tr>
</tbody>
</table>

### Panel B: Specific Negative Effects and Firm Quality

<table>
<thead>
<tr>
<th></th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent Variable</td>
<td>Decreasing Demand</td>
<td>Shuttering Down</td>
<td>Supply Chain Problems</td>
<td>Staff Shortages</td>
<td>Liquidity Shortages</td>
</tr>
<tr>
<td>Labor Productivity</td>
<td>-0.103***</td>
<td>-0.041**</td>
<td>-0.025</td>
<td>-0.067**</td>
<td>-0.145***</td>
</tr>
<tr>
<td>5th vs. 1st quintile</td>
<td>(0.036)</td>
<td>(0.020)</td>
<td>(0.033)</td>
<td>(0.029)</td>
<td>(0.035)</td>
</tr>
<tr>
<td>Mean Y</td>
<td>0.55</td>
<td>0.16</td>
<td>0.24</td>
<td>0.17</td>
<td>0.29</td>
</tr>
<tr>
<td>Full Set of Controls</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Establishments</td>
<td>6525</td>
<td>6525</td>
<td>6525</td>
<td>6525</td>
<td>6525</td>
</tr>
</tbody>
</table>

**Notes:** The table shows regression coefficients, where the full set of controls refers to the variables and fixed effects from (1). The dependent variable in Panel A is an indicator whether an establishment was in STW in 2020 or not. The dependent variables in Panel B are indicators whether an establishment reports the specific negative effect of the pandemic or not. Mean Y in Panel B denotes the overall share of establishments reporting the respective negative effect type. Labor productivity is measured in 2019. Survey weights are applied. Standard errors are clustered at the establishment level and reported in parentheses. Significance levels: * p < 0.1, ** p < 0.05, *** p < 0.01.
pandemic effects is available, which yields a very similar coefficient. Column (3) adds a dummy variable if an establishment indicates that it is negatively affected by the pandemic in 2020. Unsurprisingly, the indicator is strongly tied to STW. Interestingly, adding the pandemic effects also halves the firm quality coefficient and turns it insignificant. This suggests that a big portion of the observed adverse selection of less productive establishments can be explained by those establishments being hit harder by the pandemic downturn (also see Figure 6). Column (4) replaces the general crisis-effects indicator with the five specific types of negative ramifications to delineate which effects drive the result from column (3). Decreasing demand, shutting down and liquidity shortages are the only three factors that are positively associated with STW. Being up to three times larger than the other two estimates, the coefficient on decreasing demand is the most important one.

Panel B of Table 1 takes the five different indicators as dependent variables and shows how they are correlated with labor productivity. Higher labor productivity is negatively associated with all five, but supply chain problems and staff shortages are not predictive of STW. In absolute terms – and especially relative to the fraction of all establishments reporting each type of negative effect (denoted ‘Mean Y’) – the firm quality association is most pronounced for liquidity shortages (15 percentage points less likely), followed by decreasing demand (10 percentage points less likely). Lower quality establishments probably already had worse business models and worse liquidity when the pandemic unexpectedly hit and therefore had fewer maneuvering room and reserves to cushion the shock. With improving liquidity being one of the main reasons for STW (see Section 4.1), they took up more STW. Correspondingly, low-productivity establishments in 2020 are also significantly more likely to have taken up any type of governmental financial support other than STW.

Taken together, the finding that there is adverse selection into STW during the pandemic, but not during the financial crisis, can be attributed to establishments along the productivity distribution being affected differently by the two crises. The financial crisis disproportionately struck high-quality firms, primarily in manufacturing, but shows no clear selection pattern when adjusting for difference by narrow industries, occupations, regions and firm size. By contrast, the pandemic hit the economy more broadly and exhibits adverse selection even when conditioning on these other factors. This is driven by less productive establishments being more strongly negatively affected by the pandemic downturn and manifests itself in those establishments being more likely to report decreasing demand, temporarily shutting down and liquidity shortages. Facilitated by a worse business standing before the crisis, this pushes low-quality establishments towards using STW.
5 Employment Effects

5.1 Estimation

I estimate the effects of STW in the pandemic, utilizing data from 2016 until 2022. In general, when trying to identify the employment effects of a STW scheme with establishment-level panel data, estimation takes place via some version of the following specification (see e.g. Cahuc et al., 2021):

\[ Y_{et} = (\beta STW_e \times Post_t) + \psi_e + \tau_t + \varepsilon_{et} \] (2)

\( Y_{et} \) denotes (contributory) employment at time \( t \) (e.g., years), \( \psi_e \) are establishment fixed effects, \( \tau_t \) stand in for time fixed effects and \( \varepsilon_{et} \) denotes the error term. A matrix \( X_{et} \) could be added to (2) as an optional vector of time-varying control variables. \( \beta \) represents the parameter of interest in (2), capturing STW’s effect on employment after the shock \( Post_t \) has hit (e.g., by taking \( STW_e \) to be a dummy when there was STW in the establishment in 2020). This specification can be extended into an event-study:

\[ Y_{et} = \sum_{m \neq t-1} \beta_m (STW_e \times 1(t = m)) + \psi_e + \tau_t + \varepsilon_{et} \] (3)

This extension as spelled out in (3) has two appealing qualities. First, it allows to check for the existence of differential trends between firms using and not using STW before the event of going into STW takes place. Second, it allows for tracing the dynamics of STW by disentangling the short-, medium- and (possibly) long-run effects on employment after the period of STW has ended.

To fully capture the employment effects of STW, further consideration needs to be given to the fact that STW could affect establishment survival. Yet, to identify an employment effect, the assignment into STW needs to be as good as random, conditional on controls and the set of fixed effects. The evidence I present in Section 4.2 above, however, strongly suggests that establishment quality might be an important confounding variable that could taint employment effect estimates. This means that the firm quality dimension needs to be taken into consideration for estimating (and comparing) the employment ramifications of STW in the pandemic.

