The Cultural Cost of Overwork: Evidence from Switzerland's Röstigraben*

Giulian Etingin-Frati† Nicolas Marti‡

April 3, 2026

Abstract

Does culture shape how burdensome overtime work feels to workers? We exploit Switzerland's linguistic border, the *Röstigraben*, where French- and German-speaking workers share the same labor laws but inherit different attitudes toward work and leisure. Using data from the Swiss Household Panel (1999–2023), we show that each extra hour beyond the contract raises work-life interference by 0.038 points (on a 0–10 scale) more for German-speaking workers than for French-speaking workers, an effect that is modest in absolute terms (0.12 within-person standard deviations at mean overwork) but represents a 152% amplification of the French-speaking baseline. This cultural amplification is concentrated among part-time workers and women; for men, the effect appears exclusively in the part-time subsample, consistent with a contractual-salience mechanism that operates primarily for men; the effect vanishes when the hours gap is measured relative to habitual rather than contractual hours, consistent with a reference-point mechanism. Despite bearing higher psychological costs, German-speaking workers do not detectably adjust their labor supply differently at the annual horizon: they do not correct overwork episodes faster, do not bunch more tightly at contractual hours, and do not exit overwork situations through job changes.

JEL Classification: J22, J28, Z13, R12

Keywords: overwork, cultural norms, contractual salience, work-life interference, reference-dependent preferences

*This paper was produced as a human-AI collaboration to investigate the current capabilities of large language models in applied economics research. Claude 4.6 (Anthropic) and Gemini 3 Pro (Google, via Antigravity) were used throughout the research process, including literature review, hypothesis development, coding of the empirical analysis, drafting, and revision. AI-generated referee reports were used to stress-test the arguments and identify weaknesses. The human authors directed the research agenda and made all final decisions on methodology, interpretation, and presentation. This study uses data collected by the Swiss Household Panel (SHP), which is based at the Swiss Centre of Expertise in the Social Sciences FORS. The project is supported by the Swiss National Science Foundation.

†KOF Swiss Economic Institute, ETH Zurich, Leonhardstrasse 21, 8092 Zurich, Switzerland. E-mail: etingin-frati@kof.ethz.ch

‡KOF Swiss Economic Institute, ETH Zurich, Leonhardstrasse 21, 8092 Zurich, Switzerland. E-mail: marti@kof.ethz.ch

1 Introduction

Modern labor markets are characterized by a persistent tension between contractual obligations and the implicit expectation of discretionary effort. While economic theory typically models labor supply as a tradeoff between leisure and income,[1] the psychological cost of "overwork" (the gap between actual and contractual hours) may depend heavily on the cultural scripts through which workers interpret their professional obligations. When does an extra hour of work feel like a fair exchange, and when does it feel like a violation of the boundary between professional and private life? Answering this question requires a setting where workers share the same institutions but differ in cultural norms.

This paper asks whether cultural norms shape the psychological cost of working beyond contractual hours. We argue that they do, and that the magnitude of this cultural effect depends on how prominently the employment contract marks the work-leisure boundary. Testing this claim requires a setting in which individuals share identical institutional constraints but differ in deep-seated cultural norms. Switzerland provides precisely such a natural experiment. French-speaking and German-speaking workers operate under the same federal labor code, face similar tax considerations[2], and participate in the same macroeconomic environment, yet they inherit strikingly different historical attitudes toward work, leisure, and the boundary between professional and personal life (Brügger et al., 2009; Eugster et al., 2017). The linguistic border that separates these communities, colloquially known as the *Röstigraben*, allows us to estimate the cultural moderation of the psychological costs of labor supply.

Using 25 waves of the Swiss Household Panel (SHP, 1999–2023), we construct a longitudinal dataset of employed individuals in German- and French-speaking cantons. We define "overwork" as the gap between actual weekly hours worked and contractual hours per week. Exploiting within-person variation, we estimate how each additional hour of overwork differentially affects French-speaking workers across a battery of burnout outcomes (post-work exhaustion, work-life interference, and difficulty disconnecting) as well as domain-specific satisfaction measures.

Three main findings emerge. First, each additional hour of overwork raises work-life interference by 0.038 points on a 0–10 scale ($p < 0.001$) more for German-speaking workers than for French-speaking workers, an effect that is modest in absolute terms (0.12 within-person standard deviations at mean overwork) but represents a 152% amplification of the French-speaking baseline overwork cost. Post-work exhaustion shows a similar pattern ($p < 0.001$), though this result is sensitive to outlier trimming and should be treated as corroborating rather than definitive. German-speaking norms, which place a premium on strict contractual boundary compliance and the protection of leisure time (the "Feierabend" principle), generate a higher marginal penalty per hour of contract violation. Critically, this effect is absent among full-time men—the largest demographic subgroup ($N = 30{,}494$)—and is concentrated among part-time workers and women. For men, the cultural penalty appears exclusively in the part-time subsample ($\hat{\beta}_3 = -0.054$, $p < 0.001$), with the full-time male interaction precisely zero ($\hat{\beta}_3 = -0.002$, $p > 0.8$). For women, the cultural moderation does not depend on contract type. Within part-time workers, men show a larger cultural interaction than women. The pattern is consistent with a contractual-salience mechanism that operates primarily for men: part-time contracts more explicitly demarcate the work-leisure boundary, and German-speaking norms

[1] See, for example, Heckman (1993) and Lucas and Rapping (1969).

[2] Within bilingual cantons, municipalities on both sides of the language border share the same cantonal tax schedule and set only a scalar multiplier on it (Eugster and Parchet, 2019). These multipliers converge at the border due to fiscal competition. The Federal Act on the Harmonization of Direct Taxes (StHG) further ensures a uniform tax base nationwide. Work incentives at the Röstigraben are therefore comparable across language regions.

amplify the cost of violating it. We caution, however, that full-time men—the largest single subgroup—show zero cultural moderation, so the main result should not be interpreted as a universal feature of the German-speaking labor market.

Second, overwork also reduces satisfaction with free time ($p < 0.001$) more for German-speaking workers. Life satisfaction ($p = 0.076$) and job satisfaction ($p = 0.069$) do not reach conventional significance thresholds. The cultural penalty thus operates primarily through specific boundary-related channels rather than through broad reductions in global well-being.

Third, despite bearing higher psychological costs, German-speaking workers do not detectably adjust their labor supply differently at the annual horizon of our survey. They do not bunch more tightly at contractual hours, do not correct overwork episodes faster, and do not exit overwork situations through job changes at higher rates. We cannot rule out within-year adjustments that are resolved before the next survey wave. This absence of detectable differential behavioral response is consistent with either a "constrained preferences" view (Knaus and Otterbach, 2019) or rational non-response to welfare costs that are modest in absolute terms. A placebo test is consistent with the cultural penalty being specific to overwork and absent for underwork, while an income analysis rules out differential compensation as an alternative explanation.

We make three contributions relative to the existing Röstigraben literature. First, while prior work documents that cultural norms at the linguistic border shape unemployment duration (Eugster et al., 2017), demand for redistribution (Eugster et al., 2011), and female labor-force participation (Steinhauer, 2018), we show that culture also operates on the *intensive margin*: the psychological cost of each additional hour of work beyond the contract. This is distinct from the extensive-margin effects documented previously. Second, we develop a conceptual framework of *contractual salience*: the employment contract serves as a reference point (Kahneman and Tversky, 1979; Kőszegi and Rabin, 2006; Hart and Moore, 2008) whose psychological prominence varies with contract type and is culturally amplified (Bordalo et al., 2012, 2013). The sharpest evidence for this mechanism is that the cultural penalty appears only when the hours gap is measured against contractual hours ($p < 0.001$) and vanishes when measured against habitual hours ($p = 0.548$), a distinction the prior literature has not tested. Third, by documenting the absence of detectable differential behavioral adjustment at the annual frequency across the linguistic divide despite differential psychological costs, we extend the literature on work-hour mismatches (Knaus and Otterbach, 2019), showing that culture shapes the subjective evaluation of labor supply without generating behavioral responses observable in annual panel data.

The remainder of the paper proceeds as follows. Section 2 reviews the institutional context of the Swiss labor market and the relevant empirical literature. Section 3 develops a conceptual framework of reference-dependent labor supply with cultural heterogeneity, deriving four testable predictions. Section 4 describes the dataset and our econometric identification strategy. Section 5 presents our findings on the cultural modulation of labor supply costs. Section 6 concludes.

2 Institutional Background and Literature

2.1 The Swiss Cultural Divide

Switzerland's linguistic border runs roughly north–south through the cantons of Bern, Fribourg, and Valais, separating a German-speaking majority (about 63% of the population) from a French-speaking minority (23%). The border has been remarkably stable since the early mod-

ern period; it does not coincide with cantonal boundaries, administrative districts, or major geographic barriers, making it a credible source of exogenous variation in cultural exposure (Brügger et al., 2009).

Several studies exploit the linguistic border as a natural experiment, using spatial regression discontinuity designs to isolate cultural effects from institutional and environmental confounds. This approach complements the broader literature on culture and economic outcomes (Alesina and Giuliano, 2015) and the epidemiological approach to identifying cultural transmission (Giuliano, 2007; Luttmer and Singhal, 2011). These studies document sharp discontinuities in a range of economic attitudes and behaviors. German-speaking communities are substantially less supportive of redistributive social insurance (Eugster et al., 2011) and prefer lower taxes with less redistribution (Eugster and Parchet, 2019). They also exhibit shorter unemployment durations and stronger work-first attitudes (Eugster et al., 2017). Female labor-force participation is lower on the German-speaking side, consistent with more traditional gender-role norms (Steinhauer, 2018). Most recently, Faessler et al. (2024) find discontinuities in voting on health and fertility-related policy, suggesting that cultural differences at the border extend to choices with direct consequences for mortality and reproduction. Deopa and Fortunato (2021) document that German-speaking cantons reduced mobility for non-essential activities significantly more than French-speaking cantons during COVID-19 lockdowns, consistent with stronger norm compliance on the German-speaking side. Across domains, the border captures deep-seated differences in attitudes toward effort, independence, and the balance between market and non-market time.

All of these preference differences exist within a unified institutional framework. Federal labor law sets uniform standards for maximum working hours (45–50 hours per week depending on the sector), overtime compensation, and contract termination[3]. Unemployment insurance is governed by a single federal statute with identical replacement rates and benefit durations[4]. Tax schedules are set at the federal and cantonal levels, but the cantonal variation does not align neatly with the language border. This institutional uniformity allows us to attribute differential trends in overwork to cultural factors rather than to regulatory differences.

2.2 The Economics of Working-Time Preferences

The neoclassical labor-supply model predicts that hours are determined by the intersection of wage rates and preferences for leisure. In a frictionless market, actual hours equal desired hours. Reality departs from this benchmark in two well-documented ways. First, employers face coordination costs and fixed costs per worker that incentivize the bundling of hours into standardized schedules, creating "hours constraints." Workers may be unable to choose their preferred hours and instead select from a discrete set of employer-offered packages. Second, social norms shape the reference point against which hours are evaluated. A 42-hour work week might be perceived as "normal" in Zurich but excessive in Geneva[5].

[3]The structural parameters of employment are codified at the federal level, primarily through the Labour Act (*Arbeitsgesetz*, ArG, SR 822.11) and the Code of Obligations (OR, SR 220). The ArG prescribes a maximum weekly ceiling of 45 hours for industrial and technical workers, office staff, and large-scale retail employees, while a 50-hour limit applies to other sectors. The Code of Obligations also ensures a unified national standard for 'freedom of termination' (*Kündigungsfreiheit*) and the mandatory 25% premium for overtime, effectively neutralizing institutional variance in the legal cost of labor across the linguistic divide.

[4]Unemployment compensation is strictly harmonized under the Federal Act on Compulsory Unemployment Insurance (*Arbeitslosenversicherungsgesetz*, AVIG, SR 837.0). While the administrative execution of these benefits is decentralized through cantonal unemployment offices (KAST/OCIRT), the eligibility criteria and fiscal parameters remain invariant across linguistic regions.

[5]In 2024, the average Swiss working time in hours per week was 41.6. In Zurich, this was 41.7 hours per week. Geneva was lowest, with 40.9 (Federal Statistical Office, 2024).

The well-being consequences of work hours depend less on their absolute level than on the gap between actual and preferred hours. Wooden et al. (2009) establish this point using Australian panel data: hours mismatches reduce job and life satisfaction by magnitudes comparable to acquiring a disability, while total hours alone have little independent effect. Lepinteur (2019) similarly finds that legislated workweek reductions that narrow the gap between actual and desired hours raise worker well-being, while Bell and Freeman (2001) document large cross-country differences in hours worked between the US and Germany that cannot be explained by wages alone, pointing to cultural and institutional factors. The damage is asymmetric: a large share of British workers report a mismatch between actual and desired hours, with the majority of mismatched workers preferring fewer hours (Bell and Blanchflower, 2011). Otterbach (2010) documents the prevalence of hours constraints across 21 countries, including Switzerland, finding that the desire for additional or fewer hours is strongly related to macroeconomic conditions, income inequality, and working conditions. These findings motivate our focus on the hours gap rather than on absolute hours. We depart from the existing literature by using contractual hours rather than stated preferences as the reference point, a choice we justify in Section 3.

2.3 Cultural Norms and the Cost of Effort

Overwork represents an upward shift in labor supply at the intensive margin. In a standard framework, workers supply effort until the marginal utility of income equals the marginal disutility of labor. This disutility is not a biological constant; it is socially constructed and reference-dependent. Taxi drivers set daily income targets that generate negative labor supply elasticities (Camerer et al., 1997). [6] Workers bunch at institutionally defined thresholds such as statutory retirement ages even without financial incentives to do so (Seibold, 2021). In a field experiment with Zurich bicycle messengers, Fehr and Goette (2007) find that loss-averse workers sign up for more shifts when wages rise but exert less effort per shift. The same logic scales to aggregate outcomes: while Prescott (2004) attributes cross-country hours differences primarily to differential marginal tax rates, Alesina et al. (2005) argue that social norms determine equilibrium hours, distinguishing cultures that prioritize income from those that prioritize leisure.

If work-ethic norms create stronger social-compliance pressure to demonstrate dedication through long hours, overwork may be perceived as a virtuous signal or a necessary duty. In French-speaking regions, where cultural baselines place greater value on the protection of private time (Brügger et al., 2009; Ashforth et al., 2000; Clark, 2000), the same level of overwork may be experienced as a costly violation of the social contract. Cross-national evidence supports this reasoning: Falk et al. (2018) document substantial variation in economic preferences (including patience and willingness to work) across 76 countries, while Hamermesh and Lee (2007) show that cross-country differences in subjective time stress reflect genuine preference heterogeneity rather than mere "yuppie kvetch." In a meta-analysis of 332 studies across 58 countries, Allen et al. (2020) find that cultural values moderate the wellbeing consequences of work-family conflict. The disutility of an extra hour is higher in the French-speaking region not because the work is harder, but because it conflicts more sharply with the prevailing cultural schema of a good life.

Our contribution is to test this hypothesis using within-country variation. Rather than comparing countries with different labor laws, we compare individuals who face the same legal and economic incentives but differ in the cultural lens through which they evaluate the effort–reward bargain.

[6] However, whether income or hours constitutes the operative reference point remains contested (Farber, 2008; Crawford and Meng, 2011).

2.4 Hours Constraints and Behavioral Adjustment

A central question in labor economics is whether workers who experience a mismatch between actual and desired hours can adjust. The canonical model assumes that workers freely choose hours across jobs, so that job mobility resolves mismatches (Hamermesh, 1999). In practice, frictions are large. Knaus and Otterbach (2019) find that even among job movers in the German Socio-Economic Panel, resolution rates of work-hour mismatches remain below 40%, with a substantial share of adjustment occurring through changes in desired hours rather than actual hours. Workers adapt preferences to constraints rather than the reverse. Booth and van Ours (2013) find that partnered women in part-time work in the Netherlands report high job satisfaction and low desire to change hours, suggesting that part-time employment is a stable equilibrium rather than a transitional phase toward full-time work.

Two perspectives explain this inertia. Chetty (2012) shows that when optimization frictions exist (adjustment costs, inattention, status quo bias), the utility cost of failing to adjust is often less than 1% of earnings; small frictions rationalize large behavioral inertia. Pencavel (2016) argues that observed hours reflect employer demand at least as much as worker preferences, implying that hours are often set by firms rather than freely chosen by workers. Together, these results predict that even a large cultural amplification of overwork's psychological cost may fall below the threshold needed to trigger costly behavioral adjustment.

The behavioral response to overwork may also depend on the cultural meaning attached to the mismatch. Avgoustaki and Cañibano (2020) show that the well-being consequences of long hours depend critically on whether the effort is intrinsically or extrinsically motivated, suggesting that the same objective hours can generate very different psychological costs depending on how they are interpreted. Ichino and Maggi (2000) provide evidence that regional cultural norms within Italy predict absenteeism differentials within the same firm, demonstrating that work effort norms vary systematically within a country and affect observable behavior. The epidemiological approach (studying immigrants to isolate culture from institutions) has been applied to labor supply by Fernández and Fogli (2009), who show that source-country female labor-force participation rates predict immigrant women's work decisions in the U.S., confirming a causal role for cultural transmission.

2.5 The Röstigraben as a Natural Experiment

The Swiss linguistic border has emerged as one of the cleanest natural experiments for studying cultural effects on economic outcomes. Eugster et al. (2017) exploit the sharp change in survey-reported work attitudes at the French-German border to show that Romance-language speakers search for work almost seven weeks (22%) longer than their German-speaking neighbors, despite facing very similar labor-market conditions and identical unemployment insurance rules. Eugster et al. (2011) document that demand for redistributive social insurance is substantially lower in German-speaking communities. Cattaneo and Winkelmann (2005) show that the Swiss labor market is well integrated across the language border, with no significant earning differentials after controlling for selection, confirming that the border captures cultural rather than economic differences. Stutzer and Lalive (2004) demonstrate that stronger social work norms on the German-speaking side accelerate job search but also amplify the well-being costs of unemployment (unemployed individuals report lower life satisfaction in communities with stronger work norms), providing direct evidence that cultural attitudes at the Röstigraben shape both labor market behavior and subjective well-being.

Figure 1 provides direct evidence of the cultural divide on work-related issues. Eugster et al. (2017) similarly use referendum results to document sharp discontinuities in policy preferences at the language border; we replicate their approach with four more recent referenda directly

related to labor policy (the 1:12 Initiative on executive pay, 2013; the minimum wage initiative, 2014; the six-weeks holiday initiative, 2012; and the basic income initiative, 2016), sourced from the Swissvotes database (Année Politique Suisse, 2024), and estimate a sharp regression discontinuity at the linguistic border. In all four cases, support for the worker-protective position jumps discontinuously on the French-speaking side, with RDD estimates ranging from 4.0 to 8.2 percentage points (all $p < 0.01$). These voting patterns confirm that the Röstigraben captures deep-seated differences in attitudes toward labor regulation and the work-leisure tradeoff, precisely the cultural channel our analysis exploits. We use this variation not as a spatial regression discontinuity (the SHP lacks municipality geocodes for most observations) but as a source of cross-sectional cultural heterogeneity combined with within-person panel variation. Our design documents a robust conditional panel fact, that the within-person overwork slope differs systematically by language region, rather than satisfying the full identification standard of a geographic discontinuity.

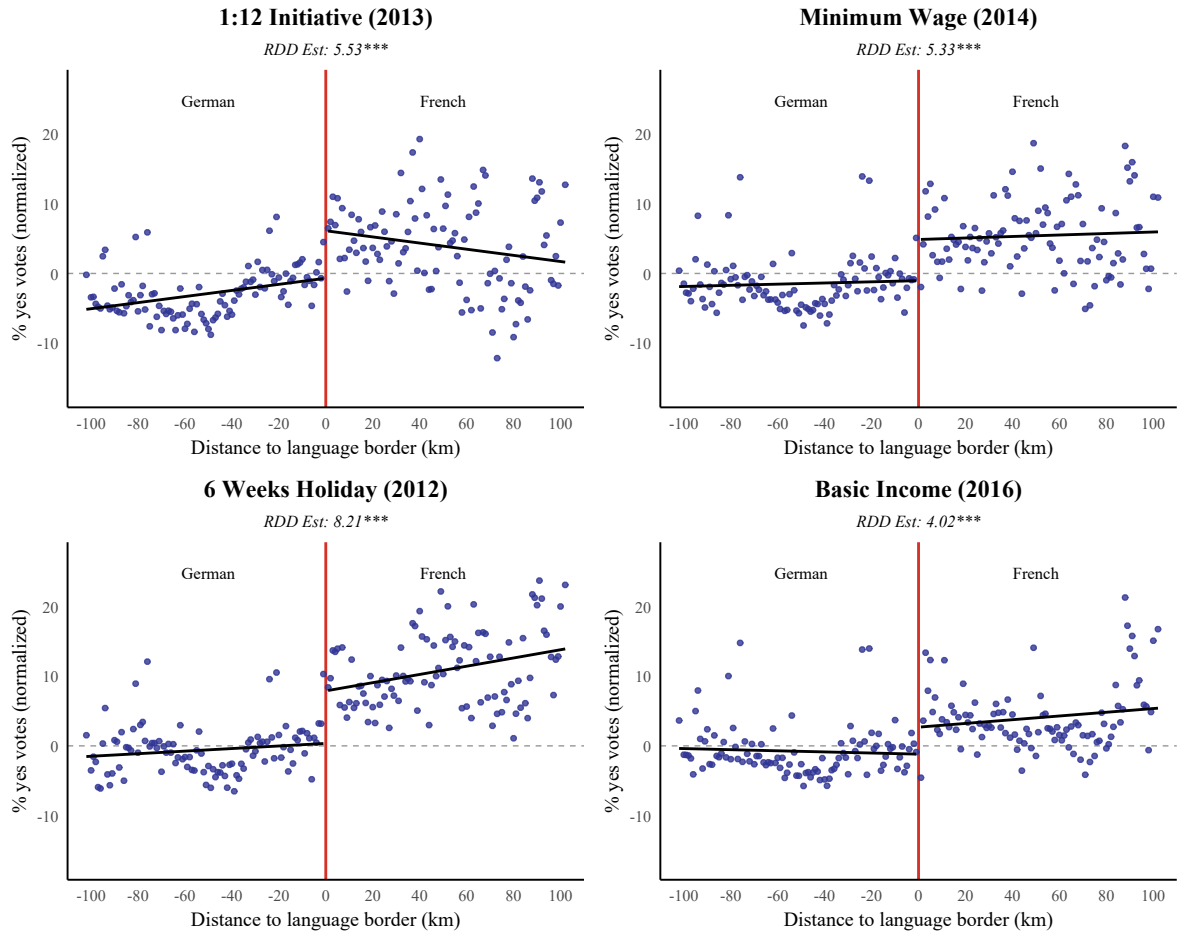


Figure 1: Regression discontinuity in municipality-level voting on labor-related referenda at the Swiss linguistic border, following the RDD visualization approach in Eugster et al. (2017). Each panel plots the share of "yes" votes against distance to the French-German language border (km), with negative distances denoting the German-speaking side. Observations binned by 1 km intervals, weighted by municipality population. Lines show local linear fits. RDD estimates and significance levels reported in subtitles. Data: Swissvotes (Année Politique Suisse, 2024); geographic boundaries and municipality language assignments from swisstopo (Federal Office of Topography , swisstopo).

3 Conceptual Framework

We develop a framework of reference-dependent labor supply in which cultural norms interact with the salience of employment contracts. The model combines three ideas: reference-dependent preferences (Kőszegi and Rabin, 2006), contracts as reference points (Hart and Moore, 2008), and salience-driven attention (Bordalo et al., 2012, 2013). The central contribution is contractual salience: the psychological prominence of the work-leisure boundary varies with contract type, and culture amplifies this variation. The framework generates four testable predictions that organize the empirical analysis.