My approach to estimate the employment effects of STW is a matched event-study specification as in (3) which explicitly factors in the firm quality dimension when matching establishments. These employment effects would then constitute an average treatment effect on the treated (ATT). Concretely, to ensure comparability of establishments with and without STW,
I use coarsened exact matching (CEM) on a range of characteristics. I match exactly on the narrow 5-digit-level industry codes (more than 800), 20 ventiles of the AKM-FE distribution prior to the pandemic, the degree of urbanity in the establishment’s county of residence (four categories; taken from the German Federal Office for Building and Regional Planning), six establishment size classes\(^{10}\) measured in 2019, three establishment age categories measured in 2019 (less than 10 years, 10-25 years, 25+ years) and how STW-eligible employment developed from 2017 to 2018 and from 2018 to 2019 (increase, unchanged or decrease are the three categories, respectively).

For the baseline model estimates, I exclude establishments that had STW in the years before the pandemic and include only those STW-using establishments in the pandemic that received support any time in 2020 (for a maximum of a total annual work loss of 10%), but not in 2021 and 2022. I impose this restriction to be able to estimate what happens to employment after STW stops, answering the question whether any effects are of a lasting nature. While this restriction could also be seen as possibly introducing a bias through conditioning on future outcomes, I later show that relaxing it does not alter the results.

Note that the matching categories combine for more than a million possible combinations (‘strata’), of which roughly 440,000 exist in the sample. Out of about 850,000 establishments\(^{11}\) that fulfill the data requirements, I can successfully match 216,619 establishments from 43,353 different strata. 67,100 of these firms are establishments with a STW spell in 2020\(^{12}\). From this results a CEM weight distribution having a standard deviation of 1.84 with a mean of one. The maximum weight is 85 (the 99th percentile is at nine), implying that – given the total number of establishments in the analysis – no unduly high weights arise.

The CEM event-study approach relies on the assumption that the treatment group of STW establishments would have experienced a similar employment development as the matched control group of establishments without STW in the absence of the German short-time work scheme, also conditional on the establishment and time fixed effects. It is important to emphasize that I match establishments exactly on the narrow-most fifth-digit industry level and on 20 ventiles of the AKM-FE distribution to ensure that establishments are affected similarly by the pandemic shock and by the measures taken to contain it (e.g., lockdowns were based on the type of economic activity). Additionally, performing the matching in conjunction with the event-study

\(^{10}\)The same as for the selection equation (1): 1-4, 5-9, 10-49, 50-99, 100-499, 500+

\(^{11}\)The loss in establishments is mainly due to the restrictions, due to a range of 5-digit-industries not or only having STW establishments and due to the availability of the AKM fixed effects, also see above.

\(^{12}\)These establishments comprise about 10% of the total STW volume in 2020. On average, each STW establishment has more than four matched controls, albeit that the distribution of matched controls is right-skewed: On third of STW establishments have exactly one matched control, the median is two, the 75th percentile equals five, the 90th percentile 10, the 95th percentile 16, and the 99th percentile 34.
approach laid out in (3) allows me to examine whether there are differential employment trends between STW and matched non-STW establishments before the pandemic hit. Reassuringly, I find no evidence for the existence of differential pre-trends, lending support to the approach’s conditional parallel trends assumption.

Further possibly confounding influences are ruled out by the variables in the matching procedure. Including establishment age is motivated by the observation that younger firms are more likely to face liquidity crunches and have worse access to financial markets compared to otherwise similar establishments and this could thus impact the decision for STW. Size classes provide a further measure (albeit rough) of firm quality as well as a proxy for access to financial markets and for management practices (which often markedly differ between bigger and smaller firms), factors which could impact the decision whether to go into STW or not. Because better firms and industries tend to cluster around themselves in more urban areas, I consider how urban the establishment’s county of residence is. Moreover, the regional information also proxies for differences in labor market tightness since cities usually have thicker labor markets. To capture not just a static snapshot of establishments in 2019, but also of their histories, employment changes in the two years up to 2019 are part of the matching to facilitate only comparisons between establishments that are on the same (employment) trajectory.

In the following, I summarize the key takeaways from my robustness checks, which I discuss in more detail after the main results. The outlined approach requires that all relevant variables are captured in the matching or by the set of fixed effects. One set of factors that could conceivably be threatening relates to the structure of the workforce. To address this concern, I perform – on top of CEM – entropy balancing (Hainmueller, 2012) on the matched sample in a robustness check, balancing on a broad set of workforce characteristics in 2019, including skill, age, nationality and gender. Applying this procedure, the results remain unchanged, demonstrating that workforce characteristics cannot explain the observed employment differential between STW and non-STW establishments.

To further ensure comparability of establishments, I restrict the sample of STW establishments to ones that have a total work loss in 2020 of 10% or less (recall that this factors in the intensive margin of employees as well). This entails keeping more than half of the original STW establishments and makes establishments with and without STW more comparable as it can be assumed that it is more likely that establishments with very high usage of STW experienced stronger a higher Covid shock intensity due to unobserved factors that could predispose them to be vulnerable to shocks. I show that dropping the threshold leads to somewhat attenuated effects, a fact which supports this assumption, but still exhibits the same pattern of employment
While CEM makes sure that like is compared with like, a loss of information comes with it as a significant portion of the sample (see above) cannot be matched on the strict criteria. To alleviate concerns arising from loss of information I check the robustness of my results against an alternative inverse probability weighting (IPW) approach that utilizes the full sample and takes into account a broad set of establishment as well as workforce characteristics. The IPW specification yields the same pattern of findings. I report all robustness checks after presenting the main results.

5.2 Results

I present (graphical) results based on the event study specification (3) while applying weights obtained from the outlined CEM approach. I consider the period from 2016 until 2022, the current maximum year of data availability. 2019 is the year of reference.

I start with the baseline specification based on (3). The left panel (a) of Figure 7 plots estimates and 95% confidence intervals for the (dynamic) employment ramifications of being in STW. Prior to the pandemic, estimates hover around zero, showing no discernible pre-trend for log employment. For 2020, the coefficient indicates a strongly significant higher employment level for STW establishments compared to their matched non-STW counterparts, on the order of about 3.5%. In the two subsequent years after STW has ended for the establishments, the point estimates fall back to about 0.5%, estimates that are however insignificant. Thus, it seems that the employment effects of being in STW are short-lived and dissipate very rapidly after an establishment exits the scheme. Figure A5 in the Appendix compares this set of baseline results to a specification where STW establishments with STW in 2020 and 2021 are matched to non-STW controls. In this case, 2022 is the only post-STW period. However, the same pattern emerges in Appendix Figure A5 with the coefficient for 2022 being close to zero and insignificant. For 2021, this approach yields an employment estimate of around 2%, somewhat lower than for the pandemic’s initial year, likely because of the reduced crisis severity and accompanying reduced STW usage (see Figure 1). A similar picture arises when defining STW establishments as those with STW in 2020 and/or 2021, as depicted in Figure A6 in the Appendix.

The absence of differential pre-trends is encouraging with respect to the matching procedure. In fact, Figure A7 in the Appendix compares the event study from the left panel of Figure 7 with a specification that uses the same sample, but without the matching weights. There, a clear positive pre-trend in terms of employment can be seen with STW establishments having

---

13Recall that the STW establishments were selected such that they do not have STW in 2021 or 2022, to measure what happens after the scheme is exited.
negative employment coefficients before the pandemic that are rising with time, providing a
result which underscores once more the negative selection into the German STW scheme during
the pandemic recession.

One shortcoming of using logged employment as the dependent variable is the existence
of zeros in employment levels that would be caused by an establishment closing down. To
avoid looking solely at surviving establishments, panel (b) of Figure 7 plots results for two
transformations that do not drop zero employment observations, the inverse hyperbolic sine and
the log of employment plus one. Although the interpretation of the point estimates themselves
is no longer straightforward, the general pattern of estimates across time is the same, suggesting
that there is little happening in terms of differential establishment deaths between STW and non-
STW establishments. This, however, is not surprising as the German government radically eased
bankruptcy regulations at the start of 2020, which led to firm closures in the pandemic actually
being lower than before the pandemic (Müller, 2021). Hence, in this context with the current
end of the observation period in 2022, the ‘extensive margin’ of establishment employment plays
no significant role.

The swift dissipation of the positive employment relationship after the end of STW begets
the question what this is driven by. Based on the theoretical considerations, one would expect
STW establishments initially being able to hold on to their workforce while their non-STW
competitors shed employees. Afterwards, the re-convergence of employment could be due to
non-STW firms rehiring to their previous employment level or due to STW establishments also

Notes: The figure shows estimates and 95% confidence intervals based on (3), applying CEM weights. Standard
errors are clustered on the establishment level. 2019 is the base year. The dependent variable is log regular
employment in panel (a). In panel (b) the inverse hyperbolic sine of regular employment and the log of 1
plus regular employment are the dependent variables to capture zeros in employment (for instance, caused by
establishment closure).
shedding workers after the end of their STW usage. To investigate, I study employment changes as well as inflows and outflows. Figure 8 provides the results for these outcome variables.

Figure 8: Effects of STW: Employment Changes, Inflows and Outflows

Notes: The figure shows estimates and 95% confidence intervals based on (3), applying CEM weights. Standard errors are clustered on the establishment level. 2019 is the base year. The dependent variable is log employment change in panel (a). In panel (b) the outcome variables are inflow and outflow rates. These are, respectively, defined as the number of flows between two time periods divided by the average employment across the two time periods. The rates are, further, winsorized to limit the impact of outlier values.

Panel (a) of Figure 8 contrasts log employment changes between STW and non-STW establishments. In the years before the recession’s onset, both types of establishment have (after matching) very similar growth rates, but STW establishments have a 3-4 log points higher employment change in 2020 than establishments without STW. This does not necessarily imply that they are growing, it is merely the difference to non-STW establishments, which – as a matter of fact – experience employment declines in 2020. Yet, after STW ends, the estimated difference in employment change sharply turns into the opposite the following year, to about negative 3 log points, a coefficient that is in the same ballpark range in absolute terms as the previous year’s coefficient. In 2022, there is again no difference in employment growth between the two groups, the adjustment taking place directly the year after exiting STW. This produces the pattern observed in Figure 7.

Panel (b) of Figure 8 expands the analysis by separately looking at establishments’ inflow and outflow rates. The event study for inflows suggests that STW-using establishments consistently have marginally higher inflow rates after the pandemic strikes, being able to keep up hiring at a higher clip in the downturn. Even more interesting is the time line of coefficients for the outflow rate. As expected, in 2020, STW establishments have comparatively decreased outflows consistent with STW saving jobs and worker-firm-matches. However, 2021 exhibits an estimate which is the roughly same-sized inversion of 2020, i.e. STW establishments have increased
outflow rates relative to their matched peers that have not utilized the scheme. Outflow rates show no difference for 2022. This indicates that many employment relationships at risk are not lastingly saved by the German STW program, as it merely postpones the severance of the worker-firm-match.