3.1 Setup

A worker i in cultural group $c \in \{F, G\}$ (French, German) holds a contract specifying hours $\bar{h}_i$ and a base wage w. The employer sets actual hours h_i, which may exceed $\bar{h}_i$.[7] Overtime is compensated at rate $\phi \geq 0$. Define $\Delta_i \equiv h_i - \bar{h}_i$ as the hours gap.

3.2 Contractual Salience

Not all contracts are equally prominent as reference points. A full-time contract ($\bar{h} \approx 42$) specifies the default arrangement, the unmarked case. A part-time contract ($\bar{h} < 35$) is a deliberate departure from the default: an explicit signal that the worker values limited hours. Following the salience literature (Bordalo et al., 2012, 2013), attributes that deviate from the norm attract disproportionate attention. Applied to employment contracts: the more a contract departs from the full-time default, the more psychologically prominent the contractual boundary becomes.

We capture this with a salience function $\sigma(\bar{h}_i) \in [0, 1]$, weakly decreasing in contractual hours:

$$\sigma(\bar{h}_{\text{PT}}) \;>\; \sigma(\bar{h}_{\text{FT}}) \;\geq\; 0. \tag{1}$$

σ measures how prominently the contract marks the work-leisure boundary. When σ is high, contractual hours serve as a vivid psychological benchmark; when σ is low, the contract fades into the background. A second implication is that σ applies specifically to formal contractual commitments. Habitual or "reference" hours (what a worker normally works, as opposed to what the contract stipulates) are informal and self-defined: they carry no legal weight and no explicit signal of a chosen work-leisure allocation. Because they lack the formality that confers psychological salience, deviations from habitual hours generate little reference-dependent disutility. A worker who exceeds their usual hours may feel busy; a worker who exceeds their *contractual* hours experiences a boundary breach.

3.3 Preferences

The worker's utility is

$$U_i = u(y_i) \;-\; v(h_i) \;-\; \varphi_c(\Delta_i,\, \sigma_i), \tag{2}$$

where $u(\cdot)$ is increasing and concave (consumption utility), $v(\cdot)$ is increasing and convex (disutility of labor), and φ_c is a reference-dependent penalty. The contract-as-reference-point interpretation follows Hart and Moore (2008); the piecewise-linear form follows the prospect-theory

[7] This reflects the evidence on hours constraints in European labor markets (Hamermesh, 1999; Knaus and Otterbach, 2019). Even in relatively flexible markets, within-job hours adjustments are limited.

tradition (Kahneman and Tversky, 1979; Tversky and Kahneman, 1991; Kőszegi and Rabin, 2006):

$$\varphi_c(\Delta, \sigma) = \begin{cases} \lambda_c \cdot \sigma \cdot \Delta & \text{if } \Delta > 0 \quad \text{(overwork)} \\ \eta \cdot \Delta & \text{if } \Delta \leq 0 \quad \text{(underwork)} \end{cases} \tag{3}$$

For overwork ($\Delta > 0$), the penalty is the product of three terms: a cultural parameter λ_c, contractual salience σ, and the hours gap Δ. The cultural assumption is:

$$\lambda_G > \lambda_F \geq 0. \tag{4}$$

German-speaking norms, which place a higher premium on contractual compliance and the protection of leisure time (the *Feierabend* principle), generate a higher marginal penalty per hour of contract violation (Brügger et al., 2009; Eugster et al., 2017). For underwork ($\Delta \leq 0$), the gain parameter η carries no cultural subscript and no salience weight: culture and salience operate asymmetrically, amplifying only the cost of overwork. The intuition is that working fewer hours than the contract specifies does not intrude on private life (it expands it), so the cultural boundary-protection norm is not activated.

3.4 The Cultural–Salience Interaction

The model's key mechanism is the interaction between culture and contractual salience. The marginal cultural excess cost of overwork is:

$$(\lambda_G - \lambda_F) \cdot \sigma(\bar{h}_i) \cdot \Delta_i. \tag{5}$$

This product implies that the cultural wedge between German and French speakers scales with contractual salience. For full-time workers (σ low), the cultural difference in overwork costs is attenuated; for part-time workers (σ high), it is amplified. The same cultural parameter $\lambda_G - \lambda_F$ generates different observable effects depending on contract type, a feature that distinguishes this framework from a standard reference-dependent model with a uniform cultural shift.

Our empirical outcomes (exhaustion, work-life interference, difficulty disconnecting) are proxies for the latent disutility $v(h_i) + \varphi_c(\Delta_i, \sigma_i)$. For overworkers ($\Delta > 0$), substituting equation (3) into this expression and taking a first-order linear approximation yields the regression equation:

$$Y_{it} \approx \beta_1 \Delta_{it} + \beta_3 (\Delta_{it} \times \mathbf{1}_F) + \alpha_i + \delta_t + \varepsilon_{it},$$

where Y_{it} denotes the burnout or satisfaction outcome (denoted B_{it} in the model derivation above), β_1 captures the baseline disutility of overwork for German-speaking workers (combining v' and $\lambda_G \cdot \sigma$), and β_3 estimates $(\lambda_F - \lambda_G) \cdot \bar{\sigma}$ in outcome-scale units, that is, the cultural differential in the marginal cost of overwork weighted by sample-average contractual salience. Because $\lambda_G > \lambda_F$, the model predicts $\beta_3 < 0$: French-speaking workers experience *less* disutility per hour of overwork than the German baseline. The main effect of $\mathbf{1}_F$ (i.e., β_2) is absorbed by the individual fixed effects α_i. Note that β_3 identifies the *product* of the cultural gap and salience; the contract-type heterogeneity analysis in Section 5.5 separately identifies the role of σ. This is the coefficient of interest in equation (8).

3.5 Predictions

Prediction 1 (Cultural moderation of overwork costs). For $\Delta > 0$, $\lambda_G > \lambda_F$ implies that each hour of overwork generates more disutility for German-speaking workers. The interaction

coefficient β_3 in equation (8) is therefore negative: French speakers experience less WLI per hour of overwork than the German baseline.

Prediction 2 (Contract-type heterogeneity). From equation (5), the cultural interaction scales with $\sigma(\bar{h}_i)$. Part-time workers should show a large cultural interaction; full-time workers, a small or zero one.

Prediction 3 (Asymmetry at the contract). For $\Delta \leq 0$, the penalty $\eta \cdot \Delta$ carries no cultural subscript. The cultural interaction should vanish for underwork, implying a kink in the burnout–hours relationship at $\Delta = 0$.

Prediction 4 (Contractual vs. habitual reference points). Salience σ attaches to formal contractual hours. Habitual hours (what the worker "normally" works) lack the formality that confers salience. Deviations from habitual hours should generate no cultural interaction.

3.6 Welfare Implications

The cultural welfare wedge for a German-speaking overworker is:

$$\Delta W_i = (\lambda_G - \lambda_F) \cdot \sigma(\bar{h}_i) \cdot (h_i - \bar{h}_i) \quad \text{for } h_i > \bar{h}_i. \tag{6}$$

This wedge is invisible in standard market data: it does not manifest in wages, hours, or quit rates, and is observable only through subjective well-being. The concentration of the effect among part-time workers implies that the welfare cost is borne disproportionately by German-speaking workers who have most explicitly chosen to protect the work-leisure boundary, precisely those for whom overwork is most salient.

The welfare interpretation involves normative ambiguity: whether the higher cultural cost represents a genuine preference (the boundary matters intrinsically) or a culturally induced bias. Reck and Seibold (2023) show that this ambiguity does not affect the *sign* of welfare effects of changing reference points; it affects only the channel through which welfare operates (direct effects when reference dependence is normative versus behavioral effects when it is a bias). Applied to our setting, the cultural welfare wedge in equation (6) is robustly signed regardless of whether $\lambda_G > \lambda_F$ reflects normative preferences or culturally induced loss aversion.

Two caveats qualify the welfare interpretation. First, if the *sensitivity* of burnout reporting to overwork varies culturally (an interaction-level reporting bias), then $\hat{\beta}_3$ overstates the true welfare wedge; we discuss this concern and the evidence against it in Section 4. Second, the model is deliberately stylized: it takes hours as given and does not endogenize employer behavior or equilibrium sorting. A richer framework would incorporate these margins.

3.7 Calibration

The contract-type heterogeneity in Section 5.5 permits a simple calibration. The model predicts $|\hat{\beta}_3^{\text{PT}}| = (\lambda_G - \lambda_F) \cdot \sigma_{\text{PT}}$ and $|\hat{\beta}_3^{\text{FT}}| = (\lambda_G - \lambda_F) \cdot \sigma_{\text{FT}}$. Using the estimated coefficients for work-life interference ($|\hat{\beta}_3| = 0.038$ for part-time, 0.010 for full-time; the full-time coefficient is not statistically significant at conventional levels, $p = 0.337$), the implied salience ratio is $\sigma_{\text{FT}}/\sigma_{\text{PT}} \approx 0.27$: full-time contracts have roughly one-quarter the psychological salience of part-time contracts as reference points. The cultural welfare wedge (equation (6)) for the average German-speaking part-time overworker is $|\hat{\beta}_3^{\text{PT}}| \times \bar{\Delta} \approx 0.038 \times 5.4 \approx 0.21$ points on the 0–10 work-life interference scale, about 0.12 within-person standard deviations. While modest in absolute terms, this corresponds to a 152% amplification of the baseline overwork cost experienced by French-speaking workers. BFS data show that approximately 570,000 workers

in the German-speaking Grossregionen hold part-time contracts, roughly 35% of regional employment, so the welfare wedge applies to a substantial population. Cultural norms appear to inflate the welfare cost of boundary violations for the affected workforce, though this calibration is an order-of-magnitude exercise: it depends on the assumption that $\hat{\beta}_3$ reflects genuine preference differences rather than differential reporting sensitivity, a concern we discuss and partially bound in Section 4.

4 Data and Methodology

4.1 The Swiss Household Panel

Our data come from the Swiss Household Panel (SHP), a nationally representative longitudinal survey administered annually since 1999 by the Swiss Centre of Expertise in the Social Sciences (FORS). We use all 25 available waves (1999–2023), covering the period from the introduction of the survey through the post-pandemic recovery. The SHP interviews all adult members of selected households, re-tracking them even if they move or change household composition, yielding a genuine panel at the individual level (Tillmann et al., 2022).

4.2 Sample Construction

We restrict the sample to individuals who satisfy three conditions: (i) aged 18–65 at the time of interview, (ii) currently employed[8], and (iii) interviewed in German or French.[9] Observations with non-positive values for all key variables are recoded to missing following the SHP's standard convention.[10]

The resulting panel consists of around 84,000 person-year observations from German- and French-speaking employed individuals observed over 1999–2023 (approximately 72,000 after outcome-specific item non-response). The SHP oversamples French- and Italian-speaking regions to ensure representativeness; our unweighted sample is approximately 71% German-speaking, consistent with the German-speaking population share of roughly 73% (after excluding Italian speakers). All regressions use the SHP's cross-sectional calibrated person weight (`wicss`), which adjusts for the oversampling of linguistic minorities, unit non-response, and demographic calibration to population margins.[11] Because our identification relies on individual fixed effects and within-person variation, the interaction coefficient $\hat{\beta}_3$ is identified from within-person changes in overwork regardless of the weighting scheme; results are also robust to unweighted estimation. Table 1 presents descriptive statistics by language region.

[8] i.e., working status code 1, 2, or 3, corresponding to full-time, part-time, or irregular employment.

[9] We exclude Italian-speaking respondents (approximately 5% of the employed sample) because their small sample size limits the statistical power of within-person estimators. Romansh speakers in Graubünden are interviewed in German or Italian by the SHP and thus enter the German-speaking sample or are dropped with Italian speakers, respectively.

[10] The SHP uses negative codes to indicate missing information: -1 = "don't know," -2 = "no answer," -3 = "inapplicable," and -4 through -8 for various technical errors.

[11] The weight combines inverse inclusion probabilities, non-response corrections, and post-stratification calibration to Swiss population totals by age, sex, nationality, and major region. We use cross-sectional rather than longitudinal weights because our sample pools all four SHP cohorts (1999, 2004, 2013, 2020), each with a different longitudinal base year; cross-sectional weights ensure that each wave's weighted sample reflects the contemporaneous population regardless of cohort entry. Results are robust to using cohort-specific longitudinal weights (`wilss`): all core interaction coefficients retain the same sign and significance level, with point estimates differing by less than 19%. See the SHP weighting technical report for details.

4.3 Key Variables

Our primary independent variable is the hours gap, defined as the difference between actual and contractual weekly working hours:

$$\text{Hours Gap}_{it} = \text{Actual Hours}_{it} - \text{Contractual Hours}_{it}, \tag{7}$$

where Actual Hours is the number of hours worked per week and Contractual Hours is the contractual hours per week. A positive gap indicates overwork: the individual works more than their contract specifies. Our key treatment variable is a binary indicator French_i equal to one if the individual's interview language is French, and zero if it is German.

We use five well-being outcomes as dependent variables, all measured on 0–10 scales. Life satisfaction, job satisfaction, and free-time satisfaction capture domain-specific evaluations of well-being. Work-life interference measures the extent to which work interferes with private activities or family obligations (0 = not at all, 10 = very strongly), and post-work exhaustion captures how often the respondent is too exhausted after work to do things they would like. All regressions include age, age squared, a female indicator, and a dummy for having co-resident children as controls.

4.4 Descriptive Statistics

Table 1: Descriptive Statistics by Language Region

Variable	French	German
Panel A: Demographics & Working Hours		
N	24,023	59,711
Individuals	4,818	12,202
Mean Age	42.0	42.6
Female (%)	51.0	49.7
Tertiary Edu (%)	35.9	30.0
Contractual Hrs	33.7	34.1
Actual Hrs	36.4	36.4
Hours Gap	2.7	2.3
Wants Less (%)	44.9	47.7
Panel B: Outcome Variables		
Life Sat (0-10)	7.86	8.07
Job Sat (0-10)	7.65	7.91
Exhaustion (0-10)	4.80	4.41
Work-Life Int. (0-10)	4.20	3.81
Disconnect (0-10)	3.74	3.12
Free Time Sat. (0-10)	6.28	6.71

The two regions differ in sample composition but not dramatically. German-speaking respondents are slightly younger and more likely to hold tertiary degrees. The mean hours gap is positive in both regions: workers across French and German Switzerland work approximately 2–3 hours more per week than their contracts specify. In the full analysis sample, 46.9% of observations are overwork episodes (hours gap > 0; mean gap $+5.7$ hours), 49.5% are exact adherence (hours gap $= 0$), and 3.6% are underwork episodes (hours gap < 0; mean gap -7.8 hours).[12] The work-life interference scale (0–10) has a standard deviation of 2.64, right-skewed (skewness = 0.09), with 15.1% of responses at the floor (0) and 1.3% at the ceiling (10). The share

[12]The large mass at exactly zero may partly reflect rounding in self-reported hours: if workers report both actual and contractual hours to the nearest 5 or 10 hours, a mechanical spike at zero is generated even when true overwork exists. We verified that the kink-design and bunching results are robust to excluding observations where both actual and contractual hours are multiples of five, which removes roughly 35% of the exact-zero observations; the regression interaction coefficient changes by less than 5%.

wanting fewer hours is broadly similar across regions, though the raw means mask important temporal dynamics that we explore below.

Figure 2 plots the evolution of two key outcomes by language region. Panel A shows the share of workers with a positive hours gap. Panel B shows mean life satisfaction. Both series exhibit visible disruption around 2020, but the trajectories vary by region.

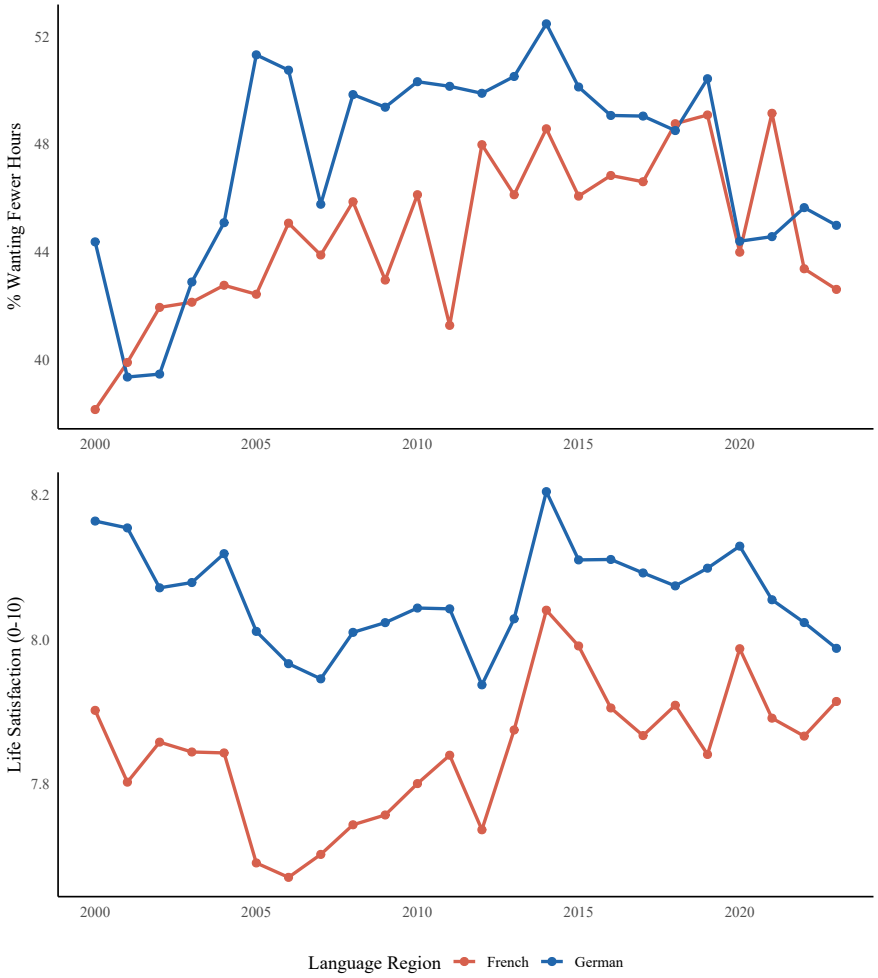


Figure 2: Trends in overwork preference and life satisfaction by language region, 1999–2023.

Figure 3 shows the geographic distribution of our key variables. Panel A maps Switzerland's language regions, highlighting the Röstigraben, a division that has been stable throughout the panel period (1999–2023). Panel B maps the post-pandemic (2021–2023) cantonal average of the overwork indicator; we focus on recent years because canton-level identifiers are most reliably available in later waves, though the cross-regional patterns are qualitatively similar across the full panel.

4.5 Empirical Strategy

Our identification exploits within-person variation in overwork and its interaction with a time-invariant cultural indicator. Unlike the spatial regression discontinuity designs used in prior Röstigraben studies (Eugster et al., 2017, 2011; Brügger et al., 2009), we compare all French- and German-speaking workers rather than restricting to those near the border. This is because the SHP (our source for longitudinal hours and well-being data) does not provide municipality-

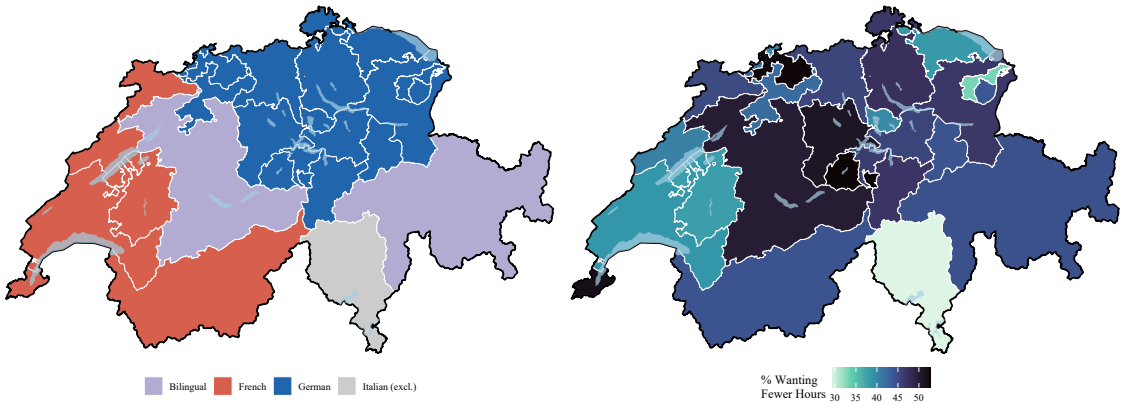


Figure 3: Maps of Switzerland: language regions (A) and post-pandemic overwork prevalence by canton (B). Ticino (Italian-speaking) is excluded from our analysis; Romansh speakers in Graubünden are interviewed in German or Italian by the SHP and thus enter the German-speaking sample or are dropped with Italian speakers, respectively. Geographic data: swisstopo (Federal Office of Topography , swisstopo).

level geocodes for most observations, precluding a geographic RDD.[13] This is not a difference-in-differences design: the linguistic border provides cross-sectional variation in cultural exposure (French_i), and $\hat{\beta}_3$ is identified from the conditional within-person slope difference across cultural groups. The core identifying assumption is: *conditional on individual fixed effects and year fixed effects, changes in the hours gap are uncorrelated with changes in unobserved factors that differentially affect well-being for French versus German speakers*. We estimate:

$$Y_{it} = \beta_1 \cdot \text{HoursGap}_{it} + \beta_3 \cdot (\text{HoursGap}_{it} \times \text{French}_i) + \mathbf{X}'_{it}\gamma + \alpha_i + \delta_t + \varepsilon_{it}, \tag{8}$$

where Y_{it} is an outcome capturing burnout (exhaustion, work-life interference, difficulty disconnecting, work stress) or domain-specific satisfaction (free time, work conditions, job, life). The individual fixed effects α_i absorb all time-invariant characteristics, including the main effect of French_i, and the year fixed effects δ_t absorb common shocks. The coefficient of interest is $\hat{\beta}_3$: the differential effect of an additional hour of overwork for French-speaking workers relative to the German-speaking baseline. A negative $\hat{\beta}_3$ on exhaustion or work-life interference, or a positive $\hat{\beta}_3$ on satisfaction, indicates that overwork is *less* psychologically costly for French-speaking workers; equivalently, that German-speaking workers bear higher costs per hour of overwork.

All standard errors are clustered at the individual level to account for serial correlation within person-year panels. The number of clusters ranges from G = 16,099 (work-life interference) to G = 16,119 (difficulty disconnecting) across the main outcomes.[14]

Two identification concerns deserve explicit discussion. First, the hours gap is self-reported and potentially endogenous: workers experiencing burnout may perceive their hours differently, and within-person changes in the hours gap are driven by job transitions and employer decisions that may independently affect well-being. We address this concern through coeffi-

[13] A border-canton restriction (Appendix Table A.7) shows the WLI interaction remains significant ($\hat{\beta}_3 = -0.030$, $p = 0.003$), somewhat attenuated relative to the full-sample estimate, consistent with reduced cultural contrast near bilingual communities.