While the employment adjustments between STW and non-STW establishments originate mainly from dynamic differences in outflows, the interpretation of this result hinges on whether these outflow differences derive mainly from involuntary outflows (i.e., firing) or voluntary outflows. On the one hand, if a majority of outflows (in STW establishments in 2021) are layoffs, this would suggest that STW saves jobs only temporarily and firms have to lay off some of their affected workers after exiting the scheme. If, on the other hand, a majority of those outflows is voluntary, STW saves jobs more lastingly but employees leave of their own volition. This could be the case, for instance, if employees treat STW as a strong signal of low firm quality and they leave because they doubt the firm’s future viability. This could be particularly salient for high-skilled workers. In the future, I plan to disentangle voluntary from involuntary outflows by classifying the nature of outflows based on how quickly those workers enter a new job. This will help me shed light on which of the two mentioned interpretations is the more important one for explaining the employment developments.

Section 4.2 above illustrates how there is negative selection into STW based on measures of establishments quality/productivity. Therefore, it is relevant to consider whether there are any differences in the employment effects for more and less productive establishments. Figure 9 compares the employment effects of receiving STW benefits between low-AKM and high-AKM establishments, where the delineation into low and high takes place via a split at the median of the AKM-FE distribution. Because 20 five-percent-intervals of the AKM fixed effect distribution are part of the CEM procedure, this simply splits the sample without changing any of CEM strata.

Figure 9 reveals differences by firm quality. While there are, for both groups, again no observable pre-trends, their estimated coefficients differ in the pandemic years. In panel (a), both groups have a positive effect in 2020, but the low-AKM establishments’ point estimate at 4.5% is nearly twice as large as the 2.5% for high-AKM establishments. Even more interesting is the divergence in 2021 and 2022, where the results for more productive establishments indicate slightly negative, but insignificant, employment effects compared to high-quality non-STW establishments. In contrast, the comparison within establishments that have low wage premia implies significant positive employment effects of about 1.5% persisting up to two years after the STW scheme was exited. When estimating the total effect of STW over the time period
Figure 9: STW Employment Effects: Low-AKM and High-AKM Establishments

Notes: The figure shows estimates and 95% confidence intervals based on (3), applying CEM weights. Standard errors are clustered on the establishment level. 2019 is the base year. Low-AKM establishments are those who are in the bottom half of the wage premia distribution as estimated from 2010 to 2017. Conversely, high-AKM establishments are in the distribution’s top half. Panel (a) takes log employment stocks as the outcome, panel (b) takes (winsorized) outflow rates as the outcome.

2016 to 2022 by comparing the time until 2019 with that from 2020 onwards (i.e., when estimating equation (2) instead of the event-study), the estimate for low-AKM establishments is 2.7% (S.E. 0.4) and 0.6% (S.E. 0.3) for high-AKM establishments\textsuperscript{14}. Correspondingly, the ability of Germany’s Kurzarbeit to save jobs in a lasting manner seems stronger for lower-quality establishments. Panel (b) of Figure 9 shows that, when considering outflow rates by AKM category as the outcome, this impact of STW is mostly driven by the reduction in outflows in 2020 being more pronounced for low-AKM establishments.

Taken together, the findings suggest a positive employment effect of the German STW scheme in the pandemic on establishments, albeit that almost all of the effect dissipates quickly when an establishment no longer is in STW. This relationship is driven by STW establishments having decreased outflows relative to non-STW establishments during their time in STW, but having relatively increased outflows afterwards that roughly cancel each other out dynamically. Employment effects during STW are more pronounced for low-AKM establishments, a group which, in contrast to high-AKM establishments, also has slight positive employment impacts even after exiting the scheme, courtesy of a stronger reduction in outflows while on STW.

5.3 Robustness

One threat to the econometric strategy could come from the assumption that all the relevant variables are captured in the matching step (or by the set of fixed effects). In particular,

\textsuperscript{14}For reference, the STW estimate based on (2) amounts to 1.8% (S.E. 0.3) for the full CEM sample.
workforce characteristics are mostly absent in the matching and could pose an issue. While there is strong evidence of assortative matching between workers and firms (e.g. Card et al., 2013), this assortative matching is not perfect and workforce characteristics (like how skilled the workforce is) could differ within matched strata by STW status. This might conceivably affect STW take-up and also employment after STW ends (for instance, higher-skilled workers being more likely to leave afterwards).

Additionally including employee characteristics is infeasible due to the fact that the number of matching strata would grow immensely and, henceforth, the curse of dimensionality would bite. Nevertheless, I consider an extension of my approach that explicitly brings in workforce characteristics and further apply entropy balancing (Hainmueller, 2012) to the matched sample. I adjust the matching weights in such a way that the first moments on a range of workforce variables are identical between the group of STW establishments and the group of non-STW establishments. Employee shares of German nationality, women, full-time workers and three skill levels as well as the mean age comprise the set of workforce characteristics.

I compare the baseline event study with the one obtained from adding entropy balancing as a second step on top of CEM in panel (a) of Figure 10. Reassuringly, this extra balancing step does little to the estimates, showing that differences in worker characteristics cannot explain the employment ramifications of STW which I find.

Figure 10: Robustness: Worker Characteristics and Time-by-Industry FE

(a) Adjusting for Worker Characteristics   (b) Including Year-by-Industry FE

Notes: The figure shows estimates and 95% confidence intervals based on (3), applying CEM weights. Panel (a) compares the baseline – from the left panel of Figure 7 – to a specification that additionally balances the obtained CEM-weights with respect to worker characteristics (age, nationality, gender, skill; see text for more details). Panel (b) compares the baseline with a specification where year-by-industry fixed effects are included instead of year fixed effects (industries are measured on the 2-digit level in this context, meaning 90 different industries).

Another concern could be that the employment development seen in the aggregate event
study could be an artefact of a misspecified regression equation, for instance if the year fixed
effects are insufficient for absorbing non-STW-related differences in time. To rule out this
possibility, panel (b) of Figure 10 overlays my baseline specification with a specification which
includes year-by-industry fixed effects for 90 2-digit industries (instead of just year fixed effects).
However, results remain unchanged. Similarly, I get quantitatively nearly identical coefficients
in Appendix Figure A8, where I include year-by-federal-state fixed effects in the estimation.

One peculiarity of the German STW scheme and how it is administered has to do with
the distinction between establishments and companies and the possibility that, for instance,
headquarters within a company could report STW for other establishments within the company
structure (see Section 3). This possibility cannot be accounted for explicitly by the admin-
istrative data, although it happens occasionally (Kagerl et al., 2022). To get a better grasp
on company networks and how they relate to establishments, I merge the administrative data
on establishments to an indicator of which establishments are single-plant establishments, and
hence not part of a larger company structure. This indicator emanates from the data of the
Mannheim Enterprise Panel which covers nearly all active firms in Germany (for more details
see Bersch et al., 2014). It should be noted that the link to the administrative establishment
data is preliminary (i.e., far from perfect), there is a sizeable portion of establishments whose
status as single-plant or multi-plant cannot be determined and it is not yet clear what selection
this restriction on establishments with a determined status induces. Nevertheless, as it is the
only available data source at this big a scale, I use this data for a robustness analysis.

Panel (a) of Figure 11 shows the baseline result and the result obtained by only consider-
ing establishments that are ascertained to be a single-plant establishments. Due to the data
constraints, this implies a reduction in sample size of about two thirds, to roughly 75,000 es-
tablishments. Correspondingly, the estimates exhibit a lower degree of precision. Yet, focusing
on single-plant establishments yields the same pattern of dissipating positive STW employment
effects, albeit that the coefficient for STW in 2020 is somewhat attenuated.

One restriction to ensure comparability that I make is the imposition of a maximum total
work loss in 2020 of 10% for STW establishments (see Section 5.1). In panel (b) of Figure 11, I
drop this restriction. Still, there is an employment effect in 2020 on the order of ca. 3% and no
strong discernible pre-trend. Merely the estimates for 2021 and 2022 show a marginally statisti-
cally significant negative employment effect of half a percent. Because dropping the restriction
increases the relative share of low-AKM establishments (also see the selection results in Section
4.2), the differences between top and bottom half AKM firms in Figure 9 cannot explain this; in
fact, this pattern actually affects the coefficients for the post-STW years positively. Therefore,
Notes: The figure shows estimates and 95% confidence intervals based on (3), applying CEM weights. Panel (a) compares the baseline – from the left panel of Figure 7 – to a specification that only considers single-plant establishments (N = 74,954 establishments; see text for more details). Panel (b) compares the baseline with a specification where the STW intensity threshold of a maximum 10% total work loss in 2020 for STW establishments is dropped (N = 257,667 establishments; see text for more details).

it seems likely that the STW group without the intensity restriction is affected more strongly by the crisis shock relative to non-STW firms. This might then be reflected by the pre-Covid coefficients, the very high STW intensity and by those firms shedding more workers after the STW spell than non-STW establishments do during the STW period.

5.4 An Alternative Approach

One disadvantage of relying on coarsened exact matching is the elimination of unmatched establishments from the analysis sample. As is evident from the numbers given in Section 5.1, the establishment sample shrinks by nearly 75%. For STW establishments, however, the decrease is relatively less severe at 45%. Still, to avoid this loss of (potentially useful) information, I consider an alternative approach to estimate the employment effects of STW that utilizes the full sample of (STW) establishments. To this end, I perform an inverse probability weighting (IPW) approach on the full sample. Because IPW does not punish nearly as much as CEM does for including more variables in the matching step (curse of dimensionality), I include a broad range of establishment variables and of workforce characteristics in the logit estimation of the underlying propensity score. Specifically, the propensity score estimation for the base year 2019 includes indicators for 5-digit industries, indicators for counties, log employment (plus three lags of it), the log median wage, establishment age, the value of the establishment AKM-FE as well as the mean age of the workforce and employee shares by gender, nationality, full-time, three skill levels and twelve occupations.
The standard deviation of the resulting IPW weights (enforcing common support) is smaller than the one for the CEM weights (1.37 versus 1.84). Similarly, a maximum value of roughly 61 indicates no outliers. Panel (a) of Figure 12 juxtaposes the event study with the IPW sample and weights with the CEM baseline. Incorporating more establishments and more information (on occupations, among others) in the IPW approach yields the same pattern of rapidly vanishing employment effects after the absence of trends prior to Corona. Merely the size of the 2020 coefficient is subdued to slightly above 2% when using IPW. Panel (b) of Figure 12 applies the IPW approach to the outcomes of inflow and outflow rates. As in the CEM case, outflows are subdued during STW and increased after STW ends (relative to non-STW establishments). Regarding inflows, the IPW coefficients actually suggest lower inflows during STW in 2020, but slightly higher inflows for STW establishments in 2022 as well.

Figure 12: Inverse Probability Weighting as an Alternative Matching Approach

Notes: The figure shows estimates and 95% confidence intervals based on (3). Panel (a) compares the baseline – from the left panel of Figure 7 – to the IPW specification (see text for more details). In panel (b), the outcome variables are inflow and outflow rates. These are, respectively, defined as the number of flows between two time periods divided by the average employment across the two time periods. The rates are, further, winsorized to limit the impact of outlier values.