[14] As robustness checks we report results with two-way clustering (individual $\times$ year) and canton-level clustering (26 cantons); all yield qualitatively identical conclusions. See Table B.3.

cient stability analysis: $\hat{\beta}_3$ is virtually unchanged as we move from no controls to the full set of time-varying controls including occupation, supervisor status, and education (Appendix Table A.2), suggesting limited selection on observables. Formally, an Oster (2019) coefficient stability analysis shows that the unrounded interaction coefficient moves by less than 0.001 as controls are added (e.g., the WLI interaction shifts from -0.03381 to -0.03316 between the no-controls and baseline-controls specifications), yielding mechanically extreme δ^* values. The rounded coefficients in Table B.1 obscure this movement: both round to -0.038, making it appear that the coefficient does not move at all. The δ^* statistic is well above the conventional threshold of 1 for all three outcomes, indicating that the interaction is stable as controls are added. We interpret this as evidence against selection on observables, though we note that coefficient stability does not rule out time-varying unobservables that are orthogonal to the controls. Second, cultural differences in reporting burnout, particularly in the *sensitivity* of reporting to overwork (an interaction-level bias), cannot be fully ruled out by individual fixed effects. The domain-specificity of our results (effects on work-life interference but not on health satisfaction, $p = 0.88$) argues against a general reporting-bias explanation: if German speakers simply report all outcomes more negatively when overworked, the effect should extend uniformly across domains. The work-stress interaction is positive ($\hat{\beta}_3 = +0.003$, $p < 0.001$), opposite in sign to the WLI interaction, further undermining a pure reporting story: a general German negative-reporting bias would predict the same direction ($\hat{\beta}_3 < 0$) for all outcomes, but the work-stress interaction breaks this pattern. We regard domain-specific reporting sensitivity as the strongest remaining identification concern: while the domain-specificity evidence and the opposite-sign work-stress result argue against a general reporting bias, we cannot definitively exclude interaction-level differences in how French and German speakers report work-leisure-boundary outcomes specifically. To bound this concern, suppose German-speaking workers' reporting sensitivity for boundary outcomes is k times that of French speakers, so the observed German baseline slope $\hat{\beta}_1 = 0.063$ overstates the true slope by factor k, while the French slope is correctly measured. The observed French slope is $\hat{\beta}_1 + \hat{\beta}_3 = 0.063 + (-0.038) = 0.025$. Under this assumption, the true cultural interaction is $0.025 - 0.063/k$. At $k = 1.5$, the implied true interaction is -0.017; at $k = 2$, it is -0.007, substantially attenuated from the observed -0.038. The cultural interaction changes sign at $k = 0.063/0.025 = 2.52$: reporting sensitivity of German speakers roughly 2.5 times that of French speakers would eliminate the estimated cultural penalty entirely. We regard $k > 2$ as implausible given the domain-specificity evidence (health satisfaction interaction is null, $p = 0.88$) and the opposite-sign work-stress result, but acknowledge that the true cultural interaction is likely smaller than the point estimate if any differential reporting exists.

A second measurement concern is that recall error in self-reported hours generates classical measurement error in the hours gap. If non-differential by language group (plausible), this attenuates both $\hat{\beta}_1$ and $|\hat{\beta}_3|$ toward zero, making the WLI and exhaustion estimates conservative lower bounds on the true cultural interaction. We note, however, that attenuation bias mechanically increases the probability of null findings, so the behavioral nulls in Section 5 could partly reflect measurement-error-induced power loss rather than a true absence of differential adjustment. The cross-equation evidence mitigates this concern: the WLI interaction is estimated with high precision ($t > 3.5$) in the same data and using the same hours-gap variable, suggesting adequate signal-to-noise to detect effects of the magnitudes documented here. We cannot rule out that a behavioral interaction, if it exists, is attenuated below our detection threshold. Weighted and unweighted estimates are nearly identical ($|\hat{\beta}_3| = 0.038$ and 0.033 respectively, both $p < 0.001$), and canton-level clustering (26 clusters) yields the same significance ($p < 0.001$), confirming that neither the weighting strategy nor the clustering choice drives the main results. As a further check on inference, we perform a permutation test by randomly reassigning language labels across individuals ($B = 500$ permutations), building a

null distribution under the hypothesis that language region is unrelated to the within-person overwork slope. The observed $\hat{\beta}_3 = -0.038$ lies far in the lower tail (two-sided permutation $p < 0.002$; permutation null: mean $= 0.000$, SD $= 0.007$), confirming that the estimated cultural interaction cannot plausibly arise by chance under random language assignment.

A third measurement concern, raised in the review process, is culturally differential recall bias in the independent variable: if German-speaking workers, with stronger boundary norms, track their hours more precisely, the measured hours gap could be less noisy for German speakers, potentially inflating the observed interaction. We test this hypothesis directly. Table D.7 reports several diagnostics. French-speaking workers show *higher* within-person hours-gap variance (SD $= 3.51$ vs. 2.72, $p < 0.001$), *higher* mean absolute hours gap (3.29 vs. 2.83 hours, $p < 0.001$), and substantially more rounding of reported hours to multiples of five (44.3% vs. 37.7% for actual hours; 27.3% vs. 18.3% for both actual and contractual hours). All three diagnostics point in the *opposite* direction from the recall-bias hypothesis: French speakers, not German speakers, exhibit noisier and more rounded hours reporting. If anything, measurement error is larger in the French-speaking sample, which would attenuate the French slope ($\hat{\beta}_1 + \hat{\beta}_3$) toward zero and *inflate* the absolute interaction $|\hat{\beta}_3|$. While this pattern does not rule out all forms of differential measurement error (e.g., mean-shifting rather than variance-shifting bias), it provides no support for the specific channel of German speakers reporting more precisely.

5 Results

5.1 Descriptive Patterns

Panel B of Table 1 reports descriptive means of the mechanism and satisfaction variables by language region. French-speaking workers report higher raw means on exhaustion (4.80 vs. 4.41), work-life interference (4.20 vs. 3.81), and difficulty disconnecting (3.74 vs. 3.12) than German-speaking workers. This unconditional ordering reflects, in part, that French speakers face larger mean hours gaps (2.7 vs. 2.3 hours above contract on average); the regression analysis reveals that this cross-sectional pattern masks a difference in the *marginal* psychological cost of overwork. The regression interaction captures the differential *slope*: how much worse each additional hour of overwork is for German speakers relative to French speakers, holding the individual's baseline level of burnout fixed via individual fixed effects. Figure 2 shows that the share of workers wanting fewer hours rose in both regions after 2020, while mean life satisfaction declined modestly and convergently. Both series return toward pre-pandemic levels by 2022–23.

5.2 Burnout and Boundary-Setting Mechanisms

Table 2 presents the core results. We estimate Equation (8) using three outcomes that capture exhaustion and boundary-setting: post-work exhaustion, work-life interference, and difficulty disconnecting.

The interaction term $\hat{\beta}_3$ (Hours Gap $\times$ French) is significant for all three outcomes in the baseline specification, with a consistently *negative* sign: French-speaking workers experience less psychological cost per hour of overwork than the German baseline. **Work-life interference** (column 3) is our primary outcome: the interaction of -0.038 ($p < 0.001$) indicates that overwork spills over into private life far more severely for German speakers (baseline $\hat{\beta}_1 = 0.063$), and this result holds across all robustness specifications (Section 5.4). **Post-work exhaustion** (column 2) is a secondary outcome: each additional hour raises exhaustion by 0.021 points more for German speakers relative to French speakers ($p < 0.001$, approximately doubling the

Table 2: Burnout and Boundary-Setting Mechanisms

	Exhaustion	Work-Life Int.	Disconnect
Hours Gap	0.042***	0.063***	0.039***
	(0.003)	(0.004)	(0.003)
French Region	0.173	0.251	0.079
	(0.259)	(0.331)	(0.386)
Hours Gap × French	−0.021***	−0.038***	−0.009*
	(0.005)	(0.006)	(0.005)
Observations	71,985	71,950	72,022
R^2	0.576	0.533	0.625
FE: Individual	X	X	X
FE: Year	X	X	X

* p<0.1, ** p<0.05, *** p<0.01. Individual and year fixed effects. Standard errors clustered at the individual level. Controls: age, age^2, female, has children. Exhaustion, work-life interference, and disconnecting are scaled 0–10.

French baseline), but this result attenuates substantially under outlier trimming and the low-contractual-hours restriction (see Section 5.4 and Appendix Table B.9); it should be treated as fragile corroborating evidence rather than a firmly established effect. **Difficulty disconnecting** (column 4) is a tertiary outcome: the cultural interaction ($\hat{\beta}_3 = -0.009$, $p = 0.071$) is borderline significant and not claimed as an established effect.

The magnitudes are modest in absolute terms. A worker transitioning from zero overwork to five additional hours per week (approximately the interquartile range of the hours gap) would experience a $5 \times 0.021 = 0.11$-point larger increase in exhaustion and a $5 \times 0.038 = 0.19$-point larger increase in work-life interference if German-speaking rather than French-speaking, on a 0–10 scale. The within-person standard deviation of work-life interference is 1.76, so the 5-hour cultural penalty corresponds to 0.11 within-person standard deviations, small in absolute terms. At mean overwork among overworkers (5.4 hours), the figure is 0.12 within-person standard deviations, the benchmark used elsewhere in the paper.

However, in proportional terms the cultural amplification is large. The German baseline slope $\hat{\beta}_1 = 0.063$ exceeds the French slope $\hat{\beta}_1 + \hat{\beta}_3 = 0.025$ by 152%: the marginal cost of an extra hour of overwork is more than twice as high for German-speaking workers as for French-speaking workers on average across the full sample. As shown in Section 5.5, this pooled figure is driven primarily by part-time workers: the part-time-specific interaction is large (-0.038, $p < 0.001$), while the full-time interaction approaches zero (-0.010, $p = 0.337$). At mean overwork levels among overworkers, the total predicted work-life interference effect is 152% higher for German speakers. The effect is thus small in absolute terms but represents a substantial proportional cultural amplification of the baseline overwork cost.

5.3 Satisfaction Outcomes

Table 3 extends the analysis to five domain-specific satisfaction measures. The cultural moderation is strongest for satisfaction with free time ($\hat{\beta}_3 = +0.029$, $p < 0.001$): each hour of overwork reduces free-time satisfaction by 0.029 points less for French-speaking workers (equivalently, German speakers experience 0.029 points more reduction per hour). Satisfaction with work conditions also shows a significant positive cultural interaction ($\hat{\beta}_3 = +0.014$, $p < 0.001$),

Table 3: The Effect of Overwork on Domain-Specific Satisfaction

	Life Sat.	Job Sat.	Free Time	Work Cond.	Work Amount
Hours Gap	−0.001	−0.003	−0.043***	−0.022***	−0.025***
	(0.001)	(0.002)	(0.003)	(0.002)	(0.003)
French Region	−0.213*	−0.069	−0.244	0.362	−0.122
	(0.121)	(0.234)	(0.315)	(0.235)	(0.320)
Hours Gap × French	0.004*	0.006*	0.029***	0.014***	0.004
	(0.002)	(0.003)	(0.005)	(0.004)	(0.004)
Observations	77,027	67,613	75,066	76,984	59,581
R^2	0.571	0.528	0.569	0.511	0.498
FE: Individual	X	X	X	X	X
FE: Year	X	X	X	X	X

* p<0.1, ** p<0.05, *** p<0.01. Individual and year fixed effects. Standard errors clustered at the individual level. Controls: age, age^2, female, has children. All outcomes scaled 0–10.

consistent with the same pattern: overwork erodes work-condition satisfaction more for German speakers. Both results survive Benjamini–Hochberg correction across the five outcomes ($p_{\text{BH}} < 0.05$ for each). Satisfaction with the amount of work ($\hat{\beta}_3 = +0.004$, $p = 0.325$) does not reach statistical significance.

Life satisfaction ($p = 0.076$) and job satisfaction ($p = 0.069$) fall below conventional significance thresholds and are therefore not claimed as established effects; together with the stronger domain-specific results, they are consistent with the view that the cultural penalty operates through boundary-specific channels rather than through broad reductions in global well-being.

The domain-specificity of these results is striking. Table A.1 reports the cultural interaction across a broad battery of additional outcomes. Health satisfaction ($p = 0.88$), financial satisfaction ($p = 0.13$), income satisfaction ($p = 0.53$), and work atmosphere ($p = 0.98$) all show null interactions, confirming that the cultural penalty is confined to the work-leisure boundary. The health satisfaction null is particularly informative for the reporting-bias concern: if German speakers simply report all outcomes more negatively when overworked, the effect should extend to health; it does not.

Work stress shows a *positive* cultural interaction ($\hat{\beta}_3 = +0.003$, $p < 0.001$; Table B.8), opposite in sign to the WLI and exhaustion interactions. The model in Section 3 offers a coherent interpretation. Work stress ("How often do you feel stressed at work?") measures within-job cognitive demands, not boundary violations. The cultural penalty $\lambda_G \cdot \sigma \cdot \Delta$ applies specifically to the intrusion of work into private life, not to in-job task load. German-speaking workers who overwork experience the hours as a boundary violation (increasing WLI) but do not necessarily experience the cognitive task load as more stressful. The opposite sign of the work-stress interaction argues against a general German reporting-bias explanation: a bias toward negative reporting would predict the same direction ($\hat{\beta}_3 < 0$) for all outcomes; the positive work-stress interaction breaks this pattern, consistent with the effect being domain-specific to the work-leisure boundary.

5.4 Robustness of the Main Results

Table A.2 subjects the exhaustion result to a battery of robustness checks. The interaction coefficient is stable when we add ISCO major-group occupation fixed effects (column 2: $\hat{\beta}_3 = -0.020$), and all additional time-varying controls simultaneously. Two-way clustering by individual and year yields similar standard errors. However, trimming extreme hours-gap values ($|\text{gap}| > 20$ hours, affecting 1.5% of observations) attenuates the coefficient to -0.018 ($p = 0.054$), indicating some sensitivity to outliers. An Oster (2019) coefficient stability analysis confirms that the interaction coefficient is stable as controls are added: the unrounded WLI interaction moves by less than 0.001 across specifications (Table B.1; the rounded values obscure this movement). The resulting δ^* values are well above the conventional threshold of 1 for all three outcomes, indicating limited selection on observables. Standard errors are clustered at the individual level (G = 16,099–16,119 clusters across outcomes). Appendix Table B.3 reports results with two-way clustering (individual $\times$ year), canton-level clustering, and unweighted estimation; all specifications yield the same qualitative conclusions. As a more demanding test, Appendix Table B.4 adds canton-by-year fixed effects, absorbing all region-specific annual shocks; the WLI interaction is virtually unchanged ($\hat{\beta}_3 = -0.038$, $p < 0.001$ under both specifications), ruling out region-specific economic shocks as an explanation. Investigation of the extreme hours-gap observations reveals that 18% of cases with $|\text{gap}| > 15$ involve very low contractual hours (< 15 hours/week), where even moderate actual hours generate mechanically large gaps; dropping observations with contractual hours below 10 (5% of the sample) attenuates the exhaustion coefficient ($\hat{\beta}_3 = -0.012$, $p = 0.084$; Table B.9). The work-life interference result is more robust: $|\hat{\beta}_3|$ remains significant across all specifications, including outlier trimming ($\hat{\beta}_3 = -0.028$, $p = 0.002$) and the principled contractual-hours trim ($\hat{\beta}_3 = -0.019$, $p = 0.012$). We regard work-life interference as our most reliable outcome.

A further concern is whether the cultural penalty reflects the hours *gap* specifically or a general cultural disutility for working longer absolute hours. Table D.2 addresses this by adding actual hours worked and its interaction with the French indicator as controls. Controlling for actual hours attenuates the interaction from -0.038 to -0.018 ($p = 0.001$), indicating that approximately 45% of the baseline estimate reflects the level of absolute hours rather than the deviation from contract. Adding the interaction of actual hours with French (column 3) yields $\hat{\beta}_3 = -0.021$ ($p < 0.001$): the cultural penalty on the hours gap remains significant and economically meaningful after absorbing any cultural differences in the disutility of long absolute hours, though the magnitude is roughly half the baseline estimate. The actual-hours interaction itself is absorbed by the individual fixed effects (collinear with the hours-gap interaction given that contractual hours vary little within person), confirming that the identifying variation comes from within-person deviations from contract rather than cross-sectional differences in hours levels.

Controlling for perceived unemployment risk (0–10 scale) produces zero attenuation of the interaction coefficient: $|\hat{\beta}_3| = 0.038$ with and without the control (Table A.3). Overwork does not differentially change perceived unemployment risk across language regions ($p = 0.63$). This rules out the hypothesis that the cultural interaction reflects differential labor market insecurity rather than cultural norms.

As a functional form check, we estimate a pooled ordered logit for work-life interference (scaled 0–10).[15] The ratio of the interaction to the main coefficient is 0.70 under ordered logit

[15]The pooled ordered logit does not include individual fixed effects (the incidental parameters problem prevents direct estimation); it controls for observable demographics but not for time-invariant individual heterogeneity. Results are therefore not directly comparable to the fixed-effects estimates in terms of levels, but the relative magnitude of the interaction versus main effect is informative about whether the OLS linear approximation distorts the cultural amplification ratio.

and 0.71 under pooled OLS, virtually identical, suggesting the linear approximation does not distort the relative magnitudes. A more direct test of functional form (whether the squared interaction term (Hours Gap$^2\times$ French) is significant) yields $p = 0.61$ (Table C.1), confirming linearity. A fractional logit specification and a log-transformed outcome both replicate the significant cultural interaction at $p < 0.001$, consistent with the linear fixed-effects baseline.

5.5 Contract-Type Heterogeneity and Selection

Table 4 splits the sample by contract type. The cultural moderation of work-life interference is concentrated among part-time workers (contractual hours < 35): $\hat{\beta}_3 = -0.038$ ($p < 0.001$) for part-time versus $\hat{\beta}_3 = -0.010$ ($p = 0.337$) for full-time. The part-time estimate coincides with the full-sample estimate (-0.038), indicating that the pooled result is dominated by the part-time subsample; the full-time interaction contributes negligibly. The German PT baseline slope is 0.064 and the French PT slope is 0.026 (a 152% difference); the German FT baseline is 0.072 and the French FT slope is 0.062 (not significantly different). The formal triple interaction does not reach conventional significance ($p = 0.202$; Table A.4), though the triple-interaction test is less powerful than the split-sample comparison. We therefore interpret the contract-type heterogeneity as suggestive split-sample evidence for the contractual-salience channel rather than a formally identified mechanism. An alternative interpretation is that part-time schedules are structurally more rigid (less employer discretion over hours), so the hours gap is more salient for institutional rather than purely cultural reasons; we cannot distinguish this schedule-rigidity channel from the cultural-salience channel with our data. A related concern is that a given hours-gap deviation (e.g., 5 hours) constitutes a proportionally larger shock for a part-time contract than for a full-time contract (25% of a 20-hour contract versus 12% of a 42-hour contract), so the part-time concentration could partly reflect a mechanical scale effect rather than differential salience. The null for full-time workers ($\hat{\beta}_3 = -0.010$, $p = 0.337$, 95% CI spanning zero) is informative: a near-zero coefficient for full-time workers is precisely what Prediction 2 requires. An a priori power calculation indicates that our sample achieves 80% power to detect a triple interaction of approximately 0.033, which is 87% of the observed part-time effect ($|\hat{\beta}_3^{\text{PT}}| = 0.038$). The 95% confidence interval for the triple interaction ($[-0.038, +0.008]$) thus cannot rule out a true difference as large as the observed part-time effect. We regard the split-sample evidence as more informative than the underpowered pooled test.

This effect is not driven by gender. Part-time work in Europe is heavily gendered and closely linked to child penalties (Kleven et al., 2019), so it is important to distinguish the contractual-salience channel from gendered norms about working time. Table 5 further splits the part-time and full-time subsamples by gender. Among part-time workers, German-speaking men show a larger cost differential than German-speaking women: the French interaction is $\hat{\beta}_3 = -0.054$ ($p < 0.001$) in the male part-time subsample and $\hat{\beta}_3 = -0.031$ ($p = 0.008$) in the female part-time subsample, though the formal triple interaction (Hours Gap $\times$ French $\times$ Female within part-time) is not statistically significant ($p = 0.275$). Both German-speaking men and women in part-time work experience significantly higher WLI per hour than their French-speaking counterparts. However, the pattern across contract types differs by gender. Full-time men show zero cultural moderation ($\hat{\beta}_3 = -0.002$, $p > 0.8$), consistent with contractual salience: the effect appears only when the contract explicitly marks the work-leisure boundary. Full-time women, by contrast, show a marginally significant interaction ($\hat{\beta}_3 = -0.032$, $p < 0.1$) similar in magnitude to part-time women (-0.031), suggesting that for women, the cultural moderation of overwork costs does not depend on contract type. The contractual-salience mechanism thus appears to operate primarily for men; for women, the cultural penalty is present regardless of whether the contract is part-time or full-time. A dose-response analysis by part-time intensity broadly reinforces this interpretation: $|\hat{\beta}_3| = 0.010$ ($p = 0.337$) for full-

Table 5: Work-Life Interference by Contract Type and Gender

	PT Women	PT Men	FT Women	FT Men
Hours Gap	0.066***	0.058***	0.090***	0.065***
	(0.008)	(0.010)	(0.010)	(0.008)
Hours Gap × French	−0.031***	−0.054***	−0.032*	−0.002
	(0.012)	(0.015)	(0.017)	(0.013)
Observations	21,366	4,592	13,212	30,494
R^2	0.533	0.598	0.563	0.556
FE: Individual	X	X	X	X
FE: Year	X	X	X	X

* p<0.1, ** p<0.05, *** p<0.01. Individual and year FE. SE clustered by individual. PT: contractual hours < 35. FT: contractual hours ≥ 35. Controls: age, age^2, has children.

time workers, 0.051 ($p = 0.033$) for workers at 60–99% of full-time hours, and 0.033 ($p = 0.011$) for workers below 60%.

Table 4: Full-Time vs. Part-Time Workers

	Exhaust. (FT)	Exhaust. (PT)	WLI (FT)	WLI (PT)	Disc. (FT)	Disc. (PT)
Hours Gap	0.053***	0.035***	0.072***	0.064***	0.044***	0.044***
	(0.004)	(0.006)	(0.006)	(0.007)	(0.005)	(0.006)
Hours Gap × French	−0.011	−0.009	−0.010	−0.038***	0.008	−0.007
	(0.009)	(0.009)	(0.011)	(0.009)	(0.009)	(0.009)
Observations	43,721	25,980	43,706	25,960	43,743	25,995
R^2	0.599	0.602	0.558	0.544	0.645	0.651
FE: Individual	X	X	X	X	X	X
FE: Year	X	X	X	X	X	X

* p<0.1, ** p<0.05, *** p<0.01. Individual and year FE. SE clustered by individual. FT: contractual hours ≥ 35. PT: contractual hours < 35. Controls: age, age^2, female, has children.

One concern is that the part-time concentration reflects selection: if German-speaking workers with strong boundary preferences disproportionately choose part-time contracts, the larger cultural interaction among part-time workers would capture selection rather than reference-point salience. We conduct four tests that collectively provide no support for this interpretation (Table 6).