The divergent patterns by establishment quality seen in Figure 9 for coarsened exact matching are replicated for IPW in Appendix Figure A9. Here, the event study suggests significantly negative post-STW employment effects for high-AKM establishments, owing to the improved precision offered by utilizing IPW on the full sample.

Summarizing the IPW results, it is reassuring that this alternative approach, which incorporates a bigger slate of characteristics at the expense of not directly matching establishments, still can deal with the selection into STW (also see Appendix Figure A10 for a comparison of the IPW weighted event study with the unweighted one) and produces the same findings as does the CEM approach. Thus, my result of positive STW employment effects that quickly fade after

30
STW proves econometrically robust.

6 Windfall Effects

Windfall effects of STW describe job matches which are supported through the STW scheme, but which would not be lost in the (counterfactual) absence of STW. In the context of the pandemic, the German government eased access to the job match insurance. Alongside the fact that the Federal Employment Agency was straining to handle the unprecedented level of STW applications in the spring of 2020, these two factors led to many applications being examined much faster and in a less detailed manner than usually. Correspondingly, it can be expected that more applications were granted due to the special circumstances of the pandemic and the eased access rules (see Section 2). Potentially, some firms factored this in and profited off the easy STW access despite there being no risk of job destruction.

One extreme version of windfall is the fraud-like abuse of STW allowances, for example by officially applying for hours reductions for employees but unofficially still requiring or expecting employees to achieve the full-time work load. This constitutes free riding in the sense that the STW benefits are abused to subsidize wages despite no decrease in work hours. The scope of such illicit behavior is obviously hard to gauge, the Federal Employment Agency only reports few cases where firms were required to repay their STW allowances after a later review. Based on experimental methods within an employee survey, Bossler et al. (2023) find a sizeable incidence of actual working hours being longer than stated in the official notification.

Free riding in this narrow sense is, however, only a subset of windfall. Another, likely much more wide-spread, version of windfall is the support of jobs with hours reductions that would not have been lost without STW. For instance, a firm only has a modest decrease in revenue which it could have weathered even without government support. One anecdote widely reported in German media concerned the car company Mercedes-Benz in 2023, which put some of its employees in STW due to supply chain problems, despite earning billions of Euros in profit at the same time. While this might be one obvious case of windfall effects (as it stands to reason that Mercedes-Benz would not have fired those employees in the absence of STW, since layoffs are excluded under an agreement with the company’s works council until 2029), the question is what the overall scale of the windfall effects is. In the following, I try to provide approximate numbers on the windfall of the German STW scheme based on different approaches.

One suggestive indicator can be the share of firms with low (average) work loss during the

\[15\text{E.g., from the German public broadcaster: } \text{https://www.tagesschau.de/wirtschaft/unternehmen/mercedes-kurzarbeit-mitarbeiter-101.html} \]
pandemic. If the total work loss in STW is low, it becomes more likely that the firm could have held on to these workers if it wanted. For example, 7.5% of all STW firms had a total actual work loss between March and December 2020 of less than 2%. Of course, where to set the threshold is arbitrary, but the decision early in the pandemic to reduce the minimum threshold to 10% of workers (from one third) with at least 10% work loss in a STW month resulted in a higher ex-post share of establishments that have very little actual total work loss.

A superior measure of windfall is the share establishments (and their employees) that experienced no or only small actual revenue decreases (also see Cahuc et al., 2021). This relates directly to the economic reasons that need to be fulfilled for STW allowances to be granted, e.g. supply chain problems, a lack of orders or – in the pandemic – governmental orders to restrict business activities. All these reasons should manifest themselves in noticeable decreases in establishment revenue. For evidence on revenue changes and economic consequences of the the pandemic, I again turn to data collected from establishments surveys (Establishment Panel and BeCovid) that can be linked to the administrative data.

I a first step, I analyze how wide-spread the utilization of STW is among (single-plant) establishments that report no revenue decrease from 2019 to 2020\textsuperscript{16}. In the BeCovid survey, establishments were asked how their revenues in 2020 compared to those in 2019. Among those reporting revenue drops, about half have at least one STW spell in the pandemic. For those reporting that revenue stayed (roughly) the same or increased, the share of STW establishments is much lower, but still stands at 20%.

In a second step, I use survey information on the pandemic’s economic impact on establishments to delineate which establishments were in distress in 2020. I draw this information from the Establishment Panel where establishments were asked in 2020 whether the pandemic had negative economic impacts on the establishments or not. 55% of firms reporting such negative effects used STW in 2020, yet so did 10% of those without adverse impacts.

Taking together, the various indicators for windfall effects shown so far – very low work loss, no initial revenue decrease in the pandemic, no self-reported negative crisis effects – do indeed suggest that there is a sizeable amount of windfall. But, the numbers just presented merely constitute a lower bound on the windfall because of the coarseness of the indicators for the pandemic impact. Establishments that were negatively affected or experienced revenue decreases are not counted even though they might not have laid off workers in the absence of STW.

\textsuperscript{16}For establishments with low revenue decreases and STW, it is hard to disentangle whether the crisis itself led to the revenue decrease or if STW caused (parts of) it. Hence, I consider only establishments with no (or positive) revenue changes.
An alternative way to arrive at a figure for windfall effects is to take the employment effect estimates from Section 5 at face value and do a back-of-the-envelope calculation to get an amount for the number of saved jobs. Setting this amount in relation to the total number of employees in STW then gives a rough approximation of the windfall. Figure 7 yields a 3.5% employment coefficient for 2020, albeit that this a mean effect over establishments. To get a more useful estimate in terms of employees, I run the CEM baseline regression but additionally incorporate employment weights for 2019. This yields a coefficient of about 4.1% for 2020 (95%-CI bounds: 2.5% and 5.6%). Contributory employment in STW establishments in 2020 accounted for a total of nearly 11 million workers, therefore the estimated number of jobs saved by STW is about 450,000 (95%-CI bounds: 275,000 and 615,000).