First, German-speaking workers are not more likely to hold part-time contracts: in a linear probability model with year fixed effects and demographic controls, the coefficient on French is 0.012 ($p = 0.294$), meaning French speakers are slightly (but not significantly) more likely to work part-time. Part-time rates are 38.0% and 37.2% in German- and French-speaking regions, respectively. Second, among full-time workers in $t - 1$, overwork does not differentially predict transitions to part-time in t: the interaction Hours Gap$_{t-1}$ × French is essentially zero. Third, and most informatively, we identify 3,853 individuals observed in both full-time and part-time states during the panel. Restricting the triple-interaction regression to these contract-type switchers, the within-person triple interaction (Hours Gap × French × Part-Time) is −0.005 ($p = 0.743$), statistically insignificant. This within-switcher triple is too imprecise (SE = 0.015, $N = 3{,}853$) to be informative about the contractual-salience mechanism; we

Table 6: Selection into Part-Time Work

	P(Part-Time)	P(FT→PT)	P(FT→PT) FE	WLI (Switchers)
French	−0.012	0.003	0.000	0.319
	(0.009)	(0.005)	(0.050)	(0.401)
Hours Gap$_{t-1}$		−0.005***	−0.003***	
		(0.000)	(0.001)	
Hours Gap$_{t-1}$ × French		0.000	0.000	
		(0.001)	(0.001)	
Hours Gap				0.074***
				(0.007)
Hours Gap × French				−0.028**
				(0.013)
Hours Gap × Part-Time				−0.013
				(0.009)
Hours Gap × French × PT				−0.005
				(0.015)
Observations	80,523	40,826	38,209	30,022
R^2	0.233	0.051	0.380	0.468
FE: Individual			X	X
FE: Year	X	X	X	X

* p<0.1, ** p<0.05, *** p<0.01. Col. 1: LPM with year FE; dependent variable = 1 if contractual hours < 35. Col. 2–3: LPM for FT-to-PT transitions; sample restricted to FT workers in $t-1$. Col. 4: WLI triple interaction estimated on individuals observed in both FT and PT states (contract-type switchers). SE clustered by individual.

present it as a selection check rather than a mechanism test. Fourth, among full-time workers who will transition to part-time next year, pre-transition work-life interference is not differentially elevated for German speakers (interaction coefficient $= -0.058$, $p = 0.61$), providing no evidence of anticipatory selection.

5.5.1 Calibration

The contract-type estimates are consistent with a calibration of the model parameters from Section 3. The model predicts $|\hat{\beta}_3^{\text{PT}}| = (\lambda_G - \lambda_F) \cdot \sigma_{\text{PT}}$ and $|\hat{\beta}_3^{\text{FT}}| = (\lambda_G - \lambda_F) \cdot \sigma_{\text{FT}}$. Using the split-sample work-life interference coefficients ($|\hat{\beta}_3| = 0.038$ for part-time, 0.010 for full-time), the implied salience ratio is $\sigma_{\text{FT}}/\sigma_{\text{PT}} \approx 0.27$: full-time contracts have roughly one-quarter the psychological salience of part-time contracts as reference points. We caution, however, that this calibration relies on the split-sample estimates; the formal pooled triple interaction (Hours Gap × French × Part-Time; Table A.4) does not reach statistical significance ($p = 0.202$), so this ratio should be treated as suggestive rather than precisely estimated. The cultural welfare wedge (equation (6)) for the average German-speaking part-time overworker is $|\hat{\beta}_3^{\text{PT}}| \times \bar{\Delta} \approx 0.038 \times 5.4 \approx 0.21$ points on the 0–10 work-life interference scale, about 0.12 within-person standard deviations.

Figure A.5 displays the interaction coefficient estimated separately for four sub-periods. The work-life interference interaction is consistently negative and significant in 1999–2005 ($\hat{\beta}_3 =$

-0.043, $p = 0.018$), 2006–2012 ($\hat{\beta}_3 = -0.034$, $p < 0.001$), and 2013–2019 ($\hat{\beta}_3 = -0.035$, $p < 0.001$), but smaller in magnitude in 2020–2023. The within-period sub-sample estimate is $\hat{\beta}_3 = -0.019$ ($p = 0.081$); the full-sample estimate interacted with a post-2020 dummy yields a post-2020 coefficient of $\hat{\beta}_3 = -0.030$ (95% CI $[-0.050, -0.010]$, $p = 0.003$), which conditions on the 2020 level shift and uses the full sample's precision.[16] A formal test for whether the cultural interaction changed after 2020 yields $\Delta\hat{\beta}_3 = -0.009$ ($p = 0.406$): the visual attenuation is not formally distinguishable from sampling variation. One substantive explanation is that remote work blurred the boundary between contracted hours and actual work for both language groups during the pandemic, reducing the salience of formal contractual hours as a reference point across the linguistic divide equally rather than selectively for French speakers. The temporal stability prior to COVID is reassuring for identification.

Table A.8 tests for differential panel attrition. The interaction Hours Gap $\times$ French does not significantly predict panel exit ($p = 0.12$), alleviating concerns about selective attrition.[17]

5.6 Heterogeneity

Figure A.3 displays the interaction coefficient $\hat{\beta}_3$ (Hours Gap $\times$ French) for three outcomes: exhaustion, work-life interference, and ability to disconnect. We estimate effects separately for subsamples defined by gender, parental status, managerial role, and border region residence. Recall that a more negative $\hat{\beta}_3$ indicates a larger German–French gap (German workers bear more cost).

We interpret these heterogeneity results with caution. While point estimates exhibit some variation across groups, the differences between subgroups are not statistically significant, as indicated by the overlapping confidence intervals. The most informative heterogeneity result comes from the contract-type split in Section 5.5.

5.7 Behavioral Response to Overwork

The preceding results establish that cultural norms shape the psychological cost of overwork. A natural follow-up question is whether cultural norms also shape the behavioral response, that is, whether the higher psychological costs experienced by German-speaking workers translate into differential labor supply adjustments. We investigate three margins: (i) adherence to contractual hours, (ii) dynamic hours adjustment, and (iii) job mobility. The results reveal an absence of detectable differential behavioral response at the annual frequency.

One measurement caveat applies to the dynamic adjustment and job-mobility analyses: the SHP is an annual survey, so behavioral adjustment that occurs within a calendar year (e.g., a February hours renegotiation affecting the subsequent wave) would not appear as an adjustment in our data. If anything, this measurement frequency biases toward behavioral nulls: workers who adjust quickly would already appear adjusted by the time of the next interview. This limitation means we cannot distinguish between no adjustment and rapid adjustment; we can only establish that there is no differential *net* adjustment observable at the annual horizon.

[16] The two post-2020 estimates differ because the within-period estimate uses only 2020–2023 observations, while the interaction-with-post-2020 estimate uses the full sample and identifies the post-2020 change conditional on the pre-2020 trend.

[17] We also identify 142 individuals who switch interview language during the panel (49 with observations in both languages). Repeating the baseline regressions on this subsample yields interaction coefficients that are statistically insignificant for all outcomes ($p > 0.26$), consistent with the small sample size ($N \approx 700$). These within-switcher estimates are too imprecise to be informative but are consistent with the cultural effect reflecting regional norms rather than individual language identity. Full results are reported in Appendix Table A.6.

Table 7: Dynamic Hours Adjustment by Language Region

	Hours Gap$_t$	Δ Hours Gap$_t$	Wants Less$_t$
Hours Gap$_{t-1}$	0.034***	−0.966***	
	(0.009)	(0.009)	
French Region	−0.340	−0.340	0.000
	(0.477)	(0.477)	(0.044)
Hours Gap$_{t-1}$ × French	−0.012	−0.012	
	(0.014)	(0.014)	
Overwork$_{t-1}$			0.006
			(0.009)
Overwork$_{t-1}$ × French			0.016
			(0.016)
Observations	54,072	54,072	54,072
R^2	0.554	0.697	0.492
FE: Individual	X	X	X
FE: Year	X	X	X

* p<0.1, ** p<0.05, *** p<0.01. Individual and year fixed effects. Standard errors clustered at the individual level. Controls: age, age^2, female, has children. Sample restricted to |hours gap| $\leq$ 20. Column 1: level persistence of hours gap. Column 2: change in hours gap. Column 3: binary indicator for wanting fewer hours on lagged binary overwork indicator.

5.7.1 Hours Distribution

Figure A.6 displays mean outcomes by hours-gap bin and language region. Both language groups show rising WLI and exhaustion as overwork hours increase; the German-speaking group shows a steeper slope above the contract, consistent with the regression estimates. The parallel trends below zero (underwork) support the placebo result.

Table A.9 formalizes this visual impression. Using canton and year fixed effects, we estimate the probability that a worker's hours gap falls within $k = 1, 2, 3$ hours of zero. The coefficient on French Region is 0.020 ($p < 0.05$) for the $\pm$1-hour band and 0.013 ($p > 0.1$) for $\pm$2 hours, indicating that French-speaking workers are slightly more likely to cluster near their contractual hours. This means German-speaking workers show, if anything, less bunching at the contractual boundary — not more — contrary to what the cultural-salience story might predict on the extensive margin. The behavioral null is thus stronger than the simple null: German speakers bear higher psychological costs *and* are not more likely to enforce contractual adherence.

5.7.2 Dynamic Hours Adjustment

Table 7 tests whether French-speaking workers who overwork in period $t-1$ adjust their hours more aggressively in period t. Column 1 reports the level-form persistence regression: a negative interaction on Hours Gap$_{t-1}\times$ French would indicate faster mean reversion for French speakers. Column 2 presents the change-form specification. Column 3 uses binary indicators for overwork and the preference for fewer hours.

All specifications include individual and year fixed effects. The interaction coefficient in column 1 is negative and small (-0.012, SE $= 0.014$, $p = 0.394$): there is no statistically significant

difference in hours-gap persistence between French and German speakers. The binary specification (column 3) shows Overwork$_{t-1}\times$ French $= 0.016$ ($p = 0.316$), also null. Despite experiencing greater psychological costs from overwork, German-speaking workers do not appear to correct overwork episodes more rapidly than their French-speaking counterparts at the annual frequency. To formally bound these null results, we apply two-sided equivalence tests (TOST) using the absolute WLI cultural penalty ($\varepsilon = 0.038$) as the equivalence margin. The 90% confidence interval for hours-gap persistence ($[-0.035, +0.011]$) lies within the equivalence region ($[-0.038, +0.038]$), establishing that the cultural differential in dynamic adjustment is smaller than the psychological effect. The job-mobility continuous interaction (90% CI: $[-0.002, +0.002]$) establishes equivalence with ease. The remaining two interactions (binary dynamic adjustment, binary job change) produce wider 90% confidence intervals that do not formally establish equivalence at this margin, reflecting lower power in binary outcome specifications. We note that the original TOST margin ($\varepsilon = 0.038$) compared coefficients with fundamentally different units (WLI points per hour versus hours per hour-lag). To address this, we re-express the psychological and behavioral interactions in comparable terms. Standardizing the WLI outcome to within-person standard-deviation units, the cultural psychological interaction is -0.019 WLI-SD per hour of overwork (90% CI: $[-0.024, -0.014]$). For behavioral adjustment, the 90% CI for the persistence interaction ($[-0.092, -0.016]$ hours per hour-lag) lies entirely within the economically meaningful threshold of 0.185 hours per hour-lag (equivalent to 1 hour of faster annual mean reversion at mean overwork). The behavioral differential is thus bounded below an economically interpretable threshold, establishing that any cultural difference in hours adjustment is small in absolute terms.

5.7.3 Job Mobility

Table A.10 tests whether overwork differentially triggers job changes across language regions. Job change is defined broadly as any professional change (new employer, new position, or change in responsibilities) occurring in the 12 months following the survey (approximately 10% annual rate). Column 2 reveals that the relationship between overwork in $t-1$ and job change is essentially zero (-0.0001, $p = 0.828$). The interaction with French region is small and insignificant across all specifications.

5.7.4 Robustness and Heterogeneity

Table A.11 reports robustness checks. Column 1 presents Anderson-Hsiao instrumental variables estimates (Anderson and Hsiao (1981); Anderson and Hsiao (1982)) using the second lag as instrument; the estimates are imprecise, as expected given the weak first stage. Columns 2–3 restrict to border cantons. Columns 4–5 split by gender, revealing that hours persistence is driven primarily by men ($\hat{\beta}_1 = 0.037$, $p < 0.01$) with no significant cultural interaction in either subsample.

Figure A.4 displays the heterogeneity in the dynamic adjustment interaction across subgroups. All confidence intervals span zero, confirming the absence of significant cultural differences in behavioral adjustment across demographic and occupational lines.

5.7.5 Discussion

The behavioral results present an informative contrast with the psychological cost findings. German-speaking workers experience significantly greater work-life interference per hour of overwork, yet this heightened cost does not translate into detectable differential adjustment along any of the three margins we examine at the annual frequency. This absence of detectable differential adjustment is consistent with two interpretations. The first is the "constrained

preferences" view: the cultural parameter $\lambda_G > \lambda_F$ generates a welfare loss observable only through subjective well-being, but cultural or institutional constraints prevent behavioral adjustment. The second is rational inaction: the cultural welfare cost amounts to 0.21 points on a 0–10 scale at mean overwork, approximately 0.12 within-person standard deviations. If the fixed costs of job search, contract renegotiation, or hours adjustment exceed this welfare gain (Chetty (2012) documents that optimal-adjustment frictions as small as 1% of earnings rationalize large behavioral inertia), workers rationally do not adjust regardless of preferences. The data cannot distinguish between these interpretations. We return to this in Section 6.

5.8 Extensions

5.8.1 Placebo Test: Underwork

If our results capture a genuine cultural moderation of overwork, the interaction should be absent when the hours gap is negative, that is, when workers work fewer hours than their contract specifies. We split the sample into overwork episodes (hours gap > 0; $N \approx 31{,}230$) and underwork episodes (hours gap < 0; $N \approx 851$) and re-estimate the baseline specification separately for each subsample. Table 8 reports the results. We note that the underwork sample is substantially smaller ($N \approx 851$ versus $N \approx 31{,}230$ overwork episodes), resulting in standard errors approximately 5 times larger for the underwork interaction; a null result for underwork should therefore be interpreted with this power asymmetry in mind rather than as precise evidence of a zero effect. With 851 underwork observations, the minimum detectable effect (80% power, $\alpha = 0.05$) for the cultural interaction is approximately 0.11, roughly three times the observed overwork interaction magnitude; the placebo null is therefore consistent with both a true zero and an effect of similar magnitude to the overwork result.

Table 8: Placebo Test: Overwork vs. Underwork Samples

	Exhaust. (Over)	Exhaust. (Under)	WLI (Over)	WLI (Under)	Disc. (Over)	Disc. (Under)
Hours Gap	0.078***	0.064	0.127***	0.008	0.075***	−0.060
	(0.006)	(0.041)	(0.008)	(0.045)	(0.006)	(0.036)
French Region	0.442*		0.734		−0.252	
	(0.254)		(0.481)		(0.435)	
Hours Gap × French	−0.033**	−0.029	−0.035***	−0.031	−0.009	0.116*
	(0.016)	(0.058)	(0.012)	(0.076)	(0.012)	(0.064)
Observations	31,235	852	31,226	850	31,240	850
R^2	0.616	0.757	0.581	0.709	0.669	0.735
FE: Individual	X	X	X	X	X	X
FE: Year	X	X	X	X	X	X

* p<0.1, ** p<0.05, *** p<0.01. Individual and year fixed effects. Standard errors clustered at the individual level. Controls: age, age^2, female, has children. Odd columns: hours gap > 0 (overwork). Even columns: hours gap < 0 (underwork). All outcomes scaled 0–10.

The cultural interaction is negative for overwork ($\hat{\beta}_3 = -0.033$, $p < 0.05$ for exhaustion; $\hat{\beta}_3 = -0.035$, $p < 0.05$ for work-life interference), confirming the German–French gap, but small and statistically insignificant for underwork ($\hat{\beta}_3 \approx -0.029$–$-0.031$, $p > 0.6$).[18] The key contrast is between the overwork and underwork columns: the overwork interaction is consistently negative and individually significant, while the underwork interaction is consistently near zero and insignificant across all three outcomes. This asymmetry is consistent with the cultural penalty being specific to hours above the contract, not to any hours mismatch, as predicted by Prediction 3 from Section 3: German-speaking contractual norms impose a psychological cost on boundary violations in the "overwork" direction, not on underwork.

[18]The overwork-subsample estimates in the placebo test are smaller in magnitude than the full-sample estimates from Table 2 because the placebo sample is restricted to hours-gap > 0 observations only, altering the composition and size of the estimation sample ($N = 31{,}226$–$31{,}240$ versus $N = 71{,}950$–$71{,}985$).

5.8.2 Income Returns to Overwork

One concern is that the higher psychological cost of overwork for German-speaking workers reflects lower material compensation for extra hours. If German-speaking workers receive smaller wage returns to overwork, the same objective hours might generate greater subjective cost simply because the pecuniary reward is inadequate. Table 9 tests this hypothesis by regressing log monthly income on the hours gap, the French indicator, and their interaction.

Table 9: Income Returns to Overwork by Language Region

	Log Income	Income (1000 CHF)	Log Income (Over)	Log Income (Canton)
Hours Gap	0.009***	0.032***	0.007***	0.029***
	(0.001)	(0.004)	(0.001)	(0.001)
French Region	−0.013	0.129	0.071	0.038
	(0.053)	(0.303)	(0.053)	(0.037)
Hours Gap × French	−0.002	0.004	−0.003	−0.002**
	(0.002)	(0.008)	(0.002)	(0.001)
Observations	63,399	63,399	27,966	66,612
R^2	0.826	0.887	0.857	0.392
FE: Individual	X	X	X	
FE: Year	X	X	X	X

* p<0.1, ** p<0.05, *** p<0.01. Columns 1–3: individual and year fixed effects, SE clustered by individual. Column 4: canton and year fixed effects, SE clustered by canton. Controls: age, age^2, female, has children (all); education (column 4). Column 3 restricted to hours gap > 0.

Each additional hour of overwork raises log income by 0.9% ($p < 0.001$, column 1). The interaction term is negative in all log-income specifications ($\hat{\beta}_3 < 0$ in columns 1, 3, 4) and a negligible $+0.004$ ($p = 0.61$) in the income-levels specification (column 2), though not statistically significant under individual fixed effects (column 1: $\hat{\beta}_3 = -0.002$, $p \approx 0.34$). Under canton fixed effects (column 4), the interaction is -0.002 ($p < 0.05$): if anything, French-speaking workers earn *less* per hour of overwork than German-speaking workers. This rules out differential compensation as an explanation for the German–French WLI gap: German speakers are not under-compensated for their overwork. The cultural penalty for German speakers operates through boundary norms rather than through inferior wage returns to extra hours.

5.8.3 Contractual vs. Habitual Reference Points

A direct test of the contractual-salience mechanism is whether the cultural interaction depends on the type of reference point. The SHP records both contractual hours (the formal benchmark) and "reference" hours ("How many hours per week do you normally work?"), a measure of habitual practice. We construct an alternative hours gap as actual minus reference hours and re-estimate the baseline specification. The two gaps are only modestly correlated ($r = 0.354$), reflecting the distinction between formal and informal benchmarks.

The cultural interaction is significant only for the contractual-hours gap ($\hat{\beta}_3 = -0.038$, $p < 0.001$ for work-life interference) and is essentially zero for the reference-hours gap ($\hat{\beta}_3 = -0.002$, $p = 0.548$). The same pattern holds for exhaustion: the interaction is -0.021 ($p < 0.001$) with contractual hours but -0.003 ($p = 0.264$) with reference hours (Appendix Table C.4). This divergence is the sharpest evidence for contractual salience: the cultural penalty is triggered by violations of the formal contractual boundary (the explicit, legally binding commitment), not by deviations from what the worker "normally" does. Consistent with this interpretation, desired hours show no differential cultural response to overwork ($p = 0.15$),

Table 10: Regression Kink at the Contractual Boundary

	Exhaustion	Work-Life Int.	Disconnect	Free Time Sat.
Hours Gap (Below Contract)	−0.013	−0.004	−0.014	0.004
	(0.013)	(0.015)	(0.013)	(0.012)
Hours Gap (Above Contract)	0.096***	0.140***	0.101***	−0.084***
	(0.005)	(0.006)	(0.005)	(0.006)
Below Contract × French	−0.011	−0.035	−0.010	0.051*
	(0.025)	(0.029)	(0.023)	(0.027)
Above Contract × French	−0.003	−0.031***	0.009	0.032***
	(0.010)	(0.011)	(0.011)	(0.010)
Observations	69,529	69,491	69,559	72,507
R^2	0.581	0.535	0.628	0.570
FE: Individual	X	X	X	X
FE: Year	X	X	X	X

* p<0.1, ** p<0.05, *** p<0.01. Individual and year FE. SE clustered by individual. Hours Gap split into Below Contract ($\min(\Delta, 0)$) and Above Contract ($\max(\Delta, 0)$). A significant "Above × French" with null "Below × French" indicates a cultural kink at the contractual boundary. Sample: $|\text{hours gap}| \leq 15$. Controls: age, age^2, female, has children.

confirming that the cultural parameter $\lambda_G > \lambda_F$ amplifies the disutility of boundary violations without changing desired hours.

5.8.4 Regression Kink at the Contractual Boundary

The contractual salience framework predicts a kink in the relationship between the hours gap and burnout at $\Delta = 0$: the cultural interaction should "switch on" when hours exceed the contract and be absent below. We test this directly by decomposing the hours gap into its positive part ($\Delta^+ = \max(\Delta, 0)$) and negative part ($\Delta^- = \min(\Delta, 0)$) and estimating separate slopes and cultural interactions for each. Table 10 reports the results.

For work-life interference, the cultural interaction above the contract is -0.031 ($p < 0.001$) while the interaction below the contract is -0.035 ($p = 0.228$), We note, however, that the point estimates are nearly identical in magnitude (-0.031 above, -0.035 below). A formal Wald test of the equality of the two cultural interactions yields $p = 0.91$ (Table D.5), confirming that the above- and below-contract slopes are statistically indistinguishable. The asymmetric significance is driven entirely by differential sample size: the below-contract sample is much smaller, generating standard errors roughly three times larger (SE $= 0.029$ vs. 0.011). We therefore revise our interpretation: the kink evidence does not support an asymmetric cultural penalty that switches on at the contractual boundary. Rather, the regression-kink test is uninformative about asymmetry given the extreme power imbalance across the two segments. Figure A.8 visualizes the relationship: both language groups show steeper slopes above the contract than below, consistent with the overwork penalty. The visual slope difference between language groups appears similar on both sides of the contractual boundary, consistent with the formal equality test ($p = 0.91$). The original interpretation that the contractual boundary serves as a threshold that triggers the cultural penalty is not supported by the formal kink equality test. The cultural interaction appears to be of similar magnitude regardless of whether hours are above or below contract; the asymmetric significance reflects differential power, not a genuine kink.

Additional robustness checks (dose-response nonlinearity, COVID-era stability, and gender attitudes as a potential mediator) are reported in Appendix C. In brief, the squared interaction (Hours Gap$^2\times$ French) is insignificant across all outcomes, confirming that the linear specification is adequate; the cultural moderation is broadly stable before and after 2020 (though attenuated in the post-pandemic period, see Figure A.5); and controlling for gender-role attitudes produces zero attenuation of the interaction coefficient, suggesting that the cultural channel operates through some other dimension of the work-leisure schema beyond standard gender norms.

6 Conclusion

This paper has used Switzerland's linguistic border (the Röstigraben) to study the role of cultural norms in shaping both the psychological cost and the behavioral response to overwork. Exploiting 25 waves of the Swiss Household Panel and the fact that French- and German-speaking workers share identical institutional constraints, we document two complementary findings. First, overwork generates significantly more work-life interference for German-speaking workers, an effect that is robust to occupation fixed effects, additional time-varying controls, and alternative clustering strategies. This effect is concentrated among part-time workers; the cultural moderation is absent among full-time workers. Post-work exhaustion and difficulty disconnecting also show cultural moderation, though the exhaustion result is sensitive to outlier trimming. Second, despite these higher psychological costs, German-speaking workers do not detectably adjust their labor supply differently at the annual horizon of our survey: they do not adhere more tightly to contractual hours, do not correct overwork episodes faster, and do not exit overwork situations through job changes at higher rates. We cannot rule out within-year behavioral adjustments that are resolved before the next survey wave.