Average unemployment stood at 2.7 million in Germany in 2020, implying that the number of jobless would have been higher by one sixth were it nor for STW (using the point estimate). Mapped onto the unemployment rate, the estimate implies an unemployment rate in 2020 of 6.9% instead of 5.9%.17 In total, 8.75 million unique people were in STW in 2020. Further taking into account an average work loss of 42% in 2020 during the pandemic would suggest a number of 3.7 million employment equivalents that were supported by STW. In relation to the estimated number of jobs saved, this indicates substantial windfall effects.

However, these are back-of-the-envelope numbers derive from a partial-equilibrium calculation and come with some fairly strong caveats. First, the point estimate is potentially inaccurately estimated because it is derived only from the STW sample of establishments with STW in 2020 (but not afterwards), excluding some heavy and long-time users of STW. Second, the calculation relies on extrapolating from the matched sample to the entirety of STW establishments. The estimates from IPW use the full sample and are only slightly lower, but both approaches incorporate information on AKM-FE, which are missing for a sizeable chunk of establishments. Third, the calculations are only for the initial shock year of 2020 and neglect any possible contribution of STW to stabilizing the economy as a whole. Such a stabilizing effect can for example derive from the mere existence of STW, which constitutes an ex-ante known safety net and thus reduces uncertainty for firms. For 2021 and later, the dynamics of unemployment and labor market frictions start to be even more crucial for the policy’s macroeconomic implications. Therefore, I restrict my back-of-the-envelope calculations to 2020. As stressed by Balleer et al. (2016), the interactions of labor market institutions matter greatly for the dynamic effectiveness and efficiency of STW.

17Implied unemployment changes at the lower bound of the 95%-CI in the absence of STW: One tenth more unemployed and a 6.5% unemployment rate. At the upper bound of the 95%-CI: 23% more unemployed and a 7.3% unemployment rate.
To summarize, the evidence – despite some caveats – indicates that there are significant windfall effects associated with the German STW scheme during the pandemic. There were STW spells with very little work loss and many establishments were in STW despite no or limited impact of the pandemic on their operations or revenues. Moreover, while the employment ramifications translate to hundreds of thousands of saved jobs, millions were in STW. This suggests that a majority of the STW benefits were not targeted very well, leading to a great deal of windfall.

7 Conclusion

The pandemic has brought the labor market instrument short-time work into the spotlight again. Between 2020 and early 2022, the costs of STW in Germany accrued to a total of 46 billion Euros and in 2020 as well as 2021 yearly STW costs were comparable to yearly spending on unemployment insurance. This paper surveys the scope of STW in Germany, considers the reasons why establishments go into STW, analyzes the selection into STW and investigates its employment and windfall effects.

For my analysis, I use novel administrative data on STW in Germany and further draw on survey information that is linked to the official data. I find that the unprecedented use of STW in the pandemic was motivated not only by firms’ immediate financial concerns, but also by their desire to hold on to their workforce in a tight labor market. Still, I document a strong negative selection into STW based on measures of establishment quality and productivity. This selection pattern, however, is only apparent in the Covid recession, not during the financial crisis. The two crises having different ramifications along the firm productivity distribution provide an explanation. Firms with more exposure to international markets – who tend to be bigger and more productive – were hit disproportionately often during the financial crisis, while the Covid-19-downturn struck the economy very broadly. With low-quality firms being in a worse position prior to the shock, the pandemic made them more likely to face decreasing demand, temporarily shutting down and liquidity crunches, which drive the selection into STW.

Regarding the core argument of saving jobs, I demonstrate that establishments in STW have higher employment levels compared to their matched non-STW counterparts when they are receiving the benefits. After exiting the scheme, employment effects quickly become insignificant and mostly vanish, a finding supported by a plethora of econometric specification checks and robustness exercises. This result is driven by outflows among STW establishments being initially lower, but being higher after the end of STW. In addition, I present evidence that the employment-saving property of STW is more pronounced and lasts longer for low-productivity
establishments. All in all, back-of-the-envelope calculations suggest up to half a million jobs were saved by STW in 2020. Yet, with 8.75 million jobs – translating to 3.7 million employment equivalents – that were supported by STW allowances overall in 2020, large windfall effects are also a feature of Germany’s STW implementation. The existence of windfall is corroborated by some establishments receiving STW benefits despite experiencing no or little pandemic repercussions in terms of economic effects and revenue.

My results have implications for policy makers. The substantial amount of windfall effects implies that improving the targeting of the German STW scheme could make it more efficient. There should be a focus on making sure that establishments are having sizeable revenue decreases due to the shock in question before granting STW. Further, the special eased access rules during the pandemic – which in part ran until 2022, see Section 2 – could have been tied directly to the countermeasures against Covid-19, for instance to lockdown orders or social restrictions. Related, introducing an experience rating element to Kurzarbeit could be helpful, a mechanism working well in other countries (see OECD, 2021; Fitzenberger and Walwei, 2023) and suggested already by Burdett and Wright (1989). The partial paying back of STW costs at a later point through increased contributions by former STW firms provides a stronger incentive to only take up STW if necessary.

However, I stress that my analysis is reduced-form and thus in partial equilibrium. I do not include (dynamic) general equilibrium considerations, such as the possibility that STW has a stabilizing macroeconomic effect through it acting as safety net and preventing a crisis from further spiralling downwards. Similarly, my results do not reasonably lend themselves to an explicit cost-benefit calculation of short-time work versus unemployment and the scheme’s overall welfare effects. STW and unemployment insurance are not one-to-one substitutes either way and, as forms of social insurance, exhibit complementarities (Giupponi et al., 2022): For instance, unemployment insurance can deal with persistent structural changes while STW is better suited to cushioning temporary and acute shocks.