The absence of detectable differential behavioral adjustment at the annual frequency is an informative null, consistent with either constrained preferences or rational non-response to welfare costs that are modest in absolute terms (0.21 WLI points, 0.12 within-person standard deviations). A placebo test is consistent with the cultural penalty being specific to overwork and absent for underwork, while an income analysis finds no evidence that German workers are under-compensated for their overwork relative to French workers (if anything, the income interaction is negative, indicating German speakers receive at least as much income per extra hour), ruling out differential compensation as an explanation for German workers' higher costs. The finding is consistent with two interpretations. The "constrained preferences" view holds that cultural norms shape how overwork feels but not what workers do. The rational-inaction view holds that the welfare gain from adjustment (0.21 points on a 0–10 scale, or 0.12 within-person standard deviations) is smaller than the fixed costs of behavioral adjustment such as job search or renegotiation.[19] Chetty (2012) shows that optimization frictions as small as 1% of earnings can rationalize large behavioral inertia, making rational non-adjustment plausible at the magnitudes we document. The data cannot distinguish between these interpretations.

The concentration of the effect among part-time workers is the paper's most informative heterogeneity result. The gender split within contract types is informative: German-speaking part-time men show a larger cost differential relative to French speakers ($|\hat{\beta}_3| = 0.054$) than German-speaking part-time women ($|\hat{\beta}_3| = 0.031$), and full-time men show zero effect ($|\hat{\beta}_3| =$

[19] Direct CHF monetization of this welfare gap is infeasible with the available data: the within-person income–WLI relationship conflates overwork with the income returns to overwork, yielding implausibly large compensating variation estimates. We therefore quantify welfare in well-being scale units throughout.

0.002). However, full-time women show a marginally significant interaction ($|\hat{\beta}_3| = 0.032$, $p < 0.1$) similar to part-time women, suggesting that the contractual-salience mechanism operates primarily for men; for women, the cultural moderation does not depend on contract type. Four selection tests (detailed in Section 5) collectively provide no support for differential sorting into part-time work. The pattern is consistent with the cultural moderation operating through the salience of the contractual reference point: part-time contracts more clearly demarcate the work-leisure boundary, and violations of that boundary are psychologically costlier for German-speaking workers whose work culture prizes strict contractual compliance and the protection of leisure. Consistent with this interpretation, an alternative hours-gap measure based on self-reported "normal" hours rather than contractual hours produces a zero cultural interaction ($p = 0.548$ for work-life interference).

An additional finding narrows the interpretation. Work stress, a measure of within-job cognitive demands, shows a cultural interaction *opposite* in sign to the main effects ($\hat{\beta}_3 = +0.003$, $p < 0.001$), consistent with the cultural penalty operating specifically through work-leisure boundary norms rather than through general within-job cognitive demands. Consistent with this interpretation, none of eight semi-objective health measures (including days affected by health problems, $p = 0.70$; doctor consultations, $p = 0.85$; and health impediment, $p = 0.83$) show a cultural interaction that survives Benjamini-Hochberg correction for multiple testing (the only nominally significant result, weakness/weariness at $p = 0.046$, adjusts to $p_{\text{BH}} = 0.36$), confirming that the cultural cost is genuinely psychological rather than a proxy for differential physical health consequences.

Our study faces several important limitations. First, the identification design does not exploit spatial proximity to the Röstigraben. While the prior literature on Swiss cultural differences typically uses spatial regression-discontinuity designs (Eugster et al., 2017, 2011; Brügger et al., 2009), our estimator compares all French- and German-speaking workers rather than restricting to those near the border. Our within-person fixed effects eliminate all time-invariant confounders, and the addition of canton-by-year fixed effects (Appendix Table B.4) leaves the interaction coefficient unchanged, ruling out region-specific economic shocks. In unilingual cantons, language is perfectly confounded with canton identity; the cultural interaction in these cantons could partly reflect canton-specific economic conditions. Restricting to the bilingual cantons (Bern, Fribourg, Valais), where within-canton language variation exists, yields a consistent but attenuated and statistically insignificant interaction ($\hat{\beta}_3 = -0.028$, $p = 0.21$; Appendix Table D.4), reflecting the smaller sample size. Our results should therefore be interpreted as documenting a cross-language-region differential under strong controls rather than a causal estimate identified at the geographic discontinuity.

Second, the hours-gap measure is self-reported and endogenous: within-person variation is driven by job changes and employer decisions that may independently affect burnout. While the coefficient is stable across specifications with progressively richer controls (an Oster (2019) analysis confirms δ^* well above the threshold of 1 for all outcomes), we lack an instrument for hours-gap variation and cannot make strong causal claims. Third, cultural differences in the sensitivity of burnout reporting to overwork cannot be ruled out by individual fixed effects. The domain-specificity of our results (the cultural penalty appearing for work-life interference but not for health satisfaction, income satisfaction, or general work-related outcomes) argues against a general reporting-bias explanation. However, we cannot exclude domain-specific interaction-level bias: if German-speaking workers report work-leisure-boundary outcomes with systematically higher sensitivity to overwork than French-speaking workers, $|\hat{\beta}_3|$ overstates the true welfare effect. Fourth, our sample is approximately 71% German-speaking; while this does not affect point estimates under individual fixed effects, it warrants care in interpreting descriptive statistics. Fifth, we exclude Italian-speaking respondents due to their

small sample size. Sixth, the exhaustion result is fragile: it attenuates to insignificance when extreme hours-gap values are trimmed or when observations with very low contractual hours (below 10 hours/week) are excluded. Only the work-life interference result survives all robustness specifications.

Despite these caveats, the paper establishes two complementary findings with reasonable confidence. The psychological cost of overwork is culturally modulated, being statistically significantly higher for German-speaking workers in part-time contracts and for women regardless of contract type. Critically, full-time men—the largest demographic subgroup—show zero cultural moderation, so the finding should not be characterized as a universal feature of the German-speaking labor market. At the same time, this subjective cost does not translate into detectable differential behavioral adjustment at the annual frequency of our data.

The contractual-salience mechanism has practical implications. Reck and Seibold (2023) show that the welfare effects of reference-point shifts are robustly signed whether reference dependence is normative or a bias; our findings suggest that the welfare cost of overwork is culturally heterogeneous along the same dimension. If the cultural cost of overwork is amplified when part-time contracts make the work-leisure boundary explicit, then firms designing flexible or part-time arrangements should recognize that contractual hours serve not merely as administrative benchmarks but as psychologically salient reference points.

The calibration in Section 5.5 estimates a cultural welfare wedge of 0.21 WLI points (0.12 within-person standard deviations) for the average German-speaking part-time overworker. BFS data place approximately 570,000 part-time workers in the German-speaking Grossregionen, of whom roughly 47% are overworking in our sample, so the wedge applies to approximately 268,000 workers. This calibration should be interpreted as an upper bound: if differential reporting sensitivity accounts for part of the estimated interaction (Section 4), the true population-level welfare wedge is correspondingly smaller.

In culturally heterogeneous labor markets, this wedge may vary across workforce segments in ways standard productivity metrics do not capture, though the magnitudes documented here are modest in absolute terms and subject to the reporting-sensitivity caveats discussed in Section 4. Future work should investigate whether similar contractual-salience effects operate in other institutional settings, particularly in the growing gig and platform economy, where the boundary between contracted and actual work is increasingly fluid.

Data and Code Availability

The Swiss Household Panel (SHP) data are available from FORS upon application (`https://forscenter.ch/projects/swiss-household-panel/`). Analysis scripts reproducing all tables and figures are available from the authors upon request. The analysis was conducted in R 4.5.3 using the `fixest`, `modelsummary`, `tidyverse`, and `kableExtra` packages.

References

Alesina, Alberto and Paola Giuliano, "Culture and Institutions," *Journal of Economic Literature*, 2015, *53* (4), 898–944.

_ , Edward Glaeser, and Bruce Sacerdote, "Work and Leisure in the US and Europe: Why So Different?," *NBER Macroeconomics Annual*, 2005, *20*, 1–64.

Allen, Tammy D., Kimberly A. French, Soner Dumani, and Kristen M. Shockley, "A Cross-National Meta-Analytic Examination of Predictors and Outcomes Associated with Work-Family Conflict," *Journal of Applied Psychology*, 2020, *105* (6), 539–576.

Anderson, Theodore W. and Cheng Hsiao, "Estimation of Dynamic Models with Error Components," *Journal of the American Statistical Association*, 1981, *76* (375), 598–606.

_ and _ , "Formulation and Estimation of Dynamic Models Using Panel Data," *Journal of Econometrics*, 1982, *18* (1), 47–82.

Année Politique Suisse, "Swissvotes – The Database on Swiss Popular Votes," 2024. Accessed: 2024.

Ashforth, Blake E., Glen E. Kreiner, and Mel Fugate, "All in a Day's Work: Boundaries and Micro Role Transitions," *Academy of Management Review*, 2000, *25* (3), 472–491.

Avgoustaki, Argyro and Almudena Cañibano, "Motivational Drivers of Extensive Work Effort: Are Long Hours Always Detrimental to Well-being?," *Industrial Relations*, 2020, *59* (3), 355–398.

Bell, David N.F. and David G. Blanchflower, "Youth Underemployment in the UK in the Great Recession," *National Institute Economic Review*, 2011, *215* (1), R1–R11.

Bell, Linda A. and Richard B. Freeman, "The Incentive for Working Hard: Explaining Hours Worked Differences in the US and Germany," *Labour Economics*, 2001, *8* (2), 181–202.

Booth, Alison L. and Jan C. van Ours, "Part-Time Jobs: What Women Want?," *Journal of Population Economics*, 2013, *26*, 263–283.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, "Salience Theory of Choice Under Risk," *Quarterly Journal of Economics*, 2012, *127* (3), 1243–1285.

_ , _ , and _ , "Salience and Consumer Choice," *Journal of Political Economy*, 2013, *121* (5), 803–843.

Brügger, Beatrix, Rafael Lalive, and Josef Zweimüller, "Does Culture Affect Unemployment? Evidence from the Röstigraben," *IZA Discussion Paper No. 4283*, 2009.

Camerer, Colin, Linda Babcock, George Loewenstein, and Richard Thaler, "Labor Supply of New York City Cabdrivers: One Day at a Time," *Quarterly Journal of Economics*, 1997, *112* (2), 407–441.

Cattaneo, Alejandra and Rainer Winkelmann, "Earnings Differentials between German and French Speakers in Switzerland," *Swiss Journal of Economics and Statistics*, 2005, *141*, 191–212.

Chetty, Raj, "Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply," *Econometrica*, 2012, *80* (3), 969–1018.

Clark, Sue Campbell, "Work/Family Border Theory: A New Theory of Work/Family Balance," *Human Relations*, 2000, *53* (6), 747–770.

Crawford, Vincent P. and Juanjuan Meng, "New York City Cab Drivers' Labor Supply Revisited: Reference-Dependent Preferences with Rational-Expectations Targets for Hours and Income," *American Economic Review*, 2011, *101* (5), 1912–1932.

Deopa, Neha and Piergiuseppe Fortunato, "Coronagraben in Switzerland: Culture and Social Distancing in Times of COVID-19," *Journal of Population Economics*, 2021, *34*, 1355–1383.

Eugster, Beatrix and Raphaël Parchet, "Culture and Taxes," *Journal of Political Economy*, 2019, *127* (1), 296–337.

_ , Rafael Lalive, Andreas Steinhauer, and Josef Zweimüller, "The Demand for Social Insurance: Does Culture Matter?," *Economic Journal*, 2011, *121* (556), F413–F448.

_ , _ , _ , and _ , "Culture, Work Attitudes, and Job Search: Evidence from the Swiss Language Border," *Journal of the European Economic Association*, 2017, *15* (5), 1056–1100.

Faessler, Lisa, Rafael Lalive, and Charles Efferson, "How culture shapes choices related to fertility and mortality: Causal evidence at the Swiss language border," *Evolutionary Human Sciences*, 2024, *6*, e28.

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde, "Global Evidence on Economic Preferences," *Quarterly Journal of Economics*, 2018, *133* (4), 1645–1692.

Farber, Henry S., "Reference-Dependent Preferences and Labor Supply: The Case of New York City Taxi Drivers," *American Economic Review*, 2008, *98* (3), 1069–1082.

Federal Office of Topography (swisstopo), "swissBOUNDARIES3D and Amtliches Ortschaftenverzeichnis," 2024. Accessed: 2025.

Federal Statistical Office, "Durée normale de travail des salariés par canton (DNT)," Technical Report, FSO, Neuchâtel 2024.

Fehr, Ernst and Lorenz Goette, "Do Workers Work More if Wages Are High? Evidence from a Randomized Field Experiment," *American Economic Review*, 2007, *97* (1), 298–317.

Fernández, Raquel and Alessandra Fogli, "Culture: An Empirical Investigation of Beliefs, Work, and Fertility," *American Economic Journal: Macroeconomics*, 2009, *1* (1), 146–177.

Giuliano, Paola, "Living Arrangements in Western Europe: Does Cultural Origin Matter?," *Journal of the European Economic Association*, 2007, *5* (5), 927–952.

Hamermesh, Daniel S., "The Timing of Work over Time," *Economic Journal*, 1999, *109* (452), 37–66.

_ and Jungmin Lee, "Stressed Out on Four Continents: Time Crunch or Yuppie Kvetch?," *Review of Economics and Statistics*, 2007, *89* (2), 374–383.

Hart, Oliver and John Moore, "Contracts as Reference Points," *Quarterly Journal of Economics*, 2008, *123* (1), 1–48.

Heckman, James J., "What Has Been Learned About Labor Supply in the Past Twenty Years?," *The American Economic Review*, 1993, *83* (2), 116–121.

Ichino, Andrea and Giovanni Maggi, "Work Environment and Individual Background: Explaining Regional Shirking Differentials in a Large Italian Firm," *Quarterly Journal of Economics*, 2000, *115* (3), 1057–1090.

Kahneman, Daniel and Amos Tversky, "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, 1979, *47* (2), 263–291.

Kleven, Henrik, Camille Landais, and Jakob Egholt Søgaard, "Children and Gender Inequality: Evidence from Denmark," *American Economic Journal: Applied Economics*, 2019, *11* (4), 181–209.

Knaus, Michael C. and Steffen Otterbach, "Work Hour Mismatch and Job Mobility: Adjustment Channels and Resolution Rates," *Economic Inquiry*, 2019, *57* (1), 227–242.

Kőszegi, Botond and Matthew Rabin, "A Model of Reference-Dependent Preferences," *Quarterly Journal of Economics*, 2006, *121* (4), 1133–1165.

Lepinteur, Anthony, "The Shorter Workweek and Worker Wellbeing: Evidence from Portugal and France," *Labour Economics*, 2019, *58*, 204–220.

Lucas, Robert E. and Leonard A. Rapping, "Real Wages, Employment, and Inflation," *Journal of Political Economy*, 1969, *77* (5), 721–754.

Luttmer, Erzo F. P. and Monica Singhal, "Culture, Context, and the Taste for Redistribution," *American Economic Journal: Economic Policy*, 2011, *3* (1), 157–179.

Oster, Emily, "Unobservable Selection and Coefficient Stability: Theory and Evidence," *Journal of Business & Economic Statistics*, 2019, *37* (2), 187–204.

Otterbach, Steffen, "Mismatches Between Actual and Preferred Work Time: Empirical Evidence of Hours Constraints in 21 Countries," *Journal of Consumer Policy*, 2010, *33*, 143–161.

Pencavel, John, "Whose Preferences Are Revealed in Hours of Work?," *Economic Inquiry*, 2016, *54* (1), 9–24.

Prescott, Edward C., "Why Do Americans Work So Much More than Europeans?," *Federal Reserve Bank of Minneapolis Quarterly Review*, 2004, *28* (1), 2–13.

Reck, Daniel and Arthur Seibold, "Welfare Economics with Reference-Dependent Preferences," *NBER Working Paper No. 31381*, 2023.

Seibold, Arthur, "Reference Points for Retirement Behavior: Evidence from German Pension Discontinuities," *American Economic Review*, 2021, *111* (4), 1126–1165.

Steinhauer, Andreas, "Working Moms, Childlessness, and Female Identity," *SSRN Working Paper*, 2018.

Stutzer, Alois and Rafael Lalive, "The Role of Social Work Norms in Job Searching and Subjective Well-Being," *Journal of the European Economic Association*, 2004, *2* (4), 696–719.

Tillmann, Robin, Marieke Voorpostel, Erika Antal, Nora Dasoki, Hannah Klaas, Ursina Kuhn, Florence Lebert, Gian-Andrea Monsch, and Valérie-Anne Ryser, "The Swiss Household Panel (SHP)," *Jahrbücher für Nationalökonomie und Statistik*, 2022, *242* (3), 403–420.

Tversky, Amos and Daniel Kahneman, "Loss Aversion in Riskless Choice: A Reference-Dependent Model," *Quarterly Journal of Economics*, 1991, *106* (4), 1039–1061.

Wooden, Mark, Diana Warren, and Robert Drago, "Working Time Mismatch and Subjective Well-being," *British Journal of Industrial Relations*, 2009, *47* (1), 147–179.

Appendices

A Additional Figures and Tables

A.1 Figures

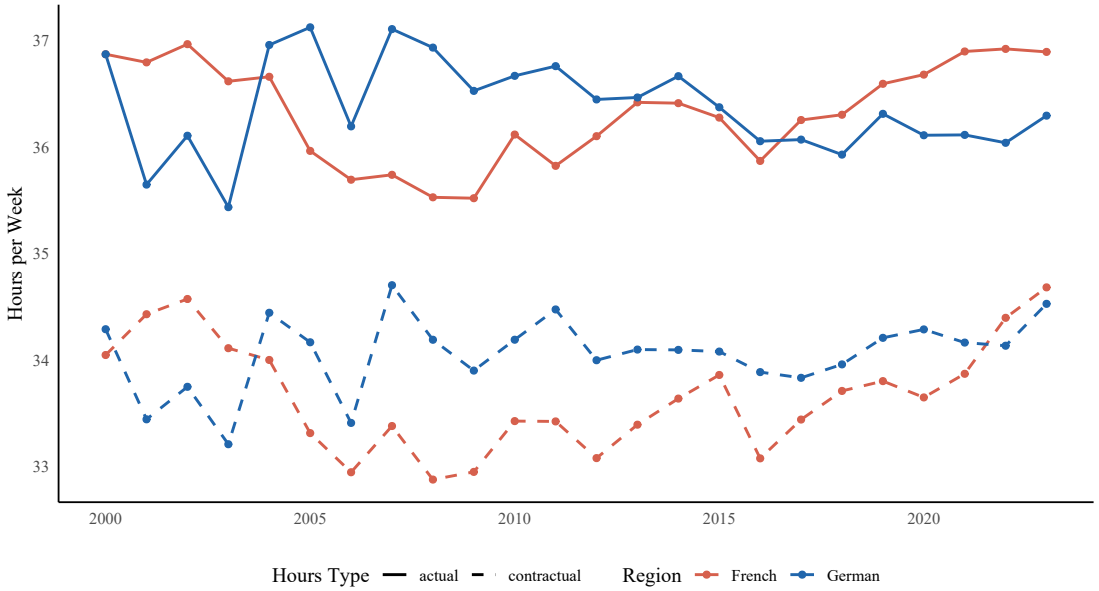


Figure A.1: Mean actual (solid) and contractual (dashed) weekly working hours by language region, 1999–2023.

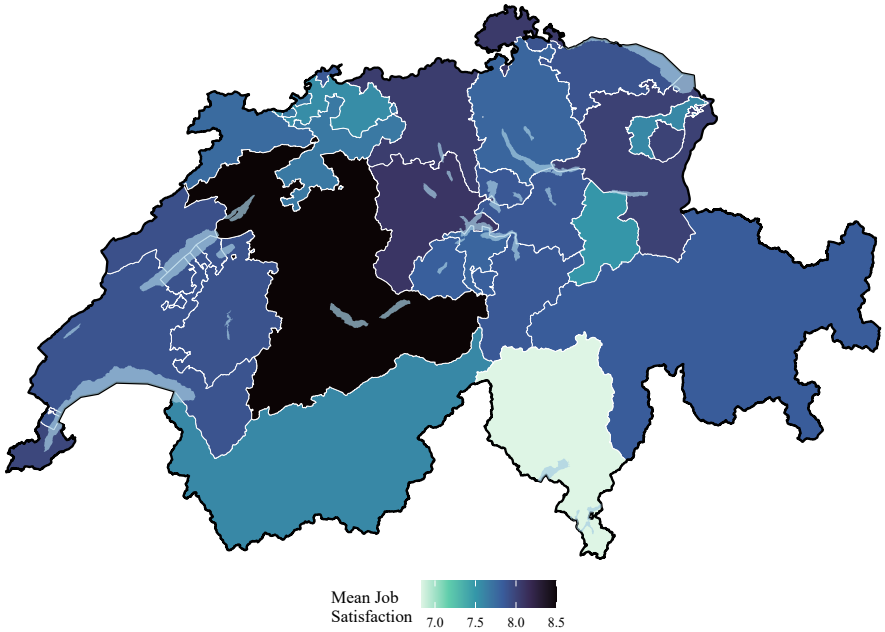
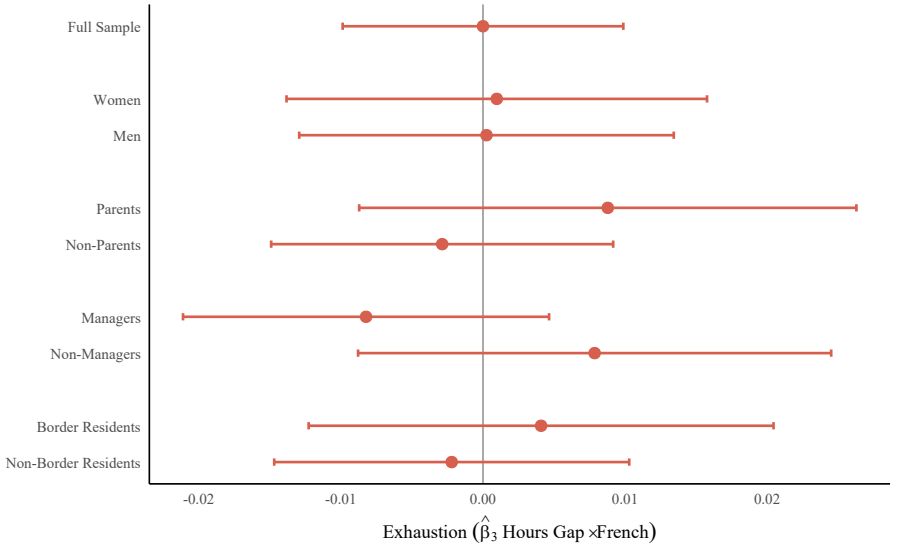
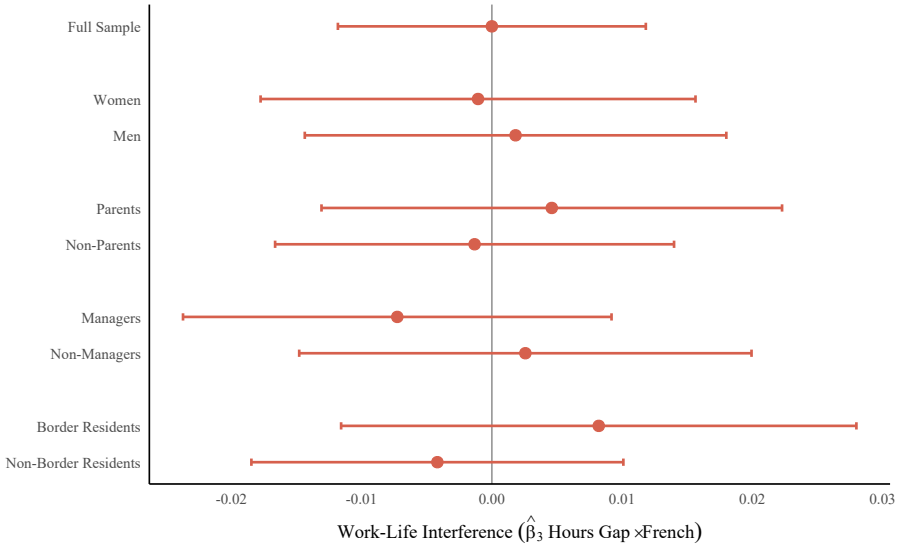


Figure A.2: Mean job satisfaction by canton (2021–2023).





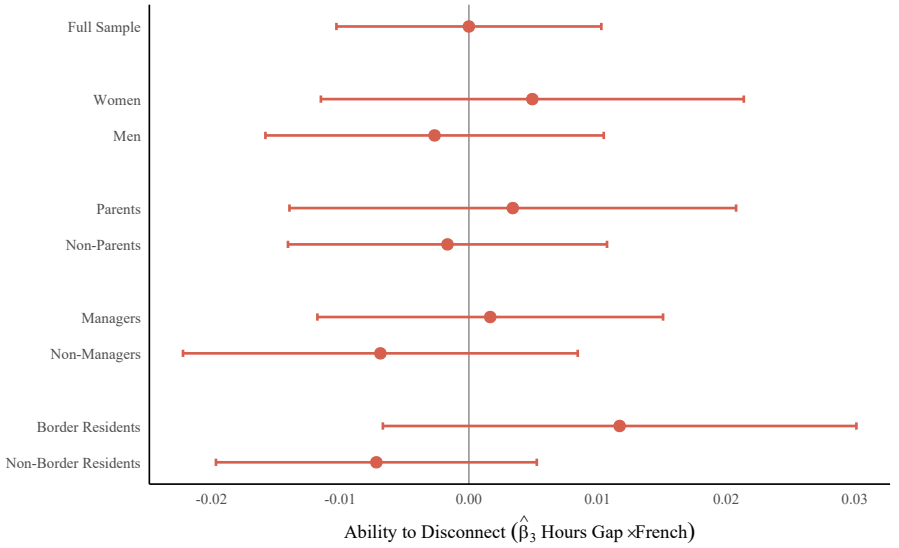


Figure A.3: Heterogeneity in the cultural effect of overwork on well-being outcomes. Coefficients on Hours Gap × French from separate regressions. 95% confidence intervals. Panels from top to bottom: Post-work exhaustion, Work-life interference, Ability to disconnect.

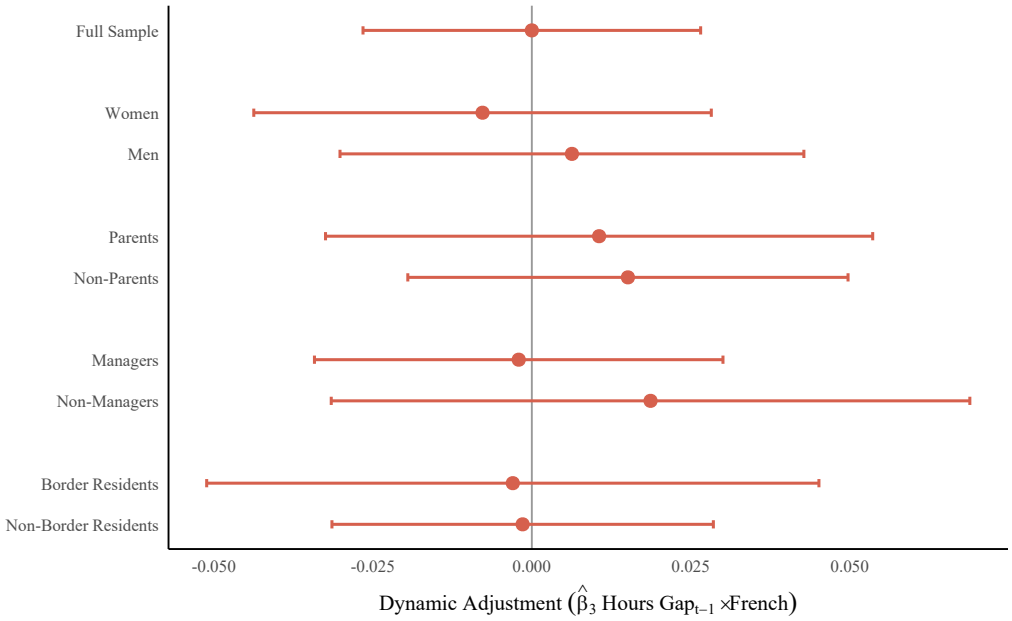


Figure A.4: Heterogeneity in the dynamic adjustment coefficient (Hours Gap$_{t-1}\times$ French) from the level-form persistence regression. 95% confidence intervals. Coefficients normalized relative to the full-sample estimate.

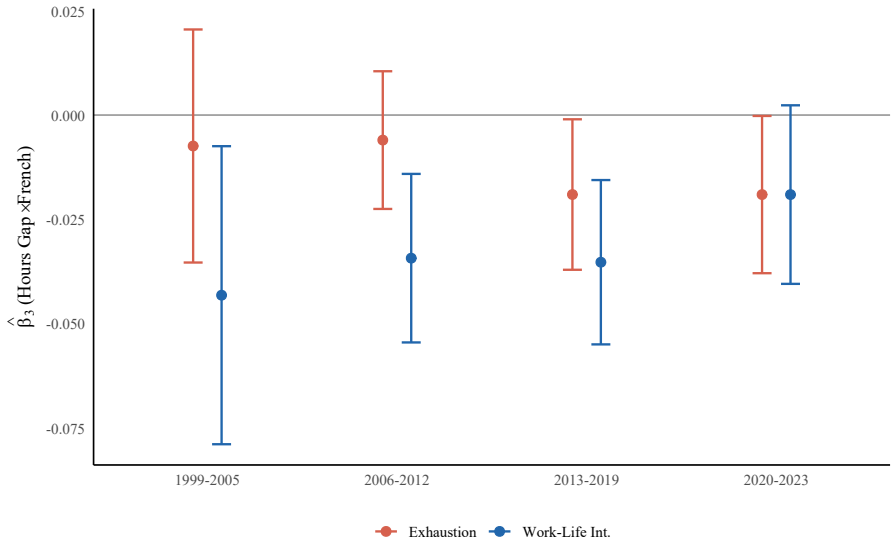


Figure A.5: Temporal stability of the cultural interaction coefficient ($\hat{\beta}_3$: Hours Gap $\times$ French) estimated separately for four sub-periods. 95% confidence intervals.

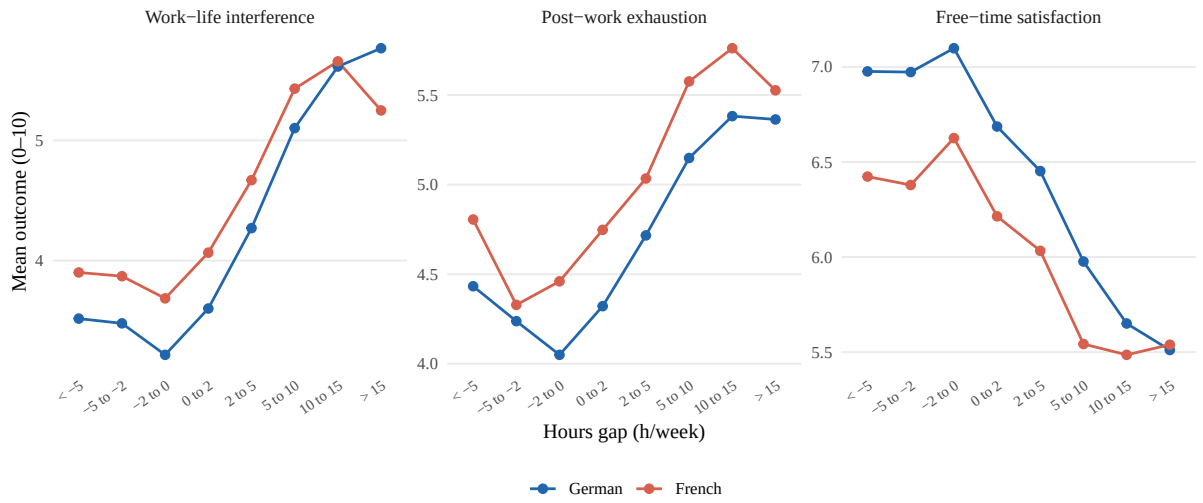


Figure A.6: Mean outcomes by hours gap bin and language region. Bins: < -5, $[-5, -2)$, $[-2, 0)$, 0, $(0, 2]$, $(2, 5]$, $(5, 10]$, > 10 hours. Cells with fewer than 20 observations are excluded. WLI = work-life interference; exhaustion and free-time satisfaction are on the same 0–10 scales. German-speaking workers (blue) show a steeper positive slope above the contract on WLI and exhaustion, consistent with the regression estimates. The parallel trends below zero are consistent with the placebo result.

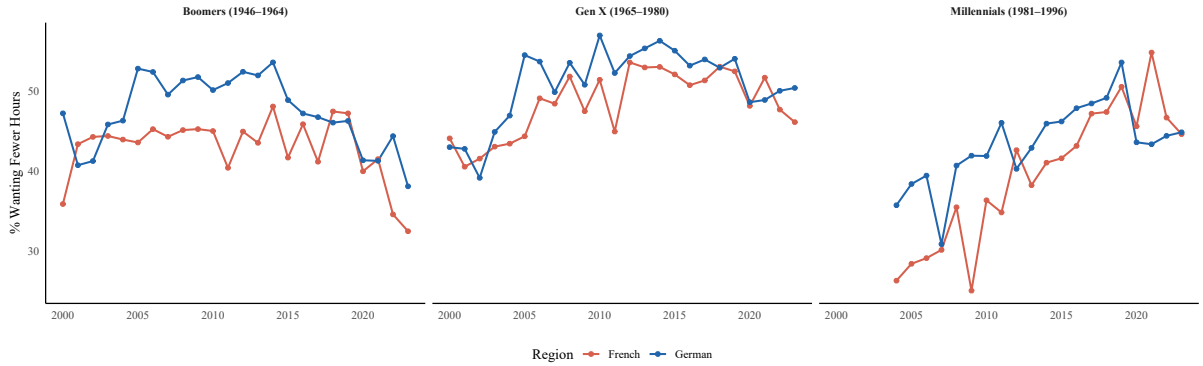


Figure A.7: Share wanting fewer hours by birth cohort and language region. The Millennial panel begins in 2004 (the cohort midpoint, 1988, plus 16, i.e. the approximate year when the median Millennial entered the labor force). Note that around 2020, the Boomer cohort midpoint (1955) plus 65 equals 2020, so many Boomers were at or near statutory retirement age during the post-pandemic period; the late-period Boomer estimates should be interpreted with this compositional shift in mind.

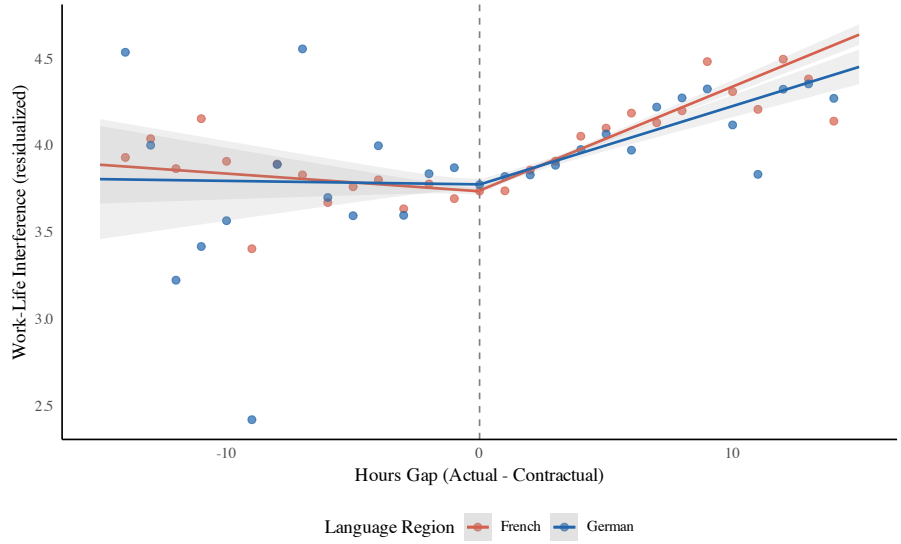


Figure A.8: Regression kink in work-life interference at the contractual boundary. Binned scatter of residualized work-life interference (individual and year FE removed) against the hours gap, with separate linear fits above and below zero. Both language groups show steeper slopes above the contract. The visual slope difference between groups appears similar above and below the contractual boundary, consistent with the formal equality test ($p = 0.91$; Table D.5).

A.2 Comprehensive Outcomes

Table A.1: Comprehensive Additional Outcomes

	Health Sat.	Financial Sat.	Income Sat.	Atmosphere	Work Stress
Hours Gap	0.004	0.005	−0.001	0.000	−0.004***
	(0.002)	(0.003)	(0.003)	(0.003)	(0.001)
Hours Gap × French	0.000	0.001	−0.003	−0.003	−0.003**
	(0.003)	(0.004)	(0.004)	(0.003)	(0.001)
Observations	77,018	76,993	76,978	76,219	67,290
R^2	0.579	0.609	0.566	0.505	0.537
FE: Individual	X	X	X	X	X
FE: Year	X	X	X	X	X

* p<0.1, ** p<0.05, *** p<0.01. Individual and year FE. SE clustered by individual. Controls: age, age^2, female, has children. All outcomes scaled 0–10.

A.3 Robustness Controls

Table A.2: Robustness: Additional Controls and Specifications (Exhaustion)

	Exhaust.	+Occ. FE	+Supv.	+All	Trimmed	2-Way Cl.
Hours Gap	0.042***	0.041***	0.041***	0.039***	0.067***	0.042***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)
Hours Gap × French	−0.021***	−0.020***	−0.021***	−0.019***	−0.018*	−0.021***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.009)	(0.004)
Supervisor			0.207***	0.192***		
			(0.030)	(0.029)		
Observations	71,985	71,953	71,869	70,822	70,797	71,985
R^2	0.576	0.579	0.577	0.582	0.579	0.576
FE: Individual	X	X	X	X	X	X
FE: Year	X	X	X	X	X	X

* p<0.1, ** p<0.05, *** p<0.01. Individual and year FE throughout. SE clustered by individual except col. 6 (two-way: individual + year). Col. 1: baseline. Col. 2: + ISCO FE. Col. 3: + supervisor. Col. 4: all controls. Col. 5: $|\text{hours gap}| \leq 20$. Col. 6: two-way clustering.

A.4 Institutional Controls

Table A.3: Institutional Controls: Perceived Unemployment Risk

	WLI (baseline)	WLI (+risk)	Unemp. Risk	WLI (+risk, FE)
Hours Gap	0.063***	0.064***	−0.002	0.064***
	(0.004)	(0.004)	(0.003)	(0.004)
Hours Gap × French	−0.038***	−0.038***	−0.002	−0.038***
	(0.006)	(0.006)	(0.005)	(0.006)
Unemp. Risk		0.088***		0.088***
		(0.007)		(0.007)
Num.Obs.	71,950	71,412	76,421	71,412
R2	0.533	0.537	0.470	0.537
FE: idpers	X	X	X	X
FE: year	X	X	X	X

* p<0.1, ** p<0.05, *** p<0.01. Individual and year FE. SE clustered by individual. Col. 1: baseline WLI. Col. 2: +unemployment risk control. Col. 3: DV = unemployment risk. Col. 4: same as 2.

A.5 Triple Interaction

Table A.4: Triple Interaction: Contract Type as Moderator

	Exhaustion	Work-Life Int.
Hours Gap	0.053***	0.071***
	(0.004)	(0.006)
Hours Gap × French	−0.013	−0.016*
	(0.008)	(0.009)
Hours Gap × Part-Time	−0.021***	−0.011
	(0.006)	(0.008)
Hours Gap × French × Part-Time	0.008	−0.015
	(0.010)	(0.012)
Observations	71,985	71,950
R^2	0.580	0.539
FE: Individual	X	X
FE: Year	X	X

* p<0.1, ** p<0.05, *** p<0.01. Individual and year FE. SE clustered by individual. Part-Time = contractual hours < 35. Controls: age, age^2, female, has children.

A.6 Selection Comparison

Table A.5: Work-Life Interference: Always-FT vs. Always-PT vs. Contract Switchers

	Always FT	Always PT	Switchers (FT obs)	Switchers (PT obs)
Hours Gap	0.069***	0.060***	0.076***	0.067***
	(0.009)	(0.012)	(0.009)	(0.008)
Hours Gap × French	−0.006	−0.031*	−0.017	−0.042***
	(0.014)	(0.018)	(0.017)	(0.011)
Observations	29,059	12,869	14,647	13,091
R^2	0.575	0.564	0.520	0.520
FE: Individual	X	X	X	X
FE: Year	X	X	X	X

* p<0.1, ** p<0.05, *** p<0.01. Individual and year FE. SE clustered by individual. Always-FT: individuals never observed in PT. Always-PT: individuals never observed in FT. Switchers: individuals observed in both FT and PT; columns split by contract type at time of observation.

A.7 Language Switchers

Table A.6 reproduces the baseline regression for work-life interference, post-work exhaustion, and difficulty disconnecting on the subsample of 142 individuals who change their SHP interview language at least once during the panel (49 with two or more observations in each language group, $N_{\text{obs}} \approx 700$). Under a strict individual-language-identity interpretation of the French indicator, these within-person estimates should capture a causal language effect rather than a regional norm effect; if the interaction merely reflects individual characteristics correlated with language, the within-switcher estimate should be large. In practice, all three interaction coefficients are statistically insignificant ($p > 0.26$), though the subsample is too small for these imprecise estimates to be informative. The null result is consistent with the main findings reflecting regional norms (absorbed by the French indicator) rather than a stable personal trait that moves with the individual.

Table A.6: Cultural interaction on language-switcher subsample

	pf50	pf51	pf52
Hours Gap	0.085***	0.033	0.042
	(0.030)	(0.020)	(0.029)
Hours Gap $\times$ French	−0.039	0.023	0.001
	(0.027)	(0.022)	(0.029)
French	0.178	−0.223	0.025
	(0.373)	(0.325)	(0.458)
Observations	673	675	675
Individual FE	X	X	X
Year FE	X	X	X

* p <0.1, ** p <0.05, *** p <0.01
Sample restricted to the 142 individuals (49 with ≥ 2 observations in each language group) who switch interview language during the panel. $N_{\text{obs}} \approx 763$. Outcomes: WLI = work-life interference (pf50); Exhaust = post-work exhaustion (pf51); Disc = difficulty disconnecting (pf52). Standard errors clustered at individual level. SHP calibration weights applied.

A.8 Border Corridor Robustness

Table A.7 replicates the main specifications restricting the sample to workers residing in cantons adjacent to the linguistic border (Bern, Fribourg, Valais, Graubünden, Solothurn, and Jura). These cantons contain the sharpest concentration of French–German cultural exposure. The border-canton WLI interaction is -0.030 ($p = 0.003$, N = 24,899), somewhat attenuated relative to the full-sample estimate of -0.038 but retaining statistical significance. The post-work exhaustion interaction also remains significant in the border sample (-0.017, $p = 0.040$). The attenuation is consistent with reduced contrast in cultural exposure near the border, where bilingual communities and cross-border commuting soften the linguistic divide; it does not undermine the main finding. Note that the French-speaking sub-sample within border cantons is concentrated in bilingual Fribourg and Valais, so the border estimate should be interpreted cautiously.

Table A.7: Border Corridor Robustness: Full vs. Border-Canton Sample

	WLI (Full)	WLI (Border)	Exh (Full)	Exh (Border)	Disc (Full)	Disc (Border)
Hours Gap	0.063***	0.056***	0.042***	0.038***	0.039***	0.032***
	(0.004)	(0.007)	(0.003)	(0.006)	(0.003)	(0.006)
Hours Gap × French	−0.038***	−0.030***	−0.021***	−0.017**	−0.009*	0.002
	(0.006)	(0.010)	(0.005)	(0.008)	(0.005)	(0.009)
Observations	71,950	24,899	71,985	24,905	72,022	24,910
R^2	0.533	0.554	0.576	0.590	0.625	0.632
FE: Individual	X	X	X	X	X	X
FE: Year	X	X	X	X	X	X

* p<0.1, ** p<0.05, *** p<0.01. Individual and year fixed effects. SE clustered by individual. Border cantons (BE, FR, VS, GR, SO, JU) are those adjacent to the Röstigraben linguistic border. French-speaking respondents in border cantons are concentrated in bilingual Fribourg and Valais.

A.9 Permutation Inference

Figure A.9 plots the null distribution of $\hat{\beta}_3$ under $B = 500$ random permutations of language labels across individuals, holding all other features of the data fixed. Under random assignment, the permutation distribution is approximately normal with mean 0.000 and standard deviation 0.007. The observed $\hat{\beta}_3 = -0.038$ (red line) lies 5.4 standard deviations below the permutation mean, yielding a two-sided permutation $p < 0.002$. This confirms that the estimated cultural interaction cannot arise from chance variation in how individual language labels map onto within-person overwork exposure.

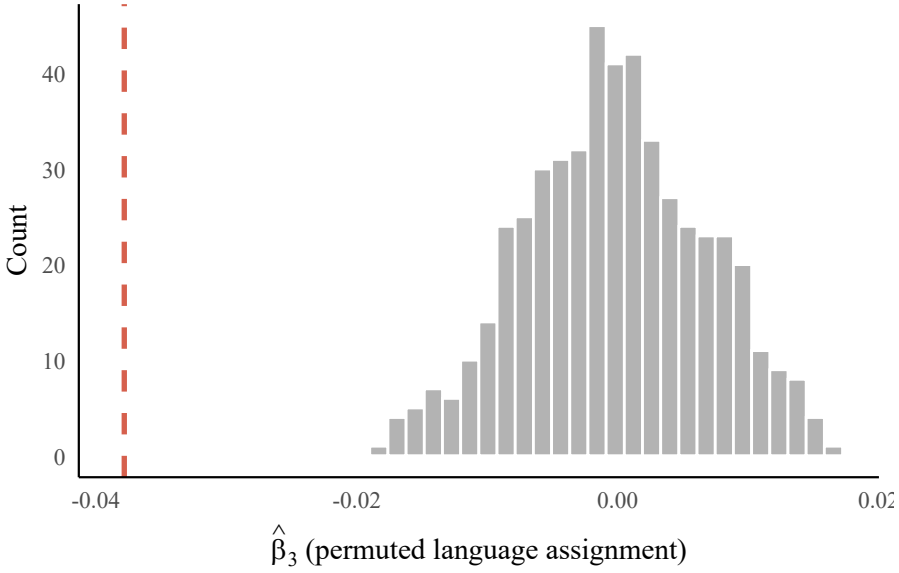


Figure A.9: Permutation null distribution of $\hat{\beta}_3$ (Hours Gap $\times$ French) under $B = 500$ random permutations of language assignment across individuals. Red dashed line: observed $\hat{\beta}_3 = -0.038$. Two-sided permutation $p < 0.002$ (0 of 500 permuted coefficients $\leq$ observed).

A.10 Period Stability

Figure A.5 plots $\hat{\beta}_3$ separately for four sub-periods (1999–2005, 2006–2012, 2013–2019, 2020–2023). Estimates are -0.043 ($p = 0.018$), -0.034 ($p = 0.001$), -0.035 ($p < 0.001$), and -0.019 ($p = 0.081$) respectively. The 2020–2023 estimate is attenuated, consistent with the post-pandemic WFH expansion blurring work-leisure boundaries and reducing the salience of contractual hours. A formal Wald test of the triple interaction (Hours Gap $\times$ French $\times$ Period) yields $F = 0.74$ ($p = 0.53$), indicating that $\hat{\beta}_3$ is statistically stable across periods and the attenuation in the final period is within sampling variation.