Worthy of further investigation are the reallocation implications of STW. On the one hand, the stronger employment results for less productive establishments could be read as foretelling negative reallocation effects. On the other hand, the existence of large windfall effects could be read as indicating that any reallocation effects are quite limited. In the future, with more data available on the post-pandemic period and special regulations on STW and bankruptcy having run out, reallocation effects of STW could provide a fruitful avenue for research.
References


Appendix

Figure A1: Selection on the Extensive Margin: Robustness (2020)

Notes: The figure shows coefficients for establishment quality based on specification (1) for 2020 with the full set of fixed effects, occupational controls. In panel (a), a full set of dummies for twenty ventiles for the median wage distribution is included additionally. In panel (b), only single-plant establishments are considered in the estimation. Used is the position in the distribution of the AKM-FE which were calculated from the period 2010 to 2017. The dependent variable is a dummy taking the value one if an establishment has at least one STW spell in the given year, and zero otherwise. The reference category is the eleventh of the twenty ventiles, i.e. between the 51st and the 55th percentile. $N(a) = 2,097,889$ establishments and $N(b) = 961,503$ establishments.
Figure A2: Selection: AKM-FE and the Intensive Margin of STW (2020)

Notes: The figure shows coefficients for establishment quality based on specification (1) for 2020 – conditioning on establishments with at least one spell in STW – with the full set of fixed effects and occupational controls. Used is the position in the distribution of the AKM-FE which were calculated from the period 2010 to 2017. The dependent variable is the average employment equivalent/work loss in STW over the whole year for the establishment. In the computation, a month without STW is given the value zero. The baseline share for STW establishments for this work loss measure is 0.18 in 2020. The reference category is the eleventh of the twenty ventiles, i.e. between the 51st and the 55th percentile.
Figure A3: Selection Over Time: Extensive Margin Robustness

(a) AKM-FE: Including Wage Controls
(b) AKM-FE: Only Manufacturing

Notes: The figure shows coefficients for establishment quality based on specification (1) with the full set of fixed effects and occupational controls. The dependent variable is a dummy taking the value one if an establishment has at least one STW spell in the given year, and zero otherwise. The coefficient values represent increasing the SD of the AKM-FE (within a 5-digit-industry) by one. This quality proxy is computed in the year prior, except in 2021 and 2022 where 2019 is used and in 2010 and 2011 where 2008 is used. Specifications in panel (a) include a full set of dummies for twenty ventiles for the median wage distribution on top of the baseline specification. Panel (b) only consider establishments in the manufacturing sector.

Figure A4: Selection Over Time: Intensive Margin

(a) AKM-FE: One SD Increase
(b) Labor Prod.: 5th vs. 1st Quintile

Notes: The figure shows coefficients for establishment quality based on specification (1) – conditioning on establishments with at least one spell in STW – with the full set of fixed effects and occupational controls. The dependent variable is the average work loss in STW over the whole year for the establishment. In the computation, a month without STW is assigned the value zero. In panel (a) the coefficient values represent increasing the SD of the AKM-FE (within a 5-digit-industry) by one. Panel (b) shows the difference obtained between the fifth and first quintile of labor productivity while applying survey weights. Both of these quality proxies are computed in the year prior, except in 2021 and 2022 where 2019 is used and in 2010 and 2011 where 2008 is used.
Notes: The figure shows estimates and 95% confidence intervals based on (3), applying CEM weights. Standard errors are clustered on the establishment level. 2019 is the base year. The baseline corresponds to panel (a) of Figure 7. The other event study performs the same matching procedure, but takes into account STW establishments that were in the scheme in 2020 and 2021. The baseline solely focuses on establishments that were in STW only in 2020. CEM abbreviates coarsened exact matching.
Figure A6: STW Employment Effects: Use of STW in 2020 and/or 2021

Notes: The figure shows estimates and 95% confidence intervals based on (3), applying CEM weights. Standard errors are clustered on the establishment level. 2019 is the base year. The baseline corresponds to panel (a) of Figure 7. The other event study performs the same matching procedure, but takes into account STW establishments that were in the scheme in 2020 and/or 2021. The baseline solely focuses on establishments that were in STW only in 2020. CEM abbreviates coarsened exact matching.

Figure A7: STW Employment Effects: Comparison of CEM to Unweighted Estimation

Notes: The figure shows estimates and 95% confidence intervals based on (3). Standard errors are clustered on the establishment level. 2019 is the base year. CEM abbreviates coarsened exact matching.
Notes: The figure shows estimates and 95% confidence intervals based on (3). Standard errors are clustered on the establishment level. 2019 is the base year. The figure compares the baseline with a specification where year-by-state fixed effects are included instead of year fixed effects (specifically, the 16 federal states). CEM abbreviates coarsened exact matching.
Notes: The figure shows estimates and 95% confidence intervals based on (3), applying IPW weights. Standard errors are clustered on the establishment level. 2019 is the base year. IPW denotes inverse probability weighting. The figure investigates differences between low- and high-AKM establishments using the IPW approach, the IPW counterpart to Figure 9. Low-AKM establishments are those who are in the bottom half of the wage premia distribution as estimated from 2010 to 2017. Conversely, high-AKM establishments are in the distribution’s top half.
Figure A10: STW Employment Effects: Comparison of IPW to Unweighted Estimation

Notes: The figure shows estimates and 95% confidence intervals based on (3). Standard errors are clustered on the establishment level. 2019 is the base year. IPW denotes inverse probability weighting.