A.11 Attrition

Table A.8: Attrition Test: Differential Panel Exit

	Panel Exit (1)	Panel Exit (2)
Hours Gap	0.000	0.000
	(0.000)	(0.000)
Hours Gap × French	0.001	0.001
	(0.001)	(0.001)
Exhaustion		−0.002
		(0.001)
Exhaustion × French		0.001
		(0.002)
Observations	73,297	68,251
R^2	0.338	0.344
FE: Individual	X	X
FE: Year	X	X

* p<0.1, ** p<0.05, *** p<0.01. Individual and year FE. SE clustered by individual. Dependent variable = 1 if individual not observed in $t + 1$ (excludes 2023). Controls: age, age^2, female, has children.

A.12 Bunching at the Contract

Table A.9: Contractual Boundary Adherence by Language Region

	$\|\text{Gap}\| \leq 1$	$\|\text{Gap}\| \leq 2$	$\|\text{Gap}\| \leq 3$
French Region	0.020**	0.013	0.008
	(0.009)	(0.008)	(0.011)
Observations	79,274	79,274	79,274
R^2	0.060	0.064	0.055
FE: Year	X	X	X
FE: Canton	X	X	X

* p<0.1, ** p<0.05, *** p<0.01. Canton and year fixed effects. Standard errors clustered at the canton level. Controls: age, age^2, female, has children, education category. Sample restricted to $|\text{hours gap}| \leq 20$.

A.13 Job Mobility

Table A.10: Job Mobility Response to Overwork by Language Region

	Job Change (FE)	Job Change (Binary)	Job Change (Canton)
Hours Gap$_{t-1}$	0.000		0.000
	(0.001)		(0.000)
French Region	0.060	0.057	−0.002
	(0.082)	(0.081)	(0.009)
Hours Gap$_{t-1}$ × French	0.000		0.000
	(0.001)		(0.000)
Overwork$_{t-1}$		−0.012*	
		(0.006)	
Overwork$_{t-1}$ × French		0.002	
		(0.010)	
Observations	54,072	54,072	56,485
R^2	0.291	0.292	0.052
FE: Individual	X	X	
FE: Year	X	X	X
FE: Canton			X

* p<0.1, ** p<0.05, *** p<0.01. Columns 1–2: individual and year fixed effects, standard errors clustered at the individual level. Column 3: canton and year fixed effects, standard errors clustered at the canton level. Controls: age, age^2, female, has children (all columns), education category (column 3). Sample restricted to |hours gap| $\leq$ 20.

A.14 Behavioral Robustness

Table A.11: Robustness: Behavioral Response to Overwork

	A-H IV	Border	Border Exit	Men	Women
Hours Gap$_{t-1}$	1.507	0.042***	−0.001	0.038***	0.023*
	(4.920)	(0.016)	(0.001)	(0.013)	(0.013)
Hours Gap$_{t-1}$ × French	−2.467	−0.015	0.002	−0.006	−0.020
	(6.332)	(0.025)	(0.002)	(0.019)	(0.018)
French Region		−1.117	0.056	0.141	−0.885
		(0.810)	(0.142)	(0.456)	(0.641)
Observations	40,675	18,658	18,658	27,729	26,342
R^2	0.527	0.581	0.298	0.591	0.477
FE: Individual	X	X	X	X	X
FE: Year	X	X	X	X	X

* p<0.1, ** p<0.05, *** p<0.01. Individual and year fixed effects. Standard errors clustered at the individual level. Column 1: Anderson-Hsiao IV (lag2 instruments lag1). Columns 2–3: border cantons only. Columns 4–5: gender split. Outcome is hours gap$_t$ except column 3 (job change$_t$). Sample restricted to |hours gap| $\leq$ 20.

B Robustness Checks

B.1 Outcome Selection and Hypothesis Registration

This paper uses three psychological outcomes: work-life interference (pf50), post-work exhaustion (pf51), and difficulty disconnecting (pf52), selected as the SHP variables most directly

measuring the work-leisure boundary. No formal pre-registration exists. Work-life interference is the **primary outcome**: it is the most direct operationalization of the boundary-violation construct, the most commonly used burnout proxy in the SHP literature, and the most robust of the three across all specifications (see Section 5.4 and Table A.2). Post-work exhaustion is the **secondary outcome**: it measures the temporal dimension of boundary costs (residual fatigue after the workday ends). Difficulty disconnecting is a **tertiary outcome**: it captures the cognitive dimension but is borderline significant ($\hat{\beta}_3 = -0.009$, $p = 0.071$) and not claimed as an established effect. The exhaustion result is *not robust across all specifications*: it attenuates when extreme hours-gap values are trimmed and when observations with very low contractual hours are excluded (Table B.9), unlike WLI. The main finding rests on the WLI result; the exhaustion result provides corroborating but fragile evidence.

B.2 Coefficient Stability (Oster, 2019)

Table B.1 reports the interaction coefficient across specifications with progressively richer control sets. The coefficient on Hours Gap $\times$ French moves by less than 2% from the no-controls to the full-controls specification for both work-life interference and exhaustion. We follow the standard convention of setting $R_{\max} = 1.3 \times R^2_{\text{full}}$ (Oster, 2019), where R^2_{full} is the R-squared from the most saturated specification (0.533 for WLI, 0.576 for exhaustion), yielding $R_{\max} = 0.693$ and 0.749 respectively. Under the proportionality assumption (selection on unobservables proportional to selection on observables), the δ^* statistic is the ratio of (i) the coefficient movement relative to the controlled coefficient to (ii) the normalized R-squared movement, and measures how many times more important selection on unobservables would need to be relative to selection on observables to drive the interaction to zero. The δ^* values are invariant to the exact choice of $R_{\max}$ when R^2 barely moves across specifications, as it does here: the coefficient changes by less than 0.001 from the no-controls to the full-controls specification (Table B.1). The resulting δ^* values far exceed the conventional threshold of 1, reflecting the exceptional stability of the coefficient across specifications. The extreme magnitudes arise because the unrounded coefficient changes by less than 0.001 (e.g., the WLI interaction moves from -0.03381 to -0.03316 between specifications), a difference that rounds away in the table but enters the δ^* denominator. A reader computing δ^* from the rounded table values would obtain much smaller values (approximately 0.9 for WLI); the discrepancy reflects rounding, not a computation error. We interpret the δ^* results as confirming that the interaction is stable as controls are added, though we emphasize that coefficient stability does not rule out time-varying confounders that are orthogonal to the included controls.

Table B.1: Coefficient Stability Across Specifications (Oster 2019)

	WLI (1)	WLI (2)	WLI (3)	Exh (1)	Exh (2)
Hours Gap	0.064***	0.063***	0.061***	0.043***	0.042***
	(0.004)	(0.004)	(0.004)	(0.003)	(0.003)
Hours Gap × French	−0.038***	−0.038***	−0.037***	−0.022***	−0.021***
	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)
Observations	71,950	71,950	71,918	71,985	71,985
R^2	0.532	0.533	0.536	0.576	0.576
FE: Individual	X	X	X	X	X
FE: Year	X	X	X	X	X

* p<0.1, ** p<0.05, *** p<0.01. Individual and year FE in all columns. SE clustered by individual. (1): no controls. (2): age, age^2, female, has children. (3): adds ISCO 1-digit occupation FE.

B.3 Occupation-Level Instrumental Variables

Table B.2 reports an instrumental variables specification that instruments the hours gap and its interaction with the French indicator using leave-one-out ISCO major-group × year mean hours gaps and their interactions with French. The Sargan-Hansen test and first-stage F-statistics are reported in the table notes.

Table B.2: Instrumental Variables: Occupation-Level Hours Trends

	WLI (OLS)	WLI (IV)	Exhaust. (OLS)	Exhaust. (IV)
Hours Gap	0.101***	0.643***	0.067***	0.530***
	(0.005)	(0.087)	(0.004)	(0.076)
Hours Gap × French	−0.028***	0.458***	−0.018*	0.345**
	(0.009)	(0.174)	(0.009)	(0.145)
Observations	70,716	70,716	70,754	70,754
R^2	0.536	0.527	0.579	0.571
FE: Individual	X	X	X	X
FE: Year	X	X	X	X

* p<0.1, ** p<0.05, *** p<0.01. Individual and year FE. SE clustered by individual. IV columns instrument Hours Gap and Hours Gap × French with leave-one-out ISCO major group × year mean hours gap and its interaction with French. Controls: age, age^2, female, has children.

We do not treat the IV estimates as informative and exclude them from the main analysis. The exclusion restriction requires that occupation-level hours trends affect individual well-being outcomes only through own hours worked. This restriction is implausible: periods of high occupational demand simultaneously push hours upward *and* generate additional stress through non-hours channels (workload intensity, supervision pressure, job insecurity), directly violating the exclusion restriction for outcomes as sensitive to job conditions as WLI and exhaustion. Consistent with this concern, the IV estimates reverse the sign of the cultural interaction (Hours Gap × French: from -0.028^{***} under OLS to $+0.458^{***}$ under IV for WLI), which we attribute to instrument invalidity rather than OLS bias. The OLS estimates, identified from within-person variation conditional on individual and year fixed effects, are our preferred specification.

B.4 Clustering and Weighting Robustness

Table B.3 compares the main results across six combinations of clustering strategy and sample weighting. Column 1 is the baseline (individual-level clustering, SHP calibration weights). Column 2 uses two-way clustering (individual × year). Column 3 uses canton-level clustering (26 clusters). Column 4 is unweighted. Columns 5–6 replicate the comparison for exhaustion. The interaction coefficient $\hat{\beta}_3$ is significant across all specifications (negative for burnout outcomes, positive for satisfaction), and the canton-level SEs are smaller than individual-level SEs because canton-level clustering is less conservative (cantons capture language-region variation already absorbed by the interaction term).

Table B.3: Robustness: Clustering and Weighting Strategies

	WLI (Indiv)	WLI (2-Way)	WLI (Canton)	WLI (Unwgtd)	Exh (Indiv)	Exh (Canton)
Hours Gap	0.063***	0.063***	0.063***	0.063***	0.042***	0.043***
	(0.004)	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)
Hours Gap × French	−0.038***	−0.038***	−0.033***	−0.033***	−0.021***	−0.018***
	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)	(0.004)
Num.Obs.	71,950	71,950	74,053	74,053	71,985	74,089
R2	0.533	0.533	0.522	0.522	0.576	0.563
FE: idpers	X	X	X	X	X	X
FE: year	X	X	X	X	X	X

* p<0.1, ** p<0.05, *** p<0.01. Individual and year FE throughout. Col. 1: baseline (individual clustering, weighted). Col. 2: two-way clustering (individual × year). Col. 3: canton-level clustering (26 clusters). Col. 4: unweighted. Col. 5–6: same as 1 and 3 for exhaustion.

B.5 Canton-by-Year Fixed Effects

Table B.4 adds canton-by-year fixed effects to the baseline specification, absorbing all region-specific annual shocks (e.g., canton-level industry demand shifts, local policy changes, or labor market conditions). This addresses the concern that the cultural interaction may partly reflect differential exposure to canton-specific economic shocks correlated with language region. The interaction coefficient is virtually unchanged: $\hat{\beta}_3 = -0.038$ ($p < 0.001$) for work-life interference under both baseline and canton-by-year FE, and $\hat{\beta}_3 = -0.022$ ($p < 0.001$) for exhaustion under both specifications. This stability demonstrates that region-specific annual shocks do not drive the cultural interaction.

Table B.4: Canton-by-Year Fixed Effects Robustness

	WLI (Baseline)	WLI (Canton×Year)	Exhaust (Baseline)	Exhaust (Canton×Year)	Disc (Baseline)	Disc (Canton×Year)
Hours Gap	0.063***	0.063***	0.042***	0.042***	0.039***	0.039***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
Hours Gap × French	−0.038***	−0.038***	−0.021***	−0.021***	−0.009*	−0.010*
	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)
Observations	71,950	71,946	71,985	71,981	72,022	72,018
R^2	0.533	0.540	0.576	0.582	0.625	0.630
FE: Individual	X	X	X	X	X	X
FE: Year	X		X		X	
FE: Canton × Year		X		X		X

* p<0.1, ** p<0.05, *** p<0.01. SE clustered by individual. Odd columns: individual and year FE (baseline, restricted to observations with non-missing canton). Even columns: individual and canton×year FE, absorbing all region-specific annual shocks. Controls: age, age^2, female, has children.

B.6 Alternative Outcomes

Table B.5 replaces the primary burnout outcomes with two additional measures: satisfaction with work conditions, and post-work exhaustion.

Table B.5: Robustness: Alternative Outcome Measures

	Work Cond. Sat.	Post-work Exhaust.
Hours Gap	−0.022***	0.042***
	(0.002)	(0.003)
French Region	0.362	0.173
	(0.235)	(0.259)
Hours Gap × French	0.014***	−0.021***
	(0.004)	(0.005)
Observations	77,001	71,985
R^2	0.511	0.576
FE: Individual	X	X
FE: Year	X	X

* p<0.1, ** p<0.05, *** p<0.01. Individual and year fixed effects. Standard errors clustered by individual.

B.7 Employees Only

Table B.6 restricts the sample to employees (excluding self-employed and other non-standard employment types). Results are virtually identical to the full sample, confirming that self-employment does not drive the main findings.

Table B.6: Robustness: Employees Only

	Exhaust. (All)	Exhaust. (Emp.)	WLI (All)	WLI (Emp.)
Hours Gap	0.042***	0.042***	0.063***	0.063***
	(0.003)	(0.003)	(0.004)	(0.004)
Hours Gap × French	−0.021***	−0.021***	−0.038***	−0.038***
	(0.005)	(0.005)	(0.006)	(0.006)
Observations	71,985	71,883	71,950	71,848
R^2	0.576	0.577	0.533	0.533
FE: Individual	X	X	X	X
FE: Year	X	X	X	X

* p<0.1, ** p<0.05, *** p<0.01. Individual and year FE. SE clustered by individual. Odd cols: full sample. Even cols: employees only (pw29=5). Controls: age, age^2, female, has children.

B.8 Comprehensive Robustness Summary

Table B.7 presents the exhaustion interaction coefficient across all robustness specifications in a single table for ease of comparison.

Table B.7: Robustness Summary: Post-Work Exhaustion

	Baseline	+Occ. FE	+All Ctrls	Trimmed	Full-Time	Part-Time	Employees
Hours Gap	0.042***	0.041***	0.039***	0.067***	0.053***	0.035***	0.042***
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.006)	(0.003)
Hours Gap × French	−0.021***	−0.020***	−0.019***	−0.018*	−0.011	−0.009	−0.021***
	(0.005)	(0.005)	(0.005)	(0.009)	(0.009)	(0.009)	(0.005)
Observations	71,985	71,953	70,822	70,797	43,721	25,980	71,883
R^2	0.576	0.579	0.582	0.579	0.599	0.602	0.577
FE: Individual	X	X	X	X	X	X	X
FE: Year	X	X	X	X	X	X	X

* p<0.1, ** p<0.05, *** p<0.01. Individual and year FE. SE clustered by individual. Col. 1: baseline. Col. 2: + ISCO FE. Col. 3: + ISCO FE, supervisor, education. Col. 4: |hours gap| $\leq$ 20. Col. 5–6: contractual hours $\geq/<$ 35. Col. 7: employees only (pw29=5).

B.9 Additional Outcomes: Work Stress

Table B.8 reports the cultural interaction for work stress alongside the main burnout outcomes for comparison. Work stress shows a *positive* interaction ($\hat{\beta}_3 = +0.003^{***}$), opposite in sign to the negative interactions for exhaustion (-0.021^{***}) and work-life interference (-0.038^{***}). This sign reversal supports the boundary-specificity interpretation: the cultural penalty amplifies outcomes related to the work-leisure boundary, not within-job cognitive demands.

Table B.8: Additional Outcomes: Work Stress

	Exhaustion	Work-Life Int.	Work Stress
Hours Gap	0.042***	0.063***	−0.007***
	(0.003)	(0.004)	(0.001)
Hours Gap × French	−0.021***	−0.038***	0.003***
	(0.005)	(0.006)	(0.001)
Observations	71,985	71,950	67,290
R^2	0.576	0.533	0.537
FE: Individual	X	X	X
FE: Year	X	X	X

* p<0.1, ** p<0.05, *** p<0.01. Individual and year FE. SE clustered by individual. Controls: age, age^2, female, has children. All outcomes scaled 0–10.

B.10 Sensitivity to Outlier Trimming

Table B.9 presents the main interaction coefficients under alternative sample restrictions. The "clean" specification drops observations with contractual hours below 10 per week (5% of the sample), where even moderate actual hours generate mechanically large hours gaps. The "strict" specification additionally restricts |hours gap| $\leq$ 20.

Table B.9: Sensitivity to Outlier Trimming

	Ex. (base)	Ex. (clean)	Ex. (strict)	WLI (base)	WLI (clean)	WLI (strict)
Hours Gap	0.042***	0.045***	0.066***	0.063***	0.066***	0.100***
	(0.003)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)
Hours Gap × French	−0.021***	−0.012*	−0.013	−0.038***	−0.019**	−0.025***
	(0.005)	(0.007)	(0.010)	(0.006)	(0.008)	(0.009)
Observations	71,985	68,267	67,415	71,950	68,232	67,377
R^2	0.576	0.582	0.584	0.533	0.536	0.538
FE: Individual	X	X	X	X	X	X
FE: Year	X	X	X	X	X	X

* p<0.1, ** p<0.05, *** p<0.01. Individual and year FE. SE clustered by individual. Base: full sample. Clean: contractual hours ≥ 10. Strict: contractual hours ≥ 10 and |hours gap| ≤ 20.

B.11 Contract Type × Gender: Full Results

Table B.10 presents the cultural interaction across all three burnout outcomes, split by contract type and gender. The pattern is consistent: the effect is concentrated among part-time workers, and within part-time workers, men show a larger point estimate than women.

Table B.10: Contract Type × Gender: All Outcomes

	Ex. (PT-F)	Ex. (PT-M)	Ex. (FT-F)	Ex. (FT-M)	WLI (PT-F)	WLI (PT-M)	WLI (FT-F)	WLI (FT-M)	Disc (PT-F)	Disc (PT-M)	Disc (FT-F)	Disc (FT-M)
Hours Gap	0.039***	0.022**	0.066***	0.048***	0.066***	0.058***	0.090***	0.065***	0.046***	0.037***	0.051***	0.041***
	(0.007)	(0.010)	(0.008)	(0.005)	(0.008)	(0.010)	(0.010)	(0.008)	(0.007)	(0.011)	(0.009)	(0.005)
Hours Gap × French	−0.009	−0.007	−0.031**	−0.004	−0.031***	−0.054***	−0.032*	−0.002	−0.003	−0.017	0.001	0.009
	(0.012)	(0.014)	(0.014)	(0.012)	(0.012)	(0.015)	(0.017)	(0.013)	(0.012)	(0.015)	(0.016)	(0.011)
Observations	21,378	4,600	13,220	30,501	21,366	4,592	13,212	30,494	21,389	4,604	13,232	30,511
R^2	0.595	0.635	0.607	0.589	0.533	0.598	0.563	0.556	0.645	0.679	0.653	0.642
FE: Individual	X	X	X	X	X	X	X	X	X	X	X	X
FE: Year	X	X	X	X	X	X	X	X	X	X	X	X

* p<0.1, ** p<0.05, *** p<0.01. Individual and year FE. SE clustered by individual. PT: contractual hours < 35. FT: contractual hours ≥ 35. Controls: age, age^2, has children.

B.12 Supervisory Status as Autonomy Proxy

Table B.11 splits the sample by supervisory status to test whether the cultural interaction on psychological outcomes varies with schedule-setting authority. The interaction is nearly identical for supervisors and non-supervisors, indicating that the cultural cost of overwork does not depend on occupational autonomy.

Table B.11: Psychological Cost by Supervisory Status

	WLI (Supervisor)	WLI (Non-Sup.)	Exh (Supervisor)	Exh (Non-Sup.)
Hours Gap	0.067***	0.046***	0.045***	0.032***
	(0.006)	(0.006)	(0.004)	(0.005)
Hours Gap × French	−0.034***	−0.036***	−0.017**	−0.024***
	(0.008)	(0.009)	(0.007)	(0.008)
Observations	35,770	32,686	35,782	32,714
R^2	0.570	0.545	0.609	0.603
FE: Individual	X	X	X	X
FE: Year	X	X	X	X

* p<0.1, ** p<0.05, *** p<0.01. Individual and year FE. SE clustered by individual. Controls: age, age^2, female, has children. Supervisor = reports supervisory responsibilities (pw87=1).

Table B.12 tests whether workers with supervisory responsibilities (and presumably greater schedule flexibility) show differential behavioral adjustment to overwork. Even supervisors do not differentially correct overwork episodes across language regions.

Table B.12: Behavioral Adjustment by Supervisory Status

	Hours Adj (Supervisor)	Hours Adj (Non-Sup.)
Hours Gap$_{t-1}$	0.037*	−0.058***
	(0.022)	(0.021)
Hours Gap$_{t-1}$ × French	−0.082***	−0.027
	(0.030)	(0.034)
Observations	28,992	23,312
R^2	0.538	0.415
FE: Individual	X	X
FE: Year	X	X

* p<0.1, ** p<0.05, *** p<0.01. Individual and year FE. SE clustered by individual. Controls: age, age^2, female, has children. DV: Hours Gap$_t$.

B.13 Semi-Objective Health Outcomes

Table B.13 tests whether the cultural interaction extends to semi-objective health measures across eight outcomes. Weakness/weariness shows a nominally significant cultural interaction (0.004, $p = 0.046$), but after Benjamini-Hochberg correction for eight simultaneous tests, the adjusted p-value is 0.36. All other outcomes are far from significant: days affected by health problems ($p = 0.70$), doctor consultations ($p = 0.85$), health impediment ($p = 0.83$), sleep problems ($p = 0.32$), burnout diagnosis ($p = 0.62$), depression diagnosis ($p = 0.55$), and seeing a doctor ($p = 1.00$). No health measure survives multiple-testing correction, confirming that the cultural cost is genuinely psychological rather than a proxy for differential physical health consequences.

Table B.13: Semi-Objective Health Outcomes

	Health Days	Doctor Visits	Burnout Dx	Depression Dx	Weakness	Sleep Problems	Saw Doctor	Health Impediment
Hours Gap	−0.218***	−0.036	0.000	0.002	0.004***	0.003***	0.002**	0.000
	(0.068)	(0.022)	(0.001)	(0.003)	(0.001)	(0.001)	(0.001)	(0.004)
Hours Gap × French	0.096	−0.014	−0.003	−0.004	−0.003	−0.004*	0.000	−0.001
	(0.110)	(0.035)	(0.003)	(0.005)	(0.002)	(0.002)	(0.001)	(0.008)
Observations	75,161	52,454	2,766	2,760	66,503	66,501	75,657	75,732
R^2	0.317	0.397	0.699	0.777	0.503	0.531	0.378	0.511
FE: Individual	X	X	X	X	X	X	X	X
FE: Year	X	X	X	X	X	X	X	X

* p<0.1, ** p<0.05, *** p<0.01. Individual and year FE. SE clustered by individual. Controls: age, age^2, female, has children. Health Days = days affected by health problems (last 12 months). Doctor Visits = number of consultations (last 12 months). Burnout Dx = self-reported burnout diagnosis. Weakness = weakness/weariness (last 4 weeks). Sleep = sleeping problems (last 4 weeks).

C Extension Analyses

C.1 Event Study Around Overwork Onset

Figure C.1 presents a regression-based event study around the first transition from non-overwork ($\Delta \leq 0$) to overwork ($\Delta > 0$). The coefficients represent the cultural interaction (Event Time $\times$ French) at each event-time relative to $t = -1$. The event study identifies 5,749 first-onset events across the panel.

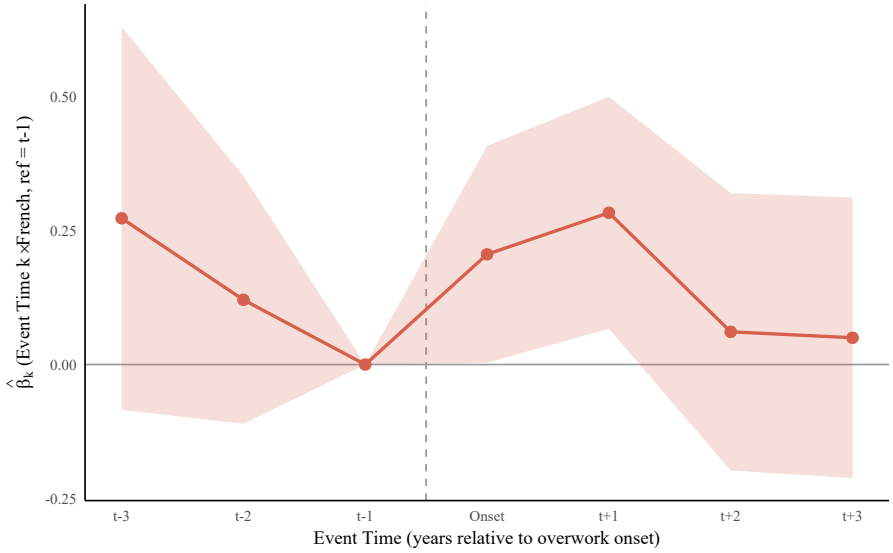


Figure C.1: Event study: cultural interaction coefficients around first overwork onset. Coefficients on Event Time $\times$ French from a regression with individual and year FE, relative to $t = -1$. 95% confidence intervals. Dashed line marks onset.

The cultural interaction jumps at onset ($\hat{\beta}_0 = 0.21$, $p = 0.046$) and is elevated at $t+1$ (0.28, $p = 0.010$) before fading at $t+2$ and $t+3$. Pre-onset coefficients at $t = -3$ (0.27, $p = 0.13$) and $t = -2$ (0.12, $p = 0.31$) are positive but individually insignificant; the joint pre-trend test does not reject the null at conventional levels ($F = 1.20$, $p = 0.30$; Table D.6). While the pre-trend coefficients are not statistically significant, their positive magnitudes are non-trivial and comparable to the post-onset estimates, warranting caution in interpreting the event study causally. One interpretation is selection: workers who will eventually become overworked may already be on diverging WLI trajectories, which would imply that the event study identifies a selection effect rather than a causal overwork effect. A second interpretation is that anticipatory stress (German-speaking workers may experience rising work-life conflict as workload increases in the months before formal overwork materializes) generates pre-trends in the event study even if the main fixed-effects specification is well-identified, because the event study conditions on a binary onset while the main specification uses continuous within-person hours-gap variation.

We emphasize that the main fixed-effects results do not rely on this event study and are identified from a different source of variation. The event study uses a binary onset indicator and compares WLI *levels* around the transition; the main specification uses *continuous* within-person hours-gap variation and its interaction with language region. The pre-trends in the event study would bias the main results only if the same time-varying unobservables that generate pre-event WLI divergence also generate differential *slopes* of the hours-gap–WLI relationship by language group conditional on individual and year fixed effects, a more demanding condition. The stability of $\hat{\beta}_3$ across specifications with progressively richer controls (Table B.1;

$\delta^* > 1{,}700$ for all outcomes) provides evidence against such confounding in the main specification. Nevertheless, we report the event study for transparency and make no causal claims based on it. Figure C.2 shows the raw means.

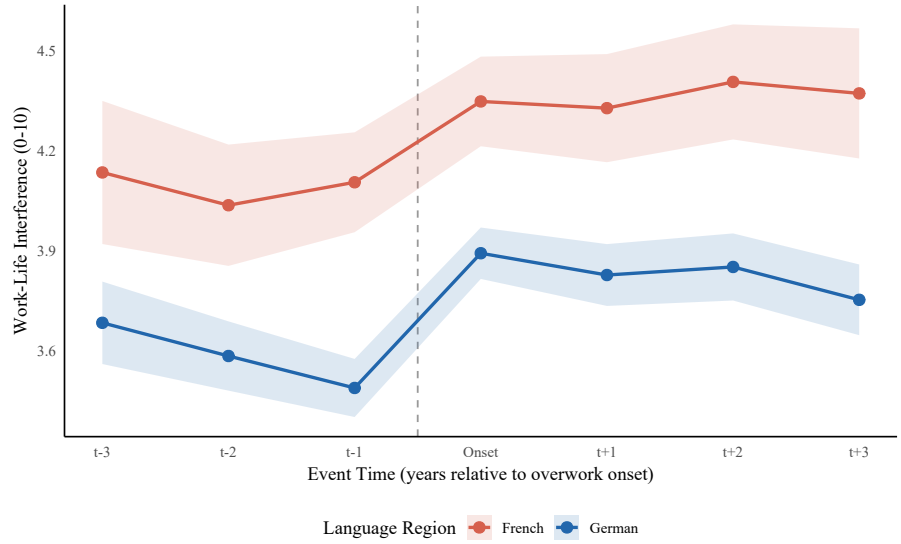


Figure C.2: Raw mean work-life interference around first overwork onset, by language region. 95% confidence intervals. Both groups show a jump at onset; the French-speaking group jumps slightly more.

C.2 Dose-Response Nonlinearity

Table C.1 augments the baseline specification with a squared hours gap term and its interaction with the French indicator, allowing the cultural moderation of overwork to vary nonlinearly with the intensity of overwork. The main squared term (Hours Gap^2) is negative and highly significant for exhaustion (-0.002, $p < 0.001$), work-life interference (-0.003, $p < 0.001$), and difficulty disconnecting (-0.002, $p < 0.001$), indicating diminishing marginal burnout effects at high levels of overwork. However, the squared interaction (Hours Gap$^2 \times$ French) is small and statistically insignificant in all columns, confirming that the concavity is the same for both language groups. The linear cultural moderation is therefore an adequate approximation.

Table C.1: Dose-Response: Nonlinear Cultural Effect of Overwork

	Exhaustion	Work-Life Int.	Disconnect	Free Time Sat.
Hours Gap	0.132***	0.176***	0.151***	−0.088***
	(0.009)	(0.011)	(0.010)	(0.009)
French Region	0.101	0.177	−0.030	−0.305
	(0.277)	(0.335)	(0.409)	(0.336)
Hours Gap × French	0.019	−0.027	0.012	0.037**
	(0.022)	(0.021)	(0.021)	(0.018)
Hours Gap^2	−0.004***	−0.004***	−0.005***	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Hours Gap^2 × French	−0.003	0.000	0.000	−0.001
	(0.002)	(0.001)	(0.001)	(0.001)
Observations	68,310	68,272	68,340	71,260
R^2	0.586	0.543	0.634	0.576
FE: Individual	X	X	X	X
FE: Year	X	X	X	X

* p<0.1, ** p<0.05, *** p<0.01. Individual and year fixed effects. Standard errors clustered at the individual level. Controls: age, age^2, female, has children. Sample restricted to hours gap ≥ 0. All outcomes scaled 0–10.

Figure C.3 plots the implied marginal effect of an additional hour of overwork on exhaustion for German- and French-speaking workers, derived from the quadratic specification. Both curves decline with the hours gap but are nearly parallel, confirming that there is no differential nonlinearity.

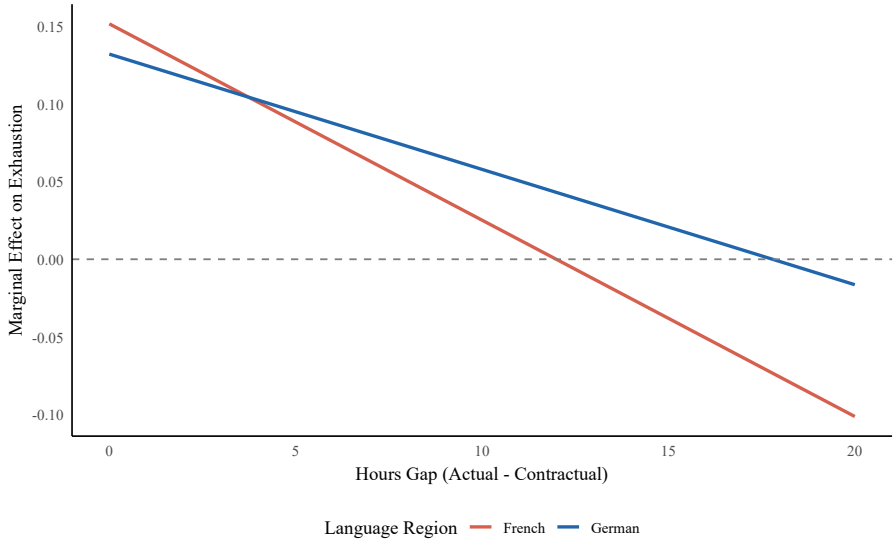


Figure C.3: Marginal effect of an additional hour of overwork on exhaustion, by language region. Computed from the quadratic specification in Table C.1, column 1.

C.3 COVID-19 Stability

Table C.2 tests whether the cultural moderation of overwork changed after the onset of COVID-19 by adding a triple interaction (Hours Gap × French × Post-2020) to the baseline specification. The triple interaction is small and statistically insignificant across all four outcomes ($|\text{coefficients}| < 0.02$, all $p > 0.3$), indicating that the cultural penalty for over-

work is temporally stable and was not disrupted by the pandemic. The French × Post-2020 term is significant only for work-life interference (0.157, $p < 0.05$), suggesting that French-speaking workers experienced a general increase in work-life conflict after 2020 independent of overwork levels.

Table C.2: Cultural Effect of Overwork Before and After COVID-19

	Exhaustion	Work-Life Int.	Disconnect	Free Time
Hours Gap	0.066***	0.100***	0.067***	−0.064***
	(0.004)	(0.005)	(0.005)	(0.005)
French	0.181	0.252	0.026	−0.333
	(0.276)	(0.342)	(0.403)	(0.322)
Hours Gap × French	−0.018*	−0.028***	−0.001	0.029***
	(0.010)	(0.009)	(0.009)	(0.008)
Hours Gap × Post-2020	0.007	0.003	−0.001	0.016*
	(0.010)	(0.010)	(0.010)	(0.009)
French × Post-2020	0.007	−0.095	−0.126	0.154*
	(0.097)	(0.092)	(0.088)	(0.082)
Hours Gap × French × Post-2020	−0.004	0.003	0.016	−0.024
	(0.020)	(0.020)	(0.019)	(0.017)
Observations	70,797	70,760	70,829	73,845
R^2	0.579	0.536	0.628	0.571
FE: Individual	X	X	X	X
FE: Year	X	X	X	X

* p<0.1, ** p<0.05, *** p<0.01. Individual and year fixed effects. Standard errors clustered at the individual level. Controls: age, age^2, female, has children. Post-2020 = 1 for years $\geq$ 2021. All outcomes scaled 0–10.

C.4 Gender Attitudes as Mediator

Table C.3 examines whether the cultural moderation of overwork operates through observable gender-role attitudes. We augment the baseline specification with two SHP attitudinal items: "Having a job is the best way to be independent" (pd91) and "A child suffers if the mother works" (pd92). Cross-sectionally, French-speaking workers score 0.36 lower on job independence and 0.54 lower on the child-suffers item, confirming that attitudes differ along the language border. Both attitude variables are predictors of burnout: valuing job independence reduces exhaustion (-0.012, $p = 0.198$), while believing children suffer from working mothers increases it (0.032, $p < 0.001$). However, adding these controls produces essentially zero attenuation of the Hours Gap × French interaction (columns 2 and 4 vs. columns 1 and 3), indicating that the cultural moderation of overwork is not mediated by standard gender-role attitudes. The cultural channel operates through some other dimension of the work-leisure schema.

Table C.3: Mediation: Gender and Work Attitudes as Cultural Channel

	Exhaust. (Base)	Exhaust. (+ Attitudes)	WLI (Base)	WLI (+ Attitudes)
Hours Gap	0.065***	0.065***	0.100***	0.100***
	(0.006)	(0.006)	(0.007)	(0.007)
French Region	0.315	0.329	0.269	0.279
	(0.434)	(0.431)	(0.392)	(0.390)
Hours Gap × French	−0.011	−0.011	−0.026**	−0.025**
	(0.011)	(0.011)	(0.012)	(0.012)
Job Independence (0–10)		−0.017*		−0.014
		(0.010)		(0.010)
Child Suffers (0–10)		0.032***		0.023***
		(0.007)		(0.009)
Observations	37,333	37,333	37,300	37,300
R^2	0.592	0.592	0.552	0.552
FE: Individual	X	X	X	X
FE: Year	X	X	X	X

* p<0.1, ** p<0.05, *** p<0.01. Individual and year fixed effects. Standard errors clustered at the individual level. Controls: age, age^2, female, has children. Sample restricted to observations with non-missing attitude variables (pd91: "Having a job is the best way to be independent"; pd92: "A child suffers if mother works"). All outcomes scaled 0–10.

C.5 Alternative Hours-Gap Measure

Our baseline hours gap is defined as actual minus contractual hours. The SHP also records a "reference" hours variable: "How many hours per week do you normally work at your job?". We construct an alternative hours gap as actual minus reference hours and re-estimate the baseline specification. Table C.4 reports the results. The two gaps are only modestly correlated ($r = 0.354$), reflecting the fact that contractual hours are a formal benchmark whereas reference hours capture habitual practice. The cultural interaction is significant only for the contractual-hours gap ($\hat{\beta}_3 = -0.038$, $p < 0.001$ for work-life interference) and is essentially zero for the reference-hours gap ($\hat{\beta}_3 = -0.002$, $p = 0.548$). The same pattern holds for exhaustion: the interaction is -0.021 ($p < 0.001$) with contractual hours but -0.003 ($p = 0.264$) with reference hours. This divergence strengthens the reference-point interpretation: the cultural penalty is triggered by violations of the contractual boundary, the formal, explicit commitment, not by deviations from habitual practice.

Table C.4: Alternative Hours Gap Measure: Reference vs. Contractual Hours

	WLI (contract)	WLI (reference)	Exhaust. (contract)	Exhaust. (reference)
Hours Gap (contract)	0.063***		0.042***	
	(0.004)		(0.003)	
Hours Gap (contract) × French	−0.038***		−0.021***	
	(0.006)		(0.005)	
Hours Gap (reference)		0.035***		0.023***
		(0.002)		(0.002)
Hours Gap (reference) × French		−0.002		0.003
		(0.004)		(0.003)
Num.Obs.	71,950	71,081	71,985	71,111
R2	0.533	0.535	0.576	0.578
FE: idpers	X	X	X	X
FE: year	X	X	X	X

* p<0.1, ** p<0.05, *** p<0.01. Individual and year FE. SE clustered by individual. Contract: actual − contractual hours (pw77 − pw74). Reference: actual − reference hours (pw77 − pw46). Controls: age, age^2, female, has children.

C.6 Desired Hours

If the cultural interaction reflects French-speaking workers' stronger preference for working fewer hours, we should observe a differential desire for reduced hours when overworked. We test this using the SHP desired-hours variable: "How many hours per week would you like to work?". In a specification with desired hours as the dependent variable, the Hours Gap × French interaction is 0.090 ($p = 0.15$). Using a binary indicator for "wants fewer hours" as the outcome yields an interaction of 0.0004 ($p = 0.41$). Both estimates are null, indicating that French-speaking workers do not differentially translate overwork into a desire for fewer hours. This result is consistent with the theoretical framework: the cultural parameter $\lambda_G > \lambda_F$ increases the disutility of boundary violations without necessarily changing desired hours, since desired hours reflect a broader set of preferences including income, career progression, and social norms beyond work-leisure boundaries.

D Variable Definitions

Table D.1: Variable Definitions and SHP Source Codes

Variable	SHP Code	Description
Actual Hours	pw77	Number of hours worked per week in current main job
Contractual Hours	pw74	Number of contractual hours per week
Hours Gap	derived	pw77 − pw74
Overwork Indicator	derived	= 1 if Hours Gap > 0
Life Satisfaction	pc44	"How satisfied are you with your life in general?" (0–10)
Job Satisfaction	pw228	"How satisfied are you with your job in general?" (0–10)
Free-Time Sat.	pa05	"How satisfied are you with your free time?" (0–10)
Work-Life Interf.	pf50	Work interferes with private/family obligations (0–10)
Post-Work Exhaust.	pf51	Exhausted after work frequency (0–10)
Disconnect Diff.	pf52	Difficulty disconnecting from work (0–10)
Work Stress	pw604	"How often do you feel stressed at work?" (0–10)
Health Sat.	pc02	"How satisfied are you with your health?" (0–10)
Financial Sat.	pw227	"How satisfied are you with your financial situation?" (0–10)
Income Sat.	pw229	"How satisfied are you with your income?" (0–10)
Work Atmosphere	pw231	"How satisfied are you with the atmosphere at work?" (0–10)
Work Cond. Sat.	pw93	"How satisfied are you with your working conditions?" (0–10)
Work Amount Sat.	pw230	"How satisfied are you with the amount of work?" (0–10)
Reference Hours	pw46	"How many hours per week do you normally work?"
Desired Hours	pw84	"How many hours per week would you like to work?"
Wants Fewer Hrs	pw85	= 1 if wants to work fewer hours
Unemp. Risk	pw101	"How do you estimate the risk of becoming unemployed?" (0–10)
Supervisor	pw87	= 1 if has supervisory responsibilities
French	plingu	= 1 if interview language is French (= 2)
Canton	canton	Canton of residence (from household file)
Age	age	Age at time of interview
Female	sex	= 1 if sex = 2
Education	educat	Highest education level (11 categories, grouped to 3)
Has Children	nbkid	= 1 if number of co-resident children > 0 (HH file)

Table D.2: Controlling for Absolute Hours Worked

	Baseline	+Actual Hrs	+Actual Hrs × Fr	Exhaustion	+Contract Hrs × Fr
Hours Gap	0.063***	0.017***	0.018***	0.011***	0.018***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Hours Gap × French	−0.038***	−0.018***	−0.021***	−0.013**	0.005
	(0.006)	(0.006)	(0.006)	(0.005)	(237.854)
Actual Hours		0.058***	0.056***	0.040***	0.056***
		(0.002)	(0.002)	(0.002)	(0.002)
Observations	71,950	71,950	71,950	71,985	71,950
R^2	0.533	0.546	0.546	0.585	0.546
FE: Individual	X	X	X	X	X
FE: Year	X	X	X	X	X

* p<0.1, ** p<0.05, *** p<0.01. Individual and year FE. SE clustered by individual. DV: WLI (cols 1–3, 5), Exhaustion (col 4). Actual Hours = self-reported weekly hours worked. Contractual Hours = contractual weekly hours. Hours Gap = Actual − Contractual.

D.1 Actual Hours Control

D.2 Canton–Language Distribution

Table D.3: Canton–Language Distribution

Canton	Name	N (French)	N (German)	Ind (French)	Ind (German)	% French	Total
SO	Solothurn	420	0	114	0	100.0	420
GL	Glarus	3693	32	800	8	99.1	3725
BL	Basel-Landschaft	3913	55	722	22	98.6	3968
VS	Valais	9103	129	1877	38	98.6	9232
NW	Nidwalden	2841	827	614	191	77.5	3668
NE	Neuchâtel	2143	713	505	176	75.0	2856
TI	Ticino	56	66	11	27	45.9	122
UR	Uri	1488	9563	315	2083	13.5	11051
JU	Jura	229	14280	63	3011	1.6	14509
SZ	Schwyz	18	1474	13	428	1.2	1492
GE	Genève	10	1029	4	225	1.0	1039
OW	Obwalden	23	3004	9	683	0.8	3027
ZH	Zürich	43	7608	16	1601	0.6	7651
AI	Appenzell I.Rh.	17	4763	8	1103	0.4	4780
AG	Aargau	6	1743	3	358	0.3	1749
BS	Basel-Stadt	14	4852	7	1012	0.3	4866
SG	St. Gallen	2	700	1	164	0.3	702
ZG	Zug	1	462	1	97	0.2	463
FR	Fribourg	1	1373	1	339	0.1	1374
GR	Graubünden	2	3299	2	683	0.1	3301
AR	Appenzell A.Rh.	0	350	0	91	0.0	350
BE	Bern	0	85	0	24	0.0	85
LU	Luzern	0	675	0	134	0.0	675
SH	Schaffhausen	0	447	0	108	0.0	447
TG	Thurgau	0	1844	0	517	0.0	1844
VD	Vaud	0	338	0	70	0.0	338

Table D.4: Within-Canton Language Variation: Bilingual Cantons

	WLI: Full	WLI: Bilingual	WLI: Biling+Ct×Yr	WLI: Full+Ct×Yr	Exh: Bilingual	Exh: Biling+Ct×Yr
Hours Gap	0.063***	0.056***	0.057***	0.063***	0.037**	0.043***
	(0.004)	(0.021)	(0.021)	(0.004)	(0.016)	(0.016)
Hours Gap × French	−0.038***	−0.028	−0.030	−0.038***	−0.017	−0.023
	(0.006)	(0.022)	(0.023)	(0.006)	(0.017)	(0.017)
Observations	71,950	9,170	9,165	71,946	9,170	9,165
R^2	0.533	0.571	0.573	0.540	0.592	0.594
FE: Individual	X	X	X	X	X	X
FE: Year	X	X			X	

* p<0.1, ** p<0.05, *** p<0.01. Individual FE throughout. SE clustered by individual. Bilingual cantons: those with >5% of observations in each language group. Controls: age, age^2, female, has children.

D.3 Bilingual Canton Robustness

D.4 Kink Equality Test

Table D.5: Formal Test of Kink Equality: Above vs. Below Contract

Outcome	Below Contract	Above Contract	Difference	SE	p-value
Work-Life Interference	−0.035	−0.031	0.004	0.032	0.906
Exhaustion	−0.011	−0.003	0.008	0.028	0.772

Notes. Coefficients on Hours Gap × French from the split-slope kink specification. Difference = Above − Below. H_0: The cultural interaction is equal above and below the contract. Sample: |hours gap| ≤ 15. Individual and year FE. SE clustered by individual.

D.5 Pre-Trend Tests

Table D.6: Pre-Trend Tests for the Cultural Interaction

Test	Estimate	SE	p-value
Joint F-test ($t = -3$, $t = -2$)	—	—	0.301
Lag Hours Gap × French (continuous spec)	−0.008	0.006	0.239
Contemporary Hours Gap × French (with lag)	−0.040	0.006	< 0.001

Notes. Row 1: joint Wald test of pre-trend coefficients at $t = -3$ and $t = -2$ in the event study specification. Rows 2–3: from a continuous FE specification including both contemporaneous and lagged hours-gap interactions.

D.6 Differential Measurement Error

Table D.7: Differential Measurement Error Tests

Diagnostic	French	German	Difference
Mean hours gap	2.68	2.29	—
SD hours gap	6.35	4.98	—
Mean |hours gap|	3.29	2.83	0.452 ($p < 0.001$)
Within-person SD of hours gap	3.51	2.72	0.794 ($p < 0.001$)
% reporting round-5 actual hours	44.3	37.7	—
% reporting round-5 both hours	27.3	18.3	—

Notes. Within-person SD computed for individuals with ≥ 3 observations. Rounding = hours reported as multiple of 5. Difference in |hours gap| from regression with year FE, clustered by individual